



**Search for the associated production of a
single top quark and a Higgs boson
in the $H \rightarrow b\bar{b}$ decay channel
at 8 and 13 TeV with the CMS experiment**

Zur Erlangung des akademischen Grades eines
DOKTORS DER NATURWISSENSCHAFTEN
von der Fakultät für Physik des
KIT - Karlsruher Institut für Technologie genehmigte

DISSERTATION

von

Dipl.-Phys. Benedikt Maier
aus Baden-Baden

Mündliche Prüfung: 22. April 2016

*Referent: Prof. Dr. Thomas Müller
Institut für Experimentelle Kernphysik*

*Korreferent: Prof. Dr. Ulrich Husemann
Institut für Experimentelle Kernphysik*

Foreword

Since the discovery of a Higgs boson-like particle in 2012, it is the goal of physicists all over the world to measure its properties as precisely as possible in order to pin down possible deviations from the predictions of the standard model of particle physics, which ultimately would give a hint at the existence of new physics. Characteristics like the particle spin and color have been successfully measured and seem to be in perfect accordance with the expectations. Another important property to probe – and the topic of this thesis – is the Yukawa coupling mechanism to fermions, which predicts that the coupling strengths are proportional to the fermion masses.

Especially the coupling parameter characterizing the interaction of the Higgs boson with the top quark is crucial for the verification of the electroweak sector of the standard model, as the top quark is the heaviest elementary particle known to exist. Hence, this thesis tries to shed light on the parameter associated to this coupling by searching for the associated production of a Higgs boson together with a single top quark at center-of-mass energies of 8 and 13 TeV with the Compact Muon Solenoid detector. Higgs boson decays into a pair of bottom quarks are considered, yielding the largest expected signal rates among all possible decay modes. The process is sensitive to the relative sign of the coupling strength to fermions and to weak gauge bosons as well as to their magnitudes, which is a unique feature of this production mode. Deviations from the coupling parameter values as predicted in the standard model would lead to altered production rates, which can be looked for in the collision data. In particular, the channel lifts the degeneracy between the standard model scenario and the scenario with a flipped sign for the coupling to the top quark, in which case the production rate is strongly enhanced.

Chapter 1 sets the theoretical scene of the associated production of Higgs boson and single top quark by first introducing the standard model and the Higgs boson as the consequence of electroweak symmetry breaking. Top quark production is briefly addressed, before the process of interest is presented in all its respects.

The analysis presented in this thesis heavily relies on statistical methods for analyzing and interpreting the data. Multivariate techniques help in separating the signal process from background contributions. They are explained in Chapter 2 together with other fundamental methods such as Monte-Carlo simulation and cross section calculation, giving a conceptual overview of the methods commonly applied in an analysis in the field of high energy physics.

In Chapter 3 the experimental setup is introduced. Descriptions of the Large Hadron Collider and the Compact Muon Solenoid detector are provided, and the digitization and reconstruction of recorded events is explained, covering the entire chain from the recording of the proton-proton collisions to providing the final processed datasets to the

analysts.

Chapter 4 presents the search for associated production of a single top quark and a Higgs boson in the data collected at $\sqrt{s} = 8$ TeV, which amounts to an integrated luminosity of 19.7 fb^{-1} . The analysis focusses on the decay of the Higgs boson into two b quarks. The purpose of this chapter is to recapitulate an analysis conducted in a group within the CMS collaboration and made public by the latter through a Physics Analysis Summary. The author has contributed significantly to these efforts at each step of the analysis chain. A paper combining the $H \rightarrow b\bar{b}$ analysis with single top quark + Higgs boson analyses performed in other Higgs boson decay channels has been submitted to JHEP at the end of 2015.

Prospects for what is possible in the area of single top quark + Higgs boson physics at the Large Hadron Collider in Run-II, which has started in the second half of 2015, are discussed for the first time in this thesis: in Chapter 5 a scan in the two-dimensional plane spanned by the couplings of the Higgs boson to fermions and vector bosons is performed. First expected upper limits on single top quark + Higgs boson associated production that correspond to 2.2 fb^{-1} of recorded collision data at a center-of-mass energy of 13 TeV are presented.

The thesis concludes in Chapter 6, providing also an outlook to larger integrated luminosities at the High-Luminosity-Large Hadron Collider.

Contents

1. Theoretical introduction	1
1.1. Standard model	1
1.1.1. Bosons	1
1.1.2. Fermions	2
1.1.3. Vices and virtues	4
1.2. The Higgs boson	4
1.2.1. Gauge invariance	5
1.2.2. The Higgs mechanism	6
1.2.3. Main production modes	9
1.2.4. Decay modes	10
1.3. The top quark	10
1.4. Associated production of single top quark and Higgs boson	10
1.4.1. tHq	12
1.4.2. tWH	13
1.4.3. tHb	14
1.4.4. CP mixing state	14
1.4.5. Experimental status	14
2. Conceptual overview	19
2.1. Cross section calculation	19
2.1.1. Short distance calculations	19
2.1.2. Renormalization and factorization scales	20
2.1.3. Cross sections at NLO in QCD	21
2.1.4. Description of b initiated processes	22
2.2. Monte-Carlo event generators	23
2.2.1. MadGraph5_aMC@NLO	24
2.2.2. Powheg	26
2.3. Parton shower and hadronization	27
2.3.1. Pythia6/8	28
2.3.2. Herwig++	28
2.4. Multivariate analysis tools	29
2.4.1. Artificial neural networks	29
2.4.2. Boosted decision trees	31
2.4.3. Interpreting training results: overtraining and ranking of variables	32
2.5. Fitting and statistical inference	33
2.5.1. Parameter estimation with the maximum-likelihood method	33

Contents

2.5.2.	Incorporation of uncertainties by means of nuisance parameters .	34
2.5.3.	Hypothesis testing, likelihood ratio and p value	34
2.5.4.	CL_S exclusion limits	35
3.	Experimental setup	39
3.1.	Large Hadron Collider	39
3.1.1.	Experiments placed along the LHC ring	42
3.1.2.	LHC operation	43
3.2.	Compact Muon Solenoid	44
3.2.1.	Tracker	46
3.2.2.	Electromagnetic calorimeter	47
3.2.3.	Hadron calorimeter	49
3.2.4.	Muon system	51
3.2.5.	Trigger system	54
3.2.6.	Detector simulation	55
3.2.7.	Computing infrastructure	55
3.3.	Reconstruction of events	56
3.3.1.	Hit reconstruction	57
3.3.2.	Track reconstruction	57
3.3.3.	Vertex reconstruction	59
3.3.4.	Muon candidates	59
3.3.5.	Electron candidates	60
3.3.6.	Photon and hadron candidates	61
3.3.7.	High-level objects	61
4.	Search for associated production of single top quark and Higgs boson at 8 TeV	67
4.1.	Channel topology and search strategy	67
4.2.	Object definitions	70
4.2.1.	Primary vertices	70
4.2.2.	Muons	71
4.2.3.	Electrons	72
4.2.4.	Jets	72
4.2.5.	Missing transverse energy	73
4.2.6.	W boson reconstruction	73
4.3.	Selection of events	74
4.3.1.	Consequences for signal and background modelling	75
4.3.2.	Controlling the multi-jet background	76
4.4.	Corrections to simulation and data	77
4.4.1.	Pile-up reweighting	77
4.4.2.	Top p_T reweighting	77
4.4.3.	Correction of b tagging efficiencies	78
4.4.4.	Lepton identification	80
4.4.5.	Jet η treatment	80

4.5. Event reconstruction	81
4.5.1. Reconstruction of events under the tHq hypothesis	82
4.5.2. Reconstruction of events under the $t\bar{t}$ hypothesis	83
4.6. Event classification	86
4.6.1. Choice and optimization of NN input variables	86
4.7. Systematic uncertainties	94
4.7.1. Theoretical uncertainties	100
4.7.2. Experimental uncertainties	102
4.8. Statistical uncertainties	104
4.9. Discussion of systematic uncertainties	104
4.10. Results	106
4.11. Possible improvements to the $H \rightarrow b\bar{b}$ analysis	110
4.11.1. Inclusion of the 2 tag region in the limit calculation	110
4.11.2. Using the full CSV shape information	111
4.12. Combination with other decay channels	114
4.13. Chapter summary	115
5. Prospects for a search for associated production of single top quark and Higgs boson at 13 TeV	117
5.1. Changes in signal and background modelling	118
5.2. Phenomenological studies	120
5.3. Event selection	123
5.4. Event reconstruction	125
5.5. Event classification	129
5.6. Systematic uncertainties	130
5.7. Results	141
5.8. Chapter summary	144
6. Summary and outlook	147
A. Appendix: Search for single top quark + Higgs boson at 8 TeV	151
A.1. Recorded and simulated datasets	151
A.1.1. Corrections	153
A.2. MVA training	154
A.2.1. Validation figures	154
A.2.2. Used settings	162
A.3. Pre-fit distributions	162
A.4. Pulls for fit to Asimov data	163
B. Appendix: Search for single top boson + Higgs boson at 13 TeV	165
B.1. Scalings and couplings	165
B.2. Production cross sections at NLO in QCD	166
B.3. Recorded and simulated datasets	168
B.4. BDT output shapes for the SM scenario	169

B.5. Nuisance pulls for an $s + b$ fit to an Asimov dataset	170
B.6. Expected limits for all 51 signal scenarios	171

1. Theoretical introduction

The purpose of this first chapter is to discuss the physics of the two heaviest particles of the standard model of particle physics (SM) – the top quark and the Higgs boson – and the relation between them. Discovered at the Tevatron and the Large Hadron Collider in 1995 [1,2] and 2012 [3,4], respectively, they mark the missing pieces that were needed for completion of the picture of quark generations and electroweak symmetry breaking as foreseen in the SM. At the same time, new physics beyond this established theory, generically referred to as beyond-Standard-Model (BSM) physics, is likely to manifest itself in the area of “top-Higgs” physics. The latter therefore is an important field to test and probe the SM and check for possible deviations from its predictions.

The topic of this thesis, the previously unexplored associated production of a Higgs boson and a single top quark, has many interesting aspects up its sleeve: firstly, it provides a special handle on one of the key parameters for the verification of the SM, namely the coupling strength of the Higgs boson to the top quark, and its sign relative to the coupling parameter characterizing the interaction with gauge bosons. In addition, BSM models, such as the scenario of CP violation in the top-Yukawa coupling, can be investigated in this channel. These aspects, which are unique to tHq production, are explained at the end of this introduction. The chapter starts with a general introduction to the SM, the Higgs boson and the top quark.

1.1. Standard model

As of now, the SM is the best theory for describing the interactions between elementary particles. It has been developed over decades in the second half of the 20th century and was tested experimentally to highest precision, so far revealing no single sign of discrepancy between its predictions and corresponding measurements. The particle content of the SM is briefly described below, followed by a discussion about what missing pieces in this theory are and about how it impressively succeeds in describing the generation of masses for the elementary particles.

1.1.1. Bosons

Table 1.1 lists the mediators of the forces that are described by the SM. Due to their bosonic nature, the so-called gauge bosons – a terminology that will become clear in Sec. 1.2 – have an integer spin of 1 (natural units with $\hbar = c = 1$ are used). They can be statistically described by means of Bose-Einstein statistics, which implies that arbitrarily many of them can have the same quantum numbers and occupy the same state.

1. Theoretical introduction

Table 1.1.: Interactions, gauge bosons and some of their properties [5], according to the SM. Although being separated into electromagnetic and, respectively, weak interactions in this table, the SM unites the two as a common electroweak interaction. This is for instance reflected in the fact that the Z boson couples to a mixture (as in a linear combination) of electric charge and weak charge, i.e. the third component of the weak isospin.

Interaction	Gauge boson	Mass [GeV]	Couples to
electromagnetic	photon (γ)	-	electric charge
strong	gluon (g)	-	color charge
weak	W^\pm	80.385	weak isospin
weak	Z	91.188	electric charge & weak isospin

Gauge bosons interact with other particles by coupling to specific charges: the photon, i.e. the carrier of the electromagnetic (EM) force, couples to particles with a non-zero electric charge. Since the photon is massless, the EM field propagates over long ranges (one can easily get a feeling for the large distance interaction by looking at the stars in the sky). Foundations of a quantum field theoretical formulation of the force, referred to as quantum electrodynamics or QED, were laid around 1950 in references [6–12].

Just as the photon, the gluon also has zero rest mass. However, the force it transmits is bound to act over very short ranges due to the property of confinement: gluons couple to the color charge of a particle, and carry both color and anticolor themselves. Their self-interaction makes the field lines between two colored objects form a string, whose energy increases linearly with the distance between the two objects. Already at very small distances (typically the size of a nucleon), it is energetically better for the system to create new quark-antiquark pairs than to further stretch the string. The relativistic quantum field theory of the strong interaction is called quantum chromodynamics (QCD).

In the SM, transitions between fermions, which are described in the next section, are mediated via the W^\pm bosons. These couple to the weak isospin of a particle. For instance, they are responsible for mediating the β decay of nucleons. QED and the theory of weak interactions have been formally unified by the electroweak theory in the 1960's [13–15]. Within the scope of the latter, weak neutral currents mediated by the later observed Z boson have been predicted. The Z boson couples to a linear combination of the electric charge q of a particle and its isospin I_3 . The quantity $2q - I_3$ is referred to as weak hypercharge Y .

With masses of 80.4 and 91.2 GeV for the W and Z bosons [5], the carriers of the weak force have a short life time of about 10^{-25} s, and therefore can act only at short ranges. The discussion of the Higgs boson is left to Sec. 1.2.

1.1.2. Fermions

The Pauli exclusion principle dictates that two fermions, i.e. particles with half-integer spin, can never have all quantum numbers the same, which entails an asymmetric combined wave function for two identical fermions. The consequence of this behaviour is

that nature can “build things” out of them: they constitute the matter that we are familiar with from our everyday life. Atoms consist of electrons orbiting around nuclei, which in turn are comprised of u and d quarks. During the last century, a multitude of additional, increasingly heavy fundamental particles has been discovered in the laboratory and through cosmic rays, culminating in the observation of the top quark about twenty years ago. As it turns out, the fermions of the SM all have spin 1/2 and can be arranged in the following way:

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix} \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix} \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix},$$

$$\begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix}.$$

For each of the particles there exists also an antiparticle with inverted charges. The above representation suggests that the structure in which the leptons (the charged e , μ and τ and the corresponding, electrically neutral neutrinos) and the six quarks are ordered corresponds to three generations of fermions. The heavier objects to the right, although tending to stay within their generation at first order, will eventually decay (via W bosons) to the group of stable particles formed by the first, lightest generation. The mass hierarchy problem (the world average of top quark mass measurements gives $m_t = 173.34$ GeV [16]; this is about 75 000 times larger than the one of the u quark) cannot be explained with the SM, which does not predict the masses of the fermions; they are free parameters and must be measured in experiments. Within a generation they are arranged in doublets of the weak charge (= weak isospin), and W bosons mediate transitions between the particles within a doublet.¹

While the electrically charged leptons interact via the electromagnetic and weak forces, neutrinos only carry weak isospin and therefore only take part in the weak interaction.

Quarks carry color charge and thus are subject to the strong interaction; in fact, a quark can have three different colors – red, blue or green – and cannot be observed as a free particle but only in bound states with zero net color charge. These can either be quark-antiquark pairs (mesons) or states formed by three (anti)quarks (baryons). This is owed to the concept of confinement which will be briefly described later. It is also reflected in the fact that only integer numbers of electric charge (in units of the electron charge) can be measured, despite the quarks having charges of $+2/3e$ and $-1/3e$ for the “up”- and “down” types. Finally, quarks also interact weakly; a subtlety worth mentioning is that the quarks’ mass eigenstates (in the creation due to the strong force) are mixing in order to form the quark eigenstates participating in the weak interaction. The mixing is parametrized in the so-called Cabibbo-Kobayashi-Maskawa (CKM) matrix [17, 18].

¹It should be noted only left-handed fermions – that is the $1 - \gamma^5$ projection of their spinor, where γ^5 is given by the four gamma matrices, $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$ – carry weak isospin and consequently take part in the weak interaction mediated by W bosons. This implies that also the coupling strength of the Z boson to fermions depends on their chirality.

1. Theoretical introduction

1.1.3. Vices and virtues

As has become apparent from Tab. 1.1, the SM does not provide a field-theoretical description of the gravitational force. This is owed to the fact that the attempt to quantize the gravitational field and introduce the hypothetical graviton leads to a theory with divergences that cannot be renormalized. For this reason, General Relativity [19] is still the common theory to describe gravity. The observation of neutrino oscillations [20, 21] implies the existence of massive neutrinos, for which there is no mechanism in the SM. Furthermore the phenomenon of dark matter (DM), which supposedly is comprised of weakly interacting particles, cannot be attributed to any of the SM particle types, at least not to the degree that is needed to explain the expected DM density in the universe indirectly inferred from cosmological observations such as gravitational lensing and rotational curves of galaxies. Dark energy, being responsible for both the expansion and the bulk of energy of the universe, is another mystery the SM does not have an answer for.

Let aside these shortcomings, the SM has been incredibly successful both in the description of precision measurements such as the one of the fine structure constant $\alpha_{e.m.}$ and in the prediction of new particles that were necessary to complement the established ones. For instance, after the discovery of the first quark (b) of the third fermion generation it was clear there would exist a heavy, sixth quark (the later observed top quark). The validity of the SM at the energy scales accessible in the laboratory therefore is striking.

Formally, the interactions are described by means of an underlying $U(1)_Y \times SU(2)_L \times SU(3)_C$ symmetry preserved under local transformations. According to Noether's theorem [22], the symmetries lead to quantities that are conserved in the respective interaction. In the case of the SM, these correspond to the weak hypercharge Y , the weak isospin, where L indicates that only left-handed states participate in the weak interaction mediated by W bosons, and the color charge C . For each of the generators of the symmetry groups (one for $U(1)_Y$, three for $SU(2)_L$ and eight for $SU(3)_C$) there is an associated field, all of which per se are massless and combine linearly in order to form the physical fields discussed above, i.e. the γ , W^\pm , Z and gluon fields with different colors.

However the symmetry is partly broken by a ground state not sharing the transformation properties of the symmetry group. As will be shown in the next section, this has deep implications on the particle masses, which are created dynamically in the process of symmetry breaking.

1.2. The Higgs boson

The particle associated with the eponymous Higgs field – the Higgs boson – was predicted in 1964 [23–26] and arises in the process of spontaneously breaking the electroweak symmetry that leaves the photon massless but gives masses to the weak gauge bosons W^\pm and Z . In the following the concept of the so-called Higgs mechanism is outlined. With the discovery of a boson fully compatible with the one predicted by this mechanism in 2012, an ultimate proof for the existence of a Higgs boson-like particle has been

delivered. Peter Higgs and François Englert consequently received the Nobel Prize in Physics in the subsequent year.

1.2.1. Gauge invariance

The way the Higgs mechanism works cannot be understood without the principle of gauge invariance: in the theoretical formulation of the SM the interactions originate from the requirement that the Lagrangian density \mathcal{L} may not change when the phase of ϕ is altered, which is referred to as gauge transformation. This concept of gauge invariance is briefly illustrated with the example of quantum electrodynamics (QED) and the $U(1)$ symmetry group in the following.

The Lagrangian density from which the Dirac equation describing massive spin-1/2 particles follows, $i\gamma^\mu(\partial_\mu\psi) - m\psi = 0$, is given by

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi. \quad (1.1)$$

Here, ∂_μ is the 4-gradient ∂/∂^μ and γ^μ are the Dirac matrices. The Lagrangian is invariant under the global $U(1)$ transformation of ψ with a constant phase θ :

$$\begin{aligned} \psi &\rightarrow \psi' = e^{i\theta}\psi, \\ \bar{\psi} &\rightarrow \bar{\psi}' = e^{-i\theta}\bar{\psi}, \\ \mathcal{L} &\rightarrow \mathcal{L}' = \mathcal{L}. \end{aligned}$$

Starting from the last equation, i.e. from $\delta\mathcal{L} = \mathcal{L}' - \mathcal{L} = 0$, it can be demonstrated that there is a current given by $j^\mu = \bar{\psi}\gamma^\mu\psi$ which fulfills the continuity equation $\partial_\mu j^\mu = 0$ and thus corresponds to the conservation of electric charge under global $U(1)$ transformations.

Causality demands that a local transformation of the field at point x in the phase space, i.e. $\theta(x) \equiv q \cdot \chi(x)$, does not change \mathcal{L} either, and information about a changed phase $e^{i\theta(x)}$ at one point should propagate to another position with finite speed. This is the reason why a global transformation of ψ with an instantaneous effect on all points in phase space cannot be regarded physical and local gauge invariance is required. However, the Lagrangian in Eq. 1.1 has no local $U(1)$ gauge symmetry under $\psi \rightarrow \psi' = e^{iq\chi(x)}\psi$, unless one introduces a term with a new field A_μ and a covariant derivative D_μ which transform in analogous ways,

$$\begin{aligned} A_\mu &\rightarrow A'_\mu = A_\mu - \partial_\mu\chi, \\ \partial_\mu &\rightarrow D_\mu = \partial_\mu + iqA_\mu. \end{aligned}$$

The photon field A_μ propagates with finite speed and thus preserves causality. The invariant Lagrangian of QED consequently reads

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi + iq\bar{\psi}\gamma^\mu\psi A_\mu - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}.$$

The first term describes the dynamics of a fermion with charge q , whose interaction with the field A_μ is controlled by the second term. Finally the kinematics of the photon field

1. Theoretical introduction

are described by a term quadratic in the anti-symmetric electromagnetic field-strength tensor given by $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. Notably there is no term quadratic in A_μ , whose prefactor could be interpreted as a mass of the photon field, as such a term would spoil local gauge invariance. Consequently the mass of the photon is zero.

1.2.2. The Higgs mechanism

In principle the same argumentation of local gauge invariance in QED applies to the $U(1)_Y \times SU(2)_L \times SU(3)_C$ symmetry groups describing the electroweak and QCD interactions. However, while the requirement for massless gauge bosons works well for the photon and the gluon fields, transferring the concept to the gauge bosons of the weak interaction strongly contradicts the observation of massive W and Z bosons (discovered in the 1980's, see [27–29]). The Higgs mechanism solves this issue by dynamically creating the mass terms for the weak gauge bosons whilst keeping the photon massless and retaining local gauge invariance.

The covariant derivative, which describes the electroweak interaction and accordingly acts on left-handed fermions, reads

$$D_\mu = \partial_\mu - ig \frac{\sigma^i}{2} W_\mu^i - ig' \frac{Y}{2} B_\mu, \quad (1.2)$$

where σ is the vector of the three Pauli matrices, Y is the weak hypercharge, g and g' are the respective coupling constants and W and B the fields of the electroweak interaction.

A new $SU(2)_Y$ doublet field Φ is introduced, which in its most general form can be written as

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}$$

and transforms according to

$$\Phi \rightarrow \Phi' = \exp(i(g\sigma^i w^i(x) + g'Yb(x))/2)\Phi, \quad (1.3)$$

where $w^i(x)$, $i = 1, 2, 3$ and $b(x)$ are real differentiable functions and are the same to the fields W and B what χ has been to the photon field A in the previous section. In the electroweak Lagrangian the new field Φ enters via a kinetic term and a potential,

$$\mathcal{L}_{\text{EW}} \propto (D_\mu \Phi)^\dagger D_\mu \Phi - V(\Phi^\dagger \Phi),$$

where D_μ is given by Eq. 1.2.

Spontaneous symmetry breaking

The potential V of the field Φ has the form

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2.$$

In order for it to have a finite minimum at all, it is required that $\lambda > 0$. On the other side, the potential has the property that for $\mu^2 < 0$, it has infinitely many degenerate minima, all satisfying $\Phi^\dagger \Phi = -\mu^2/(2\lambda)$. This symmetry is broken by picking a ground state, e.g. $\Phi_0 = 1/\sqrt{2} \begin{pmatrix} 0 \\ v \end{pmatrix}$, where $v = \sqrt{-\mu^2/\lambda}$ is called the vacuum expectation value.

Small perturbations about this minimum can be parametrized by

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1(x) + i\phi_2(x) \\ v + \eta(x) + i\phi_4(x) \end{pmatrix}.$$

As is shown e.g. in [30], terms with the fields $\phi_{1,2,4}(x)$ can be eliminated from the Lagrangian of the electroweak sector by choosing appropriate gauge transformations for the fields \mathbf{W} and B . In the so-called unitary gauge, the field Φ has the form

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}. \quad (1.4)$$

The gauging makes the fields $\phi_{1,2,4}(x)$ vanish, but it leaves terms that are quadratic in the physical fields $W^\pm = \frac{1}{\sqrt{2}} (W^1 \mp iW^2)$ and $Z = (W^3 g - Bg')/(g^2 + g'^2)^{1/2}$; consequently the corresponding bosons have received masses of $m_W = gv/2$ and $m_Z = v\sqrt{g^2 + g'^2}/2$, respectively. Furthermore, the Higgs field $H(x)$ itself has received quadratic terms and a mass equal to $m_H = \sqrt{2\lambda}v$. This is identified with the Higgs boson mass.

It should be emphasized that the ground state Φ_0 shares the symmetry of the part of the gauge transformation in Eq. 1.3 that corresponds to the electric charge. Hence, the symmetry group describing interactions of electric charges is unbroken, and the photon remains massless. Spontaneous symmetry breaking is illustrated in Fig. 1.1 for the somewhat simpler example of a complex scalar field $\phi = 1/\sqrt{2}(\phi_1 + i\phi_2)$.

Coupling to fermions²

In Sec. 1.2.1 a fermion mass term $m_f \bar{\psi}\psi$ was part of the QED Lagrangian. However, such a term is not invariant under transformations of the generalized $SU(2)_L \times U(1)_Y$ symmetry group. This can be seen by decomposing the term into different chiral states:

$$m_f \bar{\psi}\psi = m_f (\bar{\psi}_R + \bar{\psi}_L) (\psi_R + \psi_L) = m_f [\bar{\psi}_R \psi_L + \bar{\psi}_L \psi_R] \quad (1.5)$$

because of $\bar{\psi}_R \psi_R = \bar{\psi}_L \psi_L = 0$. ψ_L is an isospin doublet with isospin 1/2; ψ_R is an isospin singlet because its isospin is 0. Under the rotations defined in Eq. 1.3 they behave according to

$$\begin{aligned} \psi_L &\rightarrow \psi'_L = \exp(i(g\sigma^i w^i(x) + g'Yb(x))/2) \psi_L \\ \psi_R &\rightarrow \psi'_R = \exp(ig'Yb(x)/2) \psi_R, \end{aligned}$$

²This part is based on reference [31], which in turn mostly relies on [32].

1. Theoretical introduction

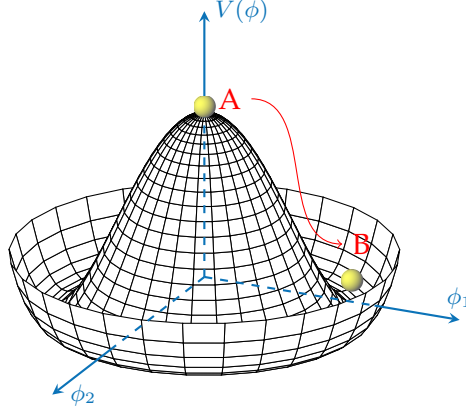


Figure 1.1.: The potential $V(\phi) = \mu^2 \phi^* \phi + \lambda (\phi^* \phi)^2$ is shaped like a Mexican hat and has degenerate minima for $\phi_1^2 + \phi_2^2 = -\mu^2/\lambda \equiv v^2$. Without loss of generality one can choose the minimum (B) to be purely real, e.g. $\phi_1 = v, \phi_2 = 0$, and expand in two directions via $\phi(x) = 1/\sqrt{2}(v + \eta(x) + i\xi(x))$. Perturbation along the real axis, i.e. in a direction where the shape of the potential is to first order quadratic, reveals a boson of mass $m = \sqrt{2\lambda}v^2$. Excitations along the direction where the potential has the same value correspond to a massless particle, a so-called Goldstone boson. Requiring gauge invariance under a local $U(1)$ symmetry allows to “gauge this boson away”, at the cost that a term corresponding to the longitudinal polarization of a new field, i.e. to a non-vanishing mass term, appears in the Lagrangian. The field can be identified with a massive gauge boson like the W or Z boson.

and therefore Eq. 1.5 is not invariant under $SU(2)_L \times U(1)_Y$ transformations of ψ . However, it can be shown that by introducing the field Φ of the previous section, there arise terms in the Lagrangian that transform like a singlet under $SU(2)_L \times U(1)_Y$ and have the structure

$$\mathcal{L} \propto -y_f [\bar{\psi}_R \Phi \psi_L + \bar{\psi}_L \Phi \psi_R],$$

where y_f is the so-called Yukawa coupling. Since Φ has a non-zero vacuum expectation value (see Eq. 1.4), terms appear that are proportional to $y_f v \bar{\psi}_{R/L} \psi_{L/R}$ and the mass of the fermion is consequently given by $m_f = 1/\sqrt{2} y_f v$. A subtlety lies in the flavor (i.e. the weak) eigenstates of quarks that are different from their mass eigenstates. Again, this mixing is expressed with the CKM matrix.

Summarizing this important concept, the Yukawa coupling mechanism predicts that masses of fermions are created by the interaction of fermions with the vacuum expectation value of the Higgs field, and that the coupling strength of the Higgs boson to fermions is proportional to the fermion mass:

$$y_f = \sqrt{2} \frac{m_f}{v}.$$

As previously mentioned, the theory does not predict the exact values of the masses. For a measured top quark mass of 173.34 GeV however, this curiously results in $y_t \simeq 1$.

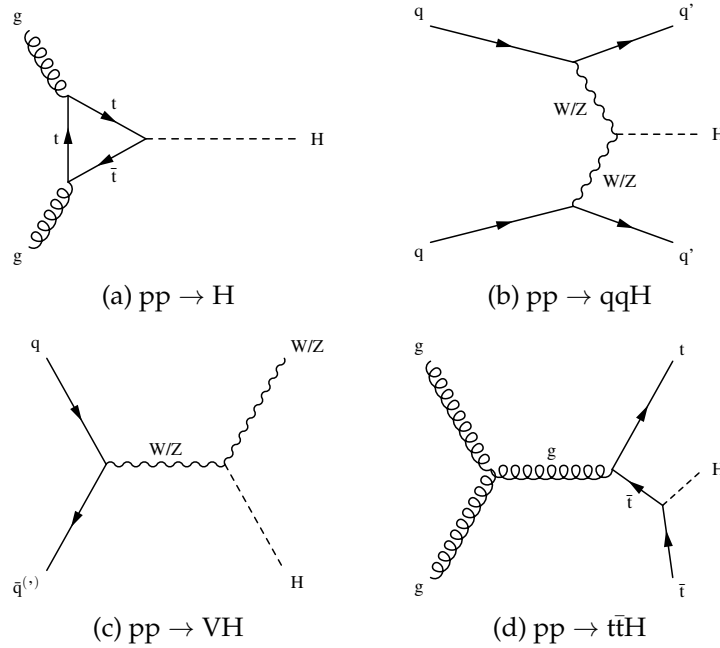


Figure 1.2.: Leading order Feynman diagrams of the most relevant Higgs boson production modes at the LHC, ordered by decreasing cross section: gluon-fusion (a), vector boson-fusion (b), Higgsstrahlung (c) and production in association with a top quark pair (d).

1.2.3. Main production modes

Apart from the associated production with a top quark (which in fact has a tiny cross section as will be shown later), the Higgs boson has four main production modes at the LHC. They are depicted in the Feynman diagrams of Fig. 1.2.

The high gluon densities in the proton make the gluon-fusion process the predominant production mode. As gluons themselves are massless and hence do not couple to the Higgs boson, a virtual top quark loop leads to an effective ggH coupling. Other quarks are also allowed to run in the loop; however their contribution is suppressed by at least a factor of $m_b^2/m_t^2 \sim 0.0007$. In experimental searches the gluon-fusion process is combined with a $H \rightarrow \gamma\gamma$ or $H \rightarrow 4\ell$ decay (in the latter case there is an intermediate state of two Z bosons) in order to have a clean signature to suppress background processes.

Vector boson fusion is the production mechanism with the second largest cross section. The Higgs boson is accompanied by two light quarks, whose presence and kinematics can be used to reject background processes.

While the Higgsstrahlung process in Fig. 1.2(c) will not be discussed here, the associated production with a top quark pair is particularly relevant, as it has a final state similar to the one that is looked after in Chap. 4. More information on the characteristics of this process can be found there.

1. Theoretical introduction

1.2.4. Decay modes

The main decay mode is the one exploited in the analysis of this thesis: around 58% of the produced Higgs bosons decay into a pair of b quarks, as they are the heaviest objects that are kinematically allowed. The Higgs boson may also decay into a pair of vector bosons. However, the available phase space is very narrow as one of them must be a virtual, very off-shell W or Z boson; this type of decay is therefore suppressed. As previously mentioned, $H \rightarrow \gamma\gamma$ or $H \rightarrow$ multi-lepton signatures provide the cleanest channels to look after Higgs boson production. They will, together with $H \rightarrow \tau^+\tau^-$, play a role at the very end of Chap. 4 in a combination of results.

1.3. The top quark

The top quark is the heaviest elementary particle in the SM. At the LHC it is mainly produced in $t\bar{t}$ pairs in gluon-fusion processes (see Fig. 1.3). Because of the large gluon densities in a proton accelerated to high energies, the $t\bar{t}$ process has a large production cross section and constitutes the main background in the analysis presented in this thesis. In addition there is the electroweak single top quark production, which is mediated by a tWb vertex; according to the virtuality of the W boson, it can be divided into t -channel-, tW - and s -channel-like diagrams (ordered by decreasing cross section). However, the separation into t -channel, tW and s -channel makes most sense at leading order in the strong coupling constant α_s . At next-to-next-to-leading order at the latest, the definitions are not unambiguous anymore, and t - and s -channels start to interfere.

The mass and width of its Breit-Wigner resonance is large enough so that the top quark decays before it can hadronize to a bound meson state or similar. In almost 100% of the cases, it will decay into a b quark and a W boson. The signature of top quark decays is therefore dominated by the presence of a b quark and, depending on the decay of the W boson, by additional light flavor quarks or a charged lepton and a neutrino.

1.4. Associated production of single top quark and Higgs boson

This section motivates the search for the associated production of a single top quark and a Higgs boson (tH) at the LHC. It starts with explicitly addressing the properties of the tHq and tWH channels and mentioning their most important aspects, and finally summarizes the experimental status in regard to the question where tH production can provide interesting information, in particular concerning the lifting of degeneracies in Higgs boson couplings. A theoretical introduction and phenomenological studies on tH production, serving as basis and motivation for large parts of the analysis performed in this thesis, are provided in [33–35]. The first papers investigating the feasibility of this channel can be found in [36–38].

1.4. Associated production of single top quark and Higgs boson

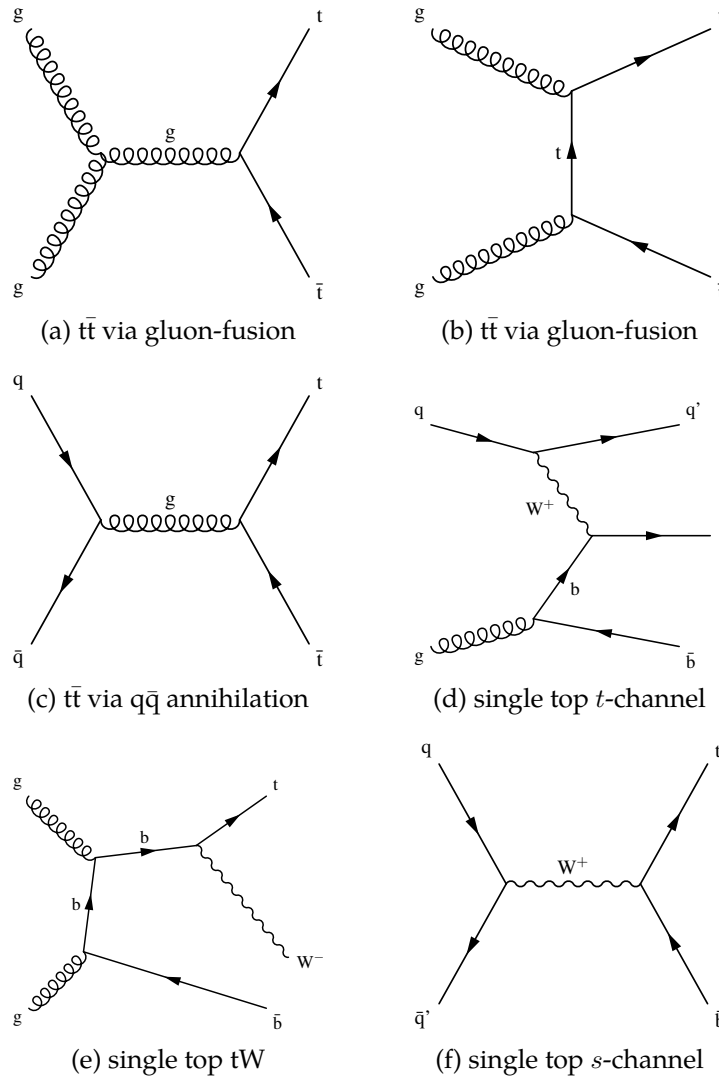


Figure 1.3.: Leading order Feynman diagrams of the most relevant top quark production modes at the LHC.

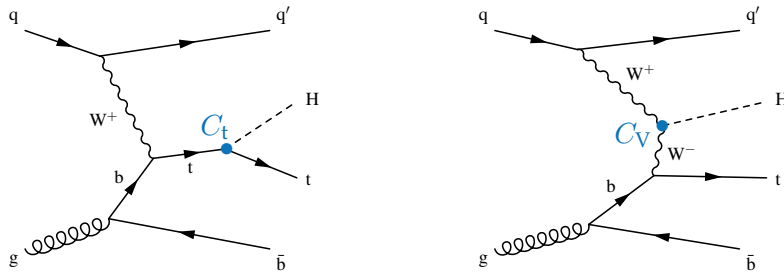


Figure 1.4.: Representative t -channel LO Feynman diagrams for the associated production of single top quark and Higgs boson.

1. Theoretical introduction

1.4.1. tHq

As is shown in the Feynman diagrams of Fig. 1.4, the major contribution to tHq production comes from t -channel single top processes, where the Higgs boson is either attached to the top quark leg or to the virtual W boson.³ In order to discuss the key aspect of this production, it is useful to introduce the two real scaling factors

$$C_f = y_f/y_f^{\text{SM}}, \quad C_V = g_{\text{HWW}}/g_{\text{HWW}}^{\text{SM}},$$

with y_f being the top-Yukawa coupling and g_{HWW} the Higgs boson coupling to W bosons. These scaling factors allow for a generalized modification of the Higgs boson couplings to the top quark and the W boson with respect to the SM values.

The Mandelstam variables characterizing the $Wb \rightarrow tH$ scattering are given by:

$$\begin{aligned} s &= (p_W + p_b)^2, \\ t &= (p_W - p_H)^2, \\ u &= (p_W - p_t)^2. \end{aligned}$$

This breakdown to the core scattering process, ignoring the upper vertices in the Feynman diagrams, is a valid approach at high energies, where in the effective W approximation [39] one treats W bosons that appear in intermediate states of a scattering process as partons in the proton. According to [34], the amplitude in the hard-scattering regime with $s, -t, -u \gg m_t^2, m_W^2, m_H^2$ is then given by

$$\mathcal{A} = \frac{g}{\sqrt{2}} \left[(C_f - C_V) \frac{m_t \sqrt{s}}{m_W v} A \left(\frac{t}{s}, \varphi; \xi_t, \xi_b \right) + \right. \quad (1.6)$$

$$\left. \left(C_V \frac{2m_W s}{v} \frac{s}{t} + (2C_f - C_V) \frac{m_t^2}{m_W v} \right) B \left(\frac{t}{s}, \varphi; \xi_t, \xi_b \right) \right], \quad (1.7)$$

where v is the vacuum expectation value of the Higgs field. The exact expressions for A and B , which are dependent on the spinors $\xi_{t,b}$ of the top and bottom quark and the angle φ (the azimuthal angle of the Higgs boson in the reference frame centered at the partonic collision point, where the z axis points along the beam direction), can be found in [34]. It is easy to see in the above equation that there is an interference between the two leading order Feynman diagrams. For the SM case of $C_V = C_t = 1$, it is of destructive nature, and the first term cancels out and with it its linear dependence on the energy of the outgoing system, \sqrt{s} .⁴ This means that the SM cross section falls with the square of the invariant mass $1/s$, as one would expect from a $2 \rightarrow 2$ hard inelastic scattering process. On the other hand, for any scenario with $C_V \neq C_f$, the amplitude increases with \sqrt{s} and therefore is enhanced with respect to $C_V = C_f$. As is demonstrated in [34], this does not pose a problem to perturbative unitarity until a cutoff scale $\sqrt{s} = \Lambda$ with

$$\Lambda = 12\sqrt{2}\pi \frac{v^2}{m_t |C_f - C_V|}$$

³Diagrams with a Hbb or Hqq coupling contribute to a negligible amount and are therefore not considered throughout this thesis.

⁴The expression s/t in the second term is connected to the angle of the scattering and is kept finite [40].

1.4. Associated production of single top quark and Higgs boson

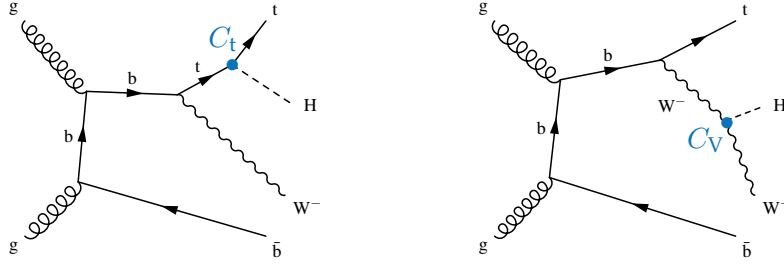


Figure 1.5.: Representative LO Feynman diagrams for the associated production of a single top quark, a Higgs boson and a W boson.

is reached. This scale has typical values of a few TeV; the invariant mass of the tH system usually is much lower, and therefore it is a relevant approach to handle the $C_V \neq C_t$ scenarios by means of an effective theory. Their cross sections are approaching constant values for increasing s because of the cancellation of proportionalities $(\sqrt{s})^2$ and $1/s$.

The SM cross section is very small due to the almost perfect cancellation between the two contributing diagrams and amounts to

$$\sigma_{\text{SM}}^{8\text{TeV}} = 18.28_{-0.38}^{+0.42} \text{ fb} \quad [34]$$

at $\sqrt{s} = 8 \text{ TeV}$. The uncertainties correspond to variations of the factorization and renormalization scales (see Sec. 2.1.2). For the case of a flipped sign for the top-Yukawa coupling, $C_t = -1$, the cross section is augmented to

$$\sigma_{C_t=-1}^{8\text{TeV}} = 233.8_{-0.0}^{+4.6} \text{ fb} \quad [34],$$

which means an enhancement by roughly a factor of 13. This brings the scenario into reach of current collider experiments, and is the basic motivation of the search presented in this thesis. A flipped top-Yukawa coupling is especially interesting as both tH and WH channels, although being sensitive to the magnitudes of C_t and C_V , respectively, are insensitive to their relative sign. The tHq channel therefore can help lifting degeneracies in the 2d plane spanned by the two parameters. There is another relevant feature that makes the search for exotic tHq production even more feasible: assuming a $H \rightarrow \gamma\gamma$ decay, the rate of this branching is further enhanced, because there is a similar interference occurring between the virtual top quark and W boson loops. The parametrization for the effective scaling factor $C_\gamma = g_{H\gamma\gamma}/g_{H\gamma\gamma}^{\text{SM}}$ is given in [41] and assumes a value of $\simeq 2.36$ for $C_t = -1$ and $C_V = 1$.

1.4.2. tWH

The associated production of t and H can also be accompanied by a real W boson. A feature of tWH production at next-to-leading order (NLO) in QCD is the interference with tH, which has the same W^+bW^-bH final state upon the top quark decay in the Feynman diagrams of Fig. 1.5. Several methods exist to eliminate the overlap with tH. A diagram removal technique has been used to derive the cross section at NLO in QCD for the first time in the course of this thesis. The cross section is found to be $\sim 16 \text{ fb}$.

1. Theoretical introduction

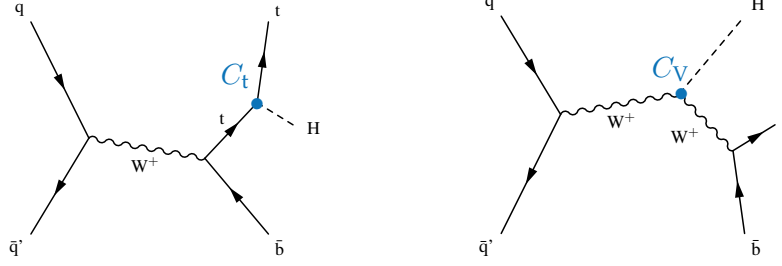


Figure 1.6.: Representative s -channel LO Feynman diagrams for the associated production of single top quark and Higgs boson.

1.4.3. tHb

The associated production of H boson and top quark in an s -channel diagrams is depicted in Fig. 1.6 for reference. It is greatly suppressed due to the highly virtual, very off-shell intermediate W boson. The cross section has been calculated in [35] for the SM at $\sqrt{s} = 13$ TeV and is so small (2.813 fb) one can safely neglect it in the search for tHq .

1.4.4. CP mixing state

In reference [35] and earlier studies [42–44], top-Yukawa couplings of a generic spin-0 particle, X_0 , that violate the CP symmetry have been investigated. The effective Lagrangian for such an interaction can be written as [45]

$$\mathcal{L}_0^t = -\bar{\psi}_t (\cos \alpha \kappa_{Htt} g_{Htt} + i \sin \alpha \kappa_{Att} g_{Att} \gamma_5) \psi_t X_0, \quad (1.8)$$

where α is the parametrization for the mixing between CP -even and CP -odd top-Higgs couplings and $g_{Htt} = g_{Att} = m_t/v$. Accordingly, the $C_V = 1$ ($C_t = \pm 1$) case can be recovered by setting $\alpha = 0^\circ$ ($\alpha = 180^\circ$), while a purely CP -odd interaction requires $\alpha = 90^\circ$. The tHq channel is in particular interesting in this regard; as the left figure of Fig. 1.7 shows, its cross section even exceeds the one of $t\bar{t}H$ from a certain value of α on. Because it behaves asymmetrically under a rotation in the scalar-pseudoscalar plane, it can help in lifting degeneracies which are there for the production of a Higgs boson with a top quark pair and the gluon-fusion process, respectively. The right-hand distributions of the same figure show that also the shapes of the differential cross sections have a dependence on the mixing angle.

A search for CP violation in tHq is not performed in this thesis; however it will provide an interesting channel for exploring the nature of the top-Higgs interaction in future experimental studies with large integrated luminosities.

1.4.5. Experimental status

There exists a variety of published ATLAS and CMS measurements and combinations thereof that aim at constraining the allowed regions for the parameters C_t and C_V (often called $\kappa_{t,V}$ in the analyses) with the data collected at center-of-mass energies of 7 and 8 TeV.

1.4. Associated production of single top quark and Higgs boson

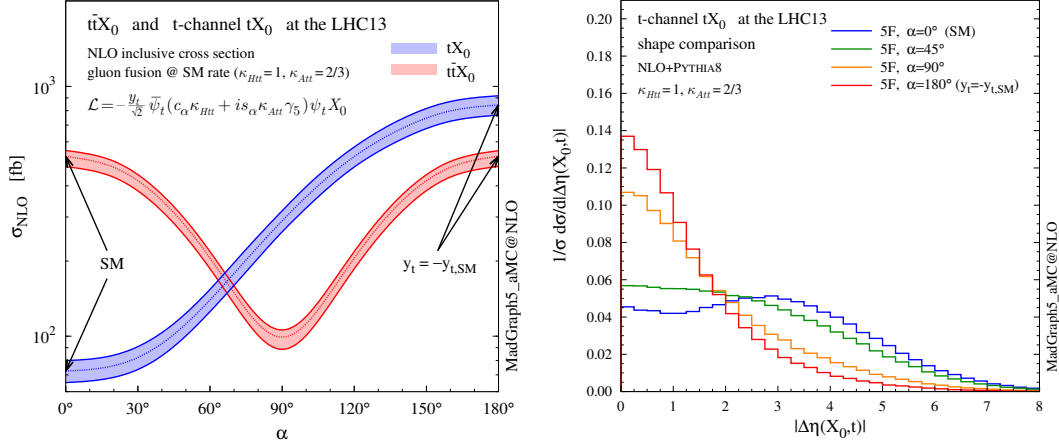


Figure 1.7.: Both inclusive (left) and differential (right) cross sections for a generic spin-0 particle produced in association with a single top quark have a strong dependence on the degree of mixing between CP -even and CP -odd components in the top-Yukawa coupling. Notably, the inclusive cross section is asymmetric under $90^\circ + \alpha \leftrightarrow 90^\circ - \alpha$, in contrast to the production together with a top quark pair. Taken from [35].

In [46, 47] global fits to the data are performed in order to simultaneously constrain the couplings of the boson at both the production and decay stage. Fig. 1.8 shows the results of a 2d likelihood scan in κ_f , κ_V . Only the $(+, +)$ and $(-, +)$ quadrants are shown, since the combined analyses are only sensitive to the relative sign between the two parameters by considering analyses looking for $H \rightarrow \gamma\gamma$, but they are insensitive to a distinction between $(-, +)$ and $(+, -)$.⁵ From the exclusion boundaries of the individual channels in the left figure one can deduce that, except for $H \rightarrow \gamma\gamma$, all decay channels do not bear interference terms at leading order and therefore have symmetric contours. As the best fit value suggests, the data favor $\kappa_f \simeq 1$ and therefore the SM scenario. This is mainly due to the incompatibility of the data with the aforementioned enhancement in $H \rightarrow \gamma\gamma$ of ~ 2.4 for $\kappa_t = -1$ and $\kappa_V = 1$. However the underlying assumption is that, apart from the modification of the coupling strengths, there is no physics beyond the SM involved, in particular there are no new particles running in the virtual loops which would alter the $H \rightarrow \gamma\gamma$ branching fraction. Fig. 1.9 shows two ways loop-induced processes can be treated in the global fit if one allows for such new loop contributions: the solid orange curve corresponds to the scenario where loop-induced couplings are fully resolved into their SM content. Then the profile likelihood ratio⁶, denoting the compatibility of the hypothesis under test with the observation, strongly disfavors $\kappa_t = -1$. However, if the loops are described by means of effective coupling parameters which are left independent and free to float in the fit, one has a minimal sensitivity on the sign of κ_t .

⁵A tHq, $H \rightarrow \gamma\gamma$ analysis will in principle be able to lift the degeneracies because the interference terms are present in the production as well as in the decay. However, large integrated luminosities will be needed to accomplish this task.

⁶An introduction to the profile likelihood ratio test statistic will be provided in Sec. 2.5.

1. Theoretical introduction

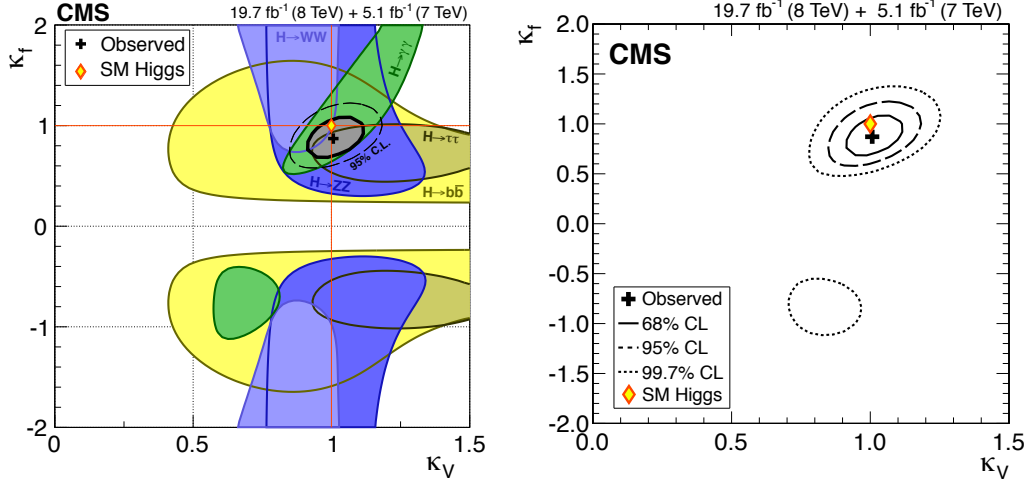


Figure 1.8.: Experimental constraints in the κ_t - κ_V plane from CMS data [47], for the individual channels (left) as well as for the global combination (right). The black cross marks the best fit value of the likelihood scan; contours in the right figure indicate the 1, 2 and 3 σ confidence regions, while in the left figure the colored areas correspond to the regions of 68% CL confidence.

This is indicated by the solid blue curve, which excludes a flipped top-Yukawa coupling sign only at 68% (the horizontal red line indicates the 1σ C.L. interval). This ATLAS measurement considers tHq as background in the $t\bar{t}H$ channel. Accordingly, the residual sensitivity on $\kappa_t = -1$ of the solid blue curve solely stems from the interference between the tree-level diagrams in Fig. 1.4. The CMS collaboration pursues a different approach that consists in dedicated searches for tHq production with $\kappa_t = -1$ in different decay channels. The $H \rightarrow b\bar{b}$ analysis is presented in Chap. 4; other decay channels that have been investigated are $H \rightarrow \tau^+\tau^-$, $H \rightarrow W^+W^-$ and $H \rightarrow \gamma\gamma$.⁷ In a paper combining the different tHq searches CMS puts an upper limit on the allowed cross section of anomalous tHq production [48]. The combination excludes a signal strength modifier $\mu = \sigma_{\text{obs}}/\sigma_{\kappa_t=-1} > 2.8$ at 95% C.L.

⁷All analyses only consider leptonic decays of the top quark.

1.4. Associated production of single top quark and Higgs boson

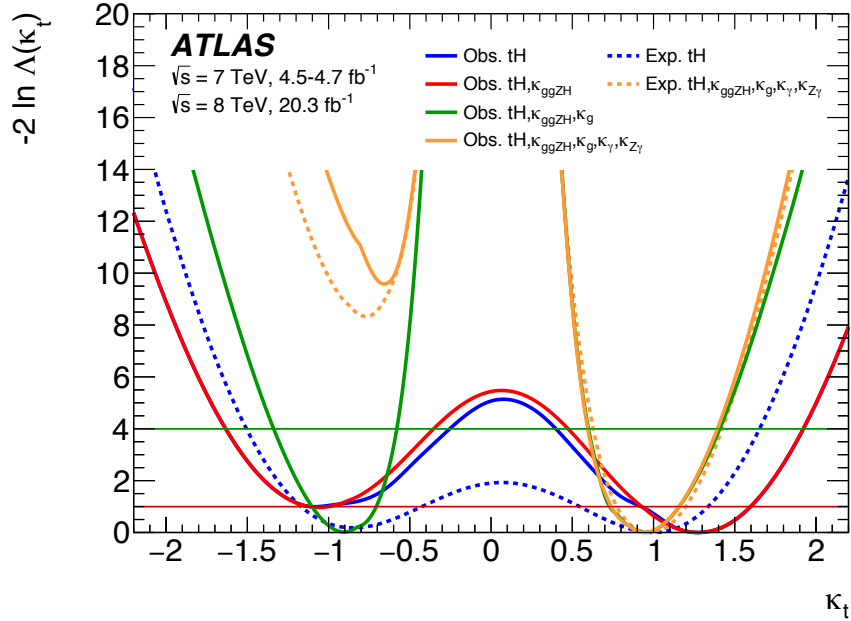


Figure 1.9.: Profile likelihood ratio of a global fit to ATLAS data as a function of κ_t [46]. The different curves describe scenarios in which different loop-induced processes can be resolved into their expected SM content. In the scenario of the solid orange line, loop processes of $gg \rightarrow H$, $gg \rightarrow ZH$ and $H \rightarrow Z\gamma$, $H \rightarrow \gamma\gamma$ can be fully resolved, and no BSM physics contributes to the loops. The free parameters in the fit correspond to all coupling strength scale factors to SM particles. The solid blue curve on the other hand disfavors κ_t only at 1σ (horizontal red line), as in this scenario interferences in loop couplings can no longer be exploited – i.e. no loop processes can be resolved – but effective scale factors have been introduced. The only sensitivity on $\kappa_t = -1$ comes from taking into account the tree-level interference effects of tHq, which is considered as background in the tH channel.

2. Conceptual overview

Conducting a search for a process like the associated production of a single top quark and a Higgs boson would not be feasible if it was not for special tools and concepts that have been established in the field of High Energy Physics over the last decades and are steadily being developed, improved and extended in their application, precision and functionalities.

This chapter introduces some of these, which are most relevant from the viewpoint of this thesis, and which have become key ingredients to both phenomenological and experimental LHC physics. The chapter starts with an overview about how cross sections and simulated events for a process can be obtained at next-to-leading order in QCD by means of Monte-Carlo event generators.¹ Considering the fact that the expected signal rate in tHq is very small, multivariate analysis tools (MVAs) like an artificial neural network or a boosted decision tree can construct a strong classifier variable which maximally separates between tHq and background processes. The different kinds of MVAs are explained in the middle part of the chapter. How to extract the signal strength from the distributions and finally set upper limits on the signal cross section is described in the last part.

2.1. Cross section calculation

2.1.1. Short distance calculations

Since high energies and large momentum transfers – or, equivalently, short distances – are involved in the hard interaction part of a pp collision at the LHC, QCD can be treated perturbatively, using the coupling constant α_s as expansion parameter. This means that the cross section for a process at given order in α_s is determined by:

$$\hat{\sigma} = \sigma^{\text{Born}} \left(1 + \underbrace{\frac{\alpha_s}{2\pi} \sigma^{(1)}}_{\text{NLO corrections}} + \underbrace{\left(\frac{\alpha_s}{2\pi}\right)^2 \sigma^{(2)}}_{\text{NNLO corrections}} + \dots \right)$$

In this case σ_{Born} is the leading-order cross section, where the first (i.e. the simplest) Feynman diagrams with a non-zero contribution to the cross section are taken into account in the scattering amplitudes. By adding higher orders – Feynman diagrams with

¹In particular with respect to Monte-Carlo event generators, this chapter focusses on the tools commonly used in CMS.

2. Conceptual overview

virtual loops or real emissions – the prediction for $\hat{\sigma}$ gets more and more precise. It should be stressed that, although real emissions are not considered in a LO calculation, the latter can still be regarded inclusive, since real emissions are absorbed into the parton density functions (PDFs) and can be made explicit by adding a parton shower (see Sec. 2.3.1). Typically, the same order is used in the PDF as in the cross section calculation, i.e. NLO PDFs are used for NLO computations etc., as they reflect the same level of accuracy.

The master equation for the cross section of $pp \rightarrow X$ at the LHC reads [49, 50]

$$\sigma_{pp \rightarrow X} = \sum_{a,b} \int f_a(x_1, \mu_F) f_b(x_2, \mu_F) d\hat{\sigma}_{ab \rightarrow X}(\hat{s}, \mu_F, \mu_R) dx_1 dx_2,$$

where $\hat{\sigma}_{ab \rightarrow X}$ is the process-specific partonic cross section for $ab \rightarrow X$. It is dependent on the center-of-mass energy of the two incoming partons and on the factorization and renormalization scales explained in Sec. 2.1.2. Interference effects between possible different intermediate states yielding the same final state X are accounted for as well as phase space constraints² and spin correlations. The quantities $f_{a,b}$ are the process-independent PDFs for the two partons and represent the probability that for a given factorization scale μ_F a certain parton carries the fraction x of the entire proton's longitudinal momentum. The functions have to be experimentally measured e.g. in deep-inelastic scattering experiments with a known energy scale $\mu_F = Q_0$. With the DGLAP evolution equations [51–53] the PDFs for other Q values can be obtained. Fig. 2.1 shows the PDF that was mainly used in the generation of samples for the 8 TeV analysis. Finally, the integral in the above equation is summed over all allowed combinations of partons (a, b) that can initiate the process with a final state X .

In other words, the partonic cross section determined by the matrix elements is convoluted with the PDF which specifies the probability the two partons can provide the needed energy.

2.1.2. Renormalization and factorization scales

The renormalization scale μ_R basically is the energy at which the running coupling constant of QCD, α_s , is evaluated. It therefore determines the strength of the strong coupling, and according to $\alpha_s(\mu_R) \propto 1/\ln(\mu_R^2/\Lambda_{\text{QCD}})$, low scales will lead to a larger coupling constant and therefore to enhanced contributions from processes with additional radiation. The parameter Λ_{QCD} is the scale until which perturbation theory is valid (typically a few hundred MeV).

The factorization scale on the other hand is less intuitive: it determines the energy below which additional radiation cannot be resolved anymore as extra hard emission in the calculations of the matrix element. Anything below this scale is resummed in the parton distribution functions, and can be recovered by applying a parton shower.

Typically, the two parameters are set to $\mu_R = \mu_F = Q$, where Q is the involved hard scale Q of the process, e.g. the momentum transfer at the strong interaction vertex.

²According to Fermi's Golden Rule, $d\sigma_{i \rightarrow f} \propto |\mathcal{M}_{if}|^2 d\zeta$, where \mathcal{M}_{if} is the matrix element given by the scattering amplitudes and $d\zeta$ is an infinitesimal piece of the allowed phase space.

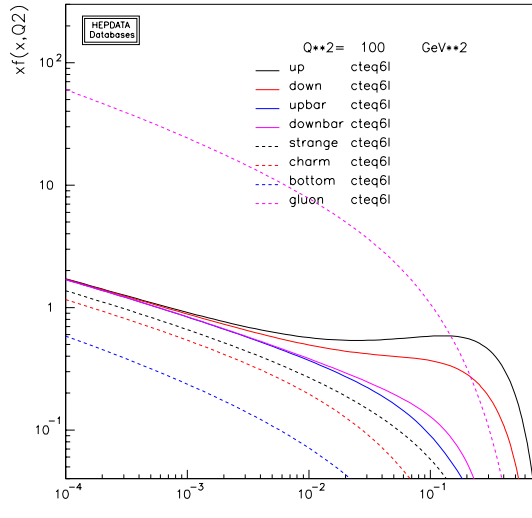


Figure 2.1.: The cteq6l1 PDF has been widely employed in the 8 TeV simulation campaign in CMS. The hard scattering of both the signal sample as well as the main background samples is modelled with this LO PDF. Here the densities $xf(x)$ for the partons u/\bar{u} , d/\bar{d} , c , s , b and g are shown for $Q^2 = 100 \text{ GeV}^2$. Created with [54].

The dependence on $\mu_{R,F}$ of the predictions is reduced if higher orders are considered in $\hat{\sigma}_{ab \rightarrow X}$ and the PDFs. This makes computations which are at least NLO accurate inevitable for a large set of processes, when small uncertainties and high precision are needed to draw conclusions from data. At the same time, it has become a standard procedure to vary the scales by factors 2 and 1/2 in order to define an uncertainty that is supposed to reflect missing higher order contributions.

Almost any cross section used in this search for tHq has been derived at least at next-to-leading order in QCD; therefore a brief description of the ingredients of an NLO cross section computation is provided in the following.

2.1.3. Cross sections at NLO in QCD

At NLO the cross section consists of the contributions from the Born diagrams, the virtual loop corrections (both of which are characterized by an n partons final state) and the corrections due to the real emission of an additional parton, leading to a final state with a parton multiplicity of $n + 1$:

$$\sigma^{\text{NLO}} = \int_n d\sigma^{\text{Born}} + \int_n d\sigma^{\text{V}} + \int_{n+1} d\sigma^{\text{R}}$$

A problem arises in the computation of the last two terms, because they are both divergent. For a soft ($k_T \rightarrow 0$) and collinear ($z \rightarrow 1$) gluon emission the branching cross section $q \rightarrow qg$ is described by

2. Conceptual overview

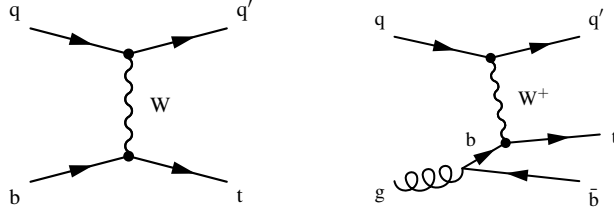
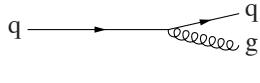
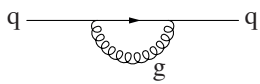


Figure 2.2.: Leading order Feynman diagrams of the single top t -channel in the 5F (left) and the 4F (right).



$$: \quad d\sigma^R \propto \frac{dz}{1-z} \frac{dk_T^2}{k_T^2},$$

where k_T is the gluon transverse momentum relative to the quark and z is the fraction of the quark momentum. The same proportionality up to the sign is present in the term describing the loop momentum of the virtual corrections,



$$: \quad d\sigma^V \propto -\frac{dz}{1-z} \frac{dk_T^2}{k_T^2}.$$

Because of the minus-sign, the singularities of real and virtual correction terms cancel to a finite value for all observables which are insensitive to the presence of an additional collinear parton in the limit $k_T \rightarrow 0$. This cancelling behaviour of infrared divergences between loop-integrals and phase space integrals has been observed in the 1960s and is referred to as the Kinoshita-Lee-Nauenberg (KLN) theorem [55, 56].

2.1.4. Description of b initiated processes

Single top production (this means also the process searched for in this thesis) involves a tWb vertex and therefore necessitates a b quark in either the initial or final state. This allows to describe the process with two different approaches referred to as the 4-flavour (4F) and 5-flavour (5F) schemes.³ The 4F adopts the following reasoning: the b quark mass is larger than the one of the proton and thus can only be produced in pairs in hard interactions with high momentum transfer. Accordingly, it cannot be an initial state particle, and the b PDF is set to zero, as b quarks do not contribute to the proton wave function. On the other hand, in the 5F, the b quark is treated as a massless particle and as a proton constituent. In terms of Feynman diagrams at LO in the t -channel single top production, this results in the $2 \rightarrow 3$ and $2 \rightarrow 2$ processes shown in Fig. 2.2.

The 4F approach is reliable if the hard scale Q of the process and the b quark mass are of the same order, which can be understood in the following way: the cross section of the single top t -channel (the same applies to tW) is enhanced by a logarithm that comes from

³A detailed discussion of these can be found in [57].

collinear splitting of the initial gluon. The t -channel propagator of the internal b quark in the right Feynman diagram of Fig. 2.2 has the form $1/(t - m_b^2)$ with the Mandelstam variable $t = (p_{\bar{b}} - p_g)^2$. This leads to collinear logarithmic contributions to the cross section, i.e.

$$\sigma \propto \int_0^{Q^2} \frac{dp_T^2}{p_T^2 + m_b^2} = \log\left(\frac{Q^2}{m_b^2}\right),$$

and large logarithms for $Q^2 \gg m_b^2$ could spoil the convergence to a finite result in perturbation calculations.

This issue is addressed and overcome by the 5F approach, where the logarithms are absorbed (and the \bar{b} radiation is resummed) into a b PDF via the DGLAP evolution equations with $\mu_F = Q$. The t -channel cross section is then not given by $\sigma(qg \rightarrow q'tb)$ but by $\sigma(qb \rightarrow q't)$. In the simulation of 5F events, the additional b quark can be recovered by the backwards evolution in the parton shower step, albeit at severely less accuracy in its kinematics than in the 4F, where the additional b quark is part of the matrix element calculation. Since m_b must be 0 in the 5F calculation, the b quark mass enters the computations only as starting scale in the evolution of its PDF up to μ_F .

As is concluded in [57], in the cases of large ratios Q^2/m_b^2 that may occur in the 4F, the size of the logarithms is reduced by accompanying, suppressing universal phase space factors. Therefore, 4F results are in general well-behaved, and should provide a better description of the b quark kinematics, taking into account mass effects already at leading order. In consideration of the above, the 4F has been widely employed in the generation of the samples for the analysis.

Fig. 2.3 shows differential distributions for the SM tHq production at $\sqrt{s} = 13$ TeV for different computations. It can be seen that going to NLO in the calculations drastically reduces the scale uncertainties, and that the 5F fails to describe b jet kinematics at LO, where the additional b quark only comes from the parton shower. Judging from these distributions, the tHq modelling employed by CMS and used for the presented analysis, 4F LO, is deemed good and reliable.

2.2. Monte-Carlo event generators

The Monte-Carlo (MC) method is a technique to obtain numerical results by randomly sampling draws according to a specific probability function. In the field of High Energy Physics it has proven to be extremely useful for the case of generating “artificial” events of the processes of interest, in order to compare the numerical predictions against real data or to measure and tune the performance of detector components.

Since with the MC method the phase space integration happens numerically, one cannot analytically solve the divergent integrals occurring in the NLO computation explained in Sec. 2.1.3.⁴ Instead, the divergence of the real emission corrections needs

⁴The so-called dimensional regularization introduces a regulator ϵ and makes the divergences of virtual and real corrections cancel to a finite value as $\epsilon \rightarrow 0$. This however is not possible in a numerical code.

2. Conceptual overview

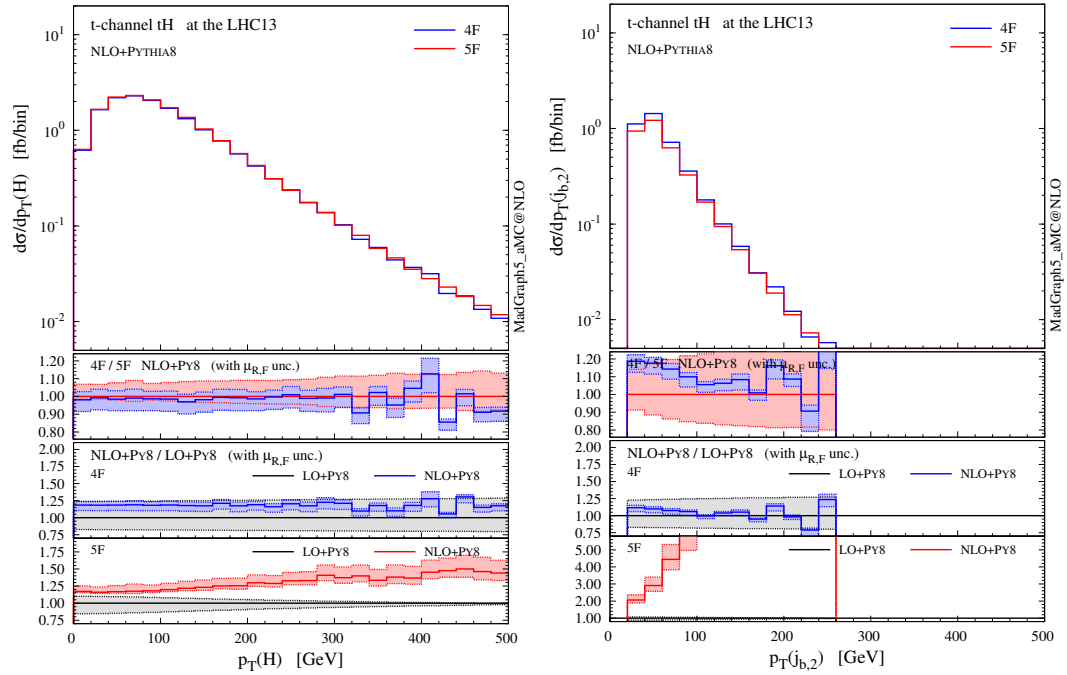


Figure 2.3.: Differential Higgs boson and 2nd b jet p_T distributions at NLO for SM tHq production at 13 TeV, for both the 4F and 5F; taken from [35]. In the ratios, 4F/5F differences and NLO/LO k factors are shown. In the 4F, the effect of going one order higher in QCD can be approximated by a single k factor, as NLO corrections do not drastically affect the shape. Only at NLO in the 5F, the additional b quark is part of the matrix element calculations and thus leads to an improved description of b jet kinematics, whereas for the 5F at LO, the additional b quark is added by the parton shower. Overall, scale dependence is reduced by performing the computation at NLO as opposed to LO, as can be seen from the size of the uncertainty bands.

to be factored out explicitly so that it cancels against the virtual correction according to the KLN theorem. This can be achieved by means of the clever use of subtraction terms, which are introduced in a way that allows to take out the divergence of the integral, rendering the latter finite and numerically solvable. Besides the so-called dipole subtraction [58,59], the FKS subtraction [60], named after its inventors Frixione, Kunszt and Signer, is commonly applied in the main NLO MC event generators, two of which will be introduced below.

2.2.1. MadGraph5_aMC@NLO

Being the NLO extension and successor of MADGRAPH5 [61], MADGRAPH5_aMC@NLO [62] is a multi-purpose matrix element and event generator, capable of performing both LO and NLO calculations. As it has been widely used in the context of this thesis, it will be explained in more detail.

Its great flexibility is grounded in the methodology of how it computes inclusive

and differential cross sections and generates events: the user only needs to provide a theory model, i.e. the Lagrangian and parameters such as particle masses and coupling strengths, and the program will calculate the cross section with process-independent building blocks, which e.g. are the FKS subtraction terms, since they can be treated in a universal manner. Once the user has defined the process that shall be generated, `MADGRAPH5_AMC@NLO` creates all possible Feynman diagrams and writes out functions for evaluating their matrix elements⁵, which will be called by the building blocks.

The user-friendliness is ensured by an interactive command line, where commands and special syntaxes can be issued and everything else happens in a fully automated way. In the case of t -channel single top + Higgs production the commands read:

```
> import model loop_sm-ckm
> generate p p > t b~ j h $$ w+ w- [QCD]
> add process p p > t~ b j h $$ w+ w- [QCD]
> output
> launch
```

The optional flag `[QCD]` makes the program calculate the process at NLO in QCD, including both virtual and real corrections. The `$$` syntax is needed for the exclusion of diagrams where a W boson appears in the s -channel. Here the model is the one determined by the Lagrangian of the Standard Model and employs a non-diagonal CKM matrix. Input parameters, particle masses and settings like scales and PDFs can be conveniently edited and handed over with a `run_card.dat` and `param_card.dat`. The decay of resonances in the NLO mode is handled via the module `MADSPIN` (see also [61]), which preserves spin correlations.

Via the so-called CKKW [64] and MLM [65] merging, it is possible to combine processes with different final state parton multiplicities at matrix element level, thus reducing the dependence on parton shower programs (which are insufficiently accurate in the generation of hard emissions) for producing extra radiations. With the FxFx [66] technique this has recently also become possible at NLO. These mergings are all available in the `MADGRAPH5_AMC@NLO` framework.

A functionality that allows to reweight events according to different renormalization or factorization scales or different PDFs is implemented [67] and comes at little extra computational cost. This feature has been deployed in the generation of the signal samples at 13 TeV, where events are reweighted even according to different values of the Higgs boson coupling parameters.

Another concept essential to NLO matrix element generators can be illustrated with `MADGRAPH5_AMC@NLO`: the same final state configuration characterized by $n + 1$ partons can be obtained by considering a real emission in the matrix element calculation or by having the parton shower add a hard emission to a Born event with n outgoing partons (see Fig. 2.4). When `MADGRAPH5_AMC@NLO` generates events at NLO, it must therefore circumvent the danger of double counting, and short distance calculations must

⁵Their evaluation is performed via so-called helicity amplitudes, see e.g. reference [63].

2. Conceptual overview

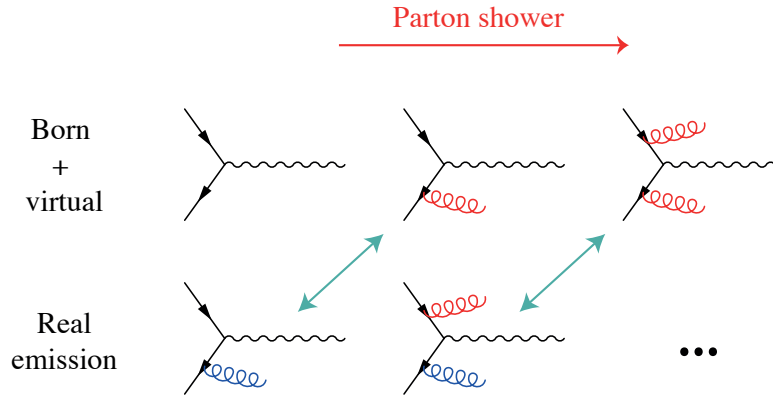


Figure 2.4.: The parton shower (red) attaches gluons to the Feynman diagrams. In the case of a matrix element Born diagram, the first emission added by the PS creates an overlap with diagrams that have a real emission (blue) from the NLO matrix element. The MC@NLO and POWHEG methods are developed to overcome this double counting in the event generation. Adapted from [50].

be matched to parton showers with dedicated techniques. In MADGRAPH5_AMC@NLO this happens with the MC@NLO formalism [68]: the parton shower contributions to get from the n -body Born final state to the $n + 1$ -body real emission final state are subtracted from the $n + 1$ -body matrix element contributions. Introducing these counterterms has the advantage of creating a smooth transition between the phase space regions which are governed by either the matrix element or the parton shower; shapes of observables related to the emission coincide with what one gets from the parton shower for the soft and collinear region and with NLO matrix element calculations for the case the emission is sufficiently hard. The major drawback of this method is the introduction of negative event weights, which degrade the statistical power of distributions when filling histograms and can even lead to unphysical, negative entries in tails of distributions populated by few events. The negative event weights are owed to the fact that the matrix elements are not necessarily positive-finite anymore in the presence of the counterterms. It remains to say these terms depend on the specific parton shower program to be used, as each parton shower has its own methodology of increasing the parton multiplicity by $+1$. In the MADGRAPH5_AMC@NLO framework, sets of counterterms are available for all popular parton shower programs; in order to ensure NLO accuracy, the appropriate parameters must also be set when showering the events. The settings to be used are given in [69].

2.2.2. Powheg

The Positive Weight Hardest Emission Generator, abbreviated POWHEG [70–72], is the other NLO MC tool that is used for simulation of a large variety of processes in the analysis presented in this thesis. It provides pre-defined processes for which the user

can generate events or calculate cross sections by running a precompiled executable `pwhg_main`. The parameter configuration is steered by a `powheg.input` card. While it also employs the FKS method for treating the divergences in the real emissions, double counting in the parton shower step is prohibited through the eponymous POWHEG method: the hardest radiation is always generated first by the matrix element generator. With this technique only positively weighted events are produced, which can furthermore be showered with any parton shower that is either p_T ordered or able to veto emissions with a p_T larger than the one of the radiation already present in the matrix element. Other than that, POWHEG is not restrictive in terms of subsequent parton showers and does not require different settings for different programs. In recent releases, some of the implemented processes come with the functionality of adding extra event weights corresponding to other PDF sets or different scales. A list of available processes can be found on the official page [73].

2.3. Parton shower and hadronization

In order to get the full, physical picture, matrix element calculations can be matched to parton shower programs to get access to the phase space where perturbative QCD breaks down (this happens at low energies and long distances because of the steadily growing α_s , which at some point is too large to allow an expansion in this parameter) and to describe the further evolution of the event. In fact, NLO matrix element events are not physical without the application of a parton shower.

In general, these programs take care of the radiation of soft and collinear colored particles from quark and gluon lines in the initial or final state by parametrizing the probability for a collinear emission with the Altarelli-Parisi splitting functions [53] and the so-called Sudakov form factors. Parton showers are unitary, which means they do not change the weight of an event (in particular not to zero⁶), and are inclusive in the sense that after the parton shower step all emissions one would expect have been generated. Emissions take place with a certain probability until the energy of the single partons is low enough that partons hadronize and form colorless, bound states. The process of hadronization can be described with either the Lund string model [74] or the cluster model [75].⁷ The former makes use of the self-interacting behaviour of the strong force by describing the gluon field between two colored partons with a color-flux string. It is stretched when the dipoles are flying apart and eventually bursts to form new quark-antiquark pairs. This happens iteratively until the energy in the string is too low to create new pairs of quarks, and the partons at the end of the string consequently form a bound state. Since the simulated samples used in this thesis are exclusively deploying

⁶This is true except for the case of CKKW, MLM and FxFx merging, where an adjustable threshold slices the emission phase space into regions which below are populated by the parton shower and above by the matrix element generator. Events can therefore also be discarded.

⁷A third approach is the cluster hadronization of the Monte-Carlo generator Sherpa [76,77]. Since this section however is intended to give an overview of the MC generators commonly employed within the CMS collaboration, Sherpa is not addressed in this chapter.

2. Conceptual overview

the Lund string model, the alternative approach of cluster hadronization will not be discussed here. Many hadrons are initially produced in excited states and consequently are not stable. Their decay, as well as the trivial (because isotropic) decay of the scalar Higgs boson, is performed within the parton shower program as well. Finally, another issue is addressed by the parton shower: until now, the proton remnants have not been considered in any simulation step. However, it is possible that multiple partons of the protons interact and lead to additional signatures in the detector, or that there is a color connection between the partons of the hard interaction and the remnants traveling down the beam pipe, which influences the step of hadronization. This is called “underlying event”, and the parameters of its simulation must be tuned against data, e.g. [78] in minimum-bias events, [79] for tuning against Drell-Yan events or taking into account $\bar{t}\bar{t}$ events [80].

2.3.1. Pythia6/8

Despite being a multi-purpose MC generator for simulating all stages of the event evolution (with leading order matrix elements), in CMS PYTHIA is mostly used as interfaced parton shower generator and for all the steps beyond, e.g. building hadrons from the partons according to the Lund string hadronization scheme. The matrix element events come mainly either from POWHEG or MADGRAPH5_AMC@NLO.

All 8 TeV simulation samples used in this thesis have been showered in PYTHIA 6 [81] with the underlying event Tune Z2* (a modification of the Z2 tune obtained from 7 TeV data [82]), using the p_T^2 of the emission as evolution scale along which the radiations take place down to the energy regime where hadronization steps in (typically at ~ 0.5 GeV). Decays of τ leptons are handled by a separate plugin, TAUOLA [83].

Its role as the main parton shower program in CMS used at 8 TeV has been taken by PYTHIA 8.2 [84] in the 13 TeV simulation campaigns. The underlying event tune in the samples analyzed in this thesis is CUETP8M1, which is based on the Monash tune [85].

2.3.2. Herwig++

Another commonly used generator is Herwig++. It is similar to Pythia in the sense that it features a library for many hard subprocesses and serves as tool for the simulation of the parton shower, hadronization and underlying event. Differences consist in the evolution scale for the successive emission of partons, which is the angle between the emitting parton and the emission in the Herwig++ case (it is therefore referred to as angular-order shower as opposed to the p_T -ordered PYTHIA), and in the use of the cluster model to model hadronization. It is frequently employed at CMS in Run-II with the underlying event tune EEC5 [86].

2.4. Multivariate analysis tools

MVAs are heavily used in this thesis in order to better discriminate between the signal and the background processes and consequently increase the search sensitivity beyond what would be possible with an approach that is purely based on imposing cuts to enrich the tHq contribution. The classification mode of the MVA (as opposed to a regression) has been utilized for assigning to each event that passes the selection a real number that can be seen as a measure for its compatibility with the signal or background hypothesis.

The MVA must be trained with simulated data from which it learns about topologies and characteristics of the respective processes. In practice, it is provided with distributions of jet transverse momenta, lepton observables etc., and the full list of input variables will be given in the analysis chapter. After the training phase, the gained knowledge is used to classify any unknown event with the same type of input information as being rather signal- or background-like.

Two types of MVAs find use in this thesis – artificial neural networks and boosted decision trees. They will be explained in the following.

2.4.1. Artificial neural networks

An artificial neural network, abbreviated NN, is inspired by the structure of the vertebrate brain with its composition of neurons that are connected via synapses and communicate by means of propositional logic [87]. In the NN the information transmitted from one node to the other via the connections between nodes can be given a variable weight, being tantamount to a minimization problem to achieve a desired outcome in the training phase. Mathematically this is possible due to the universal approximation theorem [88, 89], which enables the NN with one hidden layer of neurons between the input and output layer to approximate any continuous function. Being provided with n real-valued input variables, the NN performs an $\mathbb{R}^n \rightarrow \mathbb{R}$ mapping, where the function of the mapping depends on the weights. The so-called multilayer perceptron implementation in TMVA [90] is used in this thesis. A pictorial illustration of its architecture (with a reduced number of nodes in the input and hidden layer for the sake of lucidity) is provided in Fig. 2.5. It features one hidden layer and is feed-forward, i.e. there is no reflux of information.

Central to the NN is the neuron response function h . For a neuron in the hidden layer it is given by

$$h(\mathbf{x}) = \sum_{i=1}^n w_i x_i$$

for a given set of real inputs x_1, \dots, x_n received from the input layer with n nodes, which are weighted with weights w_1, \dots, w_n . Since h is a linear combination of the inputs, it must be transformed with a non-linear activation function S in order to introduce non-linearity into the multilayer perceptron. Otherwise it would effectively behave like a single-layer

2. Conceptual overview

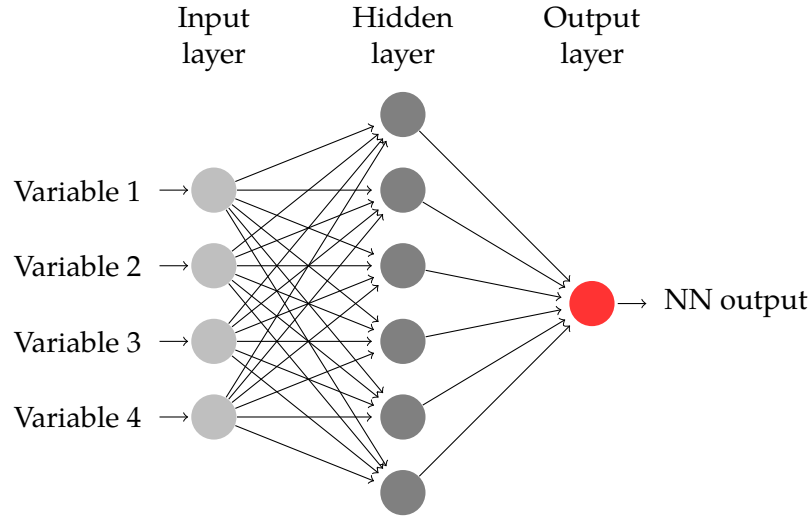


Figure 2.5.: Architecture of a multilayer feed-forward NN with one hidden layer, as employed in this thesis. Information is transported only in one direction between two nodes in neighboring layers, and combined to a single NN output, which corresponds to real value representing a signal- or background-likelihood for a given event.

perceptron when summing all layers and could not approximate any arbitrary function. A hyperbolic tangent is used as activation function,

$$S = \tanh h = 1 - \frac{2}{e^{2h} + 1}.$$

The response of the single node in the output layer – called “NN output” from now on – for an NN with m (n) nodes in the hidden (input) layer is given by

$$\text{NN output} = S \left(\sum_{j=1}^m w_j S \left(\sum_{i=1}^n w_{ij} x_i \right) \right).$$

In supervised learning, an NN is provided with events of known origin, i.e. signal or background, and tries to reproduce a target value for each case (e.g. 1 for the signal, 0 for background events). This is done by minimizing an error function defined as

$$E(\mathbf{w}) = \sum_k (y(\mathbf{x}_k, \mathbf{w}) - \hat{y}_k)^2,$$

which is the mean square error with \hat{y}_k being the target value of event k that yields an NN output $y(\mathbf{x}_k, \mathbf{w})$. The Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm [89, 91–93] is employed to find the vector \mathbf{w} for which E has its minimum. Once this vector is known, the NN can classify any unknown event.

2.4.2. Boosted decision trees

Because of their speed and robustness, boosted decision trees (BDTs) are used throughout this thesis mostly for optimization studies and preliminary results. A BDT is an assembly of weighted binary classifiers [94], e.g. of weighted simple decision trees each of which has a low separation power, as it consists of a sequence of nodes performing binary decisions. This is an essential precondition, because, upon creating a forest of many weak learners, it guarantees that no single weak learner can be strong enough to significantly influence the result on its own. It guarantees robustness of the results against large weight variations during the optimization process explained below.

In a single node one aims at maximizing the gain in separation power between signal and background achieved by a cut on a certain input variable (e.g. the reconstructed Higgs boson mass). A node split into two nodes is introduced at the specific cut value that maximizes the change in the Gini functions between the parent and the daughter nodes $\Delta = Q_{\text{Gini}} - Q_{\text{Gini,dau1}} - Q_{\text{Gini,dau2}}$, where the Gini function is defined as [95]

$$Q_{\text{Gini}} = P(1 - P) \sum_i w_i.$$

Here, w_i is the weight of event i and P is the purity given by

$$P = \frac{\sum_s w_s}{\sum_s w_s + \sum_b w_b}.$$

The leaf at whose end $P > 0.5$ is labelled a signal leaf, otherwise it is called background leaf.

The concept of boosting [96] helps in making the classification more robust: training events that were misidentified in a first iteration and ended up in a background (signal) leaf despite being signal (background) receive a larger weight, and another decision tree is built. In fact, an entire forest of trees is trained, and the final score ‘‘BDT output’’ of an event is equal to the weighted average of the sum of $+1$'s and -1 's if the event has ended up in a signal or background leaf, respectively. Two boosting algorithms are utilized in this thesis.

Adaptive boosting

The adaptive boosting (AdaBoost) [97] algorithm assigns a larger weight in the training of the $m + 1^{\text{th}}$ decision tree to events that have been misclassified by the m^{th} decision tree. The latter's misclassification rate found among N events, where each event has a weight $w_i = 1/N$, is defined as

$$r_{\text{mis}}^m = \frac{\sum_i w_i q_i}{\sum_i w_i},$$

2. Conceptual overview

where ϱ_i equals +1 if the event i has been misclassified and 0 if it has been assigned to a correct leaf. For the m^{th} weak learner one can then compute the quantity

$$\alpha_m = \ln \left(\frac{1 - r_m^{\text{mis}}}{r_m^{\text{mis}}} \right),$$

which is used to scale the event weights in the next, i.e. the $m + 1^{\text{th}}$ decision tree:

$$w_i^{(m)} \rightarrow w_i^{(m+1)} = \frac{w_i^{(m)} e^{\alpha_m \varrho_i^{(m)}}}{\sum_i w_i^{(m)} e^{\alpha_m \varrho_i^{(m)}}}.$$

The final score is composed of the scores h_m of the individual trees, which can either be +1 or -1 depending on whether or not the event has been classified as signal or background, weighted by the factors α_m :

$$\text{BDT output} = \sum_{m=1}^N \alpha_m h_m.$$

Gradient boosting

The second boosting algorithm is the so-called gradient boosting [98]. Instead of iteratively updating the weights as for the AdaBoost, it optimizes the training results via minimizing a differentiable loss function by approaching its minimum with the method of steepest descent. For a detailed discussion of this alternative algorithm see [99].

2.4.3. Interpreting training results: overtraining and ranking of variables

If the training samples have too few events, it can happen that the NN or BDT interprets statistical fluctuations occurring in the input distributions as genuine features of the respective process. This unwanted behaviour, which biases the results in the application step, is referred to as overtraining. It is checked by employing statistically independent test samples and comparing their results with the ones from the training. In the case of overtraining, the MVA is expected to perform worse in the test samples. This phenomenon is illustrated in Fig. 2.6.

A ranking of the variable importance will be given in the analysis chapter for the case of the NN. According to [100], the importance of variable k is determined by summing the weights attached to the connections between the respective input node and all the J nodes in the hidden layer:

$$I_k = \bar{x}_k^2 \sum_{j=1}^J w_{jk}^2.$$

In the above formula, \bar{x}_k is the sample mean of variable k .

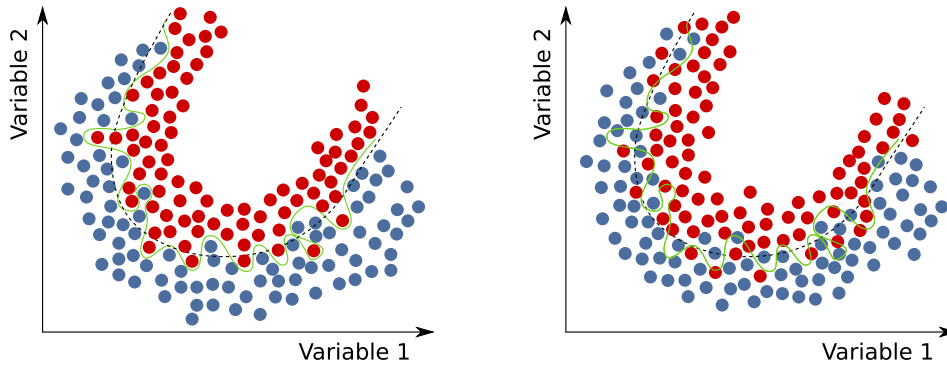


Figure 2.6.: Illustration of overtraining for an MVA trained with two input variables. The left figure represents the signal (red) and background (blue) training events. An overtrained MVA (green solid line) follows even small statistical fluctuations perfectly, while an MVA with a sound training is insensitive to fluctuating events (black dotted line). In the case of the test sample to the right, which consists of statistically independent events, the overtrained MVA tries but fails to reproduce the statistical fluctuations and consequently leads to bogus results. The not-overtrained MVA on the other side displays the same performance as in the training case. Taken from [100].

2.5. Fitting and statistical inference

In the previous section tools have been discussed which allow for a very good separation between signal and background processes. However, despite the signal being enriched in a certain region of phase space, one will not be able to make definite statements on its existence from a comparison of data to simulation by eye. Statistical methods are needed to quantify the signal strength or – in the absence of a clear indication of signal contributions to data – the search sensitivity (i.e. which signal strength the analysis is able to exclude at a certain level of confidence) by means of exclusion limits. Central to the limit setting procedure is the calculation of the a test statistic that quantifies the compatibility between observation and the assumed model. The test statistic employed in this thesis in turn relies on the parameter estimation with the maximum-likelihood method. These ingredients of the statistical inference are explained in the following.

2.5.1. Parameter estimation with the maximum-likelihood method

The likelihood function corresponds to the probability for obtaining the statistically independent measurements $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ for parameters \mathbf{a} that characterize the underlying statistical model:

$$\mathcal{L}(\mathbf{a}) = f(\mathbf{x}_1|\mathbf{a}) \cdot f(\mathbf{x}_2|\mathbf{a}) \cdot \dots \cdot f(\mathbf{x}_N|\mathbf{a}) = \prod_{i=1}^N f(\mathbf{x}_i|\mathbf{a}).$$

Equivalently, the set of parameters \mathbf{a} that maximizes \mathcal{L} belongs to the model which is most compatible with the observations. Finding this set via a maximum-likelihood fit is a

2. Conceptual overview

pivotal element in the determination of the best parameters, which denote e.g. systematic uncertainties or the signal strength, when comparing the model with measurements in terms of binned histograms. The expression of the likelihood function for a hypothesis with a generic signal strength modifier μ is the product of Poisson probabilities over all bins, with μs_i , b_i and n_i denoting the expected signal and background yields and the actually observed entries in bin i :

$$\mathcal{L}(\text{data}|\mu, \boldsymbol{\theta}) = \prod_i \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-\mu s_i - b_i},$$

It is implied that $s_i = s_i(\boldsymbol{\theta})$ and $b_i = b_i(\boldsymbol{\theta})$, so the expected signal and background contributions to a specific bin depend on the nuisance parameters $\boldsymbol{\theta}$ encoding systematic uncertainties the analysis is afflicted with.

2.5.2. Incorporation of uncertainties by means of nuisance parameters

Some of the nuisance parameters (like the uncertainty on the luminosity) apply to all processes considered, others are only there for a specific background or the signal, or vary in size for the different contributions. The nuisance parameters are incorporated into the model through up and down templates⁸ which have different numbers of events per bin than the nominal template for the respective process. The exact parametrization of the expected number of events for a specific bin in the presence of nuisance parameters is given in Eqs. (6.19)-(6.21) of [40]. It should be noted that shape uncertainties allow for a bin-wise variation of the yields, while rate uncertainties change the bin contents in a fully correlated way, i.e. the up and down templates are only shifted in their normalization. Rate uncertainties can accordingly be regarded a special case of shape uncertainties.

2.5.3. Hypothesis testing, likelihood ratio and p value

For the testing of a certain hypothesis (in this case the hypothesis corresponds to the signal scenario defined by a specific μ value) a test statistic needs to be defined in order to derive confidence intervals on μ . For the sake of simplicity and demonstration, a simple hypothesis is assumed for the moment, i.e. no nuisance parameters are considered. One possible test statistic in such a case is the likelihood ratio

$$\lambda(\mu) = \frac{\mathcal{L}(\text{data}|\mu)}{\mathcal{L}(\text{data}|0)}.$$

With the above definition, a larger value $\lambda(\mu)$ signifies a larger (smaller) compatibility of the observation with the $s + b$ (b -only) hypothesis. The test statistic (the term “statistic” implies that values of λ are following a certain probability distribution function f), can be used to compute a confidence interval for μ by means of the p value under the $s + b$ hypothesis. The p value answers the question: for a hypothesis with signal

⁸A template in this case is a histogram in a certain observable and typically contains the observed event counts or the event yields expected from the model according to its probability distribution function.

strength μ under test, how large is the probability to obtain a measurement which has a compatibility with the assumed hypothesis that is of the same size as or worse than the actual observation? If p_{s+b} is smaller than some predefined value – the significance value α – the $s + b$ hypothesis is rejected. If it is larger, the probability to obtain at least the observed level of incompatibility with the $s + b$ hypothesis is high enough that one claims one cannot exclude the hypothesis, which consequently is accepted. The entity of values for which p_{s+b} with

$$p_{s+b} = \int_{-\infty}^{\lambda^{\text{obs}}} f(\lambda|\mu) d\lambda$$

is larger than α constitutes the confidence interval for μ at $(1 - \alpha) \times 100\%$ confidence level (C.L.). Typical values are $\alpha = 0.05$ and 95% C.L.

As mentioned above, this procedure employs the probability density function (pdf) $f(\lambda|\mu)$ belonging to the signal + background ($s + b$) hypothesis; practically, the pdf is constructed by dicing pseudoexperiments according to the $s + b$ hypothesis for the μ under test and computing $\lambda(\mu)$ for each pseudoexperiment. If for this μ the p_{s+b} value happens to be smaller than 0.05, the $s + b$ hypothesis will be rejected. However, such an approach is ill-defined (not from the statistical, but from the experimental point of view) for the case in which the pdf constructed from background-only (b -only) pseudoexperiments, $f(\lambda|0)$, looks very similar to $f(\lambda|\mu)$. In such a scenario, one would claim to have excluded μ although the analysis is not sensitive to the $s + b$ hypothesis at all.

2.5.4. CL_S exclusion limits

The above issue is overcome by introducing the CL_S criterion [101–104] according to

$$\text{CL}_S = \frac{p_{s+b}}{1 - p_b},$$

which contrasts the p -value of the $s + b$ hypothesis with the one of the b -only hypothesis for a given μ . The p values for the $s + b$ and b -only hypotheses in this case read:

$$p_{s+b} = P(\tilde{q}_\mu \geq \tilde{q}_\mu^{\text{obs}} \mid \text{signal+background}) = \int_{\tilde{q}_\mu^{\text{obs}}}^{\infty} \int_{\theta} f(\tilde{q}_\mu | \mu, \hat{\theta}_\mu) d\theta d\tilde{q}_\mu,$$

$$p_b = P(\tilde{q}_\mu < \tilde{q}_\mu^{\text{obs}} \mid \text{background}) = \int_{-\infty}^{\tilde{q}_\mu^{\text{obs}}} \int_{\theta} f(\tilde{q}_\mu | 0, \hat{\theta}_0) d\theta d\tilde{q}_\mu.$$

The inverted sign in the integral of the definition of p_b reflects the fact that for the hypothesis μ under test, a smaller value of the test statistic \tilde{q}_μ (whose definition will be given below) indicates a better compatibility with the data. In turn, the b -only scenario has a worse compatibility for smaller values. As can be seen immediately, there is no conceptual problem anymore in the case of low-sensitivity searches, because if the shapes of pdfs are very similar, the CL_S value will be large enough so that an exclusion of the

2. Conceptual overview

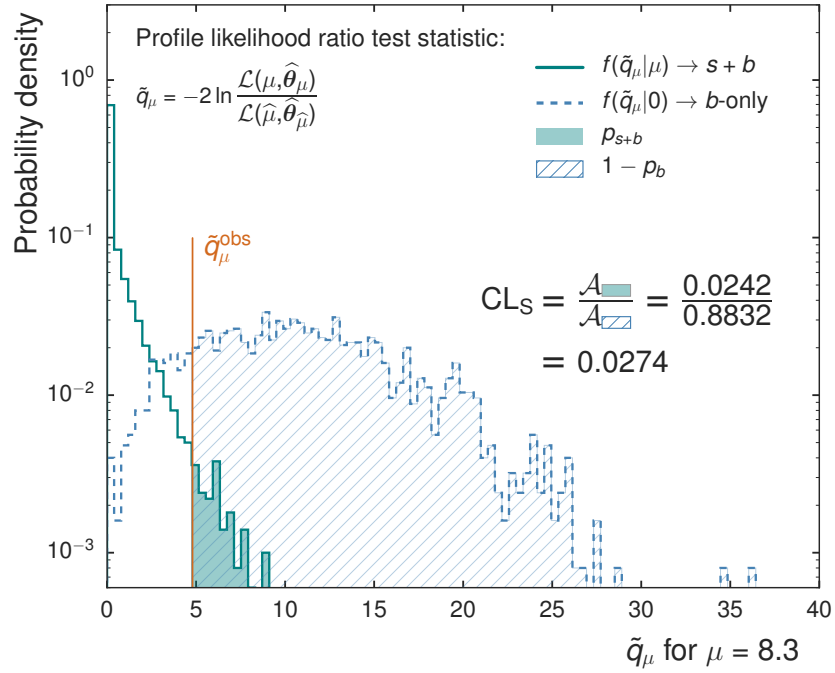


Figure 2.7.: Illustration of the concept for computing the CL_S exclusion limit for a given $\mu = 8.3$. This is an actual example of a test statistic distribution taken from Chap. 4. Since $CL_S < 0.05$, this μ is rejected.

$s + b$ scenario is prohibited. The functions $f(\tilde{q}_\mu | \mu, \hat{\theta}_\mu)$ are the probability distribution functions for the test statistic \tilde{q}_μ , which are obtained from dicing pseudoexperiments for the respective hypothesis and its best-fit nuisance parameters. The test statistic exploited in this thesis is referred to as profile likelihood and is defined as [105]

$$\tilde{q}_\mu = -2 \ln \frac{\mathcal{L}(\text{data} | \mu, \hat{\theta}_\mu)}{\mathcal{L}(\text{data} | \hat{\mu}, \hat{\theta}_{\hat{\mu}})}, \quad \mu \geq \hat{\mu} \geq 0, \quad (2.1)$$

where $\hat{\mu}$ and $\hat{\theta}_{\hat{\mu}}$ are the estimators that maximize the unconditional likelihood function, and $\hat{\theta}_\mu$ is the vector of nuisance parameter values for which, at a given and fixed μ , the likelihood function takes the maximum value. This is a practical and convenient feature of the profile likelihood method. For typically, the parameter one is interested in is the signal strength modifier μ . Making inferences about μ however is possible only if the systematic uncertainties of the measurement are determined and constrained. Instead of determining them in external measurements, the profile likelihood ratio allows to incorporate the systematic uncertainties in the statistical model via nuisance parameters that are profiled, i.e. they become a function of μ . One could also calculate an interval of allowed μ 's for every fixed θ . But making inferences on μ for every allowed point in the θ space would require a large amount of computational resources. Furthermore, the true value of θ is of no real interest; therefore, profiling the nuisances against the

parameter of interest μ is an efficient way of obtaining a tight confidence interval for the signal strength modifier.

Accordingly, in the analyses presented in Chapters 4 and 5 the CL_S criterion is chosen for setting upper boundaries on μ . A hypothesis with signal strength μ is excluded at $(1 - \alpha) \times 100\%$ confidence level if its CL_S value equals α . Conversely, the analysis is able to exclude signals with a cross section larger than $\sigma \times \mu$ at 95% CL for the μ that satisfies $CL_S = 0.05$. An illustration of the ingredients for the computation of the CL_S limit is provided in Fig. 2.7 with the help of the actual test statistic distribution belonging to $\mu = 8.3$ in Chap.4.

In practice

CMS has a tool for statistical inference in order to calculate the observed CL_S value for a specific μ . The `combine` program fits the nuisance parameters to data under the $s + b$ hypothesis (leaving the signal strength fixed). From the value of this likelihood the observed value of the test statistic, $\tilde{q}_\mu^{\text{obs}}$, can be calculated with Eq. 2.1. The program then generates a set of pseudoexperiments which are diced according to the fitted nuisances. Another ensemble of pseudoexperiments is obtained by doing the same after fitting under the b -only assumption. The pdf $f(\tilde{q}_\mu)$ of the test statistic can be constructed by computing Eq. 2.1 for each of the pseudoexperiments under both the $s + b$ and b -only hypothesis, and the CL_S value can be calculated. The above procedure is repeated for several values of μ provided to `combine` by the user, and the specific value of μ for which $CL_S = 0.05$ is obtained via interpolation.

Calculating the expected limits

The b -only pseudoexperiments as well as the nuisances $\hat{\theta}_0, \hat{\theta}_\mu$ and the pdfs f obtained in the calculation of the observed limit are saved, as they are needed for the expected limits. Every pseudoexperiment is treated as if it was real data, and the procedure of obtaining the observed limit is repeated. However, `combine` does not refit the nuisances for every pseudoexperiment, but uses the set of nuisances that maximized the fit to data, and the distributions of the test statistic derived thereof. This is done in order to save computation time. Finally, the 2.5%, 16%, 50%, 84% and 97.5% quantiles of the CL_S distribution obtained from the pseudoexperiments is reported as (the median of the) expected CL_S limit and its $\pm 1\sigma$ and $\pm 2\sigma$ deviations.

Asymptotic limits

Asymptotic limits make use of the fact that the test statistic distribution can be expressed analytically in the large sample limit [105, 106]. As a consequence of this behaviour, one can – instead of throwing computationally expensive pseudoexperiments – employ a so-called Asimov dataset which represents the entity of pseudoexperiments by one single dataset. This is a quick and easy way to obtain upper limits; however, the approach is only valid if the number of events is large enough (otherwise it will yield biased,

2. *Conceptual overview*

more stringent limits than the CL_S method), which is not in general the case for the tHq analysis. Therefore, this approach is only used when the differences between the CL_S and asymptotic limits are found to be small enough.

3. Experimental setup

Particle colliders operating at previously unexplored energy ranges are able to look for new phenomena and new, heavy particles. This rationale is essentially following Einstein's identity $E = mc^2$ and has shaped the field of High Energy Physics and the way experiments have been designed for decades. The Large Hadron Collider (LHC) has pushed the energy frontiers far into regions that thitherto had been inaccessible, and marks the beginning of a new era of experimental particle physics. Just as the collider itself, the experiments positioned along the collider are characterized by an extraordinary design, engineering and performance: the data analyzed in this thesis have been recorded with the Compact Muon Solenoid (CMS) detector. It will be described in detail in this chapter after a short introduction to the LHC, which is based on references [107, 108]. Finally the chapter explains the algorithms that translate detector signals into physics objects that an analysis can finally be performed on.

3.1. Large Hadron Collider

With a design center-of-mass energy of $\sqrt{s} = 14$ TeV for two colliding protons, the Large Hadron Collider of the Organisation Européenne pour la Recherche Nucléaire (CERN) is the most powerful particle accelerator and collider ever constructed. It is a ring of 27 km circumference and is located 45-170 m underground¹ in the area of Geneva, Switzerland, crossing the border to France. Its shape however is not perfectly circular but composed from bent and straight segments, with cavities for accelerating the protons in high-frequency electric fields and sectors of dipole magnets that force them on a circular path of radius $r = 2803.98$ m. The eight cavities are cooled down to 4.5 K and accelerate the protons with an E field of frequency 400.79 MHz, transferring an energy of 2 MeV to a proton with each passing. The strength of the dipole magnets, of which a total number of 1232 is installed along the ring, is essentially the limiting factor for the energy the particles reach. The design magnitude of the provided B field is 8.33 T (the magnets created a field of $B = 4.76$ T at a beam energy of 4 TeV during Run-I in 2012) and can only be reached with very high currents: 11850 A are flowing through their niobium-titanium cables at a temperature of 1.9 K, i.e. well below the point the material becomes superconductive and the current can flow at zero resistance. This is accomplished by cooling the magnets with liquid helium. The substance has no inner friction at this temperature and hence, as a superfluid and with its high thermal conductivity, is well suited for cooling the large magnets with a length of 15 m and a weight of 35 tons each. The cables that are wound

¹The varying depth is owed to the inclination of the layer of rock the tunnel has been built onto.

3. Experimental setup

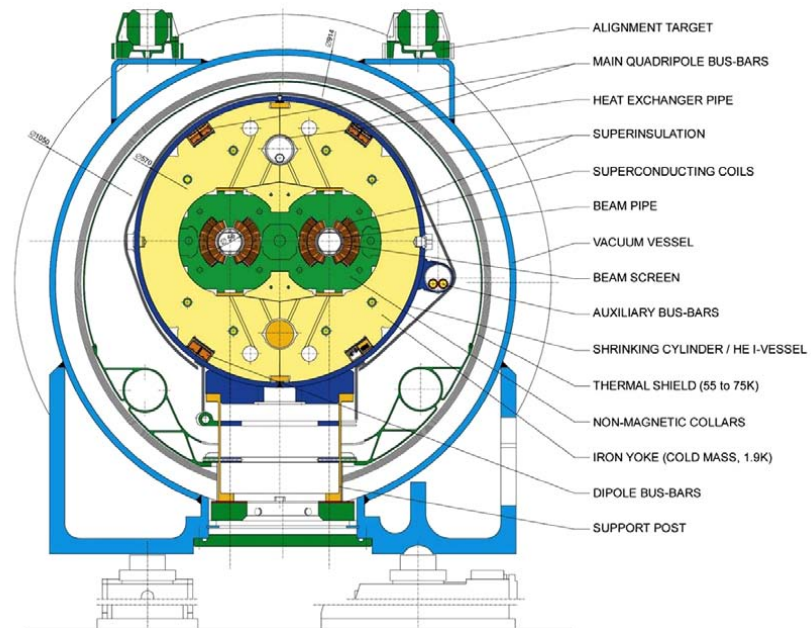


Figure 3.1.: Cross section of an LHC dipole magnet, taken from [107].

up inside the magnets consist of 36 strands which in turn are made up by thousands of filaments. The total length of the strands amounts to roughly 270000 km.

Fig. 3.1 shows the cross section of such a dipole magnet. For economical and spatial reasons (the collider has been placed inside the same tunnel previously used by the LEP experiment), a so called “two-in-one” design has been employed: two neighboring beam pipes reside in the same dipole magnet and are cooled by the same cryostat. They are surrounded by the superconducting coils cooled by the cold mass (iron) held at 1.9 K by the liquid helium.

Along the straight sections of the LHC 392 quadrupole magnets have been installed. These are responsible for focussing the beams which, according to the nominal design, consist of 2808 bunches, each made up of roughly 10^{11} protons, with a bunch separation of 25 ns. In particular at the interaction points this happens in a single pipe where the beams are squeezed to a minimal diameter ($16 \mu\text{m}$) to reach maximal luminosities. Besides sheltering the collision points and the cavities, also beam dumping and beam injection happen in the straight sections.

The LHC is not capable of accelerating the protons from rest. At the time of injection into the main LHC ring, the protons already travel with an energy of 450 GeV. This is done with a chain of pre-accelerators which are part of the accelerator complex as depicted in Fig 3.2: the journey starts in the linear accelerator LINAC 2 where, after the electrons have been stripped from the initial hydrogen atoms, the protons are accelerated to $E = 50 \text{ MeV}$ before reaching the Proton Synchrotron Booster (PSB). Here they reach 1.4 GeV and are injected to the Proton Synchrotron (PS). By the time they leave the PS

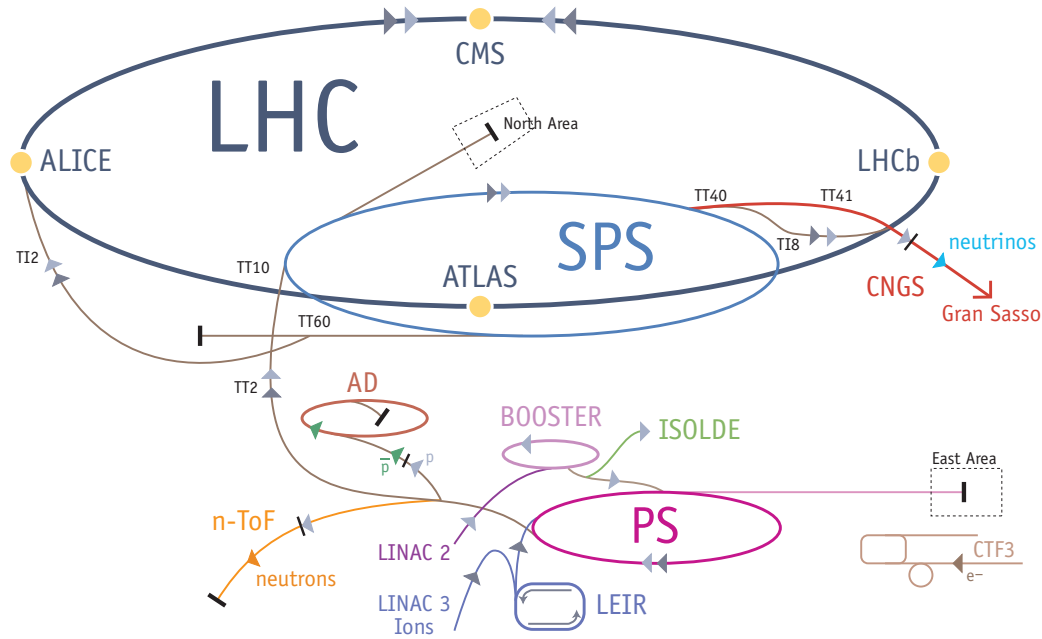


Figure 3.2.: The LHC accelerator complex [108].

the protons have an energy of 25 GeV and have adopted the bunch spacing needed by the LHC. As a last pre-acceleration step, they are brought to an energy of 450 GeV in the Super Proton Synchrotron (SPS). It is worth mentioning that the pre-accelerators have small experiments attached to them, investigating proton physics at lower energies than the ones of the LHC program. The time it takes to fill the LHC ring with a proton beam is 4 min 20 s. Finally the protons are accelerated to their target energy for about 20 minutes before bringing them to collision. The standard filling scheme is explained in [109].

The intensity of the beam will decrease as pp collisions continuously are taking place. The nominal period the LHC can operate with a beam is ~ 15 hours, which is the time by which the luminosity is expected to have dropped by a factor of $1/e$. The luminosity is given by the equation

$$L = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \varepsilon_n \beta^*} \cdot \frac{\overbrace{1}^{:=f}}{\sqrt{1 + \left(\frac{\theta_c \sigma_z}{2r_{\text{beam}}}\right)^2}},$$

where it is assumed the two colliding beams share the same beam parameters. Consequently, N_b is the number of protons in a bunch and n_b is the number of bunches per beam. The revolution frequency of a bunch is given by f_{rev} . In order to account for the relativistic effects occurring close to the speed of light, γ_r is the relativistic gamma factor of the proton. Both ε_n and β^* are parameters related to the beam size and are determined by the process of bunch preparation and the magnet configuration, respectively. The normalized transverse beam emittance ε_n is a measure of parallelism of the

3. Experimental setup

beam constituents and of their transverse speed. The amplitude function β^* , specifying how strongly the beams get squeezed by the focussing magnets, gives the distance from the collision point where the beam is twice as wide. The nominal LHC configuration foresees $\beta^* = 55$ cm and $\varepsilon_n = 3.75$ μm . The factor f describes the relative geometry of the two colliding beams. They are not collided head on. If they were totally parallel, this would result in many unwanted pp collisions along the straight pieces of the beam pipe extending from the collision point. Instead they are led together at a small angle $\theta_c \sim 300$ μrad , at the cost of a slightly smaller likelihood for two colliding protons, i.e. at the cost of a reduced luminosity. This angle depends on the length of a bunch, σ_z , and on the beam radius $r_{\text{beam}} = \sqrt{\varepsilon_n \beta^* / \gamma_r}$. The different experiments placed along the LHC ring operate at different luminosities, ranging from 10^{27} $\text{cm}^{-2}\text{s}^{-1}$ to the maximal design value of 10^{34} $\text{cm}^{-2}\text{s}^{-1}$. The former number refers to heavy ion runs, where $^{208}\text{Pb}^{82+}$ lead ions, fully stripped from their electrons, are brought to collision. When operating with a magnetic field in the dipoles of 8.33 T, each of the 208 nucleons carries an energy of $\sqrt{s}/2 = 2.76$ TeV.

3.1.1. Experiments placed along the LHC ring

A short description of the seven experiments placed along the LHC ring is provided below:

ALICE: Shortly after the Big Bang quarks and gluons did not form bound states. The ALICE experiment [110] investigates this so called quark-gluon plasma in remnants of heavy ion collisions.

ATLAS: Being one of the two multi-purpose detectors installed at the LHC, the ATLAS experiment [111] both precisely measures well known SM processes and searches for new physics such as dark matter, supersymmetry or microscopic black holes.

CMS: The oeuvre of the CMS experiment is similar to the one at ATLAS; together they discovered the 125 GeV boson in 2012 and are looking for physics beyond the SM as well as probe the theory up to ultimate precision. The data analyzed in this thesis have been recorded with the CMS detector; it will be described in detail in Sec. 3.2.

LHCb: The CP violation and exotic BSM phenomena occurring in B hadron decays are the primary subjects of the LHCb experiment [112]. Recently it has also reported the observation of bound states formed by five quarks, the so-called pentaquarks [113].

The following experiments are smaller in terms of the collaboration and more specialized in their goals. They are located in the vicinity of the larger experiments described above.

LHCf: Studying neutral particles that emerge from a pp collision into the very forward region and thereby improving the modelling of highly energetic cosmic rays is the aim of LHCf [114]. It is located near the

3.1. Large Hadron Collider

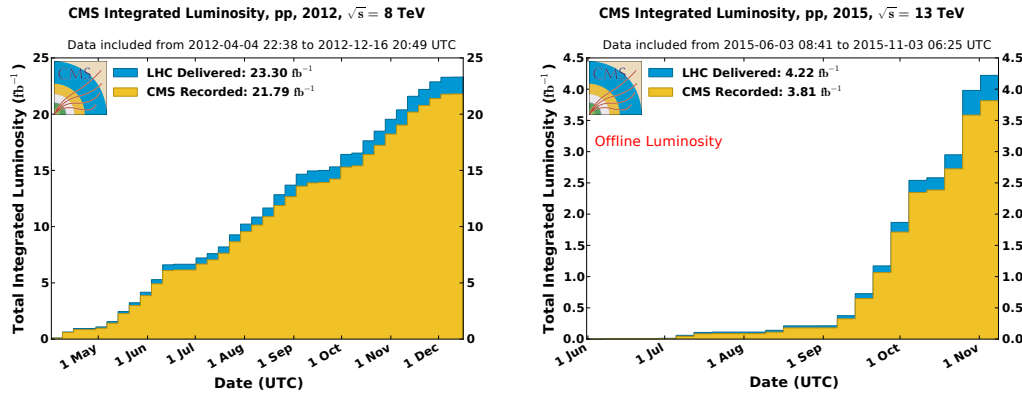


Figure 3.3.: Left: during the 2012 pp data taking period from April to December 2012, 23.3 fb^{-1} have been delivered by the LHC. CMS recorded the products of 93.5% of these collisions, which can be used for physics analyses. Right: the total integrated luminosity recorded at 13 TeV amounts to 3.8 fb^{-1} . Taken from [117].

ATLAS detector.

MoEDAL: This experiment is directly searching for magnetic monopoles and highly ionizing massive stable particles [115]. It is placed next to LHCb.

TOTEM: The TOTEM apparatus is integrated into the CMS experiment, but it is an independent detector aiming at measuring the total pp cross section and studying diffractive and elastic scattering [116].

3.1.2. LHC operation

After an incident during tests in 2008 and a long repair time, it has been decided to start the LHC program with reduced energies and not reach ultimate values until after the first years of a successful operation. During the year 2012 the LHC delivered pp collisions at $\sqrt{s} = 8$ TeV with an integrated luminosity of 23.3 fb^{-1} . The time interval separating two bunches was 50 ns, which is twice as large as the design bunch crossing time of 25 ns. The amplitude function β^* was 60 cm. With a peak instantaneous luminosity of $0.7 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ the interaction rate was already close to the design value. Fig. 3.3 shows how the total integrated luminosity grows with time. The CMS detector was able to record 21.8 fb^{-1} , out of which 19.7 fb^{-1} fulfilled certain quality criteria and could be analyzed for this thesis. The efficiency (recorded/delivered) is not 100% because of downtime of various detector components and down- and dead time of data acquisition systems.

The first stable beams at $\sqrt{s} = 13$ TeV circulated on June 3rd, 2015. After a short period of data taking with a bunch spacing of 50 ns, the protons collided every 25 ns, delivering 4.2 fb^{-1} of data in total (see also Fig. 3.3). The maximal instantaneous luminosity was $0.513 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. For some intervals of the data taking period, the magnetic field in the CMS detector had been turned off for reasons related to problems in the cooling

3. Experimental setup

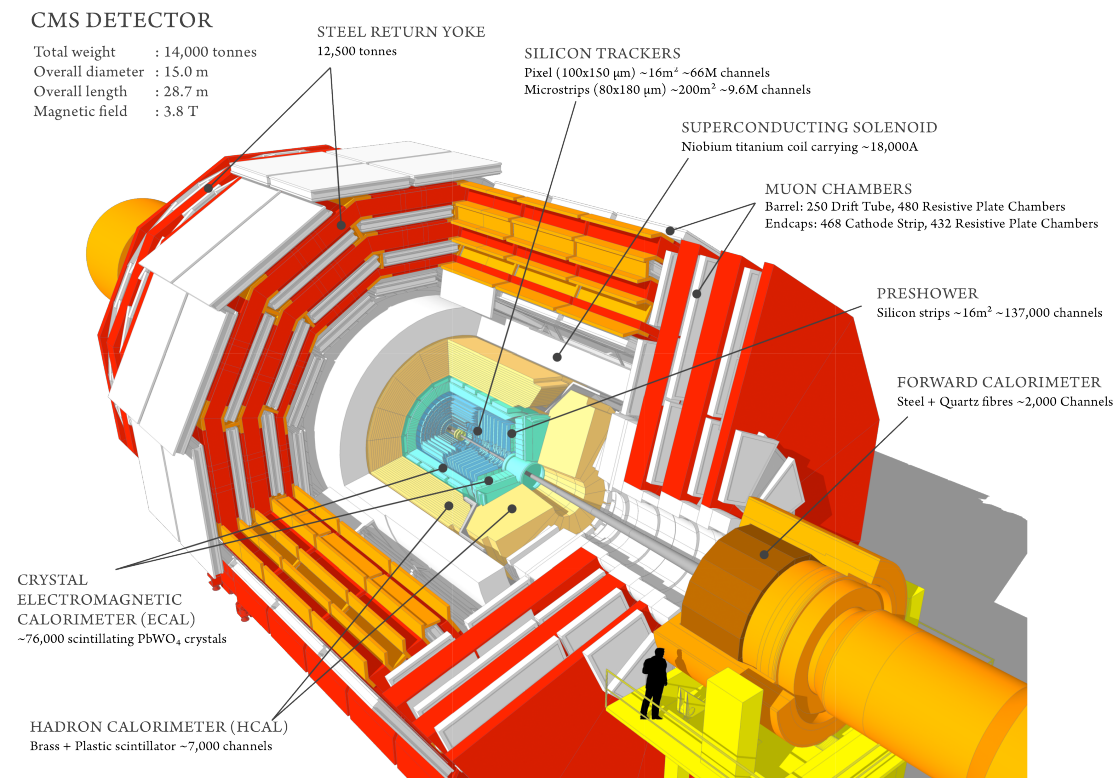


Figure 3.4.: Illustration of the CMS detector [118]. A person is shown in order to get a feeling for the dimensions of the detector systems. The beam pipe is vanishing into the LHC tunnel to the lower right.

system. These data thus cannot be used for the physics analysis conducted in this thesis. Only 2.1 fb^{-1} can be analyzed until the middle of 2016, when the LHC schedule foresees the next proton-proton collisions.

3.2. Compact Muon Solenoid

This section is based on references [119–121]. It portrays one of the two general-purpose detectors at the LHC, the CMS experiment, with which the data analyzed in Chap. 4 have been recorded. Located 100 m underground in Cessy, France, the CMS detector lies in between the Jura mountains and the Lake Geneva.

The apparatus has a cylindrical shape and is comprised of several individual elements that are assembled in layers around the beam pipe: the innermost part consists of a silicon-based tracking system for reconstructing the trajectories of charged particles. Then follow electromagnetic and hadron calorimeters for measuring the energy of electrons, photons and hadrons. In order to not disturb the particle's original path it is important for the tracking system to have interactions with the particle that are as minimal as is necessary for the reconstruction of its moving direction and its momentum. The

calorimeters on the other hand aim at stopping them completely and measuring the released energy. A solenoid made up of a superconducting niobium-titanium coil is the next subelement. A current of 18000 A is required to provide a B field of 3.8 T with force lines parallel to the beam axis in order to bend the particles on a helical path which allows momentum measurements transverse to the field lines of B according to the Lorentz force law. Operating with such high currents means the coil must be cooled down to 4.5 K and become superconducting. The solenoid is embedded in a 12500 t steel yoke returning the magnetic flux and creating an inverted polarity of the B field in the outer detector elements with respect to the tracker system. Finally the gaseous detector layers of the muon system are interleaved with the iron layers of the yoke. The muon chambers are placed in the outer regions of the detector, since muons are the only detectable particles which are expected to penetrate the entire detector material because of their low ionization potential. Altogether the CMS detector weighs 14000 t and has transverse and longitudinal dimensions of 15.0 m and 28.7 m, respectively. An illustration is shown in Fig. 3.4. As can be seen from this figure, the detector provides an almost- 4π coverage around the interaction point. The ample instrumentation and the precision of both tracking and calorimetry systems are owing to the experiment's goals – the investigation of the Higgs boson and the search for rare exotic BSM particles. These phenomena are expected to occur at low expected rates and therefore ask for a highly efficient particle detection and an accurate measurement of momenta and energies.

The righthanded coordinate system is chosen such that the x axis points inwards to the center of the LHC ring, the y axis points vertically towards the surface and the z axis specifies the direction along the ring in counterclockwise orientation. Given the experimental geometry, it is also useful to introduce the three coordinates r , ϕ and θ . The first two coordinates are defined in the x - y plane. The radial distance with respect to the nominal detector center is $r = \sqrt{(\Delta x)^2 + (\Delta y)^2}$; ϕ denotes the azimuthal angle from the x axis. The polar angle θ measures the angle towards the beam axis. It has furthermore become convenient to measure the forward direction of a particle by means of the rapidity or pseudorapidity. They are defined as

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \quad \eta = \frac{1}{2} \ln \left(\frac{|\mathbf{p}| + p_z}{|\mathbf{p}| - p_z} \right).$$

The rapidity y is a practical quantity since the differential cross section in this observable, $d\sigma/dy$, is invariant under Lorentz-boosts in the z direction. The pseudorapidity η diverges for a particle with a direction of flight parallel to the beam axis, while it is 0 for a particle moving perpendicular to the beam. This means it can also be expressed by

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right).$$

For a massless particle with $E = \sqrt{|\mathbf{p}|^2 + m^2} = |\mathbf{p}|$ the two quantities are identical. A detailed description of the tracker, the calorimeters and the muon system is provided below.

3. Experimental setup

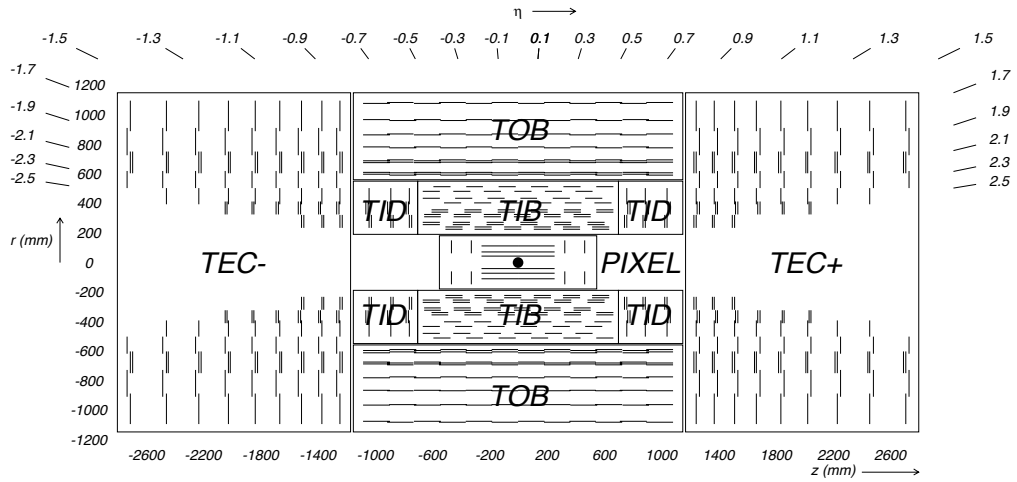


Figure 3.5.: Illustration of the CMS tracking system in the r - η plane. The part closest to the beam spot is a pixel detector, surrounded by the tracker inner barrel (TIB) and tracker outer barrel (TOB). The forward regions are equipped with the tracker inner disks (TID) and the tracker end caps (TEC). Each of the short lines represents a pixel or strip module. Some of them are mounted back-to-back (double lines) at a small angle and provide stereo hits from which the coordinate parallel to the alignment of the modules can be inferred.

3.2.1. Tracker

Being the detector element closest to the interaction point, the tracker is subject to large particle fluxes and highly energetic particles and thus to severe potential radiation damage. The material must be stable enough to operate in this environment while at the same time meeting its original goals: precisely and efficiently measure the tracks of charged particles, which also implies a resolution high enough to resolve secondary vertices as they occur if B hadrons are present; this is particularly relevant for the search for tHq , $H \rightarrow b\bar{b}$.

A silicon detector is the reasonable choice because of the granularity that can be achieved with a silicon-based design, the fast read-out time and the radiation hardness.

The innermost pixel detector covers a range of $-2.5 < \eta < 2.5$ (see Fig. 3.5). There are 66 million pixels, which are ordered in three cylindrical layers at radii 4.4, 7.3 and 10.2 cm and two times two endcap disks placed at ± 34.5 cm and ± 46.5 cm in the z direction. Most pixels have an area of $100 \times 150 \mu\text{m}^2$. However a spatial resolution in the transverse (longitudinal) dimension of $10 \mu\text{m}$ (20 - $40 \mu\text{m}$) [122] can be achieved via interpolation of the signal pulses caused by a single initial charge, which are spread over multiple cells because of its Lorentz drift in the strong B field (and E field in the sensor). Since the third dimension is determined by the position of the sensor plane, a three-dimensional (3D) reconstruction of the hit is possible. In total the pixel detector covers an area of about 1 m^2 .

About 9.6 million silicon strips with a combined active area of 200 m^2 are used in the

strip detector. Strips with n-type doping are placed onto a bulk material of p-type silicon, the ionizing particle creates a hit in a bulk material, and the coordinate orthogonal to its length can be read out in terms of a signal in the strips. Some layers are made up of two modules mounted back-to-back at stereo angle in order to also measure the second coordinate.

Four different subsystems are employed: the tracker inner barrel (TIB) has four concentric cylinders extending from $z = -70$ cm to $+70$ cm, which are placed at r values from 20 cm to 50 cm; they are equipped with strips parallel to the beam axis. The two innermost cylinders deploy the kind of double-sided modules as was explained previously. The distance between two strips (strip pitch) ranges from $80 \mu\text{m}$ for double-sided to $120 \mu\text{m}$ for single-sided modules. At z positions of ± 80 cm and ± 90 cm sit the tracker inner disks (TID), which are assemblies of radially orientated strip modules organized in three rings. The TIDs stretch across the same r range as the TIB and extend the region that is hermetically covered with silicon strips to $|\eta| = 2.5$. Similar to the TIB, the two innermost rings consist of double-sided modules. From $r = 55$ cm to 110 cm the tracker outer barrel (TOB) is installed. Six cylinders with strip sensors are adjusted in a way that guarantees an overlap of the strips of two neighboring cylinders and thus eliminates dead spots. Again a back-to-back design of the innermost layers ensures a 3-D position reconstruction.

The tracker endcaps (TEC) are mounted in the range $124 \text{ cm} < |z| < 182 \text{ cm}$ and accommodate nine disks with silicon strip modules arranged in rings around the beam axis. The number of rings is 7 for the disks closest to the interaction point and decreases to 4 for the outermost disk. The strip lengths on the other hand increase from 10 cm to 25 cm. The strip pitch is of the order of $150 \mu\text{m}$. As any other element the TECs are equipped with back-to-back modules in inner regions.

This layout guarantees that a charged particle within $|\eta| < 2.5$ creates at least three pixel cell hits and nine hits in strips, the majority of which allow for a 3-D reconstruction of the exact hit position. The entire system is cooled down to $+4^\circ\text{C}$ in order to prolong the life time of the sensors exposed to intense radiation.

3.2.2. Electromagnetic calorimeter

Moving away from the interaction point, a particle passing the tracker then enters the electromagnetic calorimeter (ECAL). The core part of the ECAL is composed of ~ 76000 scintillators made of lead tungstate (PbWO_4) crystals. The choice of material is driven by the following reasoning: with a density of $\rho = 8.28 \text{ g/cm}^3$, a relatively short radiation length² of $X_0 = 0.89 \text{ cm}$ and a Molière radius³ of $R_M = 2.2 \text{ cm}$, the calorimeter can maintain a compact design at a good spatial resolution, meeting the task of the ECAL to completely stop electrons and photons and measure their initial energy. Moreover, the photons are emitted quickly during the absorption of the particles, delivering 80% of the

²The length at which the energy of an electron traversing the material has dropped to E_0/e due to radiation.

³A material constant denoting the radius of a hypothetical cone in which 90% of the energy of the electromagnetic shower initiated by a relativistic photon or electron is contained.

3. Experimental setup

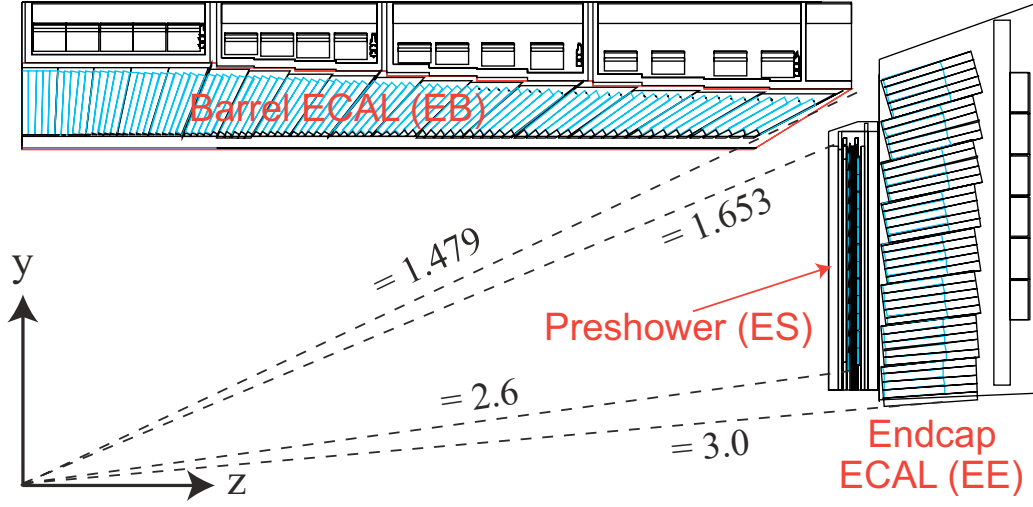


Figure 3.6.: The electromagnetic calorimeter consists of a barrel component and active material in the endcap. A preshower detector (ES) is employed to decide whether or not photons stemmed from π^0 decays. Taken from [120].

light within 25 ns, which is the order of magnitude in scintillation decay time needed when the LHC is operated under nominal conditions. The entire ECAL system is held at constant 18.0 °C, as the number of scintillation photons, N_γ , has a strong dependence on the temperature with $\partial N_\gamma / \partial T < 0$.

As can be seen in Fig. 3.6 the ECAL system consists of the barrel (EB) and the endcap (EE) part. The EB crystals have a length of 23 cm, which corresponds to 25.8 radiation lengths, and a cross section of $22 \times 22 \text{ mm}^2$ at their front face oriented towards the inside of the detector. The 61200 crystals in the barrel region ($|\eta| < 1.479$) are not exactly oriented towards the beam spot but build a small angle with a fictitious vector originating from it. This way the situation where a photon travels between two crystals for its entire path and is not detected is circumvented. The scintillation light is collected by photodiodes on the $26 \times 26 \text{ mm}^2$ large rear face.

The 7324 crystals in each endcap (EE) cover the region between $1.479 < |\eta| < 3.0$. They have a cross section of $\sim 30 \times 30 \text{ mm}^2$, a length of 22.0 cm ($24.7 X_0$) and are only minimally tapered. In order to avoid dead spots, they are installed with increasing angle with respect to the beam axis the further one goes away from it. Their common focus point lies roughly 1 m beyond the nominal detector center. Vacuum phototriodes collect the photons and the rear face of the crystals.

According to [123] the energy resolution of the ECAL has been measured with electron test beams and is

$$\left(\frac{\sigma(E)}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 \oplus \left(\frac{12\%}{E}\right)^2 \oplus (0.3\%)^2,$$

where the first term on the right-hand side accounts for the fact that the evolution of the

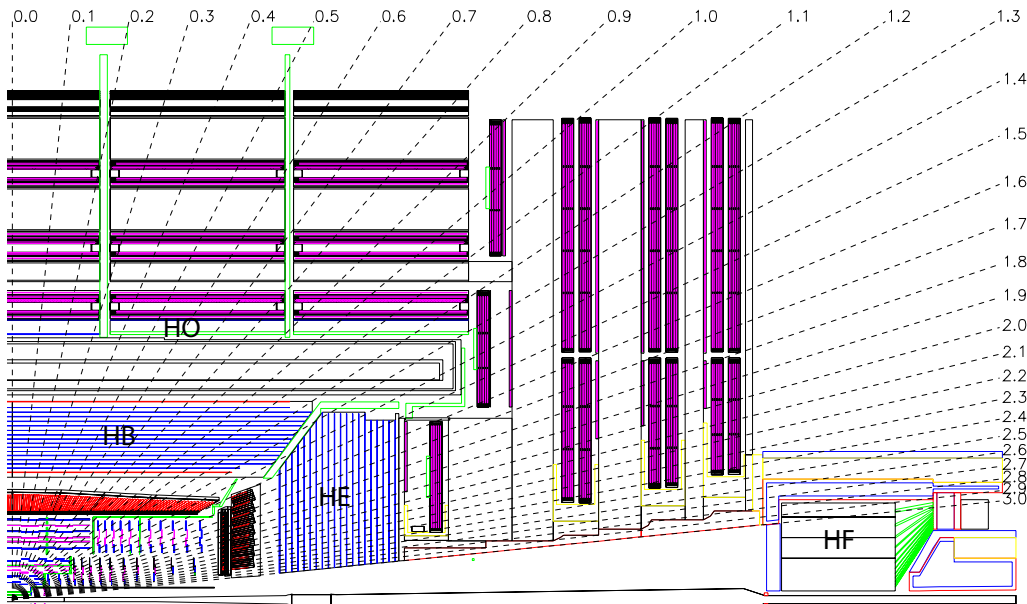


Figure 3.7.: Four calorimeters (HB, HO, HE and HF) are installed for measuring the energy of hadronic showers [119].

electromagnetic shower is not exactly the same from event to event but is a stochastic process. The second term expresses the electronic noise that is not dependent on the energy E of the initial particle; the third term describes effects like the loss of photons at the rear end of the crystals which further decrease the resolution and are proportional to E (so that $\sigma/E = \text{const.}$).

The spatial resolution of the EE is not good enough for resolving two closely neighbored photons as they could emerge in a neutral pion decay. For identifying such cases the pre-shower detector (ES) is installed in a region $1.653 < |\eta| < 2.6$. It is 20 cm thick and is made up of two layers consisting of lead radiators with a plane of silicon strip sensors placed at their ends. The strips in the second plane are oriented orthogonally to the ones in the first plane in order to be able to reconstruct the impact point in two dimensions. Photons are expected to shower before reaching the second plane in 95% of the cases, and with the high granularity of the silicon strips placed at the ends one has a good handle for resolving the photons and thus distinguishing them from the photons of a pion decay.

3.2.3. Hadron calorimeter

In contrast to the ECAL, for which the same material is used for creating effects of absorption and scintillation, the hadron calorimeter (HCAL) is designed as a sampling calorimeter with interleaved layers of different materials optimized for the two purposes. In this case the aim is to absorb and measure the energy of hadrons, which typically

3. Experimental setup

penetrate the innermost elements of the tracker and the ECAL. Hadrons entering the material create a hadronic shower of secondary particles such as pions via processes involving inelastic hadron interactions and deexcitation of nuclei. The longitudinal evolution of the shower happens considerably slower than for electromagnetic showers⁴ and is characterized by the nuclear interaction length λ , denoting the mean free path of the hadron before it undergoes an inelastic interaction with the material.

Brass is mainly used as absorber. It has an interaction length of $\lambda = 16.42$ cm and a density of 8.53 g/cm³. Plastic scintillation tiles are responsible for bringing out the light which is then collected in wavelength shifting fibers and finally collected by silicon-based, hybrid photodiodes suited for operating in a 4 T magnetic field [124].

The detector is instrumented with hadron calorimeters in four different regions: in the barrel the HCAL, which is labelled HB in Fig. 3.7, covers the range in $|\eta|$ up to 1.3 and is mounted between the ECAL and the solenoid at a radial distance between 1.81 m $< r < 2.86$ m. The geometry dictates how much absorbing material can maximally be used; for this reason the hadron outer (HO) is installed, catching tails of hadronic showers which have not been totally absorbed before. The HB latter consists of 36 azimuthal wedges made of brass absorber plates which are aligned parallel to the beam. The plates are interleaved with the plastic scintillator material; 70000 scintillator tiles are used in the EB. Stainless steel plates mounted on the innermost and outermost layer make the HB structurally more stable. The 14 layers of brass (with varying thicknesses from 50.5-56.5 mm) result in an effective absorber thickness of 5.82λ at $\eta = 0$ and 10.6λ at the end of coverage around $\eta = 1.3$.

As mentioned above, showers that were not entirely contained in the EB and HB can be additionally registered in the HO. In the central wheel of the barrel the HO is made up of two scintillator planes enclosing a 19.5 cm thick iron piece as absorber material. The four outer wheels only have one plane of scintillator tiles. This is owed to the fact that the solenoid serves the purpose of an additional absorber, and the effective length a particle has to traverse in the solenoid is equal to $1.4/\sin\theta$ interaction lengths and thus grows the more forward the particle's direction is. The system consisting of ECAL, HCAL and the solenoid has a combined thickness in absorbing material of 11.8λ . The same photon read-out system is employed as in the HB.

In the region between $1.3 < |\eta| < 3.0$ the hadron endcap (HE) is installed. Again a combination of brass as absorber material and scintillator trays equipped with 20916 plastic tiles and optical fibers for routing the light to photodiodes is used. The brass plates, staggered in a way there is no dead material in the projection on the beam spot, have a thickness of 79 mm with gaps of 9 mm for the scintillators. In this pseudorapidity range the absorbing layers of the two calorimeters add up to ~ 10 interaction lengths.

The main challenge for any material when being installed in the hadron forward (HF) covering a range of $3.0 < |\eta| < 5.2$ is the extremely high particle flux in this region. Quartz fibres have a high resistance to radiation and are consequently used as active medium. Charged particles arising in a shower in the 1.65 m thick steel absorber

⁴However, a hadronic shower has also a non-negligible fraction of electromagnetic energy, since about 1/3 of the secondaries consists of neutral π^0 mesons subsequently decaying to two photons.

can be detected in the fibres in terms of Cherenkov light if their energy exceeds the threshold determined by the speed of light in quartz. For instance, the energy of an electron entering the fibres must exceed 190 keV in order to generate a signal. This is then amplified and detected in photomultipliers. Given its design the HF is predominantly sensitive to the electromagnetic part of the shower. This means a method is needed for distinguishing between hadronic and purely electromagnetic showers. Since the latter evolve more quickly, they tend to deposit their energy in the first centimeters of the steel plate. Therefore only half of the quartz fibres stretch across the entire range of absorber material. The other half starts only after the first 22 cm of steel and is read out independently.

The resolution of the HCAL can be parametrized as

$$\frac{\sigma(E)}{E} = \frac{S}{\sqrt{E}} \oplus C,$$

and was measured in [125] as $S = 198\%$, $C = 9\%$ for the HF. In [126] and [127] it has been determined for the HB and HE, respectively. The numbers are $S = 120\%$ (153%), $C = 9.5\%$ (6.3%) for the HB (HE). These numbers refer to resolutions achieved with pion test beams.

3.2.4. Muon system

All detectable particles except for muons are expected to have deposited their entire energy in the calorimeters. Due to the large rest mass of muons they do not lose a significant amount of energy via bremsstrahlung, and also interact only minimally with the detector materials, i.e. they hardly ionize the atoms. Therefore muons can still be detected in the outermost regions of the detector, where the gaseous detectors are installed exactly for this purpose. As the name of the experiment suggests, a precise measurement of these particles is crucial for a vast range of analyses involving heavy, weakly decaying resonances like top quarks, W or Z bosons, since muons typically are possible final products at the end of the decay chain. As such their precise detection in the muon system, which complements information of the inner tracker and thus improves the momentum resolution, is fundamental for SM measurements as well as for searches motivated by BSM theories.

Three different types of detectors are employed and assembled as shown in Fig. 3.8: 250 drift tube chambers (DTC) are placed in the barrel region of the detector ($|\eta| < 1.2$), grouped in four concentric cylinders (MB1-4), called muon stations, around the beam axis. They are interleaved with the steel return yoke. Five such wheels are mounted along the z direction. The three inner stations consist of 12 drift tubes chambers, the outermost has 14 DTCs (see the turquoise elements of Fig. 3.9). The smallest elements of each DTC are the drift cells (DC) with a cross section of $13 \times 42 \text{ mm}^2$. Fig. 3.10 illustrates the design of such a cell. It is filled with an Ar/CO₂ gas mixture (85%/15%) which is ionized by a traversing muon. In the electric field created by the anode, the electrode strips and the cathode strips, both of which are aluminum tapes placed on insulators, the

3. Experimental setup

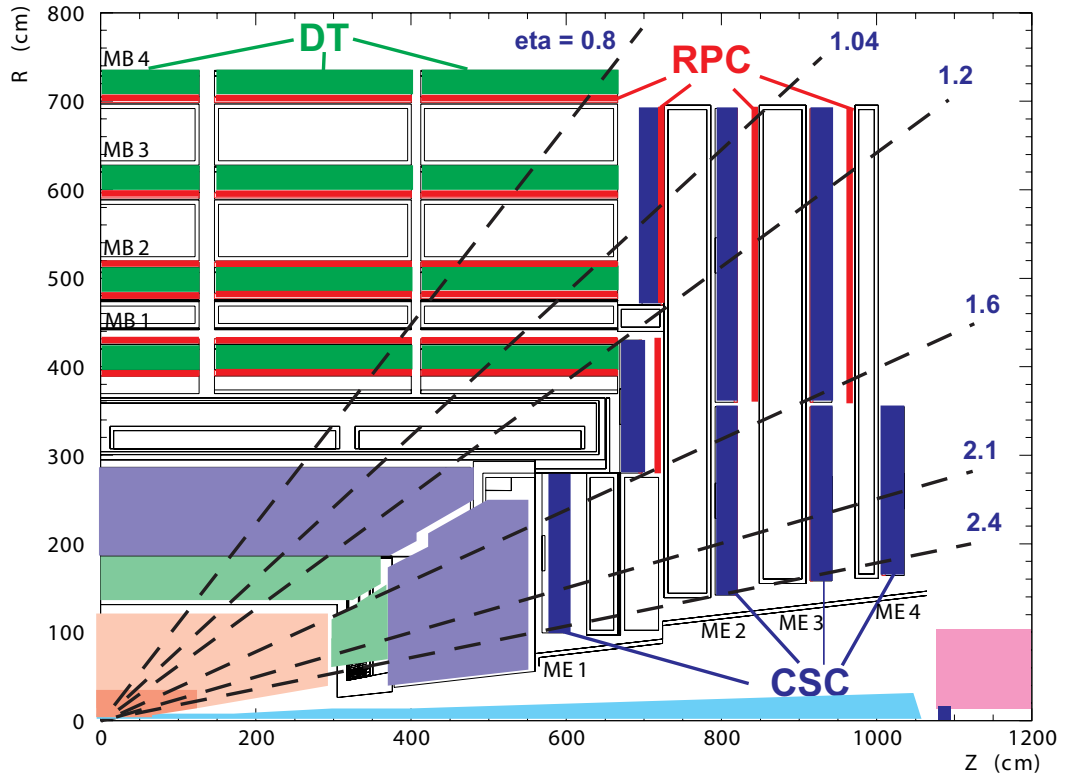


Figure 3.8.: The muon system consists of three different types of gaseous detectors: the drift tubes (DT) in the barrel, the cathode strip chambers (CSC) in the endcap regions and the resistive plate chambers (RPC). Taken from [120].

released electrons are pulled to the wire (gold-plated stainless steel) and lead to a charge avalanche which can be measured as a current. The cells are stacked into superlayers (SL) with each SL consisting of four layers of DCs staggered by half a cell to eliminate efficiency dead-spots. The DTCs in the three inner muon stations deploy three SLs. The orientation of the wire of the inner and outer SL is along the beam line. Consequently these SLs can measure the particle's r - ϕ coordinates. In between them the third SL is installed which measures the z coordinate, since its wires are perpendicular to the beam axis. This allows for a 3D reconstruction of the trajectory. In the outermost station one can only measure the muon's r - ϕ coordinates since this station lacks the SL orthogonal to the beam. About 170000 sensitive wires are installed in total, achieving a resolution of $100\ \mu\text{m}$ in the r - ϕ plane.

The cathode strip chambers (CSC) cover ranges in pseudorapidity of $0.9 < |\eta| < 2.4$. In this forward region the expected higher number of charged particles and the non-uniformity of the magnetic field discourage the use of drift tubes and demand a different design: a CSC is comprised of seven trapezoidal panels stacked on top of each other,

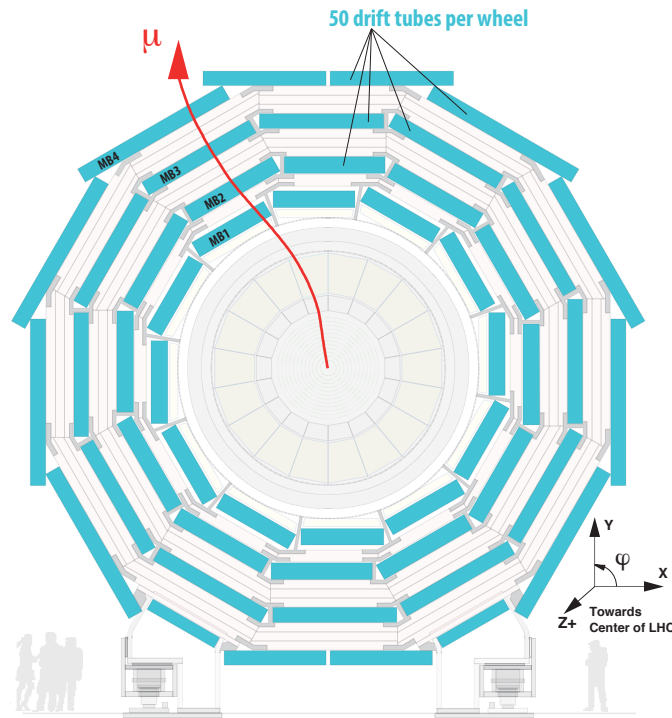


Figure 3.9.: Illustration of a drift tubes wheel installed in the barrel region of the CMS detector. It consists of four concentric chambers labelled MB1-4. The red bent arrow indicates the flight direction of a muon originating from a hard interaction at the beam spot. Adapted from [119].

creating six gaps filled with 40%Ar, 50% CO₂ and 10% CF₄.⁵ Sensitive anode wires (gold-plated tungsten) are wound azimuthally around three of the panels and provide a measurement of the r coordinate. The winding distance between the wires is 3.2 mm. The measurement in ϕ comes from 80 radial cathode strips (with widths of 8.4 mm to 16 mm) milled onto each panel by interpolating the detected charges induced on the individual strips. The entire system of sensitive planes has an area of 5000 m² and 220000 and 180000 read-out channels for the cathode strips and the anode wires, respectively. For the two CSCs of the ME1 section closest to the interaction point (see again Fig. 3.8), the spatial resolution is $\sim 75 \mu\text{m}$; it is about twice as large in the other CSCs.

The resistive plate chamber (RPC) system complements the two formerly described muon detectors, which have a very high spatial granularity, by providing excellent time resolution in ~ 1000 read-out channels. RPCs are assembled from two parallel plates, a positively-charged anode and a negatively-charged cathode, both made of a very high resistivity plastic material and separated by a gas volume. They are able to resolve two signals within a time much smaller than the nominal bunch separation of 25 ns dictated

⁵The reader is deferred to [128] for a study on the optimization of the gas composition.

3. Experimental setup

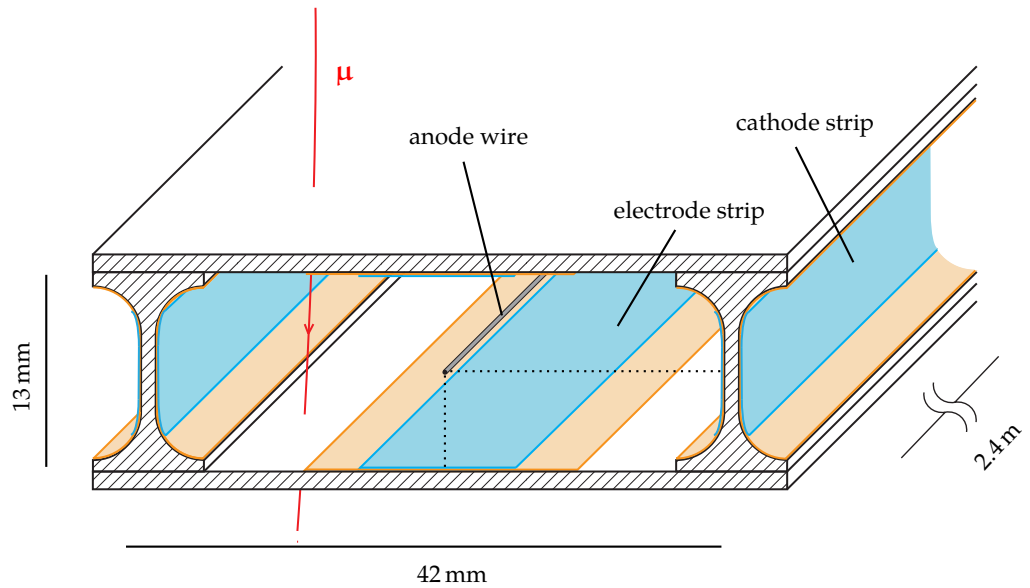


Figure 3.10.: Schematic view of a drift cell. It is filled with a gas mixture of Argon and CO₂. The applied voltages are +3600 V for the anode wire, +1800 V at the electrode strips and -1200 V for the cathodes. The brownish plates are made of mylar and serve the purpose of insulators. Adapted from [119].

by the operation of the LHC. This results in an unambiguous assignment of detected muons to the bunch crossing. The central region is covered by six layers of RPC chambers encompassing the muon stations, with the detecting strips aligned with the beam. In the endcap RPCs are oriented radially and are installed up to $|\eta| = 1.6$. During the shutdown of the LHC from 2013-2015, the coverage was extended to $|\eta| < 2.1$.

The entire muon system has 100% coverage of the region with $|\eta| < 2.4$. Studies on a sample with simulated single muons resulted in a reconstruction efficiency of 95-99% except for the transition regions, where the efficiency is decreased. The momentum resolution measured by the muon system alone was found to be 15-40% depending on the η of a muon with a transverse momentum of 1 TeV, and 9% for small η and a p_T lower than 200 GeV. If one combines with information from the tracker, the resolution is improved by roughly one order of magnitude.

3.2.5. Trigger system

With a rate of 40 MHz, the expected number of pp collisions per second produces an amount of information which can impossibly be processed and stored with available CPU and storage capacities. Hardware and software based solutions have to be designed in order to significantly reduce the rate to a reasonable degree the computing infrastructure is able to cope with. At the same time the experiment cannot afford that possibly

interesting events are discarded. A two-level trigger system has been optimized to meet these challenges and can therefore be seen as the first step of the event selection.

The Level-1 Trigger (L1T) [129] is based on programmable hardware such as ASICs (application specific integrated circuit) and configured FPGAs (field programmable gate arrays) and exploits coarse information received from the calorimeters and all three muon detector types (DT, CSC, RPC). During the decision-making, which takes $3.2 \mu\text{s}$, the high-resolution data and the entire tracker information of the bunch crossing are pipelined in appropriate memories. An event is accepted by the L1T if its global trigger level has a positive response (GT). This GT combines information from the global calorimeter trigger (GCT) and the global muon trigger (GMT) and evaluates properties like the missing transverse energy or the scalar sum of energy deposits in the transverse plane provided by the GCT, and tracks of up to four muon candidates in the muon detectors.

With a maximum output rate of 30 kHz ⁶, the L1T accounts for about 50% of the desired reduction factor of 10^6 with respect to the initial rate. Any further reduction must happen via a software-based rejection of events: the high-level trigger (HLT) is a computer cluster consisting of $O(1000)$ CPUs and discarding events on the basis of dedicated algorithms similar to the ones used in the standalone, offline reconstruction software CMSSW [120]. A reduction to below 400 Hz is achieved by the HLT.

3.2.6. Detector simulation

In order to compare the data collected with CMS to theoretical predictions, the particles of simulated events obtained with the Monte-Carlo technique (see Sec. 2.2 and 2.3) must undergo interactions with the detector material in the same way as the real particles that have left their footprint in terms of signals in the detector components.

Therefore the entire CMS detector including all of its sub-systems and read-out electronics has been simulated precisely, taking into account their respective material and geometries. The detector model is based on the `GEANT4` package [130]. This allows for a 1:1 comparison of “artificial” to “real” event, because their output format and content handed over to the event reconstruction explained in Sec. 3.3 is exactly the same. In fact the algorithms there do not know about the origin of an event – whether it comes from Monte-Carlo simulation or data sample.

3.2.7. Computing infrastructure

Several types of data are stored at different locations, the entity of which is referred to as Worldwide LHC Computing Grid (WLCG) [131]. The computing grid realized and maintained by the CMS collaboration is organized in so-called Tiers as illustrated in Fig. 3.11: the raw, unprocessed detector information is stored at Tier-0 centres at CERN and Budapest, respectively, as are primary datasets which have undergone a first reconstruction. The application of the full reconstruction algorithms explained in the next section is done at Tier-1 centres like the GridKa at the Karlsruhe Institute of Technology

⁶The design output rate is 100 kHz , however a safety factor of ~ 3 is assumed.

3. Experimental setup

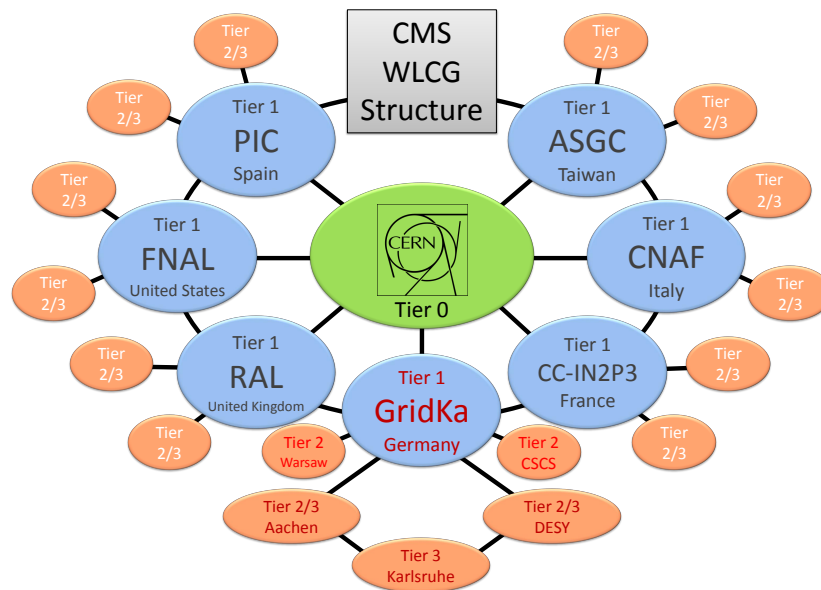


Figure 3.11.: The CMS computing grid [132] has a hierarchical design: Tier-0 centres store the unprocessed data and run first reconstruction algorithms. Tier-1 centres are big facilities for full event reconstruction and simulation. Tier-2/3 centres save and provide copies for the end user.

in Germany. These national computing facilities also provide the CPU needed for the simulation of datasets for the comparison with recorded data. Tier-2 centres are facilities situated all over the world which store and process data on a regional level in order to guarantee local universities quick and easy access to copies of the datasets. The term Tier-3 refers to the computing and storage infrastructures the analyst has immediate access to. Typically these are the machines the user runs on locally to analyze highly customized datasets.

3.3. Reconstruction of events

A high energy inelastic proton-proton interaction and the particles emerging from it leave certain patterns in the different detector elements discussed in the last chapter. Algorithms must be developed which combine the information collected in the subsystems and interpret the signatures in terms of electrons, photons and hadrons such as pions or neutrons. This challenge is inordinately complicated by typically around 20 additional hard interactions taking place in the same bunch crossing, leading to a proliferation of signals in the various components, which is referred to as in-time pile-up. Additional pp collisions from neighboring bunch crossings (before or after) can add to this effect, since their secondaries can still or already occupy the read-out channels and consequently be misidentified as belonging to the bunch crossing of interest. This is known as out-of-time pile-up. In particular the identification of charged particle tracks is a delicate task, as

around 1000 of these are expected to occur in one bunch crossing.

The reconstruction of hits, tracks and vertices will be explained below, followed by a description of the procedure to identify a muon, electron, photon or hadron candidate. Finally, algorithms clustering jets from the hadrons are an important element for analyses searching for processes with charged particles in the final state and will be described together with the concept of missing transverse energy. The CMS collaboration employs a particle flow (PF) algorithm, on whose description in [133–135], together with the details of track and vertex reconstruction in [136], this following section is largely based.

3.3.1. Hit reconstruction

Charged particles create signals in the read-out units of the silicon pixels and strips of the tracker. How a hit is reconstructed in the pixel detector will be briefly explained in the following; for a description of hit and cluster reconstruction in the strip detector see [122]. A zero-suppression algorithm [137] poses minimal requirements on the strength of the signal read out from a pixel so that it is considered in the reconstruction algorithms at all, and thereby eliminates noise effects and fake hits. This threshold corresponds to an equivalent charge of 3200 electrons⁷, which is well below the 21000 electron charges a minimally ionizing particle would create. A cluster is formed from such a pixel by considering also all of its neighboring cells; in total the cluster must have a charge equivalent to at least 4000 electrons.

A first algorithm, the so-called first-hit pass reconstruction, measures a hit position by projecting the cluster, separately on the u and v directions by summing the charges of pixels with the same u (v) coordinate. Here u is the direction transverse and v longitudinal to the sensors. The position is given by the center of the projected cluster if the cluster is only one pixel wide. If it is wider, the position is proportional to the term $(Q_{\text{last}} - Q_{\text{first}})/2(Q_{\text{last}} + Q_{\text{first}})$, where Q_{first} (Q_{last}) is the charge collected in the first (last) pixel of the cluster. The position is also dependent on the relative orientation of the track (if known) to the plane of sensors. From the boundaries of the charge distribution and its mean (which is corrected for the Lorentz drift of the electrons in the B field of $\sim 60 \mu\text{m}$), the hit position can be reconstructed to a first approximation.

The full shape of the charge projection is utilized by the more sophisticated template-based hit reconstruction, which compares the shape to charge distributions of a detailed simulation of a such pixel cluster. This reconstruction obtains the hit position by fitting to various templates, where also radiation damage effects are taken into account. The resulting set of positions and their uncertainties is updated with information about detector geometry and known misalignment of certain silicon modules.

3.3.2. Track reconstruction

A dedicated procedure is applied for the reconstruction of tracks, referred to as the combinatorial track finder (CFT). It is based on the Kalman filter algorithm [138] and

⁷In fact, from the electron-hole pairs created by the charged traversing particle, it is the holes that are collected in the read-out units.

3. Experimental setup

is repeated six times. It first reconstructs the easiest, i.e. prompt tracks originating from (very close to) the interaction point and with $p_T > 0.8 \text{ GeV}$, and then successively addresses more complicated possible tracks: tracks with only two pixel hits, with a low p_T or tracks outside the pp interaction region.⁸ Hits matched to a track from previous iterations are removed, reducing the complexity of the combinatorial problem. Each iteration can be divided into four steps:

1. Two or three hits according to the first-pass algorithm constitute the initial track candidate and serve as seed to the subsequent parts of the algorithm. Only 3D hits and an optional constraint that the track starts near the interaction point are the input for the reconstruction of the trajectory. The hits must be found in the so-called seeding layers, which are pairs or triplets of detector layers. Additionally the track candidate must fulfill certain requirements such as a minimum p_T or a maximal distance with respect to the assumed point of interaction.
2. Tracks accepted as a seed are input to a Kalman filter. First it performs a coarse estimation of the trajectory parameters from the seed and consecutively adds hits from successive detector layers, where the parameter uncertainties determine which adjacent detector layers the track can be extrapolated to, taking into account bremsstrahlung and multi-scattering effects. From all the hits in the silicon detector modules of the next layer that are compatible with the trajectory exactly one hit is added. This can also be an artificially added ghost hit to account for the possibility the particle has not created a signal in this layer. Since the parameters and their uncertainties are updated each time a new hit subjoins the track, the accuracy of the track is steadily improved. This results in multiple track candidates for a given set of hits. The candidates least compatible with the hits are discarded. Additionally, tracks may not have too many ghost hits or a too low p_T ; these are also not considered any further.
3. The track fitting – using a Kalman filter and smoother algorithm – increases the accuracy of the track and removes possible biases which may arise e.g. due to the additional beam spot constraint in case only two 3D hits were used in step 1. Starting from the innermost hit, the three next template-based hits of a track are fitted, and track parameters are derived thereof. Subsequently all of the remaining hits along the trajectory are added in the outward direction, each time updating the track and its parameters. Finally these are handed over to a smoother which runs in backward direction. The smoothing effect comes by averaging the track parameters received from going the two directions (in \rightarrow out, out \rightarrow in), since in the first case and for a given position along the path, the fit is performed using the information of all the hits found before, and the other time the fitter takes into

⁸The procedure starts from the inner part of the tracker and builds the tracks outwards, mostly because the channel occupancy is lower in the highly granular pixel part and because of the fact electrons lose energy on their way via bremsstrahlung and the algorithms profit from good quality 3D hits to start with. This leads to a higher reconstruction efficiency.

account only the hits beyond that position. This third step employs more advanced algorithms for solving the equations numerically and to extrapolate from one hit to another. Also the inhomogeneity of the magnetic field is accounted for. Spurious hits are removed a posteriori by calculating their compatibility with the track by means of a χ^2 test with thresholds varying from first and later iterations of the CTF. If such outliers are found and removed, the track is refitted until all hits are compatible with the trajectory.

4. The last step does not feature quality requirements on the hits but on the tracks themselves, as a significant amount of fake tracks not belonging to a charged particle are still expected among the list of candidates. These requirements again depend on the iteration and ask for e.g. a minimum number of layers with hits, a maximum number of layers without hits and a good fit quality (expressed with a χ^2 measure).

After the last iteration, the different collections are combined and duplicate tracks are removed.

3.3.3. Vertex reconstruction

Reconstructing the positions of the inelastic proton-proton interactions is essential for disentangling the tracks, for an increased precision in the track fitting and for assigning them to either the interaction of interest or to pile-up.

The reconstruction consists of three steps: first, candidates for prompt tracks are identified by applying certain selection criteria similar to the requirements in the last step of the CTF. Then the tracks are grouped according to the z coordinate of their closest approach to the beam spot. Here an ambiguity arises, since there are arbitrarily many ways the clustering can be performed: the algorithm must be able to resolve pile-up and distinguish it from the “signal” vertex, but a too high granularity would split up tracks that belong to a common origin. A deterministic annealing algorithm tries to create a reasonable balance by solving a multidimensional minimization problem. A detailed description is given in [139].

The adaptive vertex fitting algorithm [140] finds the x , y and z positions of the interaction point for each cluster of tracks. It iteratively minimizes a loss function which depends on the distances of the tracks to the current estimate of the position, on the number of degrees of freedom for the vertex, $n_{\text{d.o.f.}}$, and on weights assigned to each track. After the minimum is found tracks which are compatible with the vertex have a weight close to 1, while outliers have $0 < w < 1/2$. The quantity $n_{\text{d.o.f.}}$ can be identified with the number of compatible tracks and will later serve as a quality requirement ($n_{\text{d.o.f.}} > 4$) in the search for tHq.

3.3.4. Muon candidates

Three types of reconstructed muon candidates are typically used in the CMS collaboration, which are briefly described in the following. A comprehensive discussion can be found

3. Experimental setup

in [141].

The so-called tracker muon is obtained in a straightforward way: the CTF tracks are extrapolated into the muon chambers, taking into account the most accurate propagation model including multiple scattering and the inhomogeneous magnetic field, and are tried to be matched to hits in muon chamber segments, if present. If at least one such hit is compatible with the trajectory, a tracker muon is reconstructed. This method is particularly suited for low- p_T muons (< 5 GeV), because there is no stringent requirement on the activity in the muon system.

The standalone muon relies on information from only the muon detectors [142]. First, track segments are built by clustering individual hits. The reconstruction happens via algorithms which are very similar to the CTF: starting from the innermost layer the track is extrapolated to the next muon station and updated with compatible track segments. After the outermost layer is reached, the Kalman filter is applied in the other direction, and the trajectory is extrapolated to the interaction point, requiring a good-enough compatibility with the primary vertex.⁹ The position of the interaction point is also a constraint in the final fit of the track parameters.

This analysis employs global muons: after selecting a subset of all the CTF tracks in the silicon detector which is loosely compatible with a standalone muon detected in the muon system, the CTF tracks and the muon track are extrapolated to a common surface and their hits are refitted using a Kalman filter in order to obtain the track parameters. Some of the resulting global muon candidates are iteratively removed because they are not compatible enough; the final track is chosen based on the best χ^2 value in the global fit. If the transverse momentum is smaller than 200 GeV, the p_T of the muon as determined from the tracker alone is chosen as final numerical value.

Once the muon candidate is identified, energy deposits in the ECAL and HCAL in a cone around the muon track are summed, and the candidate is discarded if they exceed a certain threshold. This reflects the chance of detecting a punch-through hadron in the muon system and misidentifying it as a muon, and of tagging muons from B hadron decays, all of which would be characterized by additional activity in the calorimeters.

3.3.5. Electron candidates

A detailed description of the state-of-the-art reconstruction of electrons is provided in [143]. It is outlined here in the most relevant aspects. The main challenge in the reconstruction owes to the fact that electrons significantly interact with the tracker material and can lose a substantial part of their energy already in the silicon layers via bremsstrahlung. The emitted photons spread over a large detector volume, which makes an efficient reconstruction of the initial electron a delicate task. Two approaches are used, specially optimized and complementing each other for different ranges of the electron p_T .

Since electrons are expected to leave a signature in the ECAL, the first method starts

⁹The primary vertex typically is the one of the vertex collection with the largest scalar p_T sum of the tracks emerging from it.

with grouping the crystals around a seed crystal (the one with the largest recorded transverse energy) to so-called superclusters (SC). The clusters do not have the same size for the ECAL barrel and endcap, because the geometry of the crystal arrangement as well as the magnetic field strengths are different and consequently the spread of charge is expected to vary in the two regions. If an SC in the endcap or the seed crystal in the barrel has an energy $E > 1 \text{ GeV}$, the SC serves as seed in the search for tracks in the pixel detector, which first need to be reconstructed. The CTF algorithm is suboptimal for this purpose: the Kalman filter (KF) describes fluctuations with a Gaussian shape, but since the process of bremsstrahlung is non-Gaussian, the KF would either not find the correct hits or be not accurate enough. For this reason KF tracks with few hits or large χ^2 are refitted using a Gaussian sum filter (GSF) [144,145], in which several Gaussian distributions, each provided with an adjustable weight, are mixed in order to better model the energy loss. The above procedure works well for electrons with p_T larger than 10 GeV. For electrons with smaller p_T , a multivariate analysis helps in the electron identification: starting from seeds in the tracker, the track is extrapolated to energy deposits clustered in the ECAL using both a KF and a GSF. The χ^2 values of both tracks as well as information about the geometrical and energetical compatibility of tracker and ECAL information are used to construct a multivariate discriminator, which helps in identifying low- p_T electrons.

3.3.6. Photon and hadron candidates

After the tracks and hits for electron and muon candidates have been removed from the collections, the remaining tracks are tried to be assigned to photon and hadron candidates by matching them to energy depositions in the ECAL and HCAL. Depending on the relative sizes of the energies as estimated by the tracker or calorimeter information alone, there are two possibilities: (a) the energy of the particle candidate in the tracker is comparable (or slightly larger) than the energy in the calorimeters. In this case the track parameters are obtained fitting both tracker and calorimeter information.¹⁰ If the calorimeter energy is much higher than the tracker momentum, it must be an electrically neutral particle which hardly interacts with the tracker material but showers in the ECAL/HCAL. It is labelled a photon if the total energy deposit in the ECAL is larger than the difference between the calorimeter and tracker estimates. Otherwise there must also be energy deposits in the HCAL, and an additional neutral hadron is tagged. Energy deposits which could not be matched to tracks are not discarded but are added to the photon/neutral hadron collection.

3.3.7. High-level objects

The term high-level object refers to a structure which can be derived from the individual candidates explained above. This applies in particular to the description of collimated

¹⁰If the energy in the tracker is much larger than in the calorimeters, only the momentum measured in the tracker is used and the mass is taken as the pion mass.

3. Experimental setup

bunches of hadrons, arising in the showering and hadronization process of an initial colored particle due to confinement. The hadrons can be clustered to so-called jets with dedicated algorithms which are described in the following.

Jets

Any jet algorithm must respect an essential principle of jet clustering: it must be infrared-collinear-safe (IRC). This means the set of final state jets must be stable against the presence of soft and collinear radiation.

In particular the four-vector, i.e. the direction and energy of a jet must not depend on extra soft emissions inside of its cone. Accordingly the requirement of infrared safety reflects the lack of knowledge of long-distance properties of QCD, where perturbation theory cannot be applied anymore. The collinear safety requires the number of jets to be not dependent on collinear splittings of the particles inside a jet, i.e. it must be robust against the case where one particle emits collinear radiation. If the IRC criterion is not respected, the jet algorithm will lead to objects that cannot be easily compared to theoretical predictions which rely on the cancellation of infrared divergences appearing in the perturbative expansion of QCD.

The anti- k_t algorithm [146] is IRC-safe and is the standard way of clustering jets in CMS. It is also easy to use for theoretical and phenomenological studies and has manageable computational costs. Together with the k_t [147, 148] and Cambridge/Aachen [149, 150] algorithms it belongs to the family of sequential recombination jet algorithms. They are implemented in the FASTJET package [151, 152] and feature the calculation of the metric $d_{i,j}$, which is defined for two objects (particles or pseudo-jets) i and j in the following way:

$$d_{i,j} = \min(p_{T,i}^{2n}, p_{T,j}^{2n}) \frac{\Delta_{i,j}^2}{R^2},$$

where $p_{T,i}$ is the transverse momentum of the i^{th} object and $\Delta_{i,j}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$. R is a resolution parameter, which is set to 0.5 in the search for tHq at 8 TeV and to 0.4 for the 13 TeV analysis. In addition, $d_{i,\text{beam}}$ denotes the distance between object i and the beam and is given by

$$d_{i,\text{beam}} \equiv p_{T,i}^{2n}.$$

In the above equations, the power n can have different values depending on the algorithm: $n = 1$ yields the k_t algorithm, $n = 0$ recovers the Cambridge/Aachen way of clustering, and finally $n = -1$ corresponds to the case of anti- k_t . The clustering however always happens in the following way: in each iteration the smallest distance found among all ($d_{i,j}$) and ($d_{i,\text{beam}}$) is identified. If the two objects i and j have the smallest metric, they are merged to a pseudo-jet and the next iteration starts with recalculated distances; if $d_{i,\text{beam}}$ is the smallest one, the object i is labelled a jet and is removed from the object list before the procedure is repeated until there are no objects left.

3.3. Reconstruction of events

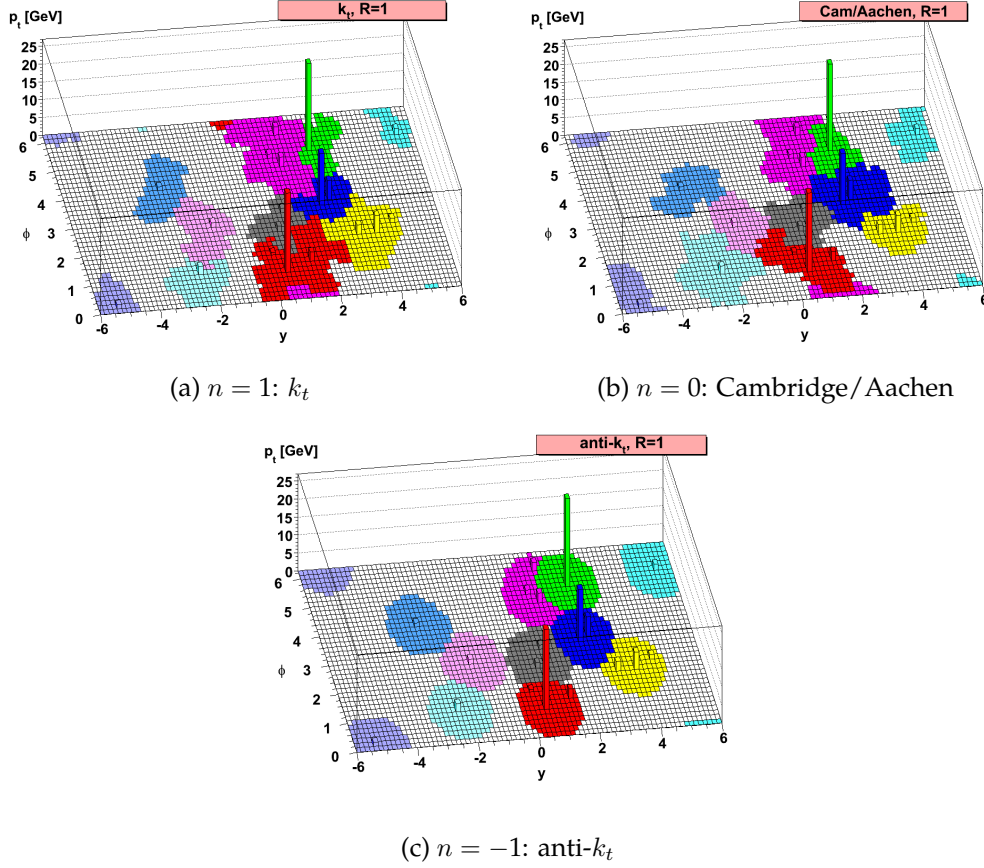


Figure 3.12.: Illustration of jet areas after a clustering with (a) the k_t , (b) the Cambridge/Aachen and (c) the anti- k_t algorithm, for a resolution parameter of $R = 1$. Taken from [146].

As can be seen from Fig. 3.12, this results in almost conical jets for the anti- k_t algorithm, a behaviour which can be understood in the following way: for $n = -1$, the distance between a hard particle h and a soft particle s_1 is given by

$$d_{h,s_1} = \min(1/p_{T,h}^2, 1/p_{T,s_1}^2) \frac{\Delta_{h,s_1}^2}{R^2} = \frac{\Delta_{h,s_1}^2}{p_{T,h}^2 R^2} < \min(1/p_{T,s_1}^2, 1/p_{T,s_2}^2) \frac{\Delta_{s_1,s_2}^2}{R^2} = d_{s_1,s_2},$$

which is smaller than the measure d_{s_1,s_2} between two soft particles if $\Delta_{s_1,s_2} \simeq \Delta_{h,s_1}$. As a consequence, soft particles tend to be primarily clustered around a hard one, and if this hard particle does not have comparably hard neighbors within $d = 2R$, the clustered jet will be conical. If there is another hard particle in the vicinity of object h , the resulting jets will be trimmed according to algorithms explained in [146].

The clustering can be performed with any kind of input objects. In the case of reconstructed jets, these are the PF particles. On generator level they can be stable Monte-Carlo truth particles; the outcome is then referred to as generator jet.

3. Experimental setup

Calibration of the jet energy

Various processes related to hardware and pile-up effects dilute the true energy and thus distort the measurement of jet energies, leading to intrinsically biased results if no corrections are applied. CMS uses factorized jet energy corrections with three levels [153], which increase the precision of jet energy measurements:

Level 1: The jet energy is increased by detector noise and pile-up. This energy offset is accounted for by subtracting the median of the distribution describing the energy density per unit area.

Level 2: The calorimeter performance is not uniform but depends on the pseudorapidity, as some detector regions lack instrumentation and outer regions have a worse response than in the central region. For this reason, relative correction factors as a function of η are applied to flatten the jet responses.

Level 3: The transverse momentum of a jet also influences the response of the calorimeters. Absolute corrections make their performance uniform in p_T .

Reconstructed jets in data undergo an additional correction referred to as L2L3Residual. This correction mitigates chances of a possible residual bias because the corrections of L2 and L3 have been derived only from a simulated QCD sample. As will be mentioned in Sec. 4.2.4, a dedicated jet smearing is used on top of the corrections to eliminate differences in the observed energy resolutions of data and simulation.

Identification of b jets

The concept of b tagging is essential to a vast range of analyses involving final state b quarks, as their presence leads to certain jet characteristics which can be exploited in the discrimination against processes without genuine heavy quarks, ultimately leading to better background suppression and sensitivity. This makes it particularly relevant for the search for tHq presented in Chap. 4, where three central b quarks from the decay of both the Higgs boson and the top quark are expected. The following logic applies in the identification of jets originating from a b quark: upon hadronization, b quarks produced in the hard interaction form a bound state – a B hadron – whose decay is CKM-suppressed because of $V_{cb}, V_{ub} \ll 1$. As a consequence, the hadron has a long lifetime and propagates in the detector for up to several millimeters (L_{xy} in Fig. 3.13) until it decays. This results in several tracks displaced by an impact parameter d_0 with respect to the primary vertex, which originate from a vertex inside the jet. The presence of such a secondary vertex (SV), its geometrical parameters and the p_T of the tracks emerging from it are distinct characteristics of a true b jet and can be used in an MVA to create a likelihood for a reconstructed jet to stem from a b quark.

The combined secondary vertex (CSV) algorithm [155] is the b tagging algorithm of choice in this thesis. It combines information on the interaction point, the SV, track multiplicities and jet kinematics and transforms them to a response value for each jet

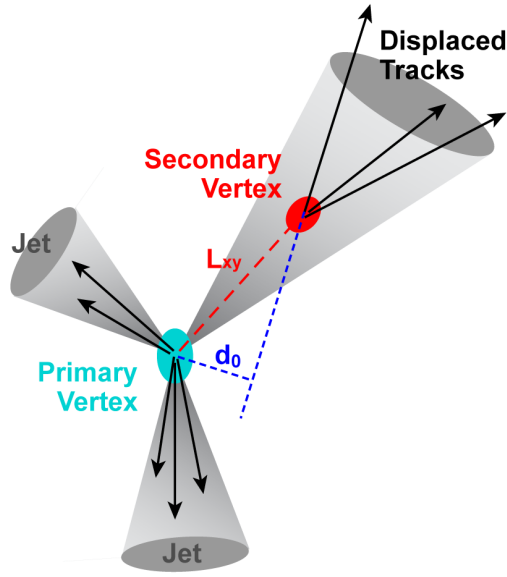


Figure 3.13.: Illustration of the concept of b tagging [154].

which can be interpreted as a probability it is a b jet. The performance is typically judged by the calculation of b tagging efficiencies and mistagging rates (the fraction of light jets wrongly identified as b jet), by means of which also working points, i.e. specific CSV discriminator cut values, are defined after measuring the (in)efficiencies in simulation. The analysis in Chap. 4 relies on the tight working point ($CSV > 0.898$) corresponding to the threshold value which one in a thousand true light jets passes.

Missing transverse energy

Some particle types, like SM neutrinos and certain supersymmetric or dark matter particles, are expected to not interact with the detector material at all; they leave the experiment undetected. However, their kinematics in the transverse plane can be inferred indirectly from all the observed momenta and energies which should not balance to 0 if such a particle has been produced. Because of momentum conservation and the initial transverse momentum of ~ 0 , the missing transverse energy vector can be defined as

$$\vec{E}_T = - \sum_{i=1}^N \vec{p}_{T,i} + \Delta \vec{E}_{JEC}.$$

Here $i = 1 \dots N$ denotes the N particle candidates the PF algorithm has reconstructed, and $\Delta \vec{E}_{JEC}$ is a term owing to the additional jet energy corrections described previously, which of course have to be propagated to the missing transverse energy. In the search for tHq, the absolute value $E_T = |\vec{E}_T|$ is identified with the transverse momentum of the neutrino from the leptonic top quark decay.

4. Search for associated production of single top quark and Higgs boson at 8 TeV

In this chapter the search for a Higgs boson produced in association with a single top quark is presented, considering Higgs boson decays into a pair of b quarks. Collision data at a center-of-mass energy of 8 TeV are analyzed and an upper limit on the possible scenario in which one of the important properties of the SM Higgs boson – the Yukawa coupling to the top quark – has a flipped sign with respect to its predicted value is provided. The motivation for conducting this search was already given in Chapter 1. This chapter describes the search in detail, providing a comprehensive overview from the characteristic signature of a tHq event in the detector, to the definition of the objects, the event selection, reconstruction and classification and finally to the uncertainties, the results and their interpretation. Special emphasis is put on the aspects that are most relevant for the analysis. These turn out to be the use of stacked Neural Networks and the modelling of the dominant background processes.

As such, the chapter is a recapitulation of the first search for single top quark + Higgs boson associated production in the $H \rightarrow b\bar{b}$ channel that has been performed within the CMS collaboration resulting in a public Physics Analysis Summary (PAS) [156], with strong contributions from the author of this thesis at all levels of the analysis. For the purpose of this thesis, the analysis is repeated and illustrated in every respect in the following. The PAS has become part of a combination paper with other Higgs boson decay channels, which has been submitted to JHEP in fall 2015 [48].

4.1. Channel topology and search strategy

An appropriate search strategy for tHq with a flipped top-Yukawa coupling, i.e. with $C_t = y_t/y_t^{\text{SM}} = -1$, must account for characteristic features in the signal and background topology as well as for the overall rates that are expected. The Feynman diagram for tHq at LO in the 4F is shown in Fig. 4.1. The most distinct feature of a single top t -channel-like process is the upper quark line, where a light quark recoils against the exchanged W boson and results in a typically very forward light jet with a substantial p_T . While pure single top analyses can be entirely based on cuts and information on this light jet (see e.g. the CMS t -channel cross section measurement at 8 TeV [157]), the search for tHq utilizes the light jet as an important object helping to identify the signal process. The analysis considers Higgs boson decays into a pair of b quarks, as this decay channel is the one with the largest branching ratio and thus gives the highest rates for a process with a tiny production cross section. The other resonance in the process, the top quark,

4. Search for associated production of single top quark and Higgs boson at 8 TeV

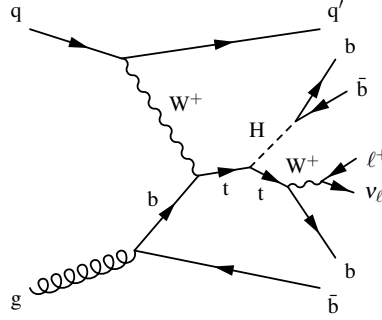


Figure 4.1.: LO tHq Feynman diagram, including the Higgs boson decay to a b quark pair and the leptonic decay of the top quark.

is required to decay leptonically. The presence of a prompt lepton (only electrons and muons are considered as final state objects; but leptonically decaying τ leptons can enter the selection) is needed to reject multi-jet background processes, for which it is difficult to produce something that can imitate a hard, well isolated lepton. Moreover, a hadronically decaying top quark would further increase the jet multiplicity, leading to an even more complex final state. The top quark decay product besides the leptonically decaying W boson is a b quark which gives rise to a b jet. Like the b jets expected from the Higgs boson decay, it appears rather central in the detector due to the large mass of its mother particle. There is an additional b quark that comes from the initial gluon splitting. It has a broad η spectrum and in general has small transverse momentum. Most of the time it therefore cannot be tagged as a b jet by the central detector elements. However a special category will be introduced to tag events in which the additional b jet could be identified. Further gluons can be radiated off colored objects in the initial or final state. To account for this possibility, the event selection employs only a lower bound on the expected number of jets.

Semi-leptonic $t\bar{t}$ production is by far the dominating background in the search for $tHq, H \rightarrow b\bar{b}$. The simulation of $t\bar{t}$ production in association with additional jets – only top quark pair processes with extra emissions can enter the selection of Sec. 4.3 – is categorized according to the flavor of the extra emissions and the number of identified heavy-flavor jets: excluding decay products stemming from an electroweak vertex (like $t \rightarrow Wb$ or $W \rightarrow cs$) all b or c quarks found after parton showering are matched to reconstructed jets within a distance $\Delta R < 0.5$. If there is a jet matched to at least one b quark, it is assigned b flavor (similarly for c quarks/jets). Events with at least two b jets fall into the $t\bar{t} + b\bar{b}$ category; if there is only one b jet found, it is labelled a $t\bar{t} + b$ event. Finally $t\bar{t} + c\bar{c}$ is the category for events with at least one jet with c quark origin. If an event has neither of the heavy flavor jets, it is a $t\bar{t} + \text{light flavor (lf)}$ jets event. This categorization follows what has been done in [158] in the $t\bar{t}H$ searches with $H \rightarrow b\bar{b}$. The rates of $t\bar{t} + \text{heavy flavor}$ processes, also in relation to the $t\bar{t} + \text{lf}$ component, are not very well known in theory and play a dedicated role in the systematic uncertainties that need to be considered in the analysis. The splitting basically allows to investigate the relative

4.1. Channel topology and search strategy

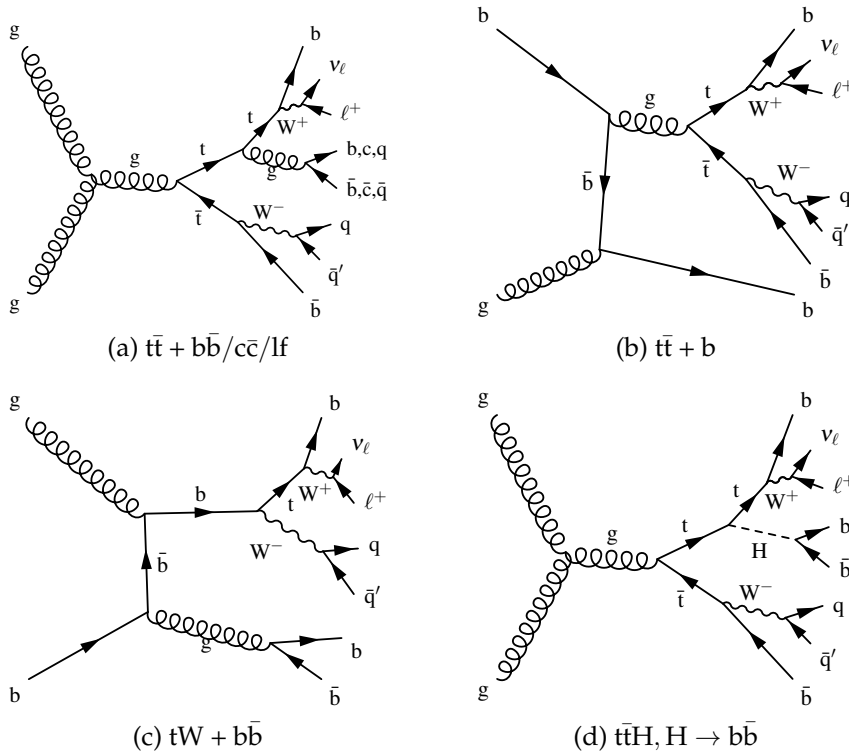


Figure 4.2.: Feynman diagrams of the most relevant background processes.

$t\bar{t} + \text{jets}$ components, as opposed to collapsing them into one category and treating the different processes fully correlated. More information on the modelling itself is provided in Sec. 4.3, the detailed discussion on the systematic uncertainties is deferred to Sec. 4.7. Feynman diagrams for the most important background processes, which all involve top quarks and extra b quarks from other origins, are depicted in Fig. 4.2. For single top production, the semi-leptonic tW production mode, despite having a smaller inclusive cross section than the t -channel process, is the more important one because of its high jet multiplicity and its resemblance with $t\bar{t}$, leading to a higher acceptance given the selection requirements.

The analysis strategy identified as providing an optimal search sensitivity is illustrated with a flow chart in Fig. 4.3. It will be explained in reverse flow direction in the following.

In order to have a discriminating distribution from which the signal (or the signal strength) can be extracted and an upper exclusion limit on the tHq process can be set, the best possible separation between signal and background must be achieved. This is attempted by using a Neural Network that is taught to distinguish between the tHq process and the backgrounds and finally classifies the observed events as being signal- or background-like by means of a continuous variable, the classification NN output. Variables that serve as input to the NN belong to three different classes: the first type of variable exploits global properties of the events. The second input variables class

4. Search for associated production of single top quark and Higgs boson at 8 TeV

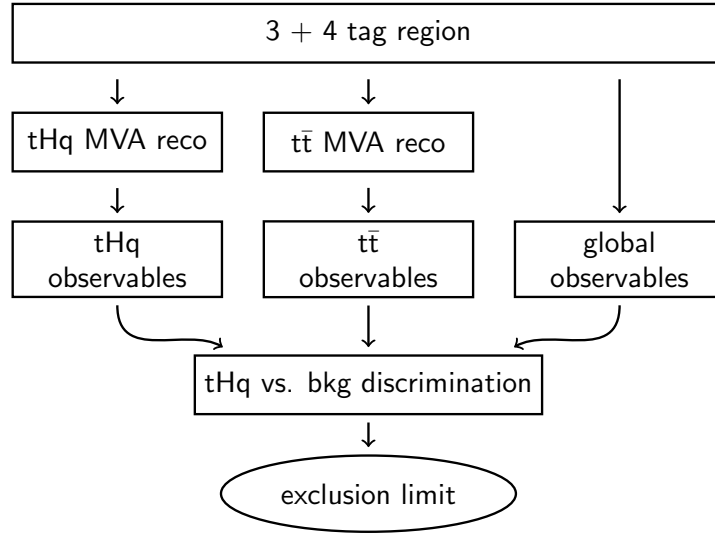


Figure 4.3.: Analysis scheme for the search for tHq with $H \rightarrow b\bar{b}$, adopted from [159].

consists of tHq observables and depends on e.g. the kinematics of the reconstructed Higgs boson or the distinct light jet. As a third class, $t\bar{t}$ observables are constructed by interpreting the signatures in the detector under the assumption the event stemmed from semi-leptonic $t\bar{t}$ production. The latter two classes however necessitate a reconstruction of the events in order to identify the various jets with (decay products of) the various objects like the Higgs boson, the hadronically decaying top quark, etc. At the cost of an increased complexity, dedicated MVA based tHq and $t\bar{t}$ reconstructions will lead to a larger flexibility in choosing the input variables for the classification NN and will improve the performance of the analysis. All of the steps outlined above are carried out in a signal-enriched region mainly defined by the requirement on the presence of either three or four b jets. A third region, which is not mentioned in Fig. 4.3, is enriched in events stemming from $t\bar{t}$ production and serves as a control region to validate shapes of important distributions. The event selection is explained in the following, after a definition of the objects that selection requirements are imposed on is given.

4.2. Object definitions

4.2.1. Primary vertices

A vertex candidate must not lie further away in the z -direction than 24 cm from the nominal detector center and furthermore must have a radial distance less than 2 cm to the beam axis. In order to reduce the impact of misidentified vertices, at least four tracks

need to be assigned to the vertex.¹

The figure-of-merit for selecting the so called primary vertex among all the reconstructed possible vertices is the metric D which is constructed as the sum of the squared p_T of all n tracks associated with the vertex candidate:

$$D = \sum_i^n p_{T,i}^2,$$

where i denotes the i^{th} track. The vertex candidate with the largest D value is chosen as primary point where the hard interaction took place. Pile-up effects are alleviated by removing all charged hadrons that are not assigned to the chosen vertex.

4.2.2. Muons

A muon PF candidate in an event that has passed the trigger HLT_IsoMu24_eta2p1 must be reconstructed as a global muon, this means tracks in both the outer muon systems and the tracker must be associated with the candidate. Six tracker layers must have identified a hit that can be assigned to the track of a muon candidate. The χ^2 for fitting the tracks needs to fulfill $\chi^2/\text{n.d.f.} < 10$. A requirement on the fit is that it uses the information of at least one hit in the muon chambers.

The tracks in the tracker are constrained by the requirement $|d_{xy}| < 2$ mm, where d_{xy} is the transverse impact parameter measured with respect to the primary vertex candidate. In addition the muon track is bound to have a z -position not further away than 5 mm from the z -position of the primary vertex.

The analysis distinguishes between tight and loose muons, based on the p_T , η and the relative $\Delta\beta$ -corrected isolation, which is defined as

$$I_{\Delta\beta} = \frac{1}{p_T} \left(\text{chargedHadronIso} + \max \left(\text{neutralHadronIso} + \text{photonIso} + \beta \cdot \text{puChargedHadronIso}, 0 \right) \right).$$

As can be seen in the above equation, the sum of transverse momenta of a certain class of reconstructed particle candidates enter the definition of $I_{\Delta\beta}$: of charged- or neutral-hadron candidates, of photon candidates or as charged hadrons identified as pile-up. All the candidates only contribute inside a cone of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.4$ around the muon. A proportionality factor of -0.5 is chosen for β , which represents a correction for the energy deposited in the isolation cone by charged particles not associated with the primary vertex [160]. The quantity $I_{\Delta\beta}$ is a relative isolation and as such does not bear a unit, as it is divided by the muon p_T .

For tight muons the requirements are: $p_T > 26$ GeV, $|\eta| < 2.1$ and $I_{\Delta\beta} < 0.12$. Loose muons are selected with less stringent cuts, particularly on the isolation: $p_T > 20$ GeV, $|\eta| < 2.5$, and $I_{\Delta\beta} < 0.2$.

¹Typically, they reflect the four degrees of freedom that need to be assigned in the vertex fit of the reconstruction algorithm.

4. Search for associated production of single top quark and Higgs boson at 8 TeV

The CMS collaboration provides muon correction factors to account for differences in observed and simulated acceptances [161]. They are applied throughout the analysis.

4.2.3. Electrons

For electron candidates, a relative isolation similar to the one of muons is defined:

$$I_\rho = \frac{1}{p_T} \left(\text{chargedHadronIso} + \max(\text{neutralHadronIso} + \text{photonIso} - \rho A_{\text{eff}}, 0) \right).$$

A smaller cone size of 0.3 around the candidate is considered, owing to the fact that electrons are more likely to be faked by QCD-like processes. In the equation, the parameter ρ is the p_T density observed in the entire event topology, averaged over $\phi = [-\pi, \pi]$. A_{eff} is a correction factor accounting for the pile-up arising from neutral hadrons. Its values, which are measured in bins of η , can be found in [162]. The electron momentum to which the isolation is normalized is obtained via the GSF algorithm as explained in Sec. 3.3.5. The relative isolation must fulfill $I_\rho < 0.1$ for a tight and $I_\rho < 0.15$ for a loose electron.

The p_T threshold for both cases is 30 and 25 GeV, respectively, while the requirement on the pseudorapidity is $|\eta| < 2.5$ both for loose and tight electrons. The latter furthermore may not stem from a photon conversion and must have an MVA ID response of larger than 0.9 [163].

The region between electromagnetic calorimeter barrel and endcap ($1.4442 < |\eta_{\text{sc}}| < 1.5660$, with “sc” specifies the supercluster) is excluded to prevent biasing the results with the problematic simulation of electron isolation in this gap [164]. The HLT path for electron candidates is E1e27_WP80.

4.2.4. Jets

For reasons that will be explained later, this analysis employs a p_T threshold on jets that is different for the central and the forward detector regions. Central jets ($|\eta| < 2.4$) are required to at least have a moderate p_T of larger than 20 GeV. Jets in the forward region need to be harder, $p_T > 40$ GeV. They are clustered with the anti- k_t algorithm explained in Sec. 3.3.7 and a cone size of 0.5. Prior to running the clustering algorithm, certain particle candidates are removed from the input collection: muons and electrons with isolations smaller than 0.2 (0.15) do not enter the clustering as well as particles identified with charged hadrons stemming from pile-up interactions.

The loose PF jet ID imposes minimal requirements on a reconstructed jet; it is preferred over the tight ID in order to not lose too many signal events with their multi-jet final state; the exact definitions are listed in [165]. On top of the standard jet energy corrections explained in 3.3.7, L1FastJet corrections are applied to correct for potential additional clustered energy arising from pile-up.

In 7 TeV analyses, the actual energy resolution observed in data was worse than the one observed in simulated events [166]. For this reason, the last modification of the jet energy consists of updated correction factors [167] applied to the reconstructed jets in simulated events in order to better match the resolution observed in recorded collisions.

4.2.5. Missing transverse energy

Missing transverse energy has been introduced in Sec. 3.3.7 as the negative sum of the momenta of all PF candidates; it is corrected for an x - y modulation observed in the reconstructed \cancel{E}_T .² Its amplitude is correlated with the number $N_{V,PU}$ of vertices associated with pile-up interactions in the event. Possible causes are of geometrical kind like a misalignment of the detector or a displacement of the beamspot, and/or are technical; inactive calorimeter segments could lead to a calorimeter performance that is non-isotropic in ϕ . More details can be found in [168] and in Appendix A.1.1.

4.2.6. W boson reconstruction

From the lepton and the transverse missing energy the W boson can be reconstructed. This implies that the neutrino, which cannot be detected, is identified with the missing energy and balances the observed momenta. Although this imbalance can only be measured in the transverse (x - y) detector plane, $p_{z,\nu}$ can be inferred from certain assumptions of the following rationale:

For a massless neutrino ($m_\nu = 0$ GeV), the identity $E = |\mathbf{p}|$ holds. With $m_\ell = 0$ GeV and the presumption that the W boson has been produced on-shell, the invariant W boson mass is given by:

$$m_W^2 = \left(E_\ell + \sqrt{\cancel{E}_T^2 + p_{z,\nu}^2} \right)^2 - (\mathbf{p}_{T,\ell} + \cancel{\mathbf{E}}_T)^2 - (p_{z,\ell} + p_{z,\nu})^2 = (80.4 \text{ GeV})^2.$$

This is a quadratic equation in $p_{z,\nu}$ and has the two solutions:

$$p_{z,\nu}^{(1,2)} = \frac{\zeta p_{z,\ell}}{p_{T,\ell}^2} \pm \sqrt{\underbrace{\frac{\zeta^2 p_{z,\ell}^2}{p_{T,\ell}^4} - \frac{E_\ell^2 \cancel{E}_T^2 - \zeta^2}{p_{T,\ell}^2}}_{:=\kappa}},$$

with

$$\zeta = \frac{\cancel{E}_T^2}{2} + \mathbf{p}_{T,\ell} \cdot \cancel{\mathbf{E}}_T = \frac{\cancel{E}_T^2}{2} + p_{T,\ell} \cancel{E}_T \cos \Delta\phi,$$

where $\Delta\phi$ denotes the ϕ difference between the direction of the lepton and the missing transverse energy vector. In the case where the measured transverse mass of the W boson is larger than 80.4 GeV, the discriminator κ is negative, which leads to two complex solutions $p_{z,\nu}^{(1,2)} = \frac{\zeta p_{z,\ell}}{p_{T,\ell}^2} \pm i\sqrt{|\kappa|}$. The x and y components of the neutrino momentum are then slightly modified (keeping the constraint $m_W = 80.4$ GeV and not changing \cancel{E}_T) until the square-root term vanishes³ and only one real solution $\frac{\zeta p_{z,\ell}}{p_{T,\ell}^2}$ remains. This

²Due of the rotational symmetry of the detector design, the \cancel{E}_T distribution is expected to be independent of ϕ .

³Since the discriminator is quadratic in $p_{x,\nu}$ and $p_{y,\nu}$, again two solutions exist. The one with the minimal ϕ distance between $\cancel{\mathbf{E}}_T$ and $\mathbf{p}_{T,\nu}$ is picked.

4. Search for associated production of single top quark and Higgs boson at 8 TeV

Table 4.1.: Requirements for an event to fall into either the 3 tag or 4 tag region.

	3 tag region	4 tag region
# tight leptons		1
# add. loose leptons		0
# central jets with $p_T > 30$ GeV or forward jets with $p_T > 40$ GeV		≥ 4
\cancel{E}_T		$> 35/45$ GeV(μ/e)
# jets	≥ 4	≥ 5
# jets with CSV > 0.898	$= 3$	$= 4$

procedure is justified because finite \cancel{E}_T resolution is assumed to be responsible for such a behaviour, and shifting the transverse momentum components can be seen as any of the other energy corrections in this case. If the discriminator is larger than 0, i.e. if there are two possible real solutions for $p_{z,\nu}$, the smaller absolute value is picked.

4.3. Selection of events

Table 4.1 lists the selection cuts defining the signal regions used in the analysis. They account for the expected signal topology and are minimal requirements on the reconstructed events in order to reject the bulk of the background while trying to retain as many signal events as possible. The selection efficiency in the 3 tag region is 3.7% for tHq but only 0.2% for semi-leptonic $t\bar{t}$ + jets production; the applied selection thus leads to an enhancement of the ratio tHq/ $t\bar{t}$ by a factor ~ 22 relative to the inclusive rates. For the 4 tag region, the efficiencies are 0.4% for tHq and 0.005% for $t\bar{t}$.

While in the 3 tag region the $t\bar{t}$ + light jets component dominates, where one of the light jets has been misidentified as stemming from a b quark, the $t\bar{t}$ + $b\bar{b}$ processes constitute most of the expected events in the 4 tag region. The contributions of tHq to the total expected rates is 1.5% for the 3 tag region. The 4 tag region is purer in signal events (5.2%); however, as will be shown later, the 3 tag region has a larger weight in the search. It drives the exclusion limit, while the 4 tag region only has moderate impact on the search sensitivity, because it is limited in statistical power. The exact event yields, determined in a binned maximum-likelihood fit to the NN output distributions, are given in Table 4.6.

In addition, a 2 tag region is employed that serves the purpose of a control region. It differs from the 3 tag region only in the jet requirements ($\#$ jets with CSV $> 0.898 = 2$), which means it is close yet orthogonal to the signal-enriched regions; all central analysis concepts and techniques like simulation corrections and multivariate analysis tools are validated here. With a purity of 94% it is largely dominated by the $t\bar{t}$ + jets background.

4.3.1. Consequences for signal and background modelling

Given the selection cuts, the analysis operates in a high-jet-multiplicity phase space, imposing several requirements particularly on central b jets. This has implications on the modelling of signal and background processes, since one has to make deliberate choices to make sure the simulation models the data well enough. In order to get the most correct predictions, the simulation needs to describe b jet kinematics and large jet multiplicities adequately. This is attempted by employing the 4F for modelling the tHq process: the dedicated 4 tag region tries to capture the additional b jet arising from the initial gluon splitting genuine to the single top t -channel process. For b initiated processes this is expected to model observables like the differential p_T and η of the additional b quark better than the 5F [57], where the b quark is typically softer as it is stemming from the backwards evolution of the parton shower (see also Section 2.1.4). The tHq process in this analysis is modelled at LO in QCD. In the 4F, the NLO shape corrections of the observables however are known to be small, and the differences are covered by scale uncertainties. Therefore a leading-order signal sample is deemed to give accurate enough predictions, especially for the acceptances of (additional) b jets. The renormalization and factorization scales are set to a fixed value of $\mu_R = \mu_F = 100$ GeV. The reasoning for choosing this value is that it should approximate the momentum transfer at the only vertex of the strong interaction in the tHq diagram, which is assumed to lie in the middle between the two involved energy scales – the masses of the top and b quark.

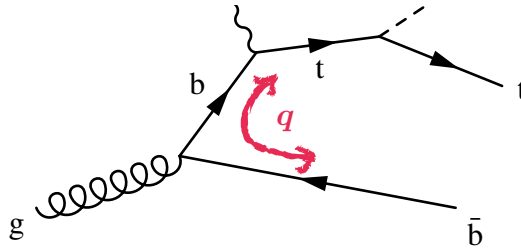


Figure 4.4.: Scale choices for t -channel-like processes in the 4F are driven by the kinematics of the additional b quark and the exchanged momentum q .

For the $t\bar{t}$ background, also a LO modelling is chosen; it is improved by merging several additional-jet-multiplicities on matrix element level into one common sample, thus increasing the accuracy of the predictions in the selected phase space as opposed to relying on the parton shower for the description of hard extra emissions. This requires a procedure to avoid double counting of real emissions, since the same final state configuration can result from either describing n jets at matrix element level where the parton shower only adds collinear radiation which the resolution of the detector and the reconstruction algorithms cannot resolve as extra jets, or from an $n - 1$ parton configuration with an extra hard emission added by the parton shower. The MLM merging technique [169, 170] is used to consistently describe $t\bar{t} + 0, 1, 2$ or 3 jets at matrix element level in this manner. The $V + \text{jets}$ background is modelled in the same way, but plays a minor role because of the large required b jet multiplicity. A similar matching procedure has been tested for the case of single top (+ Higgs) production in a bachelor thesis that has been supervised in the context of this analysis [171]. The t -channel process is delicate because of the additional light flavor jet in the matrix element. First results on the single

4. Search for associated production of single top quark and Higgs boson at 8 TeV

top + jets modelling are promising, agreeing with inclusive 4F NLO predictions within scale uncertainties; however it is crucial to have a jet matching algorithm implemented in Pythia8 which excludes jets stemming from electroweak vertices. While versions available at the time of the dedicated studies were capable of doing so, recent releases used for sample production in CMS have rolled back in this feature, which makes the production of such samples not an option for the moment. The predominant single top contribution anyway stems from tW production; it is simulated at NLO in the 5F, removing diagrams overlapping with the LO definition of $t\bar{t}$. Diboson and purely QCD-like processes are modelled at LO. The above choices guarantee accurate enough predictions, owing to the importance of the various processes; for some of them however, predictions of the inclusive cross sections at higher accuracy than the ones given by the MC generators are available. In such cases the templates have been normalized to these values in order to obtain more accurately predicted rates. The associated production of a single top quark and a Z boson which decays to a pair of b quarks could imitate the signal process almost perfectly, differing only in the peak location of the invariant $b\bar{b}$ spectrum. The tZq process has been investigated and its expected rates, which have been calculated in [172], turn out to be smaller than the signal by a factor 4.⁴ This background is therefore not considered in the following. In Tab. A.2 more details on the simulation of the various background samples and the signal sample are provided.

The recorded and simulated datasets used in this analysis are listed in Tab. A.2. The analyzed pp collisions correspond to an integrated luminosity of 19.7 fb^{-1} .

4.3.2. Controlling the multi-jet background

Considering the expected final state signature with a large jet multiplicity, concerns about possible leakage of multi-jet events into the signal region are valid. Simulations of the QCD background however suggest that the contribution is negligible. To be on the safe side, a viable data-driven cross check for estimating this background is the so called ABCD method: in addition to the signal regions given above, which are labelled A, three regions enriched in multi-jet events are defined. Region B is the signal region with inverted isolation criteria on the lepton; region C inverts the requirement on E_T ; finally, both inversions are applied in region D, which is maximally enriched in QCD-like events. Under the simple assumption

$$N_A = \frac{N_B N_C}{N_D},$$

which is verified in the simulation, the expected yields in the signal region can be estimated. They turn out to be $< 1\%$ and are ignored henceforth.

⁴The tZq process has an inclusive cross section similar to tHq , but $\mathcal{B}(Z \rightarrow b\bar{b})/\mathcal{B}(H \rightarrow b\bar{b}) \sim 1/4$.

4.4. Corrections to simulation and data

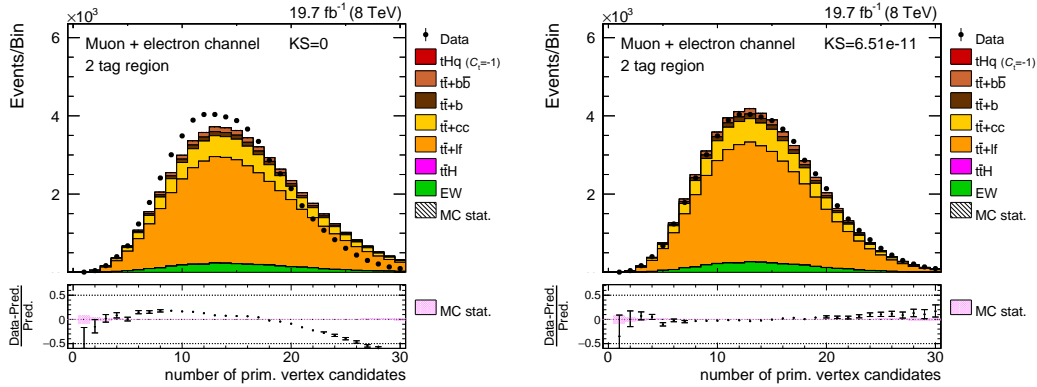


Figure 4.5.: The number of primary vertices before (left) and after (right) the reweighting of all simulations to the true distribution found in data. Simulation is renormalized to the observed data yields.

4.4. Corrections to simulation and data

After applying the selection cuts listed in the previous section, the agreement of the simulations with data is not yet satisfactory. Differences arise, as will be illustrated below, due to insufficient theoretical knowledge of the hard processes, the complexity of a bunch crossing with several points of hard interactions, or also due to imperfect detector conditions. The CMS collaboration has put a lot of work into identifying and systematically mitigating these effects. The ones relevant for this analysis are explained in the following.

4.4.1. Pile-up reweighting

The correct simulation of in-time and out-of-time pile-up effects is very challenging. Prior to starting an extended simulation campaign, a certain PU profile needs to be assumed, which is then corrected for after dedicated analyses could perform a measurement of the true number of pp interactions per bunch crossing. Measured is the luminosity-block dependent instantaneous luminosity, and the data PU profile is derived from it. More details are given in [173]. Fig. 4.5 shows the multiplicity distribution of primary vertex candidates before and after applying the correction. This observable is the most sensitive to the PU effects. Other variables like jet p_T or information on the reconstructed lepton are only mildly influenced and do not show such a drastic improvement.

4.4.2. Top p_T reweighting

Experimental results [174, 175] indicate that the p_T spectrum of top quarks is much harder in currently available simulations of the $t\bar{t}$ process (be it LO or NLO) than the one observed in data. This also affects the top quark decay products like the W boson, as is visible especially from the tail of the left distribution in Fig. 4.6. At the time this was

4. Search for associated production of single top quark and Higgs boson at 8 TeV

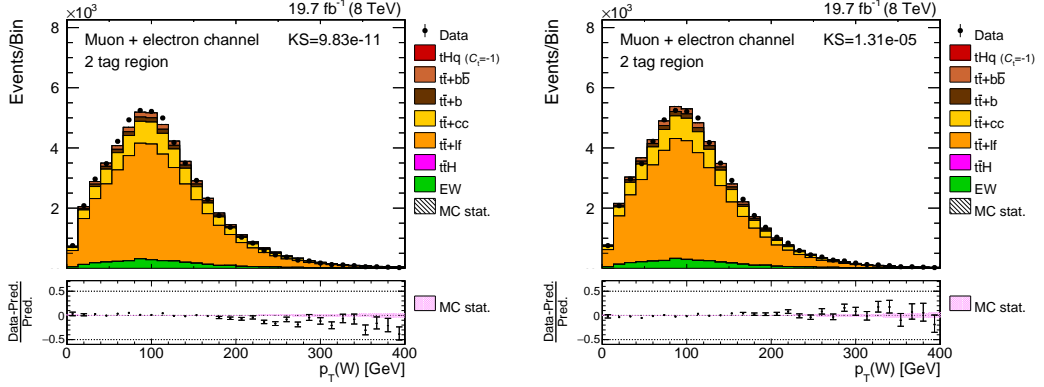


Figure 4.6.: The transverse momentum of the reconstructed W boson before (left) and after (right) the top p_T reweighting to the $t\bar{t}$ simulations, which mainly improves the data/MC agreement in the tail of the distribution. Simulation is renormalized to the observed data yields.

first observed by both ATLAS and CMS collaborations, this clearly systematic behaviour was surprising and subject to lots of discussions, but could not be attributed to a specific cause. CMS has therefore derived correction factors to fix the agreement in observables related to the top quark p_T [176].

The event weight $w_{\text{top } p_T}$ is given by the equation

$$w_{\text{top } p_T} = \sqrt{\exp(A - B \cdot p_T(\text{top})) \cdot \exp(A - B \cdot p_T(\text{antitop}))}$$

where A and B are sample-dependent (various MADGRAPH simulations for $t\bar{t}$ + jets have been used throughout the analysis) and are given in Appendix A.1.1; they have been derived in the original differential p_T measurements.

In the meanwhile, NNLO calculations of the top quark p_T spectrum have been published [177], agreeing much better with the measurements. This suggests that the likely cause of the slope has been traced down, and that higher order corrections are successively softening the transverse momentum distributions of top quarks (see also Fig. 4.7).

4.4.3. Correction of b tagging efficiencies

B tagging efficiencies are found to be different for data and simulation, for which the CMS collaboration provides correction scale factors derived on orthogonal samples [178]. With these scale factors, one can compute event weights to be applied to simulated events which lead to a better prediction of the yields observed in data.

The probability for a simulated event – with a given jet configuration determined by the p_T , η and flavor of the jets – to fall into the signal region is given by

$$P(\text{MC}) = \prod_{i=\text{tagged}} \varepsilon_i \prod_{j=\text{not tagged}} (1 - \varepsilon_j),$$

4.4. Corrections to simulation and data

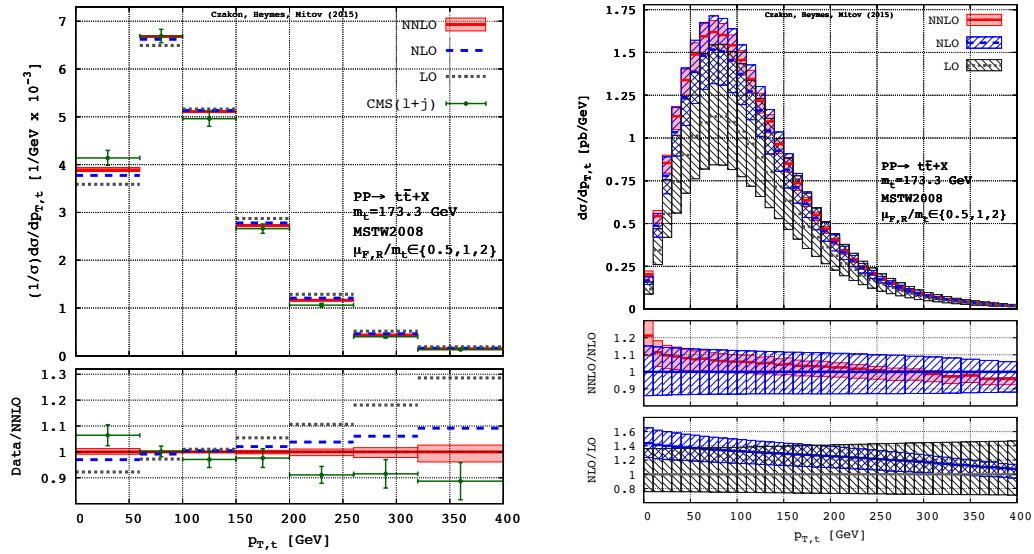


Figure 4.7.: Different predictions (LO, NLO, NNLO) for the top quark p_T distribution in $t\bar{t}$ events [177], compared to CMS data (left) and shown including their shape and normalization uncertainties (right) arising due to higher order corrections.

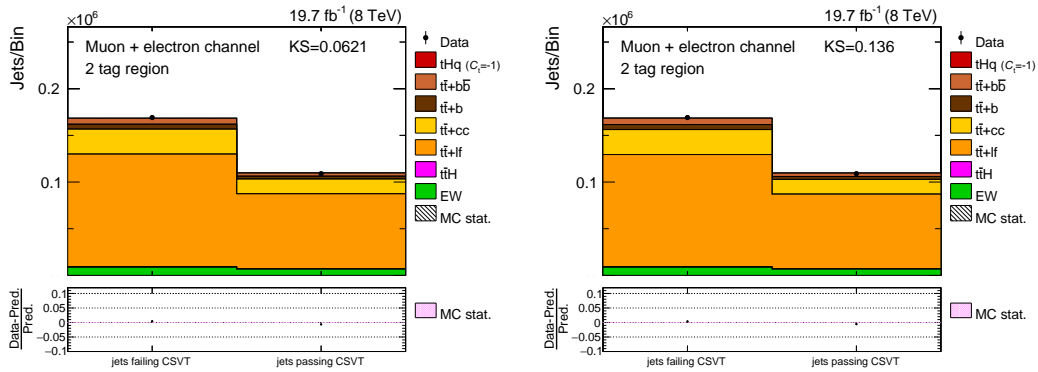


Figure 4.8.: Differences in b tagging efficiencies between data and simulation, visible in the number of jets that fail or pass the tight CSV working point (left) are corrected by assigning dedicated b tagging scale factors (right). Simulation is renormalized to the observed data yields.

4. Search for associated production of single top quark and Higgs boson at 8 TeV

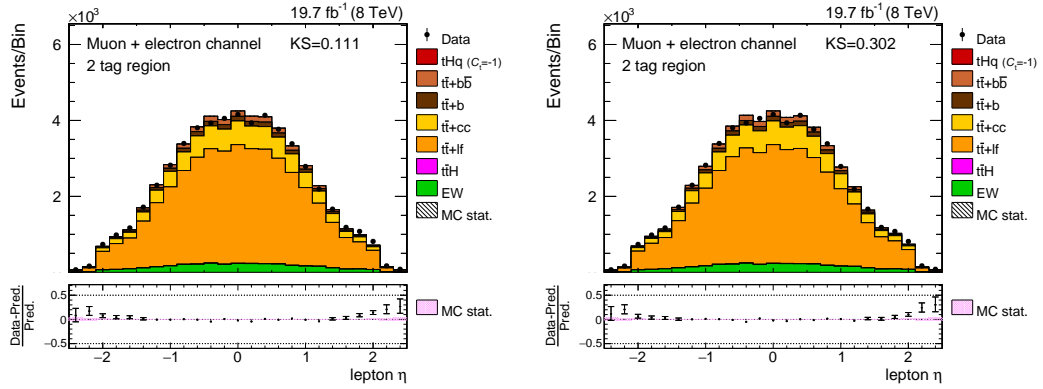


Figure 4.9.: Lepton pseudorapidity distributions in the 2 tag region before (left) and after (right) applying the lepton scale factors. The visible effect on the shape in marginal, and the KS value is shown for quantifying the improvement on the description. The steep drop at $|\eta| = 2.1$ is due to the fact that muons are considered only up to this value, while electrons are reconstructed within $|\eta| < 2.5$.

where the efficiencies ε have been determined independently for each process by applying the full selection of Tab. 4.1 and with and without the b tag requirements. Together with the provided correction factors SF a corresponding probability for data can be calculated:

$$P(\text{data}) = \prod_{i=\text{tagged}} \text{SF}_i \varepsilon_i \prod_{j=\text{not tagged}} (1 - \text{SF}_j \varepsilon_j).$$

Finally each event is assigned a weight equal to $w_{\text{b tag}} = P(\text{data})/P(\text{MC})$. The effect is illustrated in Fig. 4.8. More information can be found in [179].

4.4.4. Lepton identification

The identification efficiencies of muons and electrons and the rates of such leptons passing the selection and trigger requirements are different for data and simulation. Correction factors depending on the lepton p_T and η have been calculated by [180] for muons in a $Z \rightarrow \ell^+ \ell^-$ enriched region. For electrons the scale factors are taken from [181]. The correction factors translate into an event-specific weight that is applied to all simulated events. Fig. 4.9 is showing the pseudorapidity distributions of the lepton in the 2 tag region before and after the correction. A slightly better data/MC agreement is observed.

4.4.5. Jet η treatment

In the course of this analysis, a severe mismodelling in the η distribution of jets has been observed across most regions of phase space, i.e. in the 3/4 tag regions, the 2 tag region and V + jets enriched control regions. This suggests it is not a process-dependent mismodelling of the simulation, but rather a genuine feature of detector mismodelling

or the reconstruction. The mismodelling is particularly pronounced for low- p_T jets, as the upper distribution in Fig. 4.10 is showing for $p_T \in [20, 25)$ GeV. Besides a central “u”-like shape, the ratio data/MC displays a dip around $|\eta| \simeq 2.7$ (green shaded area) and a steeply falling slope for $|\eta| \geq 3.1$ (blue shaded area). For the presented analysis it is absolutely crucial to have an equivalent description of the pseudorapidities in both data and simulation as it heavily relies on the features of forward jets.

Therefore a detailed investigation of the possible causes for this disagreement between data and simulation has been performed. As it turns out, the problem is multifaceted and cannot be attributed to one sole source. The dip around just beyond the tracker edge in fact is part of a larger region between $2.4 < |\eta| < 3.2$. A possible bias in the η can be seen in Fig. A.1. Jets with a generator level $|\eta_{\text{gen}}|$ of ~ 2.5 and ~ 3.1 tend to be reconstructed with a value $|\eta_{\text{reco}}| \sim 2.7$. To correct for this migration in data and simulation has been considered too much effort; it was instead decided to raise the p_T threshold of the jets until the data/MC differences are covered by systematic uncertainties. This is the case for $p_T > 40$ GeV as is shown in the lower left distribution in Fig. 4.10, where variations due to uncertainties on the renormalization and factorization scales are able to cover the residual discrepancies around $|\eta| = 2.7$. The slope towards larger $|\eta|$ values from ~ 3.1 on however is still outside the assigned uncertainties and requires further attention. The likely cause for this behaviour is the binning which is employed for deriving jet energy corrections. Because of insufficient event counts in this region, they have been derived in a single bin for $|\eta| \geq 3.14$. This binning, in conjunction with collapsing the jets between $2.4 \leq |\eta| < 3.2$ into a single bin, is adopted in the analysis. Practically the pseudorapidity undergoes the transformation

$$\eta \rightarrow \eta' = \begin{cases} \eta & : |\eta| < 2.4, \\ 2.7 \cdot \text{sign}(\eta) & : 2.4 \leq |\eta| < 3.2, \\ 3.5 \cdot \text{sign}(\eta) & : |\eta| \geq 3.2. \end{cases}$$

It should be stressed this transformation, which happens prior to calculating any observable, changes the meaning of the pseudorapidity, but this is no problem in terms of interpretability: for an input variable in a Neural Network a good data/MC agreement of the shapes is most important; the variable does not have to have a physical interpretation as long as it is well modelled and provides discriminating power. In [40] the procedure is justified by arguing the same quantity would be observed with a hypothetical detector which has no granularity in $2.4 < |\eta| < 3.2$ and $|\eta| > 3.14$. After this transformation is applied, almost perfect data/MC agreement is observed (see right lower distribution in Fig. 4.10), while all remaining differences are covered by the uncertainties.

4.5. Event reconstruction

The complexity of the final state asks for a dedicated event reconstruction in order to identify the reconstructed jets with the outgoing quarks of the hard interaction in the most efficient way. A reconstruction Neural Network is used to decide which assignment

4. Search for associated production of single top quark and Higgs boson at 8 TeV

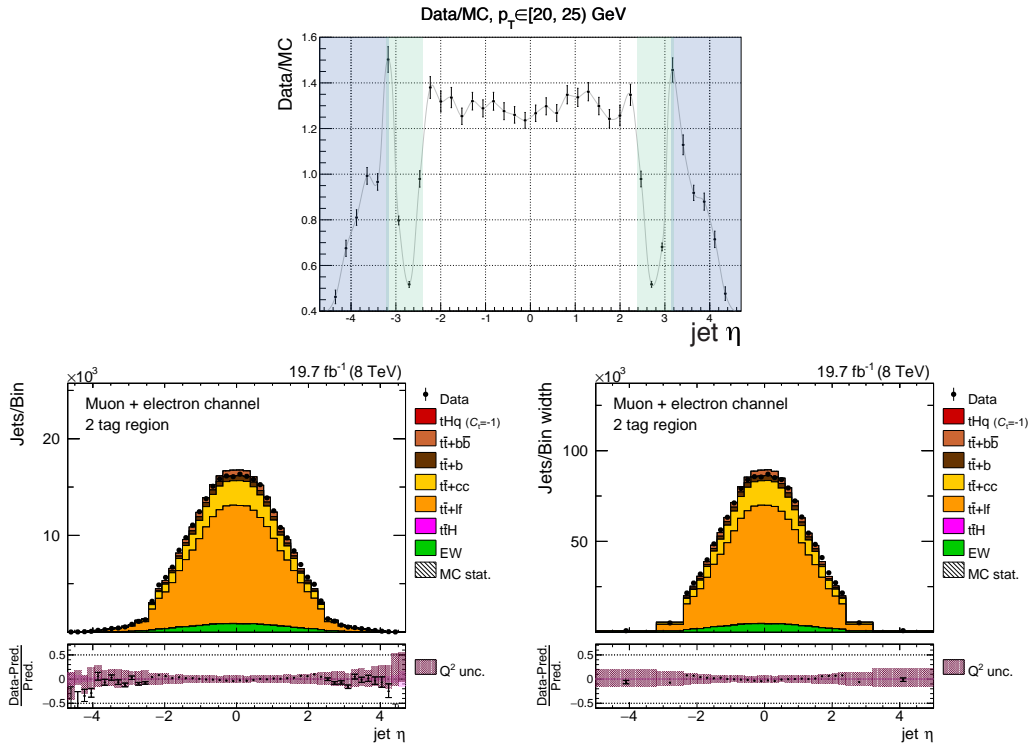


Figure 4.10.: Upper row: jet η data/MC ratio in the 2 tag control region for jets with $p_T \in [20, 25]$ GeV. The two problematic regions $|\eta| > 3.14$ and $2.4 < |\eta| < 3.2$ are shaded in blue and green, respectively. Lower row: jet pseudorapidity distribution after raising the p_T threshold to 40 GeV for forward jets (left); in the right distribution also the modified binning is applied that mitigates effects due to lacking fine grained jet energy corrections. While for the upper distribution MC is normalized to prediction, the lower row shows simulations renormalized to the observed data yields.

is the correct one among many different possible quark-jet-assignments which will be explained below.

As a cross check a comparison with a reconstruction based on minimizing a χ^2 criterion was performed in [40, 100], proving the reconstruction with a Neural Network performs significantly better in terms of final discrimination power between tHq and the backgrounds.

4.5.1. Reconstruction of events under the tHq hypothesis

First a correct quark-jet-assignment needs to be defined on a per-event basis. This is possible for the simulation of the signal, since there all the parton level information is accessible. An assignment will be referred to as “correct” if each of the four quarks (the two b quarks from the decay of the Higgs boson, the b quark from the top quark decay and the light quark) can be matched to a reconstructed jet within a distance $\Delta R < 0.3$. Other

jet assignments are called “wrong” in the following. In order to reduce the combinatorics, the number of allowed jet assignments is reduced by the following requirements: because of b tagging considerations, only central jets ($|\eta| < 2.4$) can be matched to the b quarks from the Higgs boson or top quark; only a jet which is not b tagged can be assigned to the light quark. After these limitations to the possible event interpretations, an event still has a few dozens of allowed assignments.

A Neural Network is trained to combine the separation power of jet variables which have different shapes for correct and wrong assignments into one discriminator. Per event, the correct and one randomly picked wrong hypothesis are used as training input. Events where the quarks could not be uniquely assigned to define the correct interpretation are not considered. A list of the input variables is given in Tab. 4.2. It features variables exploiting characteristics of the Higgs boson decay products, the b quark jet from the top quark decay, the light jet, as well as correlations between them. A total of 12 variables has been considered. Shapes of the first four of them are shown in Fig. 4.11 for correct and wrong assignments, while the reader is deferred to Appendix A, Fig. A.2 for the distributions of the other variables. It should be noted that a validation of the input variable shapes against data is not possible due to their definition. This would require a data sample pure in tHq events. Hence the validation of the method only happens a posteriori: statistically independent evaluation and testing templates agree in their shapes of the reconstruction NN output, as is shown in the upper figure of Fig. 4.12. There is no visible sign of overtraining, and the method is applied to all events passing the event selection. For an unknown event, all possible jet assignments are evaluated, and the interpretation that yields the largest output value for the discriminator is picked. This can also be validated with data, and the bottom row of Fig. 4.12 shows good agreement for the largest discriminator value in both muon and electron channels. The corresponding distributions for the 3 tag and 4 tag region can be found in Fig. A.3. The last piece of the validation is to check the NN output for a randomly chosen interpretation. Once more very good agreement is observed (see Fig. A.4).

4.5.2. Reconstruction of events under the $t\bar{t}$ hypothesis

The procedure is repeated under the assumption the event stems from $t\bar{t}$ production with semi-leptonic top quark decays. The quarks considered for the matching are the b quarks from the two top quark decays and the light quarks from the hadronic decay of one W boson. Combinatorics are reduced by requiring that the b quarks must be matched to b tagged jets. The Neural Network is trained with the 13 variables listed in Tab. 4.3. The most relevant variables are related to the kinematics of the reconstructed hadronically decaying W boson; information on the jet charges and on the leptonically decaying top quark also add sensitivity to the NN. Distributions of the first five variables are shown in Fig. 4.13, the remaining variables are part of the Appendix in Fig. A.5.

The validation of the procedure is verified in the same way as for the reconstruction under the tHq hypothesis: by looking at the NN output for an independent test sample, the presence of overtraining can be ruled out, as is shown in the upper plot in Fig. 4.14). The best NN output value in the 2 tag region shows good agreement in the bottom distri-

4. Search for associated production of single top quark and Higgs boson at 8 TeV

Table 4.2.: Input variables of the tHq reconstruction NN, ranked according to the importance in the TMVA training.

Rank	Variable	Explanation
1	tagged jet (b_t)	Boolean information on whether the jet assigned to the b quark from the top quark decay is b tagged.
2	$ \eta$ (light jet)	Absolute pseudorapidity of the jet identified with the light recoil quark.
3	# b tags of Higgs jets	Number of tagged jets found among the jets assigned to the Higgs boson.
4	$\log m(b_t + \ell)$	Invariant mass of the lepton and the jet assigned to the b quark from the top decay.
5	$\log m(H)$	Reconstructed Higgs boson mass.
6	$\log \min(p_T(\text{Higgs jets}))$	Minimum jet p_T found among the jets associated to the b quarks from the Higgs decay.
7	$\Delta R(\text{Higgs jets})$	Distance in the η - ϕ plane between the two jets assigned to the Higgs boson.
8	$\max \eta(\text{Higgs jets})$	Maximum pseudorapidity found among the two jets assigned to the Higgs boson.
9	$\Delta R(b_t, W)$	Distance in the η - ϕ plane between the reconstructed top quark decay objects.
10	relative H_T	$(p_T(H) + p_T(t) + p_T(\text{light jet}))/H_T$, where H_T is the sum of all transverse momenta (lepton, jets) and \cancel{E}_T .
11	$\Delta R(H, t)$	Distance in the η - ϕ plane between the reconstructed top quark and Higgs boson.
12	$q(b_t) \cdot q(\ell)$	Product of lepton charge and charge of jet assigned to the b quark from the top quark decay.

butions of the same figure. The corresponding 3/4 tag region distributions are provided in Fig. A.6. Finally, NN output distributions for randomly chosen event interpretations are shown in Fig. A.7.

For illustration purposes the complex procedure of defining a correct jet assignment is depicted in Fig. 4.15 for both hypotheses, providing also efficiencies for picking the correct quark-jet-assignments in the caption. The exact settings used in the TMVA training are given in Appendix A.2.2.

Table 4.3.: Input variables of the $t\bar{t}$ reconstruction NN, ranked according to the importance in the TMVA training.

Rank	Variable	Explanation
1	$\log m (W_{\text{had}})$	Mass of the reconstructed hadronically decaying W boson.
2	$\log(m (t_{\text{had}}) - m (W_{\text{had}}))$	Mass difference between the hadronically decaying top quark and W boson.
3	$\Delta R (W_{\text{had}} \text{ jets})$	Distance in the η - ϕ plane between the two jets assigned to the W boson of the hadronically decaying top quark.
4	$ \eta (t_{\text{had}}) $	Absolute pseudorapidity of the reconstructed hadronically decaying top quark.
5	$\log p_{\text{T}} (t_{\text{had}})$	Transverse momentum of the reconstructed hadronically decaying top quark.
6	# b tags of W_{had} jets	Number of tagged jets found among the jets assigned to the hadronically decaying W boson.
7	$\log p_{\text{T}} (t_{\text{lep}})$	Transverse momentum of the reconstructed leptonically decaying top quark.
8	$\Delta R (b_{t_{\text{had}}}, W_{\text{had}})$	Distance in the η - ϕ plane between the reconstructed hadronically decaying top quark and W boson.
9	relative H_{T}	$(p_{\text{T}} (t_{\text{lep}}) + p_{\text{T}} (t_{\text{had}}))/H_{\text{T}}$, where H_{T} is the sum of all transverse momenta (lepton, jets) and \cancel{E}_{T} .
10	$\Delta R (b_{t_{\text{lep}}}, W_{\text{lep}})$	Distance in the η - ϕ plane between the reconstructed leptonically decaying top quark and W boson.
11	$\log m (b_{t_{\text{lep}}} + \ell)$	Invariant mass of the reconstructed jet associated with the b quark of the leptonically decaying top quark and the lepton
12	$(q (b_{t_{\text{lep}}}) - q (b_{t_{\text{had}}})) \cdot q (\ell)$	Difference in charge of the two jets assigned to the b quarks from the semileptonic $t\bar{t}$ event, times the lepton charge.
13	$\Sigma q (W_{\text{had}} \text{ jets}) \cdot q (\ell)$	Sum of the charge of the jets assigned to the hadronically decaying W boson, times the lepton charge.

4. Search for associated production of single top quark and Higgs boson at 8 TeV

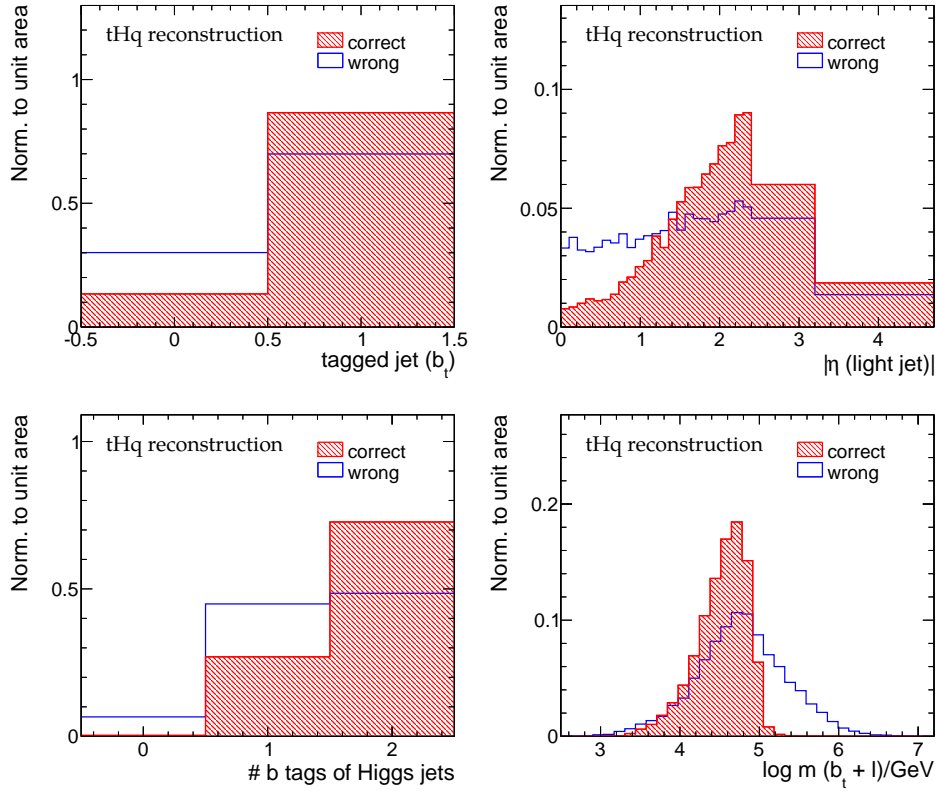


Figure 4.11.: The four most important input variables for the tHq reconstruction. A clear separation power between correct and wrong assignments is visible in all cases.

4.6. Event classification

The selection of Tab. 4.1 still leaves an overwhelmingly large $t\bar{t}$ background. In order to become more sensitive to the signal, the analysis employs another NN which is taught to classify an event as being rather signal- or background-like. Semi-leptonic and dileptonic $t\bar{t}$ events as well as $t\bar{t}H$ events contribute to the background templates fed to the NN, where each process is scaled according to its cross section. Other background processes have not been trained against because of the low number of available simulated events in the 3 tag region. The latter has been chosen as training region, and training results are applied to the 4 tag region as well as to the 2 tag region for validation purposes. As the final exclusion limits will show, the limit is mainly driven by the 3 tag region, which has a good balance between statistical power and purity in tHq.

4.6.1. Choice and optimization of NN input variables

After carrying out the two previously explained reconstructions in both the simulated and observed events passing the selection cuts, the objects are defined and the final set

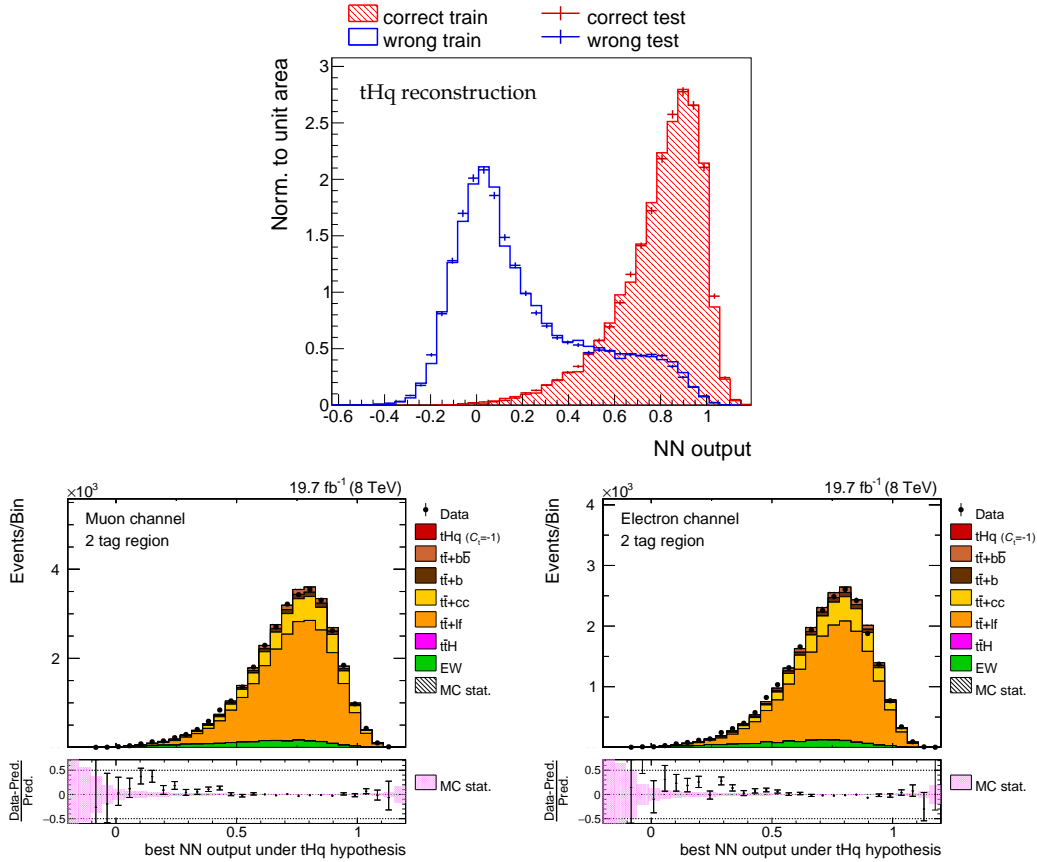


Figure 4.12.: Upper row: training results for the reconstruction NN under the tHq hypothesis; a very good correct-vs.-wrong separation can be seen. Lower row: highest BDT output of all jet assignments per event in the 2 tag region for the muon channel (left) and the electron channel (right). Simulation is normalized to data.

of input variables used in the classification NN can be constructed. It has undergone a thorough optimization and aims at addressing critical aspects in the topology of a tHq event. By means of variables resulting from the event reconstruction under the assumption of semileptonic $t\bar{t}$ events, information about intrinsic properties of the main background can also be used. The final list of input variables can be seen in Table 4.4. Alongside kinematic variables such as the p_T of the reconstructed Higgs boson or the invariant mass of the objects identified with the hadronically decaying top quark (a variable transformation by taking the logarithm of the invariant mass is used to smoothen the input distributions), information about b tagging is employed to distinguish jets containing a B hadron from light flavor jets. The most important variable is the lepton charge, i.e. a global observable that is independent from reconstruction of the event under a particular hypothesis. While this variable is symmetric for $t\bar{t}$, it is correlated to the ratio $R = \sigma_t/\sigma_{\bar{t}}$, which numerically is roughly 2 for single top t -channel-like processes.

4. Search for associated production of single top quark and Higgs boson at 8 TeV

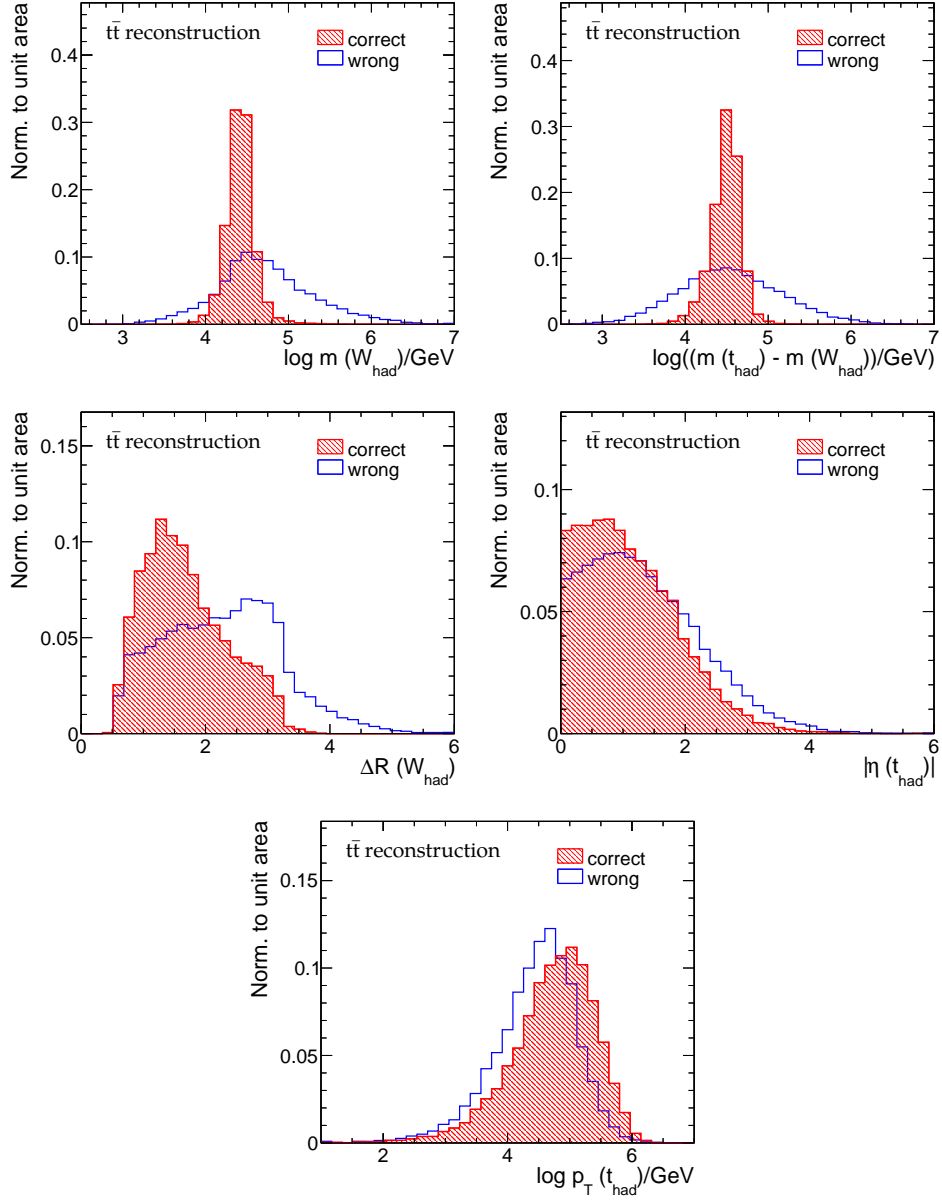


Figure 4.13.: The five most important input variables for the $t\bar{t}$ reconstruction. A clear separation between correct and wrong assignments is visible in all cases.

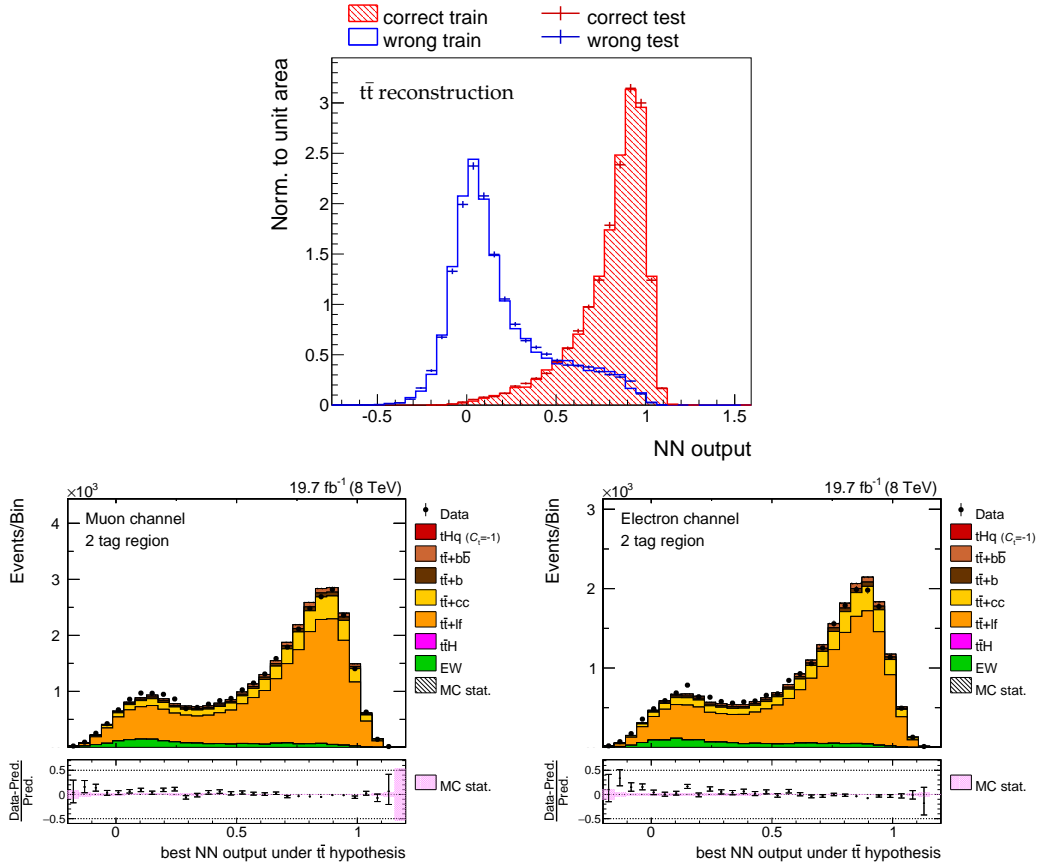


Figure 4.14.: Upper row: training results for the reconstruction NN under the $t\bar{t}$ hypothesis; a very good correct-vs.-wrong separation can be seen. Lower row: highest BDT output of all jet assignments per event in the 2 tag region for the muon channel (left) and the electron channel (right). Simulation is normalized to data.

The second-most important variable is $m(t_{\text{had}})$. Its importance can be attributed to the fact that in a semi-leptonic $t\bar{t}$ event, one expects two light jets which form the hadronically decaying W. For the signal process, there is only one light jet. This way it is reasonable that any variable related to the system of the two light jets in semi-leptonic $t\bar{t}$ would be very discriminating against the signal.

However not only the choice of variables is important, but it may also be relevant at which point in the multi-layer NN design of the analysis a certain information is exploited. In particular, the choice of $m_{b\bar{b}}$ as input variable for the reconstruction NN instead of classification NN must be scrutinized, as the reconstructed Higgs boson mass is supposedly what discriminates best between $t\bar{t}q$ and the dominant $t\bar{t}$ background. Not using the separation power of this variable in the classification NN but in the reconstruction NN might give a worse performance of the former. This can be judged in the following way:

4. Search for associated production of single top quark and Higgs boson at 8 TeV

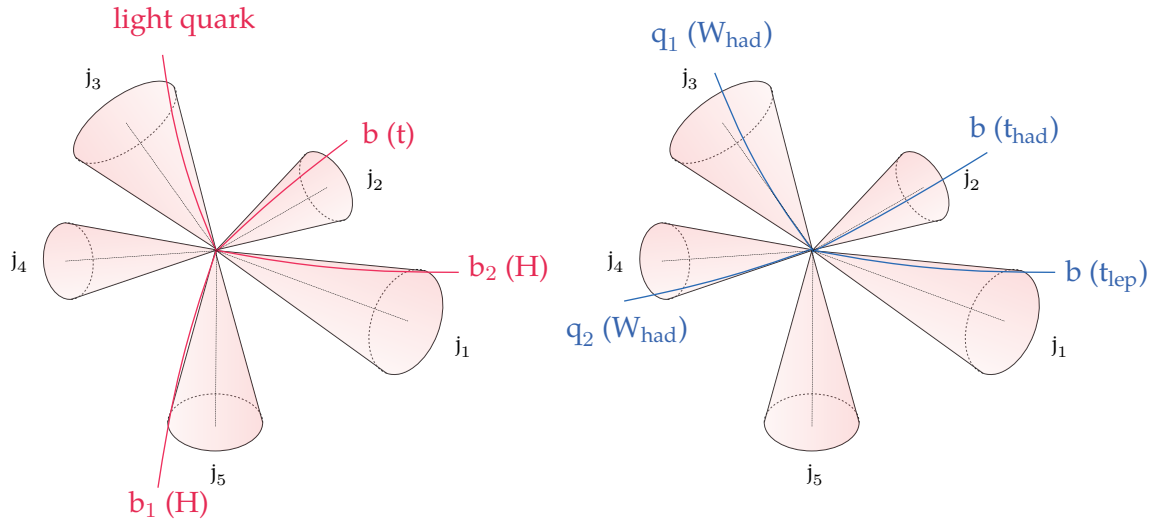


Figure 4.15.: The same reconstructed event, interpreted under the assumption it was a tHq (left) or $t\bar{t}$ (right) event. A correct assignment is found if all four quarks can be uniquely matched to reconstructed jets within a distance $\Delta R < 0.3$. The reconstruction efficiency (determined after the application of the reconstruction NNs) with respect to all events that have such a correct event interpretation was calculated in [40] and is 44% for tHq and 37% for $t\bar{t}$.

A comparison is performed between two independent analysis setups. For this test BDTs are chosen over NNs because of their speed; however they have shown comparably good performance. In one case, the Higgs boson mass is among the reconstruction NN input variables and is left out from the variable list in the classification step. The other scenario omits the Higgs boson mass in the reconstruction, but uses it in the classification NN. Finally, the ROC curves of the two classification NNs are compared to see if one gains in background rejection in one case with respect to the other.

As expected and as can be seen from the middle distribution in the top row of Fig 4.16, a well-localized Higgs boson mass peak can be observed if one uses the invariant mass of the two b jets that can be matched to the two b quarks from the Higgs boson as input in the reconstruction NN. However it does not provide a clear separation between signal and background in the subsequent classification NN anymore. The situation is different if the variable is omitted in the reconstruction and only used in the classification, employing its full discriminatory power only there for the first time. As can be seen from the upper right figure in Fig 4.16, the background distribution is much broader; this is expected for $t\bar{t}$, since there is no object decaying to a pair of quarks with a mass of 125 GeV for this process. From the ROC curve in the middle row, displaying the background rejection of the p_T of the reconstructed Higgs boson, a correlated effect is visible. This variable also separates better when the Higgs boson mass is only used in the classification. The distributions in the third row show the number of b tagged jets found among the jets assigned to the Higgs boson. Here the advantage of using the $m(H)$ information already in the reconstruction becomes visible: the reconstruction NN tries so hard to

4.6. Event classification

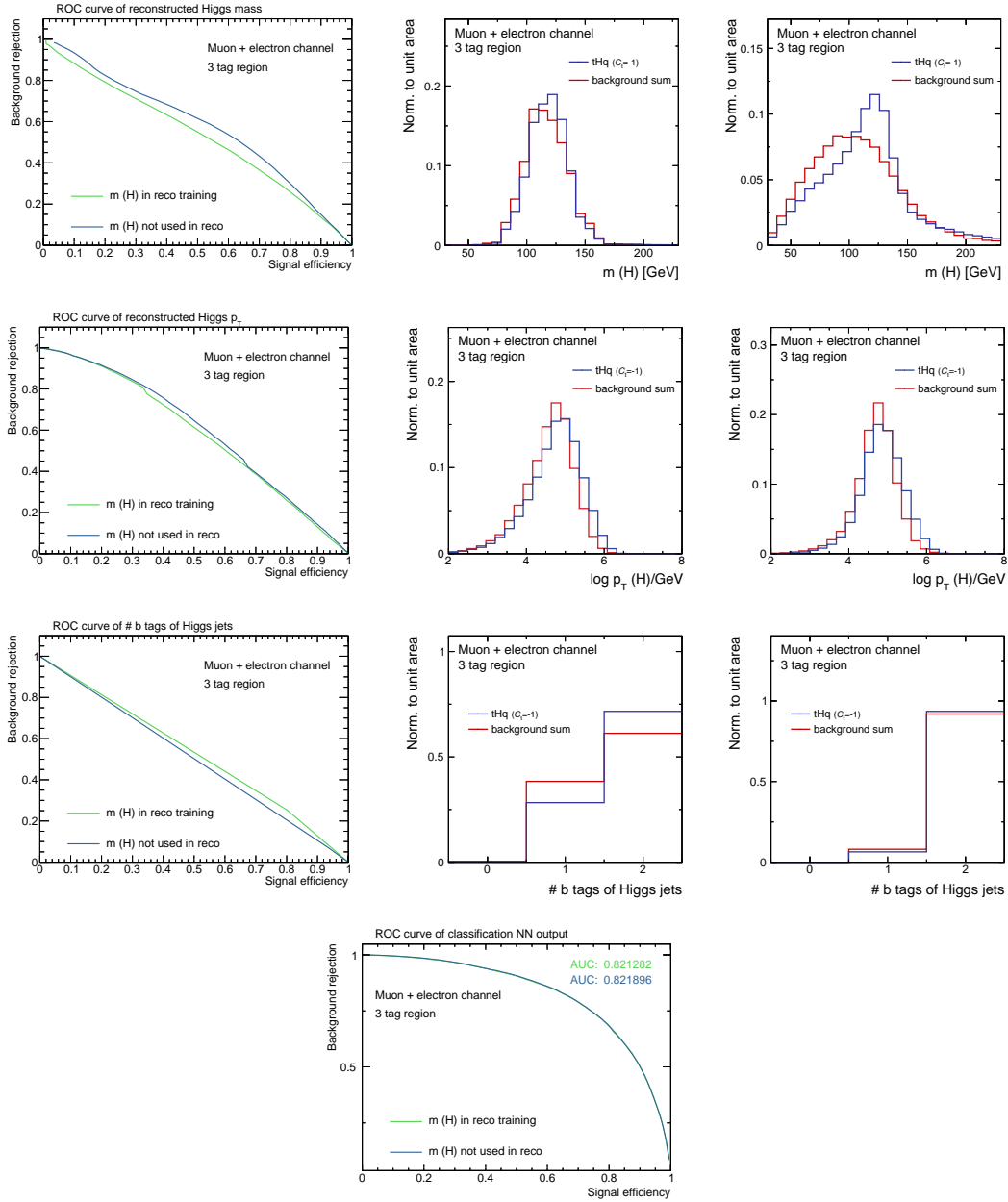


Figure 4.16.: Effect on classification NN input variable shapes of employing m_H in the reconstruction NN (middle column) or omitting it (right column). In the latter case, the reconstructed Higgs boson mass is used as input variable in the classification. The bottom figure shows it is irrelevant at which position the variable is used. The area under the curve (AUC) is identical for the ROC curves in both cases. As a consequence, it is used as input variable for the reconstruction NN.

4. Search for associated production of single top quark and Higgs boson at 8 TeV

find two jets with an invariant mass of 125 GeV that even for the case of the signal, they are not necessarily true b jets anymore, but in some cases random jets like e.g. jets arising from ISR or FSR, which then are assigned to the Higgs boson just because they have an appropriate invariant mass. This supposedly undesired behaviour however leads to the situation that the b jet multiplicity in the Higgs jets starts to discriminate between signal and background (middle figure in the bottom row), whereas the reconstruction NN would almost always pick two b tagged jets as Higgs jets in the case where the Higgs boson mass has not been used in the reconstruction (right figure in the bottom row). This is understandable, as the b tagging information is the other part that is relevant (besides the mass) in a $H \rightarrow b\bar{b}$ decay.⁵

The bottom row in Fig 4.16 displays the final ROC curves of the two scenarios. No difference can be seen in terms of background rejection or signal efficiency; the two analysis setups perform identically.

This nicely demonstrates it does not matter at which point of the stacked NNs the information of the Higgs boson mass is used. The final result is stable with respect to shifting variables from one layer to the other, as their information is always fully exploited by the NNs. This is an important test for demonstrating the architectural power of the analysis design.

Figure 4.17 shows the input shapes of the final variable set used in the classification for all processes, where it should be noted that the discriminating power against $t\bar{t} + X$ is of more importance than being able to separate between tHq and minor backgrounds such as electroweak diboson production. Nevertheless, shapes of all considered backgrounds are shown for completeness. In Figure 4.18 the correlations between the input variables can be seen. The set of variables was also optimized in consideration of having as few strongly correlated variables as possible. As it turns out, only $\Delta R(W_{\text{had}} \text{ jets})$ and $m(t_{\text{had}})$ have a larger correlation; still they add to the sensitivity of the NN due to the discriminating properties of the hadronically decaying W boson explained above. The correlations found for tHq and the backgrounds are almost identical.

Figures 4.19-4.21 finally show the data-MC comparison of all the input variables, first in the 2 tag region, then in the 3 and 4 tag regions. In these distributions, the integral of the sum of all MC receives a scaling factor such that it matches the number of events observed in data. This enables the reader to compare the shapes in data and simulation, which is the most important thing to make sure prior to training the NN. The correct normalization (also the correct relative normalization among the background processes) is adjusted for in the final fit to data in the process of deriving the exclusion limit.

Generally good agreement in shapes is found, and residual discrepancies are even covered by statistical uncertainties alone in almost all bins of all observables. This gives confidence for training the NN. The settings used for the training can be found in Appendix A. One important thing to make sure is that the NN has not suffered from

⁵A third aspect would be the color information of the Higgs boson decay objects, as the Higgs boson is a color singlet, which is in general not the case for randomly assigned jets from background processes. Variables addressing the color flow have been proposed in [182, 183] and have been tested as input variables, but did not provide additional sensitivity.

4.6. Event classification

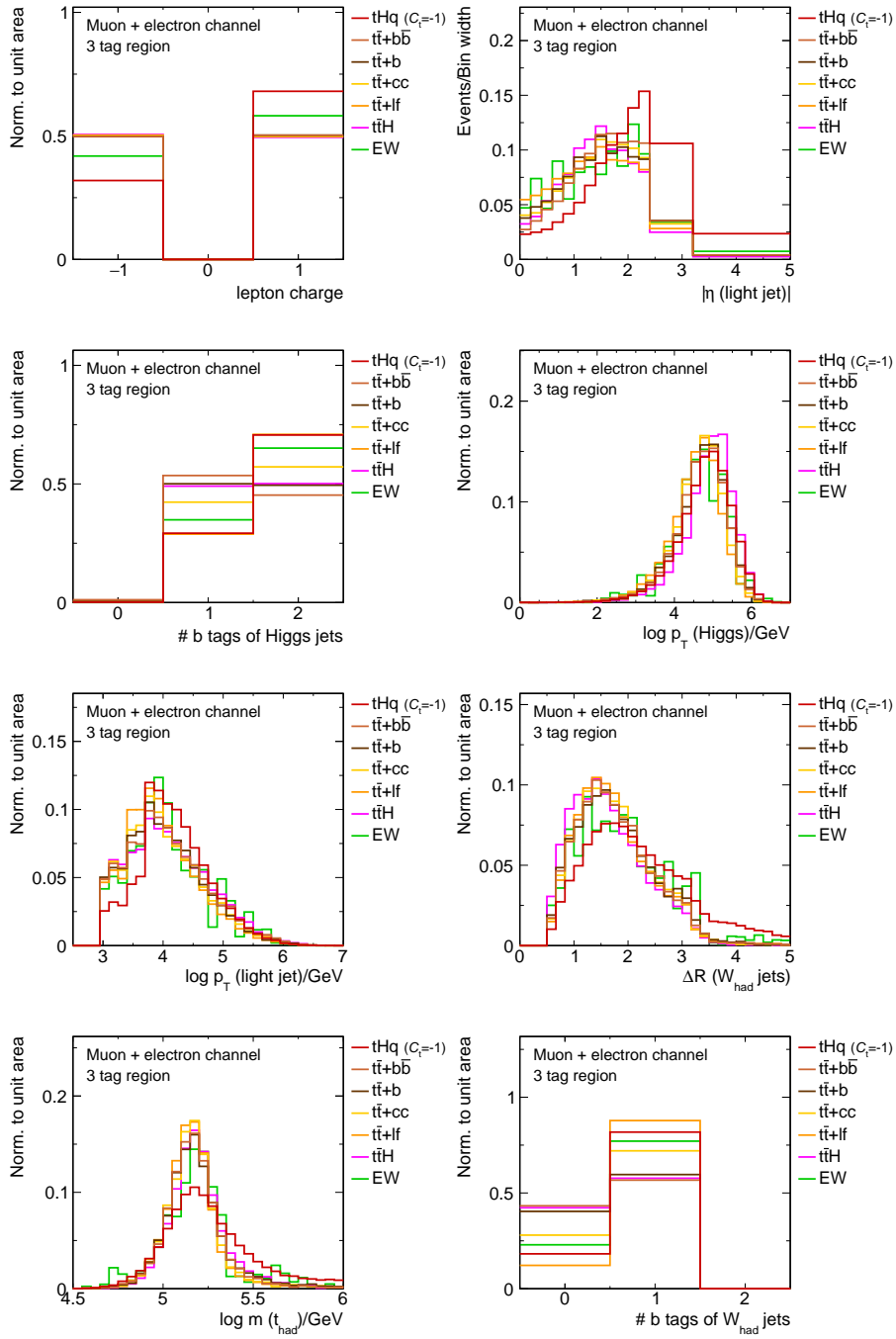


Figure 4.17.: Shape comparison of the eight classification NN input variables for the signal and background processes in the combined muon + electron channel of the 3 tag region.

4. Search for associated production of single top quark and Higgs boson at 8 TeV

Table 4.4.: Input variables of the classification NN, ranked according to the importance in the TMVA training. The row colors indicate which group the variables belong to: global (blue), tHq (red) or $t\bar{t}$ (yellow) observables.

Rank	Variable	Explanation
1	lepton charge	Electric charge of the lepton.
2	$\log m(t_{\text{had}})$	Reconstructed mass of the hadronically decaying top quark.
3	# b tags of Higgs jets	Number of tagged jets found among the jets assigned to the Higgs boson.
4	$\log p_{\text{T}}(\text{H})$	Transverse momentum of the reconstructed Higgs boson.
5	$\Delta R(W_{\text{had}} \text{ jets})$	Distance in the η - ϕ plane between the two jets assigned to the W boson of the hadronically decaying top quark.
6	# b tags of W_{had} jets	Number of tagged jets found among the jets assigned to the W boson of the hadronically decaying top quark.
7	$\log p_{\text{T}}(\text{light jet})$	Transverse momentum of the jet identified with the light recoil quark.
8	$ \eta(\text{light jet}) $	Absolute pseudorapidity of the jet identified with the light recoil quark.

overtraining. In the top figure of Fig. 4.22 no bias in the output of the NN is visible when being applied to a set of independent, unknown events. As an acid test, the events in the 2 tag region are objected to the classification. Although it has not been trained in this phase space, the NN is able to classify the bulk of ($t\bar{t}$) events as background-like. A perfect data-MC agreement is found, and the NN can be used in the signal-enriched 3 and 4 tag regions.

Pre-fit distributions for all analysis bins are shown in Fig. A.11 for completeness. It should be noted that they do not play a role in deriving the upper limit on the production cross section, hence the clear offset between data and simulation is not considered a problem, as it will be adjusted in maximum-likelihood fits during the limit setting procedure.

4.7. Systematic uncertainties

The search conducted in this analysis is afflicted with several sources of systematic uncertainty. The first group are the theoretical uncertainties, reflecting the present understanding of the involved processes in terms of modelling and accurate cross section

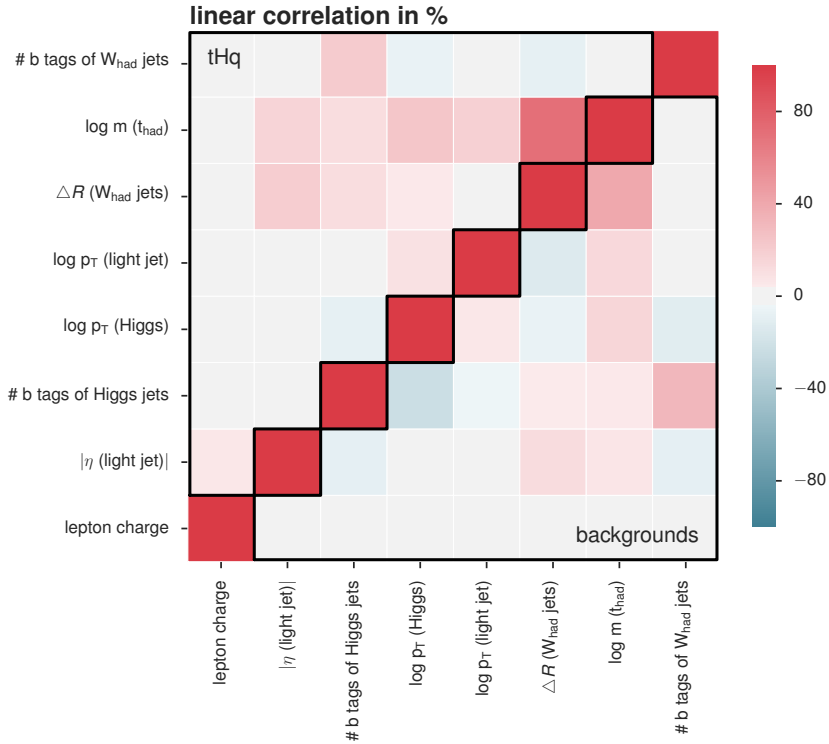


Figure 4.18.: Linear correlation coefficients between the classification NN input variables for both the signal and the background processes. The low correlation values are a result of the optimization the variable set has undergone and guarantees no redundant information is input to the Neural Network.

calculations. Secondly there are experimental uncertainties accounting for the imperfect performance of the detector and the applied reconstruction algorithms. Furthermore all these can be divided in two categories: “rate” uncertainties affect the yields of a process in a correlated way across the entire phase space, i.e. across the entire distribution of the classification NN output. The more general case however is that systematic effects can also alter the shape of the distribution; this is called a “shape” uncertainty in the following. It should be stressed that the various sources enter the analysis at the stage where they become relevant: a simple cross section uncertainty can be applied as rate variation on the final classification NN output. Uncertainties related to e.g. the jet energy scale already enter at the level of reconstructing the events. Hence both the reconstruction NNs and the classification NN are re-evaluated in such cases, and the effects of the respective systematic are propagated through the entire analysis chain.

4. Search for associated production of single top quark and Higgs boson at 8 TeV

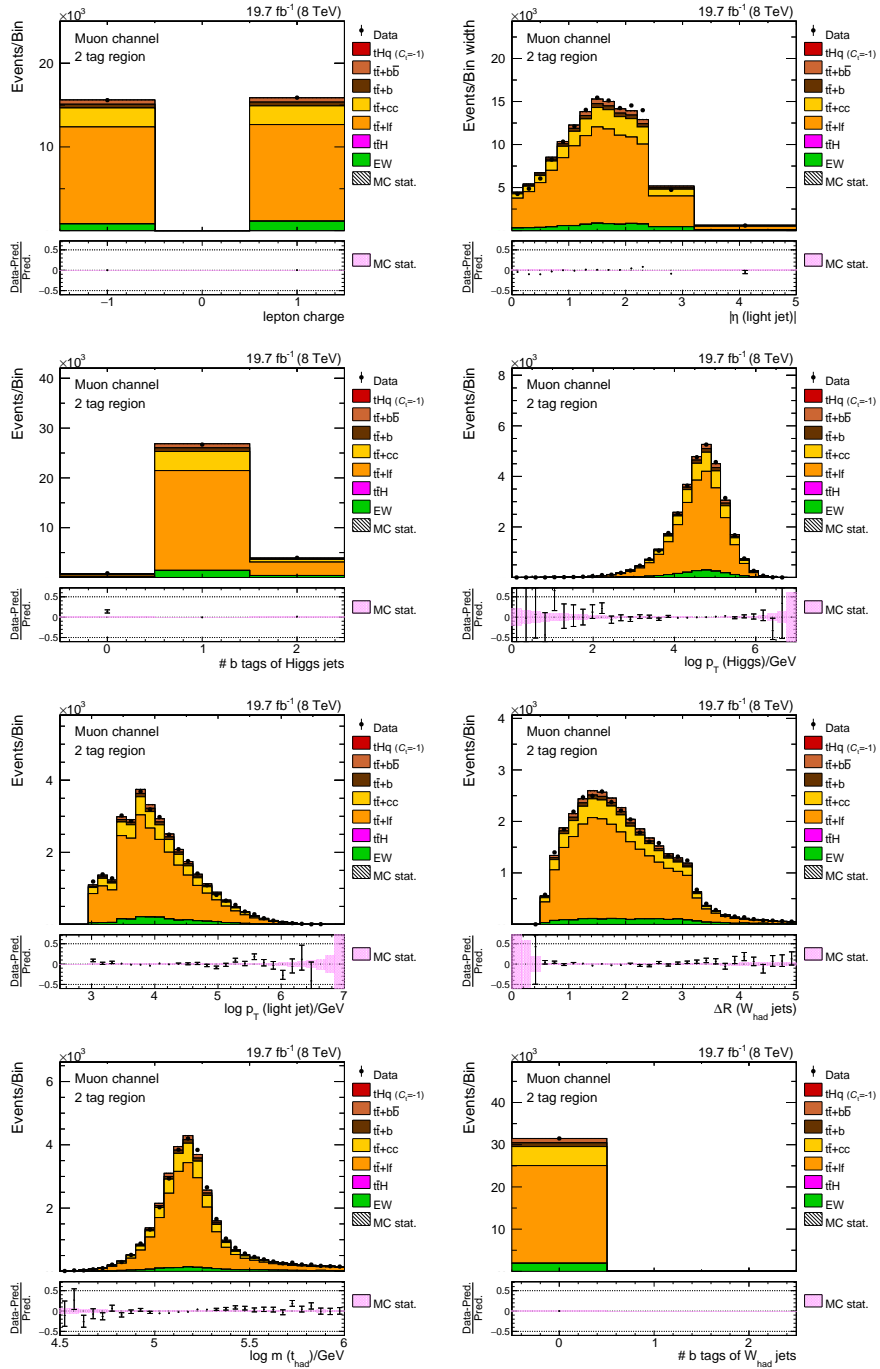


Figure 4.19.: Distributions of the input variables for the classification NN in the muon channel in the 2 tag region, which is enriched in top quark pair events. Simulation is normalized to data in order to facilitate the shape comparison. All MC weights are applied. The corresponding distribution for the electron channel can be found in Appendix A.

4.7. Systematic uncertainties

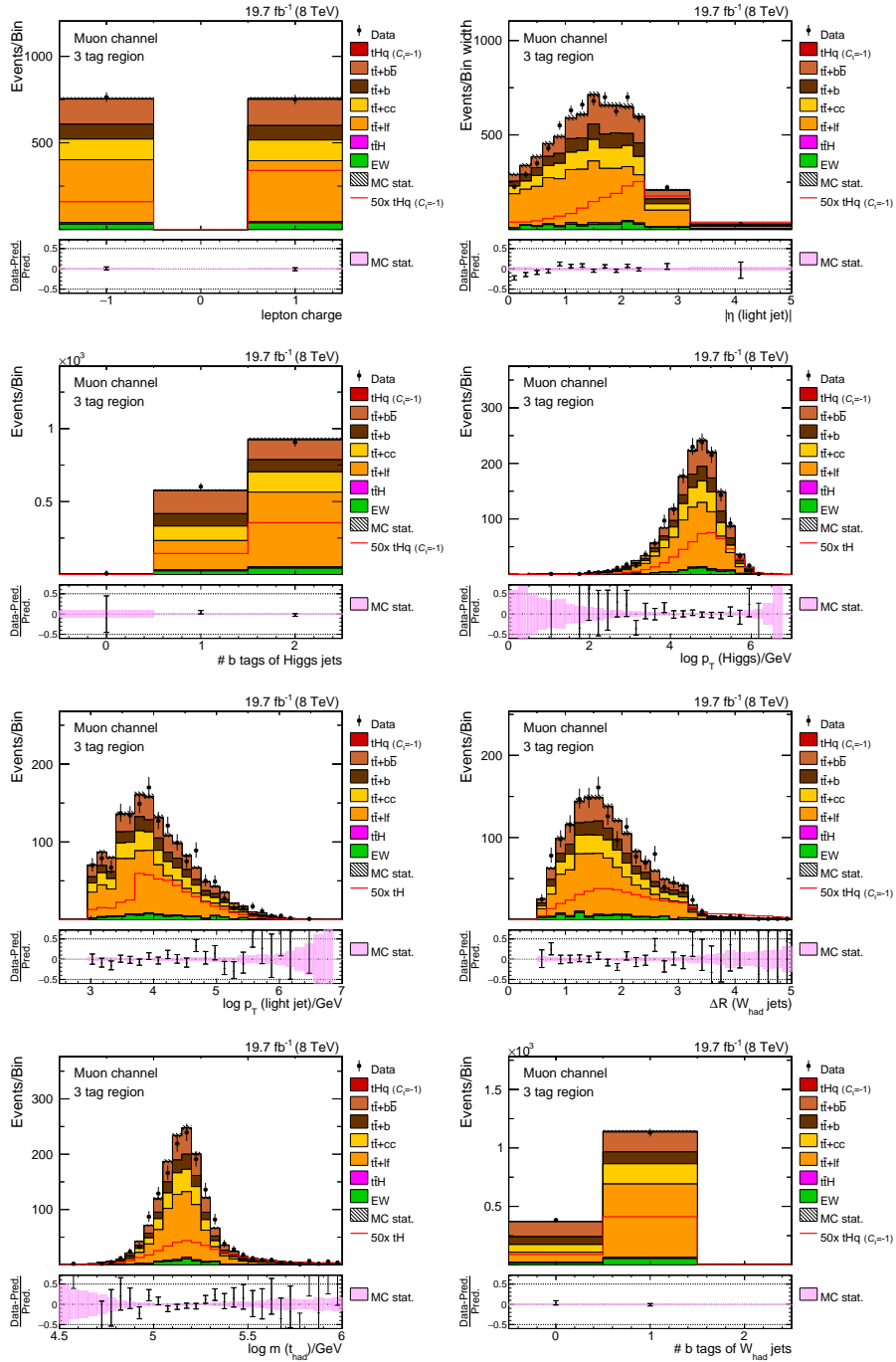


Figure 4.20.: Distributions of the input variables for the classification NN in the muon channel in the 3 tag region. Simulation is normalized to data in order to facilitate the shape comparison. All MC weights are applied. The corresponding distribution for the electron channel can be found in Appendix A.

4. Search for associated production of single top quark and Higgs boson at 8 TeV

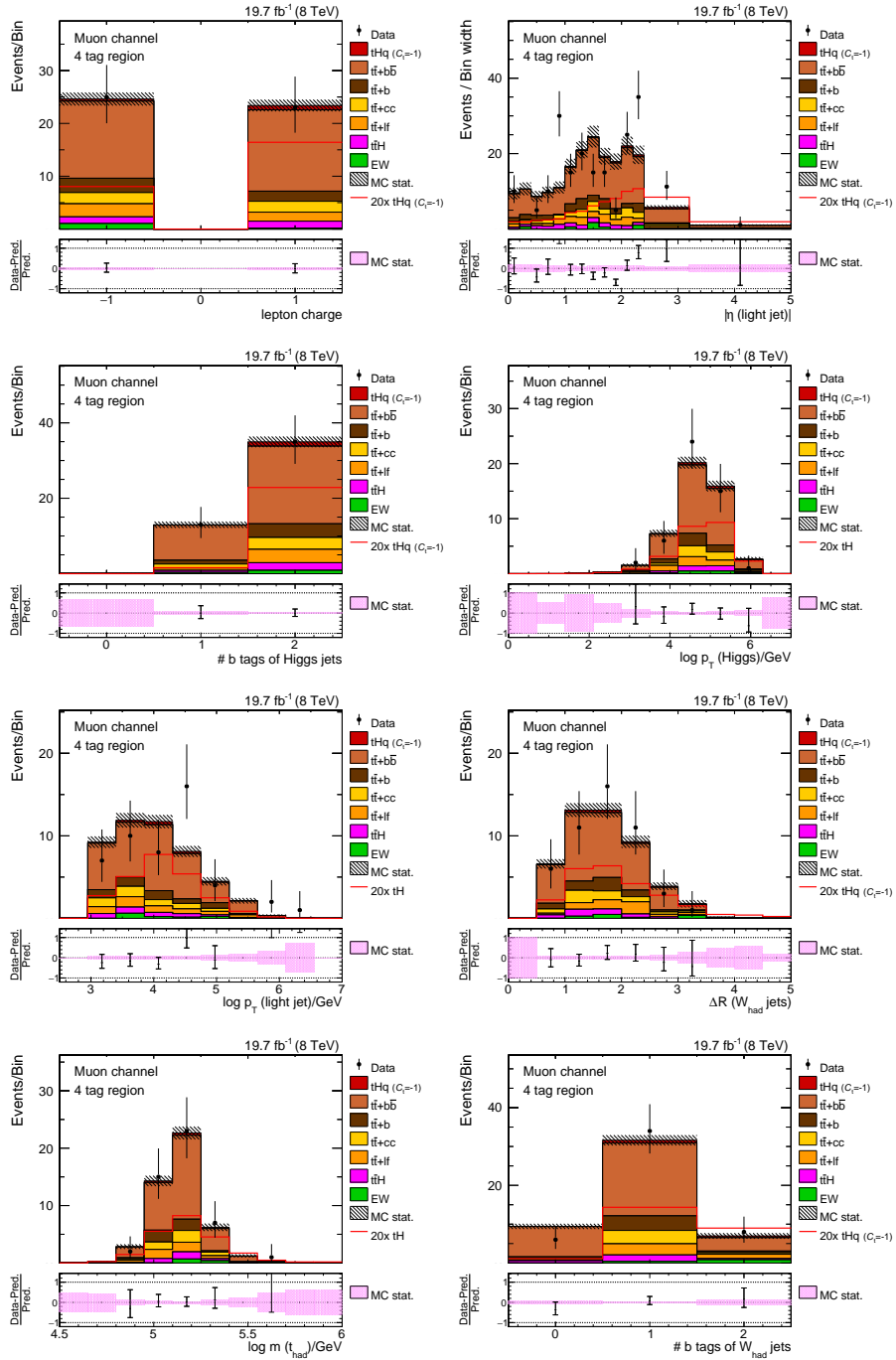


Figure 4.21.: Distributions of the input variables for the classification NN in the muon channel in the 4 tag region. Simulation is normalized to data in order to facilitate the shape comparison. All MC weights are applied. The corresponding distribution for the electron channel can be found in Appendix A.

4.7. Systematic uncertainties

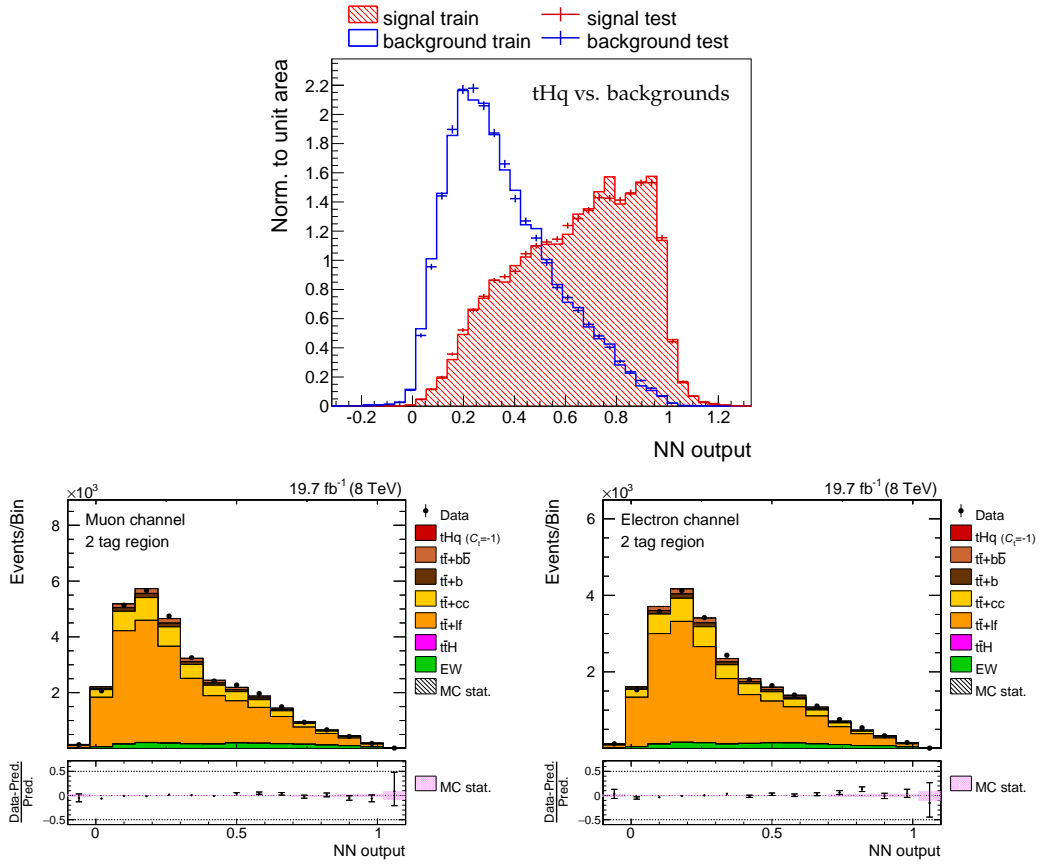


Figure 4.22.: Upper row: training results for the classification NN; a very good tHq-vs.-backgrounds separation can be seen. Lower row: NN output in the 2 tag region for the muon channel (left) and the electron channel (right). Simulation is normalized to data in order to facilitate the shape comparison; a perfect agreement is observed.

4. Search for associated production of single top quark and Higgs boson at 8 TeV

Table 4.5.: Variations of the cross sections for the considered processes due to uncertainties on the parton distribution functions, separated into the different initial states, and the Q^2 scale.

Process	PDF			Q^2 scale
	gg	q \bar{q}	qg	
tHq			2%	
t \bar{t} H	9%			12.5%
t \bar{t}	2.6%			3%
Single top			4.6%	2%
W + jets		4.8%		1.3%
Z + jets		4.2%		1.2%
Diboson				3.5%

4.7.1. Theoretical uncertainties

PDF uncertainties

The analysis assigns cross section variations as rate uncertainties, which can be grouped into two categories: for processes sharing the same initial state – a gluon pair (gg), a quark pair (q \bar{q}) or a quark and a gluon (qg) – the variations due to uncertainties on the PDFs are fully correlated. They are listed in Tab. 4.5. The cross section uncertainty for t \bar{t} is notably smaller than for t \bar{t} H, because for the former it is known at next-to-next-to-leading order in α_s [184], whereas for the latter the cross section has been calculated only at NLO [185].

Q^2 scale

Dedicated MC samples shifted in their $\mu_F = \mu_R = Q$ value by factors of 0.5 and 2.0 are used to estimate the effect of higher order corrections in QCD missing in the simulations. The scale is varied in the same direction for both the hard interaction part and the parton shower and hadronization. Such samples exist for the signal process and for the most important background, t \bar{t} . The effects due to missing higher order contributions are treated uncorrelated between signal and background. Due to different acceptances of the samples with respect to the nominal simulation and due to a different NN output value the same event can receive when its jet p_T 's are shifted, the NN output can be modified. It is therefore taken as a shape uncertainty in the model. The pre-fit NN output distributions of Fig. 4.23(a) show the effect of the Q^2 uncertainty on the shape of the discriminator for the signal and the background. As it will turn out, this is the most relevant source of systematic uncertainty. This fact is reflected in the large uncertainty bands in the ratio plot. Samples shifted in Q^2 do not exist for every background process. Instead, separate nuisance parameters as rate uncertainties for t \bar{t} H, single top quark, V + jets and diboson production are introduced. The assigned variations are given in the last

4.7. Systematic uncertainties

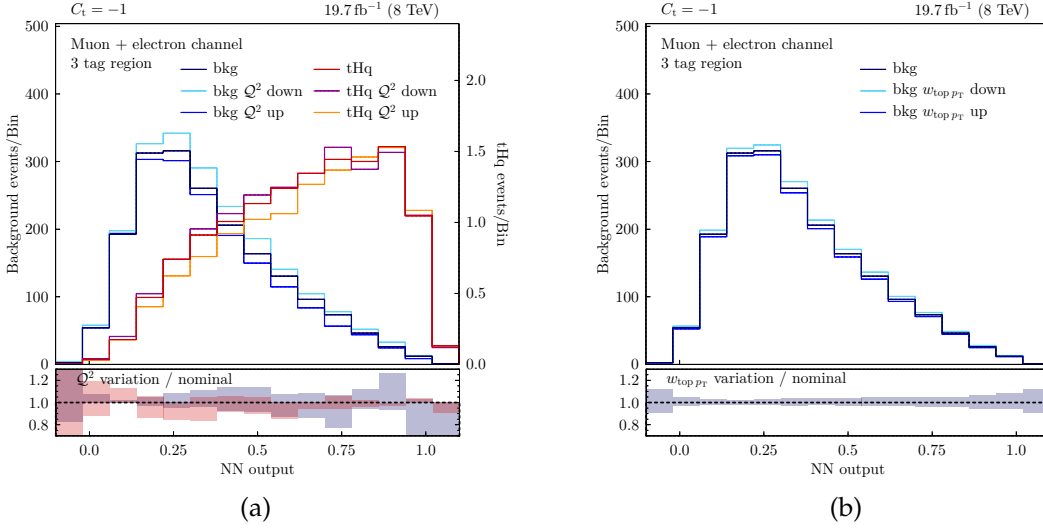


Figure 4.23.: Shape variations due to uncertainties on the Q^2 scale (a) and the top p_T reweighting (b).

column of Tab. 4.5. Although dedicated samples for the Q^2 scale variation exist for $t\bar{t}$, an additional rate uncertainty is also employed for the $t\bar{t}$ processes on top of the shape uncertainties. This is motivated by the fact that the samples systematically shifted in Q^2 are normalized to the nominal $t\bar{t}$ cross section before applying any selection cuts, and yield differences can only occur due to different acceptances. In order to account for the effect of a Q^2 variation on inclusive observables such as the total cross section, the uncertainty associated with the most accurate available calculation of the $t\bar{t}$ cross section is taken as rate uncertainty on the $t\bar{t}$ yields.

Matching threshold

In Sec. 4.3.1 it was mentioned that LO modelling was chosen for $t\bar{t}$ which incorporates additional jets on matrix element level. The threshold which divides the phase space into regions where either the parton shower or the matrix element generator is responsible for generation of the hard emissions is 40 GeV in the nominal $t\bar{t}$ sample. This value has been varied by +20 GeV and -10 GeV for accessing another uncertainty on the simulation of $t\bar{t}$ + jets processes.⁶ The Neural Networks are re-evaluated for separate matching up and down samples. The resulting variations are taken as a shape uncertainty.

⁶The exact values for the variation of the matching threshold are motivated by the range for which there is still a smooth transition in differential jet rates between the regions populated by the matrix element generator and the parton shower.

4. Search for associated production of single top quark and Higgs boson at 8 TeV

Modelling of top quark p_T

As was already mentioned in Section 4.4, the theoretical predictions for $t\bar{t}$ have undergone an ad-hoc reweighting depending on the p_T of the parton level top quarks to mitigate the disagreement between data and simulation observed in variables correlated to the true top quark transverse momentum. To this procedure, which has been introduced without a justification other than the resulting better agreement, an uncertainty needs to be assigned. The up- and down-variations correspond to applying the weight $w_{\text{top } p_T}$ twice and to not applying the weight at all. The consequence is different shapes of the NN output distributions for $t\bar{t}$, which is shown in Fig. 4.23(b).

$t\bar{t}$ + heavy flavor rates

The fraction of top quark pair production in association with heavy flavor jets is underestimated in the MADGRAPH+PYTHIA6 simulation – an issue which also seems to be there for other generator setups and which is not yet understood. Both ATLAS and CMS analyses measured the underestimation and report deficits of $\mathcal{O}(10\%)$, although with large uncertainties [186, 187]. In order to have enough lever arm for the normalization of the different components of $t\bar{t}$ + jets, a dedicated splitting is introduced in this analysis, following what was done in [158]. The following categories are defined by matching reconstructed jets to stable generator particles/hadrons: $t\bar{t}$ + $b\bar{b}$, $t\bar{t}$ + b , $t\bar{t}$ + $c\bar{c}$ and $t\bar{t}$ + light flavor jets. Extra nuisance parameters are introduced as rate uncertainties, serving the purpose of scale factors for the different heavy flavor components. Their prior uncertainty is 50%. These enter the statistical inference as rate uncertainties in addition to the previously discussed theoretical uncertainties which likewise address our imperfect knowledge of the modelling of $t\bar{t}$ events, such as the top quark p_T reweighting, etc.

4.7.2. Experimental uncertainties

Luminosity and pile-up

The total number of pp collisions that were recorded during Run-I is known only up to a certain accuracy. This uncertainty on the luminosity L amounts to 2.6% [173]. As a consequence, the expected rates at the given integrated luminosity of 19.7 fb^{-1} can only be known up to 2.6%. A corresponding nuisance parameter is introduced, controlling the rates for all considered processes in a correlated way.

The uncertainty on L also affects the expected number of pile-up interactions N_{PV} . Both the luminosity and the total inelastic cross section are input to the correction procedure for the simulated N_{PV} distribution. Uncertainties of $\pm 5\%$ on the event weights are applied and result in slightly different NN output shapes, which are treated as a separate nuisance parameter in addition to the rate uncertainty on L .

4.7. Systematic uncertainties

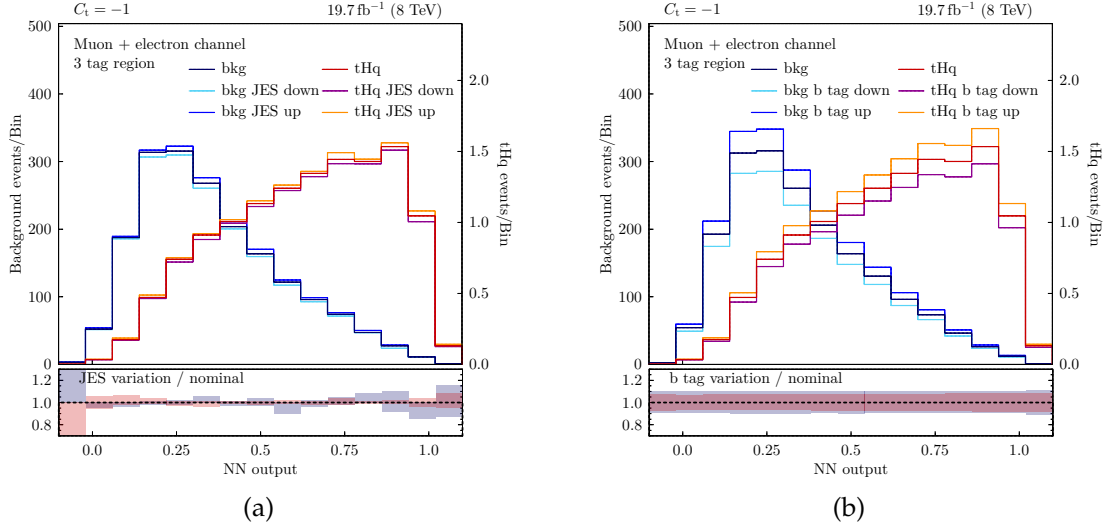


Figure 4.24.: Shape variations due to uncertainties on the jet energy scale (a) and the b tagging efficiency (b).

Jet energy

There are in total 17 uncorrelated sources of uncertainties on the jet energy scale [188]. Consequently 17 nuisance parameters enter the statistical evaluation as shape uncertainties on the discriminator after re-evaluating the NNs with updated jet energies. In order to obtain the shifted templates the jets are rescaled within $\pm 1\sigma$ of their uncertainties determined in [189]. For better illustration the variations from the different sources have been added in quadrature; their total effect on the signal and background NN output templates is shown in Fig. 4.24(b).

It has already been discussed in Sec. 4.2.4 that the jet energy resolution is smeared in order to obtain the same resolution for data and simulation. Altering the smearing factors by the $\pm 1\sigma$ uncertainties measured in dijet and $\gamma/Z + \text{jet}$ events [189] slightly changes the energy of the jets and therefore also the shape of the NN output. Only jets which can be matched to particle-level jets are subject to the systematic smearing.

Missing transverse energy

The modifications of the jet energy explained above also have an impact on \cancel{E}_T . The latter gets recomputed after the observed energies have been shifted. In addition the unclustered \cancel{E}_T from jets with $p_T < 10 \text{ GeV}$ and PF candidates which have not been clustered into jets is varied by $\pm 10\%$. The new NN output distributions are taken as shape uncertainties.

4. Search for associated production of single top quark and Higgs boson at 8 TeV

B tagging and mistagging

The weight $w_{b \text{ tag}}$ as explained in Sec. 4.4.3 is modified in order to account for imperfect b jet identification. The uncertainties in which they are allowed to vary are taken from [190]. Two independent nuisance parameters are introduced: the first one represents the actual b tagging efficiency after varying scale factors for b and c quark jets simultaneously. For the second one only scale factors for light flavor jets have been changed, which corresponds to a mistag efficiency. The altered shapes of the NN output distributions due to different event weights can be seen in Fig. 4.24(b) for the b tag efficiency.

4.8. Statistical uncertainties

The expected contribution to the yields in a bin i for a specific process represented by a simulated sample with N^{tot} total events and N^i events in bin i is given by $(\sigma L/N^{\text{tot}}) \cdot \sum_j^{N^i} w_j$. Here w_j is the event weight for the simulated event j , the inclusive cross section of the process is represented by σ and the integrated luminosity by L . N^{tot} is the number of events in the inclusive sample, while N^i is the number of events remaining in bin i after the final selection is applied. The above expression comes with an uncertainty $\delta = (\sigma L/N^{\text{tot}}) \cdot \sqrt{\sum_j^{N^i} w_j^2}$. For every process, bin and channel a separate nuisance parameter is introduced, for which the NN output shape is varied by $\pm\delta$ only in this bin. This leads to a proliferation of nuisance parameters in the statistical model. In order to reduce the dimensionality nuisance parameters are removed if the relative uncertainty on the predicted rate is $< 5\%$.

4.9. Discussion of systematic uncertainties

To quantify the importance of the sources of systematic uncertainty is crucial for understanding how one can make future tHq analyses more sensitive, besides collecting just more data or further optimizing the performance of the NNs (for instance by the use of new, additional variables).

The relevance of a specific source of systematic uncertainty and its impact on the final result can be judged in a way that is described in the following.

An asymptotic upper limit (for the many exclusion limits that need to be calculated for this study, the asymptotic method is picked because of time and CPU issues; however, tests have shown that they are very well in agreement with results derived in the full CL_S approach) is calculated one time including all systematic uncertainties (limit A) and another time omitting all the sources (limit B). This is done both after a fit to data, where the nuisance parameters are re-centered around their best-fit value and the uncertainties are constrained as from the fit result, and for a fit to an Asimov dataset thrown from the pre-fit background-only expectation. The former scenario is referred to as "post-fit", the latter as "pre-fit" in the remaining part of this Section. Every group of nuisance parameters is then either excluded from the full list, or taken as only source, and the

4.9. Discussion of systematic uncertainties

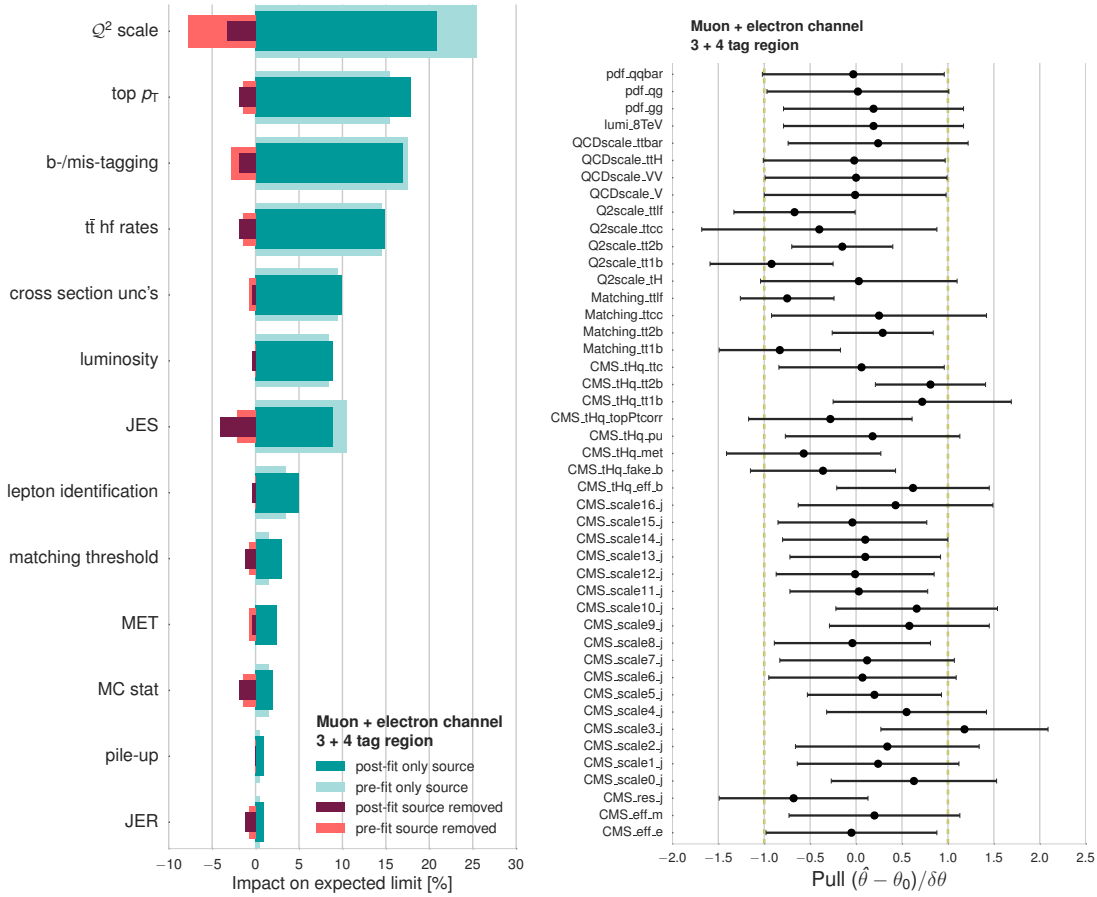


Figure 4.25.: Impact on final exclusion limit of the nuisance parameter groups (left) and pulls of the single nuisance parameters in the fit to data (right).

impact on the limit A or B is calculated. The bar chart of Figure 4.25 displays the changes in expected limits for the different scenarios, ranked according to the impact in the case where the systematic uncertainty has been used as only source after a fit to data. While the dark and light green bars give a hint on how a given uncertainty worsens the performance when put into the model as single source, the red bars, showing the impact of the uncertainty when being removed from the entire list, expresses also the correlation of the respective source with the others.

Four of the five most important systematic uncertainties are related to the modelling and theoretical understanding of the signal and the $t\bar{t}$ + jets processes and their normalizations. The Q^2 scale variation, which turns out to be the most important source, has a lower impact on the limit after the fit to data. The fit is able to constrain the uncertainties on some of the Q^2 nuisances (see their pulls in the right plot of the same figure), and the nuisances are pulled into a direction which obviously is beneficial from the search sensitivity point-of-view. The nuisances related to the matching thresholds get constrained in

4. Search for associated production of single top quark and Higgs boson at 8 TeV

a similar way, but get pulled away a lot from 0. Since they show a larger post-fit impact than in the pre-fit scenario, it needs to be assumed the related shapes are transformed in a way that is disadvantageous for the upper exclusion limit. This however does not pose a problem in general, and the impact of this uncertainty on the exclusion limit is moderate anyway.

Imperfect b jet identification is the only important experimental source; this is expected given the multi-b jet final state and selection.

Another interesting point is how the scale factors introduced for adjusting to a correct normalization of the various $t\bar{t}$ + heavy flavor components behave in the fit. Except for the case of $t\bar{t} + b\bar{b}$, they do not get overly constrained. Their pulls translate into the following scale factors: 1.03 ± 0.45 ($t\bar{t} + c\bar{c}$), 1.40 ± 0.30 ($t\bar{t} + b\bar{b}$), 1.36 ± 0.49 ($t\bar{t} + b$). This is in astonishing agreement (albeit with larger uncertainties) with the result of the dedicated measurement performed in [187], which reports $\sigma_{t\bar{t}b\bar{b}}/\sigma_{t\bar{t}jj} = 0.022 \pm 0.006$, while MADGRAPH and POWHEG predict 0.016 ± 0.002 , which corresponds to a scale factor of 1.38 for the $t\bar{t} + b\bar{b}$ component. It should be stated clearly however that the analysis searching for tHq production can be seen neither as a measurement of $t\bar{t}$ production in association with heavy flavor jets, nor as a cross-check for such. Nevertheless, the fit results are consistent with a dedicated, published measurement, hence raising the trust in the outcome. The fitting procedure is also fortified by looking at the pre-fit pulls in Fig. A.12 in Appendix A, which are not pulled away from their nominal value when fitting to an Asimov dataset.

As a last point it is worth mentioning that the pulls for the hundreds of nuisance parameters for the bin-by-bin statistical uncertainties are not shown here for convenience of the reader. A more elegant implementation using the Barlow-Beeston lite method [191], where the statistical uncertainties of all contributing processes are transformed into one single nuisance parameter per bin (as opposed to one parameter per process and bin), has not yet been available in the tools of the ROOSTATS project that have been used for the statistical inference in this analysis.

4.10. Results

The final yields after a simultaneous $s + b$ maximum-likelihood fit to data in the classification NN output distributions of the 3 tag and 4 tag regions (muon and electron bins) are given in Tab. 4.6. The fit has been performed with the combine tool⁷ and includes all 511 nuisance parameters of the model.

In the fit, the normalization of the signal is left free to float, and the best-fit signal-strength modifier r comes out of the fit as $r = 2.7^{+2.5}_{-2.3}$. The corresponding post-fit distributions are shown in Fig. 4.26, where a shaded band is representing the effects of both statistical and systematic uncertainties on the predictions. It covers the data points within their uncertainties for all analysis bins. No consistent excess across all distributions is visible in data in the regions where the signal is expected to peak (i.e. at

⁷A CMS internal tool based on ROOSTATS [192].

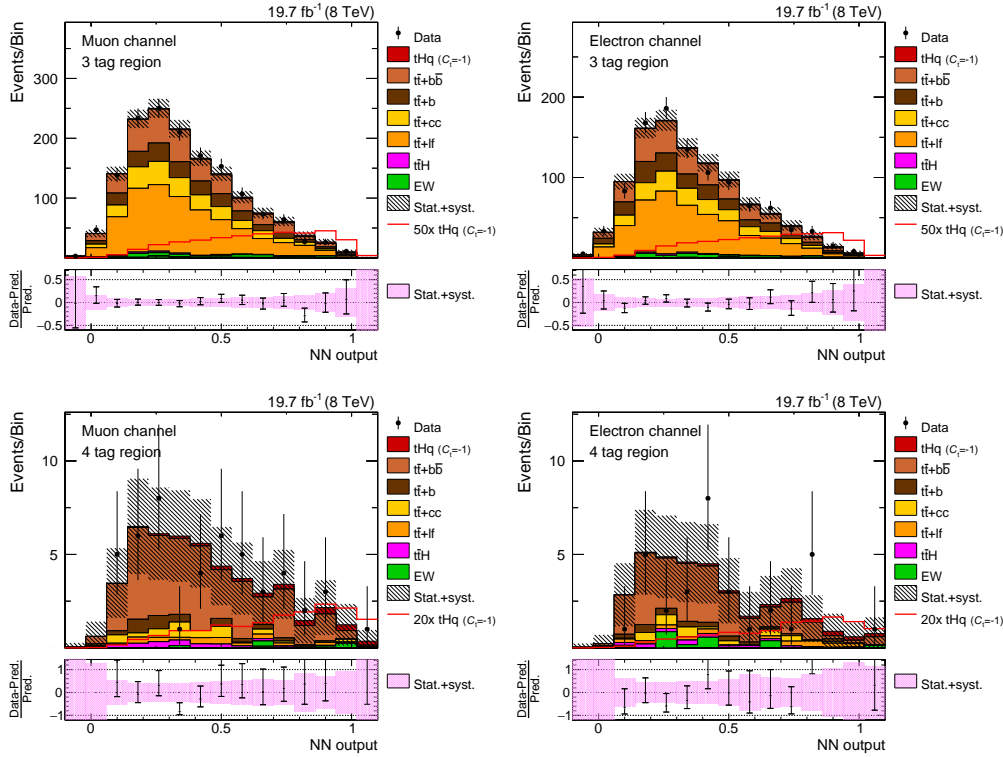


Figure 4.26.: Post-fit distributions of the final classification NN output. Uncertainties include statistical and systematic uncertainties. Good agreement between data and simulation is observed, differences are covered by the uncertainty band. The final upper boundary on the cross section of the $C_t = -1$ scenario is derived from similar distributions during the limit calculation.

high NN output values). A different representation of the outcome is shown in Fig. 4.28. All bins of the four distributions of Fig. 4.26 are sorted according to their expected S/B ratio (S (B) denotes the signal (background) yields), indicating that even for the bins with a large S/B ratio the b -only hypothesis is able to explain the observation by its own. Consequently, CL_s exclusion limits are derived from the NN output distributions in order to set an upper boundary on $tHq \rightarrow \ell\nu b\bar{q}b\bar{b}$ at 95% C.L.

A total number of ~ 45000 pseudoexperiments has been thrown in order to obtain the expected and observed limits in the 3 tag, 4 tag and the combined 3 + 4 tag region: an asymptotic limit has been calculated first, giving a feeling for the exclusion range and the search sensitivity of the analysis. Finally 200 pseudoexperiments have been thrown for various signal strengths r as suggested by the asymptotic limit calculation and combined to a grid from which the 2.5%, 16%, 50%, 84% and 97.5% quantiles and the observed limit have been determined. The results are given in Tab. 4.7. The analysis is able to exclude tHq production with subsequent decay of the Higgs boson to a pair of bottom quarks with a cross section larger than $7.6 \times \sigma_{C_t=-1}$ at 95% C.L., with an expected search

4. Search for associated production of single top quark and Higgs boson at 8 TeV

Table 4.6.: Final event yields and combined (statistical plus systematic) uncertainties at 8 TeV after the fit to data for the 3 and 4 tag regions. Shown are the expected numbers of signal and background events as well as the observed event counts.

Process	3 tag		4 tag	
	Muon	Electron	Muon	Electron
$t\bar{t} + b\bar{b}$	346 ± 62	245 ± 46	33 ± 6	23 ± 3.8
$t\bar{t} + b$	200 ± 78	152 ± 63	3.3 ± 1.2	2.5 ± 1.4
$t\bar{t} + c\bar{c}$	223 ± 106	150 ± 67	3.6 ± 1.5	2.7 ± 3.7
$t\bar{t} + lf$	645 ± 69	421 ± 46	2.3 ± 0.7	2.1 ± 0.7
single top	44 ± 4	29 ± 3	0.8 ± 0.1	1.4 ± 0.3
W/Z + jets	11 ± 2	3.6 ± 0.6	0.0 ± 0.0	0.9 ± 0.9
VV	1.1 ± 0.2	0.9 ± 0.1	0.1 ± 0.0	0.0 ± 0.0
$t\bar{t}H$	14 ± 2	10 ± 2	1.9 ± 0.3	1.6 ± 0.3
Σ	1483 ± 161	1013 ± 113	46 ± 7	34 ± 6
tHq ($C_t = -1$)	22 ± 16	15 ± 11	2.4 ± 1.8	1.7 ± 1.3
data	1514	1028	48	32

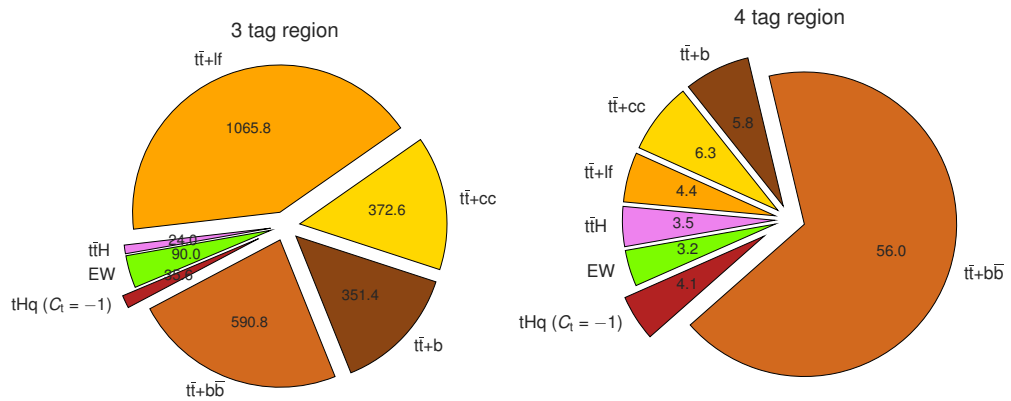


Figure 4.27.: Post-fit yields for 8 TeV in 3 and 4 tag regions.

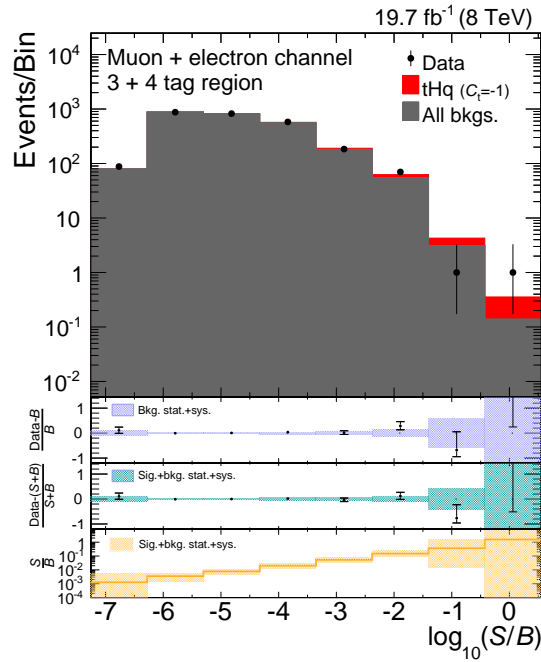


Figure 4.28.: The bins of all four NN output distributions of Fig. 4.26 are sorted according to their S/B ratio and collapsed into one single distribution with increasing S/B . The lower pads show the compatibility of data with the background-only or $s + b$ hypotheses together with the respective statistical and systematic uncertainties, and the S/B values.

Table 4.7.: Expected and observed 95% confidence level upper limits (8 TeV) on $\sigma/\sigma_{C_i=-1}$. The expected numbers are quoted together with the $\pm 1/2\sigma$ uncertainties on the limit.

Region	Expected	Observed
3 tag	$6.2^{+2.6/5.9}_{-1.8/2.8}$	7.0
4 tag	$9.7^{+4.3/10.3}_{-2.5/3.8}$	20.0
combined	$5.0^{+2.0/4.6}_{-1.3/2.2}$	7.6

sensitivity of $5.0 \times \sigma_{C_i=-1}$.⁸ This translates into an excluded production cross section of > 1.78 pb, which is a surprisingly tight upper boundary given the complexity of the final state and the overwhelming size of the background. The analysis presented above has been made public in [156].

⁸Although sizeable at first glance, the difference between observed and expected upper limits corresponds to a 1.3σ effect and therefore is not worrisome.

4. Search for associated production of single top quark and Higgs boson at 8 TeV

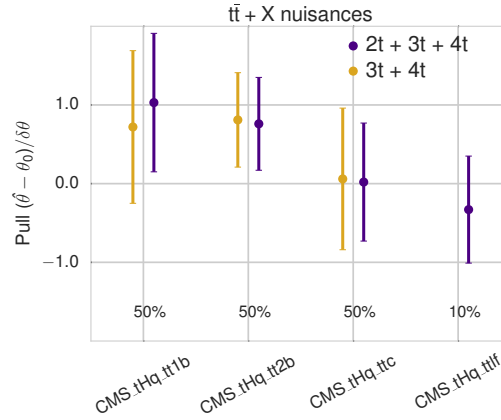


Figure 4.29.: Comparison of pulls of $t\bar{t} + X$ nuisances if the maximum-likelihood fit has been performed in the 3 + 4 tag region (brown) or adding the 2 tag region (purple). The numbers in % give the prior rate uncertainties for the processes.

4.11. Possible improvements to the $H \rightarrow b\bar{b}$ analysis

In particular with respect to the 13 TeV data taking period, it is important to identify possible ways for improvements of the achieved search sensitivity. This will become especially relevant when ultimately looking for SM-like single top + Higgs production, where, because of the smaller cross section, ultimate performance is desired. First hints have already been found in the context of this thesis:

The first strategy seeks to include more phase space in order to have more signal events and at the same time to be able to better constrain relevant sources of systematic uncertainty; the second element is to further optimize the sensitivity of the NNs by employing additional information.

4.11.1. Inclusion of the 2 tag region in the limit calculation

Despite being a control region with only a very small S/B ratio of $112/84000 \sim 0.1\%$, it is expected that the overall sensitivity of the analysis is increased when including the 2 tag region in the limit calculation. It should help in constraining the $t\bar{t}$ background in a region that is largely dominated by events where a $t\bar{t}$ pair is produced in association with light flavor jets. Hence the statistical evaluation and limit calculation is performed in six analysis bins – the electron and muon channels of the 2, 3 and 4 tag regions – assuming a correct modelling of the correlations between regions with different jet multiplicities. In addition, trading the control region for a smaller fit error is deemed a relevant approach, since the purpose of the control region is only to validate shapes of NN input and output distributions. A determination of the normalization of the background processes happens in the limit fit via the constraining of corresponding nuisance parameters. The 2 tag region is not employed as a sideband for an a-priori determination of background contributions, which would prohibit its use in the signal extraction.

4.11. Possible improvements to the $H \rightarrow b\bar{b}$ analysis

The pre-fit distributions of Fig. A.11 show a large offset between predictions and observed yields for the 2 tag region, which is dominated by $t\bar{t} + l\bar{l}$ processes. To be on the safe side an additional rate uncertainty on this process of 10% is introduced and applied in all regions. This provides enough lever arm to adjust for the discrepancies of the overall yields. Fig. 4.29 compares the pulls of the nuisance parameters for the nominal fit of Sec. 4.10 and for the fit including the 2 tag region. The central value changes notably only for the $t\bar{t} + b$ component, while the uncertainties slightly shrink for all nuisance parameters of the $t\bar{t} + h\bar{h}$ rates. The post-fit distributions of Fig. 4.30, compared to Fig. 4.26, do not indicate large differences for the 3 and 4 tag regions, while the agreement in the 2 tag region between data and prediction is found to be very good once this region has been included in the maximum-likelihood fit. A red hollow histogram shows the expected distributions of the signal process scaled by 300, which illustrates how small the S/B ratio is.

Exclusion limits are calculated and compared to the previously achieved search sensitivity. The median expected exclusion limit improves from 5.0 to 4.4 in units of $\sigma_{C_t=-1}$, which means the sensitivity is 10% better than for the nominal setup. Also the observed limit improves, as can be seen in Fig. 4.30, leading to an exclusion of $6.5 \times \sigma_{C_t=-1}$ or, equivalently, a cross section larger than 1.5 pb at 95% C.L.

4.11.2. Using the full CSV shape information

Throughout the analysis, Neural Networks and selection criteria rely only on the boolean information whether or not a jet passes a (tight) working point of the CSV b tagging algorithm. Event weights are applied to correct for observed differences in tagging (in)efficiencies between data and simulation. The method was described in Sec. 4.4.3, and their effect could be seen in Fig. 4.8. The absolute CSV value assigned to a jet however is a continuous variable and is a strong classifier for distinguishing b from non-b jets. The information hidden in the entire CSV shapes of the considered jets can bring more discriminating power for the Neural Networks and hence make the analysis more sensitive.

However, the distribution in Fig. 4.32(a) shows that the CSV shape of jets is badly described by simulation. The disagreement is worst at low CSV values, where simulation overshoots the data. Due to the nature of the $t\bar{H}q$ topology in the t -channel with the distinct light recoil quark, it is of high importance to especially have the region at low values of the CSV distribution correctly described, such that the analysis can rely on the full CSV shape of the jet identified with the light jet. The CSV shapes of b jets from the Higgs boson and top quark are also expected to help the NNs gain in performance because of the genuine multi-b jet final state; many background events enter the selection due to misidentified light quark or gluon jets.

For this reason a dedicated reweighting is employed to fix the agreement in the CSV shape, which has first been introduced in [158] for the $H \rightarrow b\bar{b}$ decay channel and is provided by the CMS $t\bar{t}H$ group. The method derives scale factors for light and heavy flavor jets in two dedicated control regions via tag-and-probe: in a region dominated in $t\bar{t}$ (2 jets, ≥ 1 b tag) heavy flavor correction factors are derived; a region enriched in

4. Search for associated production of single top quark and Higgs boson at 8 TeV

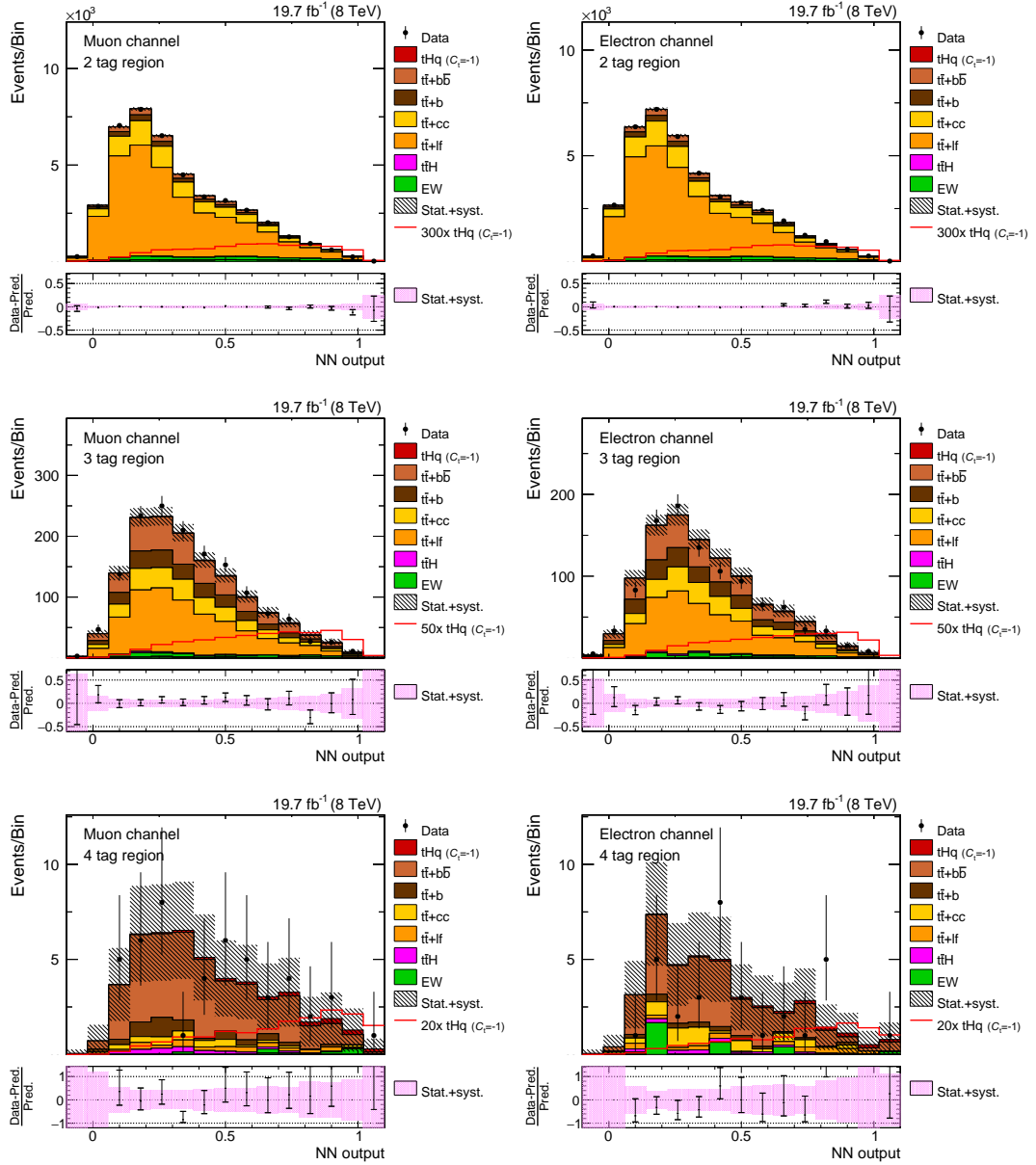


Figure 4.30.: Post-fit distributions of the final classification NN output, this time including the 2 tag region (e, μ) in the fit. Uncertainties include statistical and systematic uncertainties.

4.11. Possible improvements to the $H \rightarrow b\bar{b}$ analysis

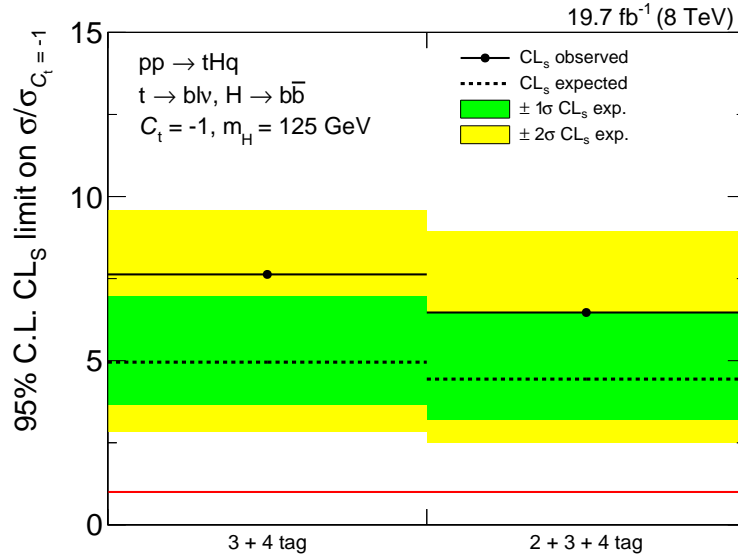


Figure 4.31.: Comparison of combined limits for the nominal scenario (derived in the 3 + 4 tag region) and with including the 2 tag region in the limit calculation. An improvement in the search sensitivity of 10% is achieved.

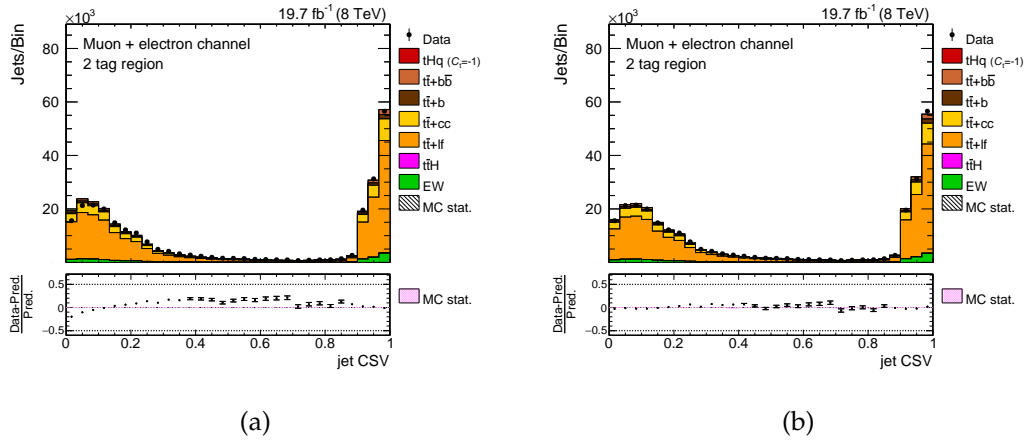


Figure 4.32.: CSV values of all jets with $|\eta| < 2.4$ and $p_T > 20$ GeV entering the selection before (a) and after (b) the CSV reweighting. A clear improvement in terms of data/MC agreement is observed when employing dedicated scale factors derived from control samples.

4. Search for associated production of single top quark and Higgs boson at 8 TeV

Table 4.8.: Observed and expected upper limits on $\sigma/\sigma_{C_t=-1}$ for other tHq channels and for the combination of all searches.

tHq channel	95% C.L. upper limits on $\sigma/\sigma_{C_t=-1}$			
	Observed	Expected		
		Median	68% C.L. range	95% C.L. range
$\gamma\gamma$	4.1	4.1	[3.7, 4.2]	[3.4, 5.3]
Multilepton	6.7	5.0	[3.6, 7.1]	[2.9, 10.3]
$\tau\tau$	9.8	11.4	[8.1, 16.7]	[6.0, 24.9]
Combination with $b\bar{b}$	2.8	2.0	[1.6, 2.8]	[1.2, 4.1]

di-leptonic Z + jets events (2 jets, ≤ 1 b tag) is employed to calibrate the CSV discriminator of light flavor jets, by requiring one either tagged or antitagged (i.e. not passing a certain cut on the CSV discriminator) and building the ratio $(\text{observation} - \text{MC}_{\text{lf/hf}})/\text{MC}_{\text{lf/hf}}$. This ratio is then equal to the scale factor. It is calculated in various p_T and η bins and then interpolated to obtain continuous correction factors. The distribution in Fig. 4.32(b) impressively demonstrates how well the predictions match the data after the correction. The downside is that this method comes with many and large uncertainties (e.g. limited statistical power in the control regions, or different assumptions on the interpolation) entering the statistical inference.

Preliminary studies [193] for $C_t = +1$ show a clear improvement in terms of ROC curves for a classification NN that uses the full CSV shape of jets associated with the objects of both tHq and $t\bar{t}$ reconstructions. This effect is attenuated in the final exclusion limits due to new sizeable systematic uncertainties entering the statistical model. However, the improvement of the CL_S limit is still in the range of $\sim 11\%$.

4.12. Combination with other decay channels

A combination with other analyses looking for different final states of the Higgs boson ($H \rightarrow \tau^+\tau^-$, $H \rightarrow W^+W^-$ [194] and $H \rightarrow \gamma\gamma$ [195]) has been performed in the context of [48] and allows to set more stringent limits on tHq production than any of the decay channels alone. The $H \rightarrow \gamma\gamma$ analysis is the most sensitive one, mainly due to the further enhancement in the decay because of the interference between the top- and W-loop for $C_t = -1$. The $H \rightarrow b\bar{b}$ and the $H \rightarrow W^+W^-$ (i.e. multilepton) channels have the same expected search sensitivity, followed by the analysis looking into $H \rightarrow \tau^+\tau^-$ with at least one subsequent hadronic τ decay. A detailed description of the separate analyses can be

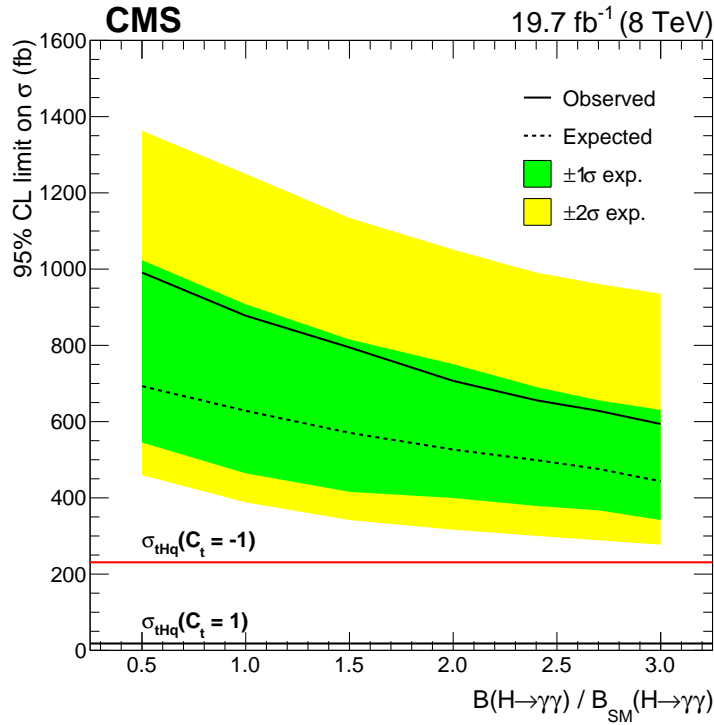


Figure 4.33.: Combined observed and expected upper limit on the production rate of tHq as a function of $B(H \rightarrow \gamma\gamma)/B_{SM}(H \rightarrow \gamma\gamma)$. Taken from [48].

found in [48]. Here only the results will be quoted. Combining the analyses summarized in Tab. 4.8 with the search for $H \rightarrow b\bar{b}$ presented in this thesis gives an expected upper limit on $\sigma/\sigma_{C_t=-1}$ of 2.0, while the observation allows to exclude 2.8 times the cross section for $C_t = -1$. These are numbers for a pure $C_t = -1$ scenario. This implies the enhancement of the branching fraction $BR(H \rightarrow \gamma\gamma)$ has been fully taken into account even for the processes $t\bar{t}H$, $H \rightarrow \gamma\gamma$ and VH , $H \rightarrow \gamma\gamma$. However it is also interesting to present results assuming the branching ratio is fixed to its SM value. Finally several BSM scenarios can alter $BR(H \rightarrow \gamma\gamma)$. For this reason Fig. 4.33 shows the upper exclusion limits at 95% C.L. for a range of $BR(H \rightarrow \gamma\gamma)$ normalized to the SM scenario. The red horizontal line illustrates the $C_t = -1$ case.

4.13. Chapter summary

The presented analysis corresponds to the first dedicated search for anomalous production of a Higgs boson in association with a single top quark in the $H \rightarrow b\bar{b}$ decay channel developed within the CMS collaboration and made public through a Physics Analysis Summary [156]. In this chapter all the aspects of the analysis have been recapulated and addressed. Special emphasis has been put on the development of the analysis strategy

4. Search for associated production of single top quark and Higgs boson at 8 TeV

and event selection, on the strategy in regard to Monte-Carlo simulation and the optimization of the analysis performance, and on the statistical evaluation and interpretation, as the author of this thesis has made key contributions to all of the aforementioned analysis ingredients.

The tHq production mode provides a special handle on one of the key parameters of the SM, the top-Yukawa coupling y_t : already at leading order there is an interference between the amplitudes of Feynman diagrams where the Higgs boson couples either to the top quark or to the space-like W boson. The tHq channel is therefore sensitive to the relative sign of the two couplings. In particular for the case of $\sigma_{C_t=-1}$ (with $C_t = y_t/y_t^{\text{SM}}$), which is not yet excluded by LHC data, the cross section is largely enhanced with respect to the SM. By analyzing 19.7 fb^{-1} of pp collision events at a center-of-mass energy of 8 TeV with sophisticated multivariate techniques – NNs have been used for event reconstruction and classification – an upper limit on the scenario of a flipped top-Yukawa coupling is set, excluding a cross section $\sigma_{C_t=-1}$ larger than 1.78 pb at 95% C.L. This corresponds to a signal strength modifier $\mu = \sigma/\sigma_{C_t=-1}$ that is smaller than 7.6 at the same level of confidence. A combination with analyses investigating other Higgs boson decay channels narrows down the range allowed for this parameter to $\mu < 2.8$.

The upper boundary on μ suggests that a direct exclusion of the anomalous $C_t = -1$ scenario is within reach of dedicated searches for tHq production. As such, they can be regarded as a complementary approach to the one followed by the ATLAS collaboration that treats tHq solely as background in the search for the $t\bar{t}H$ process and incorporates this channel in a global fit to ATLAS data. Being the only production mode without top quarks in virtual contributions at lowest order that is sensitive to the sign of C_t , the tHq contribution helps in lifting the $C_t = -1 \leftrightarrow C_t = 1$ degeneracy that is there if the virtual corrections involving top quarks of all other production and decay modes are described by effective couplings (see Fig. 1.9). The latter assumption reflects the possibility that new particles enter the loop diagrams. While the ATLAS approach aims at optimal sensitivity in a global coupling fit, the dedicated searches for tHq in CMS are more direct and eventually address the exclusion or discovery of the (anomalous) associated production of single top quark and Higgs boson. It is worth mentioning that, since the single tHq searches in CMS have been optimized with respect to their own search sensitivities, there naturally exists a huge overlap in event selections with corresponding $t\bar{t}H$ searches. This makes it impossible to include the tHq results in a global fit to 8 TeV CMS data.

5. Prospects for a search for associated production of single top quark and Higgs boson at 13 TeV

For the first time in accelerator history, protons collided at the LHC with a center-of-mass energy of $\sqrt{s} = 13$ TeV on July 3, 2015. By the end of the same year, around 3.8 fb^{-1} of collision data have been collected with the CMS detector, from which 2.2 fb^{-1} are analyzed in this thesis. The purpose of this chapter is to quantify for the first time the search sensitivity achievable already at such low integrated luminosities. The chapter can therefore be seen as an outlook on the feasibility of a search for tH at 13 TeV. The analysis is performed similarly to the one presented in the previous chapter, but features both technical and conceptual improvements developed in the course of this thesis. Corrections of a data-driven procedure improving the data-MC agreement in the b tag discriminators are applied, making it possible to fully exploit their discriminating potential. Moreover, multiple points in the 2d plane of C_V and C_t are investigated, extending the scope of the search compared to the single scenario of $C_t = -1$ in Run-I, whose inclusive production cross section is augmented from $\sigma^{8 \text{ TeV}} = 234 \text{ fb}$ to $\sigma^{13 \text{ TeV}} = 793 \text{ fb}$ at $\sqrt{s} = 13$ TeV. While reaching out to points where $|C_t|, |C_V| \neq 1$, the top quark and W boson masses are taken to be SM-like. This assumption implies the existence of physics beyond the SM and of another mass generation mechanism and is realized within the Higgs characterization model [45], which is used for the generation of the signal samples in Run-II. The analysis employs an effective field theory approach postulating new BSM particles as they may arise from additional Higgs fields. These must be much heavier than SM particles so that Higgs boson interactions are still characterized by 4-dimensional operators, and effects of the additional fields are absorbed into the 4-dimensional couplings, which can then be varied to probe the coupling against the experimentally determined top quark mass of around 172.5 GeV. It should be stressed that only the effect of a change in the top-Yukawa coupling and in the coupling to gauge bosons on the production cross section is accounted for. The interaction of the Higgs boson with bottom quarks is taken to be purely SM-like. As another conceptual extension and improvement in the search for single top quark + Higgs boson associated production, the tWH production mode is studied and considered as an additional signal process for the first time in the $H \rightarrow b\bar{b}$ channel.

The chapter starts with an overview of the changes in terms of signal and background modelling, which has been refined in the context of this thesis, and the event selection with respect to Run-I. Event reconstruction and classification are briefly addressed. Finally a preliminary expected upper limit on each point in the C_t - C_V plane is set, and a projection to larger integrated luminosities, which are expected and eagerly awaited during the next years of LHC data taking, is provided. Observed limits are not quoted

5. Prospects for a search for associated production of single top quark and Higgs boson at 13 TeV

Table 5.1.: Simulation settings for the signal processes at 8 and 13 TeV. The tWH process has not been considered in the 8 TeV analysis. ME: matrix element; PS: parton shower; UE: underlying event. The cross section have been obtained at NLO QCD.

\sqrt{s}	8 TeV	13 TeV	
Process	tHq	tHq	tWH
ME generator	MadGraph 5	MG5_aMC@NLO	
ME precision	LO	LO	
PS	Pythia6	Pythia8	
UE tune	TuneZ2*	CUETP8M1	
QCD description	4F	4F	5F
$\mu_{R,F}$	fixed, 100 GeV	dynamical, $\Sigma_{H,t,q} m_T/6$	fixed, 40 GeV
PDF	cteq611	NNPDF30_lo_as_0130_nf_4	NNPDF30_lo_as_0130
Normalized to	5F NLO	4F NLO	5F NLO
$\sigma_{C_t=-1}$ [fb]	233.8	792.7	147.2

due to the blinding of the analysis, i.e. data are removed from the distributions that are most sensitive to the presence of the signal (the BDT outputs in the signal regions).

5.1. Changes in signal and background modelling

Due to recent developments in the field of MC simulation and new phenomenological studies on tHq production, the way of modelling the signal process has been revised for the first Run-II analysis. It has been adapted to the latest phenomenological studies on tHq performed in [35]. While the accuracy of the matrix element calculation is LO in QCD as for Run-I, the main changes in the signal modelling are the use of the new NNPDF3.0 set [196] that incorporates recent LHC data, as well as switching to a dedicated dynamical scale $\mu = (m_T(t) + m_t(H) + m_T(b))/6$ to better account for the event kinematics at large transferred momenta. Moreover, Pythia8 with the CUETP8M1 underlying event tune has replaced the Run-I default Pythia6 TuneZ2*. The settings and additional information on the modelling are given in Table 5.1.

Figures 5.1-5.2 show a comparison of the final state objects kinematics at both energies for $C_t = -1$.¹ As expected from the higher energy in Run-II and the larger incoming

¹Despite extending the scope of the analysis to 51 points in the 2d plane spanned by C_V and C_t , the scenario of a flipped top-Yukawa coupling with respect to the SM will be treated as *primus inter pares* throughout

5.1. Changes in signal and background modelling

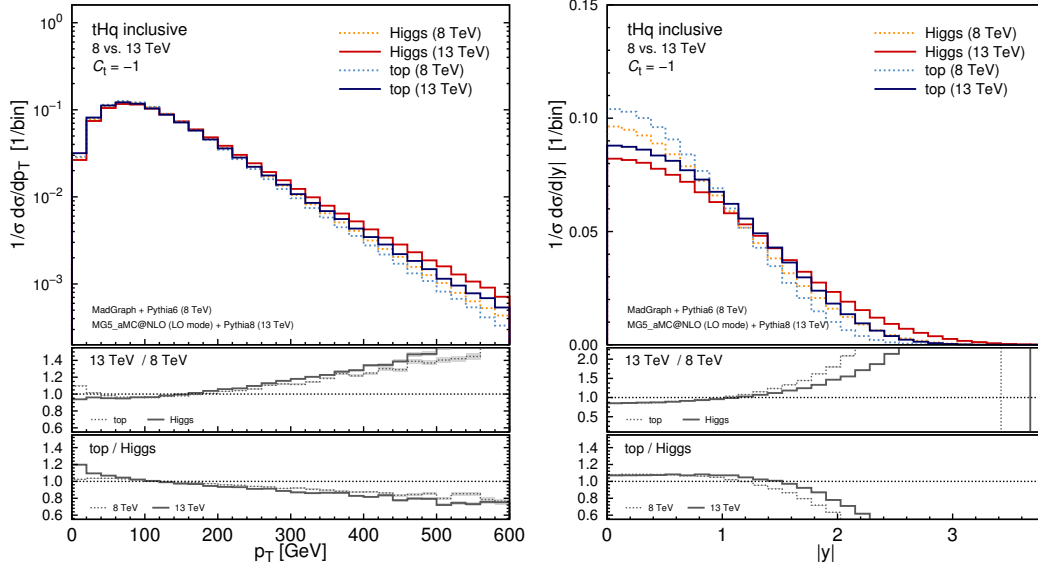


Figure 5.1.: Transverse momentum and rapidity distributions of the Higgs boson and the top quark at 8 and 13 TeV, normalized to unit area; shown are LO parton level quantities prior to the decay of the objects, including parton shower effects like ISR and FSR. The ratios 13 GeV/8 GeV and t/H are also shown.

longitudinal boost, the objects tend to lie more often in forward directions than at $\sqrt{s} = 8$ TeV and have higher p_T . This is particularly relevant for the light quark, as the pseudorapidity distribution of its associated jet is expected to be even more discriminating than at 8 TeV, where it has already been one of the most efficient variables in rejecting the backgrounds.

For the first time, the tWH process is also considered and will contribute to the signal yields. Due to its small cross section, the analysis chain however will be optimized (and BDTs trained) for the tHq process only. Fig. 5.3 contrasts the two signal processes in terms of the kinematics of the Higgs boson and the top quark. In general, the tWH process has harder p_T spectra and the objects are more central, which applies to both the Higgs boson and the top quark, and propagates to their decay products. The topology of the process resembles the one of top quark pair production (in association with a Higgs boson). Events stemming from tWH production are therefore expected to yield scores in the final BDT output which range somewhere in between the tHq signal and the bulk of $t\bar{t}$ plus $t\bar{t}H$ events.

The background processes considered are the same as for the 8 TeV analysis. The splitting of the $t\bar{t}$ background on the other hand is more fine grained, utilizing an additional category $t\bar{t} + 2b$ for events where at least 2 B hadrons are contained in the same jet. Before these were part of the $t\bar{t} + b\bar{b}$ category; now the distinction owes to the differences in the modelling of a collinear gluon splitting as opposed to two clearly separate b quarks. The

this chapter; most of the distributions are therefore only shown for $C_t = -1$ for the sake of brevity.

5. Prospects for a search for associated production of single top quark and Higgs boson at 13 TeV

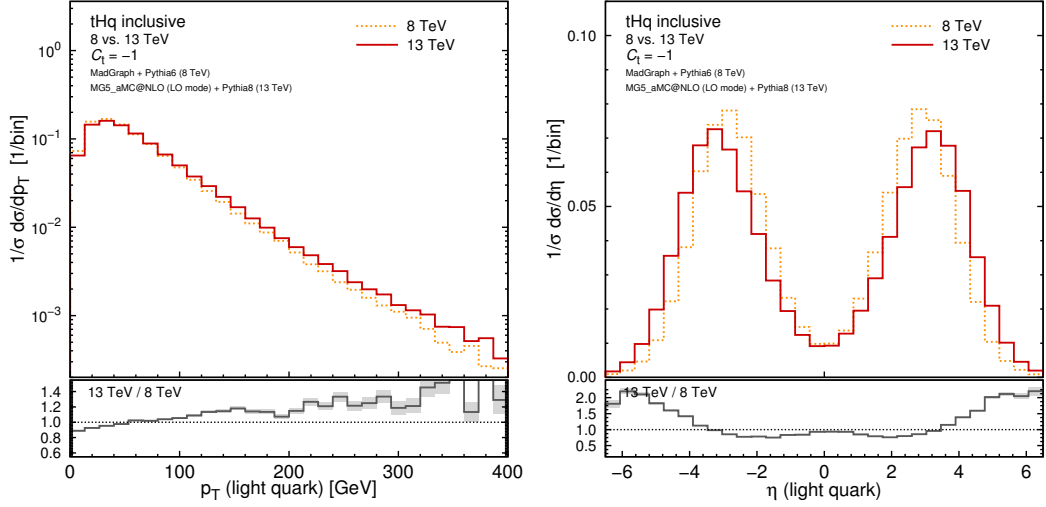


Figure 5.2.: Transverse momentum and pseudorapidity distributions of the light quark at LO parton level (no parton shower effects) at 8 and 13 TeV, normalized to unit area.

matrix element generator POWHEG is used for $t\bar{t}$, modelling the two top quarks at NLO and an additional emission at LO accuracy. The simulation of further radiation is left to the parton shower (Pythia8). This will lead to relatively large uncertainties on predictions for $t\bar{t}$ in the phase space determined by the event selection, both for shapes and the overall normalization. An alternative approach for simulating the major background employs the FxFx merging and models more jets already on matrix element level at NLO QCD. On the one side, this should help in reducing the scale uncertainties, but since this approach creates a large fraction of negative event weights, distributions are afflicted with large statistical uncertainties. Consequently, such an MC sample is not utilized in the analysis to prohibit overtraining due to statistical fluctuations and make the limit calculation more robust, until sufficient statistical power is available. A comparison of the two different $t\bar{t}$ modellings, which illustrates the previously mentioned points, is presented in Fig. 5.4, showing the ΔR difference between the jets of the hadronically decaying W boson in the 3 tag region, as this variable is sensitive to the treatment of additional jets in the simulation.

5.2. Phenomenological studies

The genuine feature of tH production is the interference between the Feynman diagrams of Fig. 1.4 (and of Fig. 1.5). Depending on the values of the couplings, it is either of constructive or destructive nature. This is illustrated in Fig. 5.5 in terms of the tHq and tWH cross sections for a center-of-mass energy of 13 TeV in the 2d plane for a range of $[0.5, 1.5]$ in C_V and $[-3.0, 3.0]$ in C_t . The exact values are given in Tab. B.1. One can deduce from the numbers that there are (C_V, C_t) configurations displaying a much larger

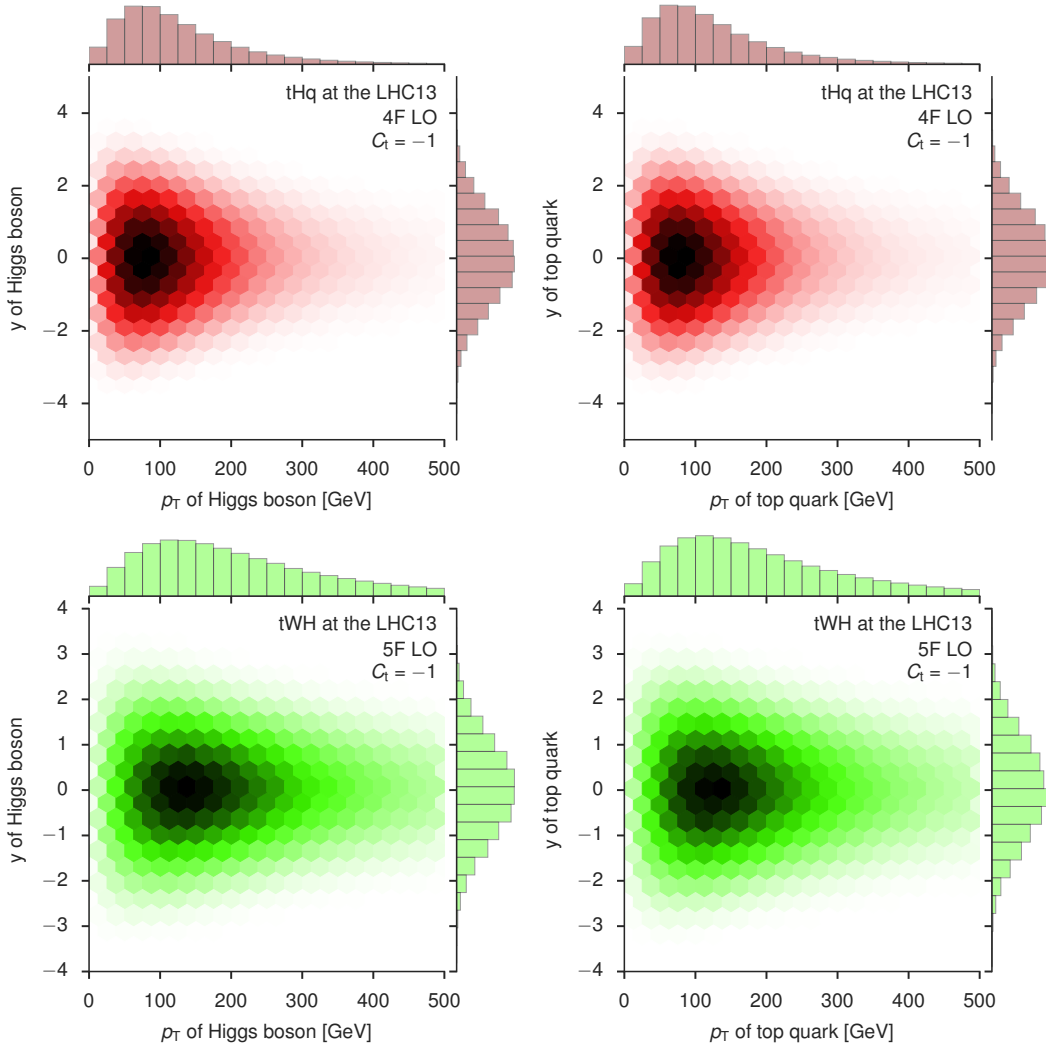


Figure 5.3.: Kinematics (p_T vs. the rapidity y) of the Higgs boson and the top quark at LO for tHq and tWH at $\sqrt{s} = 13$ TeV and for $C_t = -1$.

5. Prospects for a search for associated production of single top quark and Higgs boson at 13 TeV

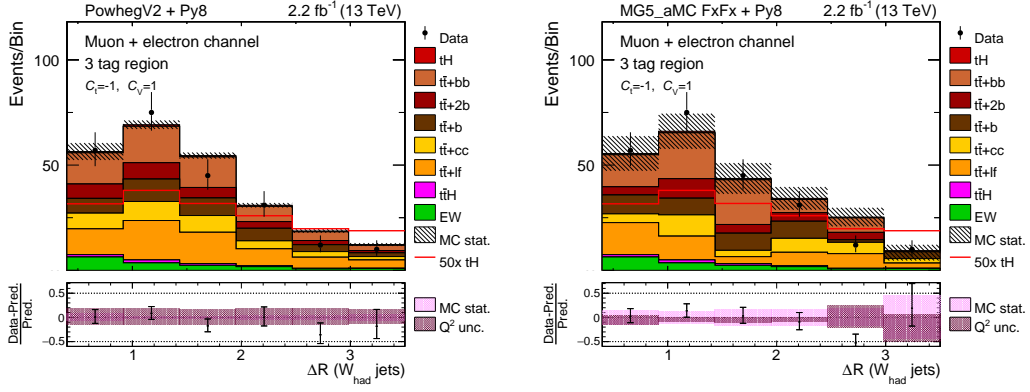


Figure 5.4.: Comparison of the $t\bar{t}$ modelling; shown is the ΔR difference between the jets assigned to the hadronically decaying W boson in the 3 tag region. Simulation is normalized to prediction. The modelling of $t\bar{t}$ processes in the right distribution incorporates more jets at matrix element level and displays reduced scale uncertainties, but suffers from low statistical power.

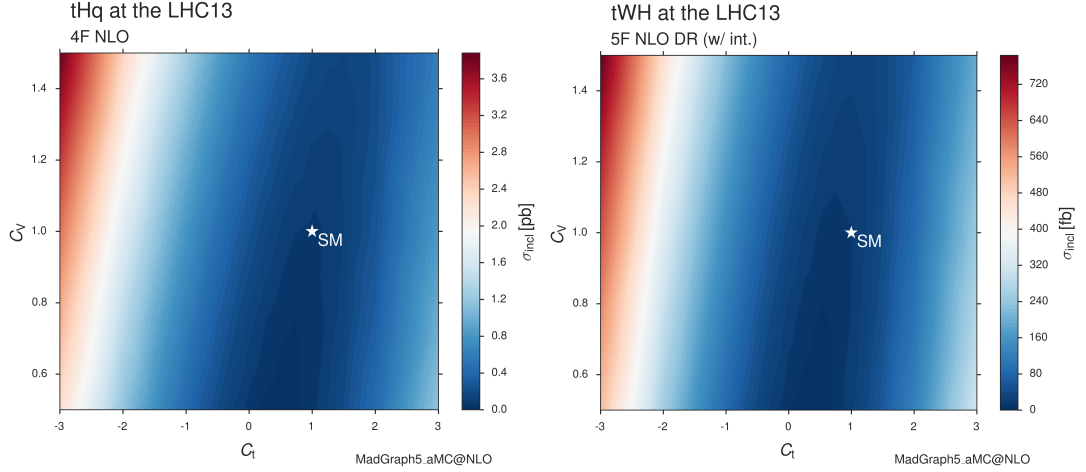


Figure 5.5.: tHq and tWH cross section in the C_V - C_t plane at 13 TeV. For tHq , the points with $C_V = C_t$ define a valley of minimal cross sections according to Eq. 1.7. For deviations from these cases and in particular from the SM, the tHq production cross section can be drastically enhanced. The tWH interference pattern looks similar.

production cross section than even the exotic scenario searched for in Run-I, providing a promising perspective for setting stringent exclusion limits on these scenarios already for low integrated luminosities. How the branching fractions of $H \rightarrow b\bar{b}$ change with C_V (a change in C_t only affects the branching fraction of $H \rightarrow \gamma\gamma$, which is so small that its effect is neglected here) is demonstrated in Appendix B.1.

Not only the total cross section is affected by a modification of the couplings. Fig. 5.6 presents the normalized differential distributions for inclusive tHq production at $\sqrt{s} = 13$ TeV for a scan in the top-Yukawa coupling and for the interaction of the Higgs boson

with the W boson fixed to SM prediction ($C_V = 1$). The more the coupling of the scenario deviates from C_t^{SM} , the harder are the p_T spectra of the top quark and the Higgs boson, and the more central are the objects. The light quark on the other hand does not display a large dependence on the value of C_t . Distributions for the additional b quark are not presented, but do not show salient differences either. This can be understood by recalling that at first order the scattering amplitude bearing the interference terms can be effectively reduced to the one of a $Wb \rightarrow tH$ process (see Eq. 1.7), in which neither the additional b quark nor the light quark play a role.

5.3. Event selection

The selection has been simplified with respect to the 8 TeV analysis by considering jets with a non-variable jet threshold of $p_T > 30$ GeV.² Muon candidates must fulfill $p_T > 26$ GeV and $|\eta| < 2.1$. For electron candidates, the threshold for the transverse momentum is $p_T > 30$ GeV; they are likewise considered up to a pseudorapidity of $|\eta| = 2.1$. Given the low expected and observed event counts, the muon and electron channel are considered as a combined lepton channel both in the presentation of distributions and in computing upper exclusion limits. The signal enriched regions require either three or four b tagged jets plus the presence of at least one untagged jet that can be assigned to the light quark in the event reconstruction. The tight working point of the CSV discriminator is employed for defining a b jet.³ The working point has changed to 0.97 as the b tagging algorithm has undergone slight modifications and a new calibration. In addition, at least three central jets are required in all events passing the selection because of b tagging considerations in the reconstruction of tHq events with at least three expected genuine b jets.

The \cancel{E}_T and lepton criteria of the 2 tag region are identical to the ones of the signal regions; the jet requirements differ from the 3 tag region only in the number of b tagged jets (= 2). In Fig. 5.7-5.8 key variables in this control region are presented, displaying overall good agreement between data and prediction. As becomes clear from Fig. 5.13 showing the pseudorapidity assigned to the light quark, a severe mismodelling affects the description of forward jets. In fact, events of both simulation and data face an issue in the reconstruction in the hadron forward calorimeter. This is understood and has been corrected in a re-reconstruction. Samples where this problem is resolved however have not yet been ready by the time this thesis has been written. The jet multiplicity distribution reveals a harder spectrum in simulation, i.e. fewer hard jets (with a p_T greater than 30 GeV) are observed in data than in simulation. This issue is present for different

²This anticipates improved jet energy corrections for forward jets with respect to 8 TeV. As will be shown later, the jet kinematics for $|\eta| > 2.4$ is affected by a reconstruction issue. Hence, no definite statement about the data-MC agreement of jet variables in this region can currently be made, but the chosen threshold of 30 GeV seems a reasonable value in the long term after the problem with the reconstruction has been fixed.

³In order to increase the overall event yields and the statistical power of the samples, relaxing the cut on the CSV discriminator is worth considering. An optimization of the cut however is beyond the scope of this thesis, which aims at applying the concept of the 8 TeV analysis to 13 TeV data in a manner as straightforward as possible.

5. Prospects for a search for associated production of single top quark and Higgs boson at 13 TeV

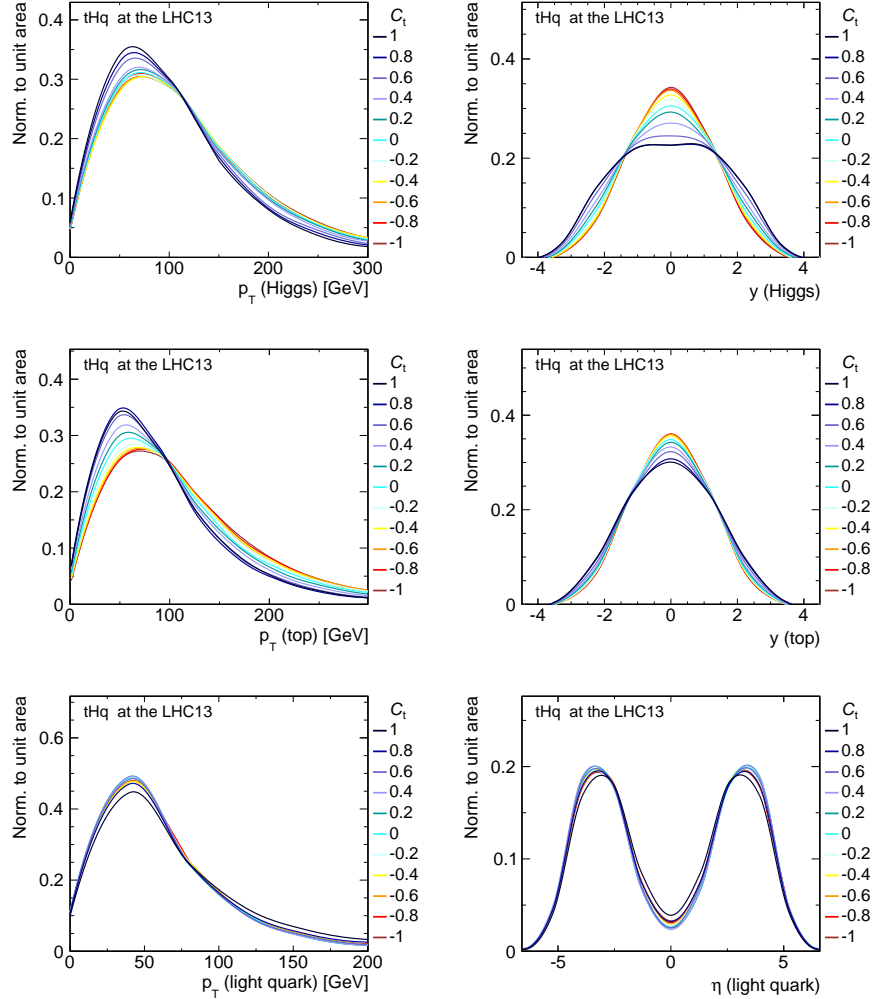


Figure 5.6.: Normalized simulated distributions of objects kinematics for the tHq process at NLO in QCD for $\sqrt{s} = 13$ TeV, obtained with a MadGraph5_aMC@NLO + Pythia8 setup. While the light quark kinematics does not show a significant dependence on C_t , both the Higgs boson and the top quark are affected by deviations from $C_t = +1$. The objects become more central and distributions display a harder p_T spectrum. The Higgs boson and top quark mass are set to 125 and 172.5 GeV, respectively.

Table 5.2.: Requirements for an event to fall into either the 3 tag or 4 tag region. Jets are considered above 30 GeV in p_T and up to $|\eta| = 4.7$.

	3 tag region	4 tag region
# tight leptons		1
# add. loose leptons		0
\cancel{E}_T		$> 35/45 \text{ GeV}(\mu/e)$
# jets	≥ 4	≥ 5
# central jets	≥ 3	≥ 4
# jets with CSV > 0.97	$= 3$	$= 4$

types of $t\bar{t}$ modelling. Unlike the problem with the HF calorimeter response, the reasons for this discrepancy are suspected to be hidden in the interfacing of the matrix element generator to the parton shower, and are not expected to be resolved until some progress on the phenomenological side takes place.

Corrections to simulation

The p_T of the W boson candidate in Fig. 5.7, which is reconstructed from the lepton and the missing transverse energy, does not possess a significant slope, but shows the same tendency as in Run-I. In absence of correction factors, which have not yet been determined by the collaboration at 13 TeV, no top p_T reweighting is applied. However, this happens on the understanding that the residual shape difference is covered almost by statistical uncertainties alone.

The distribution of the number of primary vertex candidates, being strongly correlated to the amount of pile-up, undergoes a reweighting in order to match the profile of the simulations to the one observed in data.

The CSV reweighting introduced in the last chapter corrects the entire shape of the b tag discriminator and allows to use the full shape instead of sticking to dedicated working points of the algorithm. Fig. 5.10 displays the improvement in the agreement of CSV shapes in data and simulation.

5.4. Event reconstruction

The same type of MVA based event reconstruction is performed as in the Run-I analysis. Parton level objects are matched to reconstructed jets, yielding a “correct” hypothesis in the case of a successful match; a “wrong” hypothesis is chosen randomly for each event,

5. Prospects for a search for associated production of single top quark and Higgs boson at 13 TeV

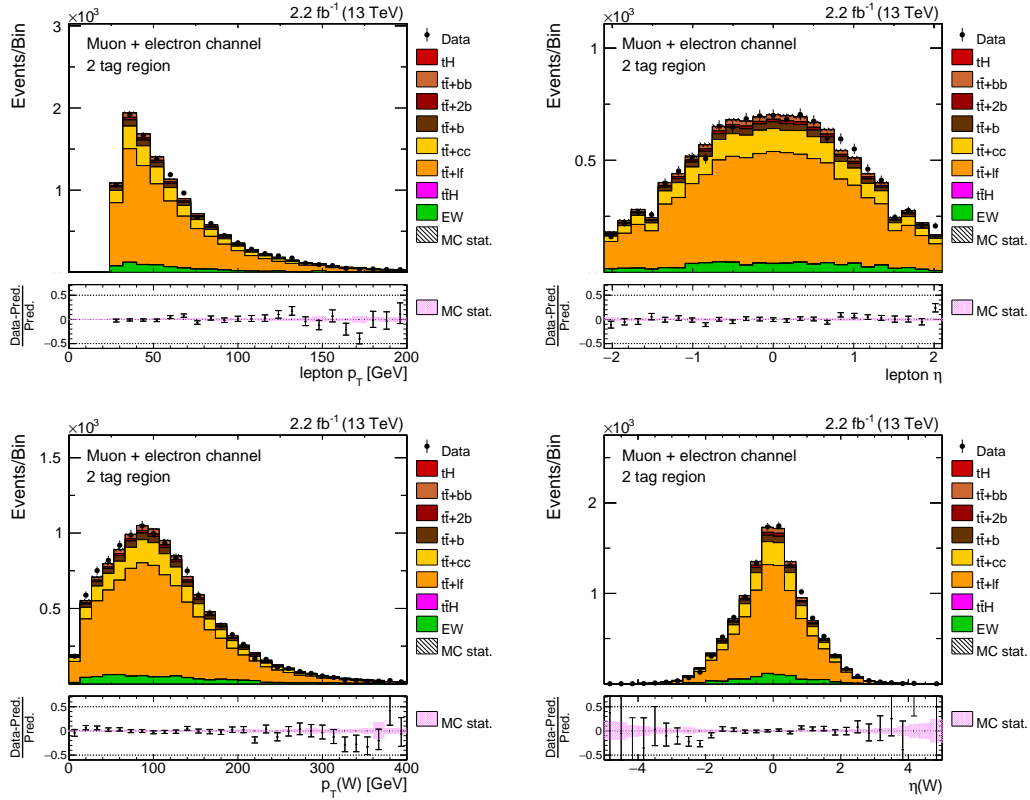


Figure 5.7.: Kinematics of the charged lepton candidate and the reconstructed W boson in the 2 tag region. Simulation is normalized to data in order to facilitate the shape comparison. Good agreement of the shapes is observed.

and a BDT is trained to distinguish between correct and wrong event interpretations. The variable sets of Tab. 4.2 and 4.3, respectively, are modified slightly. In both cases, the boolean variables on the b tag status of a reconstructed jet (or e.g. the number of b tags found among the jets assigned with the Higgs boson candidate) are replaced by the shape of the continuous CSV discriminator, which is supposed to provide maximal information on the b tagging. Four more variables are added to the input variable list for the 51 tHq reconstructions:

- the difference in pseudorapidity of the light jet and the jet identified with the b quark from the top quark decay,
- their difference in energy,
- the pseudorapidity of the jet identified with the b quark from the top quark decay,
- the pseudorapidity of the reconstructed top quark.

5.4. Event reconstruction

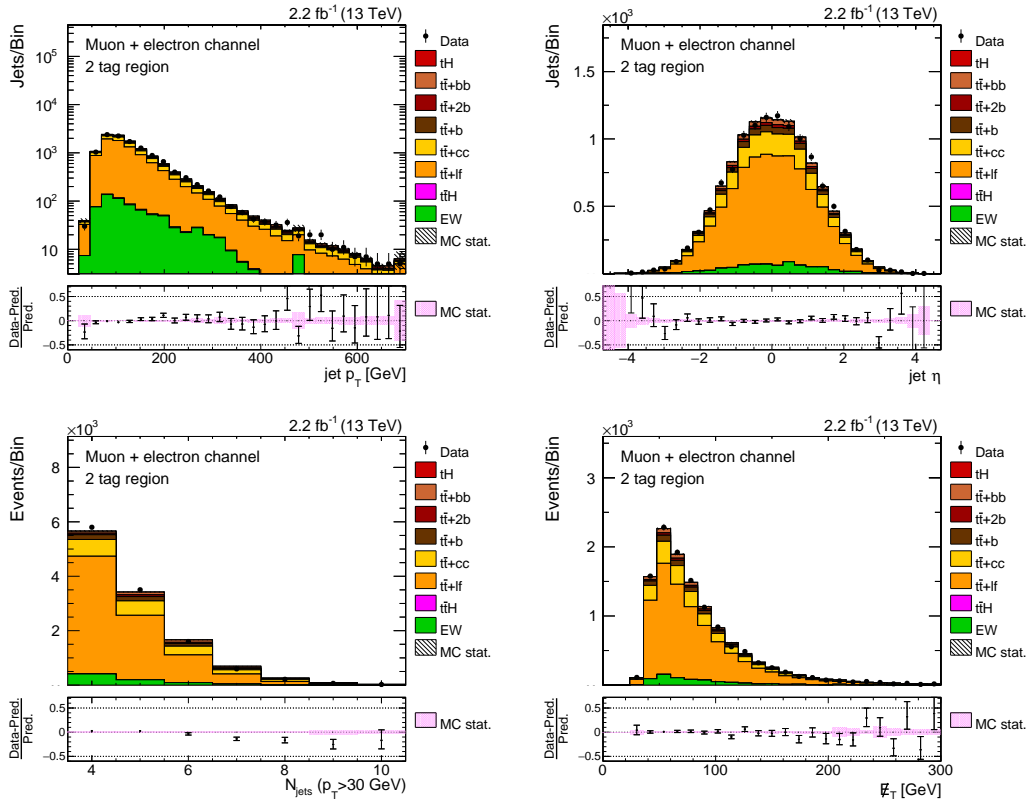


Figure 5.8.: Jet kinematics and E_T in the 2 tag region. Simulation is normalized to data in order to facilitate the shape comparison. Good agreement of the shapes is observed.

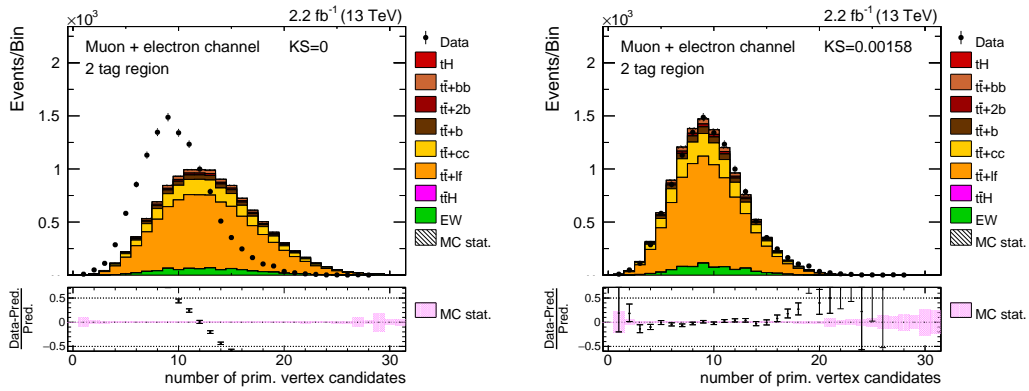


Figure 5.9.: The number of primary vertices before (left) and after (right) the reweighting of all simulations to the true distribution found in data. Simulation is normalized to data in order to facilitate the shape comparison.

5. Prospects for a search for associated production of single top quark and Higgs boson at 13 TeV

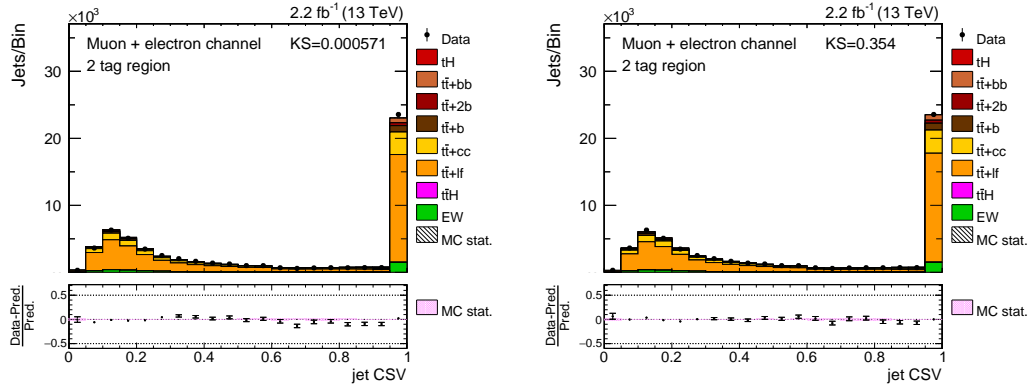


Figure 5.10.: CSV values of all considered jets entering the selection before (left) and after (right) the CSV reweighting. MC is normalized to data in order to facilitate the shape comparison. A clear improvement in terms of data/MC agreement is observed when employing dedicated scale factors derived from control samples.

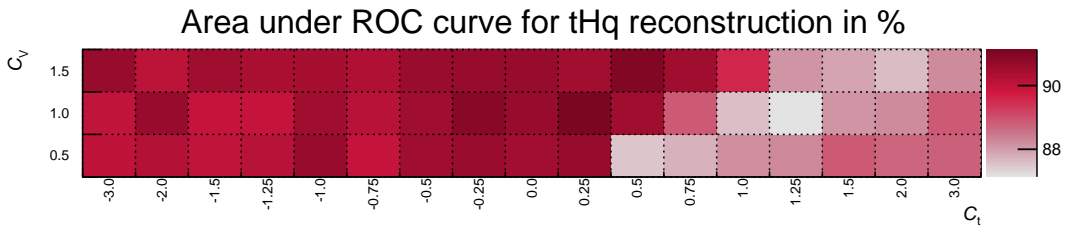


Figure 5.11.: Shown is the area under the ROC curve for the 51 reconstruction BDTs trained under the different (C_t, C_V) tHq hypotheses.

No variables have been added in the training for reconstructing the event under the assumption it has been a top quark pair event in the semi-leptonic decay channel.

It is visible from Fig. 5.11 that the reconstruction BDTs under the tHq hypothesis perform well and the area under the ROC curve varies between 87-91%. The residual differences might be attributed to the increased centrality of the Higgs boson and the top quark for negative C_t values and the therefore better reconstructability of the event, but further studies will be needed for verification of this assumption.

The validation of the reconstruction training is performed by comparing the compatibility of the distributions of BDT scores obtained with the training sample and a statistically independent test sample. The top row distributions of Fig. 5.12 reveal no signs of overtraining. In the bottom row, the data-MC distributions showing the largest BDT output that is finally picked in the application of the reconstruction BDTs are presented and show good agreement in shapes.

5.5. Event classification

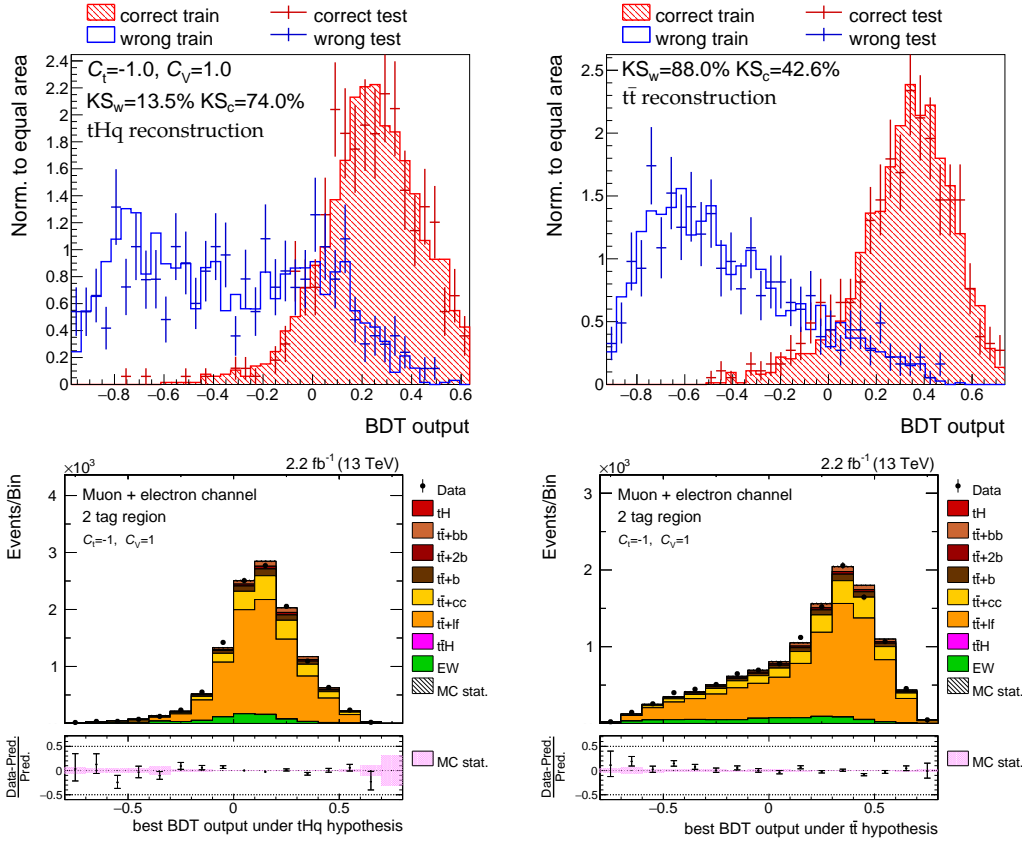


Figure 5.12.: Validation of the reconstruction BDTs at 13 TeV. Upper row: distributions comparing the reconstruction BDT outputs for the training and a statistically independent test sample do not indicate signs of overtraining. This is also affirmed by the lower row distributions revealing very good data-MC shape agreement in the best reconstruction BDT score under the signal and the $t\bar{t}$ hypotheses in the 2 tag region. Simulation is normalized to data.

5.5. Event classification

For isolating the signal from the background, in total 51 BDTs are trained for the different coupling scenarios. The input variable set (compare Tab. 4.4) likewise has undergone the same modification as in the reconstruction BDTs. Instead of counting the b tags among the jets assigned to the decay products of the Higgs boson and the hadronically decaying W boson, the whole CSV discriminator shape of the jets is used directly as input. This is possible since the CSV reweighting corrects for shape differences between data and simulation over the entire range of possible values. The jet pseudorapidity is not coarsely binned anymore, and the BDT can make use of the full discriminating potential of the $|\eta|$ distribution of the light jet, which turns out to be the most important variable in the training.

In Fig. 5.13-5.18 the input variables of the classification BDT are shown for the three

5. Prospects for a search for associated production of single top quark and Higgs boson at 13 TeV

analysis bins. For the tHq reconstruction dependent variables, the representative case of ($C_t = -1$, $C_V = 1$) has been chosen. Except for the light jet kinematics in the 2 tag region, a more than satisfactory data-MC agreement is observed in all the distributions.

In Fig. 5.19, the validation distributions for checking against overtraining and for good data-MC agreement in the 2 tag region are presented. They raise trust in applying the BDT in the signal regions for maximizing the analysis sensitivity. However, the 51 signal scenarios do not always allow for the same level of BDT performance in rejecting background processes. Fig. 5.20 shows the area under the ROC curve for the 51 classification BDTs and suggests the BDT is able to discriminate between tHq and the background processes in a more efficient way for scenarios with a small absolute value of C_t . For the benchmark case of $C_t = -1$ (and $C_V = 1$), the BDT output shape for both signal and background processes are shown in Fig. 5.21. Although the tWH process shares the same final state as $t\bar{t}H$, the BDT seems to be able to distinguish between them to some extent, with BDT scores for tWH ranging in between the bulk of the background – mostly $t\bar{t} + hf/lf$ – and the tHq signal process. This implies a certain correlation of one or more input variables with the invariant masses of the reconstructed top quarks, which are the variables providing the clearest separation power between doubly- and singly-resonant processes with a $bW^+\bar{b}W^-H$ final state. Another reason is that for $C_t = -1$, the kinematics for tWH differs more from the one of $t\bar{t}H$, which can be seen by comparing Fig. 5.21 and Fig. B.1. By including tWH events in the training signal sample or by introducing a dedicated BDT for the isolation of the tWH contribution, an improvement in terms of signal-vs.-background separation is expected to be achieved. As it lies beyond the scope of the thesis presented here, this is deferred to future studies.

5.6. Systematic uncertainties

A rich set of systematic uncertainties is considered in the statistical inference at 13 TeV, resembling the one of the 8 TeV analysis in most of its aspects. The list incorporates shape variations of the BDT output due to uncertainties on the Q^2 scale, the jet energy and the performance of the b tagging algorithm. A rate uncertainty of 4.5% reflects the imperfect knowledge about the exact value for the integrated luminosity. PDF uncertainties on the cross sections of the processes considered are taken to have the same size as in Run-I.

A subtle change has happened in the treatment of the aforementioned Q^2 scale uncertainties. In the 8 TeV analysis, inclusive distributions corresponding to predictions with a systematically shifted scale have been renormalized to match the integral of the one with nominal scale choice. Yield differences in the final templates can then only arise from acceptance differences due the influence of the scale on shapes. On top of the differences due to the shapes, a rate uncertainty corresponding to the scale uncertainty of the most accurate cross section calculation for the respective process has been used, implying that rate effects of a shifted scale are fully correlated across the phase space.

While there is still no definite prescription from the theory side as to how to incorporate scale uncertainties in a fiducial phase space, the tendency is now to consider also the normalization effects on inclusive distributions when using a shifted scale, and

5.6. Systematic uncertainties

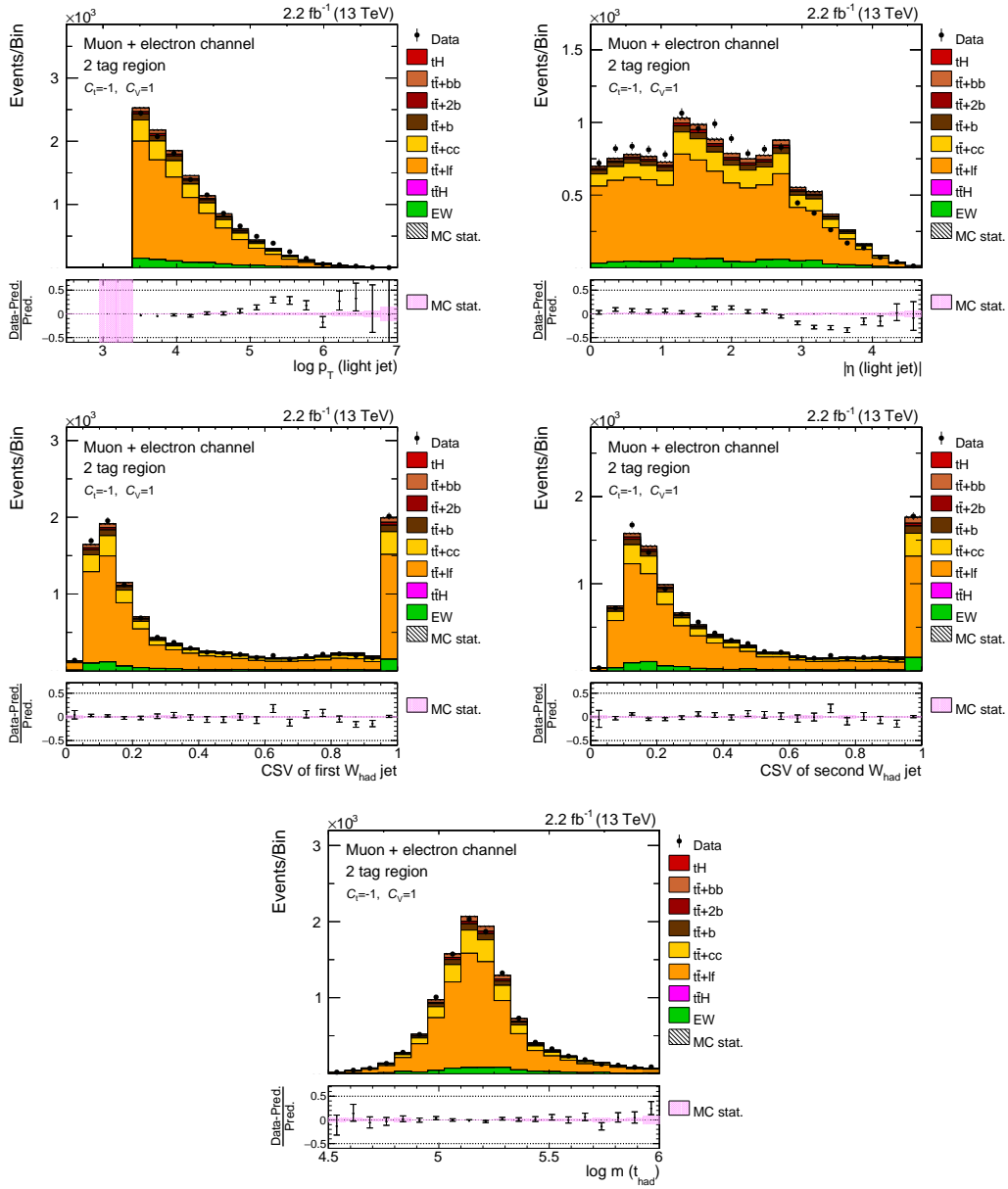


Figure 5.13.: Distributions of the input variables for the classification BDT in the muon + electron channel in the 2 tag region. Simulation is normalized to data in order to facilitate the shape comparison.

5. Prospects for a search for associated production of single top quark and Higgs boson at 13 TeV

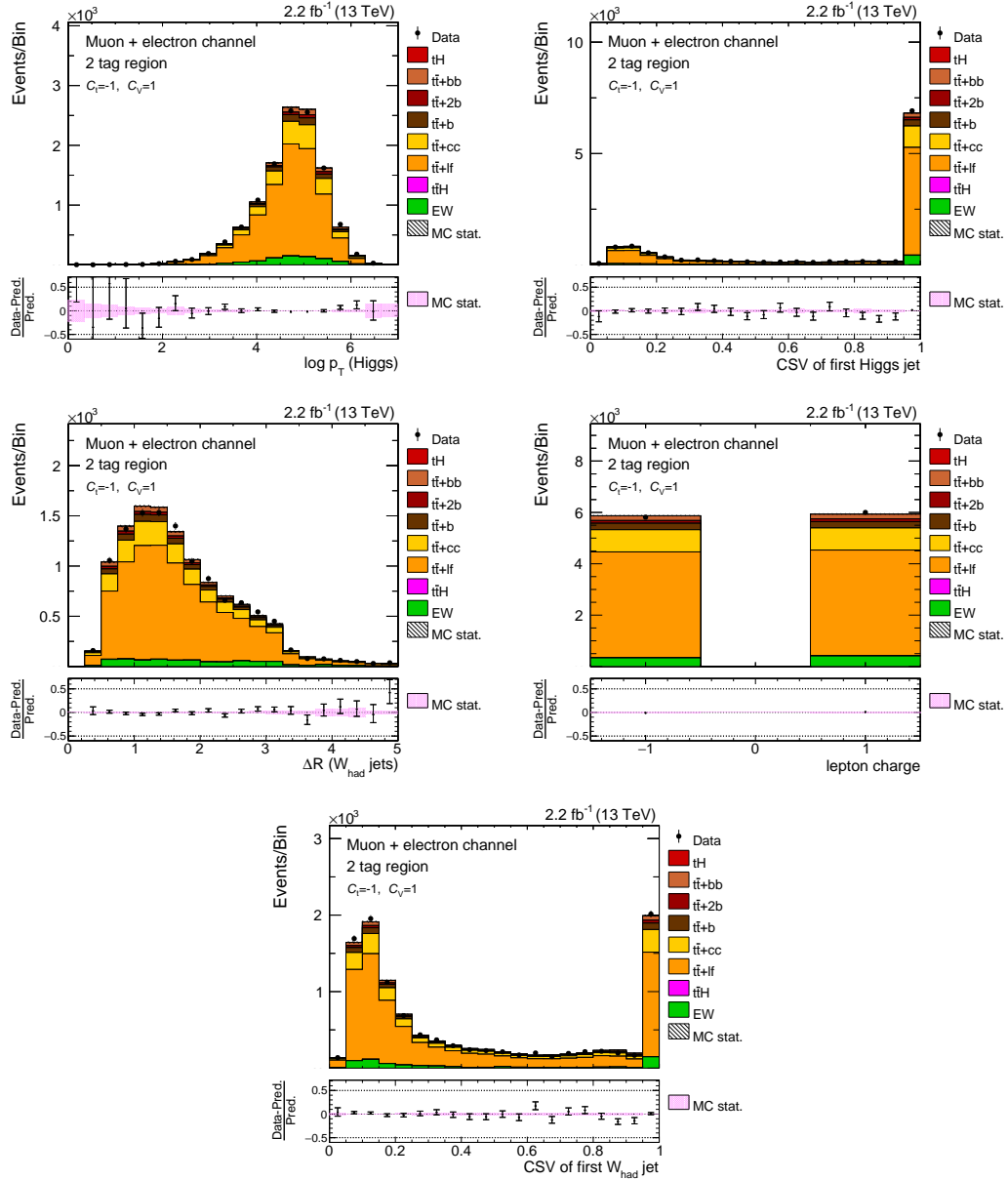


Figure 5.14.: Distributions of the input variables for the classification BDT in the muon + electron channel in the 2 tag region. Simulation is normalized to data in order to facilitate the shape comparison.

5.6. Systematic uncertainties

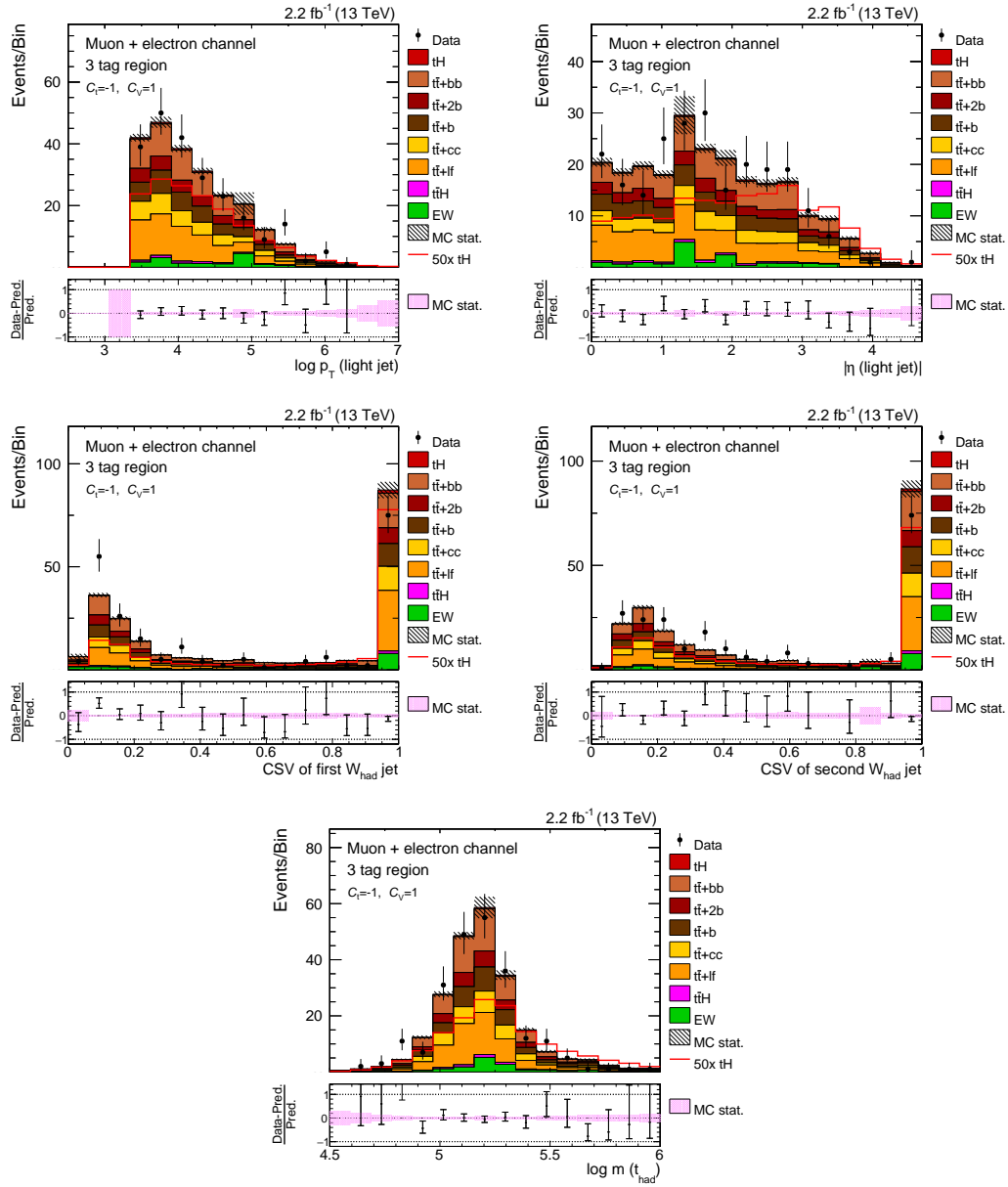


Figure 5.15.: Distributions of the input variables for the classification BDT in the muon + electron channel in the 3 tag region. Simulation is normalized to data in order to facilitate the shape comparison.

5. Prospects for a search for associated production of single top quark and Higgs boson at 13 TeV

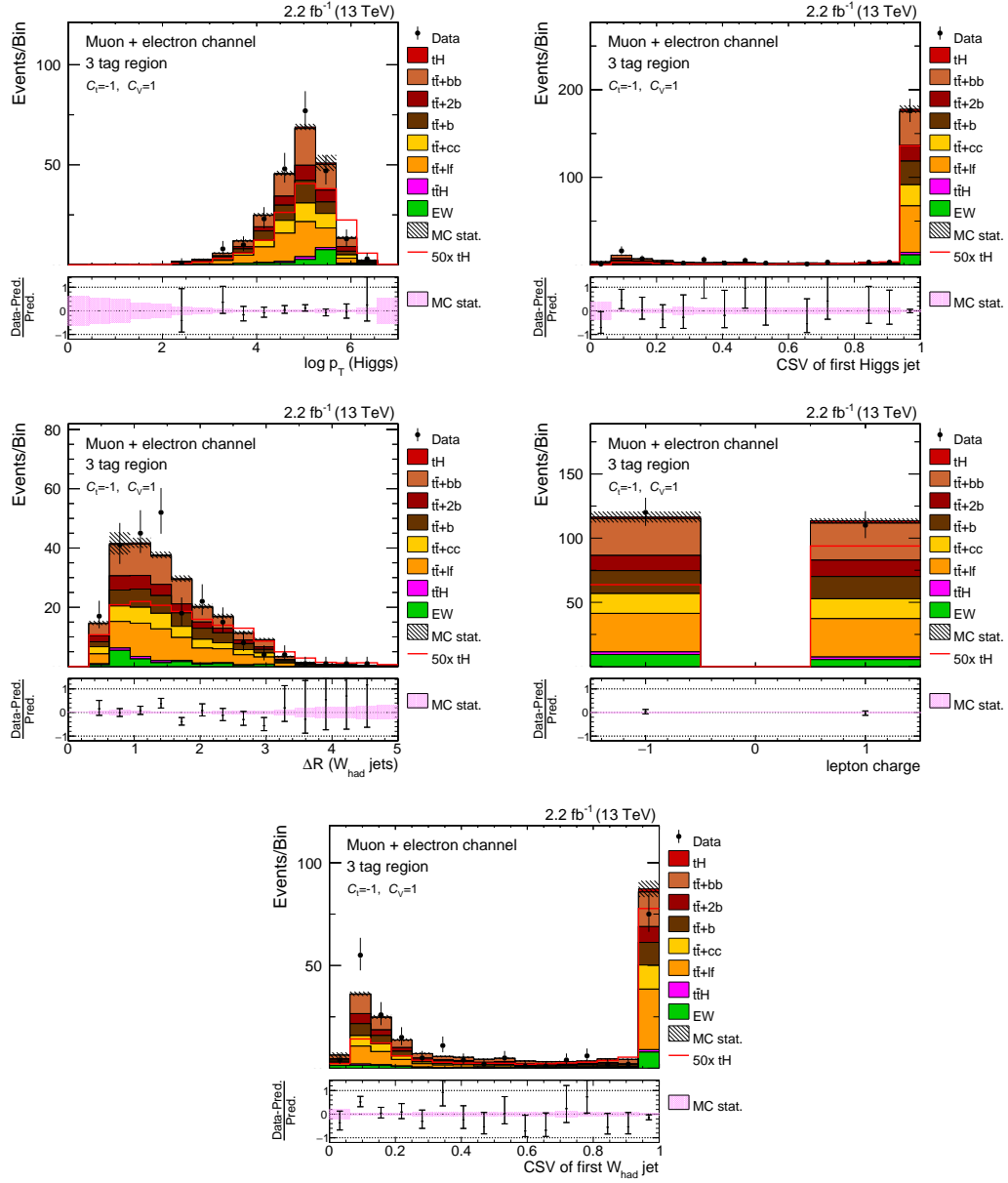


Figure 5.16.: Distributions of the input variables for the classification BDT in the muon + electron channel in the 3 tag region. Simulation is normalized to data in order to facilitate the shape comparison.

5.6. Systematic uncertainties

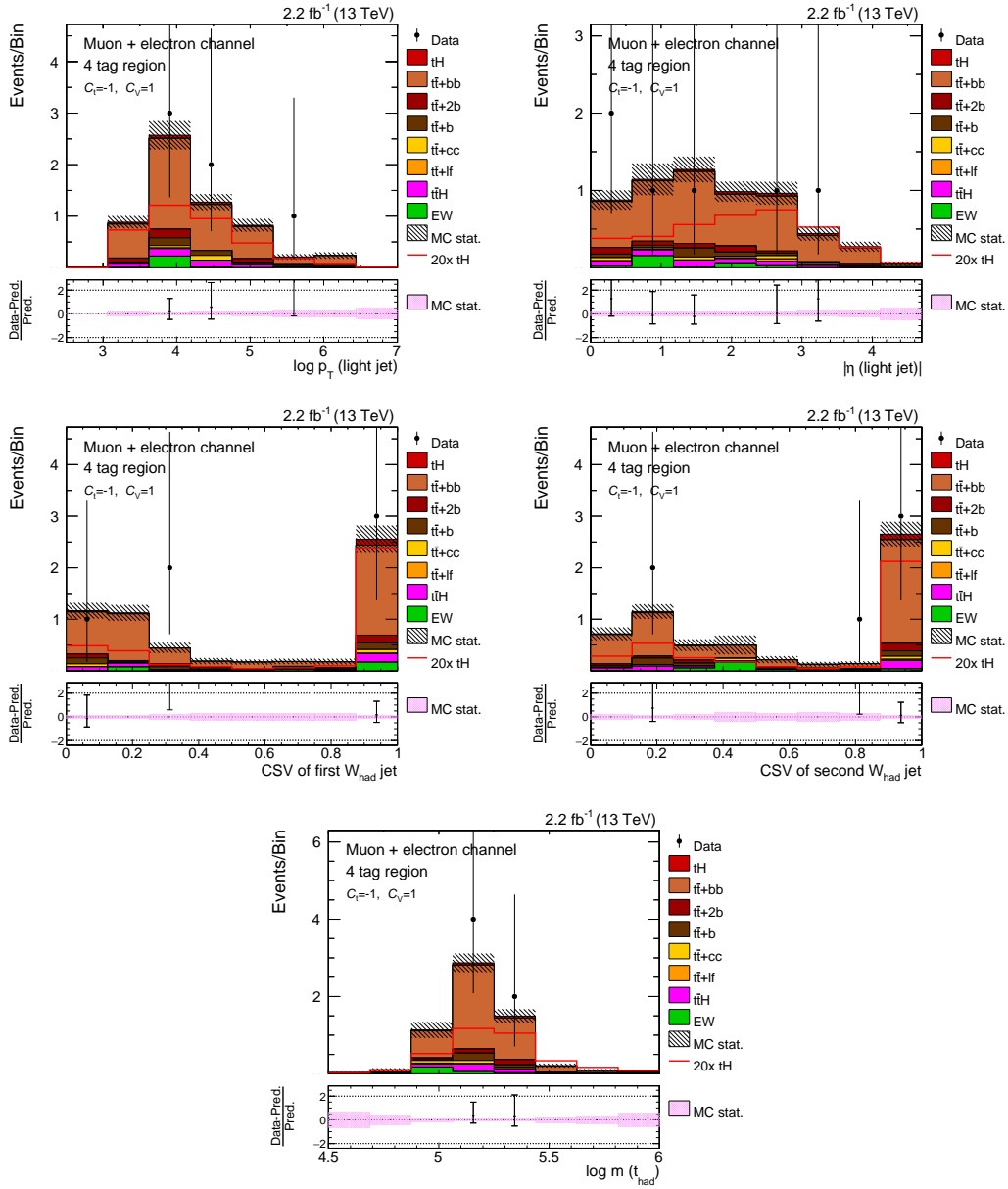


Figure 5.17.: Distributions of the input variables for the classification BDT in the muon + electron channel in the 4 tag region. Simulation is normalized to data in order to facilitate the shape comparison.

5. Prospects for a search for associated production of single top quark and Higgs boson at 13 TeV

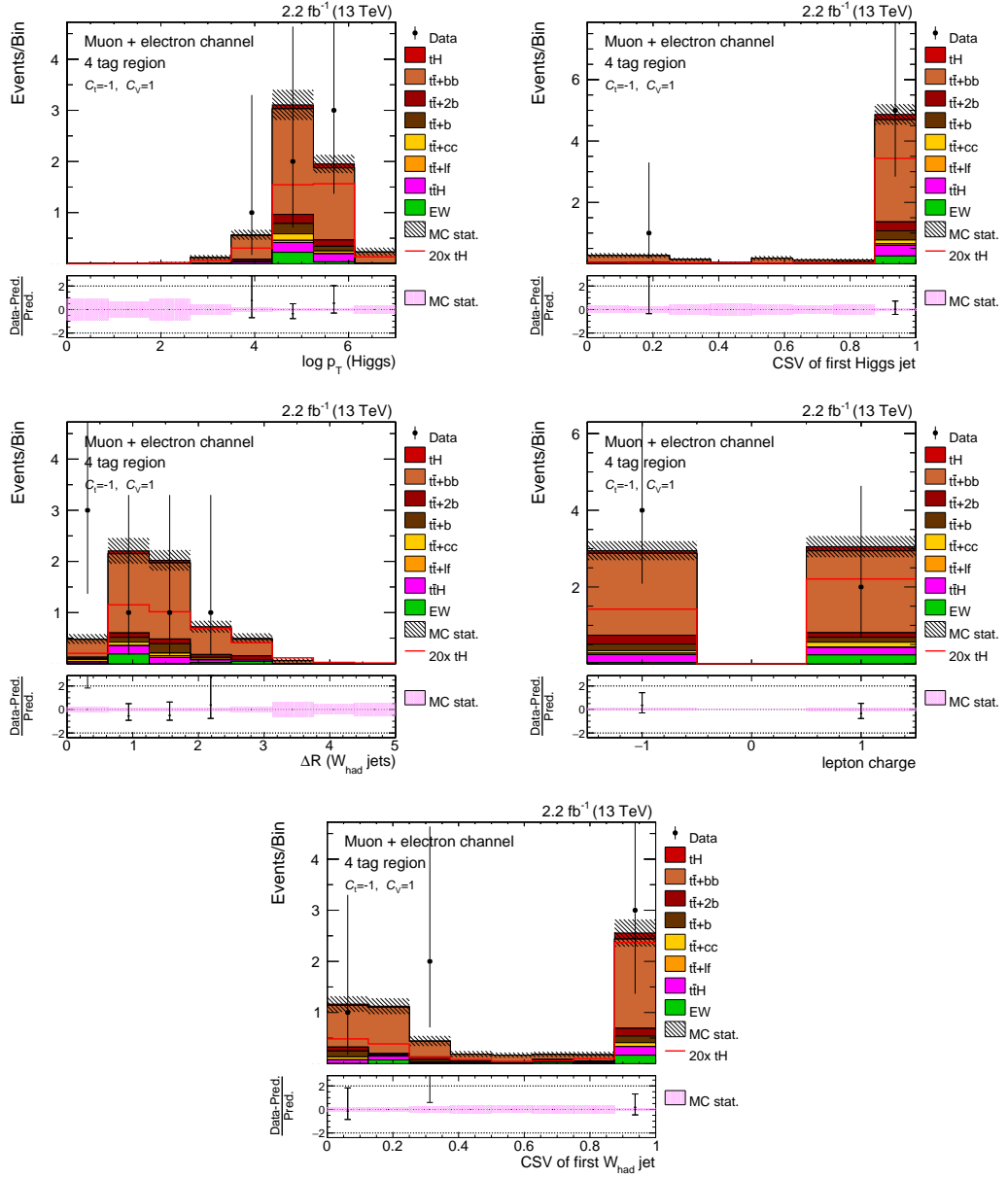


Figure 5.18.: Distributions of the input variables for the classification BDT in the muon + electron channel in the 4 tag region. Simulation is normalized to data in order to facilitate the shape comparison.

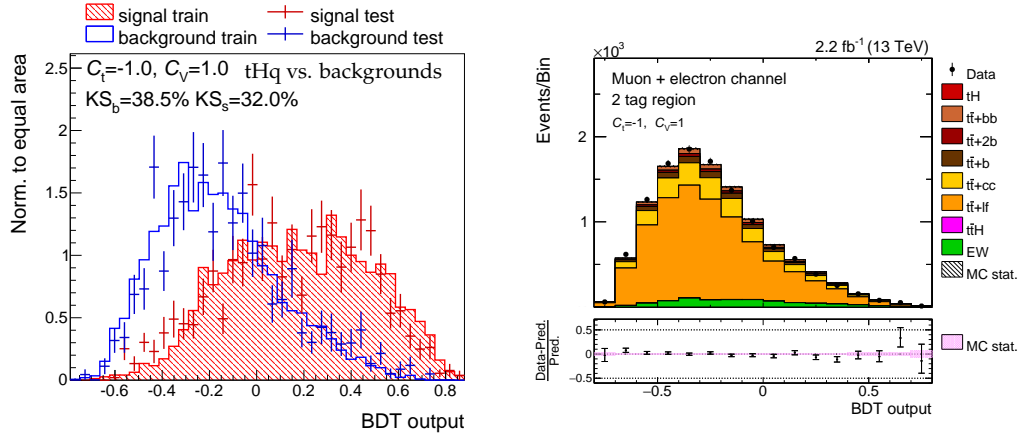


Figure 5.19.: Left: training results for the classification BDT and comparison of shapes between the training sample and a test sample. The KS values do not indicate the presence of overtraining. Right: BDT output in the 2 tag region. Simulation is normalized to data in order to facilitate the shape comparison; very good agreement is observed.

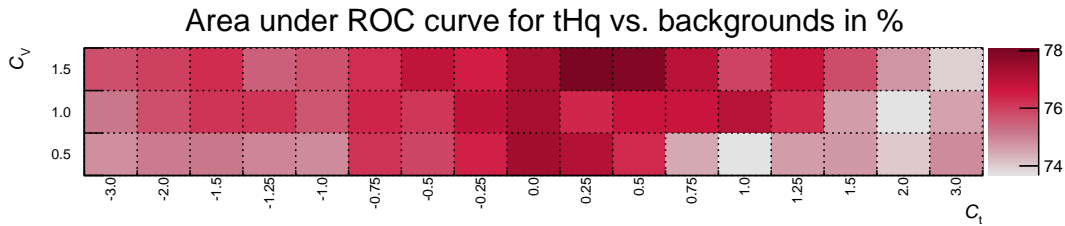


Figure 5.20.: Shown is the area under the ROC curve for the 51 classification BDTs trained under the different (C_t, C_v) tHq hypotheses. For C_t values around 0, the BDT is more able to reject background processes.

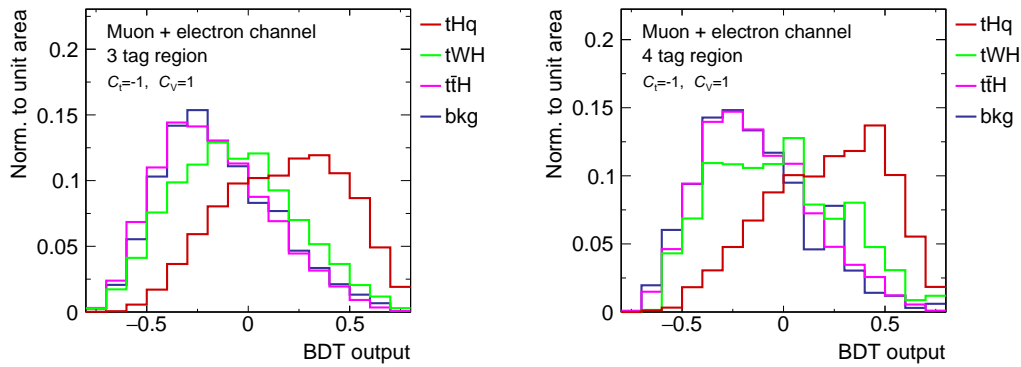


Figure 5.21.: Shapes of the classification BDT output for tHq, tWH, $t\bar{t}H$ and the remaining background processes. As expected, tWH events yield scores in the BDT that tend to lie in between the tHq and background shapes, because they have the same final state as the $t\bar{t}H$ process, which is included as background in the training.

5. Prospects for a search for associated production of single top quark and Higgs boson at 13 TeV

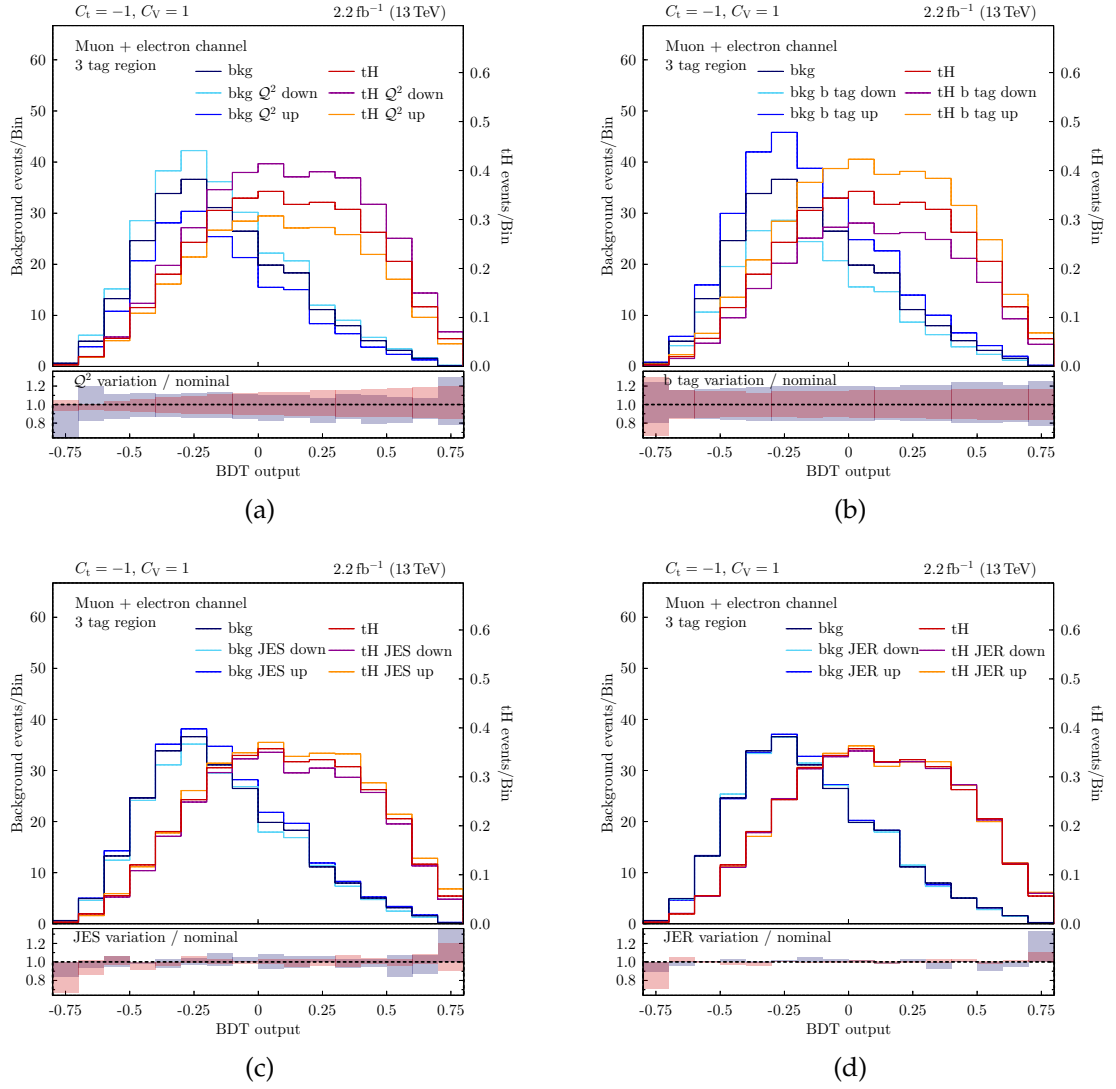


Figure 5.22.: Shape variations due to uncertainties on the Q^2 scale (a), b tagging (b), jet energy scale (c) and jet energy resolution (d).

propagate them into the fiducial phase space, which in general will lead to larger Q^2 uncertainties. However, it can be seen from Fig. 4.7 that not considering the uncertainty on the normalization of an inclusive NLO prediction would leave uncertainty bands that do not include the NNLO prediction and thus can potentially lead to a misinterpretation of the Q^2 uncertainty representing missing higher order contributions. The effect of the Q^2 uncertainties on the BDT output distributions can be seen in Fig. 5.22(a). Notably, the uncertainties increase towards larger BDT output values in the case of the signal, but not for the background (which is mainly constituted by $t\bar{t} + hf/lf$). This is an interesting observation, suggesting that a very important variable in the BDT is sensitive to the Q^2 scale. Indeed, the second most important variable – the invariant mass of the reconstructed hadronically decaying top quark stemming from the $t\bar{t}$ reconstruction – relies on the modelling of additional light flavor or gluon emissions in the parton shower for the signal process. While the b quarks from the top quark decays are mostly assigned to b jets, the decay objects of the hadronically decaying W boson are identified with light flavor jets most of the time. However, there is only one genuine light flavor jet in the tHq process, which, because of its kinematics, is not a perfect candidate for central light flavor jets as expected in the $t\bar{t}$ process. Therefore, the BDT will choose jets added by the parton shower, resulting in large Q^2 uncertainties on this observable that is strongly correlated to the final classification BDT output. This will impair the sensitivity of the analysis, which is particularly driven by the signal enriched bins of the BDT output distribution. At the same time a possibility to make the analysis more sensitive by employing an NLO simulation for the signal process, by which additional radiation would come from the matrix element calculation and not from the parton shower, is opened up.

Similarly large uncertainties are introduced due to the CSV reweighting (Fig. 5.22(b)). In the derivation of scale factors for heavy flavor (light flavor) jets, background contributions from the light (heavy) component are subtracted in dedicated control regions. The size of this component is given a 20% rate uncertainty, translating into systematically varied scale factors. Another uncorrelated source is of statistical nature and addresses the parametrization with first and second order polynomials in the scale factor evaluation. More details on the procedure of the CSV reweighting and its uncertainties are given in [197].

The uncertainty of the jet energy scale is evaluated by varying the energy scale correction of the jets by one standard deviation. Selection cuts are imposed again and BDTs are reevaluated, thereby propagating the effect of the shifted the jet energy scale and resolution to the final BDT output distributions which can have both shape differences as well as different overall normalizations due to modified acceptances. The variations are presented in Fig. 5.22(c)-(d).

Nuisance parameters reflecting the specific source of uncertainty are introduced into the statistical model. Their pulls and uncertainties after an $s + b$ fit to an Asimov dataset thrown from background-only expectation are shown in Fig. B.2, indicating that the prior rate uncertainties of 50% on the normalization of $t\bar{t} + hf$ processes might be slightly too conservative.

In general, as can be seen from Fig. 5.23, the impact of uncertainties on the limit does

5. Prospects for a search for associated production of single top quark and Higgs boson at 13 TeV

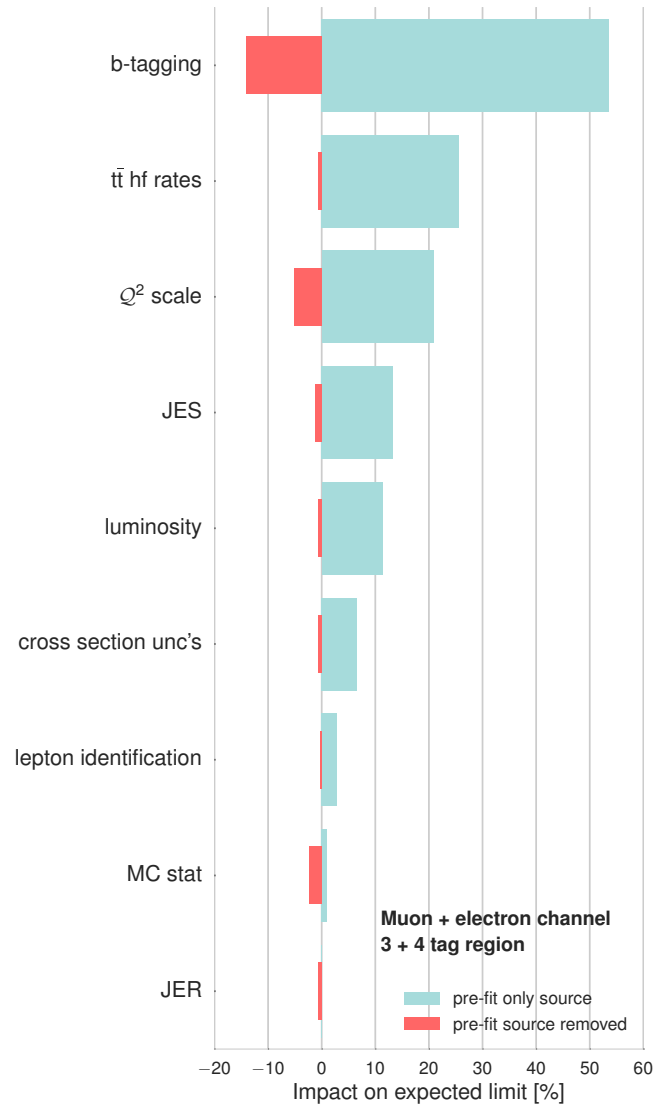


Figure 5.23.: Impact of the nuisance parameters on the final exclusion limit.

not change much compared to the 8 TeV analysis. Uncertainties on the identification of b jets are the dominant systematic source due to the complexity of the CSV reweighting procedure. The next most important uncertainties are related to the modelling of the $t\bar{t}$ process and the Q^2 scale.

5.7. Results

Since the analysis is performed in a blind way, no data is touched in calculating the expected event yields and deriving the expected upper limits for an integrated luminosity of 2.2 fb^{-1} . However, as shape comparisons of the kinematics of the light jet suggest, there is a problem in the modelling of forward jets both in data and MC, which is particularly pronounced in the 2 tag region. Given the importance of this η region for the correct reconstruction of the most characteristic feature of the signal process, a good description of objects kinematics in the forward region is essential for deriving meaningful results. For these reasons, albeit studies at 8 TeV have shown that the inclusion of this control region can make the analysis more sensitive, it is abstained from utilizing 2 tag region events at all for the final limit calculation, leaving this possible improvement to later stages after re-reconstructed MC and data samples have become available.⁴

The final classification BDT outputs in the 2 tag region and the two signal regions are shown in Fig. 5.24. Simulation is normalized to prediction and is not fit to data, which is only shown for the signal-depleted bins of the distributions in the signal region. Shaded bands on the templates and in the ratio show the effect of all shape uncertainties (CSV reweighting, Q^2 scale, jet energy scale and resolution) on the BDT output distribution. It well covers the presented data points. The templates in the 3 tag and 4 tag regions are employed to compute the sensitivity of the analysis by setting expected upper limits on the allowed signal cross section for $C_t = -1$, and similarly for the other scenarios, whose distributions are not presented here.

Exclusion limits are derived with the asymptotic CL_S method, and are given for the SM case and the scenario with a flipped top-Yukawa coupling in Tab. 5.4. The latter scenario is expected to be excluded with a cross section larger than $11 \times \sigma_{C_t=-1}$. The expected tH rates with SM-like kinematics would have to be 182 times larger than $\sigma_{\text{SM}} \times BR(H \rightarrow b\bar{b})$, i.e. larger than $182 \times (0.072 + 0.016) \text{ pb} \times 0.58 = 9.28 \text{ pb}$ so that the analysis could exclude it at 95% C.L. In this calculation, 0.072 pb (0.016 pb) is the inclusive cross section of tHq (tWH) and 0.58 is the branching fraction of $H \rightarrow b\bar{b}$.

The main pad of Fig. 5.25 presents the asymptotic expected upper exclusion limits on $\sigma/\sigma_{\text{theor.}}$ for all 51 points in the C_t - C_V plane. A cubic spline fit has been used to interpolate between the scenarios and to obtain smooth curves for the expected limits and the $\pm 1\sigma$ uncertainty bands. The analysis is least sensitive to the scenarios for which the cross section (dashed lines) has a minimum, making it hardest to distinguish between the

⁴In fact, because this is a blind analysis, predictions are to be fit to pseudo data. Since the real benefit of including the 2 tag region in the limit calculation consists in being able to better constrain the background in the fit to real data, omitting the 2 tag region in the statistical inference is not expected to worsen the sensitivity of the analysis to a significant extent.

5. Prospects for a search for associated production of single top quark and Higgs boson at 13 TeV

Table 5.3.: Expected event yields and combined (statistical plus systematic) uncertainties at 13 TeV before a fit to data for the 3 and 4 tag regions. For the sake of brevity, only the signal yields for $C_t = \pm 1$ are provided.

Process	2 tag	3 tag	4 tag	
$t\bar{t} + b\bar{b}$	325 ± 175	61 ± 34	4.3 ± 2.8	
$t\bar{t} + 2b$	192 ± 91.0	26 ± 16	0.35 ± 0.29	
$t\bar{t} + b$	444 ± 230	37 ± 19	0.30 ± 0.26	
$t\bar{t} + c\bar{c}$	1541 ± 829	33 ± 28	0.19 ± 0.58	
$t\bar{t} + lf$	7360 ± 1712	63 ± 28	0.05 ± 0.09	
single top	501 ± 78	10 ± 3	0.26 ± 0.16	
W/Z + jets	150 ± 36	5.1 ± 1.4	0.00 ± 0.00	
VV	7.2 ± 1.3	0.19 ± 0.08	0.00 ± 0.00	
$t\bar{t}H$	17 ± 8	4.2 ± 2.3	0.40 ± 0.26	
Σ	10537 ± 1929	239 ± 59	6 ± 3	
C_t C_V				
tHq				
1	1	0.89 ± 0.22	0.15 ± 0.05	0.01 ± 0.00
-1	1	12 ± 3	2.2 ± 0.7	0.12 ± 0.04
tWH				
1	1	0.61 ± 0.07	0.10 ± 0.02	0.01 ± 0.00
-1	1	6.7 ± 0.9	1.1 ± 0.2	0.06 ± 0.02

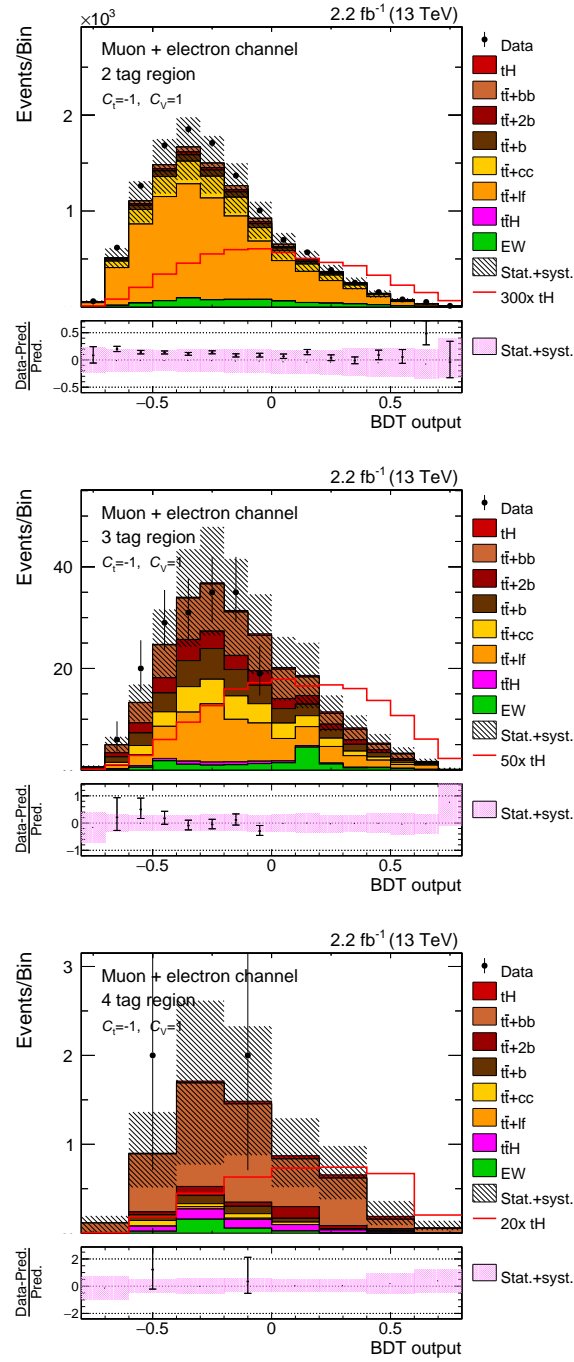


Figure 5.24.: Pre-fit distributions of the final classification BDT output. Simulation is normalized to luminosity. Data is shown only in a signal-depleted region since the analysis has been performed in a blind way. Uncertainty bands include both statistical and systematic uncertainties and cover data-MC discrepancies. The expected upper limits on the ratio $\sigma/\sigma_{\text{theor.}}$ are derived from the lower two distributions.

5. Prospects for a search for associated production of single top quark and Higgs boson at 13 TeV

Table 5.4.: Expected upper limits at 13 TeV on $\sigma/\sigma_{\text{theor}}$ at 95% confidence level. The numbers are quoted together with the $\pm 1/2\sigma$ uncertainties on the limit. The limits of the remaining signal scenarios considered can be found in Tab. B.5.

C_t	C_V	Exp. upper limit
1	1	$182^{+109/264}_{-63/98}$
-1	1	$11.0^{+6.5/16.8}_{-3.8/5.9}$

$s + b$ and the b -only hypothesis. The integrated luminosity of 2.2 fb^{-1} is not large enough to exclude any of the scenarios considered; however, going to larger negative values of $C_t = -3$, one approaches the yellow line representing an analysis that is sensitive enough to exclude exactly the respective scenario at 95% C.L.

In the middle pad of Fig.5.25 the yields in the signal region ($\sigma_{\text{fid.}}$) of tHq and tWH are compared. Here, the effect of the harder and more central kinematics of the Higgs boson in the case of tWH than for tHq (see Fig. 5.3) becomes visible. While the inclusive rate for $(C_t, C_V) = (-1, 1)$ is 5.4 times larger for tHq than for tWH, acceptance effects when imposing selection cuts on central b jets bring down the ratio to about 2. This is even more drastic for the SM case, where tHq and tWH contribute almost equally.

The bottom ratio shows the selection efficiency for the signal processes with respect to their inclusive production rate. For a given value of C_V , the selection efficiency has a minimum for the C_t values that lead to topologies where the Higgs boson and the top quark are produced least centrally. The differences between the green, blue and red curves on the other hand can be attributed to two different features: firstly, for $C_V = 0.5$, the Higgs boson and top quark are harder and more central than for the other two cases. This translates into a relatively larger acceptance for the final state objects as they appear in the detector, like the lepton from the top quark or the b jets. Secondly, the branching fraction for $H \rightarrow b\bar{b}$ decreases as C_V is increased, resulting in smaller yields for larger C_V with the employed analysis selection.

5.8. Chapter summary

An analysis strategy similar to the one at 8 TeV has been followed in the search for the associated production of a single top quark and a Higgs boson with the first 2.2 fb^{-1} of LHC data recorded at 13 TeV. However, this thesis has extended the scope of the analysis by including the tWH production mode as an additional signal process and considering multiple signal scenarios. In Run-I, only the scenario of a flipped sign for the top-Yukawa coupling, i.e. $C_t = -1$ has been considered. Now, the two-dimensional plane spanned by C_t and C_V – the latter factor denotes the scaling of the Higgs boson coupling to vector bosons – is scanned by directly searching for multiple scenarios in this plane. Some of

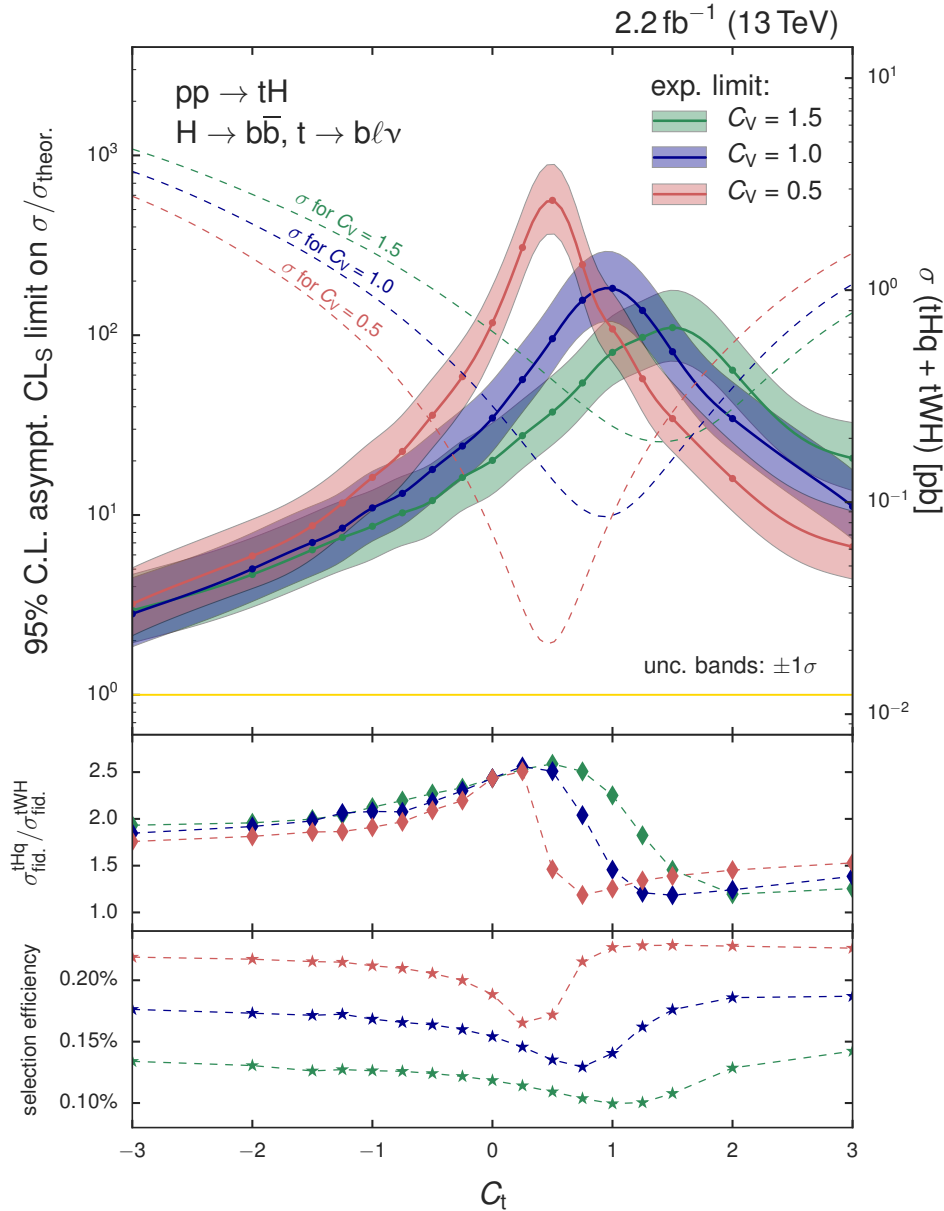


Figure 5.25.: Asymptotic expected upper limits at 95% C.L. for the associated production of a single top quark and a Higgs boson at 13 TeV with an integrated luminosity of 2.2 fb^{-1} are represented by the solid lines. Dots mark the 51 signal hypotheses considered in the analysis. A cubic spline fit has been used for interpolating between the points for a given value of C_V . Semi-transparent bands indicate the $\pm 1\sigma$ uncertainties on the exclusion limit. Dashed lines for the combined inclusive $tHq+tWH$ production cross sections allow for a direct comparison with the sensitivity on the respective scenario. To which extent either the t -channel or tW associated production of the Higgs boson contribute to the event yields in the signal region is shown in the ratio on the middle pad. The selection efficiency for the signal (the sum of tHq and tWH) with respect to their inclusive production is provided below. Tab. B.5 lists the numerical values of the limits.

5. Prospects for a search for associated production of single top quark and Higgs boson at 13 TeV

these display drastic differences in terms of inclusive and differential cross sections with respect to the SM case. With the data collected in 2015 and analyzed in this chapter, it is not yet expected that the analysis will be able to exclude any of the scenarios considered. Anomalous $t\bar{H}$ production with $C_t = -1$ -like kinematics can be excluded at 95% C.L. if its cross section is larger than 6.00 pb, translating into an upper limit on the signal strength modifier of $\mu = 11.0$. Accordingly, the search is roughly half as sensitive as the one performed with 19.7 fb^{-1} at 8 TeV. This seems reasonable given that one gains a factor of ~ 4 for the cross section (by moving from 8 to 13 TeV and including the tWH process in the signal definition) but the integrated luminosity has been nine times larger in Run-I. In the scenario with SM couplings, i.e. $C_t = C_V = 1$, $\mu > 182$ is expected to be excluded at 95% C.L.

No other search for $t\bar{H}$ production has been performed at 13 TeV yet; a statement on the combined exclusion power of all Higgs boson decay channels is therefore not possible at the moment. However, the CMS collaboration will continue pursuing the approach of a direct search for $t\bar{H}$, and all individual decay channels of the combination at 8 TeV will be covered also in Run-II. In particular, one analysis will search for events in which the Higgs boson decays into two photons and the top quark decays hadronically. This final state has not been looked for in Run-I data but seems a “low hanging fruit” given its unique signature.

Eventually, a combination with $t\bar{t}H$ searches to measure the top-Yukawa coupling at larger integrated luminosities is aimed for. First considerations as to how to make event selections orthogonal yet retain the highest possible search sensitivity for the single analyses have been made. The most straightforward way is to veto forward jets in $t\bar{t}H$ searches and require at least one forward jet for $t\bar{H}$ channels. One can also make the analyses orthogonal by dividing up the phase space into $t\bar{H}$ enriched and $t\bar{t}H$ enriched regions by means of an MVA output distribution that separates between the two processes. This is a more sophisticated possibility and would ensure that selections are not too demanding for the respective signals. However, the question will have to be answered how to account for the change in kinematics for different values of C_t (and C_V) in the case of $t\bar{H}$ with this approach, whereas only the overall rate but not the kinematics varies for $t\bar{t}H$. Further studies and discussion are therefore needed to maximize synergy effects between the two top quark + Higgs boson production modes.

6. Summary and outlook

This thesis presents the direct search for the associated production of a single top quark and a Higgs boson in the $H \rightarrow b\bar{b}$ channel. At $\sqrt{s} = 8$ TeV the first dedicated search has been developed within the CMS collaboration, with significant contributions from the author of this thesis at every step of the analysis chain, and has been made public through a Physics Analysis Summary [156] and a combination paper [48]. The analysis is recapulated here. The thesis for the first time also provides expected upper limits on tH production that are achievable with the early data collected in the LHC Run-II, where a higher center-of-mass energy, leading to an increased signal cross section compared to Run-I, is accessible.

Unlike the predominant gluon fusion or other common Higgs boson production channels, the rare tH production is sensitive to the relative sign of the top-Yukawa coupling y_t and the coupling to massive gauge bosons. Consequently, both its inclusive rates as well as the differential cross sections are functions of the Higgs boson couplings, and the channel is suited to look for deviations from their SM values. Depending on the theoretical assumptions put into a global fit to LHC data, in particular a scenario where y_t has a flipped sign with respect to the SM (the cross section of this scenario is enhanced by a factor of 13 for $y_t = -1$ at a center-of-mass energy of 8 TeV) is disfavored only at the 1σ level. This has motivated the search for the production of a Higgs boson in association with a single top quark.

The collision data analyzed in the Run-I analysis has been recorded with the CMS detector during 2012 and corresponds to an integrated luminosity of 19.7 fb^{-1} . Despite the enhancement for the exotic scenario, the expected tHq yields in a signal-enriched region defined by requirement on the presence of three or four b tagged jets and an additional untagged jet are small compared to the predominant $t\bar{t} + \text{jets}$ background (~ 40 tHq events vs. ~ 2500 background events). Another challenge consists in the multi-jet final state which hampers a correct identification and reconstruction of objects like the Higgs boson and the top quark. For these reasons, multivariate analysis tools (MVAs) have been used to maximize the performance of the analysis: an MVA aims at reconstructing the events and assigning the correct jets to the parton level quarks, which in case of the tHq process are the light forward quark and the three b quarks from the decay of both the Higgs boson and the top quark. In order to account for genuine features of the $t\bar{t}$ process with subsequent semi-leptonic decay, another reconstruction MVA is designed to identify the jets belonging to the b quarks from the top quark pair and the two light quarks from the hadronically decaying W boson. After reconstructing each event passing the event selection under the two hypotheses, a final classification MVA has been employed to separate the tHq process from the background. By means of the output distributions of this MVA, an upper CL_S exclusion limit at 95% confidence

6. Summary and outlook

level on tHq production with $y_t = -1$ has been derived, resulting in an upper boundary of 1.78 pb and a signal strength modifier of $\mu < 7.6$.

At a higher center-of-mass energy of 13 TeV, a total integrated luminosity of 2.2 fb^{-1} has been collected with the CMS detector during 2015. This data is compared to theoretical predictions and is used to give an overview on the quality of underlying distributions, revealing good agreement between observation and simulation. Expected upper limits on points in the two-dimensional phase space spanned by the generic real-valued scaling factors C_t and C_V for the top-Yukawa coupling and for the coupling of the Higgs boson to weak gauge bosons, respectively, are set. This approach represents a novelty in the scope of the analysis compared to Run-I and has been designed, developed and conducted for the first time in the context of this thesis. Assuming the existence of physics beyond the Standard Model that is responsible for the generation of SM-like top quark and gauge boson masses for $|C_{t,V}| \neq 1$, the couplings of the Higgs boson to the SM particles can be varied independently by the scale factors, and the production rates can be constrained. The amount of collected data is not large enough yet to allow for an analysis that is expected to be able to exclude any of the points considered in the two-dimensional plane. In particular, the associated production of a single top quark and a Higgs boson – where also the tWH process has been considered for the first time – with kinematics as in the two benchmark models $C_t = 1$ and -1 (with $C_V = 1$) is expected to be excluded at 95% C.L. if $\sigma \times BR$ is larger than 9.28 pb and, respectively, 6.00 pb.

A projection to larger integrated luminosities under the naïve assumption that systematic uncertainties stay the same with respect to the ones used for the first 13 TeV result mentioned above is provided in Fig. 6.1 for the two $C_t = \pm 1$ scenarios. In neither of the cases, the analysis can exclude the respective model with 3000 fb^{-1} of pp collision data at the High-Luminosity-LHC, and the limit on $\sigma/\sigma_{\text{theor.}}$ approaches a plateau which lies well above 1. In this regime, the sensitivity is purely determined by the systematic uncertainties the measurement is afflicted with. Therefore, their reduction is necessary if tH production with $H \rightarrow b\bar{b}$ shall ever be observed even at the highest luminosities. The dominant systematic uncertainties are mostly theory related and reflect the level of understanding of the $t\bar{t}$ process. Given that the phase space considered in the analysis is dominated by large jet multiplicities, incorporating more NLO-accurate jets in the matrix element calculation for the $t\bar{t}$ process seems a promising approach and can lead to an improved description and reduced theoretical uncertainties for the main background. However, generating a simulation sample with improved precision is impractical from the computational point-of-view. Progress should be made on the theoretical side to circumvent the large fraction of negative weights, also in view of a potential NLO signal sample. Along the same lines, a better theoretical understanding both of the kinematics of additional jets produced in association with a top quark pair and of the heavy flavor content among these extra jets would help constraining another dominant source of systematic uncertainty. This is also a pivotal point for other Higgs boson + heavy flavor (hf) quarks searches like $t\bar{t}H$, $H \rightarrow b\bar{b}$, where $t\bar{t} + hf$ processes represent an irreducible background.

While the points remarked above seem to provide a discouraging outlook in regard

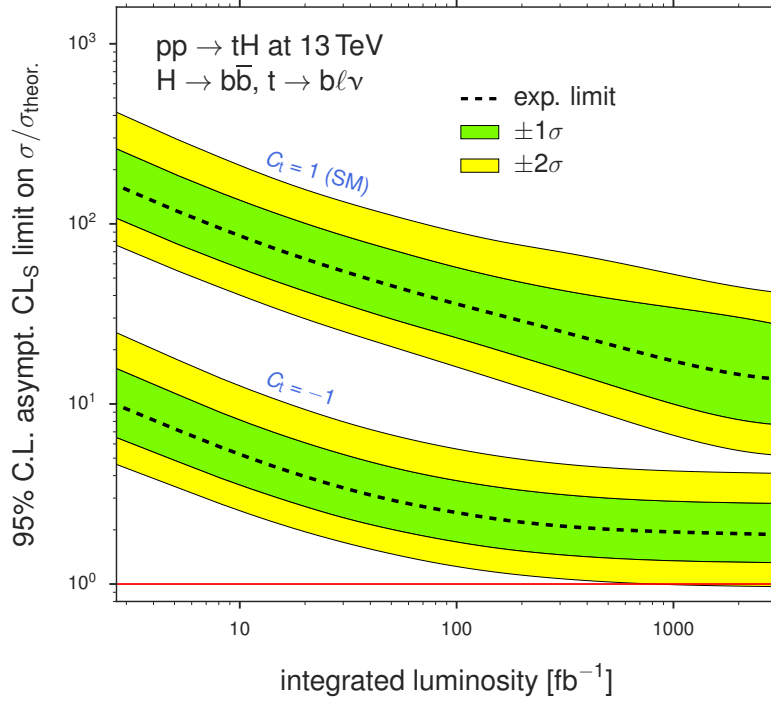


Figure 6.1.: A projection for the integrated luminosity up to 3000 fb^{-1} is performed in order to estimate the discovery potential of the channel. The red horizontal line indicates the case where the analysis is sensitive enough to exactly exclude the models; however, it is never crossed by the dashed black line representing the median expected upper limit for $C_t = -1$ and $C_t = 1$, respectively. Implications of this are discussed in the body text.

to the discovery potential of the channel alone, it is clear that the full potential of tH production in constraining Higgs boson couplings can only be released in a combination with other decay modes. There, on the other hand, $H \rightarrow b\bar{b}$ is one of the most sensitive channels, and the analysis drives the exclusion limits together with $H \rightarrow \gamma\gamma$. A first combination at $\sqrt{s} = 8 \text{ TeV}$ has been able to put stringent limits on the exotic scenario of tH production. At 13 TeV and with larger integrated luminosities, a combination of all decay channels will become even more significant: it can potentially exclude exotic single top quark + Higgs boson production, and the top-Yukawa coupling can be measured by combining tH and $t\bar{t}H$ analyses provided one succeeds in defining an efficient way how to make analysis selections orthogonal. This is subject of ongoing discussions within the community that pave the way towards a coherent strategy for the search for top quark + Higgs boson signatures.

A. Appendix: Search for single top quark + Higgs boson at 8 TeV

A.1. Recorded and simulated datasets

Table A.1.: Experimental pp collision data exploited in the analysis. The second column shows the recorded integrated luminosity.

Dataset name	Int. luminosity
Muon	
/SingleMu/Run2012A-22Jan2013-v1/AOD	876 pb ⁻¹
/SingleMu/Run2012B-22Jan2013-v1/AOD	4 412 pb ⁻¹
/SingleMu/Run2012C-22Jan2013-v1/AOD	7 055 pb ⁻¹
/SingleMu/Run2012D-22Jan2013-v1/AOD	7 369 pb ⁻¹
Electron	
/SingleElectron/Run2012A-22Jan2013-v1/AOD	876 pb ⁻¹
/SingleElectron/Run2012B-22Jan2013-v1/AOD	4 412 pb ⁻¹
/SingleElectron/Run2012C-22Jan2013-v1/AOD	7 055 pb ⁻¹
/SingleElectron/Run2012D-22Jan2013-v1/AOD	7 369 pb ⁻¹
total luminosity: 19.71 fb ⁻¹	

Table A.2.: Simulated datasets used for the 8 TeV analysis. For samples with no provided reference for the cross section, the rate as predicted by the generator is chosen (Pythia6 in the two cases). The respective branching ratios of t , W and H decays have been applied for the listed values.

Process	ME generator	Parton shower	ME precision	Num. of events	Cross section \times BR, fb
$t\bar{t}q$	MadGraph	Pythia6	LO	5 000 000	36.8 (NLO) [34]
$t\bar{t}H$	Pythia6	Pythia6	LO	995 697	130.2 (NLO) [185]
$t\bar{t}$ + jets, semi-lept. decay	MadGraph MLM	Pythia6	LO	86 814 792	107.7 (NNLO) [184]
$t\bar{t}$ + jets, full-lept. decay	MadGraph MLM	Pythia6	LO	12 119 013	25.8 (NNLO) [184]
single top t -chan., t	Powheg	Pythia6	NLO	3 915 598	18.27 (approx. NNLO) [198]
single top t -chan., \bar{t}	Powheg	Pythia6	NLO	1 711 403	9.95 (approx. NNLO) [198]
single top tW, t	Powheg	Pythia6	NLO	497 658	11.1 (approx. NNLO) [198]
single top tW, \bar{t}	Powheg	Pythia6	NLO	493 460	11.1 (approx. NNLO) [198]
single top s -chan., t	Powheg	Pythia6	NLO	3 932 710	1.23 (approx. NNLO) [198]
single top s -chan., \bar{t}	Powheg	Pythia6	NLO	1 949 667	0.57 (approx. NNLO) [198]
W + 0 jet	MadGraph MLM	Pythia6	LO	57 509 905	35 509 (NNLO) [199]
W + 2 jets	MadGraph MLM	Pythia6	LO	33 894 921	2116 (NNLO) [199]
W + 3 jets	MadGraph MLM	Pythia6	LO	15 289 503	637 (NNLO) [199]
W + 4 jets	MadGraph MLM	Pythia6	LO	13 382 803	262 (NNLO) [199]
WW	Pythia6	Pythia6	LO	9 800 431	54.8 (NLO) [199]
WZ	Pythia6	Pythia6	LO	9 950 283	12.6 (LO)
ZZ	Pythia6	Pythia6	LO	9 799 908	5.2 (LO)
Drell-Yan	MadGraph MLM	Pythia6	LO	29 909 503	3504 (NNLO) [199]

A.1.1. Corrections

The scale factors to adjust the modelling of the top quark p_T distributions to the ones observed in data are, depending on the allowed top quark decays, given by:

$$\begin{aligned} A_{\text{full-lept.}} &= 0.148, & B_{\text{full-lept.}} &= 0.00129 \\ A_{\text{semi-lept.}} &= 0.159, & B_{\text{semi-lept.}} &= 0.00141 \\ A_{\text{incl.}} &= 0.156, & B_{\text{incl.}} &= 0.00137. \end{aligned}$$

They are used in the equation for the event weight $w_{\text{top } p_T}$:

$$w_{\text{top } p_T} = \sqrt{\exp(A - B \cdot p_T(\text{top})) \cdot \exp(A - B \cdot p_T(\text{antitop}))}$$

As is mentioned in Sec. 4.2.5, the missing transverse energy needs to be corrected for an x - y modulation observed in the reconstructed \cancel{E}_T [168]. The x and y component of the \cancel{E}_T vector in data are corrected with:

$$\begin{aligned} (\cancel{E}_T)_{x,\text{corr}} &= (\cancel{E}_T)_{x,\text{uncorr}} - 0.0483642 + (0.24887 \cdot N_{V,\text{PU}}) \\ (\cancel{E}_T)_{y,\text{corr}} &= (\cancel{E}_T)_{y,\text{uncorr}} + 0.150135 + (0.0828917 \cdot N_{V,\text{PU}}) \end{aligned}$$

For simulation, the correction is:

$$\begin{aligned} (\cancel{E}_T)_{x,\text{corr}} &= (\cancel{E}_T)_{x,\text{uncorr}} - 0.162861 + (0.0238517 \cdot N_{V,\text{PU}}) \\ (\cancel{E}_T)_{y,\text{corr}} &= (\cancel{E}_T)_{y,\text{uncorr}} - 0.360860 + (0.130335 \cdot N_{V,\text{PU}}) \end{aligned}$$

A. Appendix: Search for single top quark + Higgs boson at 8 TeV

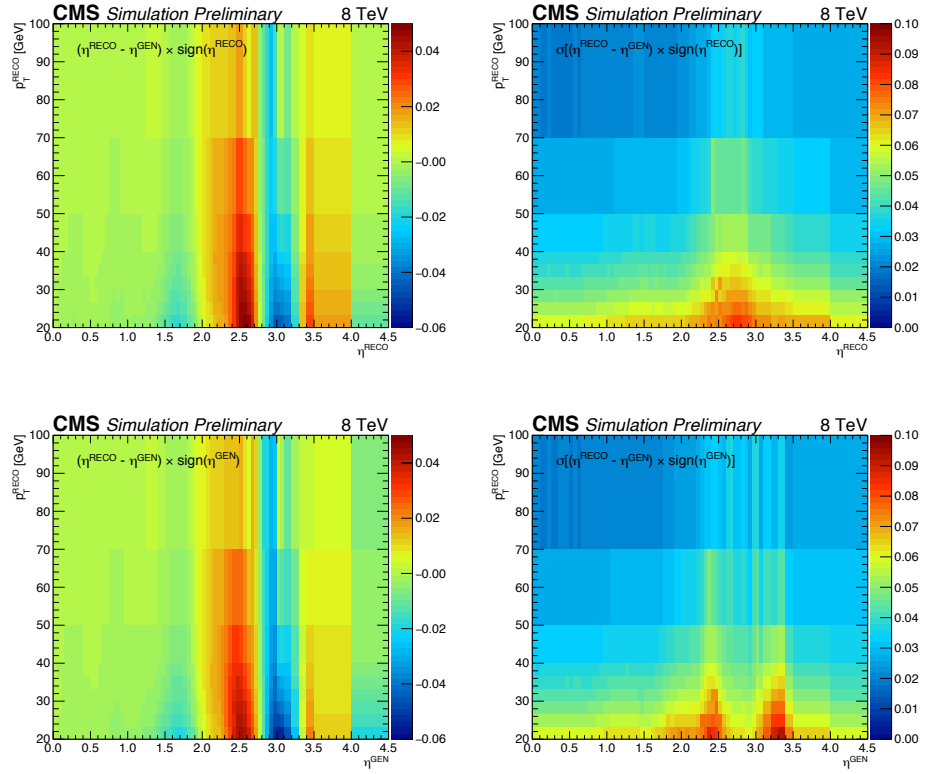


Figure A.1.: Differences (left column) and standard deviation of the differences (right column) in η between the reconstructed and matched generator level jets. Taken from [40].

A.2. MVA training

A.2.1. Validation figures

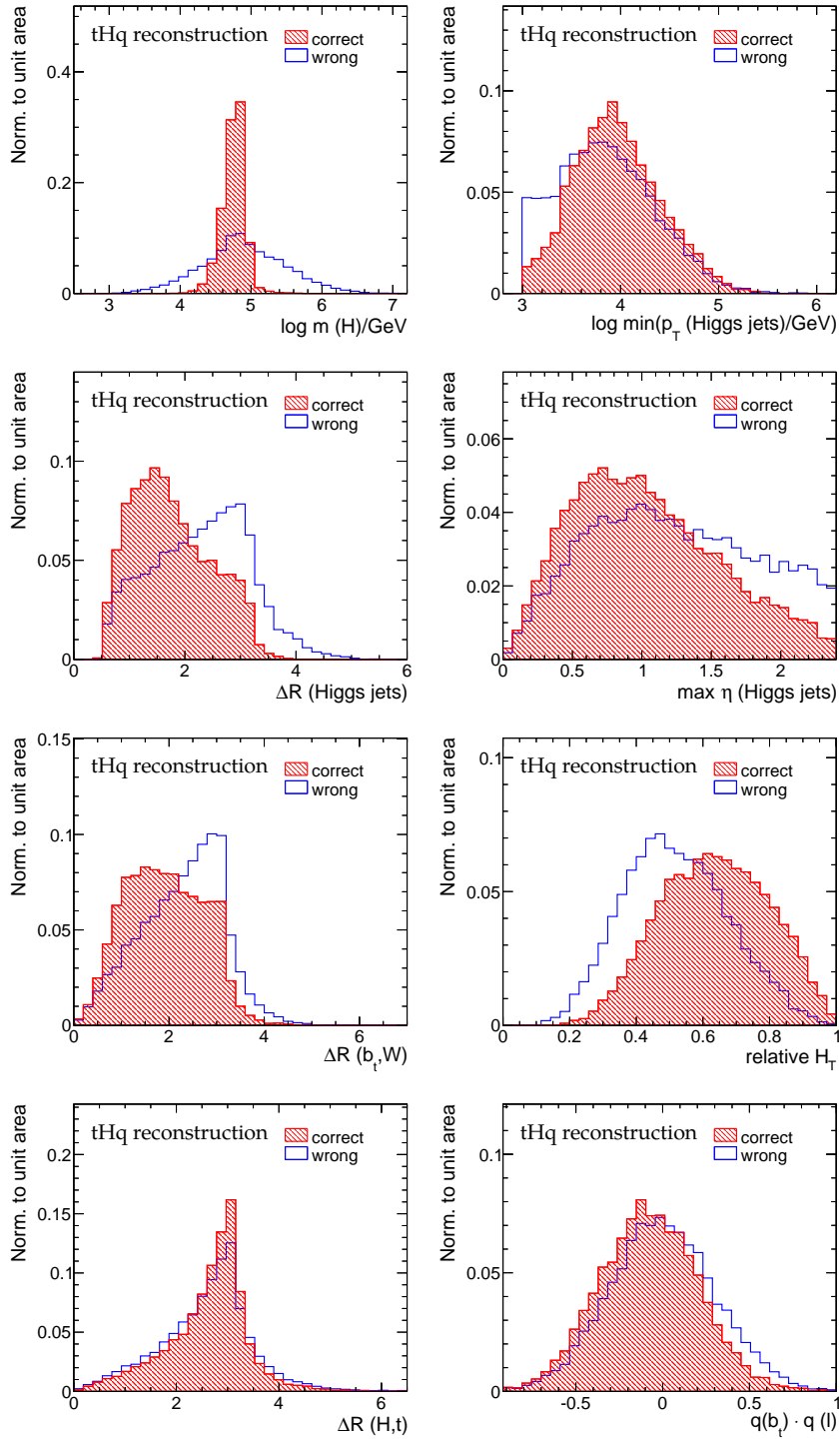


Figure A.2.: Remaining input variables for the tHq reconstruction. A clear separation between correct and wrong assignments is visible in all cases.

A. Appendix: Search for single top quark + Higgs boson at 8 TeV

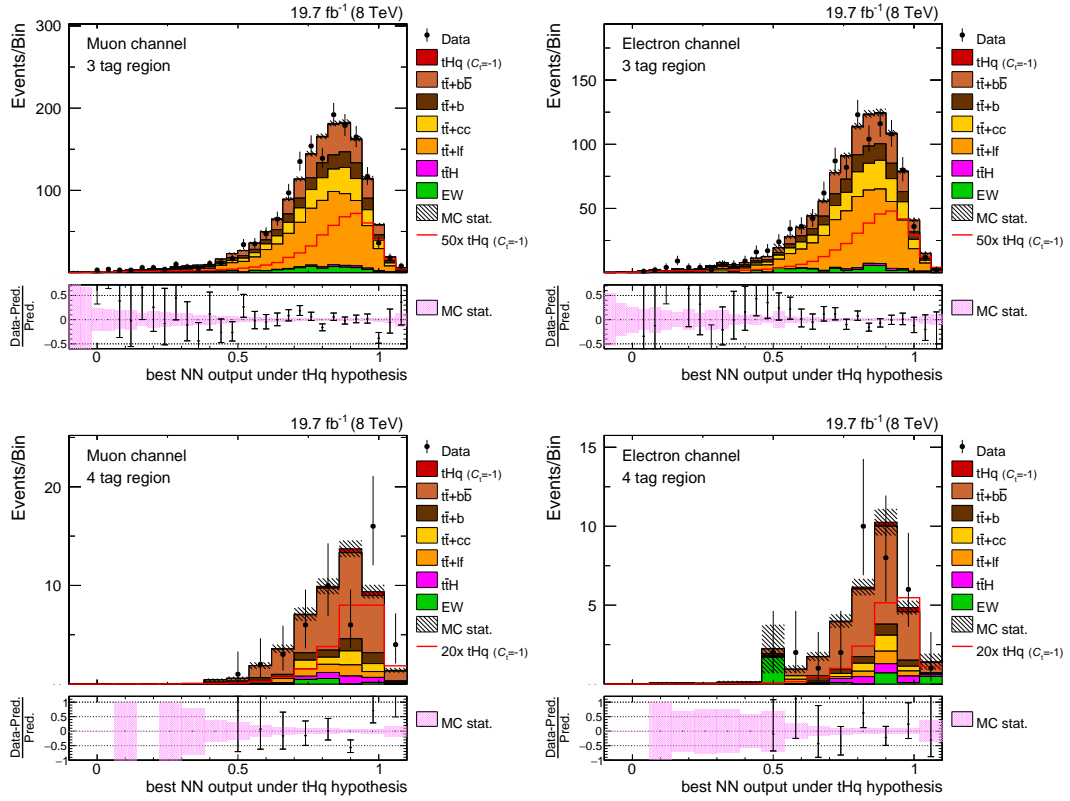


Figure A.3.: Best reconstruction NN output under the tHq hypothesis in the 3 tag and 4 tag region. Simulation is normalized to data.

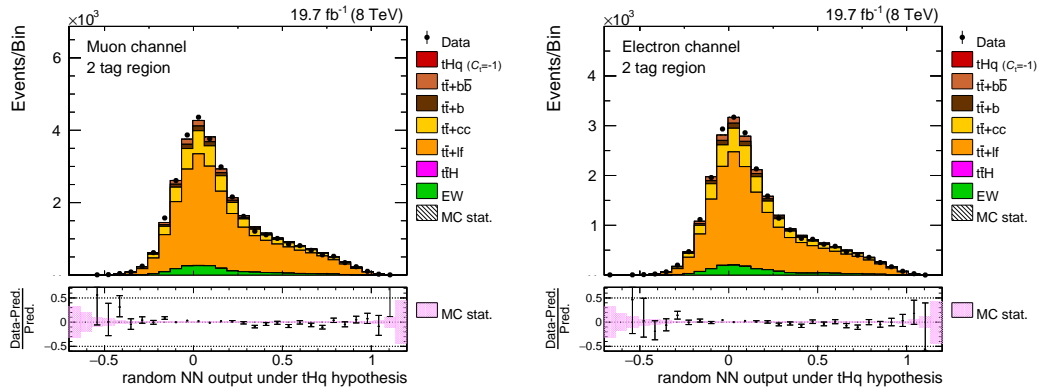


Figure A.4.: Random reconstruction NN output under the tHq hypothesis in the 2 tag region. Simulation is normalized to data.

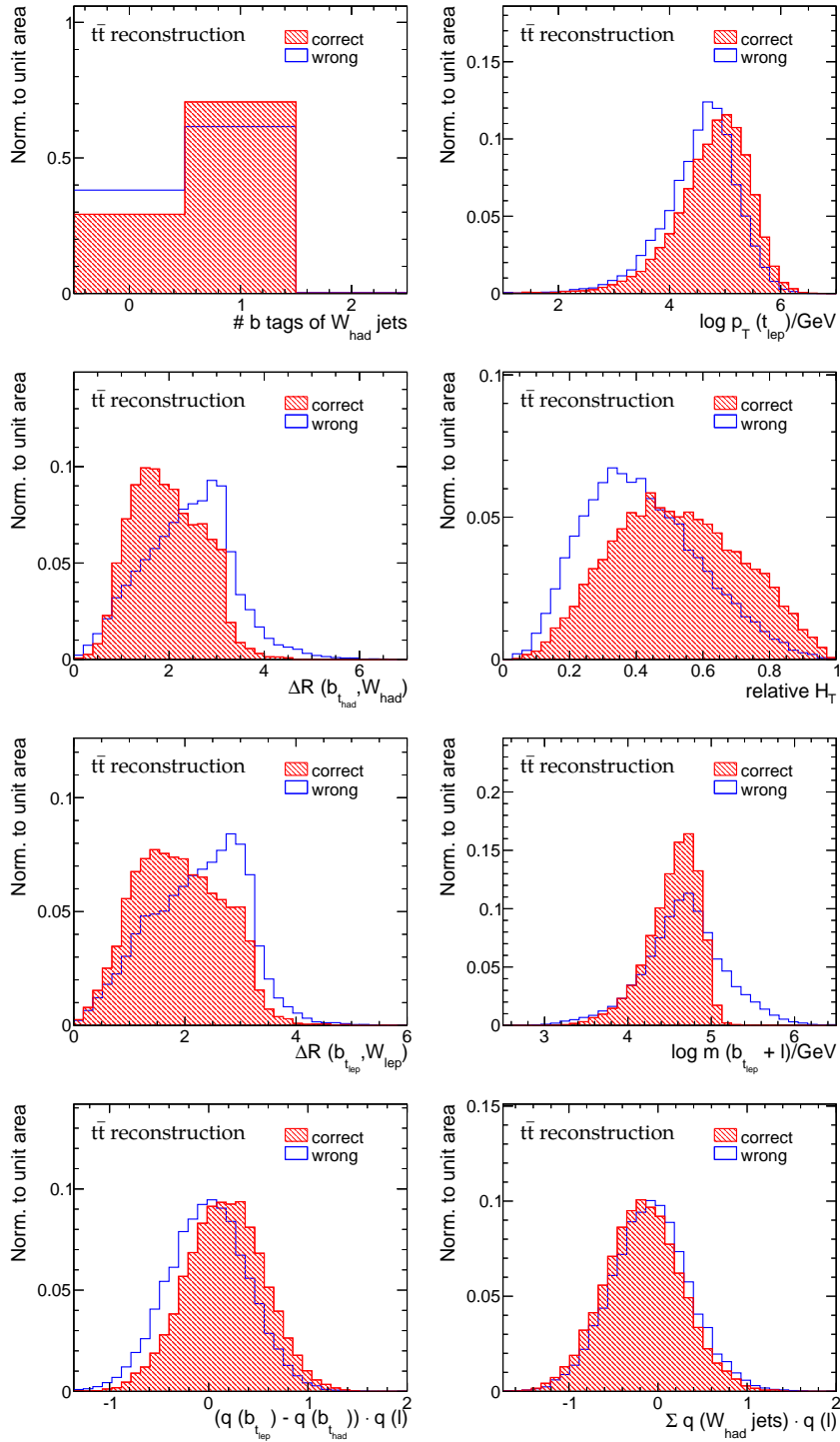


Figure A.5.: Remaining input variables for the $t\bar{t}$ reconstruction. A clear separation between correct and wrong assignments is visible in all cases.

A. Appendix: Search for single top quark + Higgs boson at 8 TeV

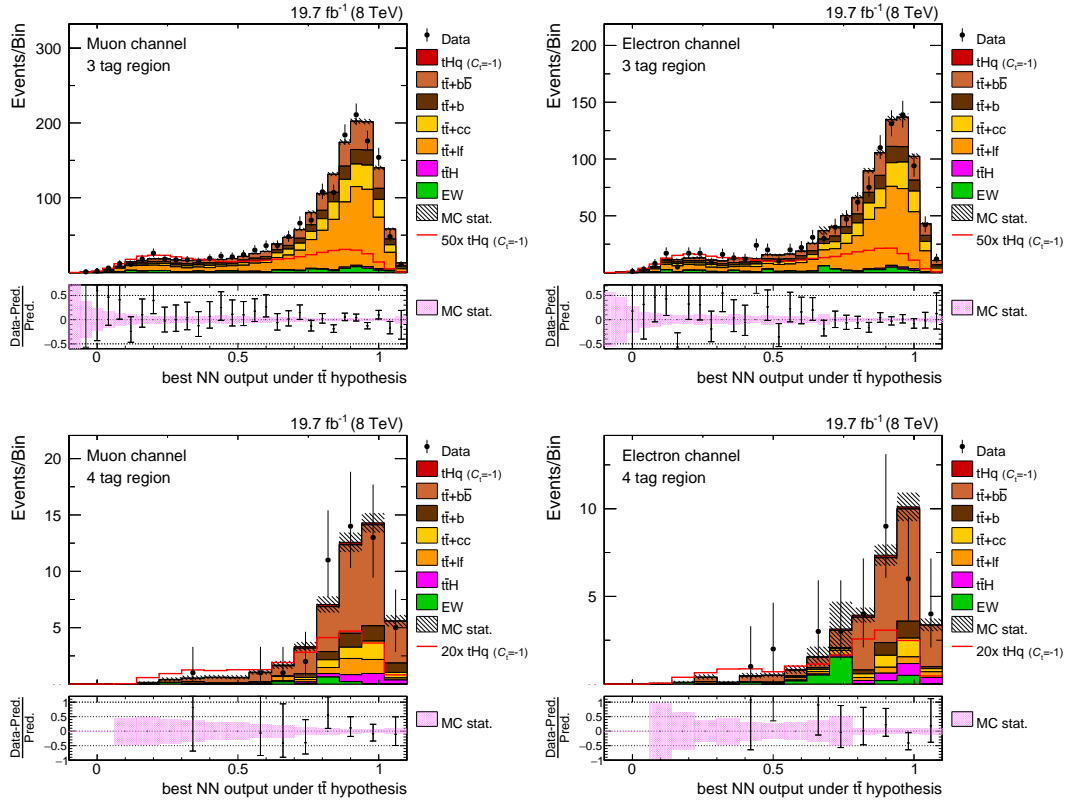


Figure A.6.: Best reconstruction NN output under the $t\bar{t}$ hypothesis in the 3 tag and 4 tag region. Simulation is normalized to data.

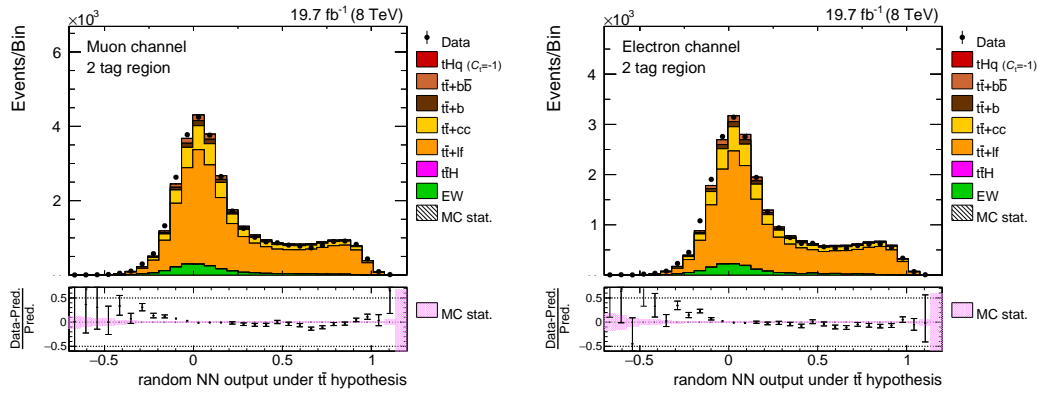


Figure A.7.: Random reconstruction NN output under the $t\bar{t}$ hypothesis in the 2 tag region. Simulation is normalized to data.

A.2. MVA training

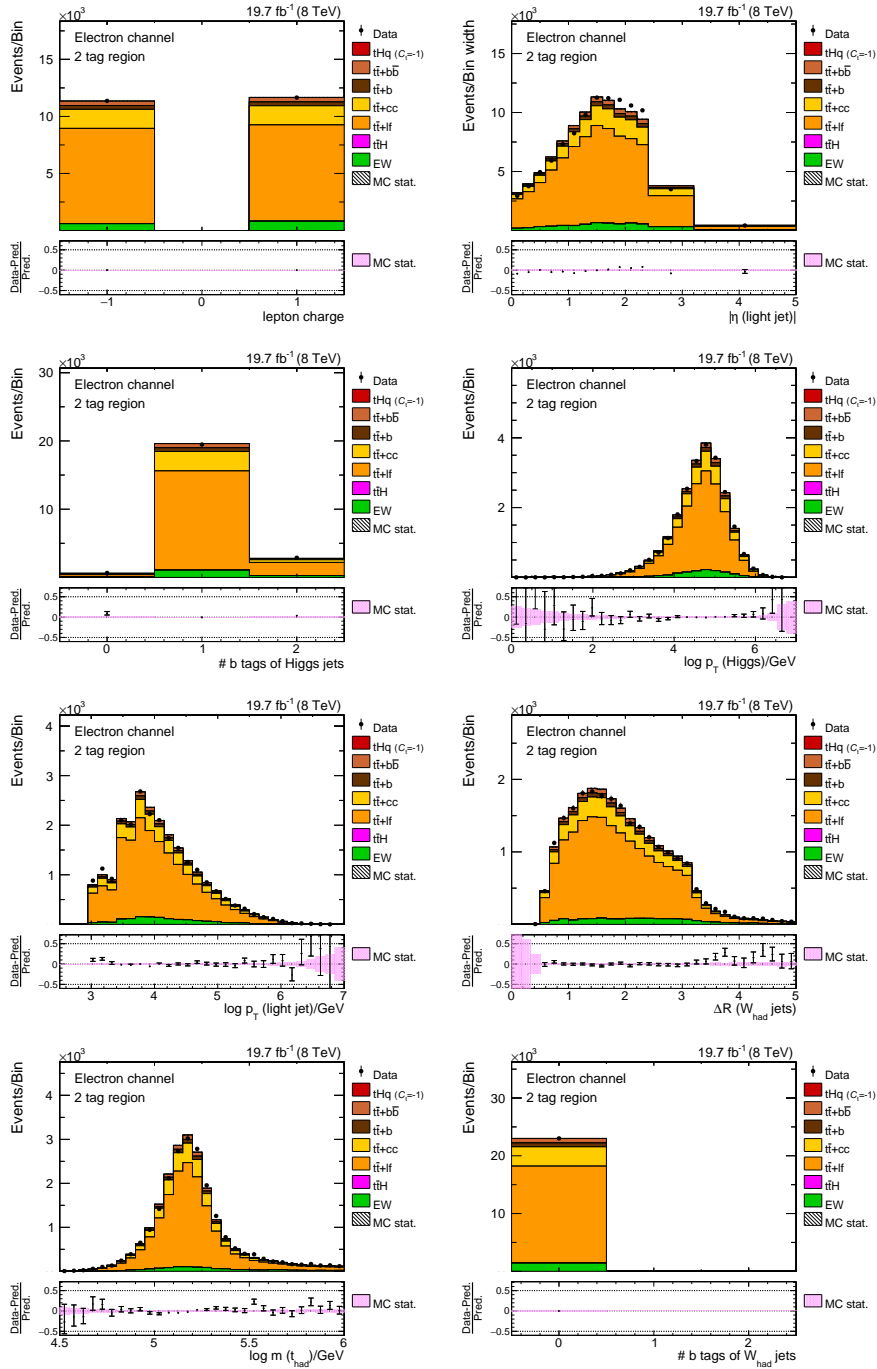


Figure A.8.: Distributions of the input variables for the classification NN in the electron channel in the 2 tag region, which is enriched in top quark pair events. Simulation is normalized to data in order to facilitate the shape comparison. All MC weights are applied.

A. Appendix: Search for single top quark + Higgs boson at 8 TeV

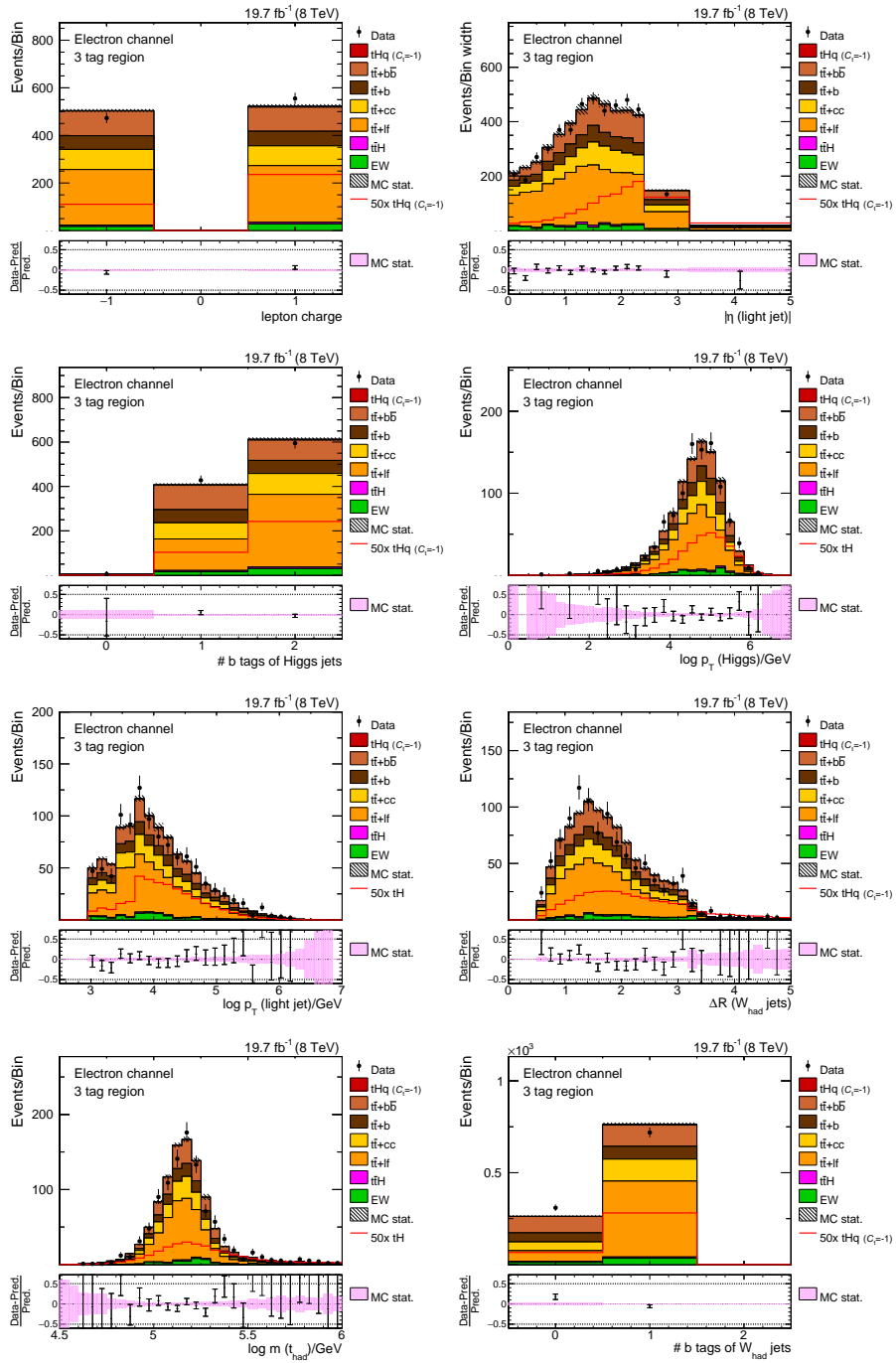


Figure A.9.: Distributions of the input variables for the classification NN in the electron channel in the 3 tag region. Simulation is normalized to data in order to facilitate the shape comparison. All MC weights are applied.

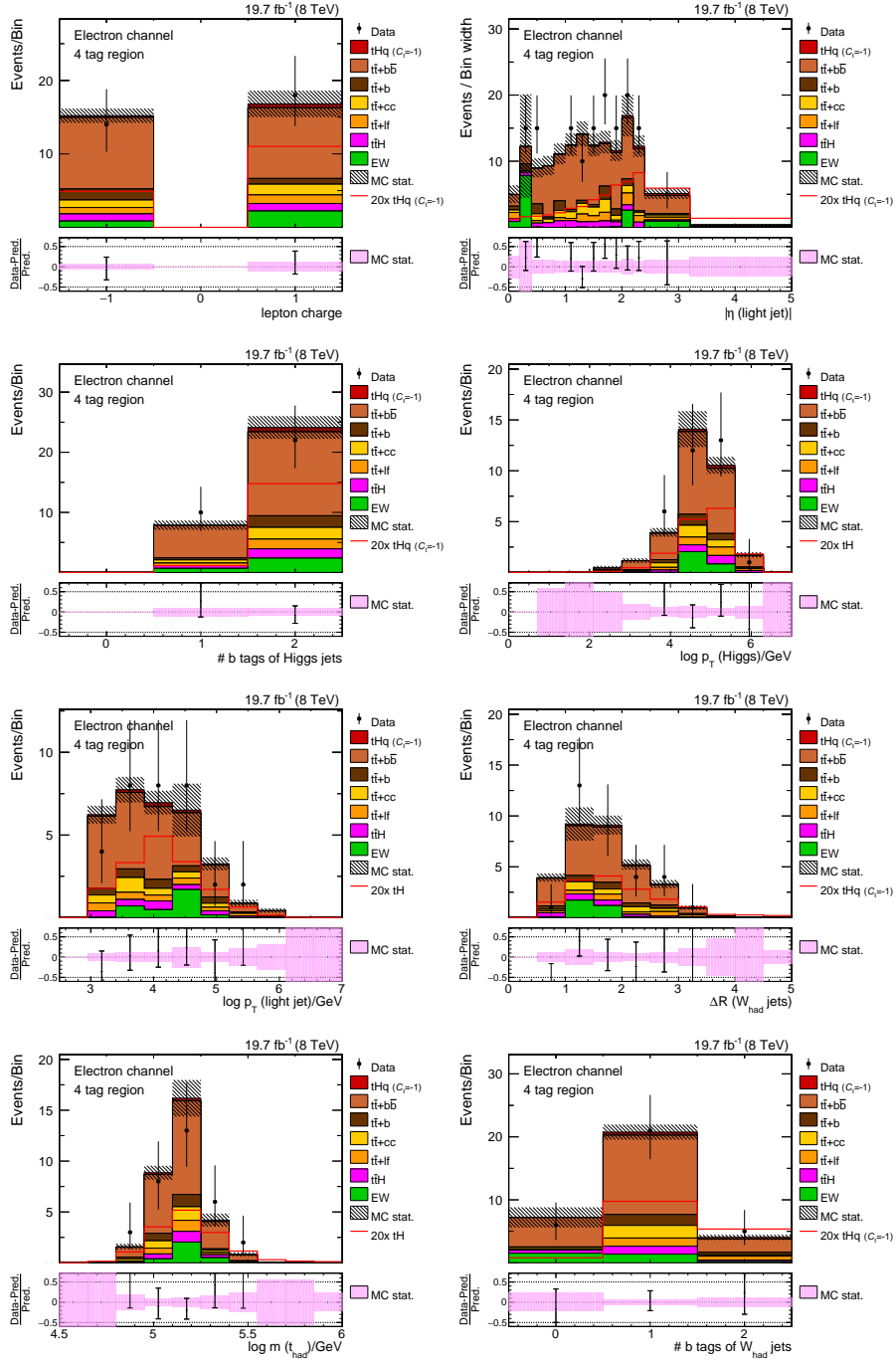


Figure A.10.: Distributions of the input variables for the classification NN in the electron channel in the 4 tag region. Simulation is normalized to data in order to facilitate the shape comparison. All MC weights are applied.

A. Appendix: Search for single top quark + Higgs boson at 8 TeV

A.2.2. Used settings

The settings used for both reconstruction NNs and the classification NN are:

VarTransform=N:NeuronType=tanh:NCycles=500:HiddenLayers=30:

TrainingMethod=BFGS:TestRate=5

A.3. Pre-fit distributions

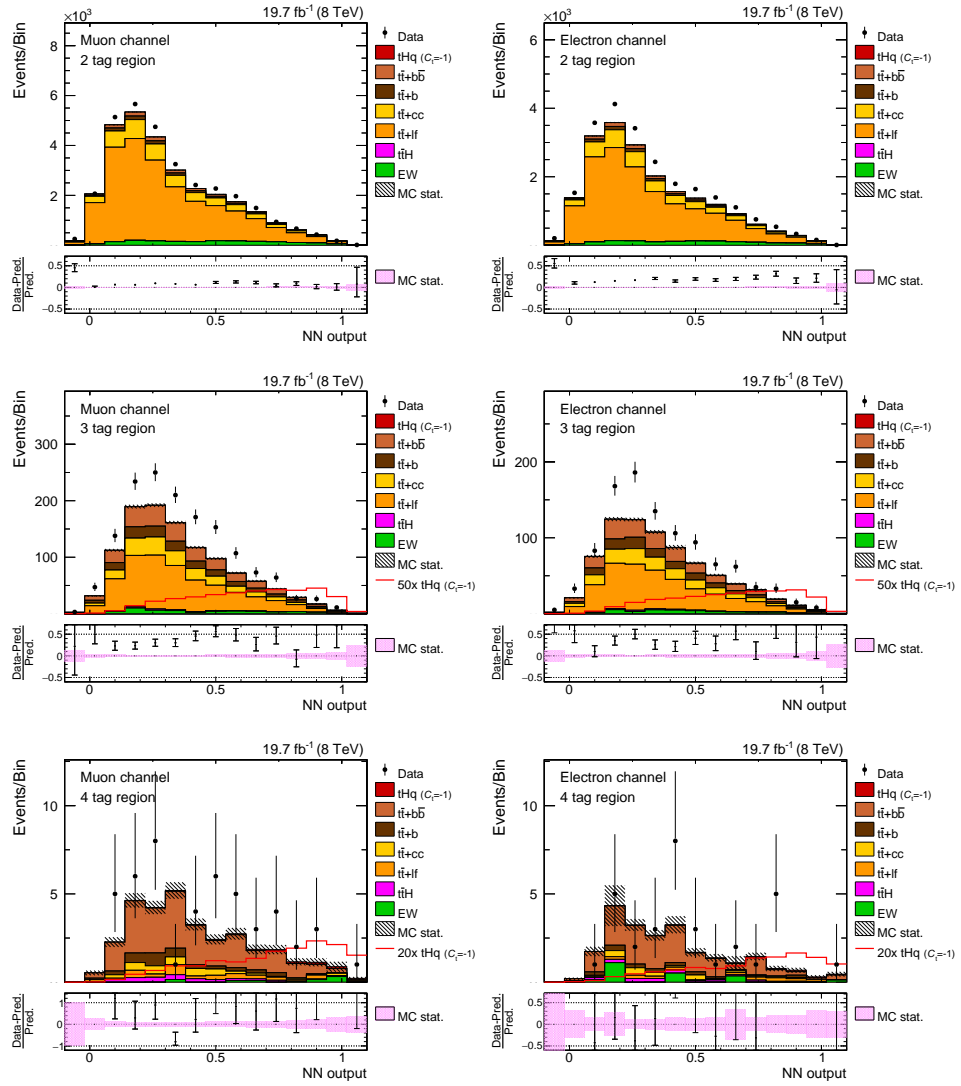


Figure A.11.: Pre-fit distributions of the final classification NN output.

A.4. Pulls for fit to Asimov data

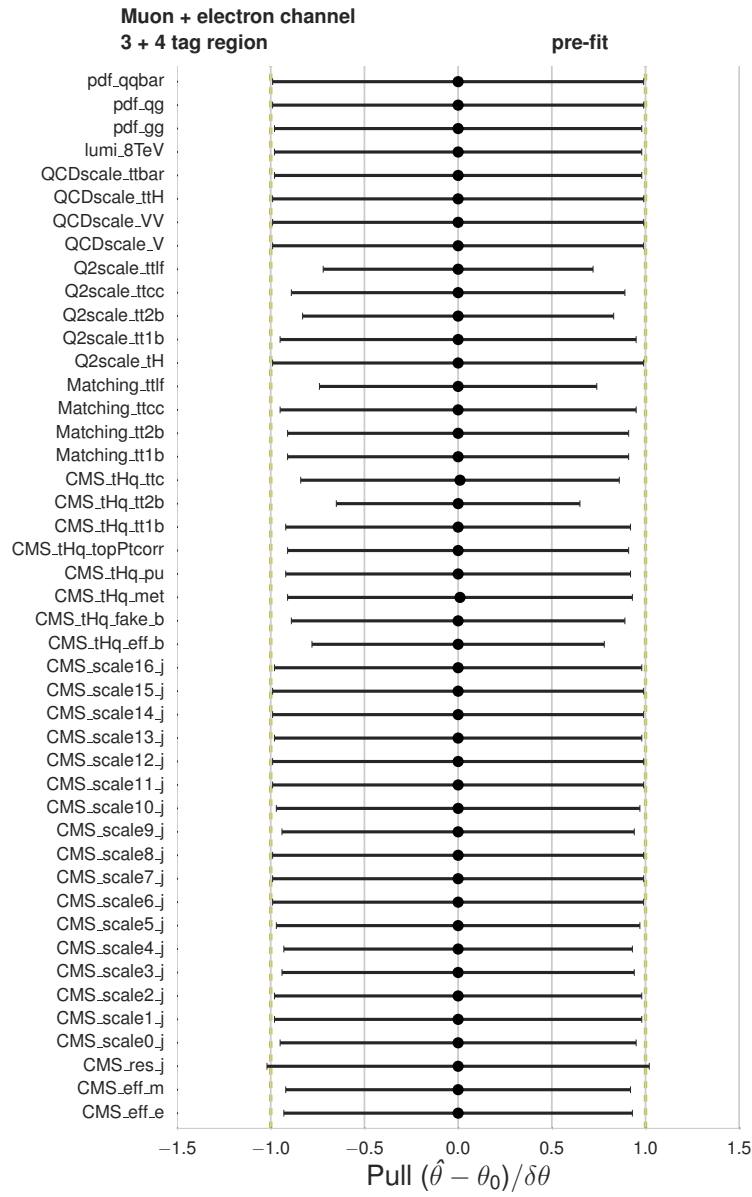


Figure A.12.: Pulls of the single nuisance parameters in the fit to an Asimov data set.

B. Appendix: Search for single top boson + Higgs boson at 13 TeV

B.1. Scalings and couplings

The branching fraction of $H \rightarrow b\bar{b}$ depends on the scaling factor C_V .

$$\begin{aligned}\Gamma_{bb}^{\text{SM}} + \Gamma_{\tau\tau}^{\text{SM}} + \dots + \Gamma_{VV}^{\text{SM}} &= \Gamma^{\text{SM}} \\ \Gamma_{bb}^{\text{SM}} + \Gamma_{\tau\tau}^{\text{SM}} + \dots + C_V^2 \Gamma_{VV}^{\text{SM}} &= \Gamma^{\text{SM}} + (C_V^2 - 1) \Gamma_{VV}^{\text{SM}} \\ BR(H \rightarrow b\bar{b}) &= \frac{\Gamma_{bb}^{\text{SM}}}{\Gamma^{\text{SM}}} \\ BR^*(H \rightarrow b\bar{b}) &= \frac{\Gamma_{bb}^{\text{SM}}}{\Gamma^{\text{SM}} + (C_V^2 - 1) \Gamma_{VV}^{\text{SM}}}.\end{aligned}$$

Therefore, the scaling factor for the branching ratio $H \rightarrow b\bar{b}$ is given by

$$\frac{BR^*}{BR} = \frac{1}{1 + (C_V^2 - 1) \Gamma_{VV}^{\text{SM}}}.$$

For varying C_t , the branching fraction $H \rightarrow \gamma\gamma$ changes as well. This would imply also a dependence on $BR(H \rightarrow b\bar{b})$ on C_t . However, since the branching fraction of the Higgs boson decaying into two photons is way smaller than into two b quarks, the dependence can be neglected.

On the other hand, the production cross section of $pp \rightarrow t\bar{t}H$ has a C_t^2 dependence that is taken into account in the analysis. Since the shape of the final classification BDT output of $t\bar{t}H$ and does not peak in signal enriched bins, the effect of changed yields for $t\bar{t}H$ on the exclusion limit for $t\bar{t}H$ is small.

B. Appendix: Search for single top boson + Higgs boson at 13 TeV

B.2. Production cross sections at NLO in QCD

Table B.1.: Production cross sections for tHq, depending on C_f and C_V . Obtained with MAD-GRAPH5_AMC@NLO at NLO in the 4F. The quoted uncertainties on the cross section correspond to scale variations in %.

C_f	C_V	σ [pb]	C_f	C_V	σ [pb]	C_f	C_V	σ [pb]
-3.0	0.5	$2.260^{+1.9}_{-2.7}$	-3.0	1.0	$2.991^{+2.1}_{-3.1}$	-3.0	1.5	$3.845^{+2.6}_{-3.2}$
-2.0	0.5	$1.160^{+2.0}_{-2.9}$	-2.0	1.0	$1.706^{+2.6}_{-3.2}$	-2.0	1.5	$2.371^{+2.5}_{-3.6}$
-1.5	0.5	$0.748^{+2.1}_{-3.1}$	-1.5	1.0	$1.205^{+2.5}_{-3.6}$	-1.5	1.5	$1.784^{+2.7}_{-3.9}$
-1.25	0.5	$0.573^{+2.1}_{-3.0}$	-1.25	1.0	$0.987^{+2.6}_{-3.4}$	-1.25	1.5	$1.518^{+2.8}_{-3.9}$
-1.0	0.5	$0.472^{+2.3}_{-3.3}$	-1.0	1.0	$0.793^{+2.7}_{-3.9}$	-1.0	1.5	$1.287^{+3.0}_{-4.3}$
-0.75	0.5	$0.300^{+2.5}_{-3.5}$	-0.75	1.0	$0.621^{+2.9}_{-4.1}$	-0.75	1.5	$1.067^{+3.1}_{-4.4}$
-0.5	0.5	$0.198^{+2.8}_{-3.9}$	-0.5	1.0	$0.472^{+3.2}_{-4.4}$	-0.5	1.5	$0.874^{+3.4}_{-4.7}$
-0.25	0.5	$0.119^{+3.1}_{-4.6}$	-0.25	1.0	$0.351^{+3.5}_{-5.0}$	-0.25	1.5	$0.703^{+3.6}_{-5.0}$
0.0	0.5	$0.062^{+3.8}_{-5.6}$	0.0	1.0	$0.248^{+3.9}_{-5.5}$	0.0	1.5	$0.558^{+3.8}_{-5.4}$
0.25	0.5	$0.028^{+5.0}_{-7.1}$	0.25	1.0	$0.169^{+4.4}_{-6.2}$	0.25	1.5	$0.437^{+4.2}_{-6.1}$
0.5	0.5	$0.018^{+4.2}_{-6.7}$	0.5	1.0	$0.113^{+5.0}_{-7.1}$	0.5	1.5	$0.334^{+4.6}_{-6.5}$
0.75	0.5	$0.030^{+1.4}_{-2.9}$	0.75	1.0	$0.081^{+5.7}_{-7.6}$	0.75	1.5	$0.256^{+5.2}_{-7.2}$
1.0	0.5	$0.066^{+1.0}_{-3.6}$	1.0	1.0	$0.071^{+4.1}_{-6.7}$	1.0	1.5	$0.200^{+5.7}_{-7.6}$
1.25	0.5	$0.124^{+0.9}_{-3.7}$	1.25	1.0	$0.084^{+2.3}_{-4.6}$	1.25	1.5	$0.167^{+5.5}_{-7.5}$
1.5	0.5	$0.205^{+0.8}_{-3.7}$	1.5	1.0	$0.120^{+1.2}_{-2.9}$	1.5	1.5	$0.159^{+4.1}_{-6.7}$
2.0	0.5	$0.436^{+1.0}_{-3.6}$	2.0	1.0	$0.260^{+1.0}_{-3.6}$	2.0	1.5	$0.211^{+2.0}_{-3.9}$
3.0	0.5	$1.177^{+1.2}_{-3.2}$	3.0	1.0	$0.821^{+0.8}_{-3.7}$	3.0	1.5	$0.589^{+0.9}_{-3.7}$

B.2. Production cross sections at NLO in QCD

Table B.2.: Production cross sections for tWH, depending on C_f and C_V . Obtained with MADGRAPH5_AMC@NLO at NLO in the 5F. The quoted uncertainties on the cross section correspond to scale variations in %.

C_f	C_V	σ [pb]	C_f	C_V	σ [pb]	C_f	C_V	σ [pb]
-3.0	0.5	$0.514^{+2.3}_{-3.0}$	-3.0	1.0	$0.641^{+2.3}_{-2.7}$	-3.0	1.5	$0.783^{+2.1}_{-2.1}$
-2.0	0.5	$0.255^{+2.3}_{-2.8}$	-2.0	1.0	$0.346^{+2.2}_{-2.5}$	-2.0	1.5	$0.457^{+2.1}_{-2.1}$
-1.5	0.5	$0.159^{+2.3}_{-2.8}$	-1.5	1.0	$0.253^{+2.1}_{-2.2}$	-1.5	1.5	$0.329^{+1.9}_{-1.8}$
-1.25	0.5	$0.120^{+2.2}_{-2.5}$	-1.25	1.0	$0.188^{+2.0}_{-2.0}$	-1.25	1.5	$0.275^{+1.9}_{-1.6}$
-1.0	0.5	$0.087^{+2.1}_{-2.3}$	-1.0	1.0	$0.147^{+2.0}_{-1.8}$	-1.0	1.5	$0.224^{+1.9}_{-1.5}$
-0.75	0.5	$0.059^{+2.0}_{-2.1}$	-0.75	1.0	$0.110^{+2.0}_{-1.7}$	-0.75	1.5	$0.180^{+1.8}_{-1.3}$
-0.5	0.5	$0.037^{+1.9}_{-1.8}$	-0.5	1.0	$0.080^{+1.7}_{-1.4}$	-0.5	1.5	$0.141^{+1.6}_{-1.2}$
-0.25	0.5	$0.020^{+1.8}_{-1.3}$	-0.25	1.0	$0.055^{+1.6}_{-1.1}$	-0.25	1.5	$0.108^{+1.6}_{-1.2}$
0.0	0.5	$0.009^{+1.6}_{-1.3}$	0.0	1.0	$0.036^{+1.5}_{-1.2}$	0.0	1.5	$0.081^{+1.5}_{-1.2}$
0.25	0.5	$0.004^{+2.1}_{-2.0}$	0.25	1.0	$0.022^{+1.6}_{-1.5}$	0.25	1.5	$0.059^{+1.5}_{-1.4}$
0.5	0.5	$0.004^{+4.6}_{-6.1}$	0.5	1.0	$0.014^{+2.1}_{-2.0}$	0.5	1.5	$0.043^{+1.8}_{-1.7}$
0.75	0.5	$0.010^{+4.7}_{-6.3}$	0.75	1.0	$0.012^{+3.2}_{-3.9}$	0.75	1.5	$0.033^{+2.1}_{-2.0}$
1.0	0.5	$0.021^{+4.0}_{-5.5}$	1.0	1.0	$0.016^{+4.6}_{-6.1}$	1.0	1.5	$0.028^{+2.8}_{-3.0}$
1.25	0.5	$0.038^{+3.7}_{-5.2}$	1.25	1.0	$0.025^{+4.8}_{-5.4}$	1.25	1.5	$0.029^{+3.6}_{-4.7}$
1.5	0.5	$0.061^{+3.5}_{-4.9}$	1.5	1.0	$0.039^{+4.6}_{-6.3}$	1.5	1.5	$0.035^{+4.6}_{-6.0}$
2.0	0.5	$0.125^{+3.0}_{-4.3}$	2.0	1.0	$0.086^{+4.0}_{-5.5}$	2.0	1.5	$0.065^{+4.8}_{-6.5}$
3.0	0.5	$0.317^{+2.8}_{-4.0}$	3.0	1.0	$0.247^{+3.3}_{-4.6}$	3.0	1.5	$0.193^{+4.0}_{-5.6}$

B. Appendix: Search for single top boson + Higgs boson at 13 TeV

B.3. Recorded and simulated datasets

Table B.3.: Experimental pp collision data exploited in the analysis at 13 TeV. The second column shows the recorded integrated luminosity.

Dataset name
Muon
/SingleMuon/Run2015D-050ct2015-v1/MINIAOD
/SingleMuon/Run2015D-PromptReco-v4/MINIAOD
Electron
/SingleElectron/Run2015D-050ct2015-v1/MINIAOD
/SingleElectron/Run2015D-PromptReco-v4/MINIAOD
total luminosity: 2.198 fb ⁻¹

Table B.4.: Simulated datasets used for the 13 TeV analysis. All samples have been showered with Pythia8. Cross sections have been taken from [200].

Process	ME generator	ME precision	Cross section \times BR, pb
tHq	MadGraph	LO	0.793 (NLO)
tWH	MadGraph	LO	0.147 (NLO)
t \bar{t} H	PowhegV2	NLO	0.2934 (NLO)
t \bar{t}	PowhegV2	NLO	831.76 (NNLO)
single top t -chan., t	Powheg	NLO	70.69 (NLO)
single top t -chan., \bar{t}	Powheg	NLO	44.33 (NLO)
single top tW, t	Powheg	NLO	35.6 (approx. NNLO) [198]
single top tW, \bar{t}	Powheg	NLO	35.6 (approx. NNLO) [198]
W + jets	MadGraph MLM	LO	61526.7 (NNLO)
WW	Pythia8	LO	118.7 (NLO)
WZ	Pythia8	LO	47.13 (NLO)
ZZ	Pythia8	LO	16.523 (NLO)

B.4. BDT output shapes for the SM scenario

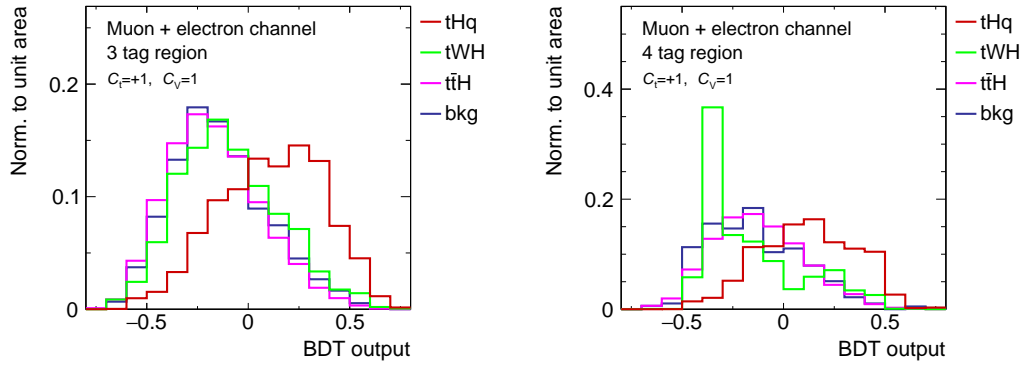


Figure B.1.: Shapes of the classification BDT output for tHq, tWH, ttH and the remaining background processes for the SM scenario. The shape for tWH more background-like than in Fig. 5.21 because tWH and ttH kinematics do not differ much for the SM.

B. Appendix: Search for single top boson + Higgs boson at 13 TeV

B.5. Nuisance pulls for an $s + b$ fit to an Asimov dataset

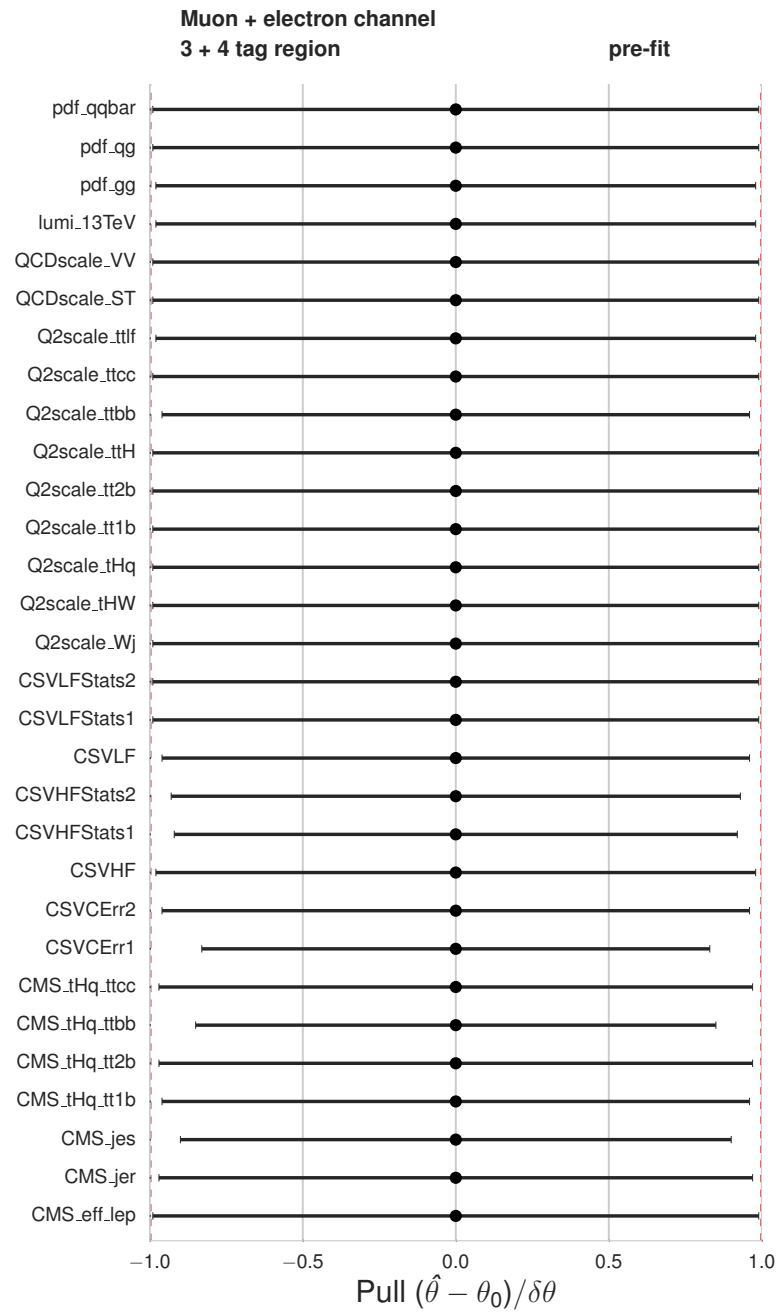


Figure B.2.: Pulls of the single nuisance parameters in the $s + b$ fit to an Asimov data set at 13 TeV.

B.6. Expected limits for all 51 signal scenarios

B.6. Expected limits for all 51 signal scenarios

Table B.5.: Expected 95% confidence level upper limits at 13 TeV on $\sigma/\sigma_{\text{exp}}$ for the different (C_t, C_V) scenarios. The limits are quoted together with the $\pm 1/2\sigma$ uncertainties.

C_t	C_V	Exp. limit	C_t	C_V	Exp. limit	C_t	C_V	Exp. limit
-3	0.5	$3.2^{+1.9/4.8}_{-1.1/1.7}$	-3	1	$2.8^{+1.7/4.2}_{-1.0/1.5}$	-3	1.5	$2.9^{+1.7/4.4}_{-1.0/1.6}$
-2	0.5	$5.9^{+3.5/8.9}_{-2.0/3.2}$	-3	1	$5.0^{+2.9/7.4}_{-1.7/2.7}$	-3	1.5	$4.7^{+2.8/7.1}_{-1.6/2.5}$
-1.5	0.5	$8.7^{+5.2/13.2}_{-3.0/4.7}$	-1.5	1	$7.0^{+4.2/10.6}_{-2.4/3.8}$	-1.5	1.5	$6.4^{+3.8/9.7}_{-2.2/3.4}$
-1.25	0.5	$12^{+7/18}_{-4/6}$	-1.25	1	$8.5^{+5.1/13.0}_{-2.9/4.5}$	-1.25	1.5	$7.5^{+4.5/11.3}_{-2.6/4.0}$
-1.0	0.5	$16^{+10/24}_{-6/9}$	-1.0	1	$11^{+7/17}_{-4/6}$	-1.0	1.5	$8.7^{+5.0/13.0}_{-3.0/4.6}$
-0.75	0.5	$23^{+13/35}_{-8/12}$	-0.75	1	$13^{+8/20}_{-4/7}$	-0.75	1.5	$10^{+6/16}_{-4/5}$
-0.5	0.5	$36^{+21/55}_{-12/19}$	-0.5	1	$18^{+10/27}_{-6/10}$	-0.5	1.5	$12^{+7/19}_{-4/7}$
-0.25	0.5	$59^{+35/90}_{-20/32}$	-0.25	1	$24^{+14/37}_{-8/13}$	-0.25	1.5	$16^{+10/24}_{-6/9}$
0	0.5	$117^{+70/177}_{-40/62}$	0	1	$35^{+21/54}_{-12/19}$	0	1.5	$20^{+12/31}_{-7/11}$
0.25	0.5	$307^{+184/464}_{-105/163}$	0.25	1	$57^{+34/86}_{-19/30}$	0.25	1.5	$28^{+16/42}_{-9/15}$
0.5	0.5	$562^{+336/875}_{-194/301}$	0.5	1	$96^{+57/149}_{-33/51}$	0.5	1.5	$37^{+22/57}_{-13/20}$
0.75	0.5	$247^{+151/385}_{-86/134}$	0.75	1	$157^{+96/244}_{-54/83}$	0.75	1.5	$54^{+32/83}_{-19/29}$
1.0	0.5	$108^{+64/164}_{-37/58}$	1.0	1	$182^{+109/283}_{-63/98}$	1.0	1.5	$80^{+48/125}_{-28/43}$
1.25	0.5	$57^{+34/87}_{-20/31}$	1.25	1	$137^{+82/214}_{-47/73}$	1.25	1.5	$97^{+58/151}_{-33/52}$
1.5	0.5	$34^{+20/52}_{-12/18}$	1.5	1	$81^{+48/122}_{-28/42.9}$	1.5	1.5	$110^{+68/175}_{-38/60}$
2.0	0.5	$16^{+9/24}_{-5/8}$	2.0	1	$34^{+21/52}_{-12/18.3}$	2.0	1.5	$64^{+37/96}_{-22/34}$
3.0	0.5	$6.7^{+3.8/9.9}_{-2.3/3.5}$	3.0	1	$11^{+7/17}_{-4/6}$	3.0	1.5	$21^{+12/31}_{-7/11}$

List of figures

1.1. Mexican hat potential	8
1.2. Feynman diagrams of the most relevant Higgs boson production modes	9
1.3. Feynman diagrams of the most relevant top quark production modes . .	11
1.4. tHq Feynman diagrams at LO	11
1.5. tWH Feynman diagrams at LO	13
1.6. tHb Feynman diagrams at LO	14
1.7. CP mixed state in tX_0q production	15
1.8. Experimental constraints in the κ_t - κ_V plane (CMS)	16
1.9. Experimental constraints on κ_t (ATLAS)	17
2.1. Parton density distribution (cteq611)	21
2.2. Leading order Feynman diagrams of the single top t -channel in the 5F (left) and the 4F (right).	22
2.3. Differential distributions for SM tHq production at 13 TeV	24
2.4. Matrix element to parton shower matching	26
2.5. Neural Network architecture	30
2.6. Illustration of overtraining	33
2.7. Illustration of the CL_S limit computation	36
3.1. Cross section of LHC dipole magnet	40
3.2. LHC accelerator complex	41
3.3. Integrated luminosity in 2012 and 2015	43
3.4. Illustration of the CMS detector	44
3.5. Tracker system	46
3.6. Electromagnetic calorimeter	48
3.7. Quarter of CMS in the r - η plane with hadron calorimeters	49
3.8. Muon system	52
3.9. Drift tubes assembled in the barrel	53
3.10. Drift cell	54
3.11. CMS computing grid	56
3.12. Illustration of jet areas after a clustering with the different jet algorithms	63
3.13. Illustration of the concept of b tagging	65
4.1. Full tHq Feynman diagram	68
4.2. Feynman diagrams of the most relevant background processes	69
4.3. Analysis scheme	70
4.4. Momentum exchange in t -channel in the 4F	75

List of figures

4.5. Pile-up correction	77
4.6. Top p_T correction	78
4.7. Top p_T NNLO calculation	79
4.8. Correction of b tagging efficiencies	79
4.9. Lepton scale factors	80
4.10. Jet η treatment	82
4.11. tHq reconstruction variables	86
4.12. Best NN output under the tHq hypothesis in 2 tag region	87
4.13. $t\bar{t}$ reconstruction variables	88
4.14. Best NN output under the $t\bar{t}$ hypothesis in 2 tag region	89
4.15. Jet assignment under the tHq hypothesis	90
4.16. Effect of m_H on the classification NN	91
4.17. Shape comparison of classification NN input variables	93
4.18. Linear correlation coefficients between the classification NN input variables	95
4.19. Input variables in the 2 tag region (muon channel) at 8 TeV	96
4.20. Input variables in the 3 tag region (muon channel) at 8 TeV	97
4.21. Input variables in the 4 tag region (muon channel) at 8 TeV	98
4.22. Classification NN validation	99
4.23. Shape variations for theoretical uncertainties	101
4.24. Shape variations for experimental uncertainties	103
4.25. Nuisance impacts at 8 TeV	105
4.26. Post-fit NN output distributions at 8 TeV	107
4.27. Post-fit yields for 8 TeV in 3 and 4 tag regions	108
4.28. Post-fit S -over- B distribution for 8 TeV	109
4.29. Comparison of pulls of $t\bar{t}$ + X nuisances for different fits	110
4.30. Post-fit NN output distributions at 8 TeV, including 2 tag	112
4.32. CSV reweighting	113
4.33. Combination	115
5.1. Object kinematics at 8 and 13 TeV – top and Higgs	119
5.2. Object kinematics at 8 and 13 TeV – light quark	120
5.3. Higgs boson and top quark kinematics for the signal processes	121
5.4. Comparison of $t\bar{t}$ modelling	122
5.5. tHq and tWH cross section at 13 TeV in the C_V - C_t plane	122
5.6. Kinematics for the tHq process at $\sqrt{s} = 13$ TeV for different C_f	124
5.7. Lepton related objects at 13 TeV in the 2 tag region	126
5.8. Jet kinematics and MET at 13 TeV in the 2 tag region	127
5.9. Pile-up correction at 13 TeV	127
5.10. CSV reweighting at 13 TeV	128
5.11. Area under ROC curve for tHq reconstruction	128
5.12. Validation of the reconstruction BDTs at 13 TeV	129
5.13. Input variables in the 2 tag region at 13 TeV	131
5.14. Input variables in the 2 tag region at 13 TeV, cont.	132

5.15. Input variables in the 3 tag region at 13 TeV	133
5.16. Input variables in the 3 tag region at 13 TeV, cont.	134
5.17. Input variables in the 4 tag region at 13 TeV	135
5.18. Input variables in the 4 tag region at 13 TeV, cont.	136
5.19. Classification BDT validation	137
5.20. Area under ROC curve for tHq vs. backgrounds	137
5.21. Classification BDT output shapes	137
5.22. Shape variations for uncertainties at 13 TeV	138
5.23. Nuisance impacts at 13 TeV	140
5.24. Pre-fit NN output distributions at 13 TeV	143
5.25. Expected upper limits at 95% C.L.	145
6.1. Luminosity projection	149
A.1. η bias of reconstructed jets	154
A.2. tHq reconstruction variables, cont.	155
A.3. Best NN output under the tHq hypothesis in 3 tag and 4 tag region . . .	156
A.4. Random NN output under the tHq hypothesis in 2 tag region	156
A.5. $t\bar{t}$ reconstruction variables, cont.	157
A.6. Best NN output under the $t\bar{t}$ hypothesis in 3 tag and 4 tag region	158
A.7. Random NN output under the $t\bar{t}$ hypothesis in 2 tag region	158
A.8. Input variables in the 2 tag region (electron channel) at 8 TeV	159
A.9. Input variables in the 3 tag region (electron channel) at 8 TeV	160
A.10. Input variables in the 4 tag region (electron channel) at 8 TeV	161
A.11. Pre-fit NN output distributions at 8 TeV	162
A.12. Pre-fit nuisance pulls at 8 TeV	163
B.1. Classification BDT output shapes for the SM scenario	169
B.2. Pre-fit nuisance impacts at 13 TeV	170

List of tables

1.1. Standard model interactions and properties of their gauge bosons	2
4.1. Event selection	74
4.2. tHq reconstruction NN input variables	84
4.3. $t\bar{t}$ reconstruction NN input variables	85
4.4. Classification NN input variables	94
4.5. Cross section uncertainties from PDF	100
4.6. Final event yields at 8 TeV	108
4.7. Expected and observed 95% confidence level upper limits at 8 TeV	109
4.8. Observed and expected upper limits on $\sigma/\sigma_{C_t=-1}$ for other tHq channels and for the combination of all searches.	114
5.1. tH simulation details at 8 and 13 TeV	118
5.2. Event selection at 13 TeV	125
5.3. Event yields at 13 TeV	142
5.4. Expected and 95% confidence level upper limits at 13 TeV	144
A.1. Recorded datasets at 8 TeV	151
A.2. Simulated datasets at 8 TeV	152
B.1. Production cross sections for tHq at 13 TeV	166
B.2. Production cross sections for tWH at 13 TeV	167
B.3. Recorded datasets at 13 TeV	168
B.4. Simulated datasets at 13 TeV	168
B.5. Expected 95% confidence level upper limits at 13 TeV	171

Bibliography

- [1] **CDF** Collaboration, *Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab*, *Phys. Rev. Lett.* **74** (1995) 2626–2631.
- [2] **DØ** Collaboration, *Observation of the top quark*, *Phys. Rev. Lett.* **74** (1995) 2632–2637.
- [3] **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** (2012) 1–29.
- [4] **CMS** Collaboration, *Observation of a new boson at a mass of 125 gev with the {CMS} experiment at the {LHC}*, *Phys. Lett.* **B716** (2012) 30–61.
- [5] K. A. Olive et al., *Review of Particle Physics*, *Chin. Phys.* **C38** (2014) 090001.
- [6] S. Tomonaga, *On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields*, *Progress of Theoretical Physics* **1** (1946), no. 2 27–42.
- [7] J. Schwinger, *Quantum Electrodynamics. I. A Covariant Formulation*, *Phys. Rev.* **74** (1948) 1439–1461.
- [8] J. Schwinger, *Quantum Electrodynamics. II. Vacuum Polarization and Self-Energy*, *Phys. Rev.* **75** (1949) 651–679.
- [9] J. Schwinger, *Quantum Electrodynamics. III. The Electromagnetic Properties of the Electron—Radiative Corrections to Scattering*, *Phys. Rev.* **76** (1949) 790–817.
- [10] R. P. Feynman, *The Theory of positrons*, *Phys. Rev.* **76** (1949) 749–759.
- [11] R. P. Feynman, *Space-Time Approach to Quantum Electrodynamics*, *Phys. Rev.* **76** (1949) 769–789.
- [12] R. P. Feynman, *Mathematical Formulation of the Quantum Theory of Electromagnetic Interaction*, *Phys. Rev.* **80** (1950) 440–457.
- [13] S. L. Glashow, *Partial Symmetries of Weak Interactions*, *Nucl. Phys.* **22** (1961) 579–588.
- [14] S. Weinberg, *A Model of Leptons*, *Phys. Rev. Lett.* **19** (1967) 1264–1266.
- [15] A. Salam. Originally printed in *Svartholm: Elementary Particle Theory, Proceedings Of The Nobel Symposium Held 1968 At Lerum, Sweden*, Stockholm (1968), 367-377.

Bibliography

- [16] **ATLAS, CDF, CMS and DØ Collaborations**, *First combination of Tevatron and LHC measurements of the top-quark mass*, (2014) [arXiv:1403.4427].
- [17] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, *Phys. Rev. Lett.* **10** (1963) 531–533.
- [18] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, *Prog. Theor. Phys.* **49** (1973) 652–657.
- [19] A. Einstein, *Die Grundlage der allgemeinen Relativitätstheorie*, *Annalen der Physik* **354** (1916), no. 7 769–822.
- [20] **Super-Kamiokande Collaboration**, *Evidence for Oscillation of Atmospheric Neutrinos*, *Phys. Rev. Lett.* **81** (1998) 1562–1567.
- [21] **SNO Collaboration**, *Measurement of the Rate of $\nu_e + d \rightarrow p + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory*, *Phys. Rev. Lett.* **87** (2001) 071301.
- [22] E. Noether, *Invariante Variationsprobleme*, *Gött. Nachr.* 1918 (1918) 235–257.
- [23] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321–323.
- [24] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508–509.
- [25] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, *Phys. Rev. Lett.* **13** (1964) 585–587.
- [26] T. W. B. Kibble, *Symmetry Breaking in Non-Abelian Gauge Theories*, *Phys. Rev.* **155** (1967) 1554–1561.
- [27] G. T. J. Arnison et al., *Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540$ GeV*, *Phys. Lett.* **B122** (1983) 103–116.
- [28] M. Banner et al., *Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the {CERN} pp collider*, *Phys. Lett.* **B122** (1983) 476–485.
- [29] G. T. J. Arnison et al., *Experimental observation of lepton pairs of invariant mass around 95 GeV/c² at the CERN SPS collider*, *Phys. Lett.* **B126** (1983) 398–410.
- [30] M. Thomson, *Modern particle physics*, Cambridge University Press (2013).
- [31] I. v. Vulpen, *The Standard Model Higgs Boson, Part of the Lecture Particle Physics II, UvA Particle Physics Master* (2011-2012).

- [32] F. Halzen and A. D. Martin, *Quarks & Leptons: An Introductory Course in Modern Particle Physics*, Wiley (1984).
- [33] F. Maltoni, K. Paul, T. Stelzer, and S. Willenbrock, *Associated production of Higgs and single top at hadron colliders*, *Phys. Rev.* **D64** (2001) 094023.
- [34] M. Farina, C. Grojean, F. Maltoni, E. Salvioni, and A. Thamm, *Lifting degeneracies in Higgs couplings using single top production in association with a Higgs boson*, *JHEP* **05** (2013) 022, [arXiv:1211.3736].
- [35] F. Demartin, F. Maltoni, K. Mawatari, and M. Zaro, *Higgs production in association with a single top quark at the LHC*, *Eur. Phys. J.* **C75** (2015), no. 6 267.
- [36] J. L. Diaz-Cruz and O. A. Sampayo, *Associated production of the Higgs boson with $t\bar{b}$ at hadron colliders*, *Phys. Lett.* **B276** (1991) 211–213.
- [37] W. J. Stirling and D. J. Summers, *Production of an intermediate-mass Higgs boson in association with a single top quark at LHC and SSC*, *Phys. Lett.* **B283** (1992) 411–415.
- [38] G. Bordes and B. v. Eijk, *On the associate production of a neutral intermediate-mass Higgs boson with a single top quark at the LHC and SSC*, *Phys. Lett.* **B299** (1993) 315–320.
- [39] S. Dawson, *The effective W approximation*, *Nucl. Phys.* **B249** (1985), no. 1 42–60.
- [40] A. Popov, *Search for anomalous higgs boson production in association with single top quarks using the cms detector*, *Université catholique de Louvain, PhD thesis* (2015).
- [41] J. R. Andersen et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, (2013) [arXiv:1307.1347].
- [42] T. M. P. Tait and C. P. Yuan, *Single top quark production as a window to physics beyond the standard model*, *Phys. Rev.* **D63** (2000) 014018.
- [43] J. Ellis, D. S. Hwang, K. Sakurai, and M. Takeuchi, *Disentangling Higgs-Top Couplings in Associated Production*, *JHEP* **04** (2014) 004.
- [44] J. Yue, *Enhanced thj signal at the LHC with $h \rightarrow \gamma\gamma$ decay and CP -violating top-Higgs coupling*, *Phys. Lett.* **B744** (2015) 131–136.
- [45] P. Artoisenet et al., *A framework for Higgs characterisation*, *JHEP* **11** (2013) 043.
- [46] **ATLAS** Collaboration, *Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s} = 7$ and 8 TeV in the ATLAS experiment*, *Eur. Phys. J.* **C76** (2016), no. 1 6.
- [47] **CMS** Collaboration, *Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV*, *Eur. Phys. J.* **C75** (2015), no. 5 212.

Bibliography

- [48] **CMS Collaboration**, *Search for the Associated Production of a Higgs Boson with a Single Top Quark in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV*, [arXiv:1509.0815](https://arxiv.org/abs/1509.0815).
- [49] J. C. Collins, D. E. Soper, and G. Sterman, *Heavy particle production in high-energy hadron collisions*, *Nucl. Phys.* **B263** (1986), no. 1 37–60.
- [50] R. Frederix, *NLO and aMC@NLO*, *Talk given at the 2nd Taipei School on FeynRules – MadGraph for LHC Physics* (Sep 04–08, 2013).
- [51] V. N. Gribov and L. N. Lipatov, *Deep inelastic $e p$ scattering in perturbation theory*, *Sov. J. Nucl. Phys.* **15** (1972) 438–450.
- [52] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and $e^+ e^-$ Annihilation by Perturbation Theory in Quantum Chromodynamics.*, *Sov. Phys. JETP* **46** (1977) 641–653.
- [53] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, *Nucl. Phys.* **B126** (1977) 298.
- [54] **The Durham HepData Collaboration**.
<http://hepdata.cedar.ac.uk/pdf/pdf3.html>. [Online; accessed 11-January-2016].
- [55] T. Kinoshita, *Mass Singularities of Feynman Amplitudes*, *Journal of Mathematical Physics* **3** (1962), no. 4 650–677.
- [56] T. D. Lee and M. Nauenberg, *Degenerate systems and mass singularities*, *Phys. Rev.* **133** (1964) B1549–B1562.
- [57] F. Maltoni, G. Ridolfi, and M. Ubiali, *b -initiated processes at the LHC: a reappraisal*, *JHEP* **07** (2012) 022.
- [58] S. Catani and M. H. Seymour, *A general algorithm for calculating jet cross-sections in NLO QCD*, *Nucl. Phys.* **B485** (1997) 291–419.
- [59] S. Catani, S. Dittmaier, M. H. Seymour, and Z. Trocsanyi, *The dipole formalism for next-to-leading order QCD calculations with massive partons*, *Nucl. Phys.* **B627** (2002) 189–265.
- [60] S. Frixione, Z. Kunszt, and A. Signer, *Three jet cross-sections to next-to-leading order*, *Nucl. Phys.* **B467** (1996) 399–442.
- [61] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, *MadGraph 5: going beyond*, *JHEP* **11** (2011), no. 6.
- [62] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079.

- [63] Z. Xu, D.-H. Zhang, and L. Chang, *Helicity Amplitudes for Multiple Bremsstrahlung in Massless Nonabelian Gauge Theories*, *Nucl. Phys.* **B291** (1987) 392.
- [64] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, *QCD Matrix Elements + Parton Showers*, *JHEP* **2001** (2001), no. 11 063.
- [65] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, *Matching matrix elements and shower evolution for top-pair production in hadronic collisions*, *JHEP* **2007** (2007), no. 01 013.
- [66] R. Frederix and S. Frixione, *Merging meets matching in MC@NLO*, *JHEP* **2012** (2012), no. 12.
- [67] R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, R. Pittau, and P. Torrielli, *Four-lepton production at hadron colliders: aMC@NLO predictions with theoretical uncertainties*, *JHEP* **02** (2012) 099.
- [68] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, *JHEP* **2002** (2002), no. 06 029.
- [69] http://amcatnlo.web.cern.ch/amcatnlo/list_detailed2.htm#showersettings. [Online; accessed 06-January-2016].
- [70] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040.
- [71] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070.
- [72] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043.
- [73] <http://powhegbox.mib.infn.it>. [Online; accessed 06-January-2016].
- [74] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, *Parton fragmentation and string dynamics*, *Physics Reports* **97** (1983), no. 2–3 31–145.
- [75] B. R. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*, *Nucl. Phys.* **B238** (1984) 492.
- [76] T. Gleisberg et al., *Event generation with SHERPA 1.1*, *JHEP* **2009** (2009), no. 02 007.
- [77] J.-C. Winter, F. Krauss, and G. Soff, *A modified cluster-hadronization model*, *The European Physical Journal C* **36** (2004), no. 3 381–395.
- [78] CMS Collaboration, *First Measurement of the Underlying Event Activity at the LHC with $\sqrt{s} = 0.9$ TeV*, *Eur. Phys. J.* **C70** (2010) 555–572.

Bibliography

- [79] **CMS Collaboration**, *Study of the inclusive production of charged pions, kaons, and protons in pp collisions at $\sqrt{s} = 0.9, 2.76, \text{ and } 7 \text{ TeV}$* , *Eur. Phys. J.* **C72** (2012) 2164.
- [80] **CMS Collaboration**, *Study of the underlying event, b-quark fragmentation and hadronization properties in tbart events*, *CMS-PAS-TOP-13-007* (2013).
- [81] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **05** (2006) 026.
- [82] **CMS Collaboration**, *Measurement of the Underlying Event Activity at the LHC with $\sqrt{s} = 7 \text{ TeV}$ and Comparison with $\sqrt{s} = 0.9 \text{ TeV}$* , *JHEP* **09** (2011) 109.
- [83] S. Jadach, J. Kühn, and Z. Was, *TAUOLA - a library of Monte Carlo programs to simulate decays of polarized tau leptons*, *Comput. Phys. Commun.* **64** (1991), no. 2 275–299.
- [84] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, *An Introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159–177.
- [85] P. Skands, S. Carrazza, and J. Rojo, *Tuning PYTHIA 8.1: the Monash 2013 Tune*, *Eur. Phys. J.* **C74** (2014), no. 8 3024.
- [86] M. H. Seymour and A. Siodmok, *Constraining MPI models using σ_{eff} and recent Tevatron and LHC Underlying Event data*, *JHEP* **10** (2013) 113.
- [87] W. McCulloch and W. Pitts, *A logical calculus of the ideas immanent in nervous activity*, *The bulletin of mathematical biophysics* **5** (1943), no. 4 115–133.
- [88] G. Cybenko, *Approximation by superpositions of a sigmoidal function*, *Mathematics of Control, Signals and Systems* **2** (1989), no. 4 303–314.
- [89] C. G. Broyden, *The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations*, *IMA Journal of Applied Mathematics* **6** (1970), no. 1 76–90.
- [90] A. Höcker et al., *TMVA - Toolkit for Multivariate Data Analysis*, (2007) [physics/0703039].
- [91] R. Fletcher, *A new approach to variable metric algorithms*, *The Computer Journal* **13** (1970), no. 3 317–322.
- [92] D. Goldfarb, *A family of variable-metric methods derived by variational means*, *Mathematics of Computations* **24** (1970) 23–36.
- [93] D. F. Shanno, *Conditioning of quasi-Newton methods for function minimization*, *Mathematics of Computations* **24** (1970) 647–656.
- [94] R. Schapire, *The strength of weak learnability*, *Machine Learning* **5** (1990), no. 2 197–227.

- [95] B. P. Roe, H.-J. Yang, and J. Zhu, *Boosted decision trees, a powerful event classifier*, *Proceedings of PHYSTAT05: Statistical Problems in Particle Physics, Astrophysics and Cosmology* (2005).
- [96] R. E. Schapire, *Using output codes to boost multiclass learning problems*, in *Proceedings of the Fourteenth International Conference on Machine Learning, ICML '97*, pp. 313–321, 1997.
- [97] Y. Freund and R. E. Schapire, *A decision-theoretic generalization of on-line learning and an application to boosting*, *Journal of Computer and System Sciences* **55** (1997), no. 1 119–139.
- [98] J. H. Friedman, *Greedy function approximation: A gradient boosting machine*, *The Annals of Statistics* **29** (2001), no. 5 1189–1232.
- [99] Y. Yuan, *Step-Sizes for the Gradient Method*, *AMS/IP Studies in Advanced Mathematics* **42** (1999) 785.
- [100] C. Böser, *Search for the Higgs Boson in WH and tHq Production Modes with the CMS Experiment*, *KIT, PhD thesis, EKP-2015-00032* (2015).
- [101] A. L. Read, *Modified frequentist analysis of search results (the CL_s method)*, *CERN-OPEN-2000-205* (2000).
- [102] A. L. Read, *Presentation of search results: the CL_s technique*, *Journal of Physics G: Nuclear and Particle Physics* **28** (2002), no. 10 2693.
- [103] T. Junk, *Confidence level computation for combining searches with small statistics*, *Nucl. Instr. Meth.* **A434** (1999), no. 2–3 435–443.
- [104] **ATLAS and CMS Collaborations**, *Procedure for the LHC Higgs boson search combination in Summer 2011*, *ATL-PHYS-PUB-2011-11* (2011).
- [105] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J.* **C71** (2011) 1554.
- [106] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, *Ann. Math. Statist.* **9** (03, 1938) 60–62.
- [107] L. Evans and P. Bryant, *LHC Machine*, *Journal of Instrumentation* **3** (2008), no. 08 S08001.
- [108] <http://cds.cern.ch/record/1165534/files/CERN-Brochure-2009-003-Eng.pdf>. [Online; accessed 23-December-2015].
- [109] M. Ferro-Luzzi, W. Herr, and T. Pieloni, *LHC bunch filling schemes for commissioning and initial luminosity optimization*, *CERN-LHC-Project-Note-415* (2008).

Bibliography

- [110] **ALICE** Collaboration, *The ALICE experiment at the CERN LHC*, *Journal of Instrumentation* **3** (2008), no. 08 S08002.
- [111] **ATLAS** Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *Journal of Instrumentation* **3** (2008), no. 08 S08003.
- [112] **LHCb** Collaboration, *The LHCb Detector at the LHC*, *Journal of Instrumentation* **3** (2008), no. 08 S08005.
- [113] **LHCb** Collaboration, *Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \rightarrow J/\psi K^- p$ Decays*, *Phys. Rev. Lett.* **115** (2015) 072001.
- [114] **LHCf** Collaboration, *The LHCf detector at the CERN Large Hadron Collider*, *Journal of Instrumentation* **3** (2008), no. 08 S08006.
- [115] **MoEDAL** Collaboration, *Technical Design Report of the MoEDAL Experiment*, CERN-LHCC-2009-006 (2009).
- [116] **TOTEM** Collaboration, *The TOTEM Experiment at the CERN Large Hadron Collider*, *Journal of Instrumentation* **3** (2008), no. 08 S08007.
- [117] <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>. [Online; accessed 26-December-2015].
- [118] <http://cms.web.cern.ch/news/cms-detector-design>. [Online; accessed 27-December-2015].
- [119] **CMS** Collaboration, *The CMS experiment at the CERN LHC*, *Journal of Instrumentation* **3** (2008), no. 08 S08004.
- [120] **CMS** Collaboration, *CMS physics: Technical design report*, CERN-LHCC-2006-001 (2006).
- [121] **CMS** Collaboration, *CMS Physics Technical Design Report, Volume II: Physics Performance*, *Journal of Physics G: Nuclear and Particle Physics* **34** (2007), no. 6 995.
- [122] **CMS** Collaboration, *Description and performance of track and primary-vertex reconstruction with the CMS tracker*, *Journal of Instrumentation* **9** (2014), no. 10 P10009.
- [123] G. Della Ricca, *Performance of the CMS Electromagnetic Calorimeter at the LHC*, in *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2011 IEEE*, pp. 1476–1479, 2011.
- [124] P. B. Cushman and A. H. Heering, *CMS HCAL Hybrid Photodiode Design and Quality Assurance Stations*, *ICFA Instrum. Bull.* **25** (2002).
- [125] **CMS** Collaboration, *Design, performance, and calibration of CMS forward calorimeter wedges*, *The European Physical Journal C* **53** (2008), no. 1 139–166.

- [126] **CMS Collaboration**, *Energy Response and Longitudinal Shower Profiles Measured in CMS HCAL and Comparison With Geant4*, CMS-NOTE-2006-143 (2007).
- [127] **CMS Collaboration**, *Design, performance, and calibration of CMS hadron endcap calorimeters*, CMS-NOTE-2008-010 (2008).
- [128] C. Anderson, V. Barashko, S. Korenblit, A. Korytov, and G. Mitselmakher, *Effect of Gas Composition on the Performance of Cathode Strip Chambers for the CMS Endcap Muon System*, CMS-NOTE-2004-033 (2004).
- [129] **CMS Collaboration**, *CMS TriDAS project: Technical Design Report, Volume 1: The Trigger Systems*, CERN-LHCC-2000-038 (2000).
- [130] S. Agostinelli et al., *GEANT4: A Simulation toolkit*, *Nucl. Instrum. Meth.* **A506** (2003) 250–303.
- [131] **WLCG Collaboration**. <http://wlcg.web.cern.ch>. [Online; accessed 31-December-2015].
- [132] A. Scheurer, *German contributions to the CMS computing infrastructure*, *Journal of Physics: Conference Series* **219** (2010), no. 6 062064.
- [133] **CMS Collaboration**, *Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET*, CMS-PAS-PFT-09-001 (2009).
- [134] **CMS Collaboration**, *Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector*, CMS-PAS-PFT-10-001 (2010).
- [135] **CMS Collaboration**, *Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV*, CMS-PAS-PFT-10-002 (2010).
- [136] **CMS Collaboration**, *Description and performance of track and primary-vertex reconstruction with the CMS tracker*, *Journal of Instrumentation* **9** (2014), no. 10 P10009.
- [137] H. Kästli, M. Barbero, W. Erdmann, C. Hörmann, R. Horisberger, D. Kotlinski, and B. Meier, *Design and performance of the CMS pixel detector readout chip*, *Nucl. Instr. Meth.* **A565** (2006).
- [138] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, *Nucl. Instr. Meth.* **A262** (1987).
- [139] K. Rose, *Deterministic annealing for clustering, compression, classification, regression, and related optimization problems*, *Proceedings of the IEEE* **86** (1998), no. 11.
- [140] W. Waltenberger, R. Frühwirth, and P. Vanlaer, *Adaptive vertex fitting*, *Journal of Physics G: Nuclear and Particle Physics* **34** (2007), no. 12 N343.

Bibliography

- [141] **CMS Collaboration**, *Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV*, *Journal of Instrumentation* **7** (2012), no. 10 P10002.
- [142] **CMS Collaboration**, *Local Reconstruction in the Muon Detectors*, *CMS-NOTE-2002-043* (2002).
- [143] **CMS Collaboration**, *Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV*, *Journal of Instrumentation* **10** (2015), no. 06 P06005.
- [144] R. Frühwirth, *Track fitting with nonGaussian noise*, *Comput. Phys. Commun.* **100** (1997) 1–16.
- [145] W. Adam, R. Frühwirth, A. Strandlie, and T. Todorov, *Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC*, *Journal of Physics G: Nuclear and Particle Physics* **31** (2005), no. 9 N9.
- [146] M. Cacciari, G. P. Salam, and G. Soyez, *The anti- k_t jet clustering algorithm*, *JHEP* **08** (2008), no. 04 063.
- [147] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, *Longitudinally-invariant k_{\perp} -clustering algorithms for hadron-hadron collisions*, *Nucl. Phys.* **B406** (Feb, 1993) 187–224. 38 p.
- [148] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, *Phys. Rev. D* **48** (Apr, 1993) 3160–3166. 15 p.
- [149] Y. Dokshitzer, G. Leder, S. Moretti, and B. Webber, *Better jet clustering algorithms*, *JHEP* **97** (1997), no. 08 001.
- [150] M. Wobisch and T. Wengler, *Hadronization corrections to jet cross-sections in deep inelastic scattering*, 1998. hep-ph/9907280.
- [151] M. Cacciari and G. P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, *Phys. Lett.* **B641** (2006) 57–61.
- [152] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet User Manual*, *Eur. Phys. J.* **C72** (2012) 1896.
- [153] **CMS Collaboration**, *Determination of jet energy calibration and transverse momentum resolution in CMS*, *Journal of Instrumentation* **6** (2011) 11002.
- [154] http://www-d0.fnal.gov/Run2Physics/top/singletop_observation/b_tagging_graphic.png. [Online; accessed 03-January-2016].
- [155] **CMS Collaboration**, *Identification of b-quark jets with the CMS experiment*, *Journal of Instrumentation* **8** (2013), no. 04 P04013.

- [156] **CMS Collaboration**, *Search for H to $b\bar{b}$ in association with single top quarks as a test of Higgs couplings*, CMS-PAS-HIG-14-015 (2014).
- [157] **CMS Collaboration**, *Measurement of the t -channel single-top-quark production cross section and of the $|V_{tb}|$ CKM matrix element in pp collisions at $\sqrt{s} = 8$ TeV*, JHEP **06** (2014) 090.
- [158] **CMS Collaboration**, *Search for the associated production of the Higgs boson with a top-quark pair*, JHEP **09** (2014) 087.
- [159] **CMS Collaboration**, *Search for H to $b\bar{b}$ in association with single top quarks as a test of Higgs couplings*, CMS-PAS-HIG-14-015 (2014).
- [160] **CMS Collaboration**, *Search for neutral Higgs bosons decaying to tau pairs in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. **B713** (2012) 68–90.
- [161] **CMS Collaboration**.
<https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffs>. [Online, CMS internal; accessed 25-November-2015].
- [162] **CMS Collaboration**.
<https://twiki.cern.ch/twiki/bin/view/CMS/EgammaEARhoCorrection>. [Online, CMS internal; accessed 25-November-2015].
- [163] **CMS Collaboration**. <https://twiki.cern.ch/twiki/bin/view/CMS/MultivariateElectronIdentification>. [Online, CMS internal; accessed 12-September-2014].
- [164] **CMS Collaboration**. <https://twiki.cern.ch/twiki/bin/view/CMS/EcalIsolationCorrection2012Data>. [Online, CMS internal; accessed 12-September-2014].
- [165] **CMS Collaboration**. https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetID#Recommendations_for_8_TeV_data_a. [Online, CMS internal; accessed 30-November-2015].
- [166] **CMS Collaboration**, *Jet energy resolution in CMS at $\sqrt{s} = 7$ TeV*, CMS-PAS-JME-10-014 (2010).
- [167] **CMS Collaboration**, *Jet Transverse Momentum Resolution Measurement using Dijet Events at $\sqrt{s} = 8$ TeV*, CMS Analysis Note AN-13-416 (2013).
- [168] **CMS Collaboration**. https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookMetAnalysis#xy_Shift_Correction. [Online; accessed 02-December-2015].
- [169] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, *Matching matrix elements and shower evolution for top-quark production in hadronic collisions*, JHEP **01** (2007) 013.

Bibliography

- [170] J. Alwall et al., *Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions*, *Eur. Phys. J.* **C53** (2008) 473–500.
- [171] L. Hilser, *Monte-Carlo-Simulation des Single-Top + Jets Prozesses und Untersuchungen zur Notwendigkeit zusätzlicher Jets auf Matrixelement-Level für eine korrekte Beschreibung von Single-Top-Ereignissen*, KIT, Bachelor thesis, EKP-2015-00019 (2015).
- [172] J. Campbell, R. K. Ellis, and R. Röntsch, *Single top production in association with a Z boson at the LHC*, *Phys. Rev.* **D87** (2013) 114006.
- [173] CMS Collaboration, *CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update*, CMS-PAS-LUM-13-001 (2013).
- [174] CMS Collaboration, *Measurement of differential top-quark pair production cross sections in pp collisions at $\sqrt{s} = 7$ TeV*, *Eur. Phys. J.* **C73** (2013), no. 3 2339.
- [175] CMS Collaboration, V. Khachatryan et al., *Measurement of the differential cross section for top quark pair production in pp collisions at $\sqrt{s} = 8$ TeV*, *Eur. Phys. J.* **C75** (2015), no. 11 542.
- [176] CMS Collaboration.
<https://twiki.cern.ch/twiki/bin/view/CMS/TopPtReweighting>. [Online, CMS internal; accessed 25-November-2015].
- [177] M. Czakon, D. Heymes, and A. Mitov, *High-precision differential predictions for top-quark pairs at the LHC*, arXiv:1511.0054.
- [178] CMS Collaboration, *Identification of b-quark jets with the CMS experiment*, *Journal of Instrumentation* **8** (2013) P04013.
- [179] CMS Collaboration.
<https://twiki.cern.ch/twiki/bin/viewauth/CMS/BTagSFMETHODS>. [Online, CMS internal; accessed 20-December-2015].
- [180] CMS Collaboration, *Single muon efficiencies in 2012 data*, CMS-DP-2013-009 (2013).
- [181] CMS Collaboration, *Electron Efficiency Measurement for Top Quark Physics at $\sqrt{s} = 8$ TeV*, CMS Analysis Note AN-12-429 (2012).
- [182] J. Gallicchio and M. D. Schwartz, *Seeing in Color: Jet Superstructure*, *Phys. Rev. Lett.* **105** (2010) 022001.
- [183] J. Gallicchio, J. Huth, M. Kagan, M. D. Schwartz, K. Black, and B. Tweedie, *Multivariate discrimination and the Higgs + W/Z search*, *JHEP* **04** (2011) 069.
- [184] M. Czakon, P. Fiedler, and A. Mitov, *The total top quark pair production cross-section at hadron colliders through $\mathcal{O}(\alpha_S^4)$* , *Phys.Rev.Lett.* **110** (2013) 252004, [arXiv:1303.6254].

- [185] S. Heinemeyer et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, arXiv:1307.1347.
- [186] **ATLAS** Collaboration, *Study of heavy-flavor quarks produced in association with top-quark pairs at $\sqrt{s} = 7$ TeV using the ATLAS detector*, *Phys. Rev.* **D89** (2014), no. 7 072012.
- [187] **CMS** Collaboration, *Measurement of the cross section ratio $\sigma_{t\bar{t}b\bar{b}}/\sigma_{t\bar{t}jj}$ in pp collisions at $\sqrt{s} = 8$ TeV*, *Phys. Lett.* **B746** (2015) 132–153.
- [188] **CMS** Collaboration.
<https://twiki.cern.ch/twiki/bin/viewauth/CMS/JECUncertaintySources>.
[Online, CMS internal; accessed 20-December-2015].
- [189] **CMS** Collaboration, V. Khachatryan et al., *Jet Energy Scale and Resolution in the CMS Experiment, to be submitted to JHEP* (2016).
- [190] **CMS** Collaboration, *Performance of b-tagging at $\sqrt{s} = 8$ TeV in multijet, $t\bar{t}$ and boosted topology events*, CMS-PAS-BTV-13-001 (2013).
- [191] J. Conway, ed., *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding*, CERN, Geneva, Switzerland 17-20 January 2011, 2011.
- [192] L. Moneta et al., *The roostats project*, arXiv:1009.1003.
- [193] N. Faltermann, *Search for Standard Model Higgs boson production in association with a single top quark with the CMS experiment*, KIT, Diploma thesis, EKP-2015-00020 (2015).
- [194] **CMS** Collaboration, *Search for Associated Production of a Single Top Quark and a Higgs Boson in Leptonic Channels*, CMS-PAS-HIG-14-026 (2015).
- [195] **CMS** Collaboration, *Search for associated production of a single top quark and a Higgs boson in events where the Higgs boson decays to two photons at $\sqrt{s} = 8$ TeV*, CMS-PAS-HIG-14-001 (2014).
- [196] **NNPDF** Collaboration, *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040.
- [197] **CMS** Collaboration, *Calibration of the Combined Secondary Vertex b-Tagging discriminant using dileptonic $t\bar{t}$ and Drell-Yan events*, *CMS Analysis Note AN-13-130* (2013).
- [198] N. Kidonakis, *Differential and total cross sections for top pair and single top production*, arXiv:1205.3453.
- [199] **CMS** Collaboration, “Standard Model Cross Sections for CMS at 8 TeV.” <https://twiki.cern.ch/twiki/bin/view/CMS/StandardModelCrossSectionsat8TeV>.
[Online, CMS internal; accessed 12-September-2014].

Bibliography

- [200] **CMS Collaboration.**
<https://twiki.cern.ch/twiki/bin/viewauth/CMS/SummaryTable1G25ns>.
[Online, CMS internal; accessed 14-January-2016].

Acknowledgements

This finally is to thank all the people without whom this work, this thesis, my last three years would not have been successful. You all have contributed in your own way, have supported me, have cheered and backed me up and made these past years unforgettable fun, for which I will be ever grateful.

Most importantly, I want to thank my supervisor, Prof. Dr. Thomas Müller, for embracing me as part of his group and for giving me the opportunity to conduct research in this fascinating field and these exciting times for particle physics. That I was able to spend extended research stays at CERN and in Belgium would have not been possible without his support; he was always there if one needed advice, and I could always count on his experience and way of seeing things.

I would like to express special gratitude also for all the guidance received from Prof. Dr. Ulrich Husemann, who not only is my co-referee, but who has greatly helped in shaping this thesis in its last stages by extensively proofreading it. In addition, his input to analysis specific questions has always been a valuable asset throughout the years, and I believe it is a very good thing the institute hosts both single top + Higgs and ttH analyses, as innumerable examples of synergy between our groups have demonstrated.

I am very thankful to Dr. Thorsten Chwalek, Dr. Jeannine Wagner-Kuhr and Dr. Hauke Held as my direct advisors for always pushing and encouraging me to do more and better and for being the immediate persons to ask if I needed help in any regard. Your navigation is the main reason for why everything went so smoothly. The same credit goes to Dr. Matthias Mozer. Both Thorsten and Matthias have been proofreading the thesis in-depth – thank you!

Thanks are due to Prof. Dr. Günter Quast as well. I share with him the interest in making particle physics graspable for the general public, and this has brought us together many times over the last years. His help and his support in my outreach activities have been substantial, and he was lending an ear to me whenever I felt the need of hearing another opinion and getting useful tips.

Prof. Dr. Fabio Maltoni as the head of CP3 at the Université catholique de Louvain accepted me as an MCnet student and made three of the most incredible months of my PhD studies possible – thank you! Working with him on a phenomenological topic helped me in becoming a more complete physicist, and his out-of-the-box thinking left a deep impression on me. Along the same lines, as big of a “thank you” goes to Federico

Demartin, whom I have worked with on the tWH project during my time at UCL and afterwards. Since then, he is also an invaluable contact for answering and discussing any theory related question I have in mind. I am therefore really lucky and happy to have gotten to know him, and I hope the outcomes of our project will be published soon. In the context of the MCnet studentship I also want to thank Priv.-Doz. Dr. Stefan Gieseke for convincing me of such a project in the first place.

I also would like to thank the members of the CMS generators group, in particular Dr. Josh Bendavid. Since at some point the lines between my analysis and my service work as a Monte-Carlo contact have become blurry, any help that I received from him and from the community on problems related to the technical aspects of MC generators also meant a crucial step forward for this thesis. Josh's work and contributions to MC simulations are precious and have made lots of things easier for so many people within CMS. Moreover, I appreciated a lot the introduction to and instructions on single top Monte-Carlo by Dr. Alberto Orso Maria Iorio!

Leading over to the persons of the CMS collaboration who were just as busy with single top + Higgs as I have been, a big thanks goes to my collaborators Dr. Andrey Popov, Dan Knowlton and Prof. Dr. Ken Bloom. I wish to accentuatedly thank Prof. Dr. Andrea Giammanco, whose winning personality and genius I have had the pleasure to experience not only in the course of the analysis but also during his time as a convener of the TOP group. I have had frequent contact with several members of the latter, and on this occasion want to thank also them for the nice and constructive atmosphere in the group.

Mrs. Bärbel Bräunling as secretary of the IEKP, the admin team for keeping the IT infrastructure going, Dr. Steffen Röcker as great office mate – they all made it very easy to find pleasure in working in 9-23 and deserve that I thank them, as do Nils Faltermann, Dr. Frank Roscher, Shawn Williamson and Michael Ziegler as examples for all the other colleagues at the institute.

This all would not have been possible, of course, if it was not for my parents and grandparents. Without their infinite support stretching way, way beyond the past three years (basically stretching to when I was born), I simply would not have gotten here. The thesis also means the end of a chapter that begun in 2007. During my studies I met guys that have become closest friends. Christoph Böhm and Florian Kassel: you are awesome!

Finally – and as concluding words of this thesis – the people I worked with the closest and longest on single top + Higgs are the people I would like to thank most. These are Dr. Christian Böser and Simon Fink. This trio has been one-of-a-kind, and I enjoyed myself so much working and sharing the office with them. It is safe to say that our contact will not break in the future, and I hope there will be many more hilarious moments that we can share together.

Hiermit versichere ich, die vorliegende Arbeit selbstständig verfasst zu haben. Alle verwendeten Hilfsmittel sind angegeben worden.

Benedikt Maier