Bergische Universität Wuppertal
Fakultät für
Mathematik und Naturwissenschaften
Fachgruppe Physik

Search for flavor-changing neutral-current interactions between the top quark and the Higgs boson in dileptonic same-charge final states with the ATLAS detector

Thesis zur Erlangung des akademischen Grades
Master of Science (M.Sc.)
im Studiengang Physik
der Bergischen Universität Wuppertal

vorgelegt von Marvin Emin Geyik

Erstgutachter: Herr Prof. Dr. Wolfgang Wagner
Zweitgutachter: Herr PD Dr. Frank Ellinghaus

Abgabedatum: 31. März 2021
Überarbeitet: 29. September 2021

Contents

1 Motivation

Since the Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN) was established in 2008, it has continuously helped to refine the current understanding of particle physics. One of its most notable achievements was the discovery of the Higgs boson in 2012, the last remaining particle to complete the Standard Model of particle physics. The discovery was carried out by the two general purpose experiments at the LHC, the ATLAS and the CMS experiment. Following the discovery, numerous additional searches for new, heavy particles that are not described by the Standard Model were carried out, but have not been successful yet. Hence, as an alternative to direct searches, indirect searches for new particles and new types of interactions are of increasing importance in the field of high energy physics. Such indirect searches do not attempt to observe particle resonances. Instead, they search for deviations from Standard Model predictions in data that can be attributed to effects of alternative theoretical models.

The analysis presented in this thesis focuses on the search for flavor-changing neutral-current (FCNC) interactions involving the top quark and the Higgs boson. Such interactions are predicted by the Standard Model. However, their predicted rate of occurrence is far below the current sensitivity of the LHC. Nevertheless, various theoretical models beyond the Standard Model predict higher rates of FCNC processes. FCNC couplings between the top quark, the Higgs boson and either the charm or the up quark are particularly compelling. The unique properties of the involved particles result in coupling strengths that are much higher than for other FCNC sectors.

This particular search is conducted in final states with two leptons that carry the same electric charge. Such final states have the advantage of containing only little contributions from other Standard Model processes. The aim of this thesis is to develop an analysis strategy which is capable of reliably detecting the considered FCNC process.

Prior to detailing approaches to this strategy, the theoretical context of this thesis is presented in Chapter 2. The experimental context is given in Chapter 3. Chapters 4 and 5 then give details on the generation of Monte Carlo samples and the definition of physical objects based on detector signals.

Chapter 6 details the **estimation of background processes** and the definition of custom background process templates. This marks the first step of the analysis strategy. Additionally, an **event selection** is performed with the aim to identify regions that maximize the signal efficiency. Chapter 7 then presents various methods used to **reconstruct signal events** from in these regions. The reconstructed objects are used in Chapter 8 to train artificial neural networks which can reliably **separate signal from background**.

Chapter 9 outlines the systematic uncertainties, which are taken into account in this thesis. Chapter 10 introduces methods of **statistical analysis**. These use the output of the neural networks to either determine the significance of an observed signal or an upper exclusion limit, should no signal be discovered. This is detailed in Chapter 11.

2 Theoretical Context

This chapter provides information on the theoretical context of this thesis. First, the main assumptions and predictions of the Standard Model of particle physics are presented. Afterwards, effective field theories are discussed as possible extensions of the Standard Model. Then the effective field theory that is taken into account in this thesis is described in more detail. More precisely, the expected features of qtH FCNC couplings in multilepton final states are examined.

2.1 The Standard Model

Particle physics has been a field of study for over a century. During this time, intensive research into the behavior of particles at ever increasing energy scales has led to the development of the Standard Model of particle physics (SM). The SM is a quantum field theory, which describes all currently known elementary particles and their interaction via three of the four fundamental forces: The electromagnetic force and the weak and strong nuclear forces. The model does not take into account the gravitational interactions between particles. Figure 2.1 provides an overview on the particles of the SM.

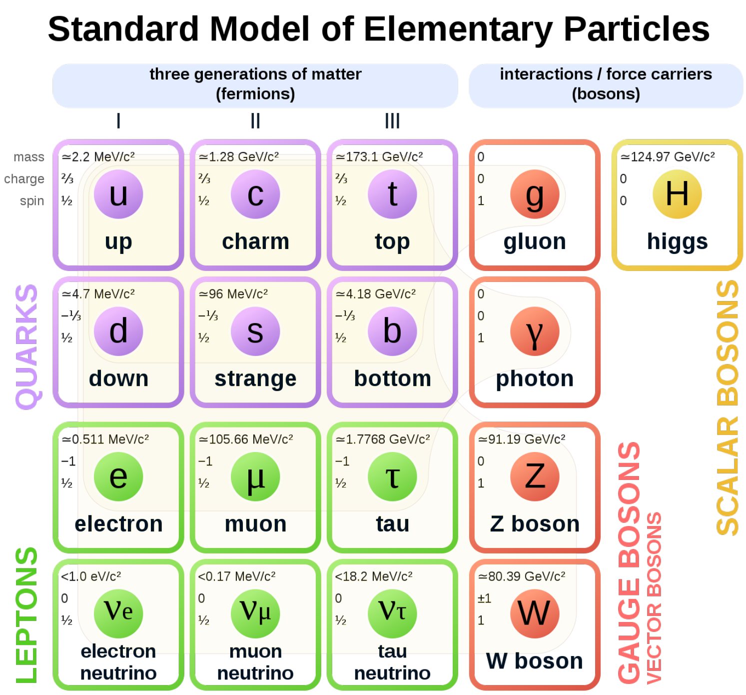


Figure 2.1: Overview of the Standard Model. All elementary particles as well as their mass, charge and spin are depicted. [1]

One of the basic assumptions of the SM is that all elementary particles are fermionic, i.e. of spin 1/2. These particles are split up into two groups: *quarks* and *leptons*. Quarks occur in six different flavors: up (u), down (d), charm (c), strange (s), bottom (b) and top (t). Similarly, leptons occur in six different flavors: electron (e), electron-neutrino (ν_e), muon (μ), muon-neutrino (ν_μ), tau (τ) and tau-neutrino (ν_τ). Both, quarks and leptons, are arranged into three generations. The three generations are depicted as columns of quarks and leptons in Figure 2.1. In general, the mass of fermions increases with their generation (except for neutrinos, which are assumed to be massless in the SM). For every fermion, an anti-fermion exists. This anti-fermion has the same mass as its original counterpart, however all charges and other quantum numbers it carries are inverted.

After this basic proposition, the SM makes the additional assumption that the physical behavior of fermions is invariant under local transformations from the group $SU(3)\otimes SU(2)\otimes U(1)$. For this invariance to hold, the existence of scalar gauge bosons is required. They represent the mediators of the three fundamental forces described by the SM [2].

The $SU(3)$ symmetry yields 8 scalar gauge fields G_a, the so-called gluon fields. Excitations of these fields manifest themselves as gauge bosons, called gluons, which transfer **the strong interaction**. Three charges, called color charges, are associated with the $SU(3)$ symmetry. Quarks are the only fermions that carry a color charge. Hence the strong force only acts on quarks and gluons, as gluons themselves carry color charges (more precisely, a mixture of color and anti-color charges). This is equivalent to saying that gluons only couple to quarks and each other. The $SU(3)$ sector of the SM is generally referred to as Quantum Chromo Dynamics (QCD).

A phenomenon that occurs within QCD is the principle of confinement. It states that only color singlets can be observed as free particles. Hence, quarks can only occur in combinations, so-called **hadrons**. There are two simple possibilities of combining quarks to form a color singlet. One of those is to combine three quarks with different color charges. These states are called **baryons**. The most commonly known baryons are protons and neutrons. They consist of u- and d-quarks and form the nuclei of atoms. The other simple possibility to form a color singlet is to combine a quark of a certain color charge with an anti-quark carrying the corresponding anti-color charge. Particles that are structured in this way are called **mesons**.

Confinement in QCD originates from the fact that the binding energy between two quarks increases with their distance. This fact also leads to the phenomenon of hadronic jets. If the distance between two quarks exceeds a certain point, the binding energy is large enough to create a new quark-anti-quark pair. This process can happen multiple times, depending on the energy of the regarded quarks. This leads to a large number of produced baryons and mesons. The collective of produced particles is called a hadronic jet. Figure 2.2 illustrates the principle of such jets. If a single quark has a high momentum in the laboratory frame, the resulting jet forms a collimated bundle of particles, which retains the original quark's direction of movement. Since the principle of confinement also applies to gluons, hadronic jets can originate from those as well [3].

Similarly to the $SU(3)$ symmetry, the $SU(2)\otimes U(1)$ symmetry yields a number of bosonic gauge fields. The $SU(2)$ symmetry yields three gauge bosons, the W_1-, W_2- and W_3 bosons, and the $U(1)$ symmetry yields one gauge boson, the B boson. At this point, an additional scalar field ϕ, the Higgs field, is introduced. Through spontaneous symmetry breaking and

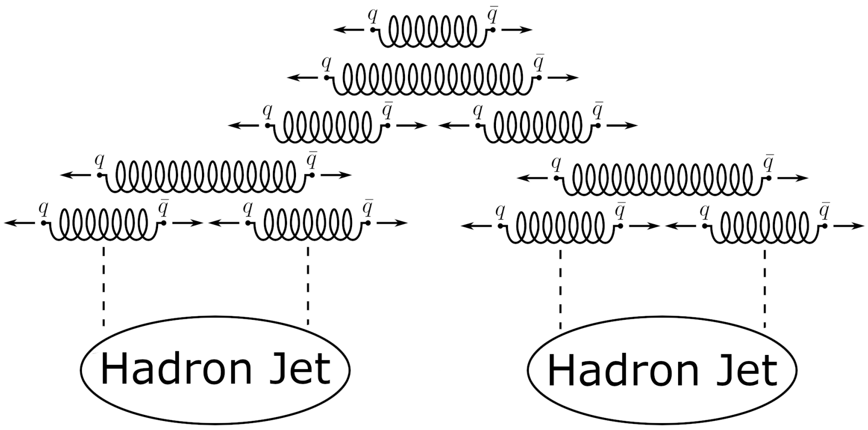


Figure 2.2: The principle of confinement. The binding energy between two quarks with opposite momenta increases, until a new quark-anti-quark pair is created. This repeated process results in two hadronic jets. The coiled lines represent gluons. [4]

the Higgs mechanism, this field causes a mixing of the above gauge bosons. The W_1- and the W_2 bosons mix into the massive W^+ and W^- bosons. The W_3 and the B boson mix into the massive Z^0 boson and the massless photon γ.
The photon itself is neutral but transfers the **electromagnetic interaction** and couples to electric charge. Charged leptons (e, μ, τ) carry an electric charge of -1, neutral leptons (ν_e, ν_μ, ν_τ) don't carry any electric charge. Quarks carry non-integer electric charges: Up-type quarks (u, c, t) carry a charge of +2/3, down-type quarks (d, s, b) carry a charge of -1/3.
The three massive bosons, $W^\pm$ and Z^0, transfer the **weak interaction**. They couple to all fermions and to each other. Additionally, the $W^\pm$ bosons carry an electric charge of ±1 and hence couple to the photon [5].
Excitations of the Higgs field manifest themselves as Higgs bosons H. The Higgs boson couples to the massive $W^\pm$ and Z^0 bosons. Additionally, it couples to all massive fermions via a Yukawa coupling. In summary, the Higgs boson couples to all massive particles of the SM. The strength of this coupling increases with the mass of the respective particle [6].
This concludes the summary of the fundamental assumptions and some important consequences of the SM. This thesis mainly focuses on the quark sector of the SM. More precisely, it focuses on specific flavor-changing processes, which are not predicted by the SM. Hence, it is useful to take a closer look at flavor-changing processes, which are predicted by the SM.

2.1.1 Quark mixing in the SM

As mentioned in the previous section, the W^+ and W^- boson couple to all fermions, including quarks. Moreover, both W bosons carry an electric charge. Thus, the only way they can couple to quarks, while conserving the overall electric charge, is to couple to one up-type quark and one down-type quark. This means that the coupling of quarks to the W bosons is intrinsically flavor-changing. Now the question arises, whether the W bosons only couple to up and down quarks of the same generation, or a cross-generational change of flavor is possible. The SM predicts the existence of both. This phenomenon is called quark mixing. However, the coupling strength among an up- and a down-type quark of the same generation to the W bosons is the strongest. The coupling strength of an up-type quark a and a down-type quark b is proportional to $|V_{ab}|^2$, where V_{ab} is an entry in the Cabibbo-Kobayashi-Maskawa (CKM) matrix V_{UD}. The structure of the matrix and the predicted values of its entries are as follows [7].

$$V_{UD} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 0.97427 & 0.22534 & 0.00351 \\ 0.22520 & 0.97344 & 0.04120 \\ 0.00867 & 0.04040 & 0.99915 \end{pmatrix} \tag{2.1}$$

One consequence of the values of the CKM-matrix is that the top quark almost exclusively decays into a W boson and a b-quark, as $|V_{tb}|$ is especially close to 1. In the SM, quark mixing only occurs via W bosons. Hence, in leading order perturbation theory all flavor-changing processes also change the charge of a given quark. Nonetheless, processes which change the flavor of a quark, but not its charge, are subject of several current research efforts [8].

2.1.2 FCNC processes in and beyond the Standard Model

Processes, which change the flavor of a quark but preserve its charge are called flavor-changing neutral-current (FCNC) processes [9]. Such processes are forbidden in the SM at leading order perturbation theory. Furthermore, FCNC processes at higher orders are suppressed due to the GIM mechanism [10]. The suppression is particularly strong for FCNC processes involving the top quark. Here, branching ratios are predicted to lie in the range from 10^{-16} to 10^{-11}. Figure 2.3 shows an example of a SM FCNC process. A top quark decays into a neutral Higgs boson and either a u- or a c-quark.

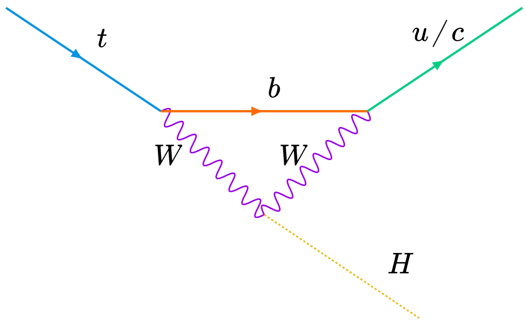


Figure 2.3: $t \to Hu$ FCNC process in the SM. The top quark decays into a Higgs boson and an up or a charm quark via an intermediate down-type quark state (here: b-quark).

Due to their strong suppression, FCNC processes as predicted in the SM, are not accessible with current experimental methods. There are, however, several theories beyond the SM, which predict higher rates of FCNC processes. These higher rates are generally introduced in two ways. The first is to postulate the existence of new particles, which causes the probability for FCNC processes to increase. The second way is to introduce direct couplings between two quarks of the same charge and a neutral SM boson in effective field theories [8]. Some theories predict rates of FCNC processes, which can be probed with current experiments. Figure 2.4 shows the current upper limits on the branching ratios of FCNC processes involving the top quark as well as theoretical predictions for the SM and other theories, that predict higher rates of FCNC processes.

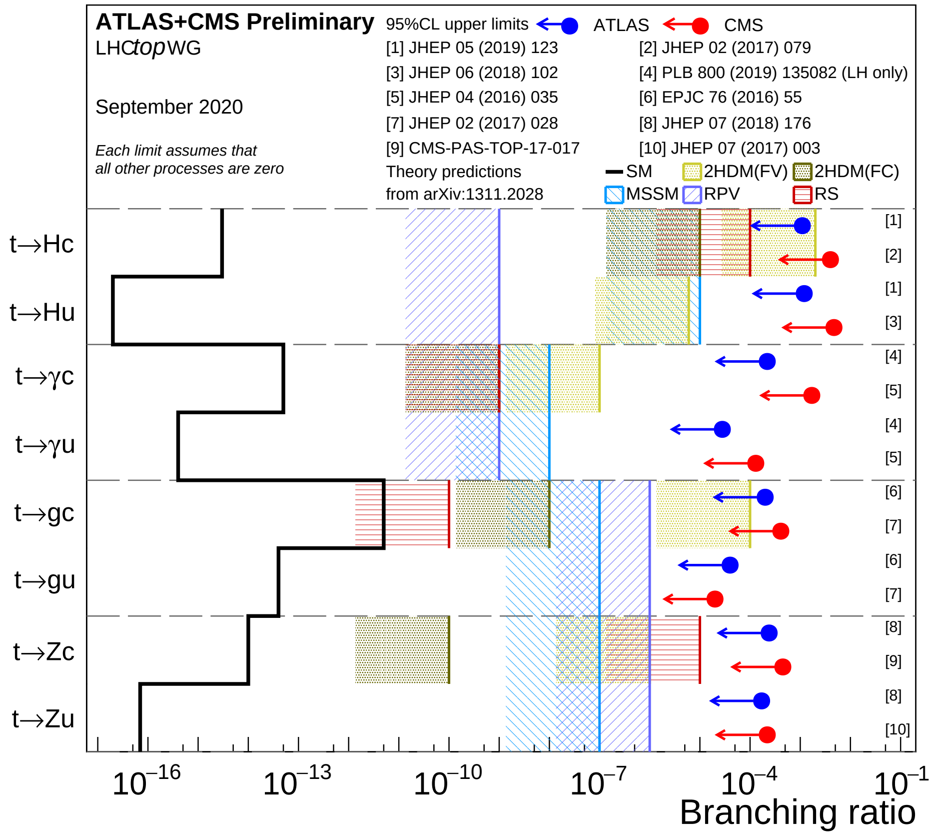


Figure 2.4: Summary of branching ratios for FCNC processes involving the top quark. For every process, the theoretical predictions of the SM (black line) as well as some theories beyond the SM (blue, yellow, red areas), are plotted. Additionally, the current upper limits on the branching ratios set by the ATLAS (blue arrows) and the CMS (red arrows) experiment are shown [11].

2.1.3 Special properties of the top quark and the Higgs boson

The FCNC coupling that will be of most interest in this thesis is a hypothetical coupling between the top quark, the Higgs boson and a different up-type quark. Both, the top quark and the Higgs boson, have characteristics that set them apart from other SM particles. The top quark is by far the heaviest particle in the SM. Its experimentally measured mass amounts to (172.69 ± 0.48) GeV [12]. This high mass leads to a very short lifetime of only approximately 5×10^{-25} s [13]. Its lifetime is not sufficient for the top quark to hadronize, as other quarks would. It almost exclusively decays into a W boson and a b-quark. This very specific signature allows for in-depth studies of the top quark at particle collider experiments.
The Higgs boson is, with a mass of (125.09 ± 0.24) GeV [14], the second heaviest particle in the SM. As mentioned, its coupling strength to other particles is proportional to their mass. Hence, it couples the strongest to the top quark. This large coupling strength propagates to theories beyond the SM, which predict direct FCNC couplings between the top quark and the Higgs boson, on average leading to higher predicted branching ratios (cf. Figure 2.4). Generally, theories beyond the SM predict effects at high energy scales. Due to the high masses of the top quark and the Higgs boson, processes involving both particles are already at a very high energy scale. Thus, processes that involve the top quark and the Higgs boson

are exceptionally sensitive to theories beyond the SM and can be studied intensively due to their characteristic event signatures.

2.2 Effective Field Theories for FCNC processes involving the top quark

Effective field theories (EFT) provide a way to search for physics beyond the SM without relying on a specific alternative theoretical model to test for. The new physical processes are generally introduced through additions to the SM Lagrangian $\mathcal{L}_{\text{SM}}$. For FCNC processes involving a top quark the standard EFT Lagrangian $\mathcal{L}_{\text{EFT}}$ consists of a sum of dimension-six operators $\mathcal{O}^{qt}_{uX}$ (for left-handed couplings) and $\mathcal{O}^{tq}_{uX}$ (for right-handed couplings), which are scaled by a corresponding, dimensionless Wilson coefficient C^{qt}_{uX}/ C^{tq}_{uX}:

$$\mathcal{L}_{\text{EFT}} = \mathcal{L}_{\text{SM}} + \sum_{q,X} \left(\frac{C^{qt}_{uX}}{\Lambda^2} \mathcal{O}^{qt}_{uX} + \frac{C^{tq}_{uX}}{\Lambda^2} \mathcal{O}^{tq}_{uX} \right). \tag{2.2}$$

Here, Λ is the energy scale, up to which the EFT describes the FCNC process. $q = u, c$ refers to the two up-type quarks, which are different from the top quark. $X = B, W, G, \phi$ denotes the SM fields, via which FCNC couplings are introduced in the EFT[1] [15]. The above EFT is implemented in the TopFCNC [16] model in the FeynRules 2.0 [17] framework. This framework is used by the MadGraph5_aMC@NLO [18] Monte Carlo (MC) generator to produce next-to leading order (NLO) calculations in QCD of the postulated processes. Hence, it is possible to conduct searches for selected FCNC processes.

2.3 Top-Higgs FCNC processes

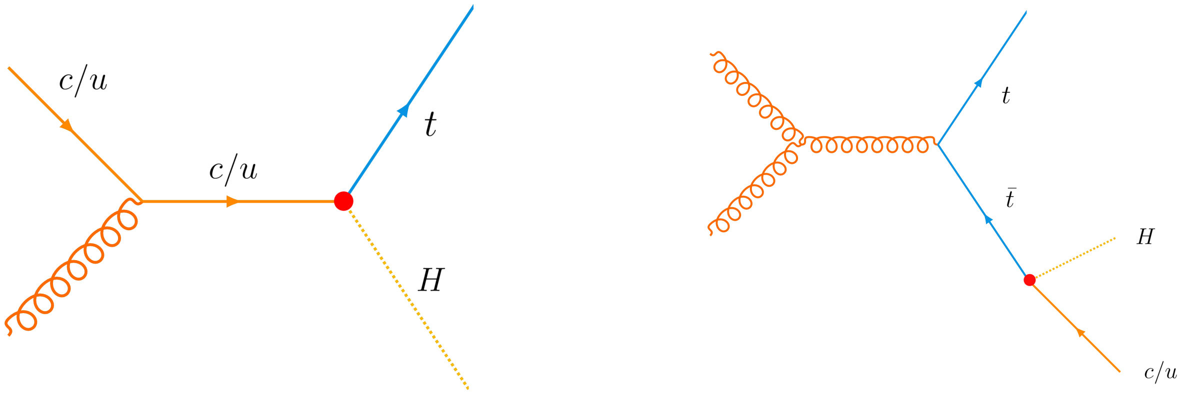
As stated previously, this thesis focuses on FCNC couplings between the top quark, the Higgs boson and a light up-type quark. Hence, only four terms in the above EFT Lagrangian are important here: The ones parameterizing the left- and right-handed utH and ctH couplings. The four corresponding operators as well as the respective Wilson coefficients are listed in Table 2.1.

Table 2.1: FCNC EFT operators and Wilson coefficients relavant for this thesis. $\bar{q}_L$ denote the left-handed doublet quark fields, q_R the right-handed singlet fields. $\tilde{\phi}$ refers to the complex conjugate of the Higgs field.

Coupling	dim-6 operator	Wilson coeff.
left-handed utH	$(\phi^\dagger\phi)(\bar{u}_L t_R)\tilde{\phi}$	$C^{ut}_{u\phi}$
right-handed utH	$(\phi^\dagger\phi)(\bar{t}_L u_R)\tilde{\phi}$	$C^{tu}_{u\phi}$
left-handed ctH	$(\phi^\dagger\phi)(\bar{c}_L t_R)\tilde{\phi}$	$C^{ct}_{u\phi}$
right-handed ctH	$(\phi^\dagger\phi)(\bar{t}_L c_R)\tilde{\phi}$	$C^{tc}_{u\phi}$

[1] Couplings to the B and the W field result in FCNC couplings to the photon and the Z boson.

The operators considered here can be studied in two main channels: The production channel $q \to tH$ and the decay channel $t \to qH$, where $q = u/c$. Feynman diagrams of both channels are depicted in Figure 2.5. This thesis, however, only takes into account the decay channel.



(a) production channel (b) decay channel

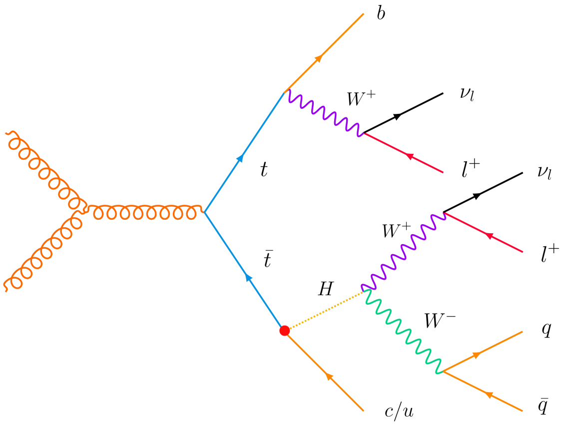
Figure 2.5: Feynman diagrams of the $q \to tH$ production (a) and $t \to qH$ decay (b) channel of tHq FCNC couplings. The production process consists of a single top quark being produced in association with a Higgs boson via the introduced FCNC vertex. The decay process consists of a $t\bar{t}$-pair production, with either the top quark or the top anti-quark decaying into a Higgs boson and a different up-type quark.

FCNC couplings between the Higgs boson and the top quark have been subject to analyses in the past, as Figure 2.4 suggests. The publication in Ref. [19] combines the results of several analyses, which analyzed qtH FCNC couplings in various Higgs-boson decay modes and final states. The limits obtained from that combination are 0.12 % and 0.11 % for the branching ratios $\mathcal{B}(t \to Hu)$ and $\mathcal{B}(t \to Hc)$, respectively. These limits will be referred to as the current upper limits of the two FCNC decays throughout this thesis.

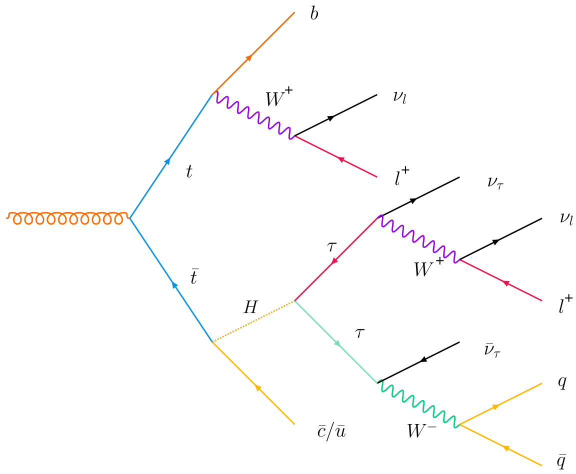
2.4 Top-Higgs FCNC processes in dileptonic same-charge final states

In the remainder of this thesis, the term lepton will refer to either an electron or a muon. Whenever a τ lepton is discussed, it will be named explicitly. The reason for this is that in current experiments, only electrons and muons are detected directly. τ leptons generally decay into other SM particles, which can then be detected.
Depending on the decay products of the Higgs boson and the top quark and their subsequent decay modes, there are several possible final states, which can emerge from the above processes. This thesis focuses on final states, which include exactly two leptons of the same charge. Such final states are referred to as dileptonic same-charge (2ℓSS) final states. For a 2ℓSS final state to be possible, the non-FCNC top quark needs to decay leptonically (i.e. into a b-quark and a W boson, which subsequently decays into a lepton and a neutrino). Hence, exactly one lepton of the same charge needs to originate from the decay of the Higgs boson. One now needs to consider various decay modes of the Higgs boson, which can produce (at least) one lepton. Based on the decay branching ratios of the Higgs boson [20], contribu-

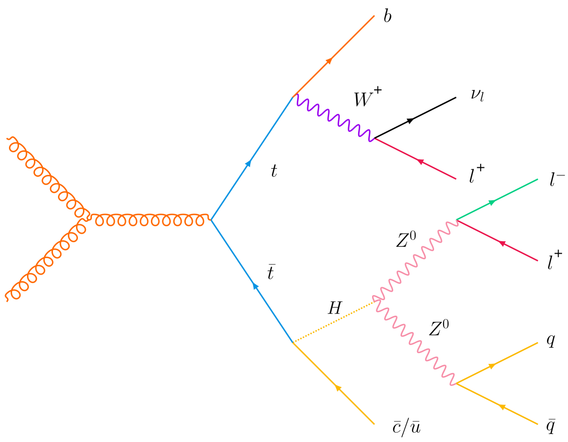
tions are expected from three decay channels: $H \to WW^*$, $H \to ZZ^*$ and $H \to \tau\bar{\tau}$. Feynman diagrams for possible contributions by each of the channels are shown in Figure 2.6. All channels require one of the Higgs boson's decay-products to decay leptonically and one to decay hadronically to contribute to the 2ℓSS final state. It should be noted that there are no events from the $H \to ZZ^*$ channel, which produce a true 2ℓSS final state. However, some events from this channel might be relevant in the analysis, if one produced lepton from the leptonically decaying Z boson is not detected or reconstructed incorrectly.



(a) $H \to WW^*$ channel



(b) $H \to \tau\bar{\tau}$ channel

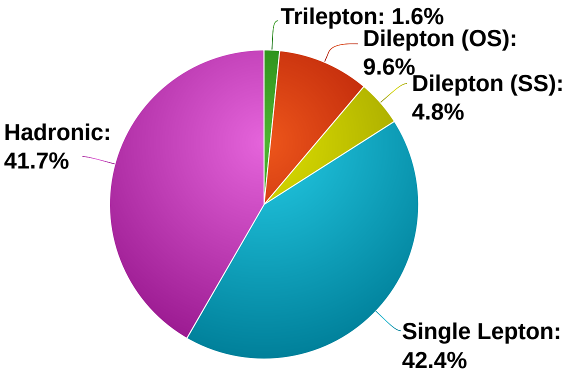


(c) $H \to ZZ^*$ channel

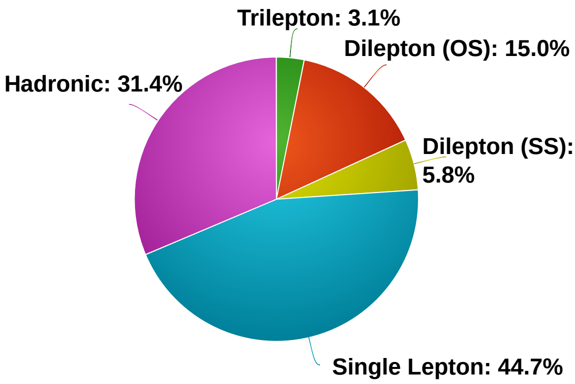
Figure 2.6: Feynman diagrams of possible dileptonic same-charge final states for the qtH vertex in the decay channel. Final states mainly originate form three Higgs decay channels: $H \to WW^*$ (a), $H \to \tau\bar{\tau}$ (b) and $H \to ZZ^*$ (c).

The branching ratios of the Higgs boson indicate that most signal events in the 2ℓSS final state fall into the $H \to WW^*$ channel, as it has a branching ratio of 21.4 %. The second largest contribution is expected from the $H \to \tau\bar{\tau}$ channel with a branching ratio of 6.3 % and only a minor contribution from the $H \to ZZ^*$ channel with a branching ratio of 2.6 % [21].
Based on the known decay branching ratios of W bosons and τ-leptons, the fraction of events in the $H \to WW^*$ and the $H \to \tau\bar{\tau}$ channel that actually produces a 2ℓSS final state is studied. Pie charts showing exact percentages of different final states, categorized by the number of

final state leptons, are depicted in Figure 2.7 for both decay channels. In the calculations it was assumed that the non-FCNC top quark decays into a b-quark and a W boson. However, all possible decay modes of the W boson were considered.



(a) $H \to WW^*$ channel



(b) $H \to \tau\bar{\tau}$ channel

Figure 2.7: Pie charts showing the occurence of various final states, categorized by the number of leptons and their charge relation, for qtH FCNC events. The fractions of final states are shown for the $H \to WW^*$ channel (a) and the $H \to \tau\bar{\tau}$ channel (b). All possible decay modes of the W boson from the non-FCNC top-quark decay are taken into account. Dilepton (SS/OS) denotes the final states with two leptons of the same or opposite charge.

Only 4.8 % of all $H \to WW^*$ events and 5.8 % of all $H \to \tau\bar{\tau}$ events contribute to the 2ℓSS final state. This corresponds to only a small fraction of all possible events. The fraction becomes even smaller, if all decay modes of the Higgs boson are taken into account. Figure 2.8 shows the fraction of different multilepton final states with respect to all Higgs-boson decay modes. Again, all decay modes of the W boson from the non-FCNC top-quark decay are taken into account.

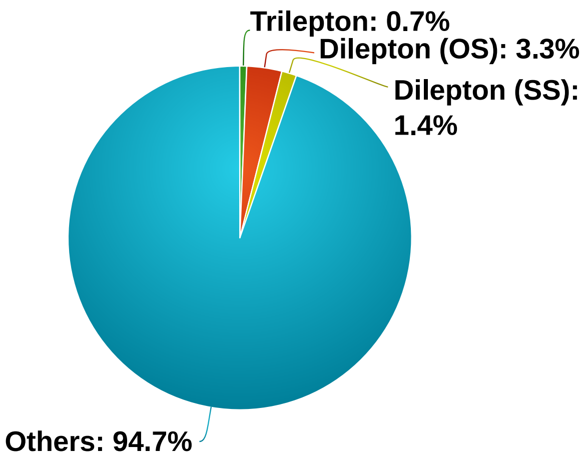
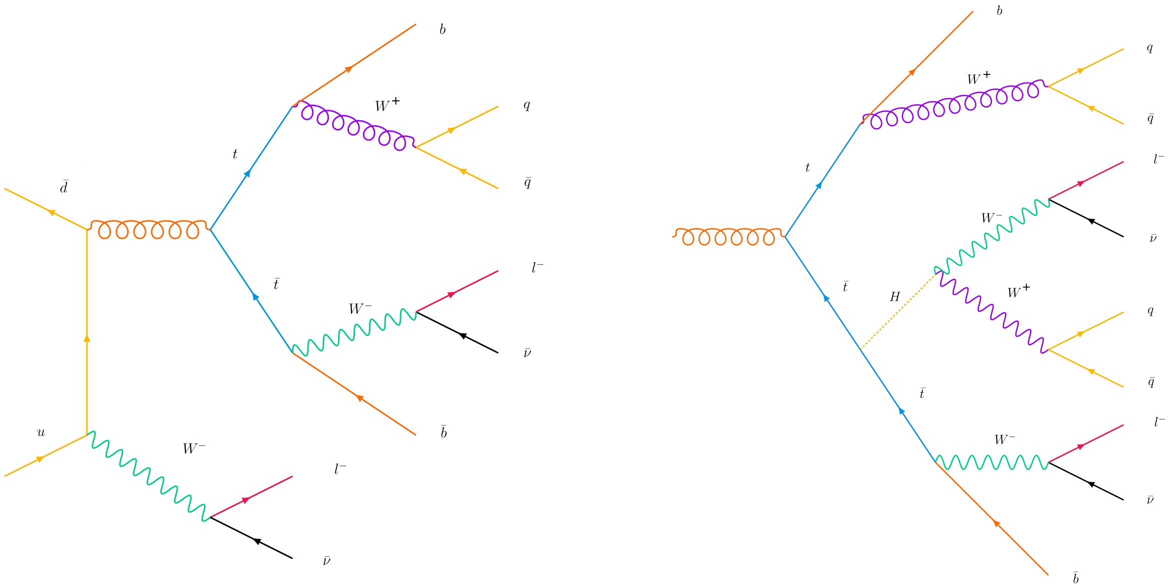


Figure 2.8: Pie chart of the occurence of various final states, categorized by the number of leptons and their charge relation, for qtH FCNC events. The fractions are given for all decay modes of the Higgs boson.

The chart shows that only 1.4 % of all $t \to Hu/Hc$ events contribute to the 2ℓSS final state. At the same time, however, only few SM processes produce 2ℓSS final states. Thus, it can be expected that a significant sensitivity to qtH FCNC processes can still be obtained. Chapter 6 discusses in detail, how the event selection is optimized to maximize the sensitivity. Prior to such discussions, however, different categories of expected background contributions need to be taken into account.

Prompt and non-prompt lepton background sources

Prompt leptons are defined as leptons, which are produced as part of the hard primary interaction at the primary vertex, in most cases via the decay of an SM vector or scalar boson. In this thesis, also leptons which originate from final state radiation or prompt quarkonium decays are defined as prompt. Generally, processes that contribute to 2ℓSS final states have relatively low cross sections [22]. Some examples include certain decay channels in the production of a $t\bar{t}$-pair in association with a W-, Z- or Higgs boson (in the following referred to as $t\bar{t}W/Z/H$). Figure 2.9 shows examplary decay modes of the $t\bar{t}W$ and the $t\bar{t}H$ process, which produce a signature of final state particles that is almost identical to the signature of the signal process (cf. Figure 2.6).



(a) $t\bar{t}W$ (b) $t\bar{t}H$

Figure 2.9: Feynman diagrams of decay modes of the $t\bar{t}W$ (a) and the $t\bar{t}H$ (b) processes, which result in a 2ℓSS final state, similar to the signal final state.

In addition to these and other, less significant prompt lepton background sources, one can expect a considerable contribution from processes, which do not produce final states with two prompt leptons of the same charge. Examples for such processes include the production of a $t\bar{t}$-pair or the production of a single Z boson in association with an unspecified number of hadronic jets (Z+jets). They can contribute to the 2ℓSS final state through so-called non-prompt leptons. In this thesis, non-prompt leptons are defined as all leptons that do not originate from a prompt source, as defined above. Exemplary non-prompt lepton sources include the hadronic jets from high-energy (mostly b-) quarks as well as the conversion of photons to electrons inside detector material. Furthermore, it is possible for electrons to be reconstructed with an inverted charge. Processes involving non-prompt leptons are generally not modeled well in MC simulations. Hence, a strategy for the estimation of the non-prompt lepton background is vital for this analysis. Prior to such a strategy a clear classification of leptons is necessary.

2.5 The IFF Truth Classifier

As stated, there is an expected, non-negligible contribution from non-prompt leptons to 2ℓSS final states. Hence, a consistent and useful categorization of different lepton sources is necessary in this analysis. Such a classification is provided by the Isolation and Fake Forum's (IFF) Truth Classifier [23]. The IFF Truth Classifier is a tool, which assigns leptons in MC simulations to a certain IFF truth class, based on their true origin in the simulation. In total, there are eleven classes defined by the tool. Their names and definitions are listed in the following.

Prompt (isolated) electron	Electrons originating from a prompt lepton source (W, Z, H, etc.) as well as final state radiation and prompt quarkonium decays
Charge-flip isolated electron	Prompt electrons, for which the charge at reconstruction-level is different from the charge of its first non-GEANT[2] mother-particle
Prompt muon	Muon originating from a prompt lepton source (W, Z, H, etc.) as well as final state radiation and prompt quarkonium decays
Prompt photon conversion	Electron from the conversion of a prompt photon through interactions near the interaction point (inner conversion) or with the detector material (external conversion)
Electron from muon	Electron originating from the decay of a prompt muon
Tau decay	Electron or muon produced in a hadronic jet, which originates from a tau lepton
B-hadron decay	Electron or muon produced in a hadronic jet, which originates from a b-quark
C-hadron decay	Electron or muon produced in a hadronic jet, which originates from a c-quark

[2]GEANT [24] is a toolkit, by which the propagation of particles through matter (here: electrons through detector material) is simulated.

Light flavor decay	Electron or muon produced in a hadronic jet, which originates from a *u*-, *d*- or *s*-quark
Known Unknown	Electron or muon, for which no classification could be found due to lacking MC information
Unknown	Electron or muon, for which no classification could be found for unknown reasons

2.6 The template fit method

With a consistent lepton classification, an approach for the estimation of non-prompt lepton background sources needs to be found. The MC modeling of events including non-prompt leptons has been found to be suboptimal in previous analyses [22]. More precisely, current MC models give a reasonably accurate prediction for the kinematic behavior of non-prompt leptons, but fail to estimate the rate of their occurrence [25]. This mis-modeling needs to be corrected for. This correction is achieved via a so-called template fit. For this, certain kinematic regions which are enriched in relevant non-prompt lepton background events, need to be identified. Normalization factors for non-prompt lepton background processes are determined via a fit to observed data in these regions. This approach to the estimation of non-prompt lepton background sources is called the template fit method. A correct choice of regions, which are enriched in non-prompt lepton events, is crucial for the applicability of this method. Chapter 6 details the selection of these regions.

3 Experimental Context

This chapter describes the experiment, which recorded the experimental data used in this thesis. More specifically, the Large Hadron Collider (LHC) and the A Toroidal LHC ApparatuS (ATLAS) experiment are presented.

3.1 The Large Hadron Collider

The LHC at the European Organization for Nuclear Research (CERN) is the largest particle accelerator in the world as well as the most powerful one in terms of its achievable particle energy. It is located at the French-Swiss border near Geneva and measures a circumference of 26.7 km. It is used to accelerate protons and heavy ions (mainly lead), but for this thesis only the acceleration of protons is of importance. This acceleration process is split into several steps. First, protons are produced by ionizing hydrogen atoms. They are then accelerated in several pre-accelerators to a particle energy of 450 GeV, before being injected into the LHC, where the final acceleration takes place. Currently, protons are accelerated to 6.5 TeV in antiparallel directions in two parallel beampipes. The LHC is designed, however, to support an acceleration of up to 7 TeV [26]. The entire accelerator complex is schematically shown in Figure 3.1.

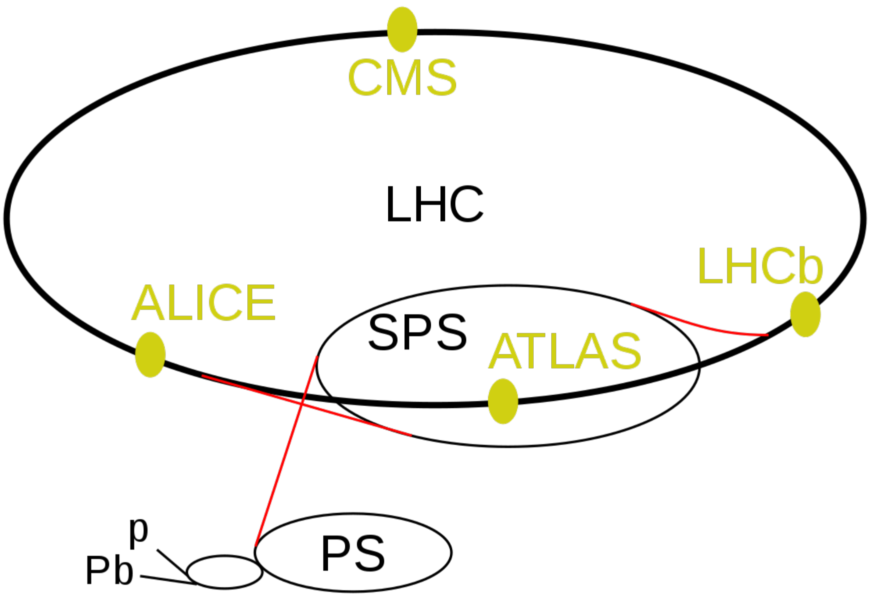


Figure 3.1: Schematic of the accelerator chain leading up to the LHC. The four experiments stationed at the LHC are depicted as dots around the LHC [27].

The protons are fed into the LHC in bunches of on average 11.5×10^{10} protons, instead of forming a continuous beam. These bunches are kept on a quasi-circular track via dipole-magnets. The LHC consists of numerous alternating sections of straight and dipole-segments, which bend the track. In total, there are 1232 dipole magnets used in the LHC [28].
The actual acceleration of the proton bunches is achieved through radiofreqency (RF) cavities. RF cavities are metallic chambers, in which an oscillating electromagnetic (EM) field is generated. The cavities are formed such that they resonate at a frequency of 400 MHz (one oscillation per 2.5 ns). The incoming bunches need to be timed at an integer multiple of the oscillation time. The bunch-frequency at the LHC amounts to 40 MHz, corresponding to one bunch every 25 ns (or one bunch per 10 RF cavity oscillations). The EM-field in an RF cavity is tuned such that protons, which arrive at the cavity with the nominal energy, experience the nominal EM-field. Protons with an energy that is slightly lower are bent stronger by the dipole magnets and thus have to cover a smaller track length, resulting in an earlier arrival. They experience a higher EM-field and are accelerated in the beam direction, while protons which arrive later (i.e. have a slightly higher energy) are decelerated. This effectively results in a focusing of the bunches in the beam direction. The focusing of proton bunches in the plane perpendicular to the beam direction (i.e. the transverse plane) is also important and is achieved with quadrupole magnets. Since quadrupole magnets focus the beam in one direction and defocus in the perpendicular direction, they are always installed in pairs. One quadrupole magnet focuses the beam in a given direction (e.g. x-direction) and the subsequent one focuses the beam in the perpendicular y-direction [29].
The two focused beams in their respective beampipes cross at four distinct places along the circumference of the LHC, where protons from two bunches are made to collide with each other at a nominal center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$. At each of these points, detector systems are placed to study the particles that are created in the collisions. The four detector systems are called ATLAS, ALICE, LHCb and CMS (cf. Figure 3.1). Bunch-crossings happen every 25 ns, based on the frequency of the RF cavities. The number of bunch crossings is, however, not identical to the number of collisions, as a varying number of protons per bunch can collide. This phenomenon is known as pile-up. These collisions can take place throughout the entire section of the beampipe, in which the two bunches cross[1]. Figure 3.2 shows an exemplary bunch crossing with high pile-up.

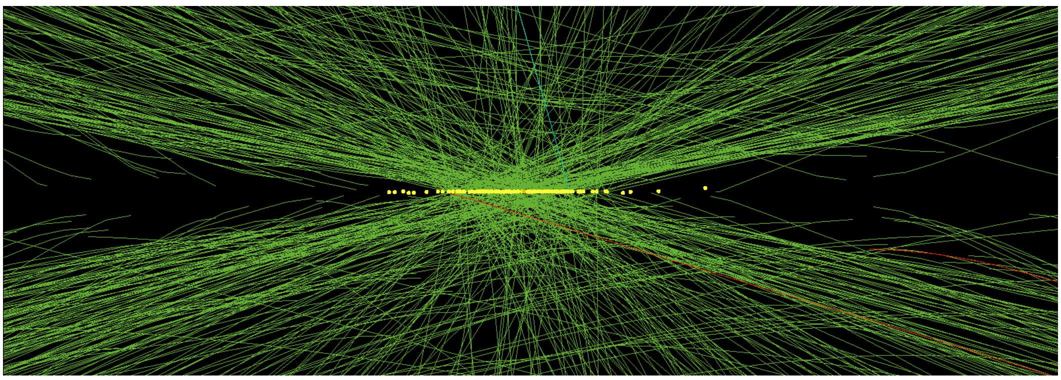


Figure 3.2: High pile-up bunch crossing in the CMS detector. The green lines correspond to reconstructed particle tracks, the yellow dots correspond to reconstructed points of collision. There are 78 reconstructed collision points in this event. [30]

[1]The bunches in the LHC have an average length of 12.5 cm [28]

For physical analyses, the number of collision events which occur in a detector is of importance, rather than the number of bunch crossings. A common measure of the number of events per time and area in a collider is the luminosity. The instantaneous luminosity is defined by [31]:

$$L = \frac{n \cdot {N_B}^2 \cdot f}{4\pi \cdot \sigma_x \cdot \sigma_y} \tag{3.1}$$

Here n denotes the total number of bunches in the collider, N_B the number of protons per bunch, f the revolution frequency and σ_x/σ_y the width of the Gaussian beam profile in the x- and the y-direction. If a particle collider like the LHC is operated for a given amount of time, the total number of expected events for a given process is important. To obtain this, the instantaneous luminosity can be integrated by time, which yields the integrated luminosity $L_{\text{int}} = \int L dt$. It corresponds to the proportionality factor between the expected number of events for a given process and its cross section. A common unit for the integrated luminosity is 1 fb^{-1}. The integrated luminosity of the full Run 2 dataset recorded by the ATLAS detector and usable for physics amounts to 139 fb^{-1}. This means that a process with a cross section of 1 fb is expected to occur 139 times within the dataset.

3.2 The ATLAS experiment

The ATLAS experiment is a general-purpose detector experiment at the LHC. This means that it aims at studying all possible proton-proton interactions, without focusing on one specific type of interaction. The ATLAS detector is the largest of the LHC's four detectors. Prior to discussing its various detector components, a general overview over the coordinate system that is used in the ATLAS detector is given. The detector is cylindrically shaped, with the cylinder axis being equal to the beampipe, in which the bunches cross. The center of the ATLAS detector is called the interaction point [32]. The ATLAS detector uses a Cartesian coordinate system, which is centered around the interaction point. The z-axis goes clockwise along the beamline, the y-axis points up-wards (meaning in opposite direction from gravitational force) and the x-axis is oriented perpendicular to both, towards the center of the LHC ring, such that x, y and z form a right-handed coordinate system. To describe the direction of particles in the detector, polar coordinates are generally more practical. For a given vector the angle ϕ is defined as the angle between the vector's projection on the x-y-plane and the x-axis. The angle θ is defined as the angle between the vector and the z-axis. An often used alternative to the angle θ is the pseudorapidity η. It is defined as a function of θ in Equation 3.2. For high-energy particles, the pseudorapidity is approximately equal to the rapidity y, which is also defined in Equation 3.2 [33].

$$\eta := -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] \overset{E\approx|\vec{p}|}{\approx} \frac{1}{2}\ln\left(\frac{E + c\,p_L}{E - c\,p_L}\right) =: y \tag{3.2}$$

Here E is the energy of a the particle in question. $\vec{p}$ is its momentum and p_L the longitudinal component (i.e. component parallel to the beamline) of its momentum.

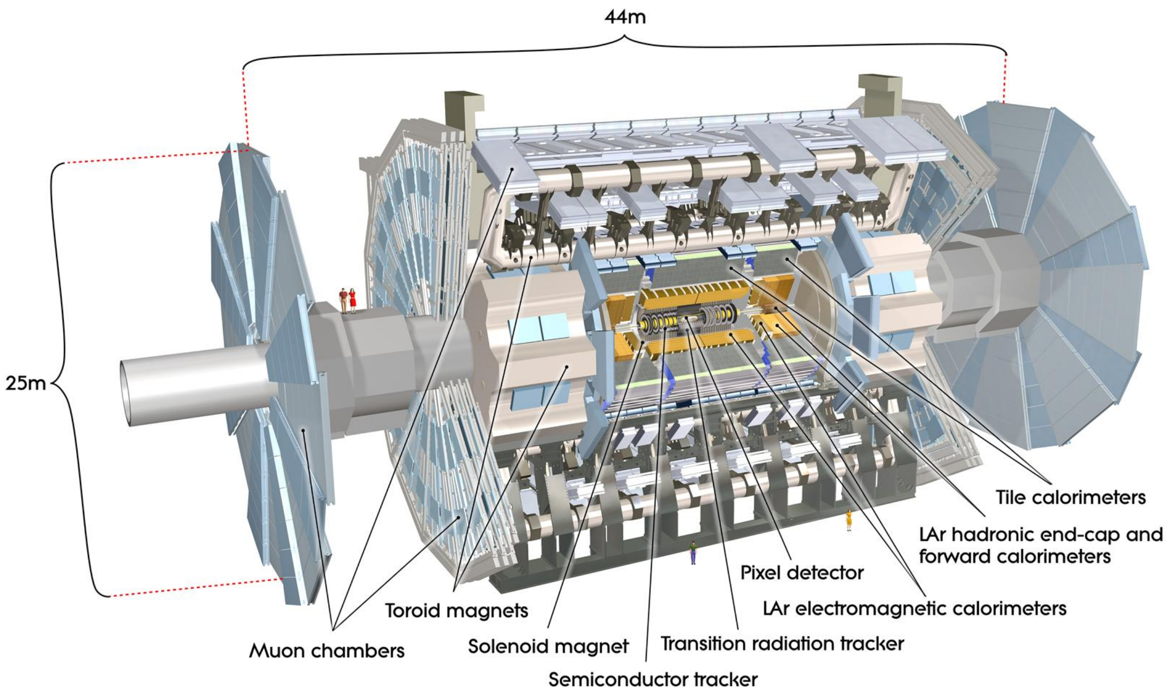


Figure 3.3: 3D-model of the ATLAS detector. All sub-detectors as well as all both magnet systems of the detector are shown, alongside the scale of the entire detector [34].

Now the structure of the ATLAS detector in terms of its various detection systems will be presented. The ATLAS detector is divided into several sub-detector systems, which are structured in cylindrical layers around the beampipe. All sub-systems are split into a barrel and an endcap region. These correspond to the cylindrical face, which wraps around the beamline, and the two faces at the end of the cylinder respectively. The entire ATLAS detector is depicted in Figure 3.3 [35].

The **inner detector** consists of a pixel detector, a silicon microstrip detector and a transition radiation tracker. It measures the tracks of charged leptons and hadrons. The pixel detector is located closest to the beam pipe, followed by the silicon microstrip detector. Both detectors are based on silicon semiconductors, in which large depletion zones are created through application of an external voltage. If a charged particle moves through such a depletion zone, a measurable number of electron-hole pairs is created. The corresponding electrical signal is used for the reconstruction of the respective particle. The transition radiation tracker works based on a different principle. It is composed of a large number of drift-tubes, filled with a Xe, O_2 and CO_2 gas mixture. The walls of the tubes are coated with a conductive material, to which a high voltage is applied. A tungsten wire is located at their center. The tubes are interleaved with radiator material, which emits X-Rays, when a charged particle passes through it. These X-Rays (in addition to the charged particle) ionize the gas in the surrounding drift-tubes, which can be measured in a change of the applied voltage. The entire inner detector is immersed in a solenoid magnetic field, parallel to the beamline, with a strength of 2 T. The field causes the track of charged particles to bend, which allows for a measurement of their momentum and the sign of their charge [36].

The detector layer surrounding the inner detector is the **calorimeter**. There are two types of calorimeters in the ATLAS detector: The EM and the hadronic calorimeter. The EM calorimeter is structured in interchanging layers of lead or stainless steel and liquid Argon (LAr) around the interaction point. Its purpose is to measure the energy of non-hadronic, electromagnetically interacting particles (mainly electrons and photons). The stainless steel/ lead layers cause the particles to loose most of their energy by creating particle showers. The showers are detected in the LAr layers and from the structure of their tracks an estimate for the energy of a given particle can be given [37]. Charged hadrons also leave tracks in the EM calorimeter, however due to their high mass their EM radiation length is too large for their full energy to be exhausted inside the EM calorimeter. Instead, hadrons loose most of their energy in the hadronic calorimeter via hadronic showers. The hadronic calorimeter is structured similarly to the EM calorimeter. However, in the barrel region only stainless steel is used as energy-absorbing material. The layers of stainless steel are generally thicker than for the EM calorimeter, to increase the cross section for an interaction of hadrons with nuclei of the material. A scintillating material is used as an active medium in the barrel region to detect the shower components and estimate the total energy of the hadron. In the endcap region, copper is used for the main energy absorption, with LAr as an active material [38].

The final detector component is the **muon spectrometer** in the outermost layer of the ATLAS detector. It is immersed in a toroidal magnetic field. This magnetic field allows for great precision in the reconstruction of detected muons. The muon detection itself is based on high precision drift chambers in the barrel region and cathode strip chambers in the endcap region. Both detection mechanisms are designed to give high-resolution tracking information on muons, which pass through them. In addition to the two high-precision detection systems, there are thin-gap chambers and resistive plate chambers, which are used for triggering [39].

4 Monte Carlo event generation

The simulation of proton-proton collisions and the subsequent particle interactions in the LHC is a fundamental aspect of this analysis. The data obtained by the ATLAS experiment is analyzed by comparing it to MC simulations of various SM processes as well as the signal process. These simulations model the hard scattering process itself, the hadronization of produced quarks or gluons and the propagation of secondary particles through the detector.

Because of their complexity and differences in the physical processes that are being modeled, the simulations of these three sub-components are decoupled from each other. For the modeling of the hard scattering process, the corresponding matrix element is calculated up to a first or second order in perturbation theory. The MC generators SHERPA v2.2.1 [40], POWHEG-BOX v2 [41] and MG5_aMC@NLO [18] are used in this thesis to perform these calculations. For the hadronization, either of the two parton shower generators PYTHIA v8.210 [42] and SHERPA v2.2.1 is used. Lastly, the interaction with the detector material is simulated using GEANT4 [24].

In this analysis, various MC samples are used, each corresponding to a certain SM or FCNC process. A list of all used samples, together with the event generators deployed for their generation, is given in Table 4.1. It should be noted that in addition to the samples listed, a small number of tWZ-events are expected to contribute to the relevant background processes in this analysis. There were, however, issues in the processing of this specific sample. Studies in preparation of this analysis have shown that tWZ-events contribute less than 0.7 % to the total number of events in regions of interest (defined in Chapter 6). Hence, the lack of this sample in the studies presented in this thesis should not alter the result in any significant way.

Table 4.1: The MC samples used in this analysis. The generators used for the calculation of the matrix element as well as for the simulation of the parton shower and hadronization are given for each sample.

Process	Matrix element	Showering + hadronization
FCNC decay process	POWHEG-BOX v2	PYTHIA v8.210
$t\bar{t}$ production	POWHEG-BOX v2	PYTHIA v8.210
Single top-quark production	POWHEG-BOX v2	PYTHIA v8.210
W+jets production	SHERPA v2.2.1	SHERPA v2.2.1
Z+jets production	SHERPA v2.2.1	SHERPA v2.2.1
Diboson production	POWHEG-BOX v2	PYTHIA v8.210
$t\bar{t}W$ production	MG5_aMC@NLO	PYTHIA v8.210
$t\bar{t}Z$ production	MG5_aMC@NLO	PYTHIA v8.210
$t\bar{t}Z$ production	POWHEG-BOX v2	PYTHIA v8.210

4.1 Decay signal samples

The decay signal samples for the utH and the ctH FCNC couplings are of particular importance in this analysis. As described in Chapter 2, the EFT used for modeling the FCNC processes contains a left-handed and a right-handed coupling for both channels. Previous studies on top-quark-Higgs-boson FCNC couplings using the same theoretical model have shown that the chirality of the coupling has little to no effect on the kinematics of $t \to Hq$ decay processes [43]. Since this thesis focuses exclusively on such decay processes, the signal samples that are considered only model the left-handed FCNC coupling. Results which are obtained using these samples are then assumed to be indicative of the left- and the right-handed couplings. Based on these insights, no distinction between the chiralities of the FCNC processes will be made in this thesis. The two signal processes considered will be referred to as the utH and the ctH signal process.

The signal samples stand out in the simulation approach used to generate them. As opposed to the other considered samples, the signal samples were created using the fast simulation technique AFII [44]. The full GEANT4 simulation of the interaction of secondary particles with detector material is very time- and CPU-consuming. The AFII simulation technique reduces the overall computation time by only simulating interactions in the inner detector with GEANT4 and using the faster FASTCALOSIM [45] simulation for the calorimeters. This has the advantage of allowing for large statistics, while only slightly reducing the accuracy of the modeling.

5 Object definition

This chapter describes the definition of different physical objects at reconstruction level based on certain signal patterns at detector level. For this thesis, electrons, muons, hadronic jets, in particular jets originating from a b-quark, and missing transverse momentum are important. In addition to the definition of those objects, an overlap removal, which avoids double counting of certain detector signals, is discussed.

5.1 Electrons

Electrons are reconstructed from topological clusters detected in the EM calorimeter, which are matched to tracks in the inner detector [46]. A veto is performed on tracks in the muon spectrometer or clusters in the hadronic calorimeter, that are matched to the cluster in the EM-calorimeter. Additionally, electron candidates are required to have a minimum p_T of 10 GeV and $|\eta| < 2.47$. Furthermore, the transition region between the barrel and the end-cap EM calorimeter $1.37 < |\eta| < 1.52$ is excluded. To ensure that reconstructed electrons originate from the primary vertex, requirements on the transverse and longitudinal impact parameters d_0^{BL} and z_0^{BL} are made. These read: $|d_0^{\mathrm{BL}}| < 5\,\mathrm{mm}$ and $|z_0^{\mathrm{BL}} \sin(\theta)| < 0.5\,\mathrm{mm}$. As a general criterion on the quality of electron quality, a likelihood-based discriminant is used, which optimizes the identification efficiency and the rejection of fake electrons[1] [47].
In addition to these definitions, a tool called the Prompt Lepton Isolation Tool [48] is used to reduce the number of non-prompt electrons being reconstructed. It is based on two different Boosted Decision Trees (BDTs): The Prompt Lepton Veto (PLV) for leptons with a p_T that is larger than 12 GeV and the low-p_T PLV for leptons with a p_T that is smaller than 12 GeV [49]. The tool is specifically tuned to reject non-prompt leptons from hadronic jets generated by a b- or c-quark. Two working points are defined for the tool: PLVTight and PLVLoose. The PLVTight working point has a higher rejection of fake electrons at the cost of a lower acceptance of real electrons. It is used as a working point in this thesis. After reconstruction, the measured energy of the lepton is scaled based on previous MC-to-data comparisons [50].

5.2 Muons

Muons are reconstructed from tracks in the inner detector that are matched to tracks in the muon spectrometer. They are required to surpass a p_T of 10 GeV and have a pseudorapidity of $|\eta| < 2.5$. Similarly to electrons, requirements on the impact longitudinal and transverse

[1] A fake electron is different from a non-prompt electron, as defined in Chapter 2, in that the detector signals used for its reconstruction are not created by a truth level electron. Instead, another truth object (e.g. jet/ muon) creates a detector signature that matches that of an electron.

parameters of muons are made to ensure their origin is the primary vertex. These requirements are $|d_0^{\mathrm{BL}}| < 3\,\mathrm{mm}$ and $|z_0^{\mathrm{BL}} \sin(\theta)| < 0.5\,\mathrm{mm}$. Requirements on the quality of muon candidates are made based on their p_T and η. Analogous to electrons, the Prompt Lepton Isolation Tool is used to reject a large number of non-prompt muons. Again, two BDTs (also called PLV and low-p_T PLV) are trained and two working points PLVTight and PLVLoose are defined. Also for muons the PLVTight working point is used. For muons, the reconstructed energy is scaled via a similar method as for electrons [51].

5.3 Jets

Jets are reconstructed from objects defined by the particle flow (PFlow) algorithm, which combines information from the inner detector and both calorimeters [52]. These PFlow objects are used as input for the anti-k_t algorithm with a radius parameter of $R = 0.4$, which performs the final reconstruction of jets [53]. Additionally, a minimum p_T of 30 GeV and $|\eta| < 4.5$ are required for each jet. The reconstructed jets are corrected in terms of their energy scale (JES) [54] and resolution (JER) [55]. This correction is based on previous MC-to-data comparisons. To reduce the number of jets that originate from pile-up and are erroneously assigned to a given event, a constraint on the Jet Vertex Tagger (JVT) discriminant is imposed. A minimum JVT discriminant of 0.5 for jets with $20\,\mathrm{GeV} < p_T < 60\,\mathrm{GeV}$ and $|\eta| < 2.4$ is required. For jets with a $2.5 < |\eta| < 4.5$ the forward JVT (fJVT) discriminant is used. For jets in this η-range with $20\,\mathrm{GeV} < p_T < 120\,\mathrm{GeV}$ fJVT< 0.4 is required [56].

5.4 b-tagged jets

Jets that originate from a b-quark (b-jets) have certain properties that distinguish them from jets originating from lighter quarks or gluons. The reconstructed origin of the jet generally lies further away from the primary vertex of a given event. This is due to the generally longer lifetime of hadrons containing b-quarks, as b-quarks can only decay via the weak interaction. Based on these and other properties, efforts are made to identify b-jets at reconstruction level. This process is called b-tagging. This analysis uses the `DL1r` algorithm for b-tagging [57]. The algorithm produces three output values, representing the probabilities for a given jet to have originated from a c-, b- or a light quark. These probabilities are then combined into a single score. This score is used as a continuous discriminant between b-tagged jets and non-b-tagged jets. Four different operating points are defined, each with a different b-tagging efficiency: 60 %, 70 %, 77 % and 85 %. The 70 % operating point is used as a default in this analysis.

5.5 Missing transverse momentum

The missing transverse momentum $\vec{p}_{T,\mathrm{miss}}$ is defined as the negative vectorial sum of the p_T of all reconstructed objects in addition to a soft term. This soft term consists of tracks in the inner detector, which are not assigned to any reconstructed object, but still point to the primary vertex. The absolute value of the $\vec{p}_{T,\mathrm{miss}}$ is denoted $\not{E}_T$ and referred to as missing transverse energy. The two components of $\vec{p}_{T,\mathrm{miss}}$ are denoted $\not{E}_x$ and $\not{E}_y$.

5.6 Overlap removal

To avoid the assignment of certain tracks or energy depositions to more than one reconstructed object, a set of overlap removal rules is defined [58]. These remove certain objects at reconstruction level in the following order. Objects are not considered in further steps, once they are removed.

- Electrons with an associated track in the inner detector that is shared with that of a muon
- Muons that are only detected within the inner detector and the calorimeter (usually due to lack of detector material in the muon spectrometer) and share a track in the inner detector with an electron
- Jets within an angular distance[2] of $\Delta R = 0.2$ of an electron
- Electrons within $\Delta R = 0.4$ of a jet
- Jets with less than three associated tracks in the inner detector within $\Delta R = 0.2$ of a muon
- Jets with less than three associated tracks in the inner detector if a muon track in the inner detector is ghost-associated[3] to it
- Muons within $\Delta R = 0.4$ of a jet

[2] $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ is a common measure of the angular distance between two objects in the detector.

[3] Ghost association is the assignment of certain objects to a jet, if the objects were not a priori included in the reconstruction of the track.

6 Background estimation and event selection

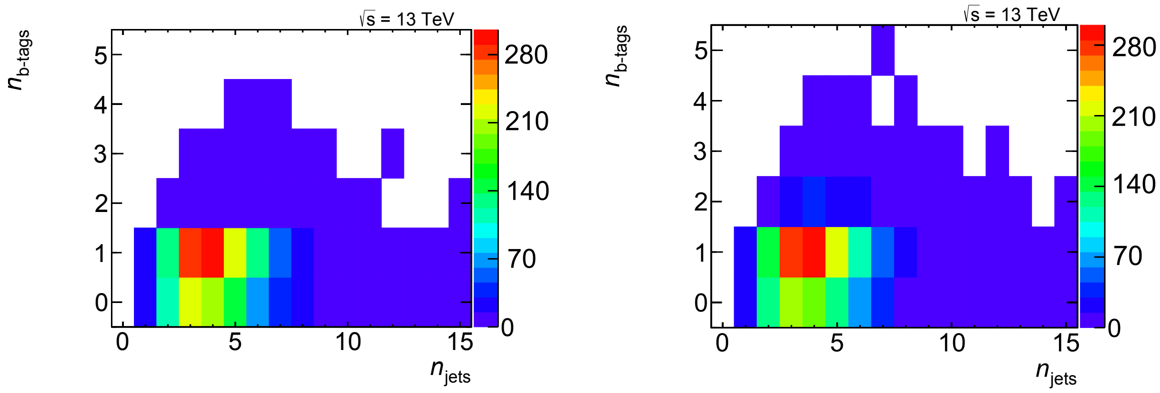
In order to study the FCNC qtH coupling as efficiently as possible, it is necessary to identify regions enriched in events originating from the signal process. Furthermore, certain background processes require especially careful attention, as the modeling of their occurrence has been observed to be suboptimal [19]. In particular, events with non-prompt leptons originating from b- or c-decays ("heavy-flavor" decays) and electrons with a misidentified charge fall into this category. To study and correct for the mis-modeling of these processes, control regions enriched in such events need to be identified.

This chapter begins with studies of the two signal processes to identify regions enriched in signal. After this, the expected background processes are taken into consideration. Based on the expected contribution from various non-prompt background sources, redefinitions of the background templates are carried out. These redefinitions allow for a categorization of the background into prompt and non-prompt background sources. The kinematics of both categories are then studied in more detail. The combination of these studies allows for suitable definitions of signal and control regions.

In the following, only events including two leptons of the same charge are taken into account. The two leptons are ordered and labeled according to their transverse momentum. The p_T-leading lepton is referred to as ℓ_0, the sub-leading lepton as ℓ_1. Both leptons need to pass a minimum p_T-threshold of 10 GeV. Additionally, the missing transverse energy is required to surpass 30 GeV. Furthermore, all signal processes are scaled to the current upper limit on the branching ratio $\mathcal{B}(t \to Hu/Hc)$. All estimated event yields are given for an integrated luminosity of $L_{\text{int}} = 139\,\text{fb}^{-1}$.

6.1 Kinematics of signal processes

The most important aspect of finding a suitable set of signal and control regions is studying the signal and its kinematic properties. Primarily it is important to understand, how many jets and b-tags are to be expected, both from the utH and the ctH signal. Hence, the two quantities n_{jets} and $n_{b\text{-tags}}$ are depicted in a two-dimensional histogram in Figure 6.1 for both signal components.



(a) *utH* signal

(b) *ctH* signal

Figure 6.1: Two-dimensional histograms of the number of jets and the number of b-tags for the two signal processes.

The histograms of both signal processes are very similar. The largest fraction of events has exactly one b-tagged jet. This is in accordance with the expected event signature of the $H \to WW^*$ channel (see Figure 2.6 a): two leptons, four jets and one b-tag. The fact that most signal events include at least three jets matches those expectations as well. Based on these observations, further studies on the signal are conducted for events with at least one jet and exactly one b-tag. The region defined by these cuts ($n_{\text{jets}} \geq 1$, $n_{b\text{-tags}} = 1$) will be referred to as testing region in this chapter. The region with zero b-tags is excluded here, because in general this region is severely contaminated by background events, resulting in a signal purity that makes meaningful studies impossible.

As the requirement for two leptons of the same charge is the most distinctive selection criterion for events in this thesis, the kinematics of those two leptons in the signal processes are studied further. Figure 6.2 shows the distribution of the transverse momenta of the two leptons for both signal processes. The distributions are normalized to unit area. Such plots will be referred to as shape plots of a certain quantity for the remainder of this thesis. It is clearly visible that the distributions of the two transverse momenta are very similar for the two signal processes.

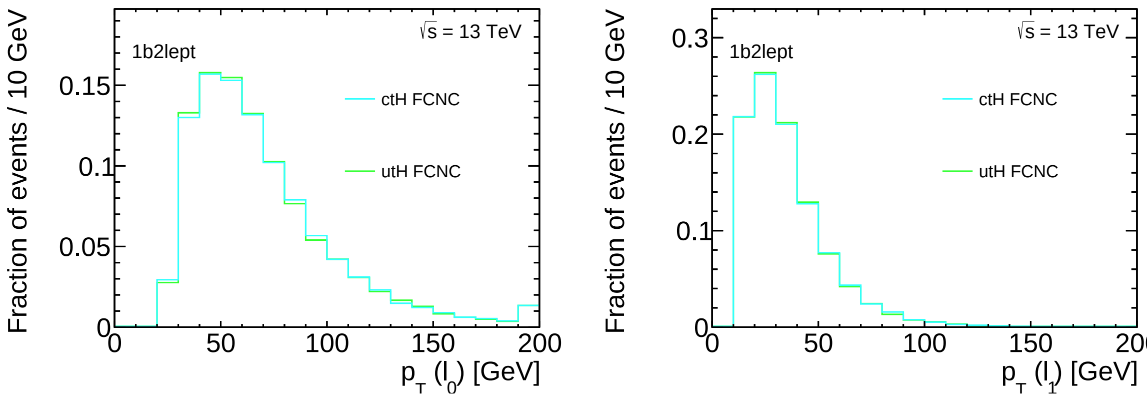


Figure 6.2: Shape plots of the transverse momenta of ℓ_0 (left) and ℓ_1 (right) for the two signal processes in the testing region.

It is observed that the leading lepton generally has a transverse momentum higher than 30 GeV. Only few fall into the bin between 20 GeV and 30 GeV. This is caused by the ATLAS single lepton trigger [59], which requires a transverse momentum of at least 28 GeV to trigger an event. Once an event is triggered, the only requirement made on the p_T of additional leptons at reconstruction level is a minimum p_T of 10 GeV. Hence, about 22 % of all signal events in the testing region include a sub-leading lepton with a p_T between 10 GeV and 20 GeV. This has to be accounted for whenever a cut on the p_T of the sub-leading lepton is considered in this thesis.

6.2 Redefinition of background templates

After a discussion of the signal, the background processes and their kinematic properties are studied. The studies are undertaken in the testing region defined in the previous section. The background from events including non-prompt leptons ("non-prompt background") is of particular importance in this thesis. Hence, it is important to study how much various non-prompt background sources contribute to kinematic regions of interest. Figure 6.3 shows the IFF Truth Class of both leptons for events in the testing region.

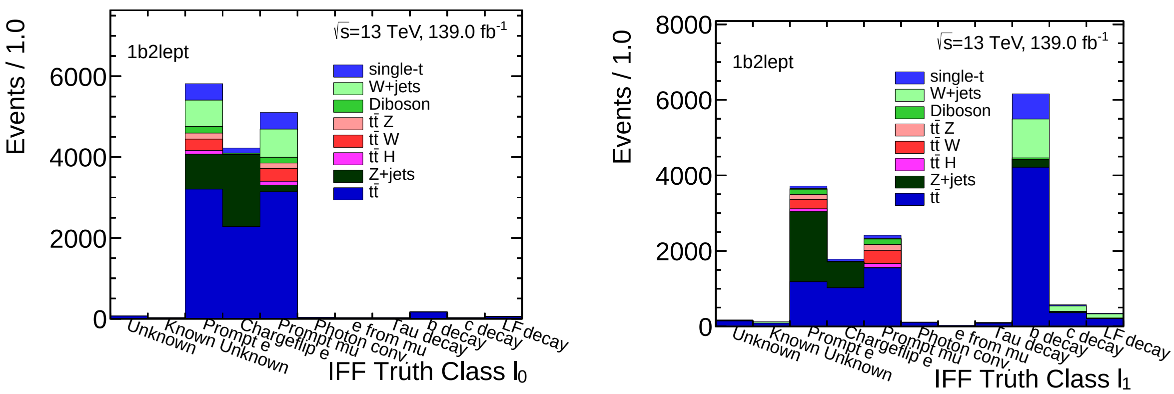


Figure 6.3: Stacked plots of the IFF Truth Class of ℓ_0 (left) and ℓ_1 (right) for events with at least one jet and exactly one b-tagged jet. The plots show the original background process templates.

As is clearly visible, the dominating non-prompt background events include either an electron reconstructed with an inverted charge (charge-flip electron), or a lepton originating from a heavy-flavor jet (HF-decay lepton). Only a small number of events includes non-prompt leptons from different sources, such as photon conversions or light-flavor decays. This leads to the conclusion that out of all non-prompt background sources charge-flip electrons and HF-decay leptons are the most important in this analysis. To be able to study those background sources more precisely, new background process templates based on the promptness of leptons are defined. These templates split the previous processes by their prompt components (i.e. the events with two prompt leptons) and their non-prompt components (i.e. the events with one or more non-prompt leptons). A definition of those templates is given in Table 6.1. It should be noted that the definition of the signal processes remains unchanged.

Table 6.1: Process templates defined based on the expected non-prompt background contributions.

Classification	**Process template**	**Definition**
Prompt background	$t\bar{t}H$	Prompt component of original $t\bar{t}H$
	$t\bar{t}Z$	Prompt component of original $t\bar{t}Z$
	$t\bar{t}W$	Prompt component of original $t\bar{t}W$
	Others (prompt)	Prompt component of all other processes ($t\bar{t}$, Single t, Z+jets, W+jets, Diboson)
Non-prompt background	Charge-flip e	Events from all processes with exactly one charge-flip electron
	HF-decay e	Events from all processes with exactly one HF-decay electron
	HF-decay μ	Events from all processes with exactly one HF-decay muon
	Others (non-prompt)	Events from all processes with exactly one lepton that is neither a prompt lepton, charge-flip electron or HF-decay lepton and events with two non-prompt leptons

6.3 Kinematics of non-prompt background processes

The redefined process templates allow for a closer inspection of the kinematics of specific non-prompt background processes. These kinematics are again studied in the testing region defined in the previous sections. Here, however, either the leading lepton ℓ_0 or the subleading lepton ℓ_1 is required to be non-prompt, while the other one is required to be prompt[1]. This ensures, that only the non-prompt background templates are included in the selection. The event yields for the two selections can be seen in Table 6.2.

Table 6.2: Event yields in the testing region for events with exactly one non-prompt lepton. Here, p_i signifies the promptness and $\not{p}_i$ the non-promptness of the i'th lepton.

Process	Total	$p_0\ \not{p}_1$	$\not{p}_0,\ p_1$
Charge flip e	6001	1782	4224
HF-decay μ	4382	4248	134
HF-decay e	2545	2486	58
Others (non-prompt)	1063	873	190

From these results, several conclusions can be made. The most obvious observation is the difference in absolute expected event yields for the various processes. The charge-flip electron process contributes the most to the non-prompt background, followed by the HF-decay muon process, the HF-decay electron process and the totality of all other non-prompt background

[1] A very small number of events falls into the category of having two non-prompt leptons. Events like this are generally very rare. For the testing region, 2.24 events are expected to fall into this category. For completeness' sake, they are included in the Others (non-prompt) process template.

sources. More notably, however, is the fact that HF-decay leptons tend to have a smaller transverse momentum than prompt leptons of the same event. This is reflected in the number of events with a sub-leading non-prompt lepton being significantly higher than the number of events with a leading non-prompt lepton for the HF-decay μ and the HF-decay e process. For the charge-flip electron background, the opposite is the case. Here, the electron with a misidentified charge is more likely to have a higher transverse momentum than the prompt lepton from the same event. Both observations are compatible with the physical understanding of HF-decay leptons and charge-flip electrons. Leptons that originate from a heavy-flavor jet only contain a fraction of the energy of the quark that was produced in the hard scattering process. This means that on average the energy and hence the transverse momentum of HF-decay leptons is smaller than those of prompt leptons. In contrast to this, electrons are more likely to be reconstructed with an inverted charge if their transverse momentum is higher than average. The charge of electrons is reconstructed using the curvature of the track in the inner detector, which is immersed in the solenoid magnetic field. This curvature decreases for an electron when its transverse momentum increases. Hence, higher transverse momenta decrease the efficiency for a correct reconstruction of an electron's charge.
The number of muons with a misidentified charge is assumed to be zero due to the elaborate muon system of the ATLAS detector. Hence, the main contribution to processes containing a non-prompt muon comes from the HF-decay μ process. This fact, in combination with the kinematic behavior studied above, leads to the conclusion that if the leading lepton of a given event is a muon, it can almost certainly be assumed to be a prompt muon.

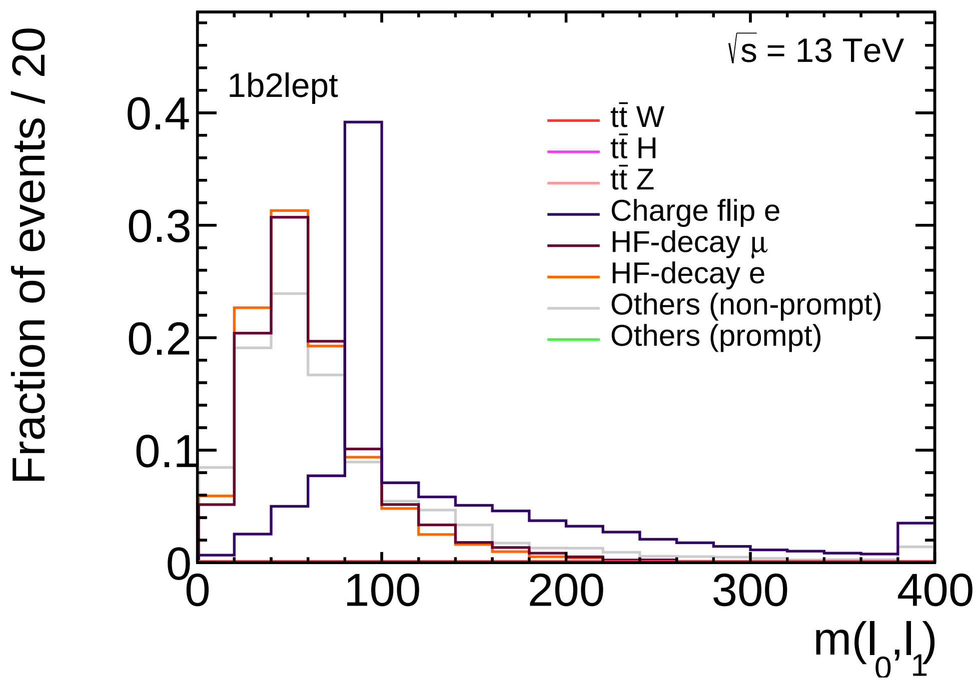


Figure 6.4: Shape plot of the invariant mass of the two leptons for events with exactly one non-prompt lepton in the testing region.

In addition to the p_T and the expected flavor-composition of different non-prompt processes, various other kinematic variables can be distinctive. The invariant mass of the two leptons

in a charge-flip e event, for instance, is very likely to fall between 80 GeV and 100 GeV, as can be seen in Figure 6.4. This is caused by Z+jets-events, in which the Z boson decays into two electrons, one of which is identified with the wrong charge[2]. This behavior is used to heavily reduce the number of charge-flip e events in the final signal region by vetoing events, in which the invariant mass of two electrons is close to the mass of the Z boson.
In addition to the p_T and the invariant mass of the two leptons, the distribution of the number of jets for non-prompt background processes is taken into account. Figure 6.5 shows the shape plot of n_{jets} for events in the testing region with exactly one non-prompt lepton, i.e. for the combination of the $p_0 \not{p}_1$ and the $\not{p}_0 p_1$ subregions.

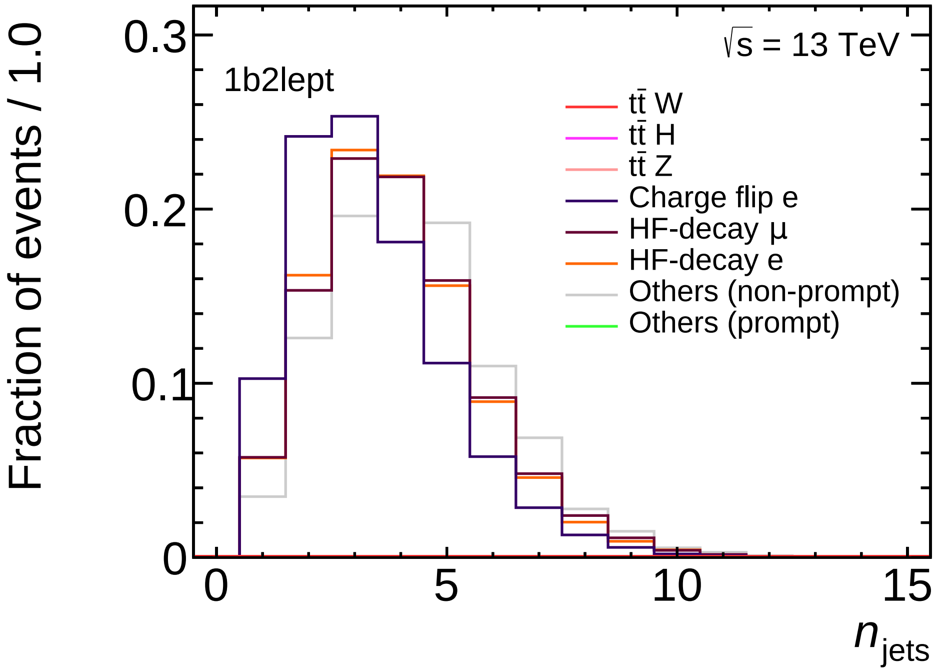


Figure 6.5: Shape plot of n_{jets} for events with exactly one non-prompt lepton in the testing region.

The plot shows that approximately 60 % of all charge-flip electron events have three or fewer jets. Likewise, approximately 45 % of all HF-decay e and μ events fall into these bins. This corresponds to a significant subset of events of the respective background processes and makes events with $n_{\text{jets}} \leq 3$ a prime candidate for a non-prompt background control region. The previous studies on the signal showed a low number of signal events with more than one b-tagged jet. By requiring $n_{b\text{-tags}} \geq 1$ instead of $n_{b\text{-tags}} = 1$ the number of non-prompt background events can hence be increased, without significantly increasing the signal contamination. Thus, a study of the region defined by $n_{\text{jets}} \leq 3$ and $n_{b\text{-tags}} \geq 1$ as a candidate control region is conducted.
As discussed above, the contribution of various non-prompt processes differs depending on the flavor-composition of the two leptons. Hence, it may be useful to split the candidate control region by the leptons' flavor. Table 6.3 shows the expected event yields for all processes

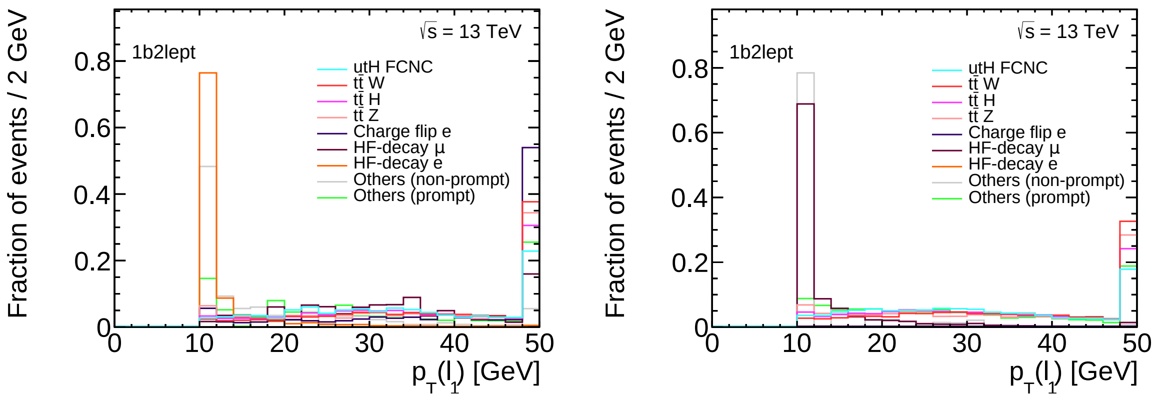
[2]The mass of the Z boson amounts to 91.2 GeV, which lies between 80 GeV and 100 GeV.

in the region. The colored entries signify enriched non-prompt processes in specific subregions. For the charge-flip e process, the ee subregion constitutes the most optimal control region. For the HF-decay e and μ processes, the μe- and $\mu\mu$ subregions are most suitable. The $e\mu$ subregion is not specifically enriched in any non-prompt background (or in signal events).

Table 6.3: Event yields split by lepton-flavor in the candidate control region defined by $n_{\text{jets}} \leq 3$ and $n_{b\text{-tags}} \geq 1$. The flavor specifications of the two leptons are ordered by p_T. The $e\mu$ subregion, for example, consists of events with a leading electron and a sub-leading muon. The red numbers signify non-prompt background processes enriched in a specific region.

Process	ee	$e\mu$	μe	$\mu\mu$
$t\bar{t}W$	40	55	46	63
$t\bar{t}H$	5	6	5	6
$t\bar{t}Z$	9	13	10	14
Others (prompt)	120	73	61	49
HF-decay e	143	15	595	0
HF-decay μ	0	260	19	1012
Charge-flip e	2888	1104	473	0
Others (non-p)	81	58	137	75
utH signal	89	118	96	128
ctH signal	87	112	91	124
S/B utH [%]	2.7	7.4	7.1	10.5
S/B utH [%]	2.6	7.1	6.8	10.2

The μe and $\mu\mu$ subregions still contain a relatively high number of signal events compared to the ee subregion. These contaminate the control regions and could, ideally, be used to gain further sensitivity to the signal. The observed tendency for HF-decay leptons to have a lower p_T than prompt leptons can be exploited here. As mentioned previously, the leading muon in the μe and the $\mu\mu$ subregions is assumed to be prompt in most cases. Hence, it is sufficient to study the p_T of the sub-leading lepton in both subregions, to find a useful cut, which separates most HF-decay events from most signal events. Figure 6.6 shows high-resolution shape plots of the low-p_T region in the μe and $\mu\mu$ subregions. The plots are shown for both components of the utH signal. The ctH signal has a similar shape. The corresponding shape plots can be found in the appendix in Figure A.1.



(a) μe subregion

(b) $\mu\mu$ subregion

Figure 6.6: High-resolution shape plots of the p_T of the sub-leading lepton in the μe (a) and the $\mu\mu$ subregions (b) of the candidate control region defined by $n_{\text{jets}} \leq 3$ and $n_{b\text{-tags}} \geq 1$. The plots only include the utH signal component. The final bin constitutes an overflow-bin, meaning it includes events with $p_T(\ell_1)$ between 48 GeV and 50 GeV as well as all events with $p_T(\ell_1) > 50$ GeV.

It is observed that the vast majority of all HF-decay events produce a sub-leading lepton with a p_T below 16 GeV. Thus, events in the μe and the $\mu\mu$ subregions with $p_T(\ell_1) \leq 16$ GeV define the control regions for HF-decay background processes. Events in the subregions with $p_T(l_1) > 16$ GeV define a signal region. The notable excess in HF-decay events in the bin between 10 GeV and 12 GeV will be discussed in detail in Section 6.5.
In summary, this section identified suitable control regions for all dominant non-prompt backgrounds and further refined, which parts of the testing region are useful as a signal region. To gain a full understanding of all processes that are relevant in this analysis, only a study on the prompt background processes remains.

6.4 Kinematics of prompt background processes

After an analysis of the signal and the non-prompt background component, a study of the prompt background sources is conducted. For this, similarly to the previous section, events in the testing region with two prompt leptons are taken into account. The contributions from various prompt background sources are listed in Table 6.4.

Table 6.4: Event yields in the testing region for events with two prompt leptons. Here, p_i signifies the promptness of the i'th lepton.

Process	p_0 p_1
$t\bar{t}W$	596
$t\bar{t}Z$	279
$t\bar{t}H$	177
Others (prompt)	478

As can be seen, the main prompt background contribution comes from the $t\bar{t}W$, $t\bar{t}Z$ and $t\bar{t}H$ (in combination: $t\bar{t}X$) processes. The measured cross sections for these processes have been found to deviate from NLO-calculations in previous analyses [60]. Thus, an additional control region to correct for the normalization of these background processes needs to be defined. Useful candidates are events with two b-tagged jets, because $t\bar{t}X$ events are expected to produce one b-jet per produced top quark. Since a large jet multiplicity and at least two b-jets per event are expected for the $t\bar{t}X$ background, only events with $n_{\text{jets}} \geq 4$ and $n_{b\text{-tags}} = 2$ are considered here. As outlined in the previous section, the contribution from different non-prompt background events strongly depends on the flavor composition of the two leptons. Hence, Table 6.5 shows the expected event yields in the region defined by $n_{\text{jets}} \geq 4$ and $n_{b\text{-tags}} = 2$ split into the different flavor combinations of the two leptons.

Table 6.5: Event yields split by lepton flavor in the candidate control region defined by $n_{\text{jets}} \geq 4$ and $n_{b\text{-tags}} = 2$. The flavor specifications of the two leptons are ordered by p_T. The $e\mu$ subregion, for example, consists of events with a leading electron and a sub-leading muon. The ratio of expected $t\bar{t}X$ events and expected events from other processes is shown as well.

Process	ee	$e\mu$	μe	$\mu\mu$
$t\bar{t}W$	64	91	71	100
$t\bar{t}H$	25	36	27	36
$t\bar{t}Z$	41	52	36	44
Others (prompt)	26	12	13	1
HF-decay e	75	1	73	0
HF-decay mu	0	127	4	131
Charge-flip e	585	406	169	0
Others (non-p)	145	84	110	53
utH signal	11	13	10	14
ctH signal	18	26	20	25
$N_{t\bar{t}X}/N_{\text{Rest}}$ [%]	15	27	35	86

It is observed that the ee and the $e\mu$ subregions are generally more contaminated by non-$t\bar{t}X$ background processes than the μe and $\mu\mu$ subregions. Hence, the latter ones are combined to the μx subregion and used as a control region for the $t\bar{t}X$ background. This concludes the search for suitable control regions for the various background processes.

6.5 Additional global cuts

As discussed in the introduction of this chapter, events in all regions are required to surpass a threshold of 30 GeV on the missing transverse energy. This section discusses two additional global cuts, which heavily reduce contributions from non-prompt background processes. The Electron Charge Identification Selection Tool (ECIDS) strongly decreases the total number of charge-flip e events, while a global cut on the p_T of the sub-leading lepton eliminates a large number of HF-decay e and μ events.

6.5.1 Electron Charge Misidentification Selection tool (ECIDS)

The ECIDS [61] is a tool that was specifically developed to reduce the amount of charge-flip *e* background events. The tool is based on a boosted decision tree, which uses information about specific kinematic variables of a reconstructed electron. It produces one output variable, which can take values between -1 and +1. Here, -1 signifies a very high charge-flip probability, whereas +1 signifies that an electron was most likely reconstructed with a correct charge. The tool uses a predefined operating point, which cuts on all electrons with an ECIDS result of less than -0.337671. To visualize the effect of such a cut on the charge-flip *e* background, the ECIDS results for both leptons are depicted in a two-dimensional histogram for charge-flip *e* events in the testing region in Figure 6.7. The ECIDS result for muons is fixed to 1 in general.

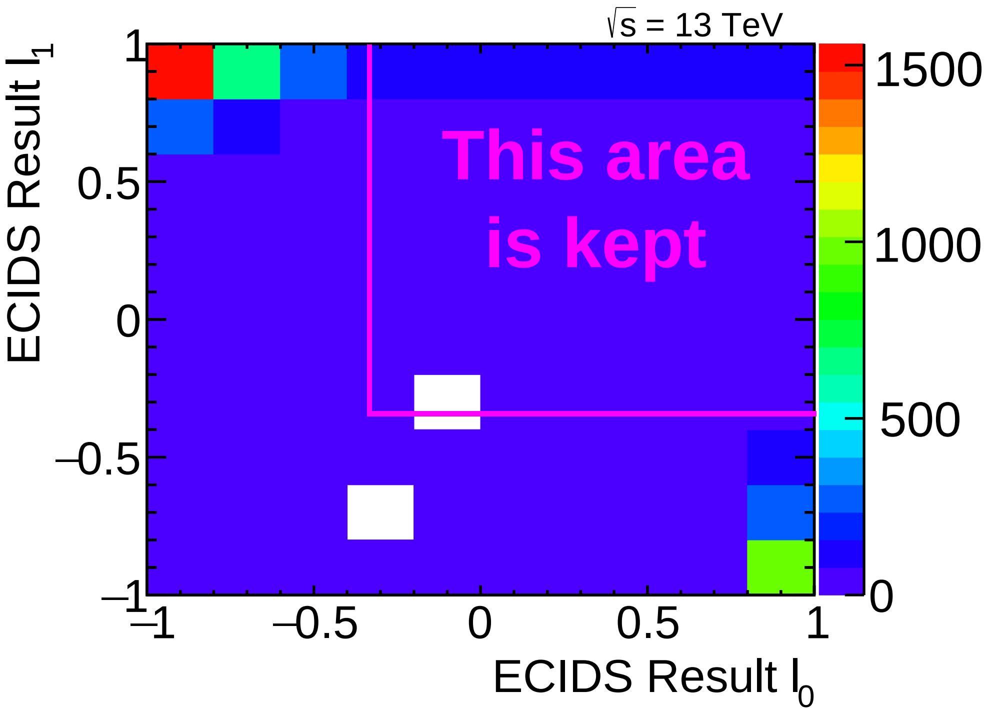


Figure 6.7: Two-dimensional histogram of the ECIDS result of both leptons for charge-flip *e* events in the testing region. The area of events that is kept after applying a cut at -0.337671 is marked.

As expected, due to the definition of the charge-flip *e* process template, most events consist of one lepton with a very high ECIDS score and one with a very low ECIDS score. Application of the cut provided by the operating point of the tool reduces the number of charge-flip *e* events in the testing region by 80 %. At the same time, the number of signal events is only reduced by less than 6 % (cf. appendix Table B.1). Thus, the cut on the ECIDS score proves to be a valuable addition to the event selection. For reasons of consistency, this cut is applied in all regions, including the candidate HF-decay control regions.

6.5.2 Minimum transverse momentum of leptons

Figure 6.6 already showed that in the candidate HF-decay control regions a large number of HF-decay events produce a sub-leading lepton with a p_T between 10 GeV and 12 GeV. A similar observation is made in the full testing region, as Figure 6.8 shows.

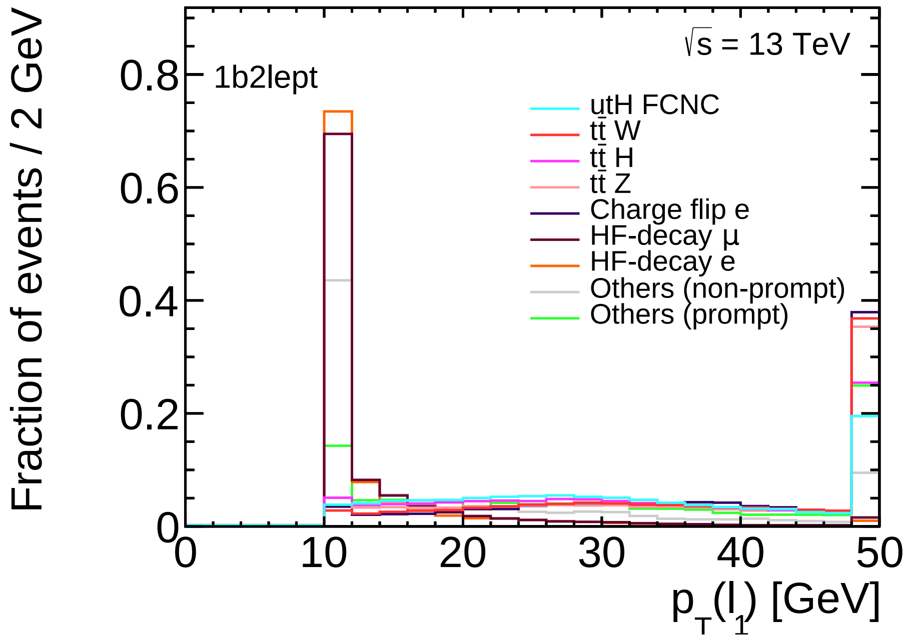


Figure 6.8: High-resolution shape plot of the p_T of the sub-leading lepton in the testing region. Only the signal contribution from the *utH* signal is shown.

As stated in Chapter 5, the prompt lepton tagger, which poses isolation criteria on electrons and muons to reduce the total number of HF-decay leptons in all regions, is based on two different BDT scores: The PLV score for leptons with a p_T above 12 GeV and the low-p_T PLV score for leptons with a p_T below 12 GeV. The low-p_T PLV does not differentiate between leptons from b- or c-jets and leptons from light flavor jets. This may cause a decrease in the efficiency of the low-p_T PLV compared to the regular PLV. To test such a hypothesis, several in-depth studies on the features of low-p_T HF-decay leptons and the PLV efficiency would be necessary. For this analysis it should suffice, however, to simply cut on all leptons with a p_T below 12 GeV. As Figure 6.8 implies, this results in a loss of less than 5 % of the signal, while removing 70 % to 80 % of all HF-decay e and μ events. Similarly to the ECIDS cut, this cut is applied in all regions.

6.6 Definition of signal and control regions

After detailed studies of all relevant background and signal processes, a set of signal and control regions is identified. An overview of these regions can be found in Figure 6.9.

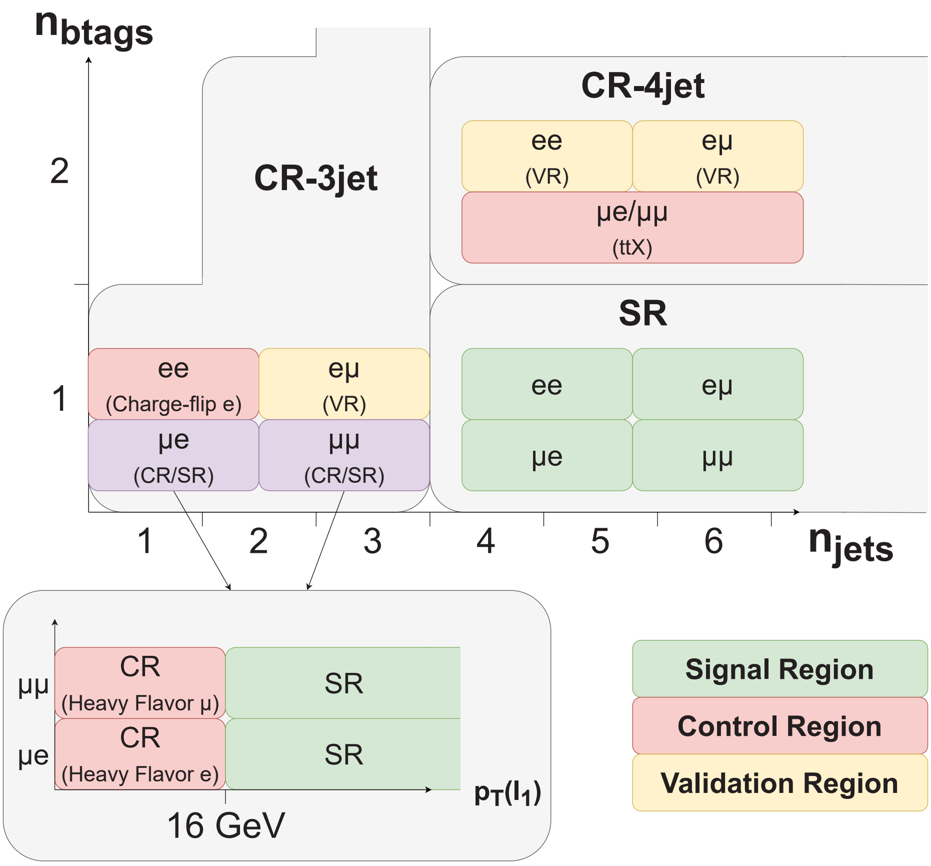


Figure 6.9: Overview of the regions used in this analysis. The abscissa shows the number of jets, the ordinate the number of b-tags. Areas in the plot define specific regions, which are further divided by lepton flavor. For the CR-3jet region the μe and $\mu\mu$ subregions are further divided by the p_T of the sub-leading lepton.

The part of the testing region with at least four jets is defined as the main signal region (SR). The control region for non-prompt background processes is called CR-3jet. As discussed, it is split by lepton flavor, because different lepton flavor compositions result in specific non-prompt background processes being enriched. The CR-3jet μe and $\mu\mu$ subregions are split by the p_T of the sub-leading lepton. The regions with $p_T(\ell_1) \leq 16\,\text{GeV}$ define the control regions for HF-decay leptons. The regions with $p_T(\ell_1) > 16\,\text{GeV}$ are used as two additional signal regions, which are referred to as SR-3jet μe and SR-3jet $\mu\mu$. Finally, the region defined by $n_{\text{jets}} \geq 4$ and $n_{b\text{-tags}} = 2$ is referred to as CR-4jet. The subregion CR-4jet μx serves as a control region for the $t\bar{t}X$ background. The three subregions CR-3jet $e\mu$, CR-4jet ee and CR-4jet $e\mu$ are defined as validation regions. These are used to validate the agreement of any corrections which are made to the various background processes with experimental data.
To reduce the contribution from charge-flip e events in the SR, a cut on the invariant mass of the two leptons is applied. As discussed previously, a large fraction of charge-flip e events comes from Z+jets and $t\bar{t}Z$ events, in which the Z boson decays into two electrons. Most of these events can be excluded by requiring $|m(ee) - m_Z| > 10\,\text{GeV}$, where $m(ee)$ is the invariant mass of two electrons and $m_Z = 91.2\,\text{GeV}$ is the mass of the Z boson. Many of these events are already excluded due to the implementation of the ECIDS in the previous

section. As this cut, however, combines information from both leptons it is still found to increase the overall signal sensitivity.
The studies on non-prompt background processes revealed that lepton flavor composition can be exploited to exclude specific non-prompt background processes and enrich others. This might be useful when constructing methods to separate the signal from the background in the main SR. Hence, in the following the SR will be considered as one combined region and as four different subregions, split by the four possible flavor compositions, simultaneously. It is possible that this splitting will prove beneficial to the overall sensitivity of the analysis. However, this can only be clearly verified once the full strategy of the analysis is outlined. Table 6.6 shows the expected event yields in the combined and the flavor-split SR.

Table 6.6: Event yields in the SR. Event yields are shown for both, the combined SR and the flavor-split SR.

Process	Total	ee	$e\mu$	μe	$\mu\mu$
$t\bar{t}W$	433	73	120	99	141
$t\bar{t}H$	150	26	44	33	46
$t\bar{t}Z$	228	47	69	50	62
Others (prompt)	105	30	29	29	17
HF-decay e	353	151	19	182	0
HF-decay mu	784	0	339	31	413
Charge-flip e	328	153	136	40	0
Others (non-p)	378	114	98	113	54
utH signal	717	127	212	151	227
ctH signal	616	106	178	139	194

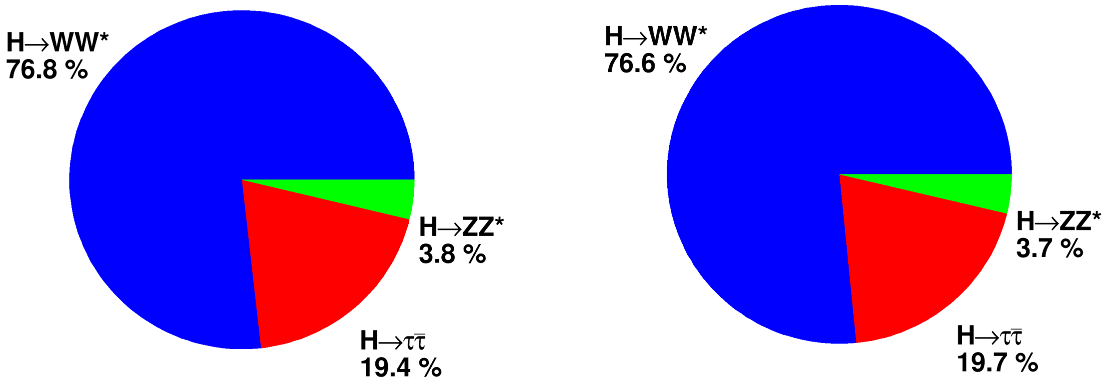
7 Event Reconstruction

For a successful separation of the signal and background processes, efficient reconstruction methods are of great benefit. This chapter outlines the reconstruction approaches taken in this thesis. The primary goal of efficient reconstruction methods is to produce kinematic variables which behave differently for the signal and background processes. These variables are then used in a multivariate analysis, which is described in Chapter 8.

As discussed in Section 2.4, various Higgs-boson decay modes contribute to the signal process. The chapter begins by quantifying the expected contributions from these decay modes to regions of interest. It is followed by the discussion of various reconstructed variables in the SR. The first variable is the four-vector of the hadronic W bosons in the $H \rightarrow WW^*$ channel. Secondly, the neutrino weighting method is introduced. It is a way to overcome challenges that arise when trying to reconstruct processes with more than one neutrino in their final state. Lastly, the so-called Fox-Wolfram moments are defined and their utilization in this analysis is described.

7.1 Signal composition in regions of interest

In the following it is assumed that in the $t \rightarrow qH$ decay channel the non-FCNC top quark decays into a b-quark and a W boson, which subsequently decays leptonically. In this case, there are several decay modes of the Higgs boson, which contribute to the 2ℓSS final state. It has already been established in Section 2.4 that these decay modes are $H \rightarrow WW^*/ZZ^*/\tau\bar{\tau}$, with the largest expected contribution from $H \rightarrow WW^*$. The definition of all signal and control regions in Chapter 6 now allows for a full assessment of the share of the various Higgs decay modes in relevant signal events. Figure 7.1 shows pie charts, which illustrate the composition of the two signal processes in the combined SR. Similar pie charts for every individual region, including the two smaller SR-3jet μe and $\mu\mu$, can be found in Figure A.2 for the utH signal and in Figure A.3 in the appendix for the ctH signal.



(a) utH signal (b) ctH signal

Figure 7.1: Pie charts illustrating the contributions from various Higgs decay modes to the combined SR. The contributions are shown for the utH signal (a) and the ctH signal (b).

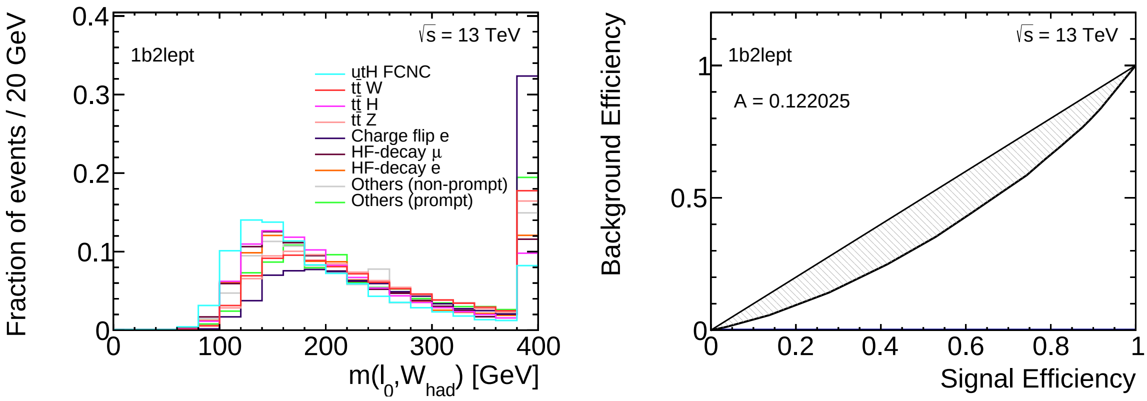
As expected, the $H \to WW^*$ decay channel dominates in the SR. Approximately 75 % of all signal events can be attributed to this decay mode. The second largest contribution comes from the $H \to \tau\bar{\tau}$ channel. However, it contributes only roughly 20 % to all signal events. An almost negligible contribution of less than 4 % comes from the $H \to ZZ^*$ decay mode. Based on these observations any efforts made in terms of event reconstruction are focused on the $H \to WW^*$ decay channel, to maximize their capability to discriminate between the signal and background processes.

7.2 Hadronic W boson

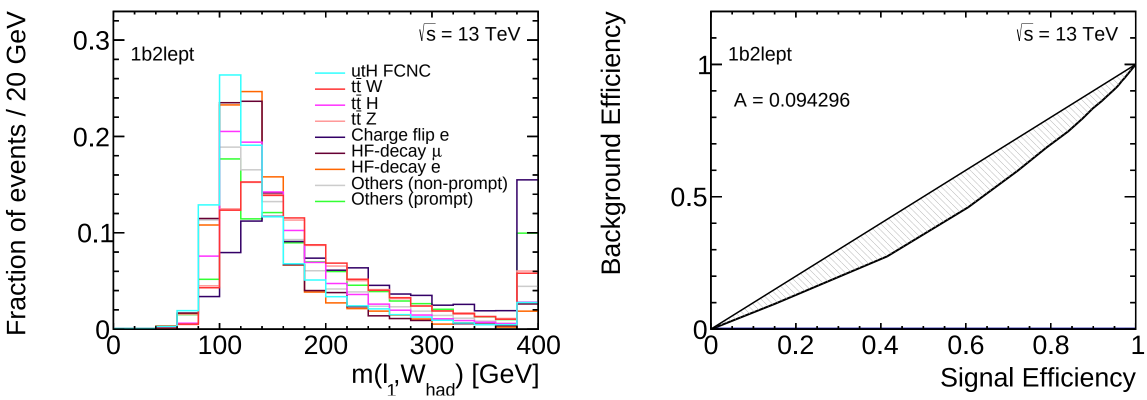
In the $H \to WW^*$ channel, one W boson is required to decay into a lepton and a neutrino to produce a 2ℓSS final state. The other one decays hadronically, producing two jets. In most cases, this hadronic W boson, W_{had}, decays into two light quarks (i.e. non-b-quarks), as the cross-generational coupling of the b-quark is highly suppressed due to the entries of the CKM matrix [7]. Hence, W_{had} is reconstructed from light jets that are detected in a given event. Out of all detected light jets, the two jets which have a combined invariant mass closest to the mass of the W boson are selected. The sum of their four-momenta is defined as the four-momentum of W_{had}.
It should be noted that this method does not necessarily select for the two light jets, which actually originate from the hadronic W boson from the Higgs decay. Firstly, as discussed in Section 7.1, approximately 25 % of all signal events do not originate from the $H \to WW^*$ decay. For those events, two random light jets are selected to reconstruct a virtually meaningless four-momentum. Variables constructed from this four-momentum cannot be used to distinguish reliably between signal and background. Secondly, for the $H \to WW^*$ channel only one W boson is on-shell[1]. If the off-shell W boson decays hadronically, the above reconstruction method likely also selects a random pair of jets which might have an invariant

[1]A particle is referred to as on its mass-shell ("on-shell"), if it satisfies the relativistic energy-momentum relations. Correspondingly, it is called off-shell, if it doesn't satisfy those relations [62].

mass close to the mass of the W boson. Nonetheless, several variables which prove to be very discriminating between signal and background processes are produced from W_{had} as defined above. The invariant masses of W_{had} and either of the two leptons of a given event, for instance, show some rather different shapes for signal and background processes in the SR. Figure 7.2 shows the shape plots of these invariant masses in the combined SR. Additionally, the signal efficiency from a given bin to the last bin is plotted against the background efficiency for the same range of bins for both variables. The resulting curve is called a receiver operating characteristic (ROC) curve [63]. The area between the ROC curve and the diagonal line can be viewed as a measure of how well a given variable separates between signal and background.



(a) $m(W_{\text{had}},\ell_0)$



(b) $m(W_{\text{had}},\ell_1)$

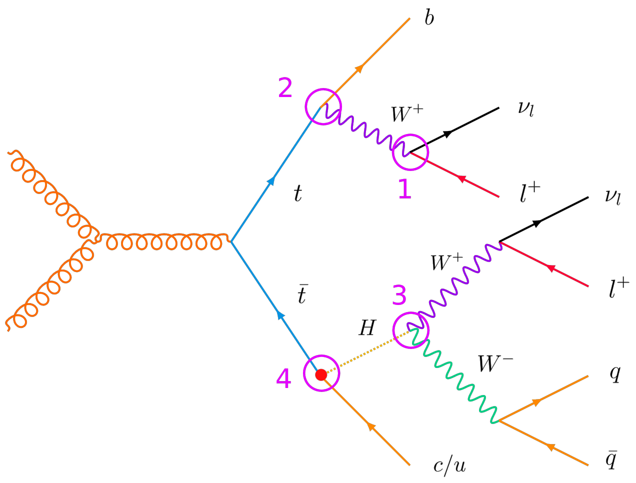
Figure 7.2: Shape plots and ROC curves of the invariant masses of W_{had} and ℓ_0 (a)/ ℓ_1 (b) in the combined SR. The area A between the ROC curves and the diagonal is given in the respective plots. Only the utH signal is shown in the plots.

As both plotting types confirm, the two variables discriminate successfully between signal and background. A large set of similar variables is constructed using the reconstructed W_{had}. These include the invariant mass and angular separation ΔR with multiple other reconstructed objects, such as the b-tagged jet in the SR.

7.3 Neutrino weighting method

A characteristic for the 2ℓSS final state is the presence of two neutrinos, as the leptonic W bosons decay into ℓ-ν_ℓ-pairs. As outlined in Section 3.2, neutrinos cannot be detected by the ATLAS detector. The only available information on the two neutrinos is the missing transverse momentum. As a result, a full reconstruction of signal events is impossible. There are, however, ways to use the knowledge about the underlying process to extract information from the kinematic variables that are available. One such method is the neutrino weighting method [64], which is explained in the following.

The method has been used in previous analyses [65], mainly to extract the mass of the top quark from dileptonic $t\bar{t}$-events. In this thesis, however, its primary purpose is to produce a variable, which helps to discriminate between signal and background. The main problem it attempts to solve is the lack of information for a successful reconstruction of the four momenta of the two neutrinos. Every four-momentum has four components. However, since the mass of a neutrino is known, each one introduces three degrees of freedom. This creates the necessity for six constraints on the neutrinos' four-momenta to reconstruct the full event. These constraints are obtained from theoretical predictions, as all available experimental information is used to reconstruct the four-momenta of other final state particles.



1) $m_W{}^2 = (p^\mu(\nu_t) + p^\mu(\ell_t))^2$

2) $m_t{}^2 = (p^\mu(\nu_t) + p^\mu(\ell_t) + p^\mu(b))^2$

3) $m_H{}^2 = (p^\mu(\nu_H) + p^\mu(\ell_H) + p^\mu(q_{H,1}) + p^\mu(q_{H,2}))^2$

4) $m_t{}^2 = (p^\mu(\nu_H) + p^\mu(\ell_H) + p^\mu(q_{H,1}) + p^\mu(q_{H,2}) + p^\mu(c/u))^2$

Figure 7.3: Feynman diagram of the signal process in the 2ℓSS final state. The four vertices at which conservation of four-momentum is used in the neutrino weighting method are marked. The constraints which these verticese pose on the four-momenta of the final state particles are depicted as well. The masses of the W boson, the top quark and the Higgs boson are denoted m_W, m_t and m_H. $p^\mu(X)$ refers to the four-momentum of a final state particle X ($X = \ell_H, \nu_t, ...$). The indices $_t$ and $_H$ denote the decay products from the non-FCNC top quark and the Higgs boson.

The most straight-forward way to extract such constraints from theoretical predictions is to rely on the fact that the total four-momentum is conserved at all vertices in a Feynman diagram. For the signal process, the four vertices that are marked in Figure 7.3 are important in particular. All of the four vertices pose constraints on the four-momenta of the final state particles, which are depicted in Figure 7.3 as well.
For a full reconstruction, two additional constraints are needed. To obtain them, the probability distributions of the pseudorapidity η_ν of the two neutrinos is taken into account. They are obtained from the MC truth information of the signal processes. The histogram of the MC truth value of $\eta(\nu_t)$ for the neutrino from the top-quark decay in the *utH* signal is depicted in Figure 7.4. Similar plots for $\eta(\nu_t)$ in the *ctH* signal as well as for $\eta(\nu_H)$ from the Higgs-boson decay for both signal processes can be found in Figure A.4 in the appendix. To obtain continuous probability density functions, χ^2-fits of non-normalized Gaussian distributions are performed. The fit function and results of the non-normalized distributions are shown in the respective figures as well.

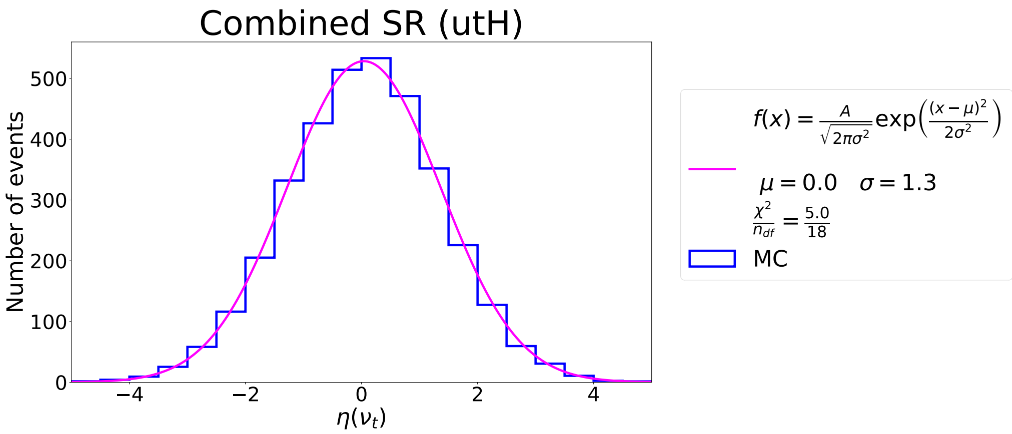


Figure 7.4: Histogram of the truth value of η_ν for the neutrino from the non-FCNC top-quark decay from the *utH* signal in the combined SR. The results of a χ^2-fit of a non-normalized Gaussian distribution are depicted as well.

Based on these distributions, values for the pseudorapidity of each neutrino are chosen. This gives the two additional constraints required for a full reconstruction of the four-momenta of the neutrinos and hence of the full event. The full calculation of the neutrinos' four-momenta based on the six given constraints is analogous to the one presented in [64].
It remains to be determined, which values are to be chosen for the pseudorapidity. An additional problem arises because different objects at reconstruction level cannot be uniquely mapped to expected final state particles of the signal process. As an example the two final state leptons are considered. It is not clear, which lepton originates from the non-FCNC top quark and which one from the Higgs boson. Both problems are solved similarly. To consider different values for the pseudorapidity, the distributions of both neutrinos are divided into 10 bins in the interval $\eta = [-4,4]$, each bin containing 10 % of the expected events. Then the four-momenta of the neutrinos are reconstructed for all $N_{\eta\text{-comb}} = 100$ combinations of $\eta(\nu_H)$ and $\eta(\nu_t)$. Likewise, the reconstruction is performed for every possible assignment of objects at reconstruction level to final-state particles. For the example of the two leptons

this means that the reconstruction is performed for $\ell_H = \ell_0$ and $\ell_t = \ell_1$, and for $\ell_H = \ell_1$ and $\ell_t = \ell_0$. Similarly, all possible assignments of reconstructed light jets to final state light quarks are considered. The total number of combinations for which a reconstruction is performed is referred to as N_{comb}, which is determined by Equation 7.1.

$$N_{\text{comb}} = N_{\eta\text{-comb}} \cdot N_{\ell} \cdot \frac{N_{l\text{-jets}}!}{(N_{l\text{-jets}} - N_{lq})!} \tag{7.1}$$

Here, $N_\ell = 2$ refers to the number of leptons, $N_{lq} = 3$ to the number of light quarks in the final state and $N_{l\text{-jets}}$ to the number of light jets in a given event. It should be noted that N_{comb} is different for every event, because $N_{l\text{-jets}}$ is not fixed in the SR.

For every combination i, reconstructed values $\not{E}_{x,i}$ and $\not{E}_{y,i}$ for the missing transverse momentum components are obtained. These are then combined into one variable, the neutrino weight ω, which is defined in Equation 7.2.

$$\omega = \frac{1}{N_{\text{comb}}} \sum_{i=0}^{N_{\text{comb}}} \exp\left[-\frac{\left(\not{E}_{i,x} - \not{E}_x^{(obs)}\right)^2}{\sigma_{\not{E}x}{}^2} - \frac{\left(\not{E}_{i,y} - \not{E}_y^{(obs)}\right)^2}{\sigma_{\not{E}y}{}^2} \right] \tag{7.2}$$

Here, $\not{E}_x^{(obs)}$ and $\not{E}_y^{(obs)}$ refer to the reconstructed values of the missing transverse momentum components in a given event. $\sigma^2_{\not{E}x}$ and $\sigma^2_{\not{E}y}$ denote the variances of both components of the missing transverse momentum taken from the simulated distributions. Both values are determined using information from the MC signal samples. A χ^2-fit of a Gaussian is performed to the histograms of $\not{E}_x$ and $\not{E}_y$ in the SR for both signal components. The fits can be found in Figure A.5 and Figure A.6 in the appendix.

It remains to be studied, how well the neutrino weight ω discriminates between the signal and the background. Figure 7.5 shows the shape of the neutrino weight in the combined SR and the ROC curve with respect to the *utH* signal. Analogous plots, with respect to the *ctH* signal, are shown in Figure A.7 in the appendix. The neutrino weight shows promising discriminatory power for both signal processes.

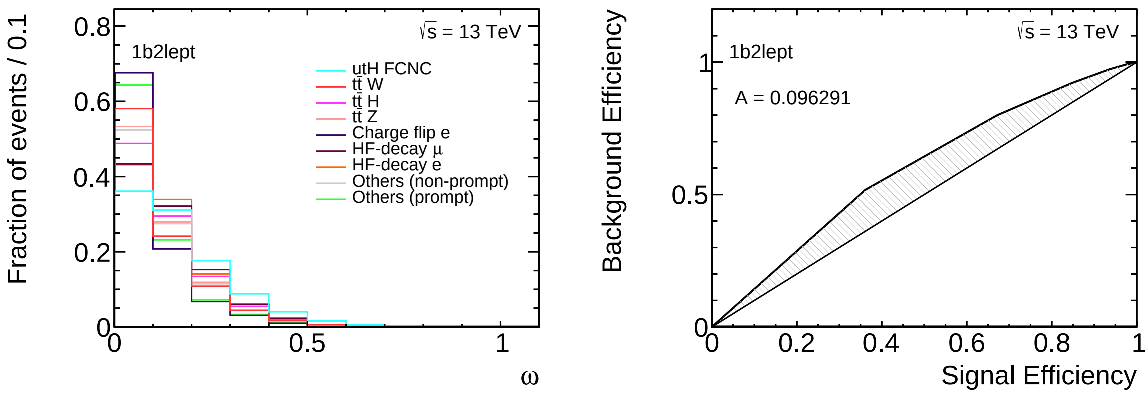


Figure 7.5: Shape plot and ROC curves of the neutrino weight in the combined SR. The area A between the ROC curve and the diagonal is given in the respective plot. Only the *utH* signal is shown.

7.4 Fox-Wolfram Moments

As a final effort of reconstruction, the so-called Fox-Wolfram moments [66] are taken into account. They are generally used to parametrize angular correlations between reconstructed objects in particle colliders such as the LHC. In this thesis, the Fox-Wolfram moments of the two leptons in the 2ℓSS final state are considered. Equation 7.3 shows the definition of the l'th Fox-Wolfram moment for N particles.

$$H_l = C \cdot \sum_{i,j=1}^{N} |\vec{p}_i||\vec{p}_j| P_l(\cos\Omega_{ij}) \tag{7.3}$$

Here C is a normalization constant, $\vec{p}_i$ and $\vec{p}_j$ are the momenta of particles $i/\ j$. P_l is the Legendre polynomial of degree l. $\cos\Omega_{ij} = \cos\theta_i \cos\theta_j + \sin\theta_i \sin\theta_j \cos(\varphi_i - \varphi_j)$ measures the angle between two particles i and j. There are several different possible choices for the normalization constant C, which are discussed in Ref. [66]. In this thesis, the normalization factor is chosen as follows, where E_i is the energy of particle i:

$$C = \left(\sum_{i=1}^{N} E_i \right)^{-1} . \tag{7.4}$$

It is found that the first Fox-Wolfram moment shows the highest discriminatory power. A shape plot of the variable in the combined SR as well as a corresponding ROC curve are depicted in Figure 7.6. Even though the distributions do not look as promising as those for the neutrino weight or the kinematic variables of W_{had}, the Fox-Wolfram moment proves useful in the following endeavor to separate signal from background.

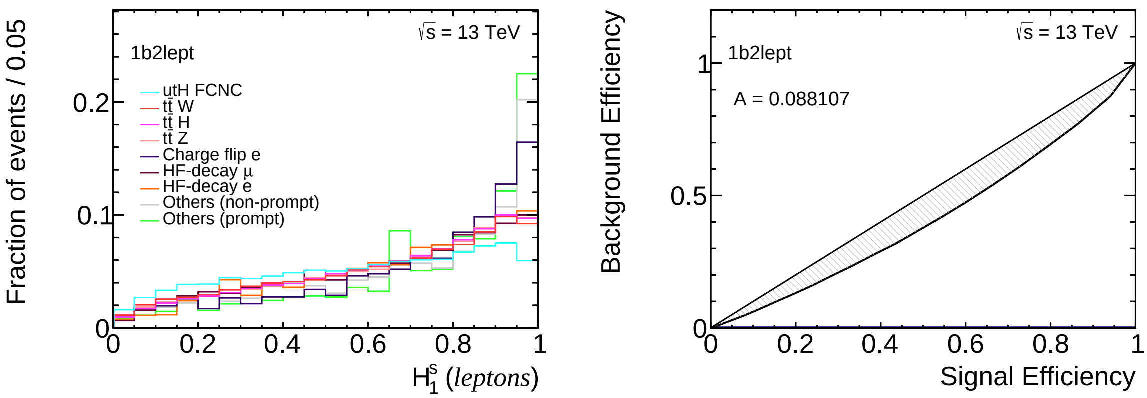


Figure 7.6: Shape plot and ROC curves of the first Fox-Wolfram moment of the two leptons H_1 in the combined SR. The area A between the ROC curve and the diagonal is given in the respective plot. Only the utH signal is shown.

8 Separation of signal and background

After a number of discriminating variables has been found in Chapter 7 via several methods of reconstruction, these variables are used in an effort to find the most optimal separation of signal and background processes in the signal regions. To combine the information from these and various other kinematic variables, artificial neural networks (NNs) are implemented in several regions.

This chapter gives an introduction into the basic functionality of feed-forward NNs and the framework Neurobayes, which is used to train NNs in this thesis. This is followed by a discussion of the NNs which are trained and used in relevant signal and control regions.

8.1 Neural Networks

NNs are used to process large amounts of information. They have a wide scope of application, but are especially useful in particle physics to combine a large number of input variables into few (often one single) discriminants [67].

An NN primarily consists of so-called nodes, which can assume values in a given finite range. In this thesis, an exemplary range between -1 and 1 is considered. These Nodes are organized in layers: One input layer, a number of hidden layers and one output layer. The values of nodes in the input layer $a_i^{(0)}$ (where $_i$ indicates the i'th node and $^{(0)}$ the 0'th layer) are given by the input variables. The values of nodes in any other layer l are determined as a linear combination of the nodes in the previous layer $l-1$. The factors connecting the i'th node of layer $l-1$ and the j'th node of layer l are called weights $\omega_{ij}^{(l)}$. They are generally not constrained and can assume any real value. Equation 8.1 shows the formula, by which the value of the node j in the layer l is calculated [68].

$$a_j^{(l)} = S\left(\sum_{i=1}^{n_{l-1}} \omega_{ij}^{(l)} \cdot a_i^{(l-1)} + \mu_j^{(l)}\right) \tag{8.1}$$

Here $\mu_j^{(l)}$ is a so-called bias, which is added to the linear combination and n_{l-1} is the number of nodes in the layer $l-1$. $S(x)$ is a monotonous activation function, which assumes values in the defined domain of the nodes[1], in this case between -1 and 1. Figure 8.1 shows a schematic of an NN consisting of several hidden layers.

[1] One example for such a function is $S(x) = \frac{2}{1+e^{-x}} - 1$. It has the additional property of being continuous.

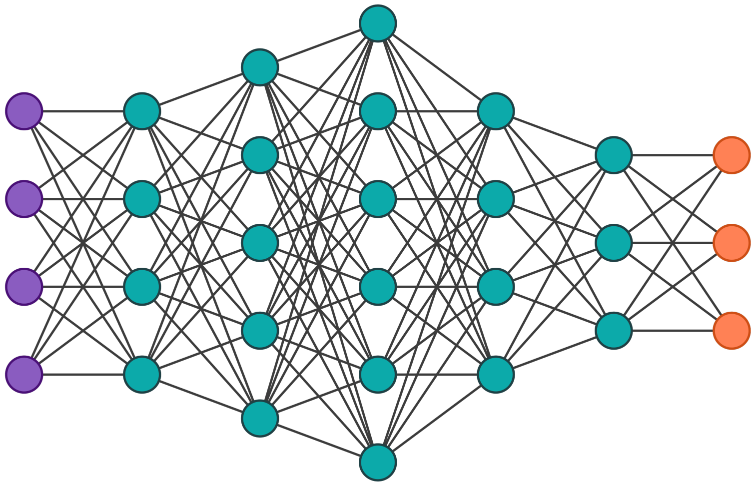


Figure 8.1: Schematic of an NN with four input nodes, three output nodes and 5 hidden layers. The input layer is depicted in purple, the output layer in orange and the hidden layers in cyan.

To produce usable results, an NN needs to be trained with (often large) sets of training data [68]. The aim of the training of an NN is to optimize the weights and biases, such that a given input produces a desired output. Hence, the training data needs to contain the correct output (target) for a given set of input values.

The training is based on an error function, which compares the output $\vec{o} \in [-1,1]^t$ that is produced by the NN to the target $\vec{\tau} \in [-1,1]^t$ for all events. Here t is the number of output nodes. A commonly used error function is given by the average of the squared differences between the target and the NN output over all training samples. It is shown in Equation 8.2.

$$E(\boldsymbol{\omega}, \boldsymbol{\mu}) = \frac{1}{N} \sum_{i=1}^{N} (\vec{\tau}_i - \vec{o}_i)^2 \qquad (8.2)$$

Here, $\boldsymbol{\omega}$ and $\boldsymbol{\mu}$ denote the vectors of all weights and biases. N is the number of training samples. The training now aims to minimize the error function. It generally does so by using a method called gradient decent, which changes each weight and bias by some number, which is proportional to the derivative of the error function with respect to that weight or bias (see Equation 8.3). Here, η is the step width, which needs to be optimized with respect to the minimization speed and the ability to find a global minimum instead of a local minimum.

$$\Delta\omega_{ij}^{(l)} = -\eta \cdot \frac{\partial E}{\partial \omega_{ij}^{(l)}} \qquad \Delta\mu_{j}^{(l)} = -\eta \cdot \frac{\partial E}{\partial \mu_{j}^{(l)}} \qquad (8.3)$$

8.2 Neurobayes

Neurobayes [69] is a framework designed for the training and application of NNs. It was developed specifically for applications in elementary particle physics and is thus very suitable for this analysis. Neurobayes is split into two large sub-structures: The teacher and the expert. The Neurobayes teacher trains an NN based on provided input data, while the expert is used to access the trained network. This section describes the training procedure of the Neurobayes teacher.

Neurobayes takes the values of various kinematic variables as input and produces a single output variable, which ideally combines the discriminatory power of all input variables. The events, from which the input variables are calculated, need to be classified either as signal or as background events for the training.

In addition to performing the training as described in the previous section, Neurobayes applies a number of pre-processing algorithms to the input. These algorithms identify the most discriminating input variables and perform a number of transformations on them. This decreases the amount of information that the NN has to process and allows for successful trainings with a relatively low amount of training data. Additionally, the overall scope of the NN is smaller due to the pre-processing, such that Neurobayes only produces NNs with one hidden layer.

At first, individual input variables are subject to the pre-processing. The variables are histogrammed with an irregular bin width. The bin width is adapted such that a flat distribution of events is created. This means that approximately the same number of events (signal + background) are included in every bin. Figure 8.2 shows an exemplary result from the automatic binning for the $p_T(\ell_1)$ input variable in an NN trained in the combined SR for the utH signal.

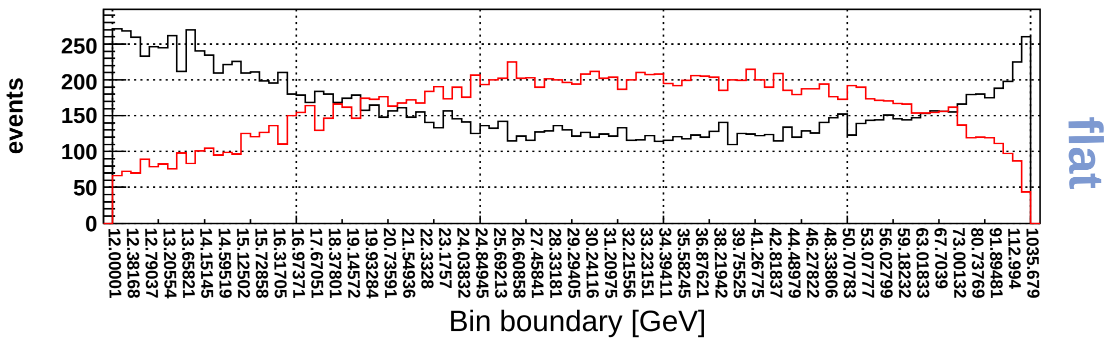


Figure 8.2: Examplary result from the binning in the single-variable pre-processing of Neurobayes. The variable used is $p_T(\ell_1)$ in the combined SR for the utH signal. The number of signal (red) and background (black) events per bin are plotted against the bin number.

After the automatic binning, the purity of signal events (i.e. the number of signal events divided by the total number of events) is calculated for each bin. Additionally, a spline fit is performed through this purity curve. Figure 8.3 shows the puritiy curve and the spline fit for the example variable $p_T(\ell_1)$. A flat line in the purity corresponds to no difference between the signal and the background in the variable. Hence, a shape similar to the one in Figure 8.3 is desired.

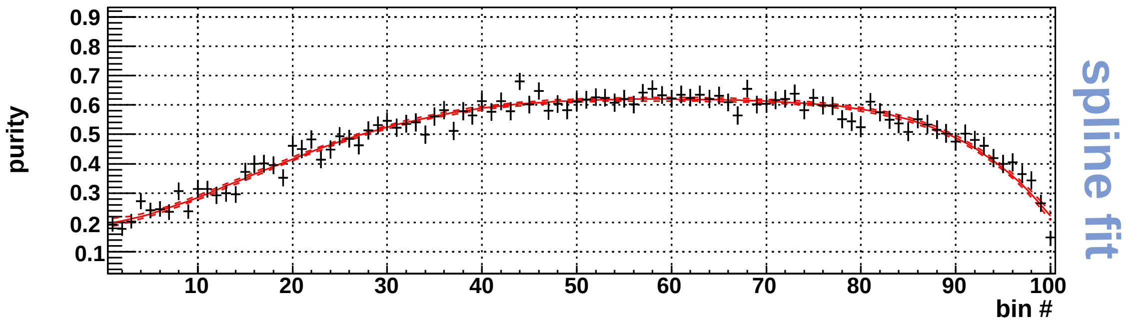


Figure 8.3: Examplary purity plot and spline fit from the single-variable pre-processing of Neurobayes. The variable used is $p_T(\ell_1)$ in the combined SR for the utH signal. The purity is given by black data points, the spline fit is depicted as a red line.

As a final step in the single-variable pre-processing, Neurobayes transforms the distribution of the variable, along with the spline fit. A distribution with a mean of 0 and a variance of 1 is created, which is then used as an input for the NN. Figure 8.4 shows the final input variable for the example variable $p_T(\ell_1)$.

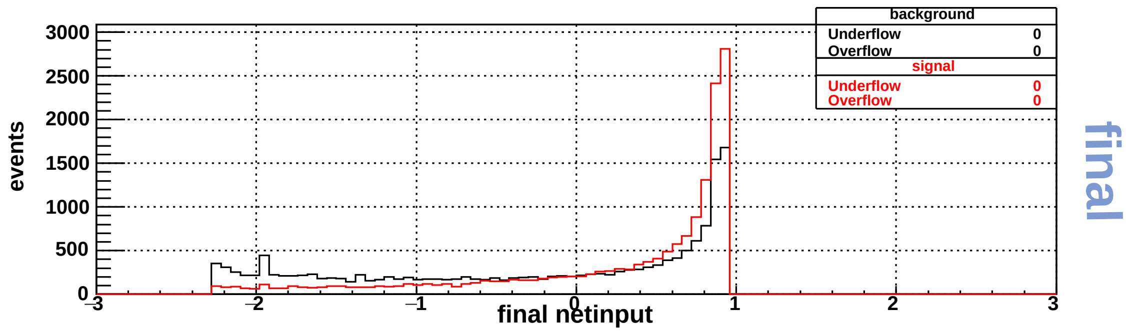
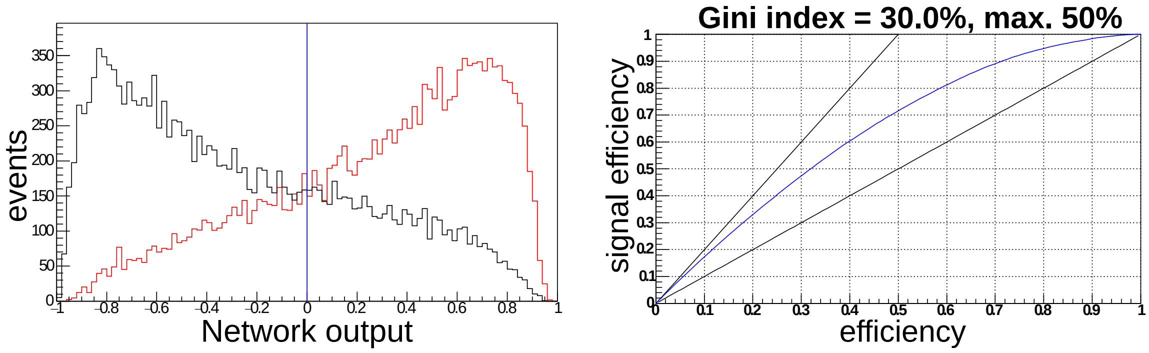


Figure 8.4: Distribution of the final NN input variable from the single-variable pre-processing of Neurobayes. The variable used is $p_T(\ell_1)$ in the combined SR for the utH signal. The distributions of the signal (red) and the background (black) are plotted separately.

After every variable has been pre-processed as described above, additional global pre-processing steps are taken to ensure an optimal training result. At first, the number of variables that are used in the final network is reduced. This process is called pruning of input variables. Prior to the actual pruning, however, Neurobayes sorts all input variables by their significance for the training. For this, the entire correlation matrix for all N input variables is calculated. Additionally, the total correlation to the target variable, which is defined as 1 for signal events and -1 for background events, is computed. Then, one by one, each variable is excluded and the correlation matrix is calculated for the remaining $N-1$ variables. The variable, for which the total correlation to the target decreases the least, when it is excluded, is identified as the least important variable. It is then removed from the set of variables, such that only the $N-1$ most significant variables remain. This procedure is repeated until only one variable remains. The process results in an implicit ranking of all variables by their added significance. The added significance is defined as the correlation to the target that is lost, when the variable is

removed in the ranking process described above, multiplied by $\sqrt{n}$, where n is the sample size. The actual pruning now excludes all variables from the training which contribute an added significance that is smaller than a predefined threshold. This threshold is defined as 3σ in this thesis.
All variables are then decorrelated as a final step before the actual training. This means, that the variable space is rotated, such that the correlations of the individual input nodes is zero. This way, the NN only has to learn effects of higher order.
After the training, Neurobayes produces information on the success and the quality of the training. Most notably, the shape of the signal and the background for the output variable o_{NN} of the NN are plotted (see Figure 8.5 a). In this plot, the signal and the background are scaled to the same total number of events for better comparison. Additionally, the so-called Gini index is given as a measure of how well o_{NN} separates signal and background from each other. It is calculated by plotting the signal efficiency against the total efficiency. The area between the curve and the diagonal corresponds to the Gini index (see Figure 8.5 b).



(a) (b)

Figure 8.5: Example output from Neurobayes for an NN trained for the utH signal in the combined SR. The distribution of the NN output o_{NN} (a) and a plot of the signal efficiency against the total efficiency showing the Gini index (b) are depicted. The signal distribution is plotted in red, while the background distribution is plotted in black.

To adapt the training process, Neurobayes includes a set of hyperparameters. These can modify the structure of the NN or details in the teacher's training procedure. The hyperparameters and their effects are as follows.

Iterations	The number of training iterations, i.e. the number of times, all training patterns are presented to the teacher. A common value for this parameter is 100.
Speed	A multiplicative factor, by which the dynamic learning speed of the teacher is modified. The speed can assume values between 0.1 and 1000.
Maxlearn	The maximum learning speed. If the (modified) learning speed of the teacher exceeds this value, it instead assumes the Maxlearn value.

Weight update	The number of events which are processed until the weights of the NN are updated. The nominal value of this parameter is 200.
Momentum	A momentum is an addition to the weight update of the NN training. The momentum term is proportional to the weight update of the previous iteration. The momentum parameter provides the proportionality factor, which can assume values in the interval $[0,1)$.
Hidden nodes	The number of nodes in the hidden layer of the NN.

8.3 Neural Networks in signal regions

Several NNs are trained in this analysis. Most of them are trained in the main SR, to optimize the separation between the utH or ctH signal and background processes. Only the flavor-split SR and the combined SR are taken into account. For the SR-3jet μe and SR-3jet $\mu\mu$ the MC-statistics for training an NN are found to be too low. As kinematics for the two signal processes utH and ctH might be different, two separate NNs are trained in each region. In total, ten NNs are trained in various signal regions. All variables presented in Chapter 7 are used as input variables. Additionally, a large number of kinematic variables is given as an input. These variables include the p_T of all reconstructed objects as well as the angular separation and the invariant masses between all pairs of reconstructed objects. Out of all variables, only a certain set is kept after the variable pruning that is done by Neurobayes. Table B.2 in the appendix shows this set of variables for the two NNs trained in the combined SR, along with the added significance for each variable. A similar table for the NNs trained in the flavor-split SR can be found in the appendix in Table B.3. In general, the p_T of the sub-leading lepton ℓ_1 is the most discriminating variable. However, various other kinematic variables, including the ones defined in Chapter 7, are important for the training. This shows that the efforts made in the separation of signal and background proved to be fruitful.

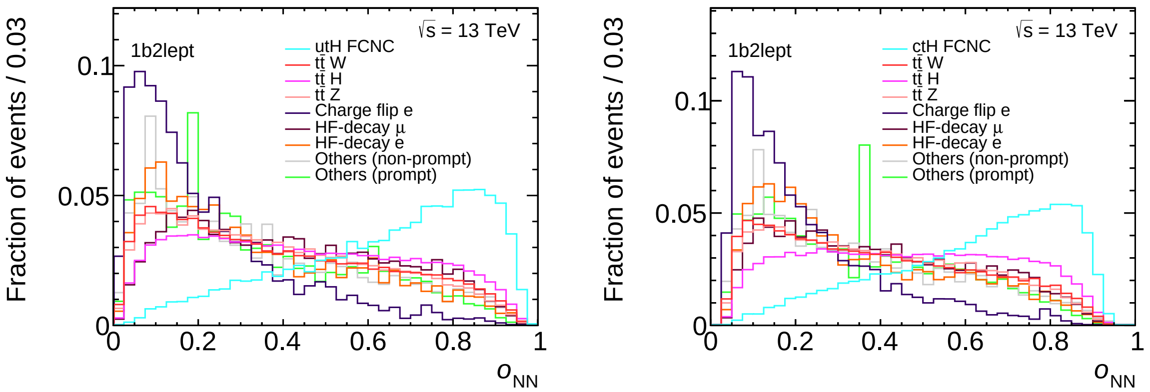
Prior to evaluating the NN-output, the set of hyperparameters that is used needs to be optimized. This is done by choosing a number of different values for each hyperparameter and training an NN for every combination. The training results are then compared. As a primary selection, the ten NNs with the highest Gini index are identified for each region. Out of these ten NNs the one with the most optimal shape and purity is chosen by hand. Table 8.1 shows the optimal set of hyperparameters for all NNs that are trained in the combined and the flavor-split SR.

The optimized neural networks are used as the final discriminator between signal and background in the various signal regions. Figure 8.6 shows the shape plots of the output variable o_{NN} for the NNs trained in the combined SR to give an estimate of their performance. The corresponding o_{NN} shape plots for the NNs trained in the flavor-split signal regions are shown in the appendix in Figure A.8 and Figure A.9. It is observed that the trained NN separates the signal process very well from the background. Even individual background processes are well separated from the signal, with some processes being excluded more efficiently than others. Especially the remaining charge-flip e background is separated against very well. A likely reason for this is the inclusion of the ECIDS result of both leptons as an input variable.

Table 8.1: Optimal set of hyperparameters for all NNs trained in the SR.

utH signal					
Hyperparameter	comb. SR	SR ee	SR $e\mu$	SR μe	SR $\mu\mu$
Gini index	30.0	31.5	30.2	30.2	29.3
Iterations	100	100	100	100	100
Speed	50	8	8	4	50
Maxlearn	1	1	1	3	1
Weight update	200	200	200	200	200
Momentum	0.85	0.25	0.80	0.85	0.85
Hidden nodes	30	25	30	25	30
ctH signal					
Hyperparameter	comb. SR	SR ee	SR $e\mu$	SR μe	SR $\mu\mu$
Gini index	29.7	31.1	29.9	30.0	28.4
Iterations	100	100	100	100	100
Speed	50	4	25	4	25
Maxlearn	1	3	1	3	1
Weight update	200	200	200	200	200
Momentum	0.85	0.90	0.85	0.90	0.80
Hidden nodes	15	25	25	30	30

The NNs appear to perform worst for the $t\bar{t}Z$-background process. This is likely caused by the promptness of the background, which makes a separation based on lepton kinematics more difficult. Overall, the performance of the NNs is satisfactory.



(a) *utH* signal

(b) *ctH* signal

Figure 8.6: Shape plots of the NN output variable o_{NN} for the two NNs trained in the combined SR. One training was done for the *utH* signal (a) and one for the *ctH* signal (b).

8.4 Neural Networks in control regions

NNs not only help in separating signal from background processes, but can be used to select for specific background processes in control regions. This analysis uses one NN to separate the $t\bar{t}X$ background from other (mainly non-prompt) background sources in the CR-4jet μx. As this NN is only loosely related to the signal sensitivity, the default values for all hyper-parameters are used. Still, the overall separation of the various $t\bar{t}X$ composites from other background sources is more than sufficient. Figure 8.7 shows the distribution of the output variable of the NN for the $t\bar{t}X$ processes and the combination of all other processes. This depiction is chosen, instead of the previous depiction with all processes plotted separately (cf. Figure 8.6), because the low statistics for some processes in the region cause the NN to be very inefficient in their individual separation, making such a plot very chaotic.

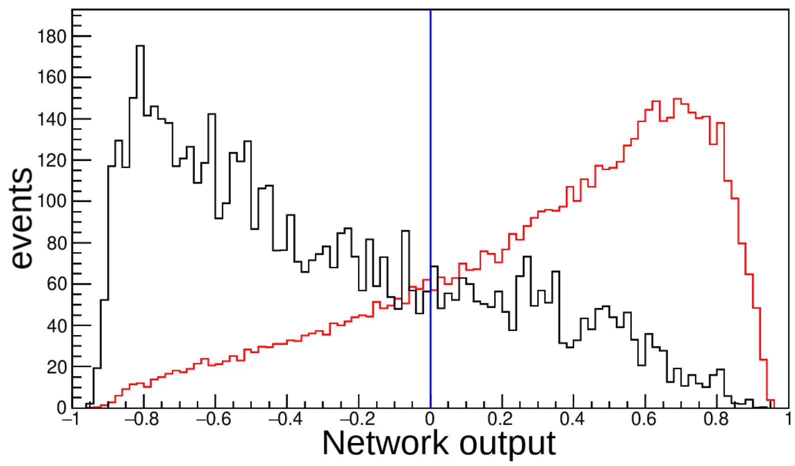


Figure 8.7: The NN output variable for the NN trained to separate the $t\bar{t}X$ background from other processes in the CR-4jet μx. The $t\bar{t}X$ processes are marked by the red line, the combination of all other processes by the black line. The number of $t\bar{t}X$-events is scaled to the number of events from other processes for better visualization of the NN's separation power.

9 Systematic uncertainties

Systematic uncertainties are uncertainties which arise from the nature of the experimental setup and the methods used to analyze experimental data. Hence, they cannot be reduced by the repetition of the experiment, as it can be done for statistical uncertainties.

In this analysis, systematic uncertainties arise from many various sources. These range from theoretical models of the hard scattering and soft hadronization processes to efficiency corrections in the final event selection. A systematic uncertainty can have an effect on the normalization as well as the shape of the final distributions of kinematic variables. To take those effects into account, a parameter that is subject to a systematic uncertainty is generally increased or decreased by one standard deviation. This is referred to as the up- and down-variation of this systematic uncertainty [70]. The entire MC-simulation is repeated for the up- and down-variation of a single parameter, while all other parameters are kept at their nominal value. All relevant kinematic variables are recalculated for these new MC samples. The resulting distributions are compared to the nominal variable distributions. In general, differences occur in their shape and normalization.

Some parameters can only be varied in one direction. They are generally referred to as one-sided systematic uncertainties. To estimate their influence, the distributions of variables are calculated for the variation in one direction and subsequently symmetrized around the nominal distribution. In the following, all systematic uncertainties taken into account in this thesis are discussed.

9.1 Luminosity uncertainties

The integrated luminosity is required for the correct normalization of MC events with respect to the ATLAS data. The nominal value of the integrated luminosity is determined using the van der Meer method [71]. Systematic variations are obtained by comparing its result to those of alternative methods. The total variation amounts to $\pm 1.7\,\%$ with respect to the nominal value of $L_{\text{int}} = 139\,\text{fb}^{-1}$. The uncertainty on the luminosity results in a normalization uncertainty exclusively, as all MC events are scaled by the same factor.

9.2 Pile-up uncertainties

To account for possible deviations from data due to the estimation of the pile-up distribution (cf. Chapter 3), a corresponding scale factor is applied to the MC samples. This scale factor is determined by previous MC-data comparisons [72]. The uncertainty on the scale factor is taken into account in the analysis.

9.3 Lepton uncertainties

Differences between MC simulation and data can originate from reconstruction, identification, isolation and trigger efficiencies of leptons. These differences are corrected using scale factors, which are determined from tag-and-probe techniques on the $Z \rightarrow \ell^+\ell^-$ resonance [73]. Uncertainties on these scale factors are taken into account by using the tools described in Ref. [74].
In addition to the respective scale factor uncertainties, systematic uncertainties for the calculation of the energy scale and resolution of electrons are considered. Both are determined using the tool in Ref. [75]. It specifies two systematic variations: `EG_SCALE_ALL` for the energy scale and `EG_RESOLUTION_ALL` for the energy resolution. The energy scale for MC simulations using the fast simulation technique AFII is varied separately and called `EG_SCALE_AF2`.
The momentum of the muon is determined using the tool in Ref.[76]. It provides several systematic variations, which need to be accounted for: `MUON_ID` and `MUON_MS` for the smearing of tracks in the inner detector and the muon spectrometer as well as `MUON_SCALE` for the scale of the muon's transverse momentum. `MUON_SAGITTA_RHO` and `MUON_SAGITTA_RESBIAS` are uncertainties on the muon sagitta[1] corrections.

9.4 Jet uncertainties

Systematic variations on the JES and the JER (cf. Chapter 5) are taken into account [78]. For the JER, seven different, one-sided systematic variations are provided for different p_T bins of the given jet. For the JES, a priori over 100 sources of systematic uncertainty are defined. These originate most notably from the energy scale calibration of jets. Additionally, they arise due to pile-up, the jet-flavor and detector effects. These sources of systematic uncertainty are partially combined into 30 different variations, which are considered in this analysis [79].

9.5 Flavor tagging uncertainties

The `DL1r` *b*-tagging algorithm provides a large number of systematic variations [80]. These include *b*-tagging and mis-tag efficiencies for a large number of bins on the continuous output variable of the algorithm. In total, 16 variations for light jet, 20 variations for *c*-jet and 42 variations for *b*-jet efficiencies are included in the analysis.

9.6 Missing transverse momentum uncertainties

As the missing transverse momentum is calculated mainly from the reconstructed transverse momenta of other objects, most sources of systematic uncertainty in the missing transverse momentum calculation are already considered for the corresponding objects. There are, however, additional uncertainties arising from the calculation of the soft term, which

[1] The sagitta is a measure of the curvature of a charged particle's track in a magnetic field. It is defined as the shortest distance between the reconstructed particle track and the line connecting beginning and end of the track (inside the magnetic field) [77].

is calculated from unassigned tracks in the inner detector [79]. The uncertainties on these soft tracks are called `MET_SoftTrk_Scale` for the p_T amplitude of soft tracks as well as `MET_SoftTrk_ResoPara` and `MET_SoftTrk_ResoPerp` for the resolution of the p_T of tracks parallel and perpendicular to the mean of all hard tracks. The uncertainties on the resolution of tracks are one-sided and correspondingly symmetrized.

9.7 Theoretical cross section uncertainties

The theoretical cross section used for the normalization of MC simulations is calculated with limited accuracy. This results in normalization uncertainties for all SM processes. As mentioned in Chapter 6, the normalization of the $t\bar{t}X$ and the non-prompt background is determined using specific control regions. Hence, cross section uncertainties for these processes are not taken into account. Furthermore, the number of prompt Z+jets and prompt W+jets is negligible in all regions. Hence, the only processes for which a cross section uncertainty is considered are the prompt $t\bar{t}$ and the prompt diboson processes. For prompt $t\bar{t}$ events, an uncertainty of ${}^{+5.5\,\%}_{-6.1\,\%}$ is considered [81]. For prompt diboson events, an uncertainty of ${}^{+2.6\,\%}_{-2.5\,\%}$ is used [82].

9.8 Modeling and parton distribution function uncertainties

Additional systematic uncertainties arise from the MC modeling of physical processes as well as the assumed parton distribution functions (PDF). The PDF are used to model the substructure of the colliding protons [83]. These sources of uncertainty are not taken into account in this thesis due to time limitations. They will be taken into account in the future of this analysis. Despite this incomplete estimation of all uncertainties, the included systematic uncertainties should still allow for a more realistic approach than a mere consideration of statistical uncertainties.

10 Statistical Analysis

This chapter outlines the methods of statistical analysis that are used in this thesis. At first, general information on binned profile likelihood fits is presented, with special focus on methods to incorporate systematic uncertainties into the fit. Additionally, details on the construction of an Asimov dataset are given. Lastly, the method by which the upper limit on a given signal strength is determined is presented.

10.1 Binned template profile likelihood fits

Maximum likelihood fits are a commonly used method to find suitable estimators for parameters of statistical distributions [84]. In particle physics, complex binned maximum likelihood fits with large numbers of parameters are performed. Most notably, the signal strength μ as well as normalization factors β_X for a given, potentially mis-modeled background process X are parameters in the fit. More precisely, they are parameters in the Poisson distribution, which is assumed for the number of entries in every bin.
Systematic variations are estimated as described in Chapter 9. For a given dataset, certain values of a systematic variation might be more suitable than the nominal one. This is taken into account in the fit via so-called nuisance parameters (NPs) [85]. These NPs estimate, how much the up- or down-variation of a systematic uncertainty improves the overall fit result. To do this, the systematic variations are interpolated by a continuous function between the up- and the down-variation. For shape uncertainties, a piecewise linear interpolation is performed, for normalization uncertainties the extrapolation is performed via a polynomial or an exponential function. The NPs are implemented into the fit via Gaussian distributions with a mean of 0 and a standard deviation of 1. A value of 0 indicates that the nominal value of the systematic uncertainty produced optimal fit results, while a value of 1/-1 indicates that the up-/ down-variation template worked best for the fit. Based on these considerations, the final likelihood function for the fit assumes the shape in Equation 10.1.

$$\mathcal{L}(\mathbf{n},\boldsymbol{\theta}^0|\mu,\boldsymbol{\beta},\boldsymbol{\theta}) = \prod_{i=1}^{N_{\text{bins}}} \mathcal{P}\left(n_i|\mu \cdot S_i(\boldsymbol{\theta}) + B_i(\boldsymbol{\theta},\boldsymbol{\beta})\right) \times \prod_{j=1}^{N_{\text{NP}}} \mathcal{G}(\theta_j^0|\theta_j,\Delta\theta_j) \tag{10.1}$$

$$B_i(\boldsymbol{\theta},\boldsymbol{\beta}) = \sum_{k=1}^{N_{\text{norm}}} \beta_k \cdot B_i^{(k)}(\boldsymbol{\theta}) + B_i^{(\text{rest})}(\boldsymbol{\theta}) \tag{10.2}$$

Here, $\mathcal{P}$ and $\mathcal{G}$ denote the probability density functions (pdf) of the Poisson and the Gaussian distribution. $\mathbf{n}$ and $\boldsymbol{\theta}^0$ refer to all observed bin contents and all nominal NPs. As mentioned above, $\theta_j^0 = 0$ is generally chosen. $\boldsymbol{\theta}$ denotes the vector of all post-fit values for the NPs. In total, there are N_{NP} NPs associated to systematic uncertainties included in the fit. μ and $\boldsymbol{\beta}$ are the signal strength and the vector of all normalization factors, by which the mis-modeling

of certain background processes is corrected. In contrast to NPs for systematic uncertainties, the normalization factors are not constrained by a Gaussian distribution and instead left free-floating in the fit. n_i refers to the measured number of events in bin i, while $S_i(\boldsymbol{\theta})$ and $B_i(\boldsymbol{\theta},\boldsymbol{\beta})$ refer to the predicted number of signal and background events in the respective bin. $S_i(\boldsymbol{\theta})$ is a function of the NPs, whereas $B_i(\boldsymbol{\theta},\boldsymbol{\beta})$ also depends on the normalization of certain process templates, as shown in Equation 10.2. There, $B_i^{(k)}(\boldsymbol{\theta})$ denotes the contribution from a background template, which is normalized by the factor β_k. $B_i^{(\text{rest})}(\boldsymbol{\theta})$ includes the combined contribution from all background processes, which are not corrected by any normalization factor. Finally, θ_j are the entries of $\boldsymbol{\theta}$ and $\Delta\theta_j = 1$ are the nominal widths of the Gaussian distributions, by which the NPs are modeled.
In addition to systematic uncertainties, further NPs are included to account for the MC statistics in every bin. These are referred to as γ factors [85]. They are implemented into the fit by complementing the Poisson pdf in the likelihood function with an additional Poisson pdf corresponding to the MC statistics. Hence, the following replacement is performed in Equation 10.1:

$$\mathcal{P}\big(n_i|\mu\cdot S_i+B_i\big)\longrightarrow\mathcal{P}\big(n_i|\gamma_i(\mu\cdot S_i+B_i)\big)\times\mathcal{P}\big(\mu\cdot S_i+B_i|\gamma_i(\mu\cdot S_i+B_i)\big)\,. \tag{10.3}$$

The fit is performed by minimizing the negative log-likelihood function. Post-fit uncertainties of all fit parameters (including NPs) are determined by the shape of the negative log-likelihood function around the minimum. The values of a given parameter, at which the negative log-likelihood function is larger than the minimum by 0.5, define the 1σ uncertainty in both directions. This approach is based on the Cramer-Rao bound [86]. In some cases the post-fit uncertainty of NPs can be different from the pre-fit uncertainty. If the post-fit uncertainty is smaller, the NP is constrained by the fit. In general, constrained NPs have a stronger effect on the fit result than unconstrained NPs. Furthermore, if the post-fit result for an NP is different from zero, it is experiencing a "pull" to either side of the distribution.

Before a fit is performed in this analysis, the following algorithms are applied to improve its general performance and reliability.

Binning

The binning of variables, in which a likelihood fit is performed, is of special importance. A high number of bins can increase the overall signal sensitivity, while it also decreases the MC statistics per bin. To obtain an optimized binning for every variable, the `TransfoD` algorithm [87] is used. It creates bins of variable width, such that every bin contains approximately the same number of MC events. The algorithm can be specified regarding how strongly it takes into account signal events in the determination of the binning. In this analysis, both the signal and the background are taken into account equally strong.

Smoothing

Often, the MC statistics for systematic variations are lower than for nominal MC samples. Because of this, distributions for the variations have relatively high statistical fluctuations. Since the above defined γ factors only take into account the MC statistics of the nominal histograms, so-called smoothing algorithms are applied for compensation of the effect. They reduce the statistical fluctuations in various ways, for instance by performing a rebinning or replacing fluctuating bins by the kernel-weighted average [88].

Symmetrization

Generally, the histograms of the up- and down-variation of a certain systematic uncertainty are asymmetric around the nominal histogram. In some cases, the overall fit stability can be increased by defining a symmetric variation in every bin i [89]. It is determined by $\Delta n_i = \frac{|n_i^{\mathrm{up}} - n_i^{\mathrm{down}}|}{2}$, where n_i^{up} and n_i^{down} are the entries in bin i of the histograms produced with the up- and down-variation. This process is called symmetrization. Whether or not the symmetrization of a systematic uncertainty is warranted is decided on a case-by-case basis, depending on the shape of the variation histograms and the overall stability of the fit.

Pruning

Some systematic uncertainties do not have any influence on the fit. If the corresponding NPs are still included, the overall duration of the fit may decrease. Because of this, such NPs can be excluded from the fit (i.e. pruned) [88]. The condition for a shape uncertainty to be pruned is a difference of less than 1 % compared to the nominal sample in every bin. A normalization uncertainty is pruned, if the total number of events differs by less than 1 % from the nominal number of events.

10.2 Asimov dataset

An Asimov dataset [90] is generated based on the predictions made by MC simulations. The dataset is constructed by taking into account the MC predictions of all processes and then summing these predictions up. It can then be used to test the outcome of fits, based on a given prediction.

An Asimov dataset is especially useful to test fit strategies and calculate expected results of an analysis, when no real data is available. This mainly concerns signal regions, which are a priori blinded from actual data to prevent biases in the analysis. At the same time, real data is used in control regions to estimate the MC mis-modeling of certain processes. This is possible, because control regions are defined such that they only contain little signal. The corrections of the MC mis-modeling are determined by a fit to data in the control regions exclusively. Afterwards, an Asimov dataset based on the corrected MC predictions and the post-fit NPs is created. Such a dataset yields different results compared to an Asimov dataset created using nominal MC predictions.

In a global fit to a nominal Asimov dataset no pulls on included NPs can be observed. This is because the fit will always determine the nominal values of all NPs to give the best MC-data agreement. However, if a fit is performed to an Asimov dataset that was created using corrected MC predictions, as described above, the pulls from the fit to data in the control regions can still be observed. This allows for an overall more realistic estimate of results that can be expected in a global fit to data. The most realistic fit setup, however, constitutes a simultaneous fit to data in control regions and to an Asimov dataset, based on the corrected

MC predictions, in the signal regions. It combines the most realistic Asimov dataset with real data, where it is available.

10.3 Exclusion limits and the CLs method

Once a fit has been performed, one of two results can be obtained. Either a signal has been detected, in which case the signal significance is determined. Should the significance be high enough, the associated cross section of the signal process is determined. The other possibility is the observation of a lack of any signal. In this case, it is possible to compute an upper limit on the signal strength. The so-called $\mathrm{CL_S}$ method [85] is commonly used to achieve this.
The Method tests a null hypothesis H_0 (e.g. lack of signal, $\mu = 0$) against a different hypothesis H_1 (e.g. presence of signal, $\mu > 0$). The test is based on the profile likelihood ratio (PLR) defined in Equation 10.4.

$$\lambda(\mu) = \frac{\mathcal{L}\left(\mu,\hat{\hat{\theta}}\right)}{\mathcal{L}\left(\hat{\mu},\hat{\theta}\right)} \in (0,1] \tag{10.4}$$

$\hat{\hat{\theta}}$ denotes the maximum likelihood estimators of all NPs for a fixed value of μ. $\hat{\mu}$ and $\hat{\theta}$ are the maximum likelihood estimators for the signal strength and all NPs that are obtained from the global fit, in which the signal strength is included in the fit. If μ is equal to the maximum likelihood estimator $\hat{\mu}$, the PLR $\lambda(\mu)$ yields a value of 1. For any deviation from $\hat{\mu}$, the value of the PLR decreases.
From the PLR a test-statistic for a given signal strength is constructed as follows:

$$q_\mu = \begin{cases} -2\ln\lambda(\mu) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \tag{10.5}$$

It should be noted, that for $\mu = \hat{\mu}$, the test statistic is equal to zero. For any deviation of μ from $\hat{\mu}$, q_μ increases. Based on this test statistic, $\mathrm{CL_B}$ and $\mathrm{CL_{S+B}}$ are defined for a given observed value of $q_{\mu,\mathrm{obs}}$ as follows.

$$\mathrm{CL_B} = \int_0^{q_{0,\mathrm{obs}}} f(q_0|0)dq_0 \qquad \mathrm{CL_{S+B}} = \int_0^{q_{\mu,\mathrm{obs}}} f(q_\mu|\mu)dq_\mu \tag{10.6}$$

Here, $f(q_\mu|\mu)$ denotes the pdf of q_μ for a given μ. Finally, $\mathrm{CL_S}$ is defined as the ratio of these two values:

$$\mathrm{CL_S} = \frac{\mathrm{CL_{S+B}}}{\mathrm{CL_B}}. \tag{10.7}$$

The value of μ, for which $\mathrm{CL_S} < p$ is fulfilled is defined as the $\mathrm{CL_S}$ upper limit with confidence level $1-p$. As a convention, a confidence level of 95 % is calculated and interpreted as the upper limit on the signal strength.

11 Results and interpretation

This chapter begins with a discussion of the fit strategy that is used in this thesis. Additionally, various possibilities of interpretation of the fit results are considered.

This is followed by a presentation of the fit results. At first, the result of a fit that only takes into account statistical uncertainties is examined. A comparison of the fit results that are obtained with the combined SR and results that are obtained with the flavor-split SR is conducted to determine, which region setup is used in the final fit.

This final fit is then performed under the inclusion of statistical and systematic uncertainties. The fit results are discussed, particularly with respect to the effect of individual systematic uncertainties on the fit result. Finally, expected upper limits on the branching ratios $\mathcal{B}(t \to Hu)$ and $\mathcal{B}(t \to Hc)$ are determined and compared to the current upper limits.

11.1 Fit strategy

The fits discussed in this chapter are performed using the statistical evaluation framework TRExFitter [91]. As discussed in Section 10.2, no data is used in the signal regions that have been defined in this analysis. However, data is available in control and validation regions, because their relative signal contamination is expected to be small, as is shown in Figure 11.1. The signal contamination is given for the utH signal, however the ctH signal shows similar behavior. All control regions have an S/B ratio of less than 5 %.

To estimate the mis-modeling of non-prompt and $t\bar{t}X$ processes, a fit of only the background templates to data is performed in the control regions. The post-fit values of all NPs (particularly the normalization factors for background processes) are then used to generate an Asimov dataset, to be used in the signal regions. As there is no knowledge about the true occurrence of the signal beyond the current upper limit, the SM prediction is used for the Asimov dataset. That corresponds to a signal strength of $\mu = 0$.

A final, global fit of all processes (including the signal) to real data in control regions and to the Asimov dataset in signal regions is then performed. This approach, as opposed to a global fit to the Asimov dataset, is chosen because it is the closest estimate to a global fit to real data that can be achieved in this setup.

The variable in which the fit is performed varies depending on the region. In all regions, in which an NN has been trained, a fit to the respective output variable o_{NN} is performed. In all other regions, a fit to $p_T(\ell_1)$ is performed. This variable is chosen, because it was identified as the most discriminating variable in most NN-trainings. For the SR-3jet μe and $\mu\mu$ regions, the variable is additionally part of the definition of the region and hence particularly important.

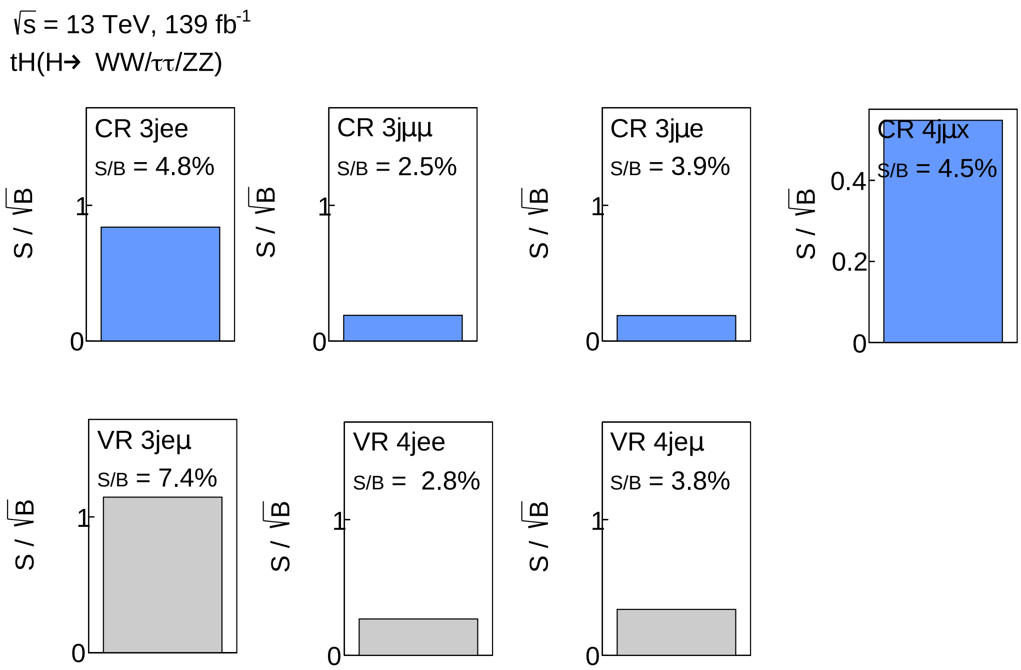


Figure 11.1: Pre-fit contribution from the utH signal process in control and validation regions. The utH signal is scaled to the current upper limit. The bars in every subplot indicate the $S/\sqrt{B}$ ratio, while the S/B ratio can be taken from the text in the plots.

11.1.1 Interpretation of fit results

As the Asimov dataset is generated with a signal contribution of zero, the post-fit value of the signal strength is expected to be close to zero as well. This value in itself does not contain any valuable information. It is, however, possible to construct an upper limit on the signal strength via the $\mathrm{CL_S}$ method. This limit corresponds to the expected upper limit, under the assumption that the SM predictions and corrections for the mis-modeling are correct. As the upper limit on the signal strength depends on the pre-fit normalization of the signal process, it does not contain any fundamental physical meaning on its own. It can, however, be transformed into a limit on the branching ratios for the utH and ctH FCNC decay. This requires knowledge of the general relation between the cross section of an FCNC process σ_{FCNC} and its corresponding branching ratio $\mathcal{B}(t \to Hq)$. This relation is given in Equation 11.1 [92].

$$\sigma_{\mathrm{FCNC}} = 2\,\sigma_{t\bar{t}} \cdot \mathcal{B}(t \to Hq)\mathcal{B}(t \to Wb) \tag{11.1}$$

Here $\sigma_{t\bar{t}} = (830 \pm 39)\,\mathrm{pb}$ [81] denotes the production cross section for a $t\bar{t}$ pair and $\mathcal{B}(t \to Wb)$ the SM branching ratio of a top quark decaying into a W boson and a b-quark.

Conventionally, the pre-fit signal strength is chosen to be $\mu = 1$. μ is proportional to the cross section of the signal process, which in turn is proportional to the branching ratio. Thus one obtains the relation in Equation 11.2.

$$\mu^{\text{lim}} = \frac{\sigma_{\text{FCNC}}^{\text{lim}}}{\sigma_{\text{FCNC}}^{\text{pre-fit}}} = \frac{\mathcal{B}(t \to qH)^{\text{lim}}}{\mathcal{B}(t \to qH)^{\text{pre-fit}}} \tag{11.2}$$

$$\Rightarrow \mathcal{B}(t \to qH)^{\text{lim}} = \mu^{\text{lim}} \cdot \mathcal{B}(t \to qH)^{\text{pre-fit}} \tag{11.3}$$

Here μ^{lim} is the obtained limit on the signal strength. $\sigma_{\text{FCNC}}^{\text{pre-fit}}$ and $\mathcal{B}(t \to qH)^{\text{pre-fit}}$ are the pre-fit values of the signal process cross section and the branching ratio. $\sigma_{\text{FCNC}}^{\text{lim}}$ and $\mathcal{B}(t \to qH)^{\text{lim}}$ are the upper limits on the cross section and the branching ratio. Thus, by knowing the pre-fit value of the signal process's branching ratio, one easily obtains a corresponding upper limit on the branching ratio. The signal is scaled to a certain cross section. Hence, it is necessary to calculate the pre-fit branching ratio using Equation 11.1 (with $\mathcal{B}(t \to Wb) = 1$):

$$\mathcal{B}(t \to qH)^{\text{pre-fit}} = \frac{\sigma_{\text{FCNC}}^{\text{pre-fit}}}{2\,\sigma_{t\bar{t}}}. \tag{11.4}$$

The pre-fit cross sections are corrected by a so-called K-factor to account for the effects of higher order corrections, since the signal samples were produced at next-to leading order in perturbation theory. Table 11.1 lists the cross sections and the K-factors for both signal processes. They are fixed by the sample production and do not correspond to the current upper limit.

Table 11.1: The next-to leading order cross sections and the K-factors to which the signal samples are scaled before the fit.

Sample	σ/ pb	K-factor
utH	11.84098	1.1945
ctH	12.19719	1.1945

11.1.2 Fits to be performed

As previously discussed, all efforts to separate signal from background have been performed in the combined SR as well as the flavor-split SR. It now remains to be seen, which of the two approaches results in a higher sensitivity to the signal. An estimate of this sensitivity can already be gained from a fit that only takes into account statistical uncertainties and ignores all sources of systematic uncertainty (a so-called stat-only fit). Hence, such a fit is performed for both configurations of the SR. The results are compared and the approach, which produces the lower expected upper limit is identified.

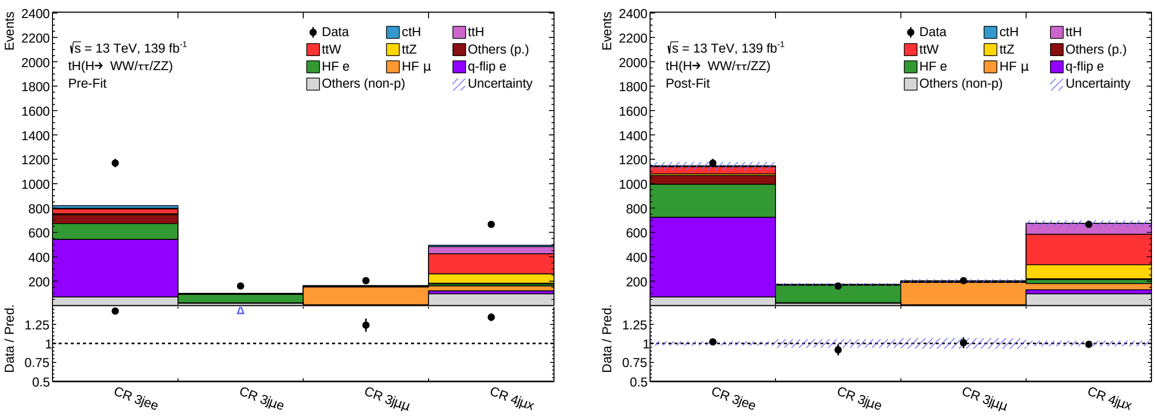
A final fit then considers only the better of the two approaches and includes all systematic uncertainties that were presented in Chapter 9. For this fit, the influence of individual systematic uncertainties is discussed. Finally, an expected upper limit on the branching ratios of the two signal processes is given.

11.2 Fit results - only including statistical uncertainties

This section describes the results of the stat-only fits. Firstly, the normalization factors of the various non-prompt and $t\bar{t}X$ templates from the background only fit are discussed. Secondly, the obtained upper limits on the branching ratios for the combined and the flavor-split SR are compared.

11.2.1 Background-only fit results

Before evaluating the differences between the combined and the flavor-split SR, it is reasonable to check how well the background-only fit results agree with data. The background-only fit is performed only in control regions, without the inclusion of any signal process. Hence, it produces the same results for both SR configurations and both signal templates.
Figure 11.2 shows pre- and post-fit summary plots of all control regions. In these plots, the regions are reduced to one bin. The fit, however, is performed with several bins in every region. The individual pre- and post-fit plots for every control region can be found in the appendix in Figure A.10 and Figure A.11. A reasonably good agreement with data in all post-fit plots is observed. This is a first indication that the template fit method works as expected and is capable of accurately correcting the MC mis-modeling.



(a) Pre-fit (b) Post-fit

Figure 11.2: Summary plots of all control regions included in the background-only fit with statistical uncertainties only. Individual bins of all regions are combined into one bin respectively. The pre-fit plot (a) includes the uncorrected MC predictions, while in the post-fit plot (b) the non-prompt and $t\bar{t}X$ templates are scaled by their respective normalization factors.

In addition to studying the agreement of the post-fit MC templates with data, the results that are obtained for the normalization factors of background processes are considered. The normalization factors are plotted in Figure 11.3. The magnitude of the correction of the $t\bar{t}X$ normalization compares well to that obtained in other analyses, which have corrected the mis-modeling of related processes in similar kinematic regions (cf. [60] for $t\bar{t}W$).

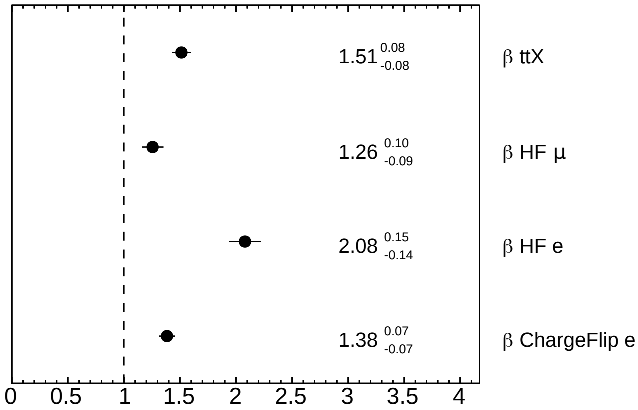
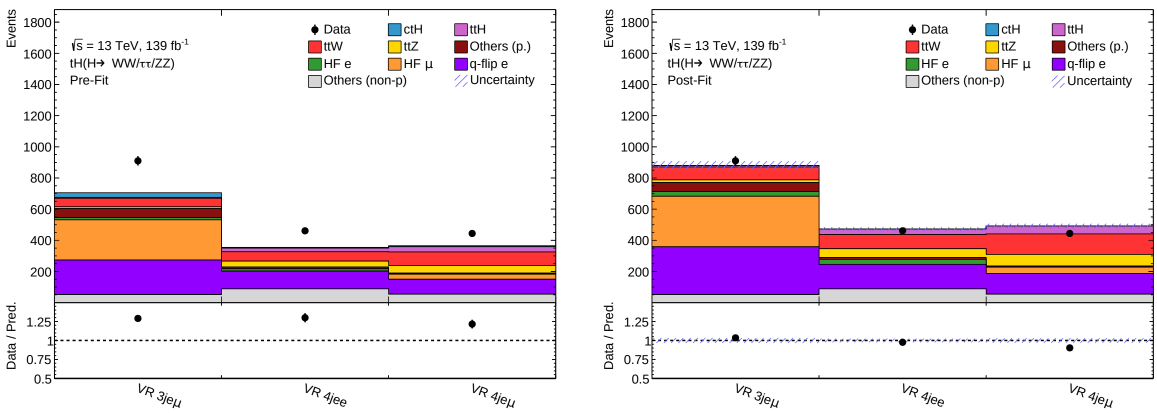


Figure 11.3: Normalization factors for the background processes that result from the background-only fit including statistical uncertainties only.

As a final check of the post-fit MC-data-agreement, the validation regions that are defined in Chapter 6 are examined. These are not included in the fit. Hence they provide a good method to check, if the obtained normalization factors reliably correct the respective processes. Especially their validity outside of the control regions is put under scrutiny in this way. Figure 11.4 shows pre- and post-fit summary plots of all validation regions. A good post-fit MC-data-agreement is observed in these regions as well. This final check serves as confirmation that an adequate strategy for the determination of background rates has been chosen in this analysis.



(a) Pre-fit (b) Post-fit

Figure 11.4: Summary plots of all validation regions with fit results from the background-only fit including statistical uncertainties only. Individual bins of all regions are combined into one bin respectively. The pre-fit plot (a) includes the uncorrected MC predictions, while in the post-fit plot (b) the non-prompt and $t\bar{t}X$ templates are scaled by their respective normalization factors.

11.2.2 Expected upper limits in combined and flavor-split signal regions

After an extensive evaluation of the background-only fit results, the expected 95 % CL_S upper limits on the branching ratio of both signal processes are computed. As mentioned previously, the upper limits are determined for two different region setups: one with a combined SR and one with a flavor-split SR.
The expected upper limits on the branching ratios of both signal processes are listed in Table 11.2 for the combined and the flavor-split SR. To obtain a better overview, all upper limits on the branching ratios are plotted in Figure 11.5. Even though the current upper limits are included in the plot, a comparison to them is not sensible at this point because of the exclusion of systematic uncertainties. Hence, only a comparison between the expected upper limits obtained with the two region setups is conducted. One can observe that there is almost no difference between the two region setups. This is true for all four signal templates. Especially when considering the uncertainty on all limits, the differences become too marginal to show any meaningful effect of the flavor-splitting.

Table 11.2: 95 % CL_S expected upper limits on $\mathcal{B}(t \to qH)$ obtained from a stat-only fit in the combined and in the flavor-split SR.

Region Setup	Couplings	$\mathcal{B}(t \to qH)$ / 10^{-4}				
		-2σ	-1σ	median	$+1\sigma$	$+2\sigma$
flavor-split SR	utH	1.8	2.4	3.3	4.7	6.3
	ctH	1.8	2.5	3.4	4.8	6.4
combined SR	utH	1.7	2.3	3.2	4.4	6.0
	ctH	1.8	2.5	3.4	4.8	6.5

It might have been expected that the flavor-split SR performs better, because the NNs in the subregions can be trained more effectively against the dominating background processes in each region. This assumption does not appear to hold. A possible reason for this is that the NN in the combined SR, which uses information on the flavor of both leptons, is able to sufficiently discriminate against all background processes simultaneously. Because of these results, the choice of which setup is used for further efforts in the analysis is ambiguous in terms of the expected results. As, however, the combined SR allows for a more manageable region setup and decreases the total number of bins, it is determined to be the more suitable choice. Hence, all further fits will only be performed for the combined SR.

11.3 Fit results - including systematic uncertainties

After the combined SR has been identified as the more adequate approach to a region setup, a final fit is performed. This fit includes the systematic uncertainties, which are listed in Chapter 9. This section begins by discussing the fit results from the background-only fit in all control regions. Especially the differences between the normalization factors determined in this fit and those from the stat-only fit are analyzed. Then, the influence of various systematic uncertainties on the fit is taken into consideration. Finally, the expected upper limits on the

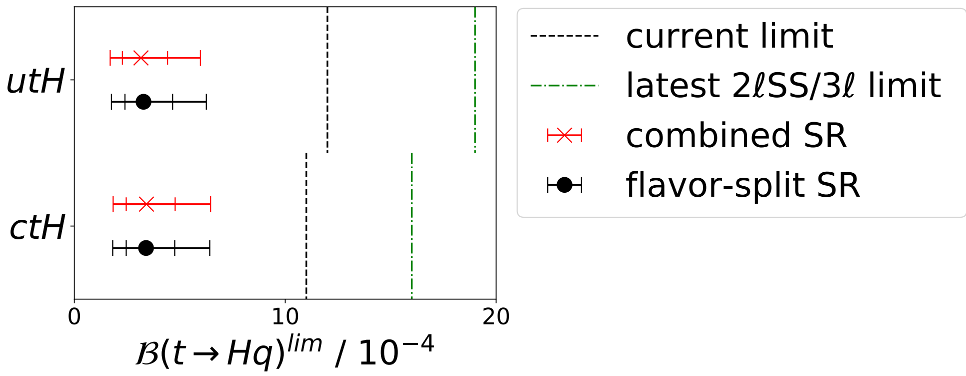


Figure 11.5: Expected 95 % $\mathrm{CL_S}$ upper limits on the branching ratio $\mathcal{B}(t \to Hq)$ for the utH and the ctH coupling from stat-only fits. The limits are given for both region setups: A combined SR (red crosses) and a flavor-split SR (black dots). The two errorbars for each limit depict the 1σ and the 2σ regions. The current upper limits [19] on the branching ratios are depicted as dashed lines. The latest upper limits obtained with multilepton final states are shown as green, dashed-dotted lines [14].

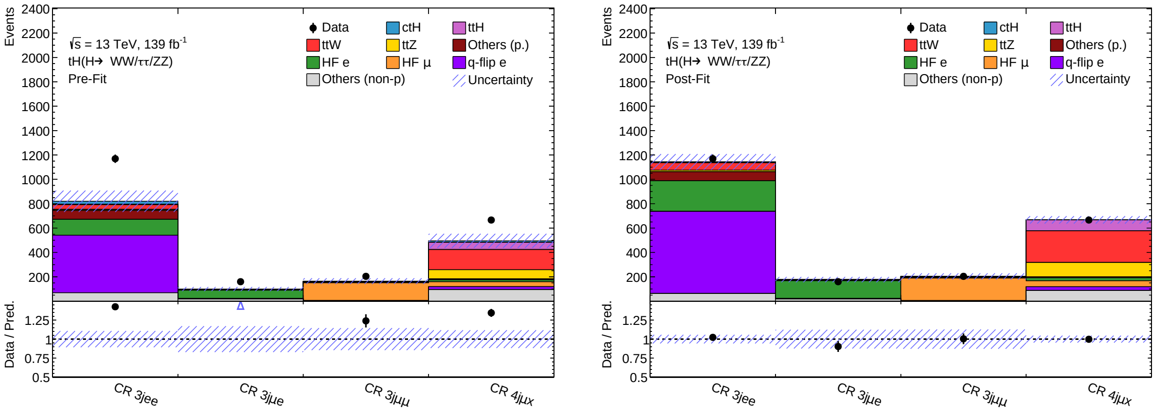
branching ratios of both signal processes are determined and compared to the current upper limits as well as the limits obtained using only statistical uncertainties. Table 11.3 shows the pre-fit expected event yields in all regions included in the fit.

Table 11.3: Pre-fit expected event yields for all regions included in the fit.

Process	SR					CR-3jet			CR-4jet
	comb	ee	$e\mu$	μe	$\mu\mu$	ee	μe	$\mu\mu$	μx
$t\bar{t}W$	433	73	120	99	141	39	2	3	165
$t\bar{t}H$	150	26	44	33	46	4	0	0	59
$t\bar{t}Z$	228	47	69	50	62	8	1	1	75
Others (prompt)	105	30	29	29	17	73	4.6	6	7
HF-decay e	353	151	19	182	0	130	69	0	16
HF-decay mu	784	0	339	31	413	0	1	147	40
Charge-flip e	328	153	136	40	0	472	3	0	24
Others (non-p)	378	114	98	113	54	69	18	5	95
utH signal	717	127	212	151	227	87	117	125	23
ctH signal	616	106	178	139	194	86	111	121	42

11.3.1 Background-only fit results

Before studying the effects of individual systematic uncertainties, the overall fit result is taken into account. As for the stat-only fit, pre- and post-fit summary plots of all control regions are produced and depicted in Figure 11.6. The pre- and post-fit plots for individual regions can again be found in the appendix in Figure A.12 and Figure A.13. A good



(a) Pre-fit (b) Post-fit

Figure 11.6: Summary plots of all control regions included in the background-only fit with statistical and systematic uncertainties. Individual bins of all regions are combined into one bin respectively. The pre-fit plot (a) includes the uncorrected MC predictions, while in the post-fit plot (b) the non-prompt and $t\bar{t}X$ templates are scaled by their respective normalization factors.

agreement between data and the corrected MC templates is observed within the given uncertainties. These are, as expected, much larger compared to the stat-only fit due to the inclusion of systematic uncertainties.
In addition to considering the agreement of the post-fit MC with data, the results for the normalization factors is again studied. They are plotted in Figure 11.7. All factors are in agreement with their counterparts from the stat-only fit within their respective uncertainties. The uncertainties, however, are much larger compared to the stat-only fit.

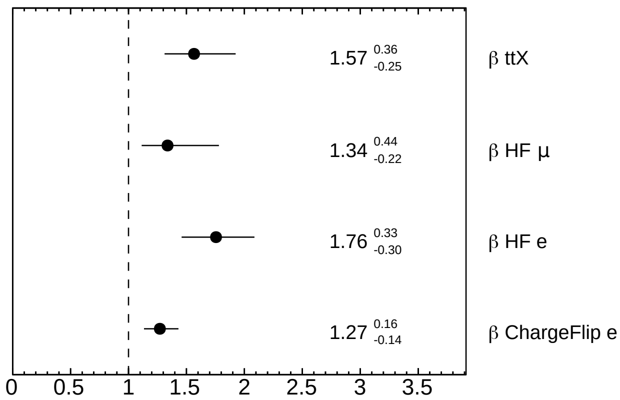
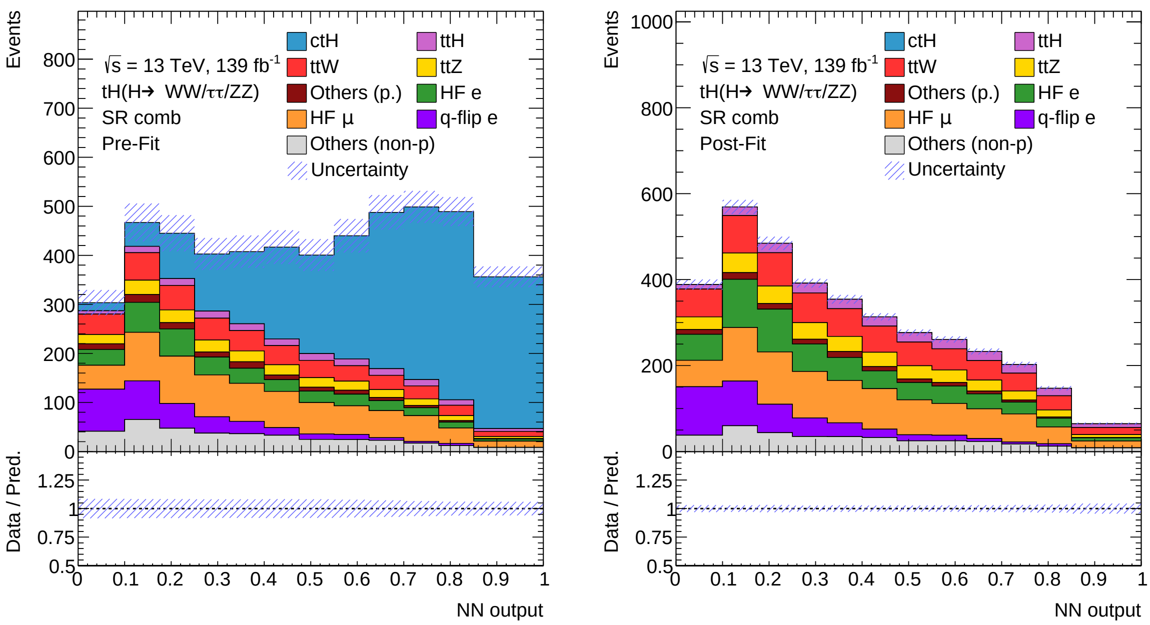


Figure 11.7: Normalization factors for the background processes that result from the background-only fit including statistical and systematic uncertainties.

Hence, the inclusion of systematic uncertainties has the primary impact of increasing the uncertainties of the pre- and post-fit MC and the normalization factors, that are determined for specific background processes. There is a notable difference between the pre-fit uncertainty and the post-fit uncertainty in all regions. The overall uncertainty decreases due to the fit. This decrease can be attributed to constraints on various NPs. To better understand this effect, detailed studies on the influence of individual NPs on the fit result are performed.

11.3.2 Influence of individual nuisance parameters

The influence of individual NPs on the fit can best be measured in terms of the NP's effect on the signal strength. Thus, fits including signal and background processes are conducted for both FCNC couplings. They are performed in all regions, with an Asimov dataset in the signal regions. The Asimov dataset is created as described at the beginning of this chapter. The signal strength is only well-defined for a specific signal process. Hence, the effects are discussed exemplarily on the basis of the ctH signal. Figure 11.8 shows the pre- and post-fit distributions of the NN output for the ctH signal in the combined SR. Similar plots for the utH signal component can be found in Figure A.14 in the appendix. The pre- and post-fit plots of the $p_T(\ell_1)$ distributions in the SR-3jet μe and $\mu\mu$ are depicted in Figure A.16 for the utH signal and in Figure A.15 for the ctH signal.
The signal processes in the pre-fit plots are scaled to the cross sections given in Table 11.1. This normalization is larger than that for the current upper limits by factors of 7.1 for the utH and 8.0 for the ctH signal. Hence, the signal templates in the pre-fit plots appear uncommonly large. There is no signal template in the post-fit plot, because the fit determined the optimal signal strength to be zero[1]. Similarly to the control regions, the overall uncertainties in the signal region are reduced by the fit due to constraints on the NPs. Furthermore, the non-prompt and $t\bar{t}X$ processes are scaled by their corresponding normalization factors.



(a) Pre-fit (b) Post-fit

Figure 11.8: Pre- (a) and post-fit (b) plots of the NN output distribution in the combined SR for the ctH signal. The signal in the pre-fit plot is scaled to the cross sections given in Table 11.1.

[1]To avoid unphysical results, the signal strength is constrained to have a minimum value of zero.

Based on these fit results, the impact of individual NPs on the fit is studied. A plot showing the pulls and constraints on all NPs is shown in the appendix in Figure A.17. The magnitude of these pulls and constraints give an estimate on which systematic uncertainties have a large impact on the fit result.

A more precise way to measure the impact of a specific NP on the fit is to determine, how much the uncertainty of the signal strength is influenced by the up- and the down-variation of the NP. To estimate this, the NP is fixed to its post-fit value $\hat{\theta}$, varied by its post-fit uncertainty $\Delta\hat{\theta}$ in the up- and down-direction, and the entire fit is repeated. The differences between the resulting signal strength from these fits and $\hat{\mu}$ from the nominal fit is then interpreted as a measure of the NP's impact on the fit result.

All NPs can then be ranked according to the magnitude of their influence on the signal strength. Figure 11.9 shows the impact of the ten highest ranked NPs. In addition to the post-fit impact of the NP, which is determined as described above, the pre-fit impact is shown in this plot. The latter is determined by fixing the NP to its pre-fit value θ, varied by its pre-fit uncertainty $\Delta\theta$ in both directions, and repeating the fit. Furthermore, the pulls and constraints of the individual NPs are depicted in the plot. Notably, the post-fit value of the γ factors is generally centered around a value of 1, with errorbars that are too small to be visible on the plot. This is in accordance with the definition of γ factors. They only scale the MC predictions in each bin slightly and are hence expected to be close to unity. Additionally, no pre-fit impact on the signal strength is shown for γ factors, as they do not have a well-defined pre-fit uncertainty. Lastly, pre- and post-fit variations for all NPs either appear to increase the signal strength or to have no effect at all. This is because the signal strength is constrained to be positive and the maximum likelihood estimator of μ is zero. Hence, only positive changes of μ are possible.

Figure 11.9 shows that the primary NP associated to the b-tagging of b-jets has the strongest impact on the signal strength. As signal and control regions are defined based on cuts on the number of b-tags, it is understandable that this NP has a strong effect. The NP is found to be specifically important for $t\bar{t}X$ events. These only fall into the SR, if one of the two associated b-jets in the event is (erroneously) not tagged.

The second most significant NP is the γ factor for the highest o_{NN}-bin[2] in the SR. The highest bin corresponds to that with the highest S/B ratio. Since the γ factors of several other high o_{NN}-bins in the SR also have a strong effect on the fit, it might be useful to consider a lower number of bins in the SR in works expanding on this thesis.

In addition to γ factors, NPs associated to systematic uncertainties on the JES, in particular the effect of jet flavor on the JES, are ranked highly in the plot. This is likely caused by the same effects that are relevant for the uncertainty on the tagging of b-jets. Finally, one NP corresponding to the efficiency of muon triggers also has a strong influence on the signal strength. This is a reasonable behavior, as lepton triggers play an important role in recording events with multiple leptons.

[2]There are 12 bins in the SR with the lowest o_{NN}-bin being numbered zero.

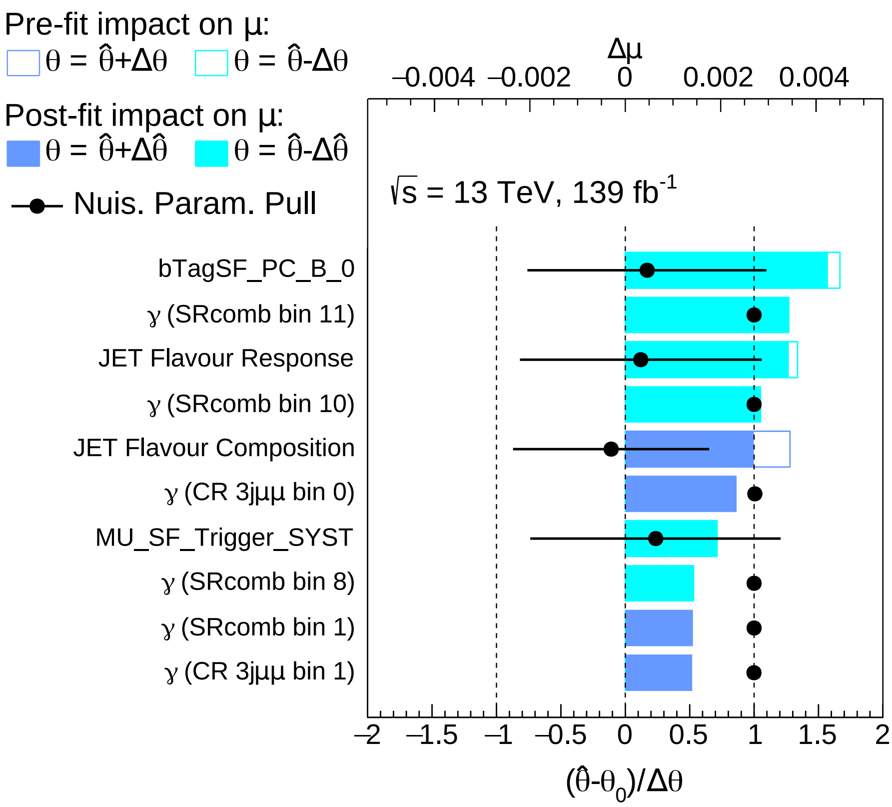


Figure 11.9: NPs ranked by their influence on the signal strength, when varied up or down by one post-fit standard deviation.

11.3.3 Expected upper limits in the combined signal region

From the fits including systematic uncertainties, new and more realistic expected 95 % CL_S upper limits on the branching ratios are calculated for both signal samples. Based on the results from the previous section, they are only calculated using the combined SR. The specific values of the limits are shown in Table 11.4. The expected upper limits on the branching ratios of all signal processes are plotted in Figure 11.10, together with the expected upper limits obtained with the stat-only fit as well as the current upper limits.

The current upper limits were determined using a combination of results from analyses on various Higgs-decay modes and final states. This makes detailed comparisons to the expected upper limits obtained in this theses less conclusive. A more meaningful comparison is one to the analysis described in [14], which analyzed the same FCNC decay process in the 2ℓSS as well as the 3ℓ[3] final state, using data corresponding to an integrated luminosity of $L_{\text{int}} = 39\,\text{fb}^{-1}$. The results obtained in that analysis are depicted as green dashed-dotted lines in Figure 11.10. They amount to upper limits on the branching ratios of 0.19 % and 0.16 % for the $t \to Hu$ and the $t \to Hc$ channel, respectively. For comparison, Table 11.4 also shows the current upper limits as well as the most recent limits achieved in the 2ℓSS/ 3ℓ final states.

[3]The 3ℓ final state is defined by containing three leptons with a sum of charges of ± 1.

Table 11.4: 95 % CL_S expected upper limits on $\mathcal{B}(t \to qH)$ obtained from a fit including statistical and systematic uncertainties with one combined SR.

Couplings	$\mathcal{B}(t \to qH)$ / 10^{-4}					Current limits	Latest 2ℓSS/ 3ℓ limits
	-2σ	-1σ	median	$+1\sigma$	$+2\sigma$		
utH	2.0	2.6	3.7	5.1	6.9	12	19
ctH	2.1	2.8	3.9	5.5	7.4	11	16

Comparing the new expected upper limits on the branching ratios to limits from the stat-only fits, a notable deterioration can be observed. For both signal components the expected upper limit increases by 15 %. Nevertheless, the expected upper limits give the prospect of a noticeable improvement over the current upper limits. This is particularly remarkable, because the current upper limits were determined from the combination of multiple decay channels of the Higgs boson, while this result only takes into account one final state of one decay mode. The result should be interpreted with some reservation, however, as systematic uncertainties on the MC modeling and PDFs are not taken into account in this thesis.

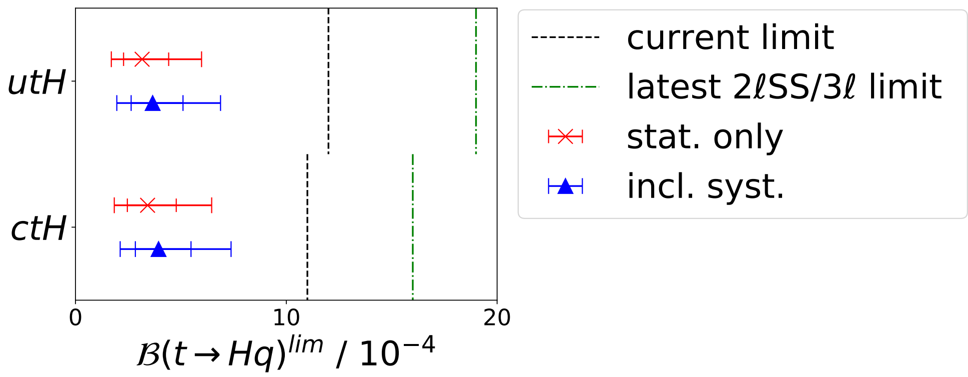


Figure 11.10: Expected 95 % CL_S upper limits on the branching ratio $\mathcal{B}(t \to Hq)$ for the utH and the ctH coupling from stat-only fits (red crosses) and the fits including systematic uncertainties (blue triangles). All fits are performed using the combined SR. The two errorbars for each limit depict the 1σ and the 2σ regions. The current upper limits [19] on the branching ratios are depicted as black, dashed lines. The latest upper limits obtained with multilepton final states are shown as green, dashed-dotted lines [14].

Compared to the previous analysis on multilepton final states, the expected improvement achieved by this analysis is even more prominent. This is especially true, since the results from the previous analysis include the 3ℓ final state, which is not taken into account in this thesis. An expected reduction of the multilepton upper limits by (81 ± 8) % for the utH signal and by (79 ± 8) % for the ctH signal is obtained. This improvement cannot be traced back to a single source. Instead, the combination of tools such as the PLV and the ECIDS and optimized NNs, together with profoundly higher statistics create a condition, in which a formidable sensitivity to the FCNC signal is achieved.

12 Summary and outlook

The aim of this thesis was the construction of an analysis strategy for the search for FCNC couplings between the top quark and the Higgs boson in 2ℓSS final states. Two couplings are considered: the utH and the ctH coupling. The most important findings of this thesis are summarized in the following.

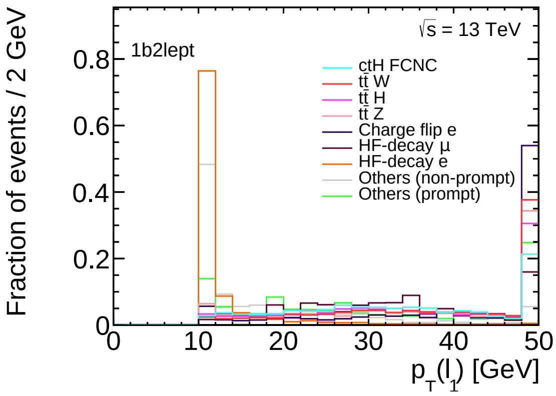
For the correct estimation of non-prompt background components, the SM process templates are redefined, to decouple prompt background processes from non-prompt background processes. Most notably, templates for the HF-decay e and μ background as well as for the charge-flip e background are introduced. Based on these definitions, specific signal and control regions are defined. The signal regions are selected by requiring a high signal efficiency. Each control region was specifically designed to be strongly dominated by one of the non-prompt background processes. These regions are used at a later stage of the analysis to correct the rates of the respective background process, which are expected to be not well predicted in MC simulations. One additional control region is defined to correct the rate of the $t\bar{t}X$ processes.
In total, three signal regions are defined: the main SR and the two smaller SR μe and $\mu\mu$. In the main SR, an additional splitting by the flavor of the p_T-leading and sub-leading lepton is conducted. This step is motivated by the observation that different non-prompt background processes dominate in the various subregions. The differences between the combined and the flavor-split SR are analyzed in the final stage of the analysis.
Several approaches to the reconstruction of signal events are taken in this thesis. Based on the observation that the signal process is dominated by $H \to WW^*$ decays, the hadronic W boson in the signal process is reconstructed. Additionally, the neutrino weighting method is implemented. It provides a way to use reconstruction techniques despite the presence of two neutrinos in the 2ℓSS final state. Lastly, the Fox-Wolfram moments of the two final state leptons are implemented into the analysis.
The reconstructed kinematic variables are used to train artificial NNs in the main SR with the aim of reliably discriminating between signal and background processes. Separate NNs are trained for the utH and the ctH signal component. Additionally, one NN is trained for each signal process in all four subregions generated by the flavor-splitting. In total, ten NNs are trained in the combined and the flavor-split SR. To better estimate the rate correction of the $t\bar{t}X$ background, an additional NN is trained in the $t\bar{t}X$ control region.
Finally, several maximum likelihood fits are performed in signal and control regions to provide an estimate on the expected sensitivity of this analysis. The first fits include statistical uncertainties only and are conducted to compare the two setups of the SR. It is found that both region setups produce similar results. Hence, the more manageable approach of a combined SR is chosen for the final fit. This final fit takes into account a number of systematic uncertainties. These include all expected sources of systematic uncertainties with the exception of PDF and MC modeling uncertainties.

The fit is performed by first determining the rate corrections of all mis-modeled processes via a background-only fit to data in the control regions. Subsequently, an Asimov dataset is created based on these corrections. Then a global fit to the Asimov dataset in the signal regions and to real data in the control regions is performed. The Asimov dataset is created under the assumption that there is no contribution from the signal processes. Because of this, expected upper limits on the branching ratios of both signal processes can be determined. These amount to $\mathcal{B}(t \to uH) = 3.7 \times 10^{-4}$ and $\mathcal{B}(t \to cH) = 3.9 \times 10^{-4}$. They correspond to a tremendous improvement over the current limits on the strength of FCNC couplings between the top quark and the Higgs boson. However, they are not fully indicative of the true sensitivity of this analysis, as two sources of systematic uncertainty are not included in the fit. Nonetheless, these results show that a capable analysis strategy has been developed in this thesis.

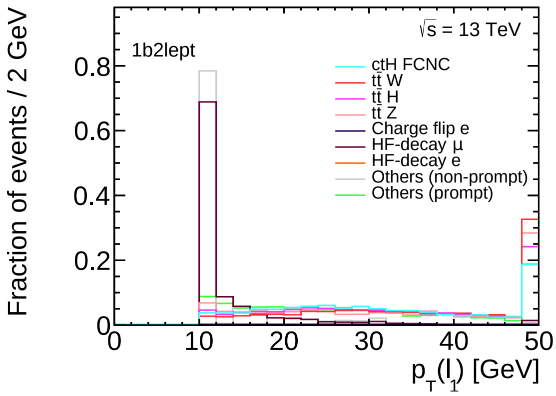
There are several possible expansions on this thesis. The most obvious next step is the inclusion of modeling and PDF uncertainties in the final fit. This allows for a much better estimation of the expected upper limits on the signal branching ratios. Furthermore, additional Higgs-boson decay channels and final states can be incorporated into the analysis. The $q \to tH$ production channel allows for even higher sensitivity to the overall FCNC ctH and utH couplings. Additionally, a sensitivity to the chirality of the FCNC couplings can be obtained in this channel. The inclusion of a final state with three leptons in the decay and production channel is likely to further improve the overall performance of the analysis.

A Figures

Background estimation and event selection



(a) μe subregion



(b) $\mu\mu$ subregion

Figure A.1: High-resolution shape plots of the p_T of the sub-leading lepton in the μe (a) and the $\mu\mu$ subregions (b) of the candidate control region defined by $n_{\text{jets}} \leq 3$ and $n_{b\text{-tags}} \geq 1$. The plots only include the ctH signal component. The final bin constitutes an overflow-bin, meaning it includes events with $p_T(\ell_1)$ between 48 GeV and 50 GeV, as well as all events with $p_T(\ell_1) > 50$ GeV.

Event reconstruction

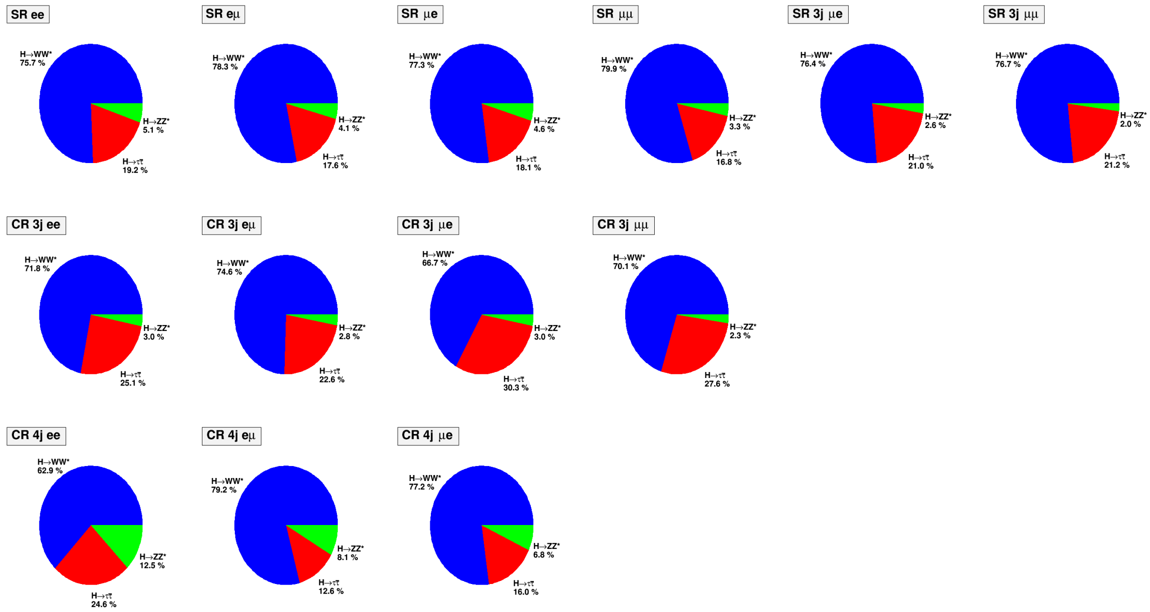


Figure A.2: Contribution of various Higgs-boson decay modes to the *utH* signal process in all defined regions, except for the combined SR.

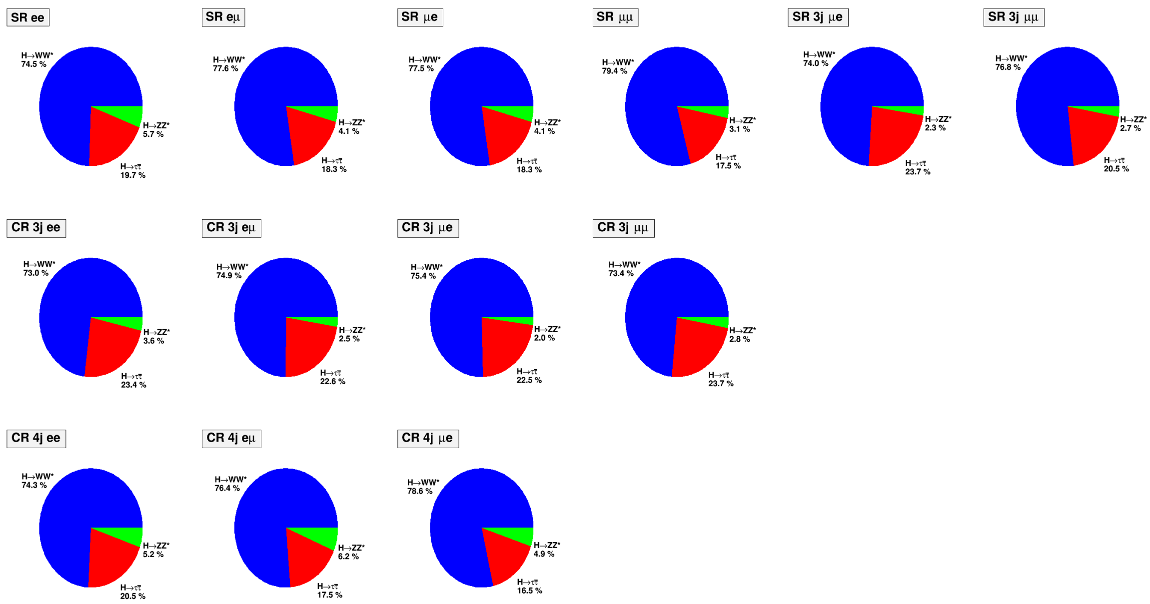
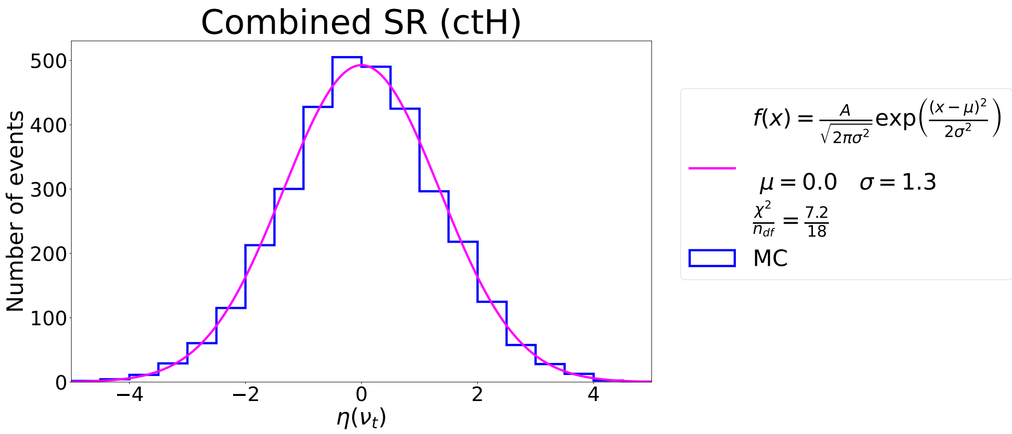
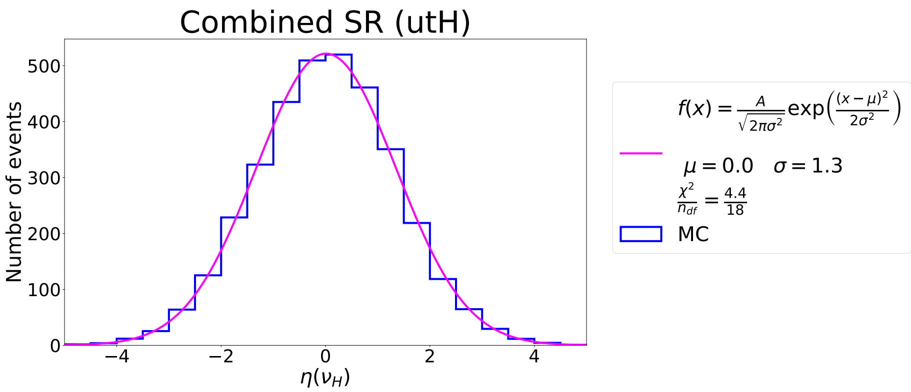


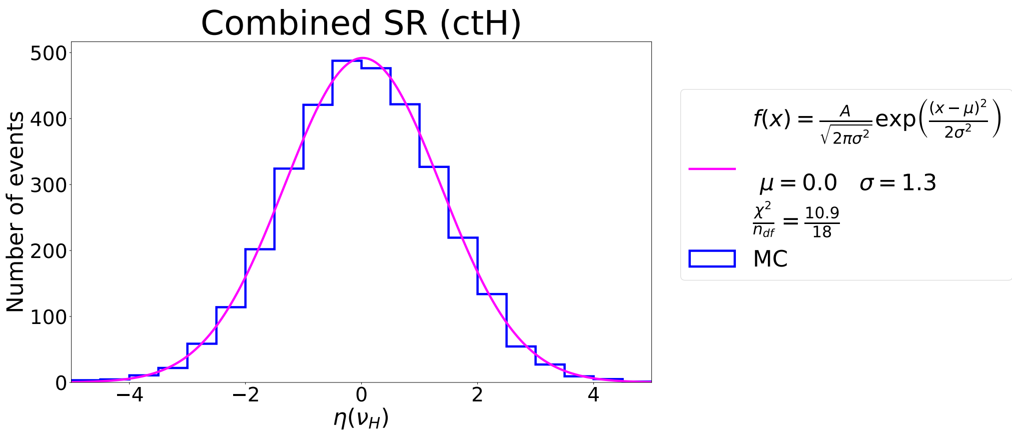
Figure A.3: Contribution of various Higgs-boson decay modes to the *ctH* signal process in all defined regions, except for the combined SR.



(a) ν from top-quark decay (ctH signal)

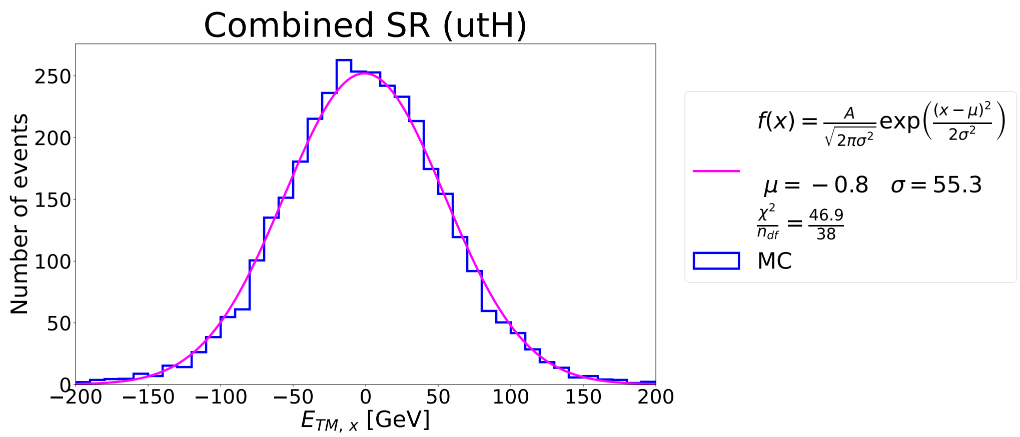


(b) ν from Higgs-boson decay (utH signal)

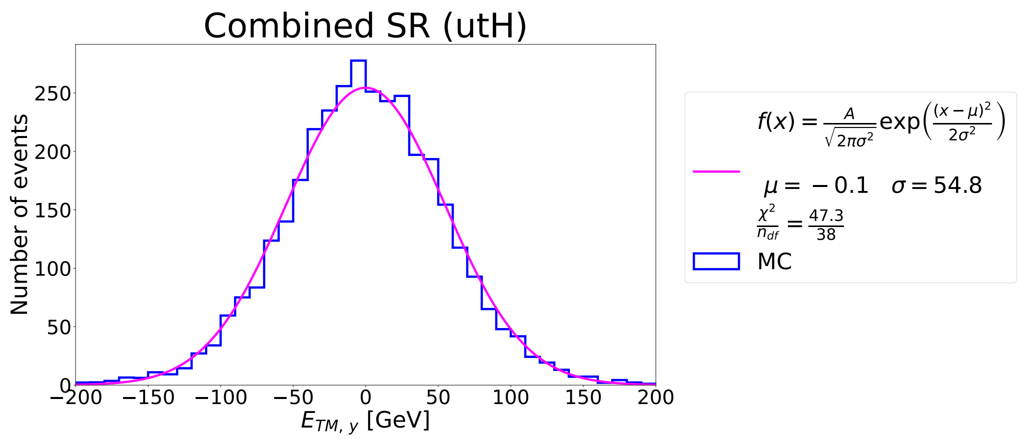


(c) ν from Higgs-boson decay (ctH signal)

Figure A.4: Histograms of the truth value of η_ν for the neutrino from the top-quark decay in the ctH signal (a) as well as for the neutrino from the Higgs-boson in the utH (b) and ctH (c) signal. Results of a χ^2-fit of a non-normalized Gaussian distribution are depicted as well.

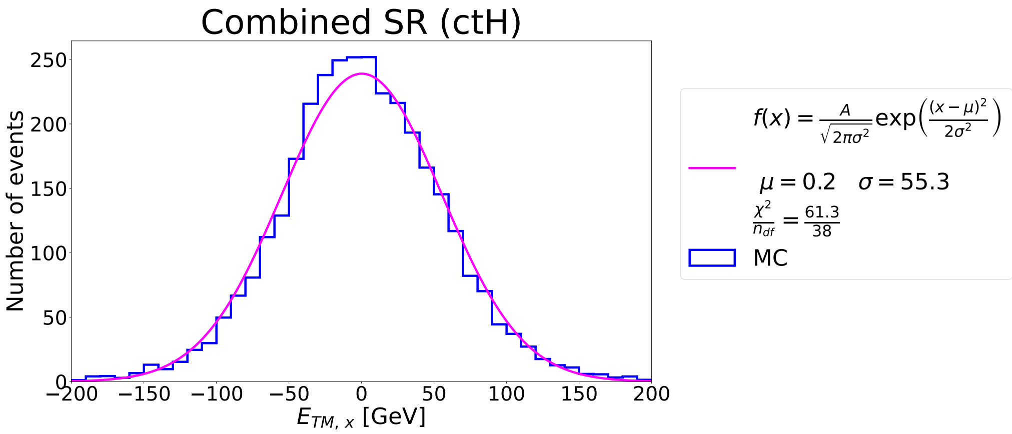


(a) $\not{E}_x$

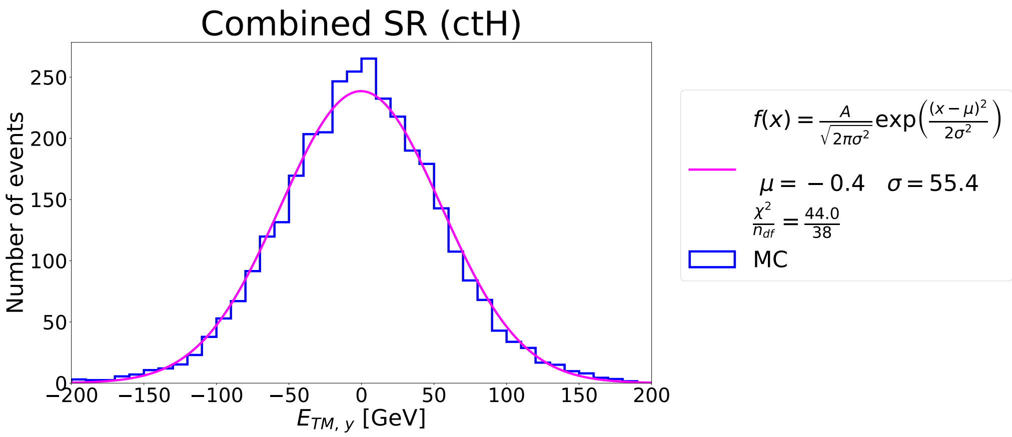


(b) $\not{E}_y$

Figure A.5: Gaussian fits through the $\not{E}_x$ (a) and $\not{E}_y$ (b) distributions for the utH signal. The fit function is identical to the one in Figure A.4.



(a) $\not{E}_x$



(b) $\not{E}_y$

Figure A.6: Gaussian fits through the $\not{E}_x$ (a) and $\not{E}_y$ (b) distributions for the *ctH* signal. The fit function is identical to the one in Figure A.4.

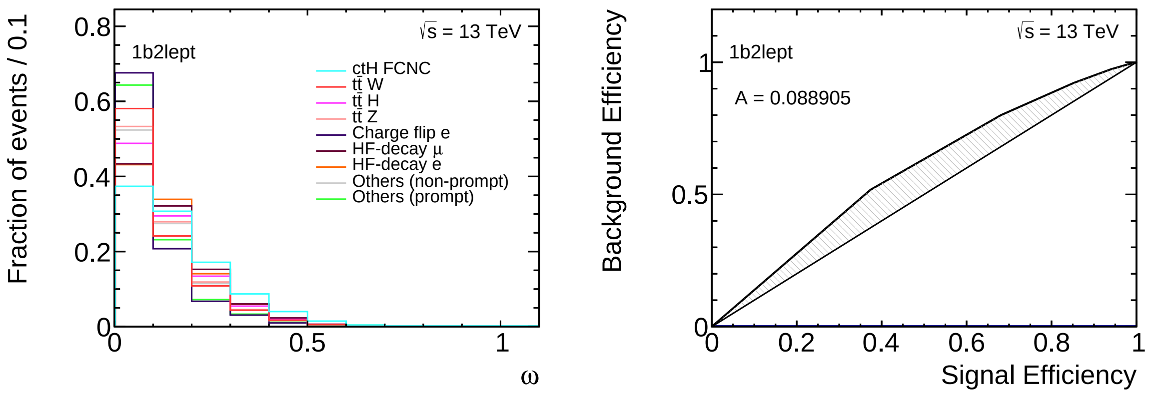
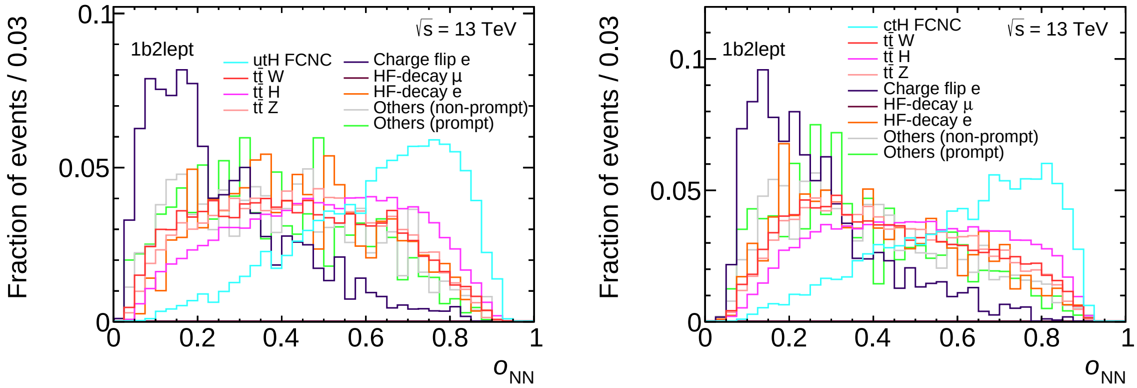


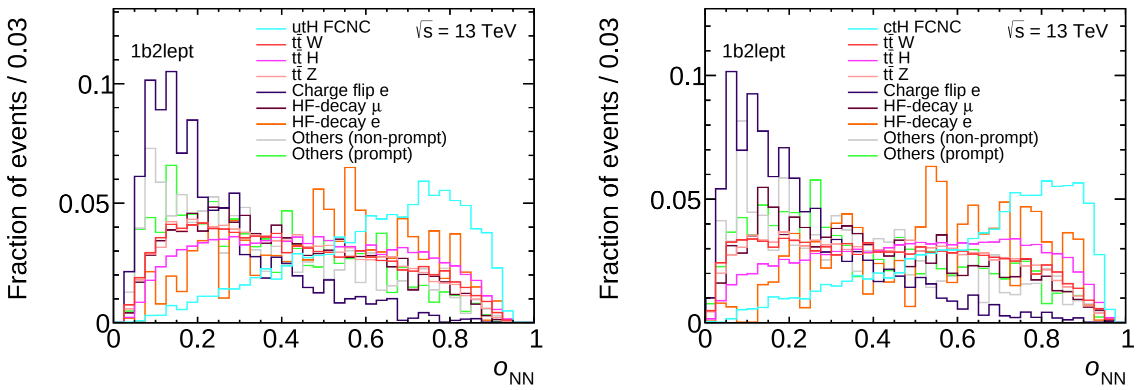
Figure A.7: Shape plot and ROC curves of the neutrino weight in the combined SR. The area A between the ROC curve and the diagonal is given in the respective plot. Only the *ctH* signal is shown.

Separation of signal and background



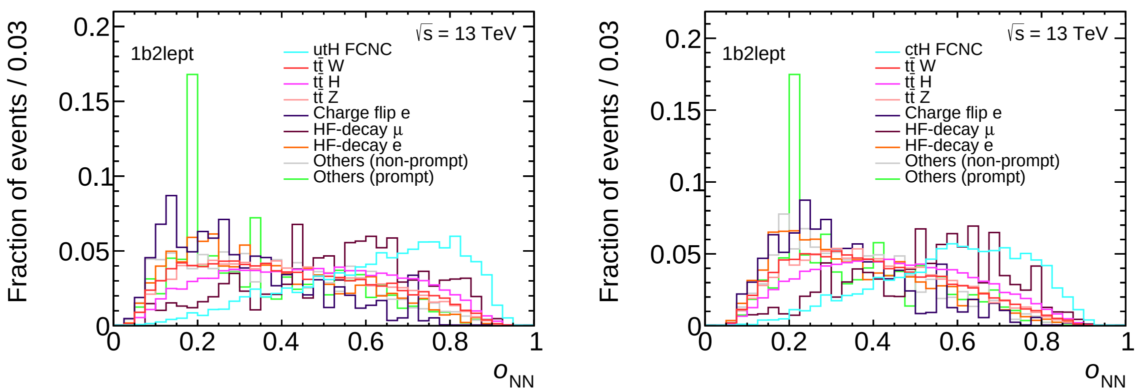
utH signal   *ctH* signal

(a) SR *ee*



utH signal   *ctH* signal

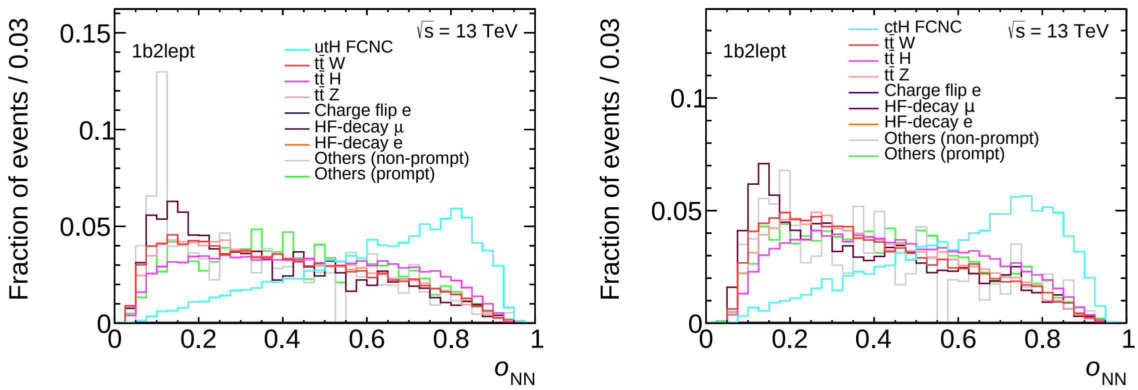
(b) SR $e\mu$



utH signal   *ctH* signal

(c) SR μe

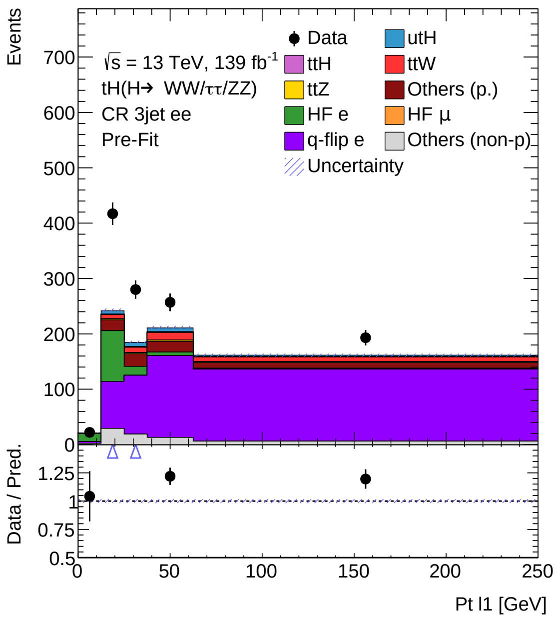
Figure A.8: Shape plots of the NN output variable o_{NN} for the NNs trained in the flavor-split SR ee, SR $e\mu$ and SR μe. Trainings were performed for the utH and the ctH signal.



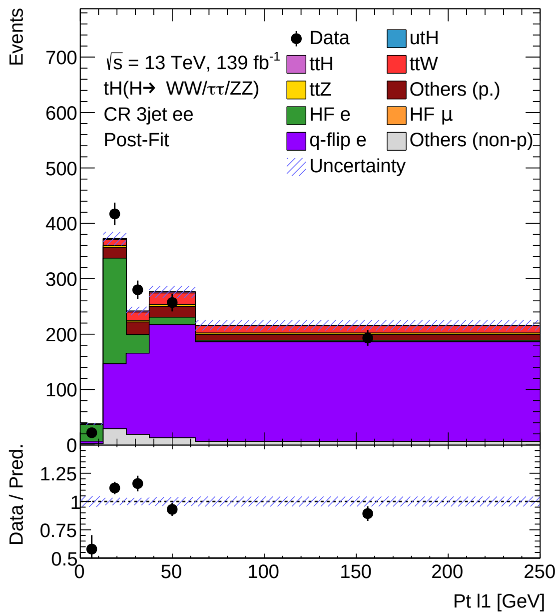
utH signal *ctH* signal

(a) SR $\mu\mu$

Figure A.9: Shape plots of the NN output variable o_{NN} for the NNs trained in the flavor-split SR $\mu\mu$. Trainings were performed for the utH and the ctH signal.

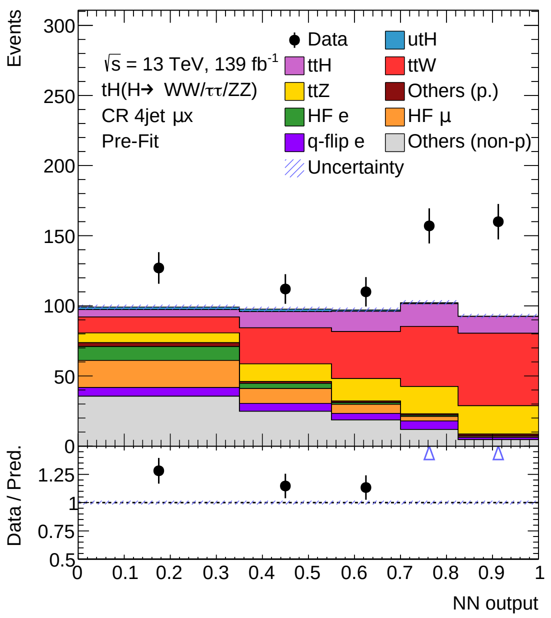


Pre-fit

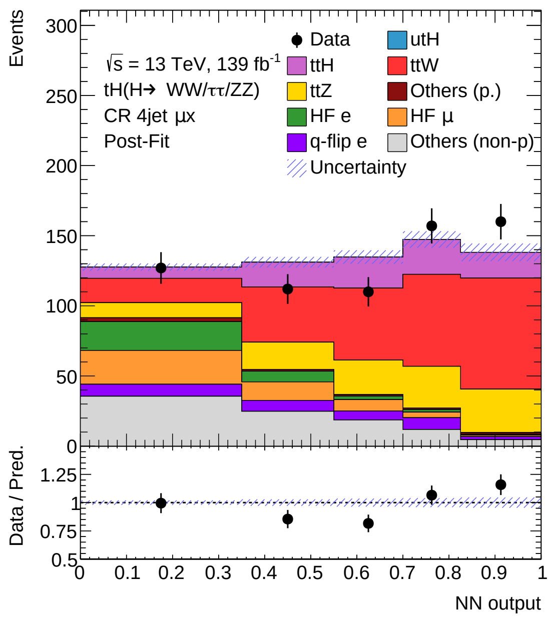


Post-fit

(a) CR-3jet *ee*



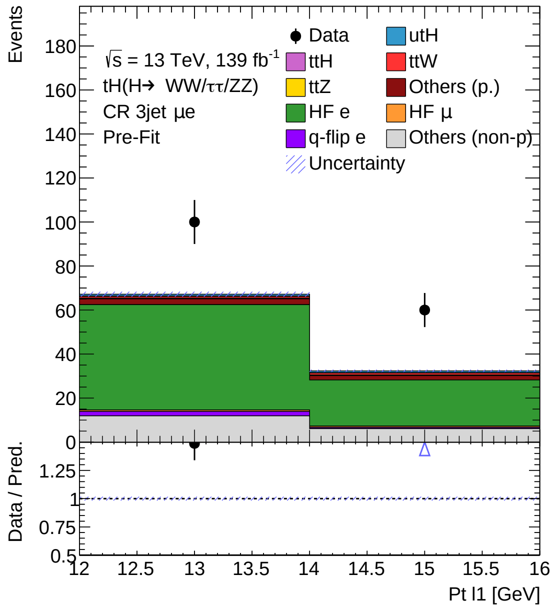
Pre-fit



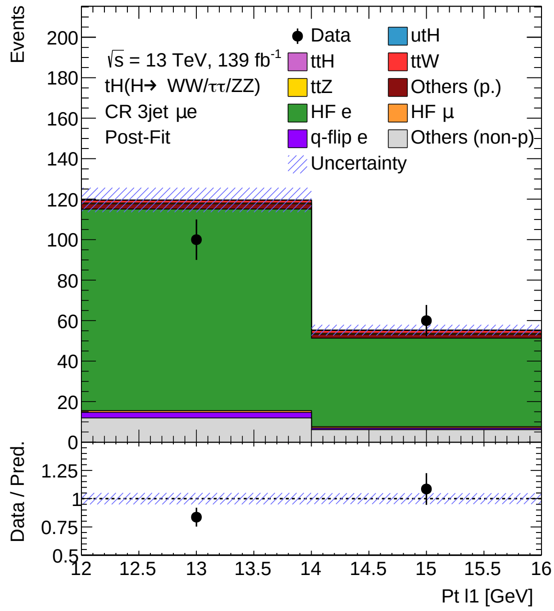
Post-fit

(b) CR-4jet μx

Figure A.10: Pre- (left) and post-fit (right) plots of the $p_T(\ell_1)$ distribution in the CR-3jet *ee* (a) and the NN output variable distribution in the CR-4jet μx (b) for the stat-only background-only fit.

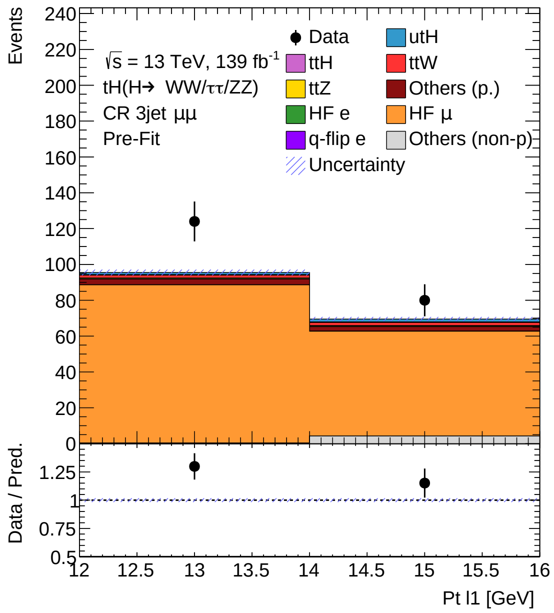


Pre-fit

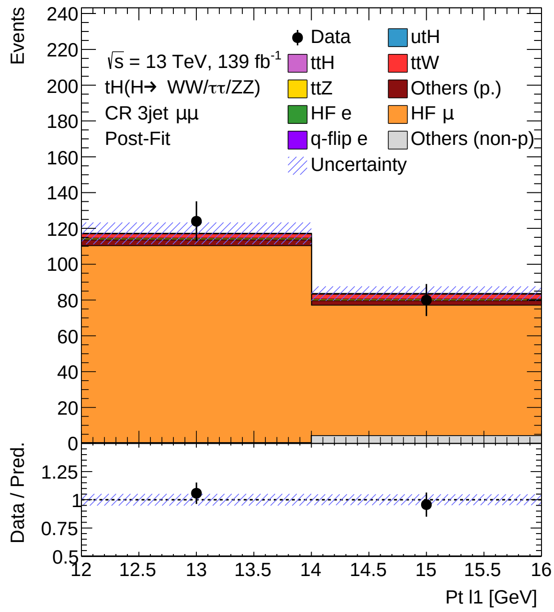


Post-fit

(a) CR-3jet μe



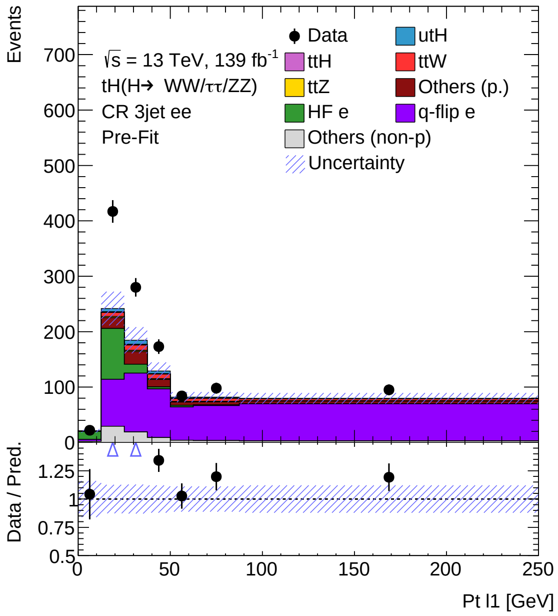
Pre-fit



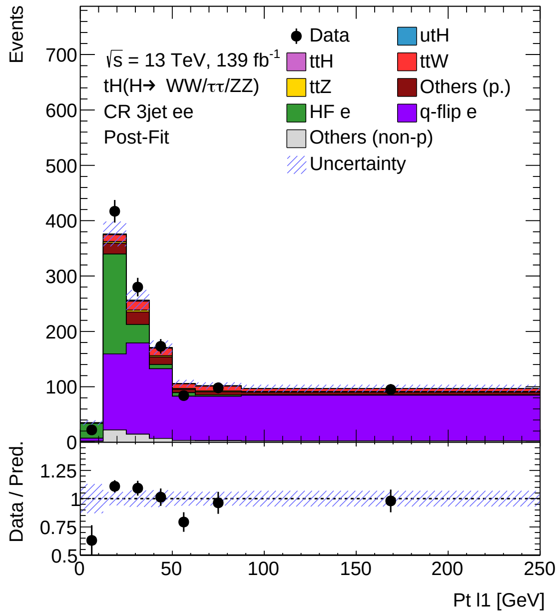
Post-fit

(b) CR-3jet $\mu\mu$

Figure A.11: Pre- (left) and post-fit (right) plots of the $p_T(\ell_1)$ distribution in the CR-3jet μe (a) and in the CR-3jet $\mu\mu$ (b) for the stat-only background-only fit.

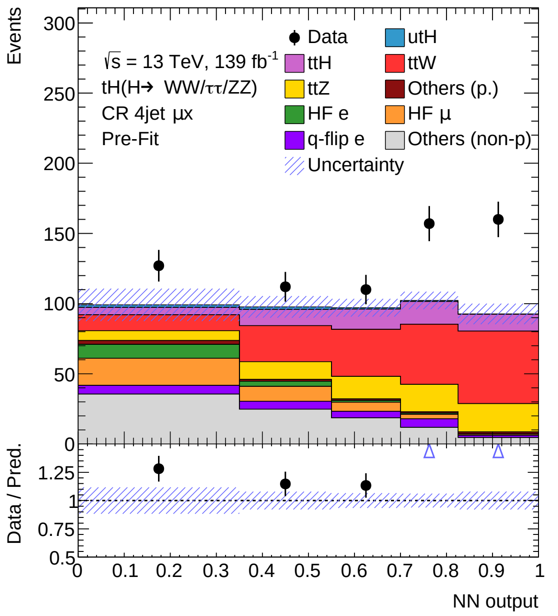


Pre-fit

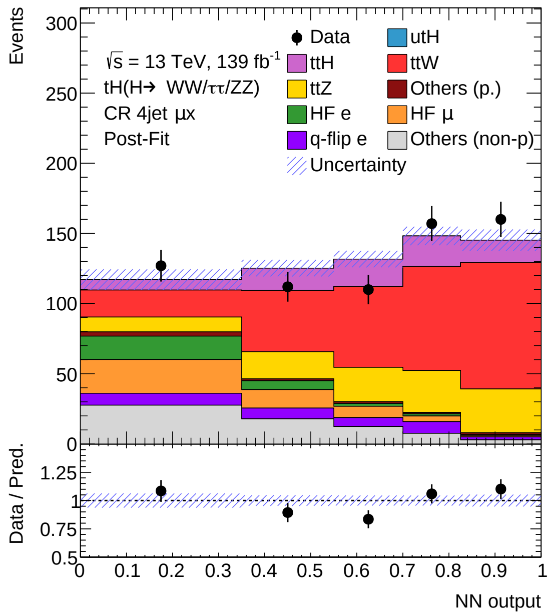


Post-fit

(a) CR-3jet *ee*



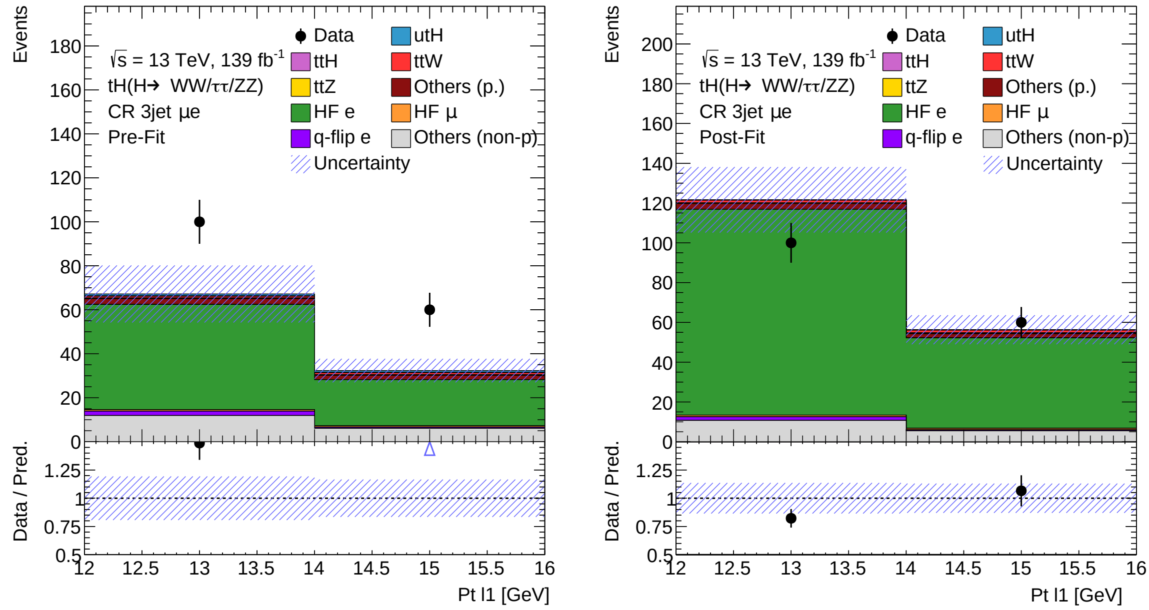
Pre-fit



Post-fit

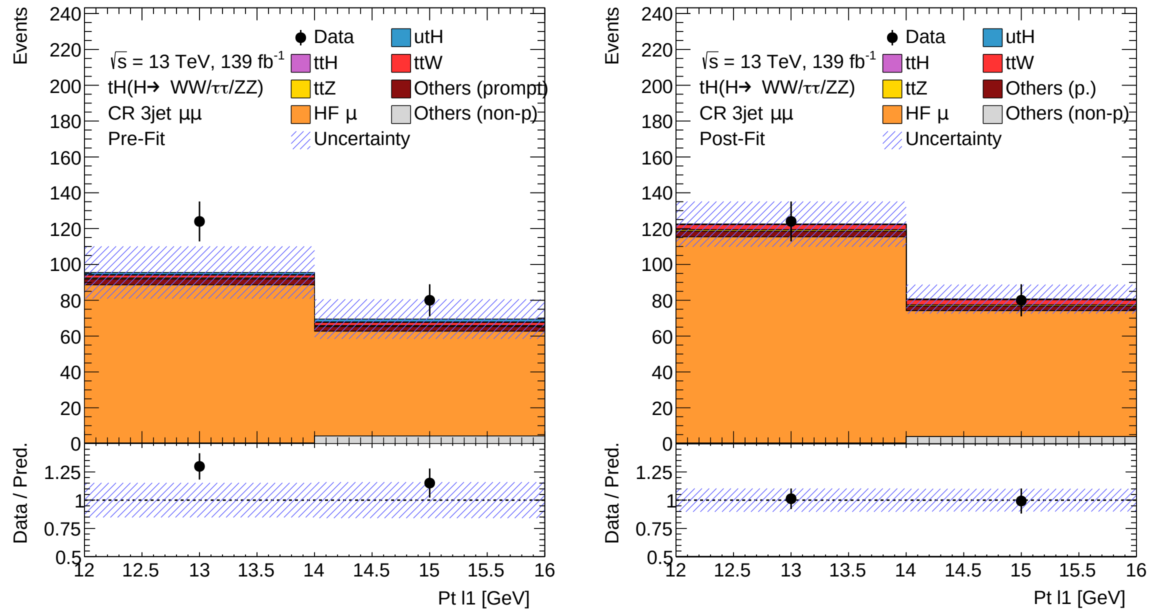
(b) CR-4jet μx

Figure A.12: Pre- (left) and post-fit (right) plots of the $p_T(\ell_1)$ distribution in the CR-3jet *ee* (a) and the NN output variable distribution in the CR-4jet μx (b) for the background-only fit including statistical and systematic uncertainties.



Pre-fit Post-fit

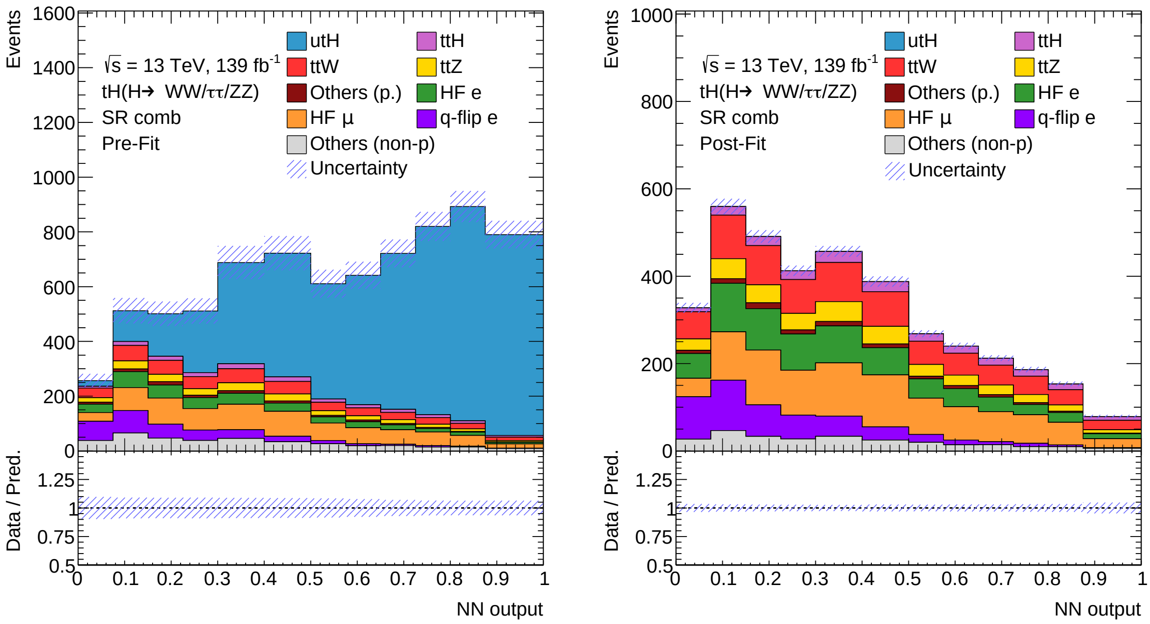
(a) CR-3jet μe



Pre-fit Post-fit

(b) CR-3jet $\mu\mu$

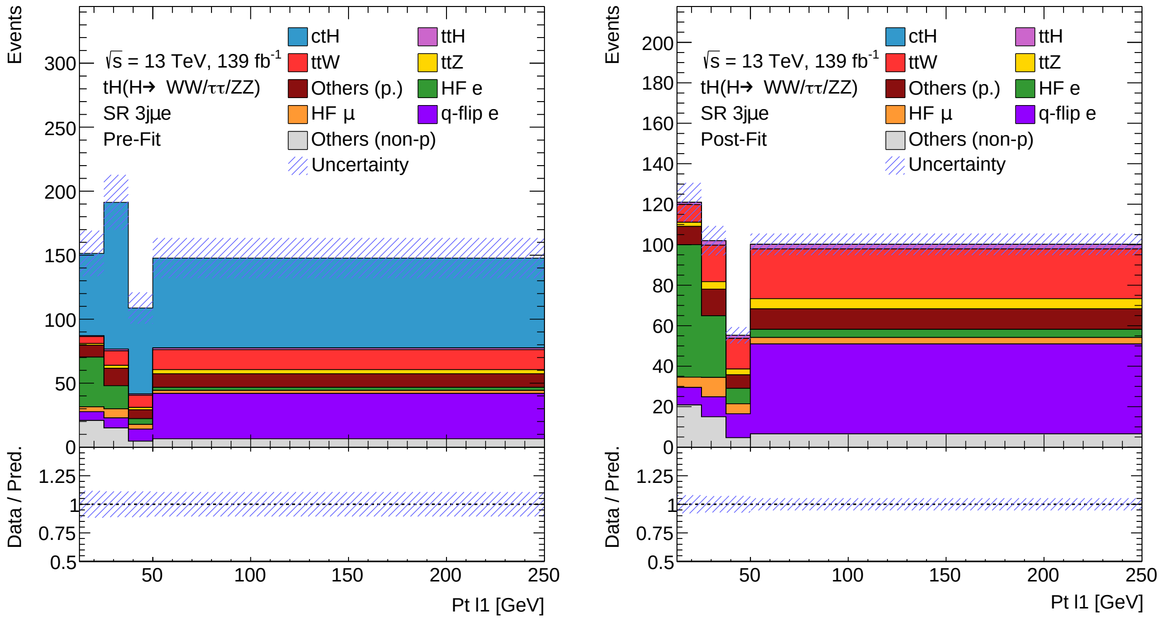
Figure A.13: Pre- (left) and post-fit (right) plots of the $p_T(\ell_1)$ distribution in the CR-3jet μe (a) and in the CR-3jet $\mu\mu$ (b) for the background-only fit including statistical and systematic uncertainties.



(a) Pre-fit

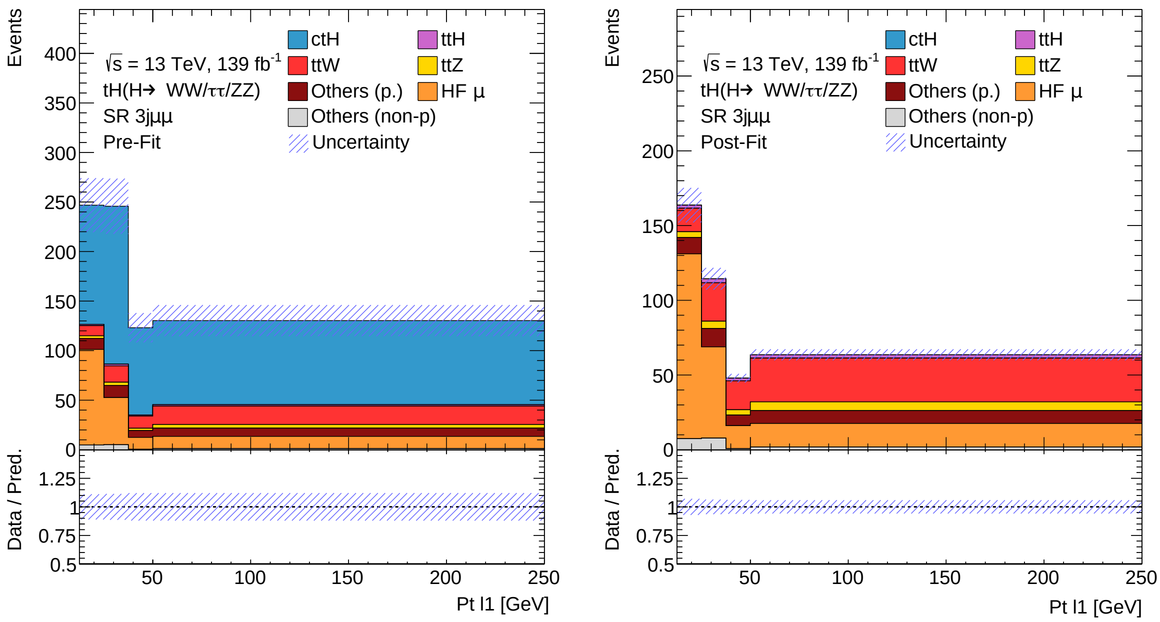
(b) Post-fit

Figure A.14: Pre- (a) and post-fit (b) plots of the NN output distribution in the combined SR for the utH signal. The signal in the pre-fit plot is scaled to the cross sections given in Table 11.1.



Pre-fit Post-fit

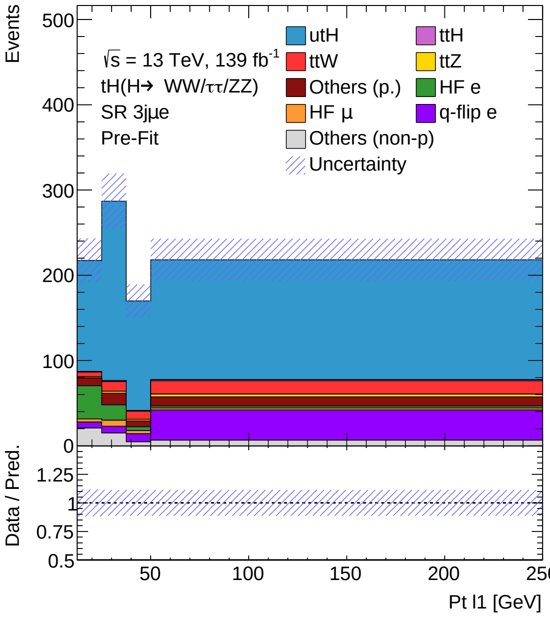
(a) SR-3jet μe



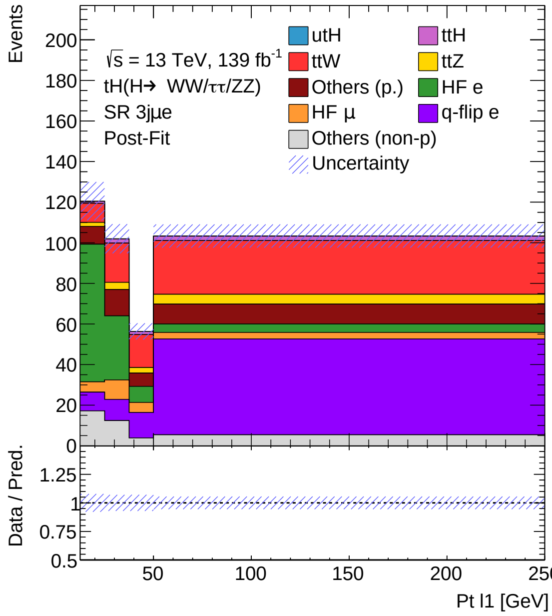
Pre-fit Post-fit

(b) SR-3jet $\mu\mu$

Figure A.15: The distributions of $p_T(\ell_1)$ in the SR-3jet μe (a) and the SR-3jet $\mu\mu$ (b). The pre- (left) and post-fit (right) plots are shown for the ctH signal. The signal in the pre-fit plots is scaled to the cross sections given in Table 11.1.

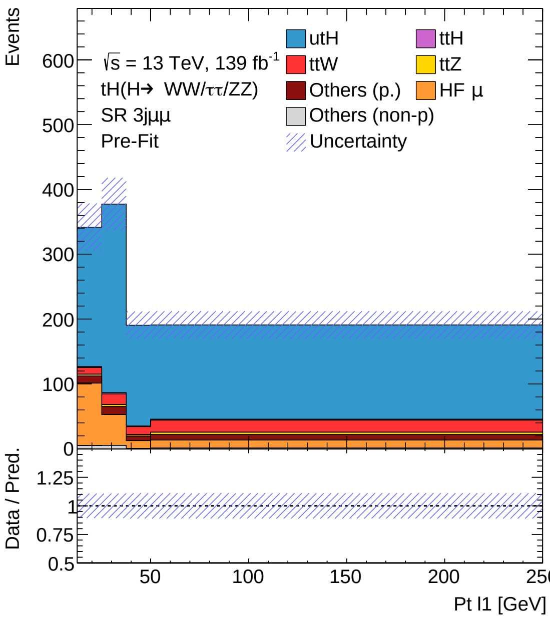


Pre-fit

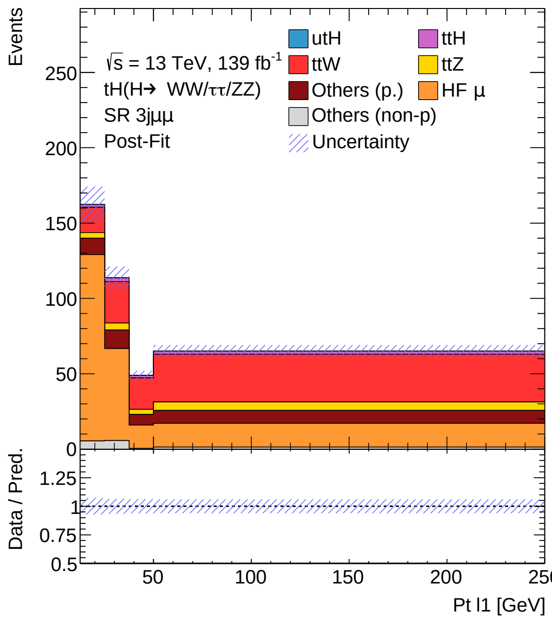


Post-fit

(a) SR-3jet μe



Pre-fit



Post-fit

(b) SR-3jet $\mu\mu$

Figure A.16: The distributions of $p_T(\ell_1)$ in the SR-3jet μe (a) and the SR-3jet $\mu\mu$ (b). The pre- (left) and post-fit (right) plots are shown for the utH signal. The signal in the pre-fit plots is scaled to the cross sections given in Table 11.1.

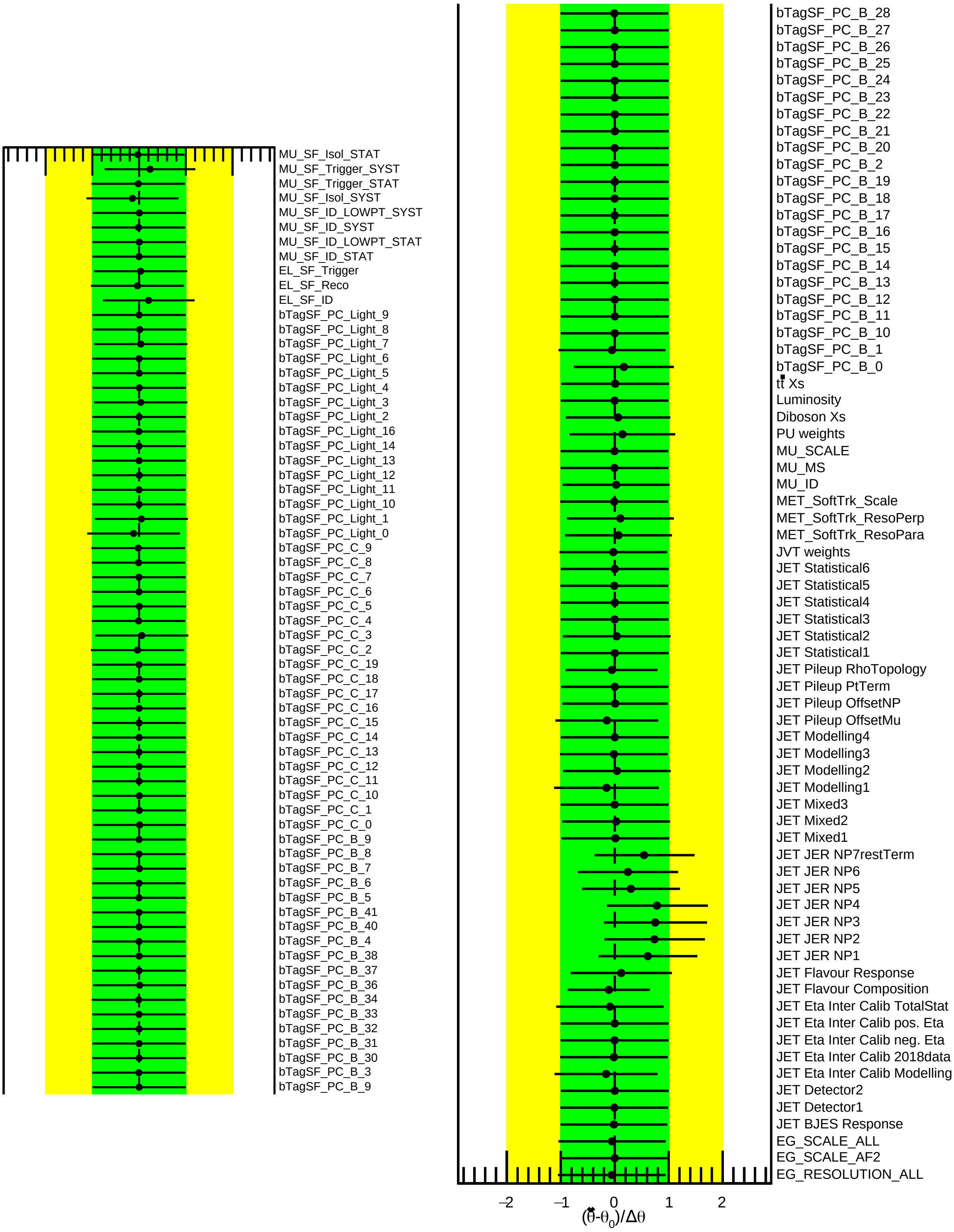


Figure A.17: Pulls and constraints of all NPs associated to systematic uncertainties. The green and yellow bands visualize the pre-fit 1σ and 2σ uncertainties. The dashed black line signifies the NPs' pre-fit value. The post-fit values and uncertainties are given by the black data points.

B Tables

Background estimation and event selection

Table B.1: The expected number of events in the combined SR before and after the ECIDS cut. The relative change in events due to the cut is shown as well.

Process	pre ECIDS cut	post ECIDS cut	relative change [%]
$t\bar{t}W$	452	433	-4.1
$t\bar{t}H$	160	150	-6.5
$t\bar{t}Z$	243	228	-6.4
Others (prompt)	164	105	-36.1
HF-decay e	1383	353	-74.5
HF-decay mu	2576	784	-69.6
Charge-flip e	1694	328	-80.6
Others (non-p)	681	378	-44.4
utH signal	759	717	-5.5
ctH signal	654	616	-5.8
Total background	7353	2759	-62.5

Separation of signal and background

Table B.2: Variables used in the NNs, which are trained in the combined SR, alongside their added significance. The NNs trained for the utH- and the ctH-signal are included. Non-b-tagged jets are denoted as l-jet and ordered by their p_T. l-jet$_0$ is the leading light jet, l-jet$_1$ the sub-leading etc. The X signifies that a variable was pruned.

Variable	Added signi. utH	Added signi. ctH	Description
$p_T(\ell_1)$	41.84	40.77	p_T of ℓ_1
$p_T(W_{\text{had}})$	X	9.35	p_T of W_{had}
$p_T(l\text{-jet}_0)$	6.51	3.19	p_T of l-jet$_0$
$p_T(l\text{-jet}_1)$	28.62	21.49	p_T of l-jet$_1$ p_T
$p_T(l\text{-jet}_2)$	6.56	3.69	p_T of l-jet$_2$
ECIDS ℓ_0	28.28	26.83	ECIDS result of ℓ_0
ECIDS ℓ_1	16.28	16.45	ECIDS result of ℓ_1
$H_1(\ell_0,\ell_1)$	12.99	11.11	1st Fox-Wolfram moment of ℓ_0 and ℓ_1
$m(\ell_0,W_{\text{had}})$	7.30	X	Invariant mass of ℓ_0 and W_{had}
$m(\ell_0,b\text{-jet})$	16.78	15.52	Invariant mass of ℓ_0 and the b-jet
$m(\ell_0,l\text{-jet}_0)$	5.98	X	Invariant mass of ℓ_0 and l-jet$_0$
$m(\ell_1,b\text{-jet})$	6.39	4.60	Invariant mass of ℓ_1 and the b-jet
$m(\ell_1,l\text{-jet}_0)$	22.24	30.37	Invariant mass of ℓ_1 and l-jet$_0$
$m(b\text{-jet},l\text{-jet}_0)$	X	4.31	Invariant mass of the b-jet and l-jet$_0$
$m(b\text{-jet},l\text{-jet}_1)$	9.34	6.74	Invariant mass of the b-jet and l-jet$_1$
$m(b\text{-jet},l\text{-jet}_2)$	5.66	4.29	Invariant mass of the b-jet and l-jet$_2$
$m(l\text{-jet}_0,W_{\text{had}})$	5.24	X	Invariant mass of l-jet$_0$ and W_{had}
$m(l\text{-jet}_0,l\text{-jet}_1)$	4.25	X	Invariant mass of l-jet$_0$ and l-jet$_1$
$m(l\text{-jet}_1,W_{\text{had}})$	3.66	X	Invariant mass of l-jet$_1$ and W_{had}
$m(W_{\text{had}})$	3.54	4.13	Mass of W_{had}
ω	15.44	12.95	Neutrino weight
$\eta(\ell_0)$	3.69	X	η of ℓ_0
$\eta(l\text{-jet}_1)$	9.95	3.65	η of l-jet$_1$
$\Delta R(\ell_0,W_{\text{had}})$	X	5.46	Angular separation between ℓ_0 and W_{had}
$\Delta R(\ell_0,l\text{-jet}_0)$	14.80	16.23	Angular separation between ℓ_0 and l-jet$_0$
$\Delta R(\ell_0,l\text{-jet}_1)$	15.90	15.57	Angular separation between ℓ_0 and l-jet$_1$
$\Delta R(\ell_0,l\text{-jet}_2$	X	9.38	Angular separation between ℓ_0 and l-jet$_2$
$\Delta R(\ell_1,W_{\text{had}})$	8.60	14.83	Angular separation between ℓ_1 and W_{had}
$\Delta R(\ell_1,b\text{-jet})$	5.15	5.69	Angular separation between ℓ_1 and the b-jet
$\Delta R(\ell_1,l\text{-jet}_0)$	3.86	3.11	Angular separation between ℓ_1 and l-jet$_0$
$\Delta R(\ell_1,l\text{-jet}_1)$	14.47	7.20	Angular separation between ℓ_1 and l-jet$_1$
$\Delta R(\ell_1,l\text{-jet}_2)$	4.60	3.92	Angular separation between ℓ_1 and l-jet$_2$
$\Delta R(b\text{-jet},l\text{-jet}_0)$	4.67	9.28	Angular separation between the b-jet and l-jet$_0$
$\not{E}_T$	5.04	4.07	Missing transverse energy
n_{jets}	5.17	6.94	Total number of jets (l and b)

Table B.3: Variables used in the NNs, which are trained in the flavor-split SR, for the utH and the ctH signal. The description of variables is analogous to Table B.2. The X signifies that a variable was pruned.

Variable	Added significance							
	ctH signal				utH signal			
	SR ee	SR $e\mu$	SR μe	SR $\mu\mu$	SR ee	SR $e\mu$	SR μe	SR $\mu\mu$
$p_T(\ell_1)$	16,17	22,39	11,17	23,26	9,87	22,10	20,14	24,42
$p_T(W_{\text{had}})$	4,32	4,19	6,08	X	X	4,67	4,06	X
$p_T(l\text{-jet}_0)$	X	X	X	3,42	8,94	X	X	3,70
$p_T(l\text{-jet}_1)$	10,08	10,64	16,70	X	X	15,15	11,52	X
$p_T(l\text{-jet}_2)$	X	X	X	7,75	X	X	X	8,83
ECIDS ℓ_0	18,11	20,46	X	X	19,22	20,56	X	X
ECIDS ℓ_1	9,72	X	18,20	X	13,89	X	11,69	X
$H_1(\ell_0,\ell_1)$	8,10	6,59	7,50	X	7,17	9,01	6,20	X
$m(\ell_0,b\text{-jet})$	8,10	9,12	8,14	8,23	6,82	9,11	X	X
$m(\ell_0,W_{\text{had}})$	4,16	X	X	X	14,62	X	X	13,05
$m(\ell_0,l\text{-jet}_0)$	X	X	4,24	X	3,32	X	X	X
$m(\ell_0,l\text{-jet}_1)$	X	X	X	4,96	X	9,40	7,58	X
$m(\ell_1,b\text{-jet})$	3,46	X	X	X	X	4,57	7,04	5,18
$m(\ell_1,l\text{-jet}_0)$	X	16,80	10,54	15,20	4,40	14,42	X	15,83
$m(\ell_1,l\text{-jet}_1)$	X	X	4,83	11,20	X	X	5,45	X
$m(\ell_1,l\text{-jet}_2)$	3,26	X	X	X	X	X	X	4,66
$m(b\text{-jet},W_{\text{had}})$	3,50	X	X	X	6,12	X	X	X
$m(b\text{-jet},l\text{-jet}_0)$	X	X	X	5,39	X	5,12	5,05	4,55
$m(l\text{-jet}_0,W_{\text{had}})$	X	X	X	X	3,98	X	X	X
$m(l\text{-jet}_0,l\text{-jet}_1)$	X	X	X	4,57	X	X	X	4,99
$m(l\text{-jet}_1,l\text{-jet}_2)$	X	X	X	X	3,73	X	X	X
ω	5,47	7,91	4,03	9,84	3,73	8,49	5,78	10,74
$\Delta R(\ell_0,\ell_1)$	X	X	4,17	X	X	X	X	X
$\Delta R(\ell_0,W_{\text{had}})$	X	3,92	8,88	8,82	X	4,32	4,33	7,56
$\Delta R(\ell_0,b\text{-jet})$	X	X	X	X	X	X	X	3,77
$\Delta R(\ell_0,l\text{-jet}_0)$	6,79	7,50	X	6,18	4,83	X	7,44	7,95
$\Delta R(\ell_0,l\text{-jet}_1)$	5,59	9,47	4,25	X	3,56	7,77	3,42	5,36
$\Delta R(\ell_1,W_{\text{had}})$	7,18	7,47	5,77	5,84	8,03	X	7,21	3,77
$\Delta R(\ell_1,b\text{-jet})$	X	X	3,18	X	3,72	3,78	X	X
$\Delta R(\ell_1,l\text{-jet}_0)$	X	X	X	X	X	3,82	X	4,64
$\Delta R(\ell_1,l\text{-jet}_1)$	X	3,99	X	X	4,72	7,08	X	10,20
$\Delta R(\ell_1,l\text{-jet}_2)$	X	4,63	3,02	X	X	6,75	X	X
$\Delta R(b\text{-jet},W_{\text{had}})$	X	X	3,75	3,59	X	X	3,41	3,78
$\Delta R(b\text{-jet},l\text{-jet}_0)$	X	3,66	5,48	7,71	X	X	X	6,43
$\Delta R(b\text{-jet},l\text{-jet}_1)$	X	X	X	4,97	X	X	X	X
$\Delta R(l\text{-jet}_0,l\text{-jet}_1)$	X	3,27	4,54	5,08	X	X	3,04	5,15
$\Delta R(l\text{-jet}_1,l\text{-jet}_2)$	X	X	X	X	4,12	X	X	X
$\not{E}_T$	X	X	3,19	4,04	3,97	X	4,49	X
n_{jets}	X	3,81	X	X	X	4,64	3,51	X

Bibliography

[1] "Standard Model of Elementary Particles". In: (Sept. 2019). URL: https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg.

[2] Yorikiyo Nagashima. *Elementary Particle Physics: Foundations of the Standard Model V2*. John Wiley and Sons, 2013. ISBN: 978-3-527-64890-0.

[3] Eckart Stein Walter Greiner Stefan Schramm. *Quantum Chromodynamics*. Springer-Verlag Berlin Heidelberg, 2007. ISBN: 978-3-540-48534-6.

[4] "Color Confinement". In: (Oct. 2019). URL: https://en.wikipedia.org/wiki/Color_confinement#/media/File:Quark_confinement.svg.

[5] Bernd Müller Walter Greiner. *Gauge Theory of Weak Interactions*. Springer-Verlag Berlin Heidelberg, 2009. ISBN: 978-3-540-87842-1.

[6] Roger Wolf. *The Higgs Boson Discovery at the Large Hadron Collider*. Springer International Publishing, 2015. ISBN: 978-3-319-18511-8.

[7] Y. Saka A. Ceccucci Z. Ligeti. "The CKM Quark-Mixing Matrix". In: (Feb. 2008). URL: https://pdg.lbl.gov/2008/reviews/kmmixrpp.pdf.

[8] K. Agashe et al. *Snowmass 2013 Top quark working group report*. 2013. arXiv: 1311.2028 [hep-ph].

[9] S. N. Mukherjee T. Morii C. S. Lim. *The Physics of the Standard Model and Beyond*. World Scientific, 2004. ISBN: 978-9-812-79560-1.

[10] Luciano Maiani. "The GIM Mechanism: origin, predictions and recent uses". In: *48th Rencontres de Moriond on Electroweak Interactions and Unified Theories*. Mar. 2013. arXiv: 1303.6154 [hep-ph].

[11] *Top Quarks + X Summary Plots Summer 2020*. Tech. rep. ATL-PHYS-PUB-2020-022. Geneva: CERN, Sept. 2020. URL: http://cds.cern.ch/record/2730251.

[12] *Top Quark Mass Summary Plots*. Tech. rep. ATL-PHYS-PUB-2019-036. Geneva: CERN, Sept. 2019. URL: http://cds.cern.ch/record/2690526.

[13] A. Quadt T.M. Liss F. Maltoni. "The Top Quark". In: (Sept. 2017). URL: https://pdg.lbl.gov/2017/reviews/rpp2017-rev-top-quark.pdf.

[14] M. Aaboud et al. "Measurement of the Higgs boson mass in the $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ channels with $\sqrt{s}$=13 TeV pp collisions using the ATLAS detector". In: *Physics Letters B* 784 (Sept. 2018), pp. 345–366. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2018.07.050. URL: http://dx.doi.org/10.1016/j.physletb.2018.07.050.

[15] Cen Zhang. "Effective approach to top-quark decay and FCNC processes at NLO accuracy". In: *Journal of Physics: Conference Series* 556 (Nov. 2014), p. 012030. ISSN: 1742-6596. DOI: 10.1088/1742-6596/556/1/012030. URL: http://dx.doi.org/10.1088/1742-6596/556/1/012030.

[16] Celine Degrande et al. "Automatic computations at next-to-leading order in QCD for top-quark flavor-changing neutral processes". In: *Phys. Rev. D* 91 (3 Feb. 2015), p. 034024. DOI: 10.1103/PhysRevD.91.034024. URL: https://link.aps.org/doi/10.1103/PhysRevD.91.034024.

[17] Adam Alloul et al. "FeynRules 2.0 — A complete toolbox for tree-level phenomenology". In: *Computer Physics Communications* 185.8 (2014), pp. 2250–2300. ISSN: 0010-4655. DOI: https://doi.org/10.1016/j.cpc.2014.04.012. URL: https://www.sciencedirect.com/science/article/pii/S0010465514001350.

[18] Johan Alwall et al. "MadGraph 5: going beyond". In: *Journal of High Energy Physics* 2011.6 (June 2011). ISSN: 1029-8479. DOI: 10.1007/jhep06(2011)128. URL: http://dx.doi.org/10.1007/JHEP06(2011)128.

[19] M. Aaboud et al. "Search for top-quark decays $t \to Hq$ with 36 fb^{-1} of pp collision data at $\sqrt{s}$ = 13 TeV with the ATLAS detector". In: *Journal of High Energy Physics* 2019.5 (May 2019). ISSN: 1029-8479. DOI: 10.1007/jhep05(2019)123. URL: http://dx.doi.org/10.1007/JHEP05(2019)123.

[20] M. Kado M. Carena C. Grojean and V. Sharma. "Status of Higgs Boson Physics". In: (May 2016). URL: https://pdg.lbl.gov/2016/reviews/rpp2016-rev-higgs-boson.pdf.

[21] A. Denner et al. "Standard model Higgs-boson branching ratios with uncertainties". In: *The European Physical Journal C* 71.9 (Sept. 2011). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-011-1753-8. URL: http://dx.doi.org/10.1140/epjc/s10052-011-1753-8.

[22] Georges Aad et al. "Evidence for $t\bar{t}t\bar{t}$ production in the multilepton final state in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 80.11 (2020), p. 1085. DOI: 10.1140/epjc/s10052-020-08509-3. arXiv: 2007.14858 [hep-ex].

[23] "The IFF Truth Classifier". In: (2020). URL: https://gitlab.cern.ch/atlas-physics/sm/ew/wgg/IFFTruthClassifier.

[24] S. Agostinelli et al. "Geant4—a simulation toolkit". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(03)01368-8. URL: https://www.sciencedirect.com/science/article/pii/S0168900203013688.

[25] G. Aad et al. "Observation of the associated production of a top quark and a Z boson in proton proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Journal of High Energy Physics* 2020.7 (July 2020). ISSN: 1029-8479. DOI: 10.1007/jhep07(2020)124. URL: http://dx.doi.org/10.1007/JHEP07(2020)124.

[26] Michael Benedikt et al. *LHC Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. DOI: 10.5170/CERN-2004-003-V-3. URL: https://cds.cern.ch/record/823808.

[27] Arpad Horvath. "The LHC experiment and its pre-accelerators". In: (Apr. 2006). URL: %7Bhttps://en.wikipedia.org/wiki/Large_Hadron_Collider#/media/File:LHC.svg%7D.

[28] Mike Lamont. "Performance of the LHC in 2016". In: (July 2016). URL: https://lhc-commissioning.web.cern.ch/performance/2016-performance.htm.

[29] D. Boussard et al. "The LHC superconducting cavities". In: *Proceedings of the 1999 Particle Accelerator Conference (Cat. No.99CH36366)*. Vol. 2. 1999, 946–948 vol.2. DOI: 10.1109/PAC.1999.795409.

[30] Burkhard Schmidt. "The High-Luminosity upgrade of the LHC: Physics and Technology Challenges for the Accelerator and the Experiments". In: *Journal of Physics: Conference Series* 706 (Apr. 2016), p. 022002. DOI: 10.1088/1742-6596/706/2/022002. URL: https://doi.org/10.1088/1742-6596/706/2/022002.

[31] Werner Herr and B Muratori. "Concept of luminosity". In: (2006). DOI: 10.5170/CERN-2006-002.361. URL: https://cds.cern.ch/record/941318.

[32] The ATLAS Collaboration et al. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08003–S08003. DOI: 10.1088/1748-0221/3/08/s08003. URL: https://doi.org/10.1088/1748-0221/3/08/s08003.

[33] Cheuk-Yin Wong. *Introduction to High-Energy Heavy-Ion Collisions*. USA: Oak Ridge National Laboratory, 1998. DOI: https://doi.org/10.1142/1128.

[34] M. Aleksa. "Status of the ATLAS Experiment". In: *Journal of Physics: Conference Series* 452 (July 2013), p. 012041. DOI: 10.1088/1742-6596/452/1/012041.

[35] A. Airapetian et al. *ATLAS detector and physics performance: Technical Design Report, 1*. Technical design report. ATLAS. Geneva: CERN, 1999. URL: http://cds.cern.ch/record/391176.

[36] *ATLAS inner detector: Technical Design Report, 1*. Technical design report. ATLAS. Geneva: CERN, 1997. URL: http://cds.cern.ch/record/331063.

[37] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1996. URL: http://cds.cern.ch/record/331061.

[38] *ATLAS tile calorimeter: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1996. URL: http://cds.cern.ch/record/331062.

[39] *ATLAS muon spectrometer: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1997. URL: http://cds.cern.ch/record/331068.

[40] Enrico Bothmann et al. "Event generation with Sherpa 2.2". In: *SciPost Physics* 7.3 (Sept. 2019). ISSN: 2542-4653. DOI: 10.21468/scipostphys.7.3.034. URL: http://dx.doi.org/10.21468/SciPostPhys.7.3.034.

[41] Paolo Nason. "A New method for combining NLO QCD with shower Monte Carlo algorithms". In: *JHEP* 0411 (2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146 [hep-ph].

[42] Torbjörn Sjöstrand et al. "An introduction to PYTHIA 8.2". In: *Computer Physics Communications* 191 (June 2015), pp. 159–177. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2015.01.024. URL: http://dx.doi.org/10.1016/j.cpc.2015.01.024.

[43] Atlas Top MC production. *Request FCNC in top pair production tqH, inclusive H decays*. 2020. URL: https://its.cern.ch/jira/browse/ATLMCPROD-7119.

[44] A. Salzburger. "Fast simulation in ATLAS". In: (Oct. 2011). URL: https://indico.cern.ch/event/144956/contributions/1380435/attachments/135861/192810/LPCC-Workshop-AS-2011-10-07.pdf.

[45] Flavia Dias. "The new ATLAS Fast Calorimeter Simulation". In: *PoS* ICHEP2016 (2016), p. 184. DOI: 10.22323/1.282.0184.

[46] *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data*. Tech. rep. ATLAS-CONF-2014-032. Geneva: CERN, June 2014. URL: http://cds.cern.ch/record/1706245.

[47] *Electron and Photon Selection and Identification for Run2*. Tech. rep. Geneva: CERN, Feb. 2021. URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/EGammaIdentificationRun2#Electron_Identification.

[48] Shunichi Akatsuka and Shion Chen. *PLV + LowPtPLT proposal*. Tech. rep. Geneva: CERN, Sept. 2019. URL: https://indico.cern.ch/event/848448/contributions/3564603/attachments/1908395/3152628/20190916_akatsuka_PLT.pdf.

[49] Shunichi Akatsuka and Shion Chen. *PLV Loose/Tight WP Re-Optm. and Addition of LowPtPLT*. Tech. rep. Geneva: CERN, Sept. 2019. URL: https://indico.cern.ch/event/846898/contributions/3557447/attachments/1904267/3145282/IFF_PLVWP_repotm_20190909.pdf.

[50] M. Aaboud et al. "Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data". In: *Journal of Instrumentation* 14.03 (Mar. 2019), P03017–P03017. ISSN: 1748-0221. DOI: 10.1088/1748-0221/14/03/p03017. URL: http://dx.doi.org/10.1088/1748-0221/14/03/P03017.

[51] G. Aad et al. "Muon reconstruction performance of the ATLAS detector in proton proton collision data at $\sqrt{s} = 13$ TeV". In: *The European Physical Journal C* 76.5 (May 2016). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-4120-y. URL: http://dx.doi.org/10.1140/epjc/s10052-016-4120-y.

[52] M. Aaboud et al. "Jet reconstruction and performance using particle flow with the ATLAS Detector". In: *The European Physical Journal C* 77.7 (July 2017). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-017-5031-2. URL: http://dx.doi.org/10.1140/epjc/s10052-017-5031-2.

[53] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. "The anti-ktjet clustering algorithm". In: *Journal of High Energy Physics* 2008.04 (Apr. 2008), pp. 063–063. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2008/04/063. URL: http://dx.doi.org/10.1088/1126-6708/2008/04/063.

[54] M. Aaboud et al. "Jet energy scale measurements and their systematic uncertainties in proton proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev. D* 96 (7 Oct. 2017), p. 072002. DOI: 10.1103/PhysRevD.96.072002. URL: https://link.aps.org/doi/10.1103/PhysRevD.96.072002.

[55] ATLAS Collaboration. *Jet energy scale and resolution measured in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. 2020. arXiv: 2007.02645 [hep-ex].

[56] *Tagging and suppression of pileup jets with the ATLAS detector*. Tech. rep. ATLAS-CONF-2014-018. Geneva: CERN, May 2014. URL: https://cds.cern.ch/record/1700870.

[57] Philipp Jonas Windischhofer. “Heavy-flavour jet tagging in ATLAS”. In: (Jan. 2020). URL: http://cds.cern.ch/record/2706702.

[58] Steve Farrell. “Overlap Removal Tools”. In: (Sept. 2017). URL: https://indico.cern.ch/event/631313/contributions/2683959/attachments/1518878/2373377/Farrell_ORTools_ftaghbb.pdf.

[59] The ATLAS collaboration. “Operation of the ATLAS trigger system in Run 2”. In: *Journal of Instrumentation* 15.10 (Oct. 2020), P10004–P10004. ISSN: 1748-0221. DOI: 10.1088/1748-0221/15/10/p10004. URL: http://dx.doi.org/10.1088/1748-0221/15/10/P10004.

[60] Nedaa Asbah et al. *Search for $t\bar{t}t\bar{t}$ Standard Model Production in the Multilepton Final State in proton-proton collisions with the ATLAS Detector*. Tech. rep. ATL-COM-PHYS-2019-302. Geneva: CERN, Apr. 2019. URL: https://cds.cern.ch/record/2670354.

[61] Philip Sommer. “Electron Charge ID Selector Tool ”. In: (Mar. 2019). URL: https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ElectronChargeFlipTaggerTool.

[62] Nima Arkani-Hamed et al. *Scattering Amplitudes and the Positive Grassmannian*. 2014. arXiv: 1212.5605 [hep-th].

[63] Jayawant N. Mandrekar. “Receiver Operating Characteristic Curve in Diagnostic Test Assessment”. In: *Journal of Thoracic Oncology* 5.9 (2010), pp. 1315–1316. ISSN: 1556-0864. DOI: https://doi.org/10.1097/JTO.0b013e3181ec173d. URL: https://www.sciencedirect.com/science/article/pii/S1556086415306043.

[64] Joerg Meyer. “Measurement of the Top Quark Mass using Dilepton Events and a Neutrino Weighting Algorithm with the D0 Experiment at the Tevatron (Run II)”. In: (Jan. 2007). DOI: 10.2172/935242. URL: https://www.osti.gov/biblio/935242.

[65] Victor Mukhamedovich Abazov et al. “Precise measurement of the top quark mass in dilepton decays using optimized neutrino weighting”. In: *Phys. Lett. B* 752 (2016), pp. 18–26. DOI: 10.1016/j.physletb.2015.10.086. arXiv: 1508.03322 [hep-ex].

[66] Catherine Bernaciak et al. “Fox-Wolfram moments in Higgs physics”. In: *Physical Review D* 87.7 (Apr. 2013). ISSN: 1550-2368. DOI: 10.1103/physrevd.87.073014. URL: http://dx.doi.org/10.1103/PhysRevD.87.073014.

[67] Liliana Teodorescu. “Artificial neural networks in high-energy physics”. In: (2008). DOI: 10.5170/CERN-2008-002.13. URL: https://cds.cern.ch/record/1100521.

[68] C. A. L. Bailer-Jones, R. Gupta, and H. P. Singh. “An introduction to artificial neural networks”. In: (Feb. 2001). arXiv: astro-ph/0102224.

[69] phi-t. “The Neurobayes User Guide”. In: (Apr. 2010). URL: %7Bhttps://manualzilla.com/doc/6923372/the-neurobayes-r-user-s-guide---neurobayes.de---phi-t%7D.

[70] M. J. Costa. *Experimental systematic uncertainties (and object reconstruction) on top physics, their correlations, comparison ATLAS vs CMS (vs Tevatron) and common agreements*. 2014. arXiv: 1412.0875 [hep-ex].

[71] C. Schwick J. Boyd. “Luminosity measurements and systematic uncertainties”. In: (2017). URL: https://indico.cern.ch/event/663598/contributions/2782457/attachments/1574279/2543359/Lumi-Evian2017.pdf.

[72] Will Buttinger. “Extended Pile-up reweighting”. In: (Nov. 2017). URL: https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ExtendedPileupReweighting.

[73] Georges Aad et al. “Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data”. In: *JINST* 14.12 (2019), P12006. DOI: 10.1088/1748-0221/14/12/P12006. arXiv: 1908.00005 [hep-ex].

[74] Philip Sommer Cyril Becot. “Electron Efficiency Run 2”. In: (June 2019). URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ElectronEfficiencyRun2.

[75] Philip Sommer Kristin Lohwasser. “Electron Efficiency Correction Model”. In: (Nov. 2018). URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ElectronEfficiencyCorrelationModel.

[76] Stefano Zambito. “MCP Analysis Guidelines MC16”. In: (Mar. 2021). URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MCPAnalysisGuidelinesMC16.

[77] D. J. Summers. “Hybrid rings of fixed 8T superconducting magnets and iron magnets rapidly cycling between -2T and +2T for a muon collider”. In: *eConf* C960625 (1996), ACC033. arXiv: physics/0108001.

[78] ATLAS Collaboration. “Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector”. In: *Phys. Rev. D* 96 (2017), p. 072002. DOI: 10.1103/PhysRevD.96.072002. arXiv: 1703.09665 [hep-ex].

[79] Chris Malena Delitzsch. “Jet ETmiss Recommendations Release 21”. In: (Mar. 2021). URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetEtmissRecommendationsR21.

[80] Bingxuan Liu. “B-tagging calibration recommendtions Release 21”. In: (Mar. 2020). URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/BTagCalib2017.

[81] G. Aad et al. "Measurement of the tt production cross-section in the lepton+jets channel at s=13 TeV with the ATLAS experiment". In: *Physics Letters B* 810 (Nov. 2020), p. 135797. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2020.135797. URL: http://dx.doi.org/10.1016/j.physletb.2020.135797.

[82] Lara Lloret Iglesias. "Diboson production at the LHC". In: *PoS* LHCP2016 (2016), 022. 14 p. DOI: 10.22323/1.276.0022. URL: https://cds.cern.ch/record/2276624.

[83] Stefano Forte and Stefano Carrazza. *Parton distribution functions*. 2020. arXiv: 2008.12305 [hep-ph].

[84] J. W. Fowler. "Maximum-Likelihood Fits to Histograms for Improved Parameter Estimation". In: *Journal of Low Temperature Physics* 176.3-4 (Feb. 2014), pp. 414–420. ISSN: 1573-7357. DOI: 10.1007/s10909-014-1098-4. URL: http://dx.doi.org/10.1007/s10909-014-1098-4.

[85] Michele Pinamonti. "Statistical methods at ATLAS and CMS". In: (May 2018). URL: https://indico.cern.ch/event/727396/contributions/3021899/attachments/1657532/2654085/Statistical_methods_at_ATLAS_and_CMS_2.pdf.

[86] Elad Meir and Tirza Routtenberg. *Cramer-Rao Bound for Estimation After Model Selection and its Application to Sparse Vector Estimation*. 2021. arXiv: 1904.06837 [eess.SP].

[87] Thomas Calvet Georges Aad. "Automatic binning implematation in TTHFitter". In: (Jan. 2016). URL: https://indico.cern.ch/event/455289/contributions/1953694/attachments/1209081/1762963/Calvet_binning_Htop_160108.pdf.

[88] Spyros Argyropoulos. "Systematics pruning and smoothing". In: (Jan. 2017). URL: https://indico.cern.ch/event/691683/contributions/2873279/attachments/1593521/2522846/PruningSmoothing.pdf.

[89] Tomas Dabo. "Fit Problems Tutorial". In: (Mar. 2019). URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/FitProblemsTutorial.

[90] Eilam Gross and Ofer Vitells. "Wald's theorem and the "Asimov" data set". In: (Dec. 2009). URL: https://indico.cern.ch/event/74940/contributions/2088584/attachments/1047729/1493442/Wald-Asimov.pdf.

[91] Tomas Dado and Alexander Held. "TRExFitter Documentation". In: (2020). URL: https://trexfitter-docs.web.cern.ch/.

[92] Oliver Thielmann. "Search for flavor-changing neutral currents via anomalous ctH and utH couplings". Presented 20 Dec 2019. 2019. URL: http://cds.cern.ch/record/2715194.

Name, Vorname: ....

Erklärung

gem. § 14 Abs. 7 Prüfungsordnung vom 15.04.2013

Hiermit erkläre ich, dass ich die von mir eingereichte Abschlussarbeit (Master-Thesis) selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie Stellen der Abschlussarbeit, die anderen Werken dem Wortlaut oder Sinn nach entnommen wurden, in jedem Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

....	....
Datum	Unterschrift

Erklärung

Hiermit erkläre ich mich damit einverstanden, dass meine Abschlussarbeit (Master-Thesis) wissenschaftlich interessierten Personen oder Institutionen und im Rahmen von externen Qualitätssicherungsmaßnahmen des Studienganges zur Einsichtnahme zur Verfügung gestellt werden kann.
Korrektur- oder Bewertungshinweise in meiner Arbeit dürfen nicht zitiert werden.

....	....
Datum	Unterschrift