Towards a measurement of the tHq process in channels with hadronic tau lepton decays

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Tanja Holm
aus
Troisdorf

Bonn

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. Ian C. Brock
2. Gutachterin: Priv. Doz. Dr. Tatjana Lenz

Tag der Promotion:
Erscheinungsjahr:

Contents

Introduction

Wen considered the nature of time and understood that the universe is, instant by instant, recreated anew. Therefore, he understood, there is in truth no past, only a memory of the past. Blink your eyes and the world you see next did not exist when you closed them. Therefore, he said, the only appropriate state of mind is surprise. The only appropriate state of the heart is joy. The sky you see now, you have never seen before. The perfect moment is now. Be glad of it.

– Terry Pratchett, *Thief of Time*

If you think of it, the world is a miraculous thing. If it exists or not philosophers still try to find out[1] but the way it presents itself to us in its various, colorful ways is undoubtedly marvelous. Even the most unassuming object, like a pen, consists of many, many unanswered questions. The easiest of which is usually how it works. One will find out by disassembling that a pen consist of four main parts: a main body, a liquid filled cylinder that leaves the line, some sort of vessel that holds the cylinder. This small vessel is connected to the main vessel in a way that allows it to switch between a 'writing cylinder outside' and 'writing cylinder inside' state. The last piece is a spring which holds this piece in place unless the user wants to alternate between these states. This of course just opens up more questions: why does the liquid in the cylinder leave a line on solids? Does the spring work the same in both directions? Am I able to put everything back together without destroying it or losing something? Most questions can be traced back to the next smaller unit, where new questions will surely arise. One can follow that path and will always find that some questions can only be answered by trying to understand a component of it. Following that logic most things can actually be explained by finding the most fundamental components of everything and understanding how they work. This activity is also called particle physics and this thesis tries to answer a little question about how some rarer fundamental objects interact with each other.

The objects this thesis is looking at are primarily the H boson and the top quark. When looking for an interaction between those, the most common process actually involves two instead of one t-quark, while the interaction including just one t-quark is very rare. On the flip side it is possible to receive some information on the H boson – t-quark interaction via this process which is unavailable in the process with two t-quarks (namely the relative sign of interaction). For this reason, this thesis takes the task to find events like this in the vast amount of data that has been produced and recorded at the LHC and ATLAS to answer these kinds of questions.

---

[1] This thesis follows the assumption that it in fact exists, to avoid using the conjunctive too much.

This thesis describes that process by first giving an overview of the most fundamental components we know of and how they work in chapter 2, the process this thesis wants to look at (chapter 3), the tools that are used to analyze the data in chapter 4 and the experiment that took the data in chapter 5. Second, the selection of events (chapter 6) is described. Chapter 7 describes the expected results from the statistical analysis and also presents the observation. Finally, chapter 8 summarizes how far this thesis got with answering its questions.

Theory

„Mein Freund der Dichter beschrieb nun die einfachsten Dinge, die er finden konnte", fuhr er fort, "und stellte fest, daß es das Schwierigste überhaupt war. Es war leicht, einen Palast aus Schnee und Eis zu beschreiben, aber unsäglich schwer, dasselbe mit einem einzelnen Haar zu tun. Oder einem Löffel. Einem Nagel. Einem Zahn. Einem Salzkorn. Einem Holzsplitter. Einer Kerzenflamme. Einem Wassertropfen."

– Walther Moers, *Die Stadt der träumenden Bücher*

This chapter aims at giving an overview of the base model that we use to describe the most fundamental objects that we know and their interactions, it is called Standard Model 'SM' [1, 2]. As the name says it is a model, a very well tested model and the best we have today, but not the full story. An overview of things that cannot be described by the SM is given in section 2.3 which also touches on a few extensions of the SM which might explain them.

2.1 Elementary particles

The fundamental objects that exist in the SM can be divided into 'matter constituents' and 'interactors'. They share the property of being point-like meaning they do not have a finite spread. Therefore, ranking them by size is nonsensical. Instead, another property called 'spin' can be used to classify them. The spin behaves mathematically like the torque in macroscopic objects. So one can talk about concepts like the orientation of the spin and the fact that it is preserved, but a point-like object can obviously not turn[1]. Particles that have a half-integer spin belong to the fermions and are 'matter constituents'. Bosons are particles with an integer spin and all 'interactors' are part of this group.[2]

2.1.1 Fermions

All fundamental fermions possess a spin of $\frac{1}{2}$ and can be further grouped according to their attributes, also called charges. Neutrinos, ν, are the 'easiest' particles to describe. They just exist as particles and

---

[1] Turning would require a change under rotation. In a spacious object a rotation moves the components of an object with respect to the rotation center. In a point-like object there are no components, but the only possible rotation center is the object itself. Therefore, no components can move with respect to the rotation center.

[2] Non-fundamental particles possess spins, too, which is talked about in a moment.

anti-particles[3] and belong to one of three 'generations'. Each generation holds the two neutrinos and two other leptons which have an additional attribute: the electric charge. It is -1 for charged-lepton-particles and +1 for the charged lepton anti-particles. Additionally, charged leptons also inherit a quantity called mass which is the same for each charged lepton in a generation. The lightest charged lepton is called electron, e^-, and thus the other leptons in its generation are called: anti-electron (or positron), e^+, electron-neutrino, ν_e, and anti-electron-neutrino, $\overline{\nu}_e$. The same naming scheme is used for the next heavier muon, μ, and tau(on), τ, generation, where the latter holds the heaviest existing leptons.

The last type of fundamental fermion is called a quark, and they also come in three generations of particles. As the charged-leptons they have a mass and an electric charge (in each generation one with $\frac{+2}{3}, \frac{-2}{3}, \frac{1}{3}, \frac{-1}{3}$). In addition to that, they possess a color-charge. It can take values of red, blue and green for quarks and anti-red, anti-blue and anti-green for anti-quarks.

An overview of all fermions can be found in table 2.1.

Particle	lepton	quark	generation	electric charge [e]	color charge	mass [MeV]
e^-	x	-	1	-1	-	0.51
e^+	x	-	1	$+1$	-	0.51
μ^-	x	-	2	-1	-	106
μ^+	x	-	2	$+1$	-	106
τ^-	x	-	3	-1	-	1 776
τ^+	x	-	3	$+1$	-	1 776
$\nu_e/\overline{\nu}_e$	x	-	1	$-$	-	≈ 0
$\nu_\mu/\overline{\nu}_\mu$	x	-	2	$-$	-	≈ 0
$\nu_\tau/\overline{\nu}_\tau$	x	-	3	$-$	-	≈ 0
u	-	x	1	$+2/3$	x	2.2
$\overline{u}$	-	x	1	$-2/3$	x	2.2
d	-	x	1	$-1/3$	x	4.7
$\overline{d}$	-	x	1	$+1/3$	x	4.7
s	-	x	2	$-1/3$	x	93
$\overline{s}$	-	x	2	$+1/3$	x	93
c	-	x	2	$+2/3$	x	1 270
$\overline{c}$	-	x	2	$-2/3$	x	1 270
b	-	x	3	$-1/3$	x	4 180
$\overline{b}$	-	x	3	$+1/3$	x	4 180
t	-	x	3	$+2/3$	x	172 000
$\overline{t}$	-	x	3	$-2/3$	x	172 000

Table 2.1: List of fermions in the SM. Masses are rounded from [3]. ν and $\overline{\nu}$ share the same properties and can be interpreted as either separate or the same type of particle.

2.1.2 Bosons

Bosons are roughly divided by their spin. On the one hand we have the massive Higgs boson H (suffixes like 'boson' or 'quark' are omitted when this enhances the readability of the text) which is the only particle

---

[3] In another description the neutrino particles and anti-particles are even the same objects.

of spin 0 (more on this later) and on the other hand we have vector-bosons (or 'interactors') which carry a spin of 1. The latter can again be divided by their charges and masses: gluons and photons do carry no electric charge and are also massless. Gluons differentiate themselves by being the only bosons which carry color-charge (always one color and one anti-color) which puts them in a unique connection to quarks. Z and W do have a mass, but only the W carries also an electric charge and can thus be seen as either a W^+ (electric charge of +1) or the anti-particle W^- (electric charge of -1).

All bosons can also be seen in table 2.2.

particle	couples to	electric charge [e]	color charge	spin	mass [GeV]
γ	electric charge	0	-	1	0
W^+	chirality	1	-	1	80
W^-	chirality	-1	-	1	80
Z	weak charge	0	-	1	91
g	color charge	0	x	1	0
H	mass	0	-	0	125

Table 2.2: List of bosons in the SM. Masses are rounded from [3]. More on the coupling is described in section 2.2.3.

2.2 Quantum Field Theory

Now, as we have a list of basic constituents, the next task is to describe how those constituents behave and how they interact. It is important to note that we leave the realm of our normal understanding. The objects we are talking about are way smaller and the interactions happen at so small scales that the normal rules of our macroscopic world do not hold anymore. Particles are described as a wave or field (often denoted as ψ as a function of space $\vec{x}$ and time t). The effects of this are sometimes confusing to us, but this confusion is an artifact of our macroscopic world that we are used living in and not of the particle physics realm. Therefore, to be able to describe particle physics instead of words this thesis is using the language of mathematics which is way clearer and less biased by our interpretations of the world. This thesis does not aim at giving a rigorous description of the mathematics involved, but rather a brief overview that aids the reader in understanding the following chapters.

Two words on the notation: this thesis is written using 'natural units'

$$c = \hbar = 1 \tag{2.1}$$

which is a convenient way of writing down equations. However, the connection to 'measurable units' is lost and everything is measured in units of powers of energy. At any time a sensible reconversion in metric units is possible. On another note, this section does not introduce all mathematical symbols and indices used and focuses on the fields. More complete descriptions can for example be found in [3–6].

For a start some central equations need to be introduced. A central equation of quantum mechanics is the Dirac wquation:

$$i\gamma^\mu \partial_\mu \psi(\vec{x}, t) = m\psi(\vec{x}, t) \tag{2.2}$$

It describes the change of ψ over time (left side) depending on the state of the particle and its mass (m).

With this a stable fermion can be described. From this the Lagrange density ($\mathscr{L}$) can be formed:

$$\mathscr{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \tag{2.3}$$

with $\bar{\psi}$ describing an anti-fermion. Analogous equations exist for bosons of different spins, often denoted by Φ or capital latin letters. Interactions are incorporated into $\mathscr{L}$ by adding terms which feature both fields (although these terms are sometimes hidden inside the field definition to improve readability). $\mathscr{L}$ is a concept that describes the action (S) in an infinitesimal area and time (which is why it is called a density).

$$S = \int d^4x \, \mathscr{L} \tag{2.4}$$

The evolution of ψ can be found by extremizing S under different conditions (principle of minimal action). In the end the form of $\mathscr{L}$ holds all information about our description of particle physics.

2.2.1 The Standard Model

To describe the complete standard model (SM) of particle physics we need to look at the mathematical descriptions of the particles described in section 2.1. Fermions were already introduced as ψ. To describe all existing fermions (quarks, neutral and charged leptons) we need to sum over them. The fermionic part of the full Lagrange density of the standard model ($\mathscr{L}_{\text{SM}}$) becomes:

$$\mathscr{L}_f = \sum_n \bar{\Psi}_n i\gamma^\mu \partial_\mu \Psi_n \tag{2.5}$$

For bosons a similar description can be introduced: the three different fields B_μ, W_μ^i and G_μ^a correspond to the electromagnetic, weak and strong interaction. Therefore, G_μ^a corresponds to the eight different gluons and $W_\mu^{1,2}$ correspond to the positively and negatively charged W. The photon and Z are linear combinations of B_μ and W_μ^3. This gives us the second and third part of our full Lagrange density:

$$\mathscr{L}_{QED} = j_1^\mu B_\mu + j_i^\mu W_\mu^i \tag{2.6}$$

$$\mathscr{L}_{QCD} = j_a^\mu G_\mu^a \tag{2.7}$$

These descriptions did not include the masses of our objects. To include them it is necessary to introduce another field Φ which is called the Higgs field [7–12]. All mass terms stem from an interaction with this field and take this form for fermions:

$$\mathscr{L}_{mass} = y_x \Phi \bar{\Psi}_L \Psi_R + y_x \Phi^* \bar{\Psi}_R \Psi_L \tag{2.8}$$

with y_x being a dimensionless coupling constant which determines how strong each coupling is. It depends on the mass of the particle and another factor k_x. This factor can take into account different coupling strength between different particle types (for example fermions and bosons):

$$y_x = m_x k_x \tag{2.9}$$

In the standard model all k_x have the value 1.

At last, also the H has its own term in the Lagrangian:

$$\mathscr{L}_H = 0.5 M^2 |\Phi|^2 \tag{2.10}$$

In total the SM Lagrangian can be written as:

$$\mathscr{L}_{SM} = \mathscr{L}_f + \mathscr{L}_{QED} + \mathscr{L}_{QCD} + \sum_x \mathscr{L}_{mass,x} + \mathscr{L}_H \tag{2.11}$$

2.2.2 Feynman diagrams

The Lagrangian density describes which interactions are possible, for a specific process one would want to calculate its probability

$$P = |\langle X_{out} | S | X_{in} \rangle|^2 \tag{2.12}$$

with the interaction matrix S and the incoming and outgoing state being described by $X_{out/in}$ (in this context it can include both fermions and bosons). The widely used application of this is to calculate the cross-section of a process σ which gives a measure for the prevalence of a process:

$$\sigma = \int d\Omega \frac{d\sigma}{d\Omega} \tag{2.13}$$

$$d\sigma = |M|^2 \frac{PS}{sym} (2\pi)^n \delta^n(p) \tag{2.14}$$

The factors at the end of the formula are needed to make all momenta add up ($\delta^n(p)$), take into account different contributions to the same diagram (sym) and describe the available phase space (PS). The phase space can roughly be described as the allowed behaviors of the decay particles. If for example more angles between two decay particles are kinematically allowed, the available phase space will be bigger and thus the cross-section will be higher. The most interesting part of the calculation is bound in M, also called the matrix-element. It includes the probabilities of getting from the initial state to the final state.

For a more visual description, Feynman diagrams can be used. These represent particles and interactions described in $\mathscr{L}$, which also gives the rules to translate a diagram back to a written formula. Every properly drawn Feynman diagram which connects the incoming particles with the outgoing particles will contribute to M. However, most processes described here are dominated by the simplest diagrams (the ones with the least vertices) called leading-order diagrams (LO).

(a) electroweak boson   (b) gluon   (c) fermion   (d) H boson

Figure 2.1: Feynman diagram propagators.

As a first step, the parts of a Feynman diagram possible in $\mathscr{L}_{SM}$ are described. Figure 2.1 shows the representation of the different particles, they are also called propagators because they represent a particle propagating in spacetime. The convention used here is to have time going from left to right, i.e. the lines on the left of a diagram represent the incoming particles and the lines on the right represent outgoing

(a) Fermion-*W* boson   (b) Fermion-*Z* boson   (c) Fermion-γ boson   (d) Quark-*g* boson

(e) Triple gluon   (f) Triple *W*-*Z* boson   (g) Triple *W*-γ boson
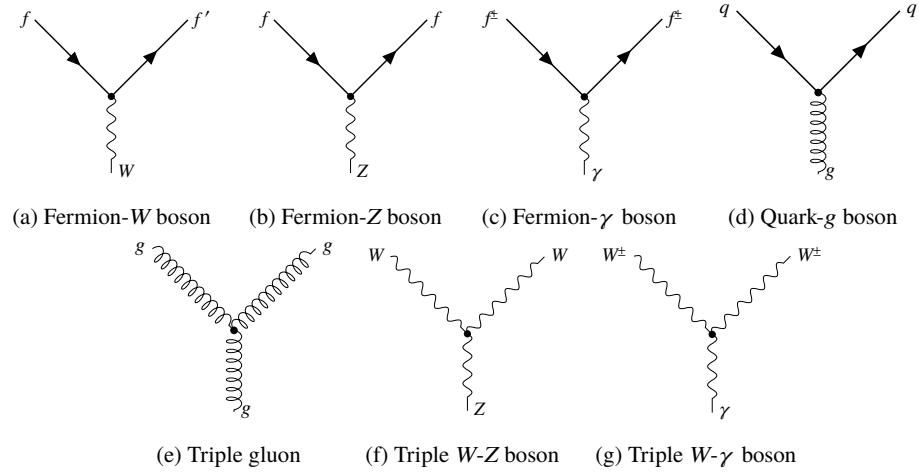
Figure 2.2: Triple Feynman diagram vertices.

(a) Quartic gluon   (b) Quartic *W* boson   (c) Quartic *W*-γ boson   (d) Quartic *W*-*Z* boson   (e) Quartic EW-boson
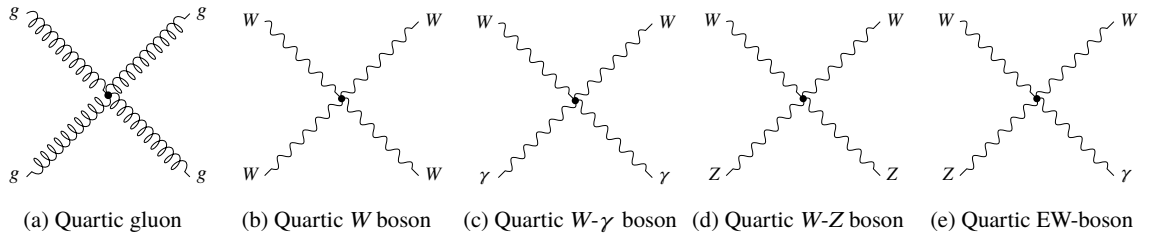
Figure 2.3: Quartic Feynman diagram vertices.

(a) Massive lepton *H*   (b) Quark *H*   (c) *W*-*H* boson   (d) *Z*-*H* boson

(e) Quartic *W*-*H* boson   (f) Quartic *Z*-*H* boson   (g) Triple *H* boson   (h) Quartic *H* boson

Figure 2.4: Vertices with *H*
Feynman diagram vertices including a *H* boson.

particles. The vertical expansion is however not attached to the spatial expansion of the process.[4] To connect propagators vertices are drawn. These represent the interactions. All $\mathscr{L}_{SM}$ possible vertices are shown in figures 2.2 to 2.4 and ordered according to the part of $\mathscr{L}_{SM}$ where they are defined.

(a) s-channel   (b) t-channel
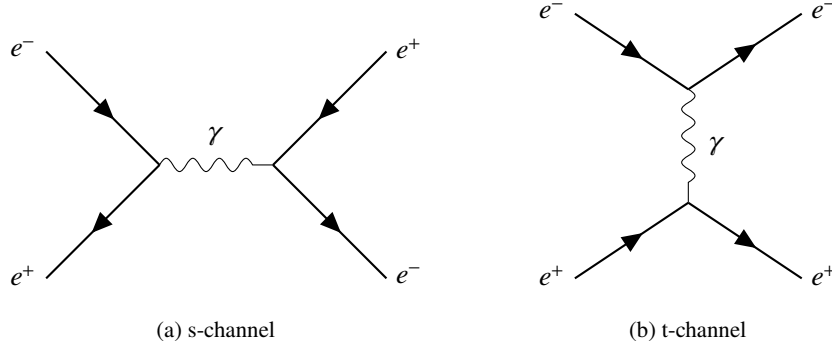
Figure 2.5: Feynman diagrams for electron-positron scattering in s-channel (figure 2.5(a)) and t-channel (figure 2.5(b)).

Two examples of complete Feynman diagrams are shown in figure 2.5. As one can see they both feature the same initial and final state particles, but the shapes are different. Figure 2.5(b) has a 'T'-like structure and is often called t-channel, the structure of figure 2.5(a) on the other hand is called s-channel.

To calculate the complete probability for the process, both diagrams have to be added first before taking the square (see equation (2.12)). This gives rise to three terms:

$$\sigma \propto |\mathscr{M}^2| = |(M_s + M_t)| = |M_s|^2 + |M_t|^2 + \mathrm{Re}\left(2M_s M_t^\dagger\right) \tag{2.15}$$

The last one is called interference term. Depending on the process (everything is a complex calculation) this term can increase or decrease the probability of the process.
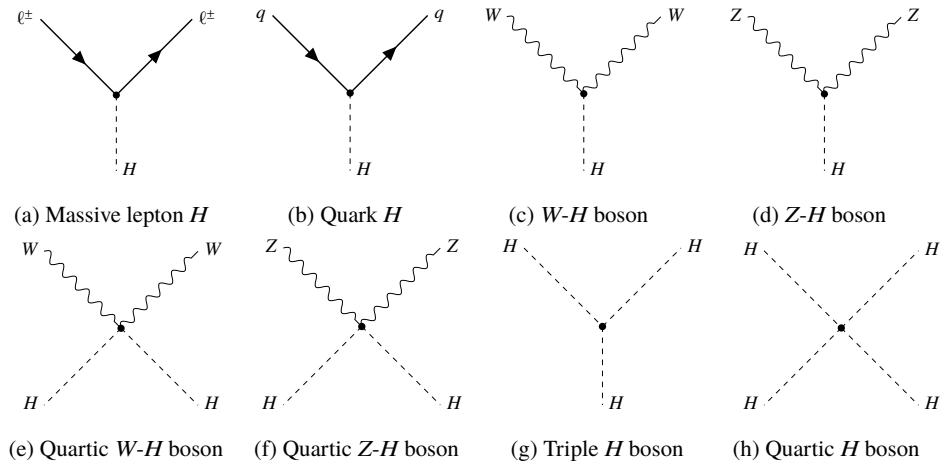
One more word about the s in s-channel: this can also be interpreted as a variable s defined as:

$$s = (p_1 + p_2)^2 = (p_3 + p_4)^2 \tag{2.16}$$

with p_1 and p_2 denoting the momentum of the two incoming particles and p_3 and p_4 of the two outgoing particles (in a $2 \rightarrow 2$ process). $\sqrt{s}$ captures the energy available for an interaction and thus available for the production of any new particle.

2.2.3 A few words on the differences between interactions

Equation (2.6) and equation (2.7) showed the minimal possible description of the electroweak and strong interaction. A more detailed description about the differences between the interactions are given in this section.

The electroweak interaction has four interactors: massless photons γ which couple to electrically charged objects and massive W^+, W^- and Z. These can interact with objects based on properties called 'chirality' and 'weak charge'.[5] Due to the mass of the latter the weak interaction has limited range (since the bosons

---

[4] The vertical ordering of external lines might be important for calculation when the 'same' particle is present twice, but as this thesis is not aiming at calculating any Feynman diagrams this is not viewed separately.

[5] Chirality refers to an intrinsic property of particles which is also called 'handedness' and in the limit of no mass describes the configuration of spin and direction. In practice all fermions, γ and other W as well as Z can interact with the W to a degree.

will decay into lighter objects). The γ on the other hand has an unlimited range, but in practice is limited by the abundance of charged objects in matter. Thus, it can travel through the universe but not through the earth. The $W^{\pm}$ play an important role in the SM as they are the only bosons which facilitate a change of flavor and thus enable the direct decay of heavier fermions into lighter fermions. Any change of flavor is predominantly done within the same generation. As they carry an electrical charge of ± 1 they will actually always change the flavor of the fermions they are interacting with. A useful example is a t decay which is shown in figure 2.6. The t will decay into a b as shown here in 95.7($\pm$3.4)% [3] of all cases, on the other hand the leptonic W decay shown will just arise in about 33.3% [3] of all t decays. Those fractions are often denoted as

$$\Gamma_{i/\text{tot}} = \frac{\Gamma_i}{\Gamma_{\text{tot}}} \tag{2.17}$$

called relative branching ratios (BR). With the total decay width of a particle with mass m

$$d\Gamma_{\text{tot}} \propto \frac{|\mathcal{M}|^2}{2m} \Gamma_{\text{tot}} = \frac{1}{\tau} \tag{2.18}$$

being the inverse of the (mean) lifetime of a particle (or bound state) τ. Any change in fermion flavor by a neutral particle (Z,γ,g,H) at leading-order is forbidden in the SM.
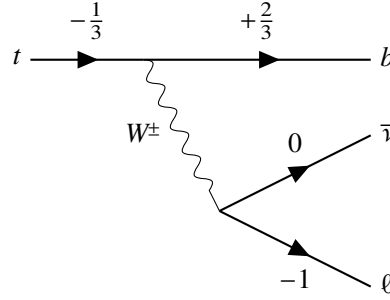
Figure 2.6: Feynman diagram for an example top decay. The t decays into a b under radiation of a W^- which carries away an electric charge of -1 and decays into a charged lepton ℓ (can be e, μ, τ). Electric charges in units of the electron charge e are indicated next to the propagators.

The strong interaction is transmitted by the massless gluons which couple to objects with color-charge. Interestingly, gluons carry color-charge themselves. This leads to gluons coupling to themselves. And while the coupling is higher in comparison to the weak force, this property leads to a limited range of the strong interaction since any colored objects form a bound state (hadron) which is color neutral to the outside (see also section 2.2.5).

That effect poses a problem for the detection of quarks (which carry color-charge) as they will hadronize before they can be detected (more on this in section 5.4.5). However, the top quark is unique in this regard as it even decays before it hadronizes – due to its high mass it has a lifetime of only 5×10^{-25} s [3].

An even bigger problem is, that QCD processes at low energies are no longer dominated by the leading-order diagrams. This imposes serious problems in calculation. Exact treatment of these processes is still

---

Since the Z is a mixing of two fields it couples to particles in a mixed behavior, too, sometimes called the weak charge. In practice, it also couples to all fermions in some way, as well as to $W^{\pm}$ and H. Keep in mind that just one coupling is needed to enable an interaction. For example, a H can couple to W^+ due to its mass, but a H does not hold chirality.

not possible to this day, but in combination with statistical methods relevant results can be obtained by limiting the calculation to a grid.

2.2.4 The Higgs boson

The H boson is the latest addition to the SM. As described before it arises from the Higgs mechanism which gives fermions and bosons their inert mass. However, the H boson itself does not 'transfer' mass related information or interaction itself – this would relate to gravitation which is not described in the SM (or by any other Quantum Field Theory). Instead, the H can be viewed as an excitation of the field which gives elementary particles their masses.[6] As it was just observed in 2012 [14–16] the study of its properties is an ongoing field [17, 18]. Therefore, by studying the H the SM can be tested once more and—maybe even more exciting—any deviations from the easiest description would lead to a more advanced description of physics and might even give a hint on how to solve some of the still existing unexplained phenomena explained in section 2.3.

Since the H is of special interest to this thesis it is useful to go into a bit more detail regarding its properties. From the tests that could be run until now, the H properties are fully consistent with the SM. One property that is not predicted from the SM was the H mass, it has been extensively studied and is measured to be (125.25 ± 0.17) GeV [3]. At this mass the dominating decay mode is the decay into two b-quarks $\Gamma_{H\to bb/\text{tot}} \approx 58\%$ followed by the two bosonic decay modes $\Gamma_{H\to WW/\text{tot}} \approx 22\%$, $\Gamma_{H\to gg/\text{tot}} \approx 8.2\%$ and another fermionic decay $\Gamma_{H\to\tau\tau/\text{tot}} \approx 6.3\%$ (all numbers from [19]). We see that $\Gamma_{H\to\tau\tau/\text{tot}}$ is already one order of magnitude smaller than $\Gamma_{H\to bb/\text{tot}}$! Further decay modes have even smaller relative branching ratios which is displayed in figure 2.7(a). Similar to the BRs the processes to create a H in the first place (production modes) depend on the H mass. Additionally, they depend on the available $\sqrt{s}$ as well as on the available particles in the initial state. Relevant for this thesis are any production modes with quarks or gluons in the initial state and their cross-sections at $\sqrt{s} = 13$ TeV (see section 5.1). These production modes can be seen in figure 2.7(b), they also span several orders of magnitude. More details on the relevant processes are described in chapter 3 including Feynman diagrams in figures 3.1, 3.9 and 3.11.

2.2.5 Bound states

As described in section 2.2.3, colored objects (namely quarks) persist in color neutral bound states, called hadrons. There are two basic ways to create a color neutral objects with several quarks: mesons, the combination of a quark and an anti-quark carrying opposite color-charges, or baryons the combination of three quarks (or three anti-quarks) with the three individual (anti-)color-charges.

As an example, the mesons with the lowest mass (and thus the most common) are a bound state of the lightest quarks: u and d. They are called pions and exist with three electromagnetic charges π^+ ($u + \bar{d}$), π^- ($\bar{u} + d$) both with mass 139.57 MeV [3] and π^0 (a superposition of $u + \bar{u}$ and $d + \bar{d}$) with mass 134.98 MeV [3]. Charged pions overwhelmingly decay into an (anti-)muon and the corresponding (anti-)neutrino while neutral pions overwhelmingly decay into two γ, the decay in lighter hadrons is not possible. On the other hand, heavier mesons may decay into pions. This includes mesons with different quark content but also

---

[6] The mathematical description of this is called 'spontaneous symmetry breaking' and refers to a field which is not minimal at 0, but has a rotational symmetry around 0, sometimes referred to as 'Mexican hat potential'. Graphically, the H boson in this field can be described by an excitation of the field perpendicular to 0. An excitation transverse to 0 or 'around' it would stay in the minimum and thus not be relevant, this is also called Goldstone boson [13].

(a) Higgs branching ratios
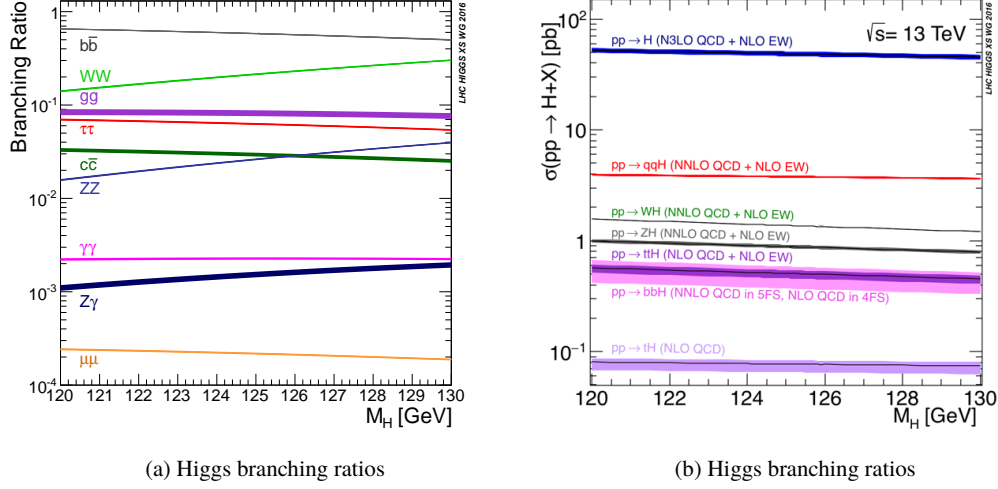(b) Higgs branching ratios

Figure 2.7: Higgs branching ratios and production modes around 125 GeV [19]. The production modes are given for the LHC Run 2 around 125 GeV. *tH* refers to Higgs production in association with a single *t*-quark t-channel and s-channel, but not associated with a *W* boson.

mesons which have the same quark content. Those receive their higher masses due to higher internal angular momentum or relative spin (for example the ρ).

The lowest mass baryons are also built from *u* and *d* and are widely known as the building blocks of our everyday matter: protons and neutrons. The neutron consists of one *u*-quark and two *d*-quarks (electrical charge $\frac{2}{3} - \frac{1}{3} - \frac{1}{3} = 0$) and has a mass of 939.57 MeV [3]. The proton has an even lighter mass of 938.27 MeV [3] and consists of two *u*-quarks and one *d*-quark. The electrical charge $\frac{2}{3} + \frac{2}{3} - \frac{1}{3} = +1$ adds up beautifully to the expected value. All other baryons are built analogously. Heavier baryons are built with different quark contents, relative angular momenta etc. Anti-quarks also form baryons as they have a composite color structure to quarks. These bosons are then called anti-proton, anti-neutron and so on.

However, a few things do not quite add up. Especially the quark masses (calculation based on rough values from [3]):

$$2m(u) + m(d) = 2 \times 2\,\mathrm{MeV} + 5\,\mathrm{MeV} \ll 938\,\mathrm{MeV} = m(p)$$

Remembering, section 2.2.3 made an explicit point of the difficulties of QCD at low momenta (which holds in bound states). Gluons are present to create the bound state, additionally they can couple with themselves, and split into quark-antiquark pairs (the energy available due to the bound states' mass is clearly sufficient for all quarks except the *t*). This gives rise to the interpretation of hadrons being built out of 'valence quarks' and a 'sea' of gluons, quarks and anti-quarks. In this interpretation the valence quarks then dominate the properties of the hadron like electric charge, spin etc. However, the proton mass is then made out of the different quark masses present in the proton all the time, as well as energies attributed to quarks, gluons and interactions [20–22].

This phenomenon is also important when performing collision experiments with protons, where the collision will actually be performed with one of the objects inside a proton, called partons. The cross-sections are just compatible with the pure valence quark view at low four-momentum exchange $Q^2 =$

$(q_1 - q_2)^2$. However, at higher Q^2 the collision can also occur with the 'sea'. In any case, the valence quarks will still carry the highest momentum fractions.

A way to visualize the density of a parton type in the proton $f(x, \mu^2)$ is to plot them at a given momentum fraction x. Any such plot is just valid for a certain Q^2 which is often equivalent to the variable μ^2. (μ^2 describes the 'factorization scale' which is an important ingredient to the calculation of $f(x, \mu^2)$.) To keep everything in a reasonable sized plot often $xf(x, \mu^2)$ is plotted on the y-axis instead. This type of plot is called a 'Parton Distribution Function' or short PDF. The values inside the PDF cannot be calculated from first principles (due to the overall issues with QCD), but need to be adjusted with experimental data.

Example PDFs can be seen in figure 2.8 for two different energy scales. The PDFs used in this thesis are based on the program 'NNPDF' and tuned with ATLAS data obtained from Run 1 [23–26].
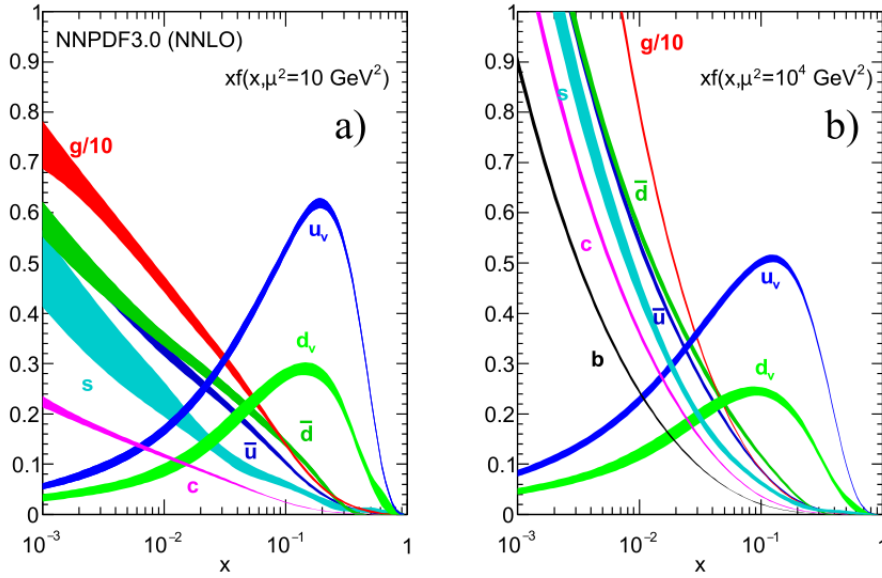
Figure 2.8: Proton PDFs at two different energy scales ($10\,\mathrm{GeV}^2$ on the left and $10.000\,\mathrm{GeV}^2$) obtained with [27] by [28]. Note that the gluon contribution is scaled with $1/10$th to fit on the plot.

2.3 Beyond the standard model

Not every phenomenon observed can be described via the standard model. Most notable the mass terms of the standard model just explain the inert mass of fundamental particles. The heavy mass and the effect of gravitation cannot be described with the standard model.

Another apparent property of the macroscopic world we live in is that it is constructed of matter while anti-matter (for example positrons) is nearly absent. One explanation of this would be an uneven distribution of matter and antimatter in the universe. However, at the boundaries of matter-enriched and anti-matter-enriched areas they would annihilate and we would be able to detect the radiated photons. Another explanation is an unequal treatment of particles and anti-particles. This can actually be observed in experiments, but the extent to which it is detected is currently not enough to explain the phenomenon [29].

There is a third big phenomenon the SM cannot describe: Astrophysics calls for the need of some heavy matter that does not emit photons (like stars do) to describe the movement of astrophysical objects we can observe. This kind of matter is called 'dark matter' and its nature is not part of the SM [3].

Different approaches are taken to extend the SM to describe those phenomena. Some popular ones are for example SUSY (supersymmetry) and string theory models [30].

This, of course is not trivial as they are often reliant on effects which are too small to measure in the particle physics realm (gravity), or at least almost too small (particle-antiparticle asymmetry) or propose particles we have not measured yet (dark matter). Of course searches for additional particles with high mass or a low interaction probability with our 'ordinary' matter are going on. Another approach is to measure SM processes and look for deviations from the expectation that arise from deviations of the SM Lagrangian density.

In order to organize the search one strategy is to test every single possible configuration by itself. This would lead to a multitude of similar tests that need to be done. To circumvent this, a popular approach is to add all allowed additional higher order terms to the SM Lagrangian density, called effective field theory EFT. Each of these terms is assigned an operator which would be 0 in the SM. Those can be tested independently of the (numerous) alternative physics models, but not independently of the correlations between the operators. Different bases are used to describe the numerous operators. A popular one to describe processes connected with the H is the Warsaw basis [31]. Figure 2.9 shows one overview of the current state of the measurement of different EFT operators most relevant to the process in this thesis. Up to now no significant deviation from the SM has been found [32–34].
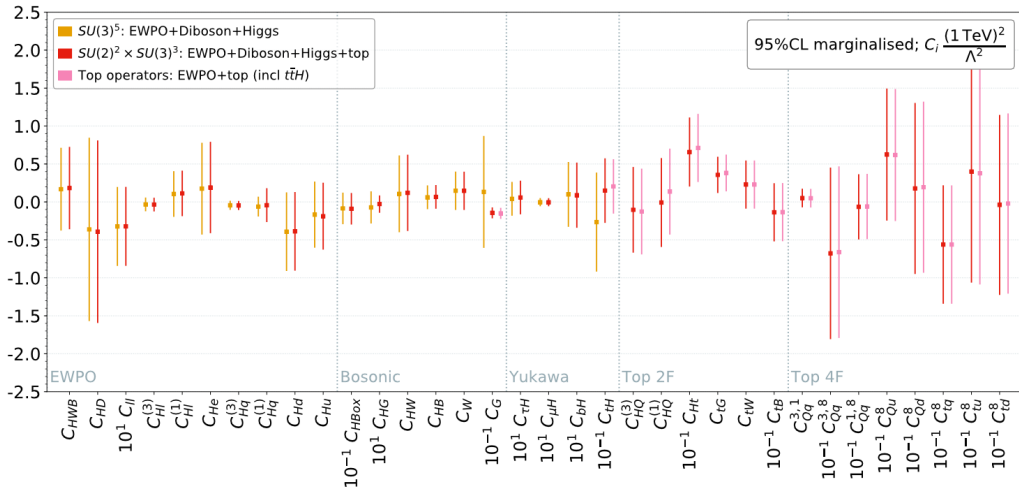
Figure 2.9: EFT parameters fitted to LHC results [34]. The small deviations from 0 are not significant.

Signal and background processes

All that is gold does not glitter,
Not all those who wander are lost;
The old that is strong does not wither,
Deep roots are not reached by the frost.
        – J. R. R. Tolkin, *The Lord of the Rings, The Fellowship of the Ring*

This chapter describes the most important processes for the analysis, starting with the target process *tHq* in section 3.1. In the following similar processes are described as well as other processes which might be confused with the target, also called background processes.

3.1 *tHq* process

The *tHq* process is the subject of this thesis. It is characterized by containing a single top quark, a *H*-Boson and having a t-channel structure (which leads to the additional quark in the final state). As this thesis is using ATLAS data, the initial state particles are constituents of two protons (more on this in section 5.1). The central Feynman diagrams that are used to describe this process are shown in figure 3.1.
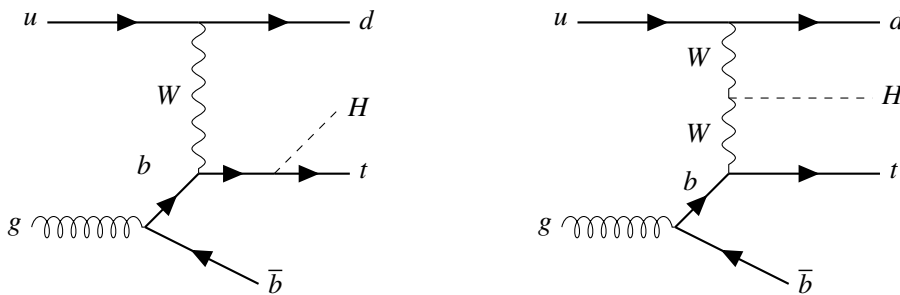
Figure 3.1: Feynman diagrams for the *tHq* process in the 4FS (see section 3.1.1). The *H* is associated with different objects: *t* or *W*.

To digest this process step by step, first the focus lies on the underlying t-channel process with a single top quark in the final state (see figure 3.2) and the consequences for the final state particles. As there is

a top quark in the final state but not in the initial state the flavor has to be changed and the only boson which allows such an interaction is the *W*. Thus, both the upper and lower part consist of two quarks with different flavor.
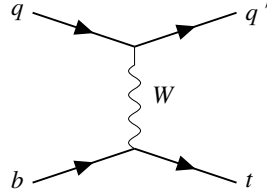
Figure 3.2: t-channel single top production in the 5FS.

3.1.1 4 and 5 flavor scheme

The most probable vertex that produces a top quark has a b-quark in the initial state. *b*-quarks are part of protons as sea-quarks and thus equally common (or uncommon) as quark and anti-quarks. An alternative way to create a *b*-quark is assuming it came from a gluon splitting (see figure 3.3). Calculation-wise one can assume that this is the only process to create *b*-quarks from a proton and thus a proton just contains four different flavors of quarks. This is called the 4-flavor-scheme (4FS) and stands in contrast to the 5-flavor-scheme (5FS, *b*-quarks emerging directly from the sea). The 'additional' *b*-quark in the 4-flavor-scheme has no influence on the measurement, as it is mostly carried away with the rest of the proton remnants and in both schemes b and $\bar{b}$ are equally likely. However, the calculation yields slightly different results for both the overall probability as well as angular and momentum distributions. The used *tHq* samples (appendix A) utilize the 4-flavor-scheme as the results have been found to be more similar to data [35]. Therefore, also the cross-section used for prediction is using the 4FS. However, for cross-section calculations a combination of 4FS and 5FS has been found more accurate [36, 37]. A comparison of the resulting cross-sections can be found in section 3.1.3.[1]

Figure 3.3: t-channel single top production in the 4FS (compare to figure 3.2).

16

(a) t-channel production with an *u* in the initial state.

(b) t-channel production with an $\bar{u}$ in the initial state.

(c) t-channel production with a *d* in the initial state.

(d) t-channel production with a $\bar{d}$ in the initial state.
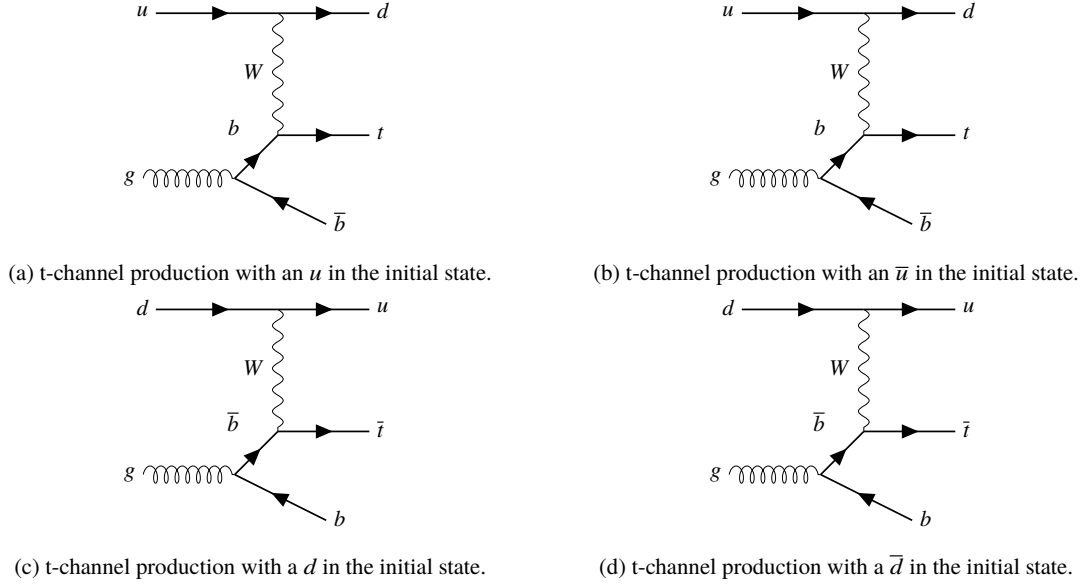
Figure 3.4: The four different charge configurations of t-channel single top production (depicted in the 4FS).

3.1.2 Top quark charge

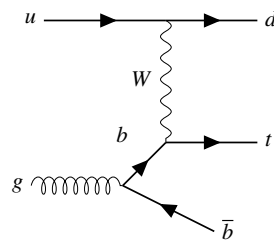The fractions of *t*- and $\bar{t}$-quarks in the final state is determined by the charge of the incoming quarks. As described above *b* and $\bar{b}$ are equally likely, therefore, the ratio is solely defined by the lighter quarks. In figure 3.4 one can see the Feynman diagrams for single top t-channel processes for up- and down-type quarks and anti-quarks as well as the resulting top-type. Combined with the information of the PDFs (see section 2.2.5) one can calculate the fractions of *t* and $\bar{t}$ quarks as seen in table 3.1. As the proton content at low Q^2 is dominated by the valence quarks the ratio is roughly $\frac{2}{3}$.

quark-type	number of valence quarks	*t*	*t* probability	$\bar{t}$ probability
u	2	*t*	x	
$\bar{u}$	0	$\bar{t}$		x
d	1	$\bar{t}$		x
$\bar{d}$	0	*t*	x	
total	3		0.64%	0.36%

Table 3.1: Possible initial state quarks and the resulting *t* and $\bar{t}$ ratios calculated from [38]. The first generation is a placeholder for the first two generations, the third generation does not impact this calculation.

In the following *t* will refer to both the particle and antiparticle if not specified otherwise. The same goes for the other particles in the process.

---

[1] For some non-signal processes described in section 3.4 a distinction is also necessary: for single-top t-channel processes the 4FS is used, while for $t\bar{t}$, tW and tW-like processes the 5FS is used, since their interference can just be accurately described in this scheme.

3.1.3 Associated *H*-Boson

As described in section 2.2.1 the *H*-Boson couples to all particles with mass, and the heavier the mass the higher the coupling strength. In the aforementioned t-channel process the heaviest particles are (in order): *t*, *W*, *b*, and light quarks. Each coupling can be described in a separate Feynman diagram. From these the two most likely (coupling to *t* and *W*) are dominating the calculation, and we will focus on those two. Both Feynman diagrams can be found in figure 3.1. As described in section 2.2.1, when two diagrams are contributing to the same final state they can interfere positively or negatively. In the case of *tHq* production the negative interference between both diagrams nearly leads to an obliteration of the cross-section. The final cross-section of $\sigma_{SM} = 71$ fb [37] is four orders of magnitudes lower than the most likely *t* production mode [39]. While this certainly does not make the analysis any easier it provides an unique opportunity to probe the SM.

In more detail the cross-section with uncertainties is[37]:

$$\sigma(tHq) = 71.3 \, \text{fb} \, (1)(\text{stat}) \pm^{+7.2\%}_{-10.9\%} (\text{scale}) \pm^{+0.5\%}_{-0.5\%} (\text{PDFs}) \tag{3.1}$$

This calculation is done at next-to-leading order. The central value is taken from the 5FS, but the uncertainties are obtained from the envelope of the 4FS and 5FS calculation. In comparison, the central value of the 4FS would be 68.1 fb which corresponds to a 4.5% difference of the central values.

In BSM scenarios where the couplings between *H* and bosons and fermions (in this case k_t and k_W) are altered the cross-section of *tHq* is enhanced as can be seen in figure 3.5. Thus, by probing the *tHq* cross-section, the possible BSM Higgs couplings can be indirectly constrained. [36, 40–43]. In case of EFT benchmark models the *tHq* cross-section might be significantly enhanced [43, 44].

$$p \, p \rightarrow t \, h \, j \quad (\text{LHC 14 TeV})$$
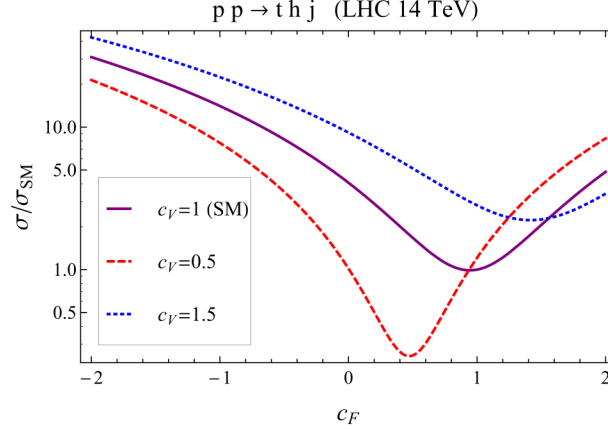
Figure 3.5: *tHq* cross-section in dependence of different k_t and k_W values(here noted as c_F and c_V) [42]. Note that this calculation was done for the LHC running at 14 TeV, the results calculated at 13 TeV by [36] are qualitatively comparable.

An important aspect of the *H* boson preferred coupling to heavier objects is that it also prefers to decay into heavier objects (see also section 2.2.4). The main Higgs decay channels and resulting *tHq* cross-sections can be found in table 3.2. As this analysis targets events with τ_{had} the most important decay channels are into a τ, *Z* or *W* pair as the *Z* and *W* may decay into τ (a further discussion can be found in section 5.4).

H decay	BR	tHq cross-section in [fb]
$H \to b\,\bar{b}$	5.82×10^{-1}	41.5
$H \to W^+\,W^-$	2.14×10^{-1}	15.3
$H \to \tau^-\,\tau^+$	6.27×10^{-2}	4.5
$H \to c\,\bar{c}$	2.89×10^{-2}	2.1
$H \to Z\,Z$	2.62×10^{-2}	1.9
$H \to \gamma\,\gamma$	2.27×10^{-3}	0.2

Table 3.2: Most important Higgs decay channels, their relative branching ratios from [3] and the resulting expected tHq cross-sections.

3.2 Single top + Higgs boson processes

There are three processes, which at LO can create a single top quark: t-channel, s-channel and single top quark production in association with a W (tW) (see figure 3.6 for the additional Feynman diagrams). All three of them can radiate off an additional H and all of them are affected by the negative interference effect. One could argue that measuring all three of them together might give better information on the H couplings. However, the predicted cross-sections of tWH ($\sigma_{SM} = 19$ fb) and s-channel+H ($\sigma_{SM} = 0.6$ fb) [19] are much smaller than the tHq cross-section. Additionally, the final states differ between all three processes and hence targeting the other two processes would make the analysis less specific. Therefore, focussing on tHq as signal allows for a well-defined target process.

(a) tWH (b) s-channel single top production in association with a H.
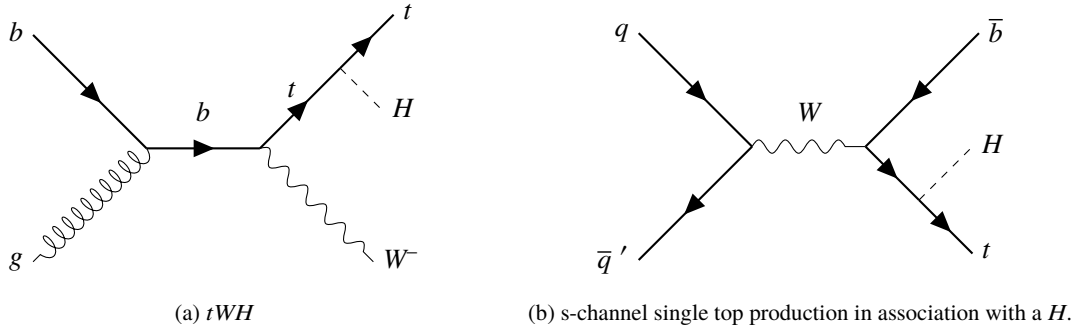
Figure 3.6: Other single top production modes with an associated H in the 5FS.

3.3 tZq process

tZq is similar to tHq as it shares the t-channel single top structure. The Z has different decay channels from the H. However, it can decay into a τ pair while the decay channels to W and Z pairs do not exist. Additional differences are the more equal coupling of the Z to the t-channel components and a cross-section which is about one order of magnitude higher than that of the tHq process. The tZq cross-section has been predicted to be 805 fb [37] and was already measured in ATLAS and CMS in [45, 46].

A different way of looking at the tZq process would also include off-shell Z/γ^*. In this case one speaks of the $t\ell\ell q$ process. Previous measurements and theory calculations have focussed on measuring $t\ell\ell q$

instead of *tZq*. Accordingly, the used background samples are of *tℓℓq* type. However, extra care has been taken to ensure events with an *H* in the off-shell contribution are excluded.
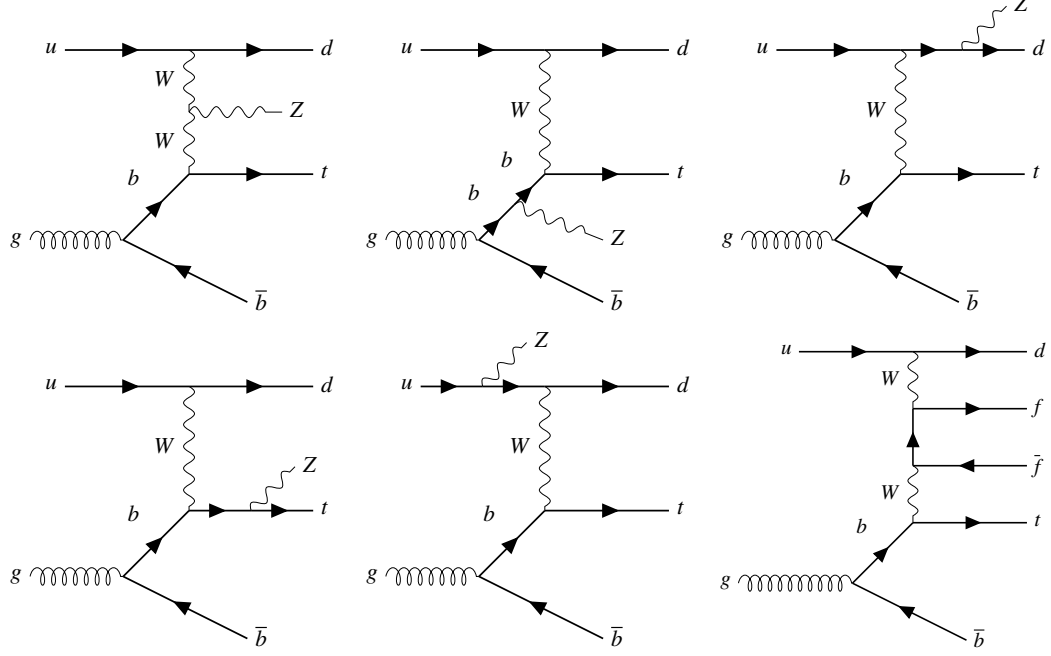
Figure 3.7: LO Feynman diagrams for the *tℓℓq* process in the 4FS. The *Z* is associated with different objects: *W*, *t*, *u* and *b*.

3.4 Other background processes

Background processes can be split into reducible and irreducible backgrounds. Reducible backgrounds have a different final state (focussing on the number of leptons) and thus can be reduced by improving the ability to recognize those objects. However, important reducible backgrounds in this analysis have a much higher cross-section than the target process. Therefore, some events are misreconstructed and end up looking similar to our target process. This analysis is dominated by reducible backgrounds. The important ones are $t\bar{t}$, *Z*+jets and *W*+jets (see figure 3.8). Smaller reducible background processes are *tW*[2] and s-channel single-top production (ordered by SM cross-section). The associated LO Feynman diagrams can be found in figure 3.8.

If a background has the same or higher number of objects in the final state, it is called irreducible. For *tHq*, the already discussed *tZq*,*tWH* and s-channel single-top production with an additional *H* are

---

[2] At NLO some processes cannot be differentiated via the final and initial state contributions, for example $t\bar{t}$ and *tW*. Therefore, both diagrams interfere with each other when calculating the ME. Different schemes are used to deal with this: Diagram Subtraction (DS) where the $t\bar{t}$ terms are subtracted after calculation, Diagram Removal without interference (DR1) where diagrams are removed before calculation including the interference terms and Diagram Removal with interference (DR2) where diagrams are removed before calculations, but interference terms are left in. The nominal samples of *tW* use DR1 and the same scheme is used for *tWH* and *tWZ* to remove the contribution from $t\bar{t}H$ and $t\bar{t}Z$ respectively [47, 48]. Due to this effect it is necessary that the affected sample cross-sections are calculated in the 5FS.
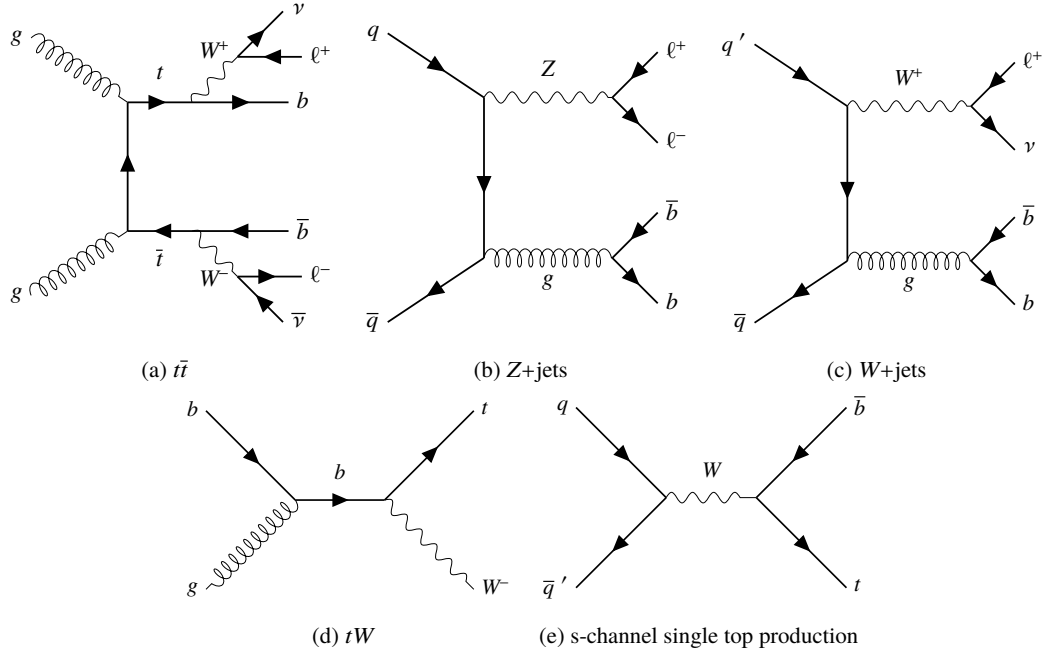
Figure 3.8: Feynman diagrams for reducible background processes. Note that for every process just one example diagram is shown. If possible the diagram that is most likely to appear as a background in the analysis.

irreducible backgrounds as well as $t\bar{t}W$, $t\bar{t}H$, $t\bar{t}Z$, diboson, triboson, 3-top production, 4-top production and tWZ and other Higgs production modes[3] see figure 3.9, figure 3.10 and figure 3.11. (The production of 3 top-quarks does not have a contribution at similar order as the other processes; therefore, it is not shown.) It is important to note that those processes might also have reducible components (for example depending on the H decay mode). An additional distinction can be made between processes with at least one b-quark in the final state figure 3.9, figure 3.10 and without any figure 3.11 with a similar argument as for the lepton requirements. However, as the 'jet-tagging' (see section 5.4.5) is not as reliable as lepton reconstruction, processes with a different number of expected of b-quarks in the final state are counted as irreducible.

---

[3] The number of leptons from the Higgs decay varies depending on the decay mode. Therefore, it can be irreducible.

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

(d) triboson one example process   (e) tZq

Figure 3.9: Leading-order Feynman diagrams for irreducible background processes with at least one b-quark in the final state. To keep the number of diagrams reasonable not the processes are not shown in all possible configurations. Additionally, heavier particle decays are just shown for a few processes to keep the diagrams readable.

(a) example 4top production process

(b) *tWH*

(c) s-channel+Higgs

(d) *tWZ*
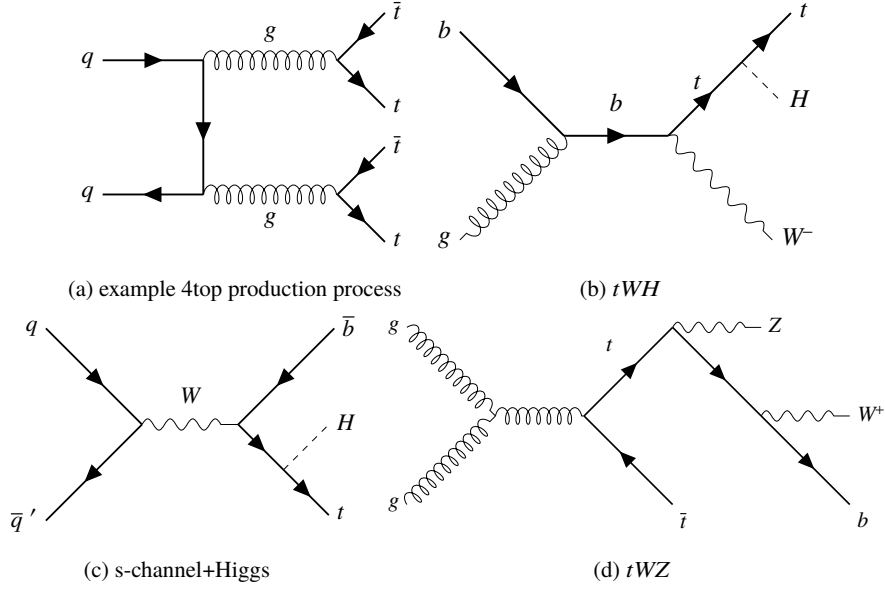
Figure 3.10: Some more Feynman diagrams for irreducible background processes with at least one *b*-quark in the final state. To keep the number of diagrams reasonable not the processes are not shown in all possible configurations. Additionally, heavier particle decays are just shown for a few processes to keep the diagrams readable.

(a) *WZ*

(b) *ZZ*

(c) Gluon-gluon-fusion (ggf) *H* production
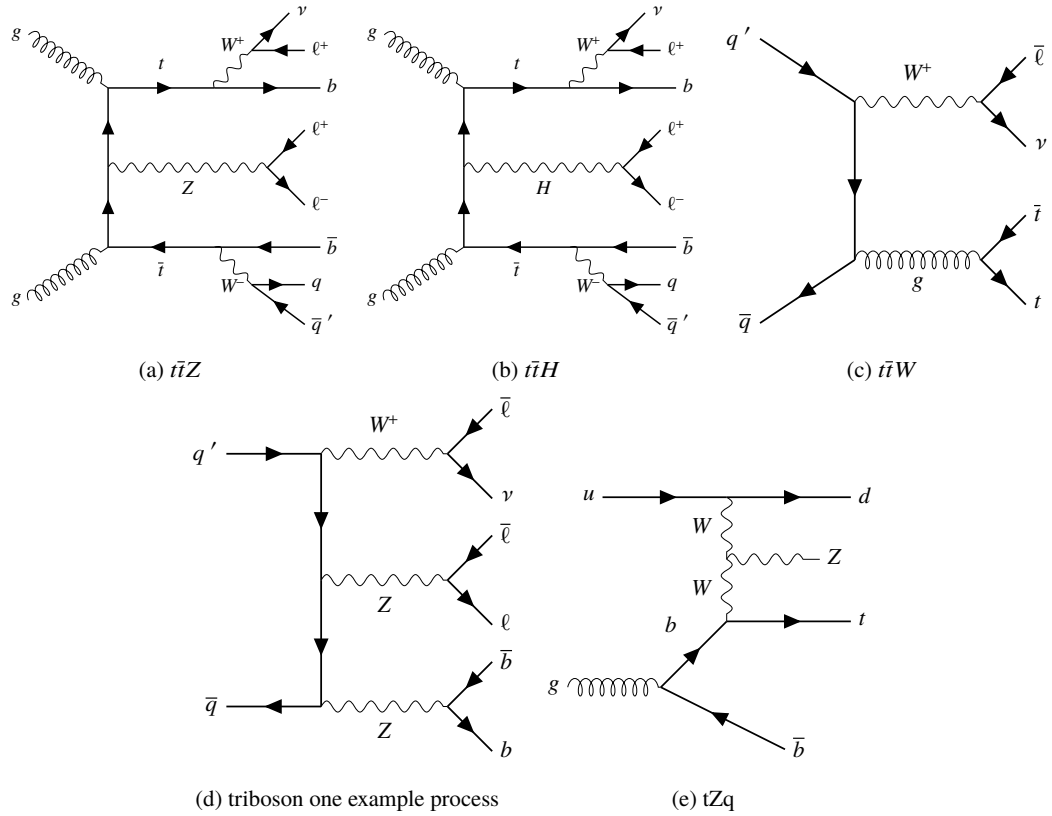
(d) Vector-boson-fusion (VBF) *H* production

(e) *WH*

(f) *ZH*
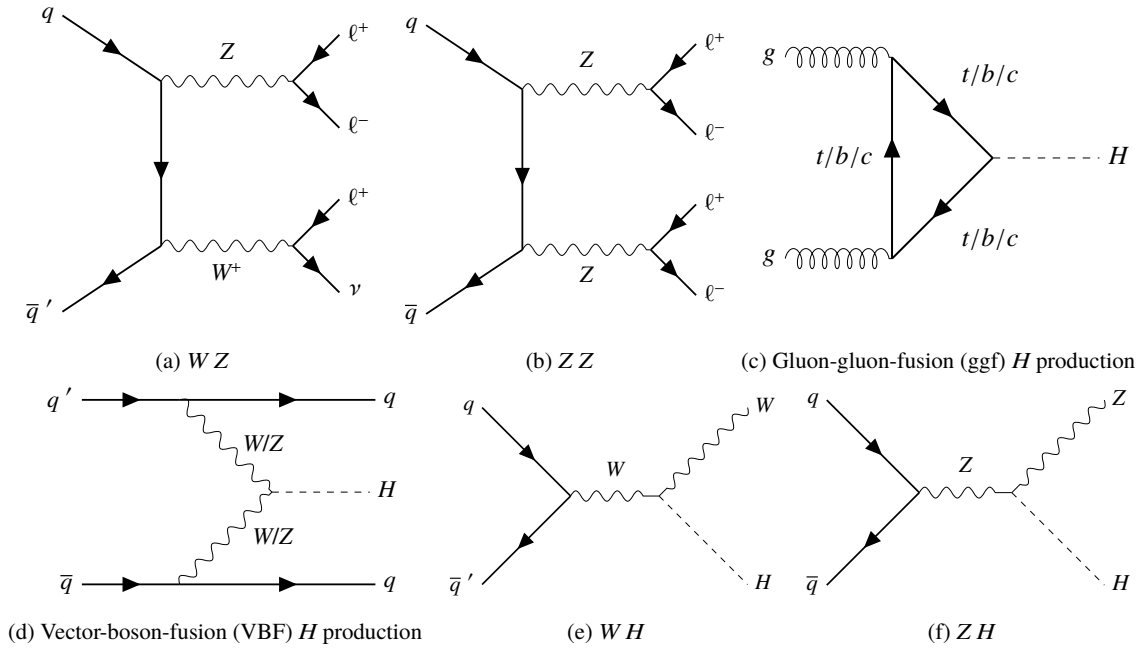
Figure 3.11: Feynman diagrams for irreducible background processes with no *b*-quark in the final state. To keep the number of diagrams reasonable not the processes are not shown in all possible configurations. Additionally, heavier particle decays are just shown for a few processes to keep the diagrams readable.

23

3.5 Previous searches for tHq

The tHq process is the main topic of this thesis. There have been previous searches targeting this process; notably the CMS collaboration published several searches. This includes two searches for multiple final states [49, 50]. Most recently a search for $t\bar{t}H$ and tHq with multiple leptons (including τ_{had}) in the final state [51] has been performed. The ATLAS collaboration published two searches for $t\bar{t}H$ and tHq when the H decays into photons [52, 53].

None of these searches has observed the tHq process, but upper limits (see section 4.4) have been set. The ATLAS result limits the tHq cross-section to 12 times the SM (at 95% confidence level) [53], the most sensitive quoted limit by CMS is 25 times the SM result [50]. However, the most recent CMS publication [51] does not quote a limit on the cross-section. Instead, the best observed fit result is given as:

$$\mu(tHq) = 5.7 \pm 2.7(\mathrm{stat}) \pm 3.0(\mathrm{syst}) \tag{3.2}$$

Statistical methods

A moment in the mind is worth nine in the fire.

– Patrick Rothfuss, *The Name of the Wind*

4.1 Monte Carlo simulation

In order to compare the SM expectation with the measured data one needs a basis for comparison. This should be an as accurate description as possible of what the data is expected to look like under the assumption that the SM describes physics completely. The technique used for this is called simulation, more accurately Monte Carlo simulation (MC). Monte Carlo techniques is a collective term for any calculation that utilizes (pseudo-)random generated numbers to some extent.

The process of MC simulation relies on calculating example solutions that in sum will converge to the analytic complete solution [54] (the easiest case to imagine this is the calculation of an integral). This process is called sampling. While it does not follow the path of an analytical solution, sampling may allow for a faster convergence, especially if one is utilizing weighting. In this case, every calculated result is assigned a number (weight) which reflects how strong the result should impact the final result. The sum is then performed as a weighted sum, while ensuring the final result is still correct. In addition to faster calculation of a combined result, MC simulation also provides singular correct examples of valid outcomes. This is on the one hand useful to study those events in more detail, on the other hand it allows several simulation steps to be combined. For example, a MC simulation of the *tHq* process will converge towards the correct cross-section, but also give valid example events with a specific momentum for each final state particle. Those events can then be used as an input for a MC simulation of the measurement procedure.

The simulations used in this thesis are separated into different steps which resemble the evolution of an event over time. However, the strict division between steps is artificial and the transition has to be done carefully. The steps can be divided like this:

1. calculation of the matrix element (ME) for the hard process
2. forming (semi-)stable particles (showering)

   - (simulation of soft processes)
   - decay of heavy particles (like *H* and *t*)
   - parton shower: including radiation in either final or initial state (FSR and ISR)

- underlying event: multiple parton interactions that overlap with the hard process (MPI)
- beam remnants
- hadronization: the showered variables form hadrons
- decays of hadrons and τ

3. detector simulation

For every step different simulation programs can be used, for example:

- **Madgraph:** a standard tool to calculate the matrix element of different processes at LO or NLO [55, 56].

- **Powheg:** a tool to calculate the ME at NLO which takes into account the parton showering and is able to provide events with only positive weights (more on negative weights below) by default used for every process where it is available [57–59].

- **Madspin:** used to efficiently calculate heavy particle decays with correct spin correlations [60, 61], interfaced for some t decays.

- **Pythia:** a multipurpose generator that allows to calculate the matrix element at LO and to simulate the showering [62]. Hadronization in Pythia is done with a string fragmentation model [63, 64]. Pythia is used as the nominal generator for everything but the purely bosonic samples in combination with Powheg or Madgraph. In addition, it is used to produce the overlapping MPI for all samples [65–67].

- **Herwig:** a multipurpose generator similar to Pythia [68], but using the cluster model for hadronization [69]. Therefore, it is used as a systematic variation to Pythia in combination with Powheg or Madgraph.

- **Sherpa:** another multiple purpose generator which does include calculations of the hard process at NLO [70]. It is used as a nominal generator for samples with mainly W and Z in the final state [71, 72] and as a systematic variation.

- **Geant4:** is able to simulate the passing of particles through matter in detail [73–75]. It is used for the detector simulation per default.

- **ATLAS fast 2 (AFII):** a software that is designed to speed up the event generation process. Therefore, most of the detector is simulated with Geant4, but the calorimeter interactions are approximated by parametrizing the shower of a photon (for electrons or photons) or a charged π (for all hadrons). This speeds up the simulation by an order of magnitude [76–78]. It is used for samples which are associated with not (yet) discovered processes (including *tHq*).

Every generated sample is associated with a dataset ID (DSID) to ensure a unique assignment. Some processes are split into several sub-processes to limit the individual sample size. Therefore, some processes are associated to several DSIDs. An overview which generator has been used for which sample can be found in appendix A.

Weights are crucial for a faster convergence of MC simulations. It has already been mentioned that some weights can even be negative. In particle physics simulation this is mostly utilized in order to correct for the overlap between NLO calculation and parton showering. The downside of this is that fully valid events are generated, which actually decrease the available statistics. The Powheg method is able to generate samples without any negative weights by dealing with the divergencies of each hard process separately [57, 58]. Unfortunately this method is not available for t-channel processes with an associated boson (as *tHq*

and *tZq*). In small quantities negative weights are not a big nuisance to the overall simulation, but in the *tHq* signal sample the fraction of negative weights is about 30%! (It is even 40% for the *tZq* sample.)

To calculate the influence negative weights have on the available statistics, the spread of a distribution (σ) is used which is associated with the possible fluctuations from that sample. For distributions following the Poisson statistic (for example in counting experiments) $\sigma = \sqrt{N}$ holds. For MC simulation we are interested in a method's relative error with respect to the generated events [79]:

$$\frac{\sigma(N)}{N} = \frac{\sqrt{\sum_{i=1}^{N} w_i^2}}{\sum_{i=1}^{N} w_i} \tag{4.1}$$

The w_i describe the individual weights. Since the weights in the nominator are squared, the nominator is independent of the samples' population with negatively weighted events. The denominator, however, shrinks when our sample contains negatively weighted events. This way we are not only simulating a lower effective number of events $N_{\text{eff}} < N$ but also keep the spread the same. If we look at the simplified case of events being just either positively weighted ($w_i = 1$) or negatively weighted ($w_i = -1$) with a fraction x of negatively weighted events we get:

$$N_{\text{eff}} = N_+ - N_- = (1 - x)N - xN = (1 - 2x)N \tag{4.2}$$

$$\frac{\sigma(N_{\text{eff}})}{N_{\text{eff}}} = \frac{\sqrt{N}}{(1 - 2x)N} \tag{4.3}$$

$$< \frac{\sigma(N)}{N} = \frac{\sqrt{N}}{N} \tag{4.4}$$

where the compared relative error describes a potential method without any negative weights. For a case of $x = 0.3$ (as for our signal sample)

$$\frac{\sigma(N_{\text{eff}})}{N_{\text{eff}}} = 2.5 \frac{\sigma(N)}{N} \tag{4.5}$$

which means the relative error rises with a factor of 2.5 relative to a sample with the same amount of generated events but no negative weights.

An additional complication arises when a sample is not modeled with enough events in the region of interest. Especially if the region is further divided into bins, the weighted sum of events in this region might be negative. This is in general not physical (an exception might be a negative interference between samples). Therefore, negative contributions in those bins are set to 0. The ideal solution would be to generate more events in the region of interest, but this is not always feasible.

4.2 Multivariate analysis methods

One way to utilize the large number of simulated events we gain from Monte Carlo generated samples is to train Multivariate Analysis tools (MVA). This refers to any tool which is able to analyze several variables at the same time. In this thesis MVAs are used to take decisions on whether an event is signal-like, or background-like. The types of MVAs used are Boosted Decision Trees (BDT) and Neural Networks (NN).

4.2.1 Boosted decision trees

Decision trees are commonly used to separate a large group of events into different subcategories. Taking the example of two categories signal and background the tool would determine the best cut. For that it has to find the variable and the optimal cut which is most powerful to separate those two. This way two subsets of events are created which pass or not pass the cut (a decision). On each of the subsets a new decision is made and this continues until a certain limit is reached. This way a tree of decisions is created. The emerging structure can be applied to additional events, for example test events. These can just undergo the already defined cuts to determine how likely they are to be classified as signal or background. One set of decisions is called a layer, each decision made can be called a node. A schematic view of a BDT with two layers can be seen in figure 4.1. In the final row we see four distinct datasets (leaves) which are subsets of the original distribution and can be ordered as "signal-like" and "background-like" (their sizes might differ, which is not reflected in the schematic view).
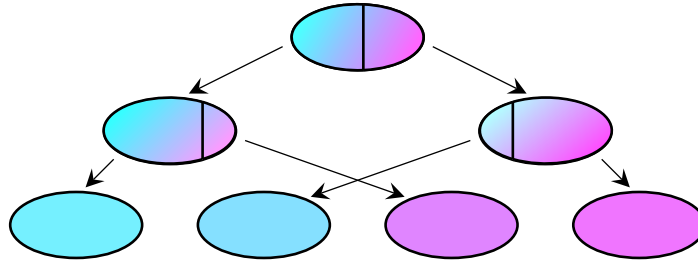

Figure 4.1: Schematic view of a BDT with depth 2. Two decisions are made on the initial distribution, resulting in 4 distinct samples that contain different fractions of signal (red) and background (blue).

One could design a decision tree which has a high number of layers to obtain a perfect decision for the original set of events (training sample). However, such deep decision trees are prone to being sensitive to random fluctuations in the training sample (overfitting). Applied to a new set the classification performance might drop drastically. This is often tested by comparing the performance on the training sample with the performance on a distinct test sample. In order to avoid overfitting, BDTs create a series of decision trees which themselves are not strong enough to figure out the optimal separation, however every decision tree is trying to correct the previous results and thus a good separation can be achieved without the high risk of overfitting, this process is called boosting.

One popular variant of boosted BDTs are gradient BDTs. They are guided by a loss-function, which is the measure for an MVA method to determine how far away from the goal it is. In our example the most straightforward loss-function for a set of events $\vec{x}$ which are either signal or background would be to compare every given prediction to the true value and add the differences up for the whole set:

$$L = - \sum_x \Delta(x_i^{\text{true}}, x_i^{\text{pred}}) \tag{4.6}$$

Loss functions are designed to be higher if the wrong classification (loss) is more common. This makes it easy to visualize that one needs to follow the gradient of the loss to find a minimum where the network performs best[1] to separate signal and background. This optimization process lies at the heart of several MVA methods.

---

[1] As good as reasonably achievable, finding the global minimum in such a large parameter space is not feasible.

Gradient BDTs operate on a differentiable loss-functions. Each tree is build to have the biggest possible improvement over its predecessor by fitting its loss-function's gradient. From this a score can be calculated for any event which tells us how signal-like an event is. Often a score of 1 means very signal-like, a score of 0 not signal-like. [80, 81]

4.2.2 Neural Networks

Neural Networks (NN) are also based on the idea of combining different small steps to achieve a goal, for example a classification. They also work on 'nodes' in different layers, but in contrast to a boosted decision tree a NN does not split the original dataset. Instead, each node in a NN scores the events. This way it keeps each event throughout the network and in the build of the network the information of all events are combined. In more detail, the network starts with all available information for a specific event and each node in the subsequent layer combines *all* outputs that have been produced in the previous layer. Each connection (edge) might have a different importance, represented by the weight of that connection.[2] The result of the linear combination is taken as input for a *nonlinear* activation function to produce an output in a predefined range (often between 0 and 1). This is performed for a number of layers, each layer may have a different number of nodes than the other layers. The last layer often consists of a single node that outputs a smooth distribution between 0 and 1 which relates to the desired classification (1 means the output is very signal-like, 0 the output is very background-like). A depiction of a small NN with just three input distributions and one internal layer can be seen in figure 4.2. Some NNs used in this thesis do contain three output nodes, each related to a different set of signal (categorical NN). The sum of the three outputs has to be 1, so the network has one additional degree of freedom for the classification, which can be helpful if either the signal or the background set (as in our case) is in itself heterogeneous.
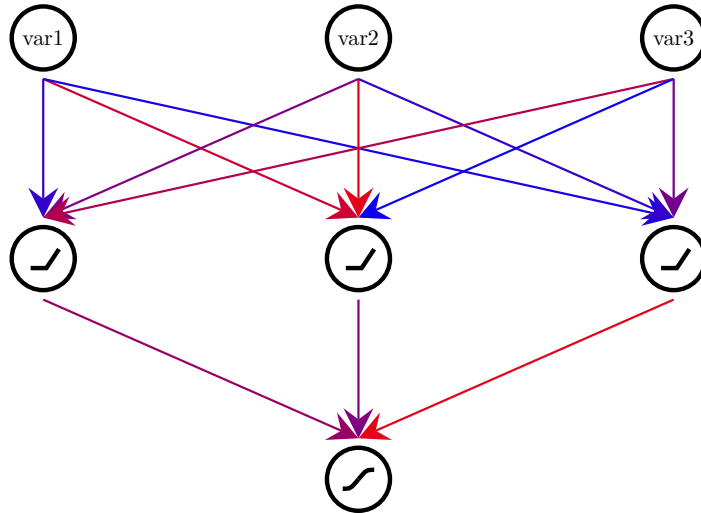
Figure 4.2: Schematic view of a simple NN. The first line of nodes represent the input variables. With the symbol in the subsequent nodes the activation functions are indicated. The different colors of the arrows indicate the different weights of the edges.

---

[2] The described architecture is also called a fully connected feed-forward NN, other designs as for example Recurrent Neural Networks (RNN) do not rely on a strict hierarchy of nodes, but may also connect the output of a node with the input of a node in a previous layer, the same layer or even the same node.

The optimization of a NN works a bit different from that of a BDT, although the decision for both are based on a loss-function. Instead of making new decisions after each other (or even new trees) one goes back and changes the weights in the network, this is called backpropagation [82]. To achieve this, the partial derivative of the end result as a function of all weights can be computed (weights related to 'deeper' connections in the network do depend on the weights towards the final decision). This gradient allows optimizer algorithms to find connections between the weights and the loss-function. Therefore, the weights are updated according to this gradient. Some optimizers (for example Adam) even take the second derivative into account, which results in a faster and more robust optimization [83].

In table 4.1 one can find a list of commonly adjusted properties of NNs, often called hyperparameters. In principle everything from the number of layers to the functions by each node can be altered. Additionally, the groups in which events are used for the different steps (like optimization and testing) can be adjusted. The adjustment of the hyperparameters is non-trivial as each problem requires the NN to be built separately. Additionally, the high degree of nonlinearity makes it very hard to predict how the optimal NN would look like. However, the more complicated a problem is, the more layers and nodes are usually needed to complete the given task.

Hyperparameter	description	example
loss-function	see text	crossentropy
optimization	see text	Adam
hidden layers	number of layers of nodes without the input and output notes	6
nodes	number of nodes in each hidden layer (can be the same or changing for each layer)	96
activation function	nonlinear weighting function for each node	relu
output activation function	activation function of the final nodes, often chosen to be smooth	sigmoid
weight initialisation	initial value of weights, starting point of the NN	lecun normalisation
dropout	ignoring a fraction of nodes in each training step, helps to avoid overtraining	20%
training epochs	optimization cycles	100
learning rate	sensitivity of the optimizer	0.001
batches	number of subsets used for training steps	16
k-folds	number of individual trainings, in each training another subset is used for testing, checks the homogeneity of the data	5

Table 4.1: Different parameters that can be optimized in the design of a NN.

4.2.3 Negative weights in MVAs

One challenge for Multivariate Analysis Tools are negative weights, especially in the quantity contained in the signal samples. There is no proposed optimal solution for this, so strategies have to be tested on a case by case basis. Given that each (negatively or positively) weighted event in particle physics simulations is representing a physical final state by itself, it could be possible to treat positively and negatively weighted events differently. Possible methods are not using negatively weighted events in the training or flipping

the sign for negatively weighted events and treating them as if they were positively weighted. However, this approach has to be applied carefully as even though negatively weighted events are physical by themselves the distribution of physical quantities might be different in the positively weighted/negatively weighted/combined set of events.

4.3 Binned profile likelihood fit

In order to determine if a process has been found or not it is necessary to compare the measured data with the simulated datasets (section 4.1). However, not every disagreement has to be linked to a discovery of the signal process. On the one hand every distribution has some statistical fluctuations, on the other hand the simulated datasets might not be completely accurate. A binned profile likelihood fit can be used to compute the signal strength or the exclusion limit for the given signal in the measured dataset (*data*) given the simulated datasets (*mc*) for signal (*sig*) and background samples (*bkg*, treated here as one sample). This section explains these fits from a users' perspective. First, the simplest setup is explained and further concepts, like multiple regions and nuisance parameters, are added when the need arises.

The first step in applying the profile likelihood fit is to select the events which should go into the fit. Of central interest are the events in the signal enriched region (SR). Ideally, this region is dominated by signal events. This way any excess in data in comparison to the background-only hypothesis can be linked to the signal strength μ:

$$n_{data} = n_{bkg} + \mu \, n_{sig} \tag{4.7}$$

For rare processes this is either not possible, or the statistical uncertainties (see section 4.1) would dominate the calculation. In these cases the SR can be subdivided into different bins, for example according to a discriminating variable like an MVA output. The distribution in each bin can be described by a Poisson distribution for both signal and background events separately:

$$f(n_{sig}|\mu, \theta) \tag{4.8}$$

$$f(n_{bkg}|\theta) \tag{4.9}$$

Both distributions can additionally depend on a variable (or a vector of variables) θ, called nuisance parameters (NP). When one wants to describe the probability of μ for a given (measured) n this is described by a likelihood, which again depends on all NPs:

$$L(\mu, \theta) = \prod_l L(\mu, \theta_l) \tag{4.10}$$

The global maximum of $L(\mu, \theta)$ is found at $\hat{\mu}$ and $\hat{\theta}$, which can be used as estimators for the true values. Since minimizing is more convenient one usually uses $-2 \ln (L(\mu, \theta))$ instead. In some cases (for example limit setting) it is beneficial to use a ratio of likelihoods. Both alternatives do not impact the statement made before [84].

An example distribution of $-2 \ln (L(\mu, \theta))$ as a function of μ can be seen in figure 4.3. The distribution is obtained from a fit which treats *mc* as *data*, called Asimov fit. In this case the best agreement between '*data*' and *mc* is given at $\mu = 1$, which can also be seen in the minimum of the distribution.

Apart from finding a valid estimate for μ and θ, it is important to know how secure that estimate is. This variance is given by the 1σ variation. For a Gaussian distribution this corresponds to the interval
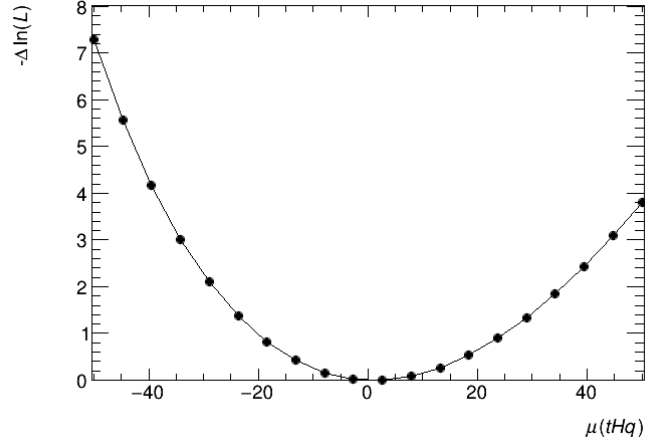
Figure 4.3: An example distribution of $-2\ln\left(L(\mu, \theta)\right)$ as a function of μ for an Asimov fit with no additional NPs.

around the mean which holds 68% of the events. In case of an asymmetric distribution this is reinterpreted as 16% of events lying on the left and right of the interval. However, not all minimization algorithms are able to calculate asymmetric errors. Migrad for example calculates σ from matrix inversion and such always arrives at a symmetric error distribution. Minos on the other hand tests the distribution and such is able to discover asymmetric errors [85]. Unfortunately, the Minos-approach takes more computation time and as such is only used for parameters which are expected to be asymmetric and important (like μ). For a successful Minos-calculation the allowed parameter space needs to be big enough to not bias the calculation.

Another complication is the presence of θ. To find the minimum of $L(\mu, \theta)$ 'simultaneously' for μ and θ one uses the profiling technique: to eliminate the explicit dependency on θ, the value of θ which maximizes $L(\mu, \theta)$ for a given μ called $\hat{\hat{\theta}}(\mu)$ is used. This is also called the profiled value of θ. The function which is minimized is then called $\lambda(\mu)$ [86]:

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\theta}})}{L(\hat{\mu}, \hat{\theta})} \tag{4.11}$$

This is especially necessary in case of a multitude of NPs (for a list and type of NPs used in this thesis see appendix B). Note that the inclusion of NPs takes into account the presence of additional uncertainties and thus increases the uncertainty on μ. Also, μ and θ variables are treated differently, although they are not necessarily inherently different. One difference in treatment is that θ are provided to the fit with a previously known uncertainty $\Delta\theta$. This uncertainty comes from a previous measurement and is usually expected to be Gaussian. Note that if the fitted dataset includes additional information on θ the post-fit uncertainty on $\hat{\theta} - \theta_0$ might be smaller than $\Delta\theta$. In some cases this constraint might be expected, in other cases this hints at a technical error.

Different types of NPs can be distinguished: for NPs connected to detector properties they are usually measured in data of the same experiment, for NPs connected to theoretical calculations like cross-sections of other processes the calculations are interpreted as measurements. Some NPs are also coming from the statistical uncertainty of a *mc* dataset, here the Poissonian uncertainty is used for $\Delta\theta$. For example, the differences between *data* and *mc* per bin (called γ) have uncertainties based on Poissonian distributions.

This example hints at an additional complication for the other types of NPs: are they universal across all bins or do they impact the samples differently from bin to bin? The first is called normalization, the latter shape. NPs describing the cross-section of a background sample do affect this sample in all bins the same, so they only have a normalization impact on that sample. A NP related to the reconstruction might affect each bin separately, so it has a normalization and shape effect. To minimize statistical fluctuation the shape effects can be smoothed. In case the NP has an overall little impact on the fit, it might actually be neglected for this fit. This can be done separately for normalization and shape impacts (pruning).
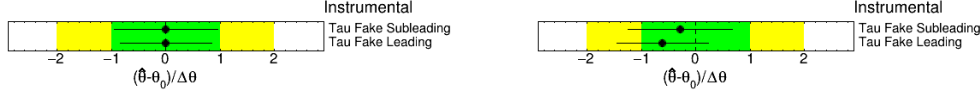

Figure 4.4: Example nuisance parameter plots for an Asimov fit (left), and a fit including some data (right). The lower NP is more constrained than the upper.

Figure 4.4 shows example plots for two NPs. The left plot is obtained from an Asimov fit, we see that $\hat{\theta} = \theta_0$. The uncertainty is comparable to $\Delta\theta$ for the upper NP, but constraint for the lower NP. The right plot is obtained from a fit with data. Here the estimate for θ is lower than the given θ_0, this is called pull. A pulled NP points to a mismeasurement of θ in the pre-measurement, or to a pre-measurement that is not fully applicable to the fitted *data* (for example because the used *data* were inherently different to the *data* used in the pre-measurement).

In some cases, one might already suspect that the pre-measurement of a parameter might not be applicable to the *data* at hand. In this case it can be useful to add an additional measurement to the fit with *data* that is more comparable to the SR. This is called a Control Region (CR). A CR contains additional *data* , which needs to be similar to the *data* in the SR but does not contain any event of the SR (called orthogonality). One example is a CR which helps to determine the cross-section of a *bkg* sample. Ideally, the CR is dominated by this sample, close to the SR, but orthogonal. The parameter measured in the CR is usually computed as a μ, called normalization-factor (NF). The fit is first performed on the CRs, then the NFs are extrapolated to the SR and at last the SR is fitted [87].

4.4 Discovery and limit setting

The interpretation of a fit as described above is usually done in terms of $\sigma(\mu)$. If the fit result for μ is bigger than $5\sigma(\mu)$ this is deemed a discovery, as the probability for a statistical fluctuation causing this excess from 0 is highly unlikely. If this degree of certainty is not reached one can instead calculate an exclusion limit x, claiming that $\mu > x$ is excluded on a 95% confidence interval. This describes that in only 5% of all statistical fluctuation cases μ would take a value higher or equal to x. Special caution has to be taken to interpret measurements with a high amount of expected *bkg* events [88]. For a given analysis strategy the possible outcomes are then scanned to find x, which is often expressed in terms of μ. For example observing $x = 5\mu$ might be excluded at a 95% confidence interval.

To obtain a more significant result, several similar measurements might be combined. The measurements in this case are referred to as channels and while they should be aimed at the same observable the events used in any of their regions should not be reused (all channels need to be orthogonal). The fit is then performed with each NP completely correlated between the channels. μ on the other hand is computed

once for each channel individually and once with a fit on all SRs together.

The complete procedure explained above is implemented in a program called TRExFitter, which is used by this thesis. It is based on RooFit [89] and HistFitter [87].

Experimental setup

I want to know why. Why everything. I don't know the answers, but a few days ago I didn't know there were questions

– Terry Pratchett, *Nation*

5.1 Large Hadron Collider

Why is so much effort made on building accelerators? A high energy of colliding particles is needed to create heavy particles (like the *t* and *H*, see also chapter 2). Highly energetic particles do exist in nature as cosmic particles which are accelerated in space (probably close to rotating stars [90]). However, those are not predictable in terms of occurrence, energy and place of interaction with 'earth-particles'. Therefore, it is beneficial to build man-made particle accelerators. This section describes the Large Hadron Collider (LHC). Therefore, first the basic design principles of particle accelerators are described. Second, important benchmark properties like the luminosity are discussed. Finally, an overview over the LHC including its pre-accelerators is given.

In the SM the only available force to accelerate elementary particles is the electromagnetic force. Therefore, accelerators use electrically charged particles (with charge q) and accelerate them via an electric field $\vec{E}$ of length L to an energy E:

$$E = q \int_0^L \vec{E} ds$$

The interesting processes can only happen at the moment highly energetic particles collide. Here, two basic design principles can be chosen from: either one source of highly energetic particles that hit a fixed target (similar to what would happen for cosmic radiation) or two sources of highly energetic particles that collide. The first configuration is easier to achieve as just one set of particles has to be accelerated, the latter configuration achieves a higher center of mass energy with the same energy:

$$\sqrt{s_{\text{fixed target}}} = \sqrt{2Em_{\text{target}}}$$
$$\sqrt{s_{\text{head-on collision}}} = 2E$$

Another differentiation can be made between linear and circular colliders. To achieve the highest energies

a circular collider is preferable as the particles can circulate multiple times and can be accelerated each time. The downside of this is that charged objects forced on a bend trajectory do emit synchrotron radiation, which for highly energetic particles results in an energy loss of

$$-\Delta E \propto \frac{q^2 E^4}{\rho m^4}$$

which is anti-proportional to the bending radius ρ, but especially rises with a power of four with respect to the energy and m^{-1}. Due to the mass of an electron being small this becomes a limiting factor of a circular electron accelerator. For heavier particles a limiting factor for circular accelerators is the strength of the bending magnets:

$$\frac{1}{\rho[m]} = 0.2998 \frac{B[T]}{p[GeV\,c^{-1}]}$$

The average magnetic field B and the particle momentum p limit the bending radius ρ (as the accelerator does not only consist of magnets ρ does not equal the accelerator radius R).

The LHC is the world's largest and highest energetic particle accelerator and is a circular collider. It is located at the Swiss French boarder and is operated via CERN, the world's biggest high energy particle physics facility bringing together thousands of researchers from countries all over the world. Up to now two run intervals have been completed, Run 1 from 2009-2013 and Run 2 from 2015-2018. This thesis describes and uses only events from Run 2. The LHC accelerates protons, but is built in the tunnel of the previous LEP accelerator which used to accelerate electrons [91]. It is a 26.7 km circumference ring. The ring design allows to accelerate the particles several times with the same structure as it circles. In order to bend protons to the course of the LHC it was designed with 8.33 T dipole magnets [92, 93]. Those were developed for the LHC and by physicists and engineers working for the LHC. For comparison the magnets used by LEP only needed to be 0.11 T strong [94] to keep the particles on track.

Another downside of protons is that they are themselves not elementary particles. Therefore, while accelerated to an energy of 6.5 TeV the colliding partons have an unknown and lower energy (see section 2.2.5). Additionally, the partons that are not included in the hard interaction can produce underlying events which are not of interest. They make it more challenging to measure the less-likely process of interest. The number of interactions per bunch crossing $\langle \mu \rangle$ for the LHC can be found in table 5.1.

5.1.1 Luminosity

To enhance the probability of interesting events in the LHC the goal is to have many collisions. Therfore, not single protons are accelerated but 'bunches' of about 10^{11} protons. The "squeezing" of the bunches to a relatively small package is one of the most challenging tasks of an accelerator. At the LHC 392 quadrupole magnets are used to squeeze the bunches of protons.[1] Since all protons carry the same sign of charge the bunches always inflate again. The bunches are squeezed to be the densest at the interaction points of the experiments. The measure β^* does roughly correspond to the distance of the focal point and the point where the bunch cross-section is doubled. Since the repelling force increases with smaller distance, a lower β^* corresponds to a more densely packed bunch. The LHC achieved values of up to $\beta^* = 0.8\,\text{m}, 0.4\,\text{m}, 0.3\,\text{m}, 0.25\,\text{m}$ in the years 2015,2016,2017 and 2018 respectively [95]. From those

---

[1] Additionally, to the quadrupole magnets higher order magnets are used to shape the bunches and correct for other smaller effects like gravity.

	$\int \mathscr{L}[\mathrm{fb}^{-1}]$	$\langle \mu \rangle$	$\beta^*[\mathrm{cm}]$
2015	4.2	13	80
2016	39.7	25	40
2017	50.6	38	40-30
2018	66	37	30-25

Table 5.1: LHC parameters over the years. Integrated luminosity and β^* are given for the ATLAS detector site [95].

numbers it becomes obvious that the performance of the LHC differed (and increased) over the years. The most important measure of this from the viewpoint of the experiments is the luminosity $\mathscr{L}$ which is a direct measure of the rate of events of a given process that are expected to be produced:

$$\frac{dN}{dt} = \mathscr{L}\sigma \tag{5.1}$$

$\mathscr{L}$ is also defined via machine parameters

$$\mathscr{L} = f\frac{N_1 N_2}{4\pi\sigma_x\sigma_y} \tag{5.2}$$

where f stands for the frequency of bunch crossings, $N_{1,2}$ for the number of protons in bunch 1 or 2 and $\sigma_{x,y}$ for the width of the bunches in both directions. The integrated $\mathscr{L}$ of the LHC can be found in table 5.1. As one can see, the highest integrated luminosity (which means highest number of expected events of a certain cross-section, for example *tHq*) was achieved in 2018, second was 2017 while 2015/2016 did not contribute as much [95].

5.1.2 Accelerator chain

One important aspect of an accelerator has almost not been discussed yet – the acceleration. For any charged object an electric field is the obvious way to accelerate it. However, when the electric field is static, a charged particle would immediately be pulled back after passing the boundary of the electric field. Therefore, a number of electric fields which change their polarization in synchronization with the passing bunches is the basic idea of most acceleration designs. Since the speed of a proton before acceleration is way lower than the speed of the protons at their maximum energy several accelerators of different technology are needed [96]:

1. **Duoplasmatron:** The proton source used is a duoplasmatron, where a hydrogen gas is heated and protons are extracted via an electric field of 91 kV [97–99].

2. **RFQ2:** The Radio Frequency Quadrupole (RFQ2) is an accelerator made out of quadrupoles which accelerates the beam from 90 keV to 750 keV and spaces the beam into bunches [99, 100].

3. **LINAC2:** a linear accelerator that accelerates protons to an energy of 50 MeV via drift tubes at increasing distances[2] [101–103].

4. **PSB:** The Proton Synchrotron Booster (or short PS Booster) accelerates protons from 50 MeV to 1.4 GeV in a synchrotron setup. In order to increase the number of protons being processed it

---

[2] Organizational also the RFQ2 and duoplasmatron are part of Linac2

consists of 4 accelerating rings vertically stacked [104, 105].

5. **PS:** a Synchrotron (also CERN Proton Synchrotron) which accelerates protons to 26 GeV and prepares the bunch structure which is kept throughout the further acceleration. Therefore, the bunches injected from PSB are split and compressed to create bunches with the desired 1.4×10^{11} protons at a distance of 25 ns. In the BCMS scheme the bunches are additionally merged after compression which creates bunches of 1.3×10^{11} protons with better beam qualities[3]. During the years either a 'continuous' stream of bunches (72 in a fill) has been produced or an alteration of eight bunches and 4 empty slots[4] [105–107].

6. **SPS:** In the Super Proton Synchrotron up to four injections of the prepared bunches are accelerated to 450 GeV while maintaining the good beam quality needed [105].

7. **LHC:** stores two beams of protons which circulate in opposite direction and accelerates them from an energy of 450 GeV to the energy of 6.5 TeV. The beam crossings (and experiments) are also hosted here [108].

A sketch of the CERN accelerator complex can be found in figure 5.1.

The full layout of the LHC can be found in figure 5.2. The LHC has eight bending sections which each host 138 dipole magnets to bend the two beams as well as higher order focussing magnets [107]. In between those are eight straight sections which each serve an individual purpose. Just one section is used for the actual acceleration (marked with RF). Two sections are used to maintain the good beam qualities and the last section is used to dump (eject) the beams. Each of the beams can be controlled separately. At four of its eight straight sections the LHC hosts experiments. Out of those four experiments ATLAS is explained in more detail in the next section.

---

[3] The bunch is more compact and the proton momenta are more similar [106].

[4] A stored proton beam over time creates electron clouds in the vacuum through processes like photoemission from Bremsstrahlung or interaction with residual objects which can disturb the beam, the 8b4e scheme mitigates that effect [106, 107].
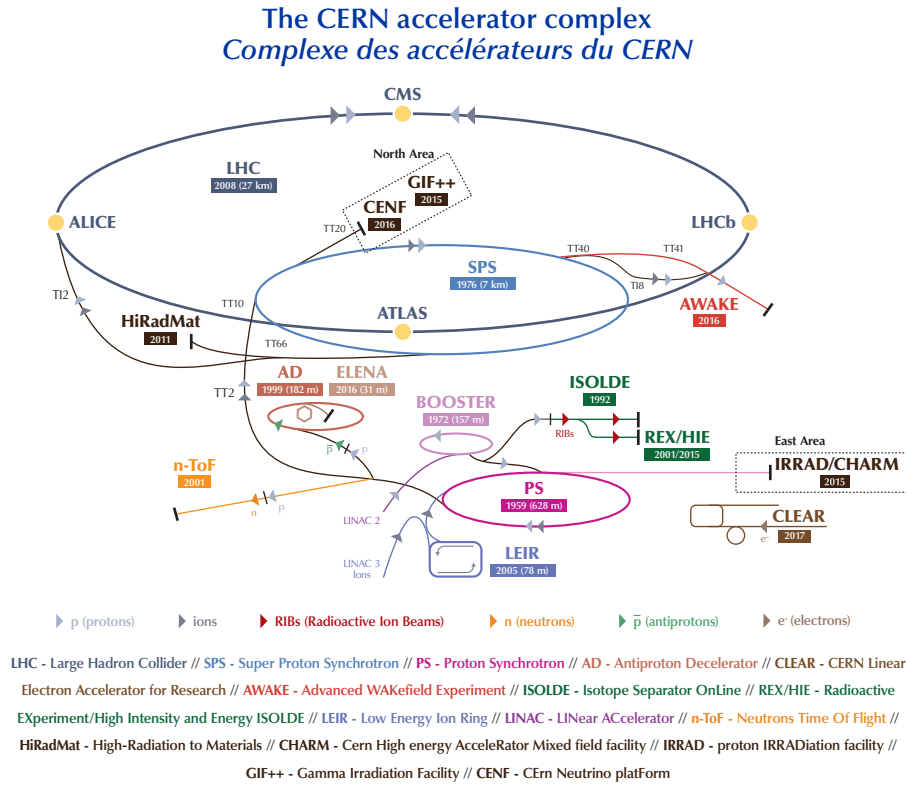
Figure 5.1: Layout of the CERN accelerator complex in 2018. The accelerator chain discussed here includes LINAC2 (purple), Booster (dusty pink), PS (magenta), SPS (light blue) and the LHC (dark blue) [109].
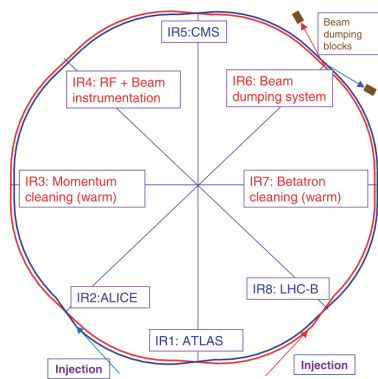

Figure 5.2: A more detailed layout of the LHC, indicating the position of the four LHC experiments and the use of the other straight sections [108].

5.2 ATLAS

The ATLAS experiment is located at point 1 of the LHC (see figure 5.2). It was designed to measure the highest energy (hard) processes that occur in proton collisions. Therefore, it consists of several subdetectors that are able to measure almost all fundamental particles[5] (or the decay products of fast decaying particles). Before its subdetectors are described, we first need to take a look at the coordinate system used in ATLAS. The main focus is on the objects that have a high momentum transverse to the beam direction (p_T) as this requires a high four-momentum transfer during the process (hard process). Therefore, from the three cartesian axes x (towards the middle of the LHC), y (upwards) and z (along the beam direction) another coordinate system is derived. It utilizes the spherical coordinate system where the azimuth angle ϕ describes a rotation around z. Instead of the polar angle θ another measure is used:

$$\eta = \ln\left(\tan\left(\frac{\theta}{2}\right)\right) \tag{5.3}$$

Additional to giving a more intuitive correlation between p_T of an object and its position with respect to the detector, in the limit of high energy η equals the Lorentz invariant variable rapidity:

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right) \tag{5.4}$$

It is therefore called the pseuodorapidity. With these definitions, distances in ATLAS are commonly measured in the η-ϕ-plane via

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \tag{5.5}$$

A sketch of the ATLAS coordinate system can be found in figure 5.3.
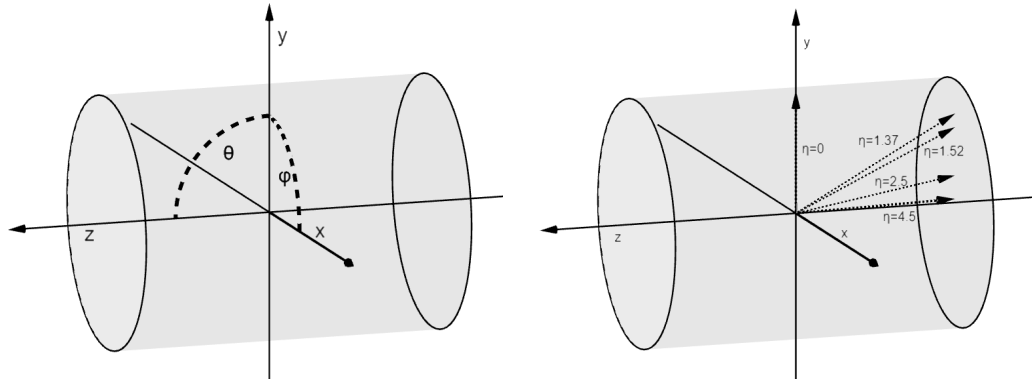
Figure 5.3: Sketch of the ATLAS coordinate system, the axis as well as θ and ϕ are indicated in the left picture. The right picture shows important η thresholds.

An overall model of the ATLAS detector indicating the placements of its subdetectors can be seen in figure 5.4. The subdetectors are mainly focused on measuring the low η regions and are therefore positioned

---

[5] Neutrino detection is not possible as their interaction probability is small due to their only weakly interacting nature and thus requires even larger detectors and a low background.

cylindrically (in the *x-y* plane) around the interaction point (IP). In order to not lose the information of objects with higher η (called forward) additional forward detectors are installed. In the following, both sets of detectors are discussed from the ones closest to the IP outwards. One important source for this section is [110]. For an easier overview section 5.2 shows an outline of all subdetectors and their components.
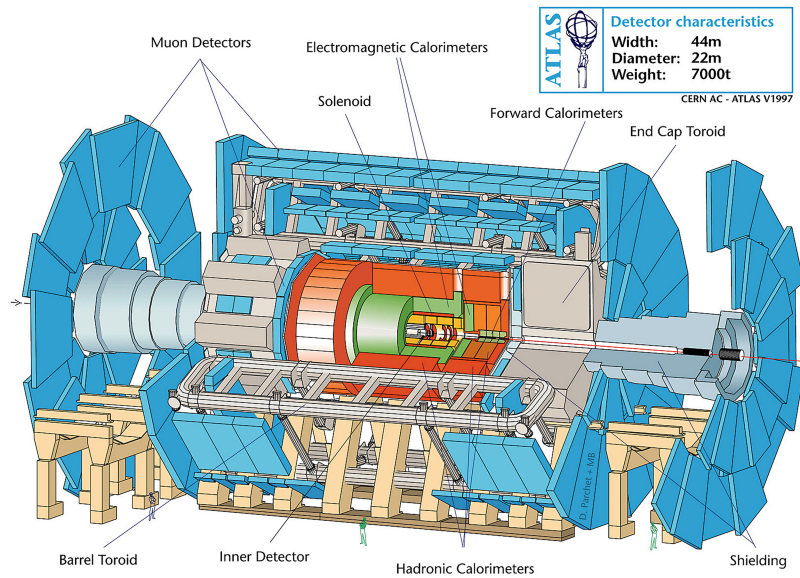

Figure 5.4: Overview of the ATLAS detector [111].

5.2.1 Tracking system

The tracking system is located closest to the interaction point (the innermost layer is just 33 mm away [113]). It is designed to track the path of charged particles (which is bend due to the magnet system described in section 5.2.5) without destroying them. The two basic designs for this are semiconductors and gaseous detectors.

Semiconductor detectors consist of a bulk of depleted semiconductor material with smaller locations of enriched semiconductor material. Applying a voltage creates a charge difference and subsequently an electric field. Therefore, the free charge carriers (electrons and holes) are depleted from the bulk of the detector. A charged particle passing through this depleted zone creates electron-hole pairs. These drift due to the electric field and can be measured at the bulks. Due to the possibility to create fine structures in semiconductors (here silicon) this can measure the finest resolution of tracks (but is also the most expensive).

Gaseous detectors follow a similar principle as semiconductor detectors, but they utilize gases as the main material. In this, an electron-ion pair is created (often via ionization). A voltage is applied to a metal structure within the gas in order to create the electric field. This not only makes the charge carriers drift, but also accelerates them. They can then ionize further atoms and create an avalanche which is easier to detect. A common geometry consists of gas filled tubes with a wire in the center.

In the central region the tracking system consists of the following subdetectors:

- **Pixel detector:** semiconductor detector made of silicon in two-dimensional pixels of size $50\,\mu\mathrm{m} \times 250\,\mu\mathrm{m}$ in the innermost *b*-layer and $50\,\mu\mathrm{m} \times 400\,\mu\mathrm{m}$ in the two outer layers [113–115].

Sytem	Detector	Detector type	η coverage	Section	# channels
Tracker	Pixel	Silicon pixel	$\eta < 2.5$	5.2.1	92 M
	SCT	Silicon strip det.	$\eta < 2.5$		6.3 M
	TRT	Transition radiation det.	$\eta < 2.0$		350 k
EM cal.	LAr	EM calorimeter	$\eta < 3.2$	5.2.2	170 k
Hadronic cal.	Tile calorimeter	Had. calorimeter	$\eta < 1.7$	5.2.3	5200
	Hadronic End-Cap	LAr	$1.5 < \eta < 3.2$		5600
	Forward LAr Calorimeter	LAr	$3.1 < \eta < 4.9$		3500
Muon	MDT Muon Drift Tubes		$\eta < 2.7$	5.2.4	357k
	CSC Cathode Strip Chambers		$2 < \eta < 2.7$		31 k
	RPC Barrel Muon Chambers	Resistive plate chambers	$\eta < 1.05$		383 k
	TGC Endcap Muon Chambers		$1.05 < \eta < 2.4$		320 k
Forward	LUCID	Cerenkov det.	17m		32
	ALFA	Roman pots	240m		10 k
	AFP	Roman pots	210m		430 k

Table 5.2: Overview of the ATLAS subdetectors. The η coverage is given in order to understand the impact on reconstruction. Particles that transverse the detector at different η might pass through a different amount of material. For the forward detectors the distance from the IP (along the beam axis) is given instead. In order to give an estimate of the precision and complexity the number of readout channels per detector is given, for the tile calorimeter the number of readout cells is used [112].

- **Strip detector:** semiconductor detector made of strips which can only read out information in one direction. Eight layers in alternating directions are used to obtain three-dimensional position information [116].

- **Transition radiation tracker:** gaseous detector with tube geometry. The initial charge is created via ionization or transition radiation. The latter occurs when a charged particle crosses the border between two materials with different refractive indices. Its intensity is proportional to γ (until a saturation is reached). This allows for a differentiation between photons and electrons [117].

Additional to the cylindric system, the central system also consists of disk-like detectors in the same order (pixel-, semiconductor- and TRT-detector) which extend the overall coverage of the system to $\eta = 2.5$.

An overview of the tracking system can be found in figure 5.5.

5.2.2 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECal) is designed to measure the energy of photons, neutral pions or electrons. This is done by stimulating an electromagnetic shower, which is a cascade of Bremsstrahlung and pair production until all the energy of the primary particle is absorbed by the detector. The probability of pair production and Bremsstrahlung are going like Z^2. Therefore, the environment in which this is facilitated consists of atoms with a high charge Z. The used measure for depth is the characteristic length X_0, which follows $X_0 \propto Z^2$, it is defined via
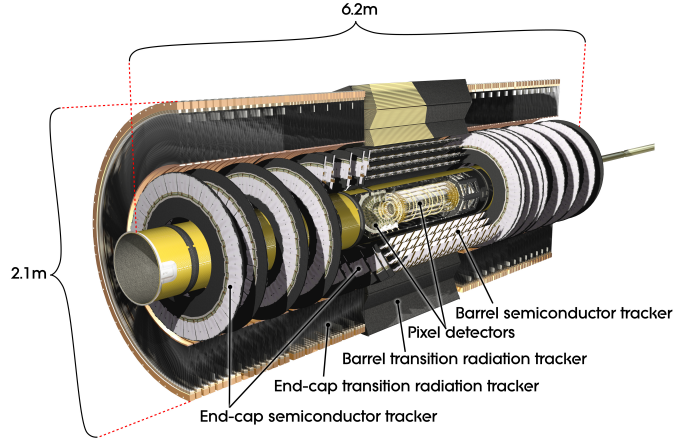
$$\frac{dE}{E} = \frac{dx}{X_0}$$

Figure 5.5: Overview of the ATLAS tracker system. The additional 'insertable *b*-layer' is not shown here (as it was only installed in 2014) [118].

for electrons. For highly energetic photons the absorption length λ is used instead:

$$\lambda = \frac{-N}{\frac{dN}{dx}} \approx \frac{9}{7}X_0$$

ATLAS chose to use lead ($Z = 82$) as the passive material (material which facilitates the interaction). However, the emerging energy also has to be read out. This is done in the active layers. ATLAS chose to use liquid argon ($Z = 18$) which gets ionized by the shower products. A high electric field draws the ionization products to the readout channels (ionization chamber). The different layers need to be ordered in a way that ensures the shower to pass both the active and passive layers in order to fully shower out and a high fraction of the original particle being measured[6]. In the ATLAS LAr calorimeter this is achieved by an 'accordion-like' structure – the different layers are folded back and forth at an angle with respect to the IP. A sketch of this can be found in figure 5.6. This allows for a frequent crossing of both types of layers and for the readout to occur at the outside of the calorimeter section without the need for a blind space where the cables are located [117].
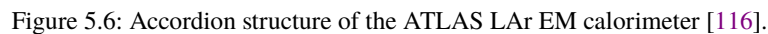
The LAr tile calorimeter is installed in two sections in the ranges of $|\pm \eta| < 1.475$. To achieve a better energy resolution an additional presampling layer of liquid argon is installed in $|\eta| < 1.52$. Together they achieve up to $33X_0$. In the more forward regions a similar configuration is installed: first a presampler ($1.5 < |\eta| < 1.8$), then the ECal covering $1.375 < |\eta| < 3.2$. The latter is installed as two disks (like a lid for the cylinder). The depth varies from $24X_0$ to $28X_0$ in the region of $1.475 < |\eta| < 3.2$ [116].

5.2.3 Hadronic Calorimeter

A hadronic calorimeter is designed to measure the energy of hadrons (mostly pions). The processes at play are hadronic and more diverse than for the electromagnetic calorimeter. They revolve around hadrons interacting with heavy nuclei, which produce other hadronic objects (like protons or pions) which

---

[6] It is not possible to measure all the energy as some of it is lost in the passive layers, therefore the output of the calorimeter has to be calibrated with particles of a known energy.

Figure 5.6: Accordion structure of the ATLAS LAr EM calorimeter [116].

again interact with further nuclei. The emerging neutral pions start an electromagnetic shower which is responsible for a large component of the measured energy fraction. To keep both shower components at a similar spatial expansion, designs with a passive component with a low Z (like iron $Z = 26$) are preferred.

The ATLAS detector has one Tile calorimeter in the central region. It uses steel plates in the passive layers and scintillators[7] in the active regions. To provide detailed information about the shower progression, it is separated in η, ϕ and r. The resulting 5200 cells are all read out separately. In order to ensure the full containment of the shower (and such the measurement of all energy) it is built to be 7.4λ deep. In the forward region two lead-liquid argon calorimeters are used, the Hadronic Endcap LAr calorimeter (HEC) from $\eta = 1.5$ to $\eta = 3.2$ and the Forward LAr calorimeter (FCal) from $\eta = 3.1$ to $\eta = 4.9$. The HEC is constructed as two wheels on each side using copper and liquid argon in plates oriented along r and ϕ.

To measure the most forward objects, the FCal is inserted even closer to the beam pipe ($\approx 5 - 45\,\mathrm{cm}$). It consists of one module designed to measure the electromagnetic shower and two modules which are designed to have a short λ. All of them are build out of copper plates with holes for the electrodes. Those mainly consist of a copper rod for the first module ($Z = 29$) and tungsten rods for the later two modules ($Z = 74$) in combination with LAr as the active layer.

A sketch of the whole calorimeter system can be found in figure 5.7.

---

[7] Scintillators consist of materials which absorb high energetic radiation and emit the energy again as light in a characteristic spectrum. This can then be read out by photomultipliers (PMT) which convert photons via the photoelectric effect into electrons. Those start an avalanche of electrons that can be readout without the need of amplifiers. Therefore, this system allows for a relatively fast detection [117].
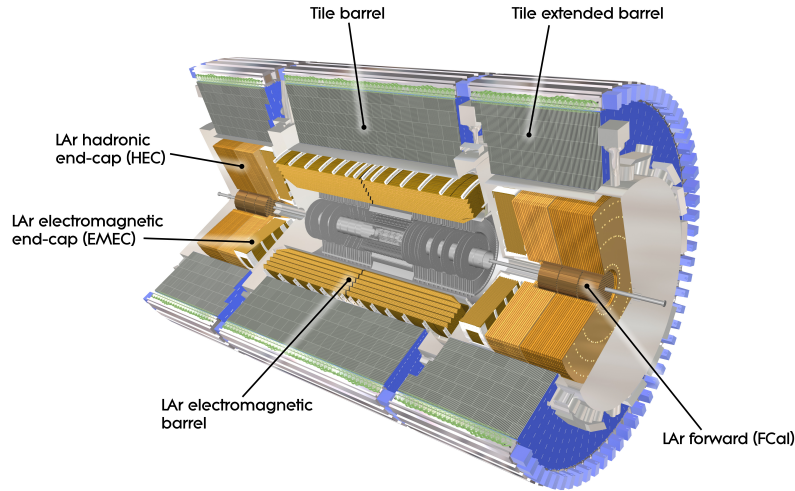
Figure 5.7: Overview of the ATLAS calorimeter system. Liquid argon and Tile calorimeter are shown as well as the forward calorimeters [119].

5.2.4 Muon detectors

The muon detectors are installed outside the calorimeters, and aim at measuring the tracks of charged objects that traverse the calorimeter (mostly muons). They consist of different types of gaseous detectors in different η-ranges. An overview of the muon detector system is shown in figure 5.8.

Monitored Drift Tubes (MDTs) cover the largest part of the detector in $\eta < 2.7$ (with a small gap at $\eta \approx 0$ for cables for the inner-lying detectors), except for the innermost layer where only $\eta < 2.0$ is covered. They are designed to enable precision tracking over a large volume. Therefore, the less expensive gas-filled tube design for a tracker is chosen (see section 5.2.1). The tubes are filled with an argon/CO2 mixture and generated electron-ion pairs are collected via a voltage of 3 080 V. At this high voltage both components are accelerated so much that they create further electron-ion pairs, those partially recombine which creates photons that can create new electron-ion pairs further away. This phenomenon is called streamer and is not wanted for a stable detection. In order to prevent it a quencher gas (3% CO_2) is added to the mixture which has a molecular structure to absorb the photons (and emits several lower energetic photons in return) alongside the 97% argon which has superior ionization and drift properties.

In the more forward ($2.0 < \eta < 2.7$) region the number of particles traversing the detector is way higher, thus the inner layer of the MDT is replaced by Cathode Strip chambers (CSC) which are able to measure several tracks at once. This is possible due to the different geometry. Instead of separated tubes a larger rectangular volume with wires spaced at the same distance between each other and the wall is used. This wall functions as the cathode and is separated into strips orthogonal to the wire direction for readout. In order to prevent high energetic photons from starting showers in other areas of the volume a higher ratio of quenching gas (80/20 Ar/CO_2) is utilized.

The detection of muons with the described systems is accurate but not the fastest. In order to be able to trigger on muons the time precision has to be smaller than the time between two bunch crossings. A gaseous detector which is designed to be fast are resistive plate chambers (RPCs). As the name suggests they consist of two plates made of a highly resistive material (phenolic-melaminic plastic laminate in

ATLAS). If a shower occurs the collected charges take some time to flow away and create an electric field which counters the shower, which leads to a faster dissolution of the shower. Additionally, a gas mixture is used which is able to collect both highly energetic photons and free electrons. As the highly resistive plates would be slow to read out a second layer of conductors (separated in strips) is installed next to them which is able to pick up the charge influence. In the more forward region TGCs (Thin Gap Chambers) are used; they are built similar to CSCs, but the gap between wire and wall is thinner than the gap between wires. Together with a highly quenching gas mixture this controls streamers better and allows for a better time resolution [120, 121].
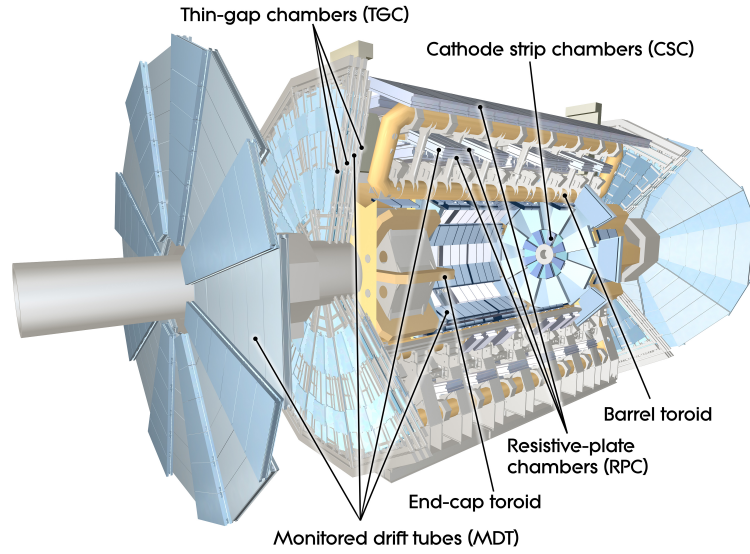

Figure 5.8: Overview of the ATLAS muon detector system [122].

5.2.5 Magnet system

A large part of the ATLAS detector is its magnet systems, which create two different magnetic fields. Both fields bend the path of traversing charged particles depending on their momentum, this allows for a p_T reconstruction (see section 5.4).

A sketch of the magnetic field lines can be found in figure 5.9. The inner magnetic field is created by a superconducting 2 T solenoid, this creates field lines along the z-axis and causes particle tracks to be bent in ϕ. The solenoid windings are contained in the LAr calorimeter vacuum which is needed to be operated at a temperature of 4.5 K. The outer magnetic field has the form of a toroid (this is the origin of the 'T' in the ATLAS acronym). It is created by one 0.5 T magnet in the central region and two smaller toroids that create a magnetic field of 1 T at higher η. The magnet field lines of the toroids follow ϕ which forces the particle tracks to bend in η which provides an additional independent measurement of p_T. It is located around the muon system and due to its size each coil of the central toroid is contained in its own vacuum vessel. The end-cap toroids are each build within one volume which is cooled to the needed 4.5 K for a superconducting magnet [123].
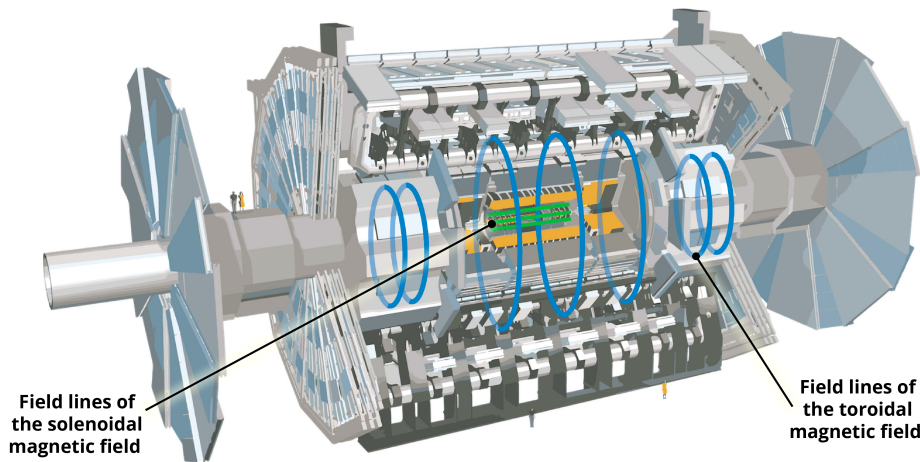
Figure 5.9: Sketch of the magnetic field lines in the ATLAS detector [124].

5.2.6 Forward detectors

Some detectors are not built to detect the high p_T particles but aim at detecting forward particles and are thus located very close to the beam line. An overview can be seen in figure 5.10. (The ZDC aims at measuring in heavy ions runs and thus is not discussed here.)

LUCID2

LUCID (Luminosity measurement using Cerenkov Integrating Detector, updated for Run 2) is a Cerenkov detector which aims at measuring the luminosity delivered to ATLAS by measuring forward components of inelastic scattering. The detector on each side (10 cm away from the beam pipe, $\eta \approx 5.8$) consists of twenty gas bottles filled with Perflubutane (C_4F_{10}) which emits light when energetic particles ($> 2.8\,\text{GeV}$ π, $> 10\,\text{MeV}$ e) cross [125–127]. This phenomenon is called the Cerenkov effect and is caused by charged particles passing through a medium at a speed that is faster than the speed of light in that medium. The emerging photons are detected by PMTs in 16 of the bottles (the other 4 are read out with another method which is studied for future use). As the amplitude of the PMT signal is proportional to the original number of particles detected this setup provides a relative measurement of the luminosity at ATLAS, which is calibrated during special low-luminosity runs of the LHC.

ALFA and AFP

ALFA (Absolute Luminosity For ATLAS) is a detector which is designed to measure products of elastic scattering. These are mostly protons which are scattered at low angles. Therefore, it is located 240 m from the IP, but only less than 1 cm away from it (located above and below the beam pipe). This is made possible by the Roman-pot design, which is a small volume where the detector volume is not part of the vacuum but is moved close to the beam pipe by bellows. In this small volume scintillating fibers pick up the signal which is read out by PMTs (located further away from the beam pipes).

The AFP (ATLAS Forward Proton detector) is located on both sides of the IP once 205 m (near) and once 217 m (far). It uses Roman-pots which are located vertically along the beam pipe. The near Roman-pots contain silicon trackers while the far trackers contain silicon trackers and time of flight detectors. It was installed during Run 2 (2016 and 2017).

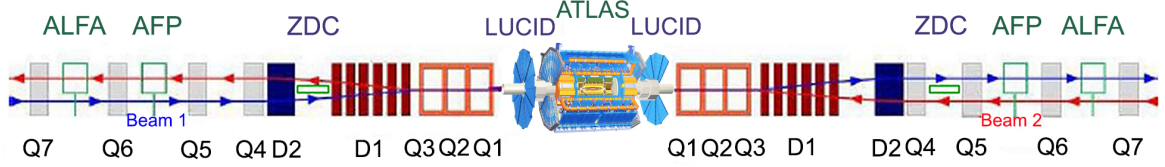Both ALFA and AFP were used in special runs to measure physics with very forward protons [128].

Figure 5.10: Sketch of the ATLAS forward detectors [129].

5.3 Trigger and data processing

5.3.1 Trigger and data flow

As discussed in section 5.1 bunch crossings occur during data taking with 40 MHz and as discussed in section 5.2 a lot of different subdetectors produce information which needs to be stored and processed. This results in an event size of about 1 MB [130] and would very shortly outgrow any storing and processing capacity. On the other hand not every bunch crossing produces an interesting hard collision, but often times just soft collisions, which are not of interest. Therefore, a system is set in place which decides which events to keep in a short timeframe, this is called trigger. It is divided into a L1 (level 1) trigger and a high level trigger (HLT) which is applied on top of the L1 decision [131].

There are several limits on the capacity and other effects which play a role in the decision taking process besides which hard process occurred:

- Read out window: Reading the detector information takes time, especially the LAr information is collected over several bunch crossings. Therefore, a dead time of about four bunch crossings is applied after every triggered event to avoid overlapping events.

- Front-end buffers: While awaiting the trigger decision event information is stored on front-end boards (FE) and deleted or passed on afterwards, these have a limited capacity and no further information is stored when they are full. This accounts to a peak inefficiency of 1% for the L1 trigger.

- Bunch crossing types: There are 3564 bunch crossing slots per LHC revolution, not all of them can be used for a physics trigger decision, for example because just one of the beams actually had a bunch in the slot. Some others are used for detector configuration, counter resets etc.

- Resonance frequencies: Reading accepted data from detectors in a magnetic field causes a Lorentz-force in the connecting wires, if this happens at the resonance frequency it can cause them to break over time. To avoid this a fixed-frequency-veto is applied. However, the impact on physics data taking is small as this happens mostly at runs with a low number of bunches and secondary triggers are cut.

- Noisy subdetectors: Some subdetectors produce more noise than others, the L1 triggers which depend on them just take a subset of the events (every x event), this concept is called prescaling.

- Data quality: When reconstructing events a lot of things have to all go correctly to get a valid reconstruction. This is monitored directly during the run and checked again offline. If any anomalies occur, the events are not used for analysis [132].

- Data transfer: Cables between different data processing units have a limited capacity (this was actually the limit on the L1 trigger rate in Run 1 and upgraded for Run 2) [133].

- Buffer: While awaiting a HLT trigger decision the full information is stored and additionally the information necessary for the trigger decision is temporarily stored at the trigger farm.

- Trigger processing time: The L1 trigger takes about 2.5 µs and the HLT 0.4 s to process one event.

- Data storage: All final triggered events have to be stored both on tape and on the computing sites worldwide, in the raw but also in processed formats.

As a result of this the L1 trigger is set to trigger at a maximum capacity of 100 kHz and the HLT at 1 kHz. How this is achieved is described in the following and can also be seen in figure 5.11.
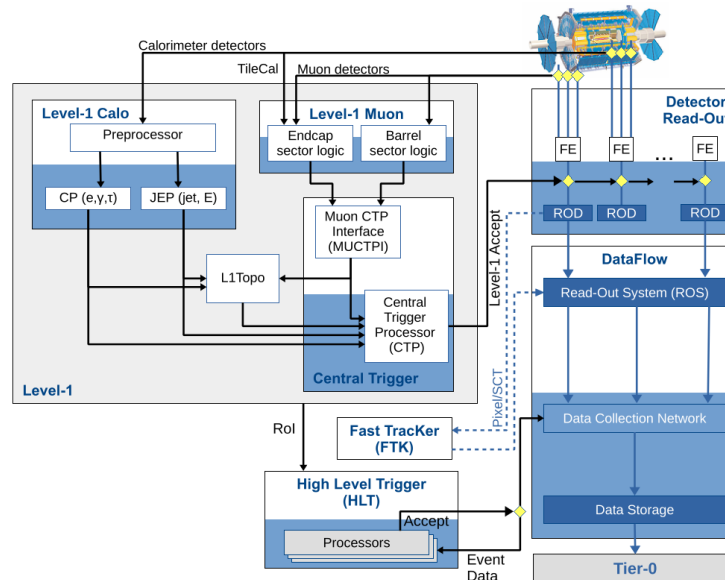

Figure 5.11: Sketch of the ATLAS trigger system in Run 2. Showing Trigger decisions and Data flow for data taken from the ATLAS detector. Taken from [131].

The L1 trigger processes events on a hardware level using electronics. Therefore, specific subdetector readouts are used. For most objects calorimeter information in reduced granularity from all calorimeters is used and processed separately to form e, γ, τ or jet and sum of energy candidates. For muons these are TPC and RPC (see section 5.2.4) information, which get processed separately for the barrel and endcap regions. In the latter, additional calorimeter information is used to mitigate triggers from signals not originating from the IP. The different information streams are combined with additional subdetector triggers (for example from LUCID) in the Central Trigger Processor [134]. The total trigger rate is 100 kHz for all output streams, this is mostly physically interesting data, but a small fraction of this is also taking not-so-interesting data which is used for evaluation and calibration.

For any event which passes the L1 trigger the FE information is sent to the Read-Out Drivers which are build specifically for each detector and format the data directly next to the detector and then send to the

Read-Out Systems. This data is then accessed as needed by the HLT which consists of a CPU farm in a cavern next to the detector. It looks into the regions of interest found by the L1 and does a fast version of the reconstruction for the different objects that can trigger [135].

The best trigger efficiency is reached for objects which have a clear signature that deviates from the all-hadronic profile of background processes. Electrons, photons and muons fall in this category (for more information on the reconstruction see section 5.4) while jets, hadronically decaying taus and also E_T^{miss} have a certain rate of failure due to their hadronic nature. Not all electron and muon candidates can be triggered as the data stream would be too big. In order to focus on electrons and muons from energetic interactions there is a p_T threshold of 27 GeV which has to be met for a single electron or muon. In case two or more objects can be reconstructed it is also possible to trigger on both of them with a lower p_T threshold [136–139]. In figure 5.12 one can see the efficiencies of single electron triggers in comparison with τ_{had} triggers.
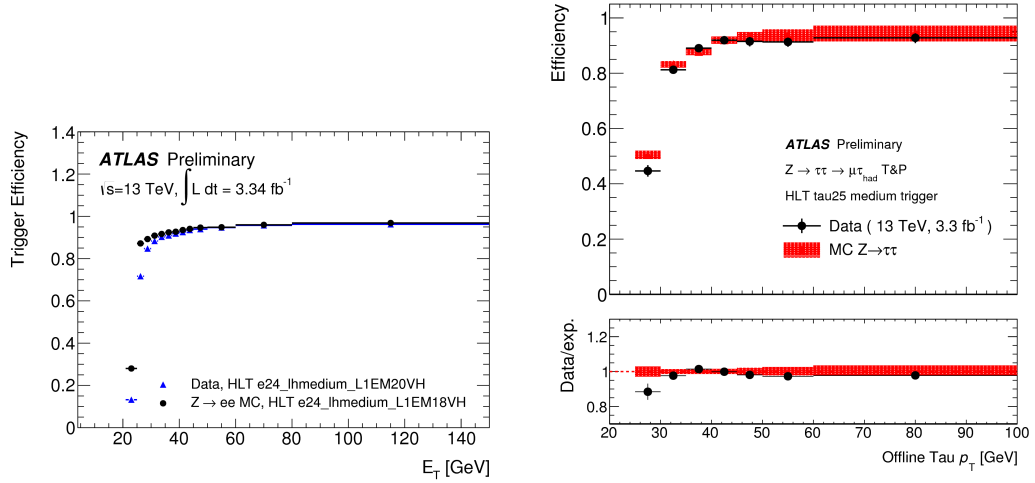
Figure 5.12: Comparison of trigger efficiencies for medium electron triggers (left) and τ_{had} triggers (right) for the 2015 dataset [136].

5.3.2 Further data processing

This section gives a brief overview over what happens with detector information from a technical point of view, more detailed information can be found for example in [130, 140–143].

Any event which is triggered by the HLT is transferred from the Read-Out System out of the caverns to Tier-0 of the CERN computing network. Tier-0 consists of servers, CPUs and data storage capacity of about 280 TB located directly at CERN [144]. It is responsible for processing raw data (in a byte-stream format) into a C++ representation which also includes reconstructed objects (ESD). As this is more than two times larger than the raw format of each event this is further converted into AODs which just contain the reconstructed objects and event information that is sufficient for most analyses. Both raw and AOD data are distributed to the Tier-1 sites of the CERN computing network (grid) worldwide, to distribute the storage load and create a backup for all used data. There are currently 15 Tier-1 sites on three continents which are responsible for storing and reconstructing the AODs, as well as performing the further processing on the AODs for example to create DAODs (derived AODs) which store only a subset of interesting events

and information for a certain analysis purpose. One example of these are TOPQ1 DAODs which are used by top-related analyses and contain events with at least one light lepton and a reduced set of calorimeter and track information (as well as a reduced set of truth information). The further processing of these to create analysis samples can then be performed at Tier-2 sites, which are currently 155 sites located around the world. The analysis samples can be accessed by Tier-3 sites which can be local clusters which just store information relevant to the people working at these universities (for example Bonn).

Additionally, to reconstruct the data collected by ATLAS the simulated datasets (see section 4.1) have to be created by the computing system. Main responsible for computing the different MC steps are the Tier-2 sites, while reconstruction is again mostly done on Tier-1 sites.

However, the whole hierarchical construction of the grid has the downside of not optimally using the computing resources available (for example at times when the detector is not taking data). Therefore, the whole task distribution became more flexible over the years. While every Tier has still the main responsibility of their main task, it is taking over other computing tasks if free. This goes so far as to using the HLT CPU farm for MC production when it is free.

5.4 Object reconstruction

This section describes the reconstruction procedure of the data collected by the ATLAS experiment [116, 145]. First, the basic ingredients tracks and clusters are described (sections 5.4.1 and 5.4.2). The later part of the section describes the objects which are associated with physical particles (for example electrons) and are used for the further analysis (chapter 6).

5.4.1 Track reconstruction

Charged particles leave signals along their trajectory in the innermost detectors as described in section 5.2.1. From these energy deposits (hits) one can reconstruct the trajectory of the particle, this is called track reconstruction. It is done first for the Pixel and strip detector and the found tracks are then later extended to the TRT.[8] As those detectors lie within the magnetic solenoid (see section 5.2.5) the particle trajectory and thus the reconstructed tracks are curved according to the Lorentz-force. This makes the reconstruction a bit more complicated than the reconstruction of a straight line, but allows to determine the p_{T} and charge (via bending direction) of single tracks, in addition to their starting point and direction in ϕ and η.

The main challenge in ATLAS is the high number of charged particles which results in a lot of tracks to reconstruct and even more candidates for tracks [146]. It can also happen that two tracks cross each other and deposit energy indistinguishably close to each other in the detector. In order to find track candidates a combinatorial Kalman filter [147] is used which starts from seeds made out of three hits in different detector layers and extrapolates them to find a matching hit in a subsequent layer, which then updates the predicted track. This is done iteratively until the hits form a complete track (expected 12 hits in the silicon trackers) [148].

These track candidates then have to pass a set of criteria to be used further: a minimum p_{T}, $|\eta| < 2.5$ (inside the coverage of the tracking detectors), the track needs to point to the IP and a sufficient amount of hits in the detectors as well as not too many non-hits where they would be expected. To mitigate the effect of indistinguishable close hits, a NN is used to identify those places [149] and subsequently the use of that hit for more than one track is allowed, otherwise the better track candidate is allowed to use that hit. Only

---

[8] Tracks in the muon detector are described in section 5.4.4.

tracks which pass the complete selection are fitted to receive the best direction, p_T and lowest uncertainty. For the tight selection efficiencies around 80% are achieved [150]. Figure 5.13 shows this efficiency as a function of η and p_T of the reconstructed track.
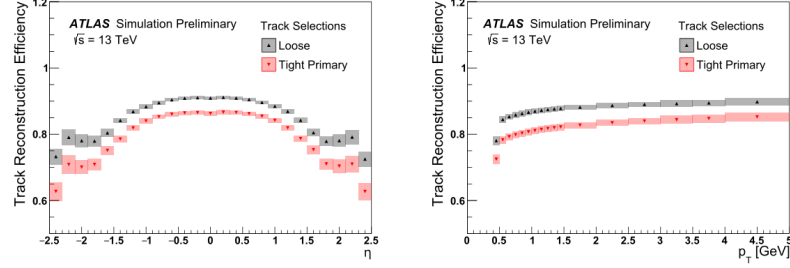

Figure 5.13: Track reconstruction efficiencies for 2015. Taken from [150].

Primary vertex

From the reconstructed tracks further properties can be determined, for example an approximation of the IP. While the IP is roughly known from collider operations the true point of interaction can vary. To find the interaction points for each event the fitted tracks are extended and a primary vertex (PV) is determined. For a candidate vertex at least two high-quality tracks have to cross. The candidate and all tracks are now fitted in an iterative procedure in which incompatible tracks play a smaller role in each iteration. After the vertex is determined, the tracks not compatible with the vertex are used to find another vertex. This is important for determining charged objects which did not emerge from the IP but from a later decay [151, 152].

The distance of any given track to the PV can be quantified to distinguish charged objects which did not emerge from the IP but from a later decay. The reconstructed track is extended to find the point of the closest approach to beamline or PV and a cartesian distance is calculated. The distance in the transverse plane $|d_0|$ is usually defined by the distance to the beamline as the width of it tends to be smaller than the uncertainty of the PV. Therefore, d_0 is often divided by this uncertainty to arrive at the d_0 signigficance. In the z-direction the distance to the PV is used to arrive at z_0. θ denotes the track angle, which leads to $|z_0|sin\theta$ denoting the distance of the track to the PV [153].

5.4.2 Clusters

While tracks can be considered the sum of information an object left in the tracking detectors, topological clusters play that role for signatures in the calorimeters. As they are a signature of electromagnetic and hadronic showers clusters spread in three dimensions but do not necessarily leave energy in all layers of the calorimeters. The unique information that is picked up from clusters is the energy and therefore determining which cells measurement to add is the main goal of cluster formation. The difficulty for that is the energy measured from electronic noise and from low energy background processes (pile-up). Therefore, just neglecting all cells that did not measure any signal does not work. Any information for cluster formation is

therefore the energy measured by a cell relative to the expected electronic noise measured by that cell [9]:

$$\zeta_{\text{cell}} = \frac{E_{\text{cell}}}{\sigma_{\text{noise,cell}}} \tag{5.6}$$

There are three ζ_{cell} thresholds used: seed ($\zeta_{\text{cell}} = 4$), neighbor ($\zeta_{\text{cell}} = 2$) and edge ($\zeta_{\text{cell}} = 0$). Any cell passing the seed threshold starts a cluster. As long as neighboring cells (next cell in the same layer or overlapping cell in the next layer) of the cluster passing the neighbor cut are found they are added to the cluster. This is done repeatedly and can also lead to merging clusters. The last set of cells that are added have to just pass the edge requirement, but this is not done repeatedly. To avoid the formation of large clusters originating from several objects clusters with two maxima (cells with $E_{\text{cell}} > 500\,\text{MeV}$ without neighboring cells with higher energy) are split [154, 155].

5.4.3 Electrons

Electrons are light charged particles and as such they are expected to leave tracks in the tracking detectors of ATLAS and additionally start an electromagnetic shower which deposits its energy in the electromagnetic calorimeter. To reconstruct this signature tracks are extrapolated and matched to electromagnetic clusters. In further steps their energy needs to be calibrated and additional measures are created to be able to select good quality electrons, which will be described in the later subsections. This section is largely based on [156–158].

Electromagnetic clusters which are used for electron reconstruction consist just from the energy deposited in the electromagnetic calorimeter. This energy has to be at least 400 MeV and has to be the majority of energy measured in all calorimeters. Tracks are matched to these clusters if they coincide in η and ϕ within a small threshold. If several tracks match the requirements the better quality track which is closer in ΔR is chosen. The final step in electron reconstruction is the formation of superclusters. Seeds for superclusters are formed by clusters which have a matched track and at least 1 GeV, starting with the highest E_T cluster of the event. Any other clusters in a window of $\Delta\eta < 0.075$ and $\Delta\phi < 0.125$ are added to that supercluster. (In case the track of the seed cluster is also associated to a different cluster that cluster might be added with a relaxed distance requirement.) Each supercluster gets matched to tracks in the way described before and forms an electron candidate.

Calibration

From the supercluster and the track all properties of an electron are determined. The track delivers the most accurate information about charge and direction. The most accurate information about the energy is given by the calorimeter information. However, this information is not just able to be readout, but needs to be calibrated. This calibration procedure starts with recalibrating the energy of the clusters themselves, and continues with a multivariate regression algorithm performed by a BDT that was trained on simulated data. This calibration only uses information from the cluster response and is performed in several $|\eta|$ and E_T bins.[10] In a second step the differences between the simulated and physical detector are tackled by performing a calibration just on data. This includes an intercalibration between different regions of the detector, namely the first and second layer of the calorimeter in different p_T bins. Other effects that are

---

[9] $\sigma_{\text{noise,cell}}$ is estimated for each cell and year separately.

[10] E_T is defined equivalent to p_T using the calorimeter information. Therefore, it is more accurate to bin calorimeter information in terms of E_T, although it can almost directly be translated to p_T.

corrected are mismodelling of the presampler energy and pile-up influence. All of these steps scale the measured energy with the goal of arriving at a more accurately measured energy; this is called electron energy scale. In a third step the difference between the energy of electron in data and MC is measured. Therefore, simulated $Z \to ee$ events are compared to data with two electrons which each have an energy of $E_T > 27$ GeV (this sample is dominated by $Z \to ee$ events). The resulting difference in scale is then applied (multiplied with the energy) to data (as the energy of electrons in MC is known).

Considering the fluctuations of the showering and measurement process it becomes clear that two electrons of the same energy do most likely not have the same measured energy. Some calibration steps like the intercalibration mitigate some of those effects. However, a certain spread of the measured energy remains. This spread is modelled by a Gaussian for which σ is the defining width, this is called electron energy resolution. It is measured together with the data/MC correction in $Z \to ee$ samples. The resulting difference in resolution is then applied (added in quadrature) to MC, as the goal is to describe the resolution of data in the simulated samples.

Photons themselves are not of interest in this thesis as the target process does not contain photons. However, photons are reconstructed in the same way as electrons just without the track requirements. Therefore, it is in principle possible to mistake a photon for an electron if the track-matching procedure failed. Additionally, photons can interact with the detector material in the tracking system and create an electron-positron pair, called converted photon. The signature of this is similar to an electron. Reconstruction aims at enabling a clear identification, but necessarily fails sometimes.

Working points

To distinguish electrons from converted or unconverted photons or other objects as jets a likelihood is defined of the electron candidate being an electron. For that several variables concerning the track (for example number of hits in different layers), the cluster (for example the energy ratio between different calorimeter layers) and the track-cluster matching (for example ratio between the energy from the cluster and momentum from the track) are used. For each of the variables a histogram of signal and background events (data which contains mostly $Z \to ee$ and data which contains mostly dijet events) in E_T and $|\eta|$ is created (and smoothed) to create a PDF. Those PDFs are multiplied for each bin to derive two likelihoods L_S and L_B. The discriminant

$$d_L = \ln\left(\frac{L_S}{L_B}\right)$$

can be cut on different values to derive several working points (WPs). Those working points can be higher to ensure an electron candidate is likely to be an electron (called tight) or lower to ensure more true electrons pass the selection (loose). For electrons additional cuts on singular variables are applied to distinguish the WPs. For example, a Loose WP which 93% of true electrons averaged over all p_T and bins do pass (93% efficiency) is defined and a LooseAndBLayer WP which uses the same cut on d_L but the additional requirement of a hit in the b-layer of the tracking detector. Two other WPs Medium with 88% efficiency and Tight with 80% efficiency are defined. While their efficiency is lower, the background rejection (percentage of background events failing the cut) increases by a factor of about 2 and 3.5 respectively. The electron identification generally works better for higher p_T electrons. Therefore, for a given WP the efficiency and rejection is higher for higher p_T. Such, electron candidates on a given WP with low p_T are more likely to originate from another object. This is called a fake electron.

Prompt leptons

An additional requirement is made on the distance to the primary vertex to ensure used electrons are originating in the PV. This requires $|d_0|$ to be at maximum 5 times as large as the uncertainty on the beamline, while the distance $|z_0 \sin \theta|$ can be maximally 0.5 mm.

To reject electrons originating in the primary interaction (prompt) from electrons that originated in a hadronic jet (see section 5.4.5) additional isolation requirements can be applied. Specifically built to reject against jets originating in heavy quarks (c or b) the PLIV (PromptLeptonImprovedVeto) utilizes a BDT to combine several already complex variables that are related to an electron being isolated. This includes the sum of p_T from tracks in a cone of $\Delta R = 0.3$ around the object and the equivalent for clusters. Another variable is itself the output of an RNN that aims at distinguishing prompt electrons using track information. Just two PLIV WPs are available: PLIVTight and PLIVVeryTight such as for looser WPs older methods as the regular PromptLeptonVeto [159] can be used.

Electron charge flip

Another misreconstruction that might affect the analysis is the wrong association of an electron charge. A wrong assignment can occur when electrons interact with the tracking detector (similar to converted photons) and create photons and electron-photon pairs, which disturb the charge measurement. Alternatively, as at high p_T the trajectory gets straighter, it can happen that small disturbances of the track lead to a wrong estimation of the curvature. To mitigate that effect a BDT is trained on electron information and a Loose WP (97% efficiency) is not allowed to be passed for any electron in the analysis.

Each step of the reconstruction and calibration procedure is estimated together with appropriate uncertainties for that particular step. Some of these are of statistical nature, as the corrections are derived from finite samples. Others are of systematic nature as some assumptions are made in the derivation of the corrections, for example in the choice of MC simulation. To make the application of uncertainties easier, several uncertainties concerning the same objects might be combined if it is possible to neglect the influence of correlations. The application of all uncertainties follows the strategy of fit systematic uncertainties described in chapter 7. A list of systematic uncertainties related to electron reconstruction can be found in table B.2.

5.4.4 Muons

Muons are electric charged leptons with a higher mass than electrons. Thus, they leave a track in the detectors, but do not shower completely in the calorimeters. As such they are the only particles which leave a trace in the muon detectors, which form a curved track in the toroidal magnetic field. Additionally, muons are detected by the tracker and leave a small amount of energy in the calorimeters. Due to this diverse detection opportunities muon can be reconstructed using four different strategies:

- Combined: Tracks are build from hits in the muon system analogous to the inner detector [11] and matched to tracks from the inner detector. Properties are determined from the combination of both objects.

---

[11] In each subdetector segments are formed from hits and combined via their rough direction. Ambiguities are resolved and p_T is determined via several X^2 fits.

- Inside-Out combined: Complementary, inner detector tracks can be extrapolated to hits in the muon system, taking into account the energy lost in the calorimeters. This is advantageous in case a muon was not detected in all layers of the muon detector (for example due to low p_T).

- Muon spectrometer extrapolated: In order to reach a higher $|\eta|$ coverage muon system tracks can be extrapolated to the beamline without any inner detector tracks.

- Segment-tagged: In case a track from the inner detector is just able to be matched to a single segment with an angular requirement the inner detector track fully determines the muon properties.

- Calorimeter-tagged: Especially in the region of bad muon system coverage $|\eta| < 0.1$ a muon can also be formed from an inner detector track and energy deposits in the calorimeter consistent with an energetic particle.

Depending on the identification requirement all muon options are reconstructed and taken into account, as long as they are not overlapping. If this is the case a priority is given to the higher quality muons, which would appear higher in the above list. For Medium muons just Combined and Inside-out combined muon candidates are used in the central region ($|\eta| < 2.5$) and additional requirements are applied regarding hits in subdetectors and corresponding momentum in inner detector and muon system. Loose muons can also contain Segment-tagged and Calorimeter-tagged muons and do not have to pass the additional requirements. This increases the efficiency for muons between 20 GeV and 100 GeV from 97% to 99% but decreases the background rejection from 99.83% to 99.75%. A tight WP is available but is not used in this thesis as it decreases the efficiency to about 93%.

To ensure muons originating in the PV d_0 divided by the beamline uncertainty is required to be smaller than 3 and $z_0 sin\theta < 0.5$ mm. To ensure they are not originating in jets an improved prompt lepton veto is built. This is done in the same way as for electrons, but everything including the training of the RNN is adjusted for muons specifically.

As muons do not shower significantly in the calorimeter, their energy is not a basic variable. Instead, the muon momentum is calibrated. This is done in bins of η and ϕ following the specifics of the inner detector, muon spectrometer and toroid magnet. The scale correction accounts for inaccuracies on the magnetic field (s_1) and the energy loss in front of the magnet system s_0 (for example in the calorimeters). The momentum resolution is expected to come in three terms which are added in quadrature:

$$\frac{\sigma(p_T)}{p_T} = r_0 \frac{1}{p_T} \oplus r_1 \oplus r_2 p_T \tag{5.7}$$

The first term denotes the effects which are independent of p_T (note the fraction on the left-hand side of the equation) as energy loss in the not active parts of the detector. The second term is scaling with p_T and is caused mostly by multiple scattering and local effects. The third term scaling quadratic with p_T is able to correct for residual mis-alignments. All the effects have been measured in simulated Z boson decays separately for inner detector and muon system and as such are also corrected for separately. The momentum correction is applied to the simulated datasets via the following (simplified) formula:

$$p_T^{corr} = \frac{p_T + s_0 + s_1 p_T}{1 + \frac{r_0}{p_T} g_0 + r_1 g_1 + r_2 p_T g_2} \tag{5.8}$$

where g_x are random variables sampled from a Gaussian with mean 0 and width of 1 to extend the simulated resolution to the resolution in data [153].

An overview over the systematic variations invoked by the described calibration steps can be found in table B.3.

5.4.5 Jets

Jets are complicated. Therefore, this section can only give an overview of the signature of jets, jet algorithms and finally the jet calibration procedure and further jet properties. Quarks and gluons carry color charge and as described in chapter 2 unpaired color charges are confined to each other by an energy scaling with r^2. For energetic partons traversing away from the beam this energy surpasses the threshold which is necessary to create additional quarks. This leads to a multitude of additional low energetic quarks and gluons which are emitted roughly along the parton trajectory.[12] The color charged quarks in close proximity recombine to mesons, from which some might decay. This fragmentation process cannot completely be described in simulation yet, but different parton shower models are in use which, after tuning with data, can model jets decently (see for example in section 4.1). The signature of multiple mesons and their decay products roughly along a trajectory contains one or more tracks in the inner detector and energy deposits in the electromagnetic and hadronic calorimeters. Therefore, the largest challenge is to find the measured tracks and clusters which originate in the jet and not pile-up or noise. Candidates are both well measured tracks and clusters, with the important requirement of not double counting energy that has been measured in both detectors (for example charged mesons). Therefore, tracks are matched in ΔR'[13] and energy-momentum ratio to clusters. In some cases a single cluster contains the energy coming from one track, in other cases several clusters are attributed to one jet. For each matched track the energy measured in the track is subtracted on a cell level from matched clusters. It can happen that not all the energy of a cluster can be attributed to a track. Smaller remnants are removed, bigger remnants are kept as their own objects. The objects remaining after this procedure are called Particle Flow objects and should describe the event without any double counting of energy [160].

A jet is going to contain several objects in a small ΔR around the originating quark. However, jet reconstruction algorithms which look for high p_T seed jets and model a cone around them (called cone algorithms) have been found to overestimate the energy in pile-up rich environments among other problems. Instead, sequential recombination algorithms are used which repeatedly iterate over the objects and combine objects closest to each other according to a p_T-dependent metric:

$$d_{ij} = \min\left(p_{T,i}^x, p_{T,j}^x\right) \frac{\Delta R_{ij}^2}{R^2} \tag{5.9}$$

$$d_{iB} = p_{T,i}^x \tag{5.10}$$

The most commonly used jet finding algorithm uses $x = -2$ (anti-k_t algorithm) and therefore prioritizes objects with high p_T. It creates almost circular jets without the need to manually choose seed jets and split cones of nearby jets. As jets are fully reconstructed when d_{iB} is the smallest, R plays the effective role of the jet radius [161, 162]. For jets built from particle flow objects $R = 0.4$ is the common choice in ATLAS. These jets are calibrated for an energy range of 20 to 1 500 GeV and are expected to originate in gluons and quarks (except the top quark) from the primary interaction [160]. Therewhile, bigger jets ($R = 1$) are calibrated at higher energies (200 to 2 000 GeV) and often originate from boosted objects like top quarks

---

[12] Some quarks and gluons may also be emitted between the two original colored objects. However, this information is not recovered and used in this thesis.

[13] There are some technicalities as to clusters not having a well-defined center, which is accounted for by the distance used.

[163].

Jet calibration

As jet measurements can vary widely (for example due to pile-up) the jet calibration consists of more steps than for the more consistent electrons and muons. An overview over the calibration sequence can be found in figure 5.14. In a first correction the average pile-up density for each event $\left\langle \frac{p_T}{A} \right\rangle$ is measured (using jets clustered with an alternative $x = 2$ which prioritizes low p_T objects in $|\eta| < 2$) and subtracted scaled to the jet area. Additional dependencies are corrected for dependent on p_T and $|\eta|$: The number of interactions during that event (in-time pile-up) is modeled by the number of reconstructed vertices N_{PV} and the number of interactions per bunch-crossing by $\langle \mu \rangle$ in that luminosity block. In total one arrives at the following corrections:

$$p_T^{corr} = p_T - \left\langle \frac{p_T}{A} \right\rangle A - \alpha \left(N_{PV} - 1 \right) - \beta < \mu > \tag{5.11}$$
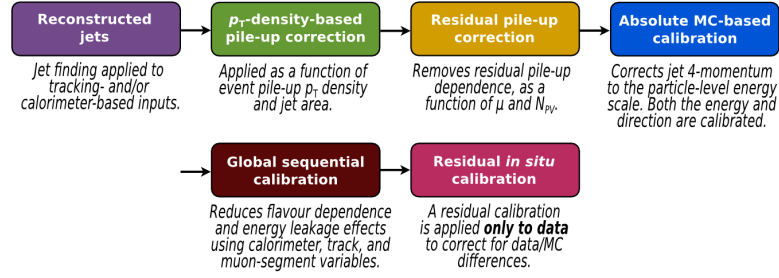
Figure 5.14: Overview of the jet calibration sequence [164].

This corrected p_T needs to be calibrated to the original energy of the depositing particle, as the ATLAS calorimeters only measure a subset of the shower energy (see section 5.2.3 and section 5.4.3). Additionally, energy loss in dead material and due to not perfect jet-reconstruction are compensated that way. To bring the jet energies to the level of their true counterparts, jets in simulated events are matched to their counterparts before detector simulation (truth jets) and the mean of the Gaussian response is corrected for in bins of p_T and $|\eta|$. Even after this correction one finds differences between jets and the accuracy of their measured energy. These differences are associated with the large variety of jets which are in turn associated with different originating partons (quarks or gluons). The correction is therefore done dependent on several jet-shape variables in succession: energy fraction carried by tracks, energy deposited in the first layer of the hadronic calorimeter, the third LAr layer, number of tracks, average distance of tracks with respect to the jet center (jet track width) and number of segments in the muon system which can be matched to the jet. All of them dependent on p_T and $|\eta|$ and form the Global Sequential Calibration (GSC). A last correction is done to correct for the differences between data and simulation (compare section 5.4.3) where jets are balanced against another object in the same event. It is called insitu calibration. This is done both in data and simulated events and the ratio between the responses is averaged and applied to data. In a first step dijet events are used to correct all jets in $0.8 > |\eta| > 4.5$ to the response of central jets ($|\eta| < 0.8$). Second, central jets are balanced against well measured Z bosons and photons. Third, residual pile-up effects in data are calibrated by balancing one high-p_T jet against several low-p_T jets.

Again, all those calibration steps create a large amount of systematic uncertainties. To reduce their

number they are partially combined to the independent variations that are listed in table B.1.

Jet energy resolution

All the previously described steps aim at having the best possible energy measurement and arriving at a preferably narrow energy distribution by mitigating differences between jet types and environments. The remaining resolution is described by three terms:

$$\frac{\sigma p_{\mathrm{T}}}{p_{\mathrm{T}}} = \frac{N}{p_{\mathrm{T}}} \oplus \frac{S}{\sqrt{p_{\mathrm{T}}}} \oplus C \tag{5.12}$$

The first term N describes the noise effect and as such is not dependent on p_{T}. S describes the stochastic term and is dependent on $\sqrt{p_{\mathrm{T}}}$, which is easy to follow if one keeps in mind that the p_{T} of a jet is roughly linearly correlated to the number of particles created in the calorimeter showers which can be described by a Poisson statistic. The third term C is a constant fraction of p_{T} and captures effects like energy lost in dead material. All three terms are measured in dijet events sorted in p_{T} by balancing the two jets against each other and fitting the width of the asymmetry:

$$\mathcal{A} = \frac{p_{\mathrm{T}}^{\mathrm{probe}} - p_{\mathrm{T}}^{\mathrm{ref}}}{p_{\mathrm{T}}^{\mathrm{avg}}} \quad \sigma_{\mathcal{A}} = \left\langle \frac{\sigma p_{\mathrm{T}}}{p_{\mathrm{T}}} \right\rangle^{\mathrm{probe}} \oplus \left\langle \frac{\sigma p_{\mathrm{T}}}{p_{\mathrm{T}}} \right\rangle^{\mathrm{ref}} \tag{5.13}$$

This asymmetry contains the width of both jets as well as the asymmetry between the objects before producing the jets (particle level). The second effect is corrected for by the width of asymmetry between truth jets, the first effect by halving $\sigma_{\mathcal{A}}$ for events with two well measured jets. As jets in the central detector region $0.2 < |\eta| < 0.7$ are measured the best, dijet events with both jets in this region are used for this step. Other $|\eta|$ regions are always balanced against this region and their jet energy resolution is found by subtracting the central resolution from the asymmetry width. As this is done in bins of p_{T} especially the S and C terms can be measured by this approach. The p_{T} independent noise term N is fixed by an additional measurement. Here instead of reconstructed jets, random cones with a similar area to the jets are pointed somewhere in the events and balanced against each other.

All steps are done independently for data and simulation samples to achieve the most accurate description (the subtracted truth imbalance is of course always taken from simulation). However, this imposes a challenge as for low p_{T} jets the resolution for simulated jets is found to be smaller than for data. Contrary to other objects a consistent description can thus not be found by smearing the simulated dataset. Smearing the measured data and thus worsening the measurement to the level of the best description in our simulation is also not desirable. To solve this dilemma no smearing is performed in those regions and a systematic uncertainty using pseudodata (a variation of simulated events) is introduced to cover this effect [164]. Additionally, the number of uncertainties is reduced by just taking the principle components of the joint uncertainty space. Therefore, the systematic uncertainties related to the jet energy resolution listed in table B.1 cannot be associated with any direct sources.

JVT

Reconstructed and calibrated jets can still originate in pile-up events, which happened to produce a high p_{T} jet (up to $60\,\mathrm{GeV}$). To mitigate that effect the jets are tested to originate from the primary vertex, this is called jet vertex tagging (JVT). In the central region where tracks are available ($|\eta| < 2.5$) this relies heavily

on the tracking information. A two-dimensional likelihood is built from the p_T ratio of tracks originating in the PV with respect to the reconstructed jet p_T and the corrected jet vertex fraction, a variable which aims at finding the ratio of tracks which originated in the primary vertex with respect to all other vertices while correcting for the pile-up effects. On this likelihood two WPs are available: medium and tight which both maintain 96% (97%) jet efficiency [165]. In the $|\eta|$ region which extends past the inner detector no tracks are available to distinguish the origin of a jet. Instead, the missing transverse momentum (see section 5.4.8) neglecting this jet is calculated for every pile-up vertex (see section 5.4.1). Then the jet momentum is projected onto this vector; if it aligns it probably originated from this vertex and not contributing to the calculation caused the E_T^{miss} for this vertex. The maximum of these projections is not allowed to be over 40% for the tight WP and 50% for the loose WP [166].

b-tagging

As described above, jets can originate from any quark or gluon (except the t if not heavily boosted) but those jets do not all look the same. A jet originating from a b- or c-quark for example still contains that heavy quark. Since this quark has a longer lifetime it will form heavy mesons. These mesons alter the form and energy distribution inside the jet. This information can be utilized to find jets originating from heavy quarks, especially the b which are unlikely to originate from pile-up, but instead might be a sign of a t or H decay. To differentiate the jet origin a categorical NN is created which has three output categories: b-jet, c-jet and light jet (a jet originating from a u, d, s or g). To combine the strength of all outputs WPs are defined by cuts on

$$D_{\mathrm{DL1r}} = \ln \left(\frac{p_b}{f_c p_c + (1 - f_c) p_l} \right) \tag{5.14}$$

where $p_{b,c,l}$ denote the three different NN outputs and f_c the fraction of c-jets in the background training sample, which is used to correct for different c fractions in the samples. The input variables available to the NN are itself high level variables: an algorithm which aims at finding secondary vertices in the jet, one fitting tool which tries to find the whole b decay chain, a likelihood based algorithm that takes the impact parameters of associated tracks into account and a complete RNN which aims at doing the same but is able to take the correlations between tracks into account. The WPs defined on the final DL1r combination of all these information sources result in an efficiency of 85%, 77%, 70% and 60% as measured in $t\bar{t}$ events [167, 168].

5.4.6 Hadronic taus

τ leptons are the heaviest leptons and have a mean lifetime of 290×10^{-15} s which corresponds to an averaged traveled distance of 87 μm in vacuum. Thus, the possible decay channels into light leptons (about 17.5% each) cannot be distinguished from prompt electrons and muons. From the reconstruction perspective only hadronically decaying τ leptons τ_{had} are considered separately. The characteristic decay of a τ goes into one or three charged mesons ($\pi^{\pm}$ or $K^{\pm}$) in combination with a varying number of neutral particles (ν_τ and light mesons). Therefore, the expected signature contains one or three tracks in the inner detector and a shower in both calorimeters.

τ_{had} candidates are seeded in jets reconstructed with the anti-k_t algorithm with a width of $R = 0.4$ taking as input only clusters which underwent an additional local energy calibration [169]. In a range of $\Delta R = 0.2$ around the jet barycenter all tracks with a $p_T > 1$ GeV are used to determine the tau vertex [170,

171]. Any jets within ΔR =0.4 are classified via a BDT as core and isolation tracks. Core tracks define the 'prongness' of τ_{had}. Just τ_{had} with 1 or 3 tracks (1 or 3 prong) are considered in the further reconstruction as a τ is expected to decay in exactly one or three charged mesons [172]. The momentum fraction carried by isolation tracks is one of the high-level input variables fed into an RNN which aims at classifying τ_{had} against quark- and gluon-jets. Additional to the high-level variables the RNN also uses low-level variables from the tracks and clusters associated to the τ_{had} candidates. This information is first processed separately, then combined and fed through three further layers of the NN. The output score performs significantly better to separate true τ_{had} from jets than for the previously used BDT and generally better for 1-prong τ_{had} than 3-prong (1- and 3-prong are trained separately). Background rejection against the efficiency for both methods is compared in figure 5.15. The indicated and available WPs are Very Loose, Loose, Medium and Tight ranging from 95% to 45% in signal efficiency [173].
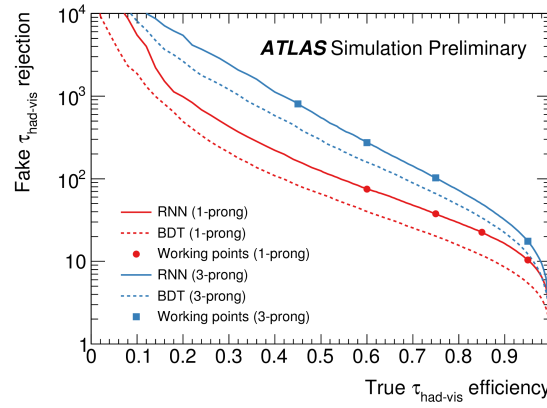
Figure 5.15: Efficiencies and rejection power of the τ_{had} selections using the used RNN and the old BDT for 1- and 3-prong τ_{had} [173].

A subdominating background for 1-prong taus are electrons which have been misreconstructed as τ_{had}. Therefore, an additional BDT is trained just to discriminate between τ_{had} and electrons [173].

Tau energy scale

Similar to jet the τ_{had} energy needs to be calibrated. This is done in four steps. The first of which is applied directly after reconstruction and adjusts the calorimeter measurement. The second step subtracts energy from pile-up (modeled as a function of N_{PV}, $|\eta|$, prongness and energy). Other effects like missing clusters are compensated by scaling the energy to the true τ energy (determined as the mean of a Gaussian fit for different energy, $|\eta|$ and prongness). This method works well for energetic τ_{had} (p_{T} >200 GeV), but τ_{had} with a lower p_{T} benefit from an additional regression performed with a BDT [172, 174]. τ_{had} are more distinctive than general jets, therefore the number of systematics that can be seen in table B.4 could be reduced to a smaller number.

5.4.7 Overlap removal

All described reconstruction mechanisms are performed independent of each other. This can lead to an event where one object has been reconstructed as two different objects. One example would be a muon which left some energy in the calorimeters and could be reconstructed as a τ_{had}. To avoid double

counting objects an overlap removal is performed, which removes all objects which overlap with an already reconstructed object. Priority is given to the objects which have a clearer signature. This leads to the removal of the following objects:

1. muons which share a track with an electron and vice versa
2. jets within $\Delta R < 0.2$ of an electron
3. electrons within $\Delta R < 0.4$ of a jet
4. jets with less than three tracks that are within $\Delta R < 0.2$ of a muon
5. jets with less than three tracks that share a track with a muon
6. muons within $\Delta R < 0.4$ of a jet
7. τ_{had} within $\Delta R < 0.2$ of a loose electron
8. τ_{had} within $\Delta R < 0.2$ of a muon is removed if the muon has a $p_T > 2\,\text{GeV}$ (in case the τ_{had} has a $p_T > 50\,\text{GeV}$ this only applies if this is a combined muon)
9. jets within $\Delta R < 0.2$ of a τ_{had}

All overlap removal is done with loose objects.

5.4.8 Missing E_T

Some particles like neutrinos are not expected to be measured with the ATLAS detector. However, taking into account that the event needs to be balanced in *x-y* direction one can get some information about those particles by measuring the imbalance. This imbalance is measured with focus on energy and called E_T^{miss}.

Besides not measured particles, some detector effects can also contribute to E_T^{miss}:

- energy lost in passive materials of the detectors
- mis-measurement of objects
- particles lost through regions that are not covered by the detector
- over- or underestimation of pileup
- irregularities in the detector
- double counting of a track or cluster in different objects

To cover all these effects E_T^{miss} is measured as the vectorial p_T sum of all reconstructed objects pointing to the PV (hard term) plus the tracks that originate from the PV and have not been taken into account for any object reconstruction (soft term). The challenge for the hard term is to not double count a track or cluster in two different reconstructed objects or miss an object due to too harsh overlap removal. Therefore, the overlap removal for the E_T^{miss} calculation is done separately from the one done for the analysis [175]. The soft term on the other hand is built by adding all tracks with $p_T > 0.5\,\text{GeV}$ that are not associated with any other object (to avoid double counting) and point to the correct vertex (to mitigate the impact of pile-up). Depending on the JVT different WPs for E_T^{miss} are available, the performance for events with a high number of jets is best for the tight WP [176].

In addition to the vectorial E_T^{miss} ($\sum \vec{p}_T$) also the sum of all p_T in the event ($\sum p_T$) is calculated ('sumet') with the same object selection. This gives a measurement of the overall activity in the event.

The systematic uncertainties for E_T^{miss} included in table B.6 are obtained by comparison with simulated datasets. The main scale is just allowed to vary the size of E_T^{miss}, not the direction. However, the resolution is given in parallel and perpendicular to the calculated E_T^{miss} [175].

Preselection

After all, is it not the way we humans shape the universe, shape time itself? Do we not take the raw stuff of chaos and impose a beginning, middle, and end on it, like the simplest and most profound of folktales, to reflect the shapes of our own tiny lives?

– Tad Williams, *Otherland: Sea of Silver Light*

The part of the analysis which specifically targets the *tHq* process is a group effort. Especially the central preprocessing is a task to which I contributed. From the more specialized tasks the focus of the description lies on the tasks I have worked on myself. Other processes like the fake estimation and NN application are just briefly described.

6.1 Centralized Preprocessing

The goal of this thesis is to find *tHq* events in the collected ATLAS dataset. The first step in finding them is to define a subset of events which is more likely to contain *tHq* events. The reduced number of events simplifies the further search. This step is done within a large framework used by several analyses which look for processes with a similar final state (all of them containing *t*-quarks) called 'AnalysisTop'. It is configurable to select the events and objects needed by each analysis. Additionally, it applies the correct configurations like scale factors, systematic uncertainties and overlap removal as described in section 5.4 and provided by the whole ATLAS collaboration. This analysis uses 'SingleTopAnalysis' which is build upon AnalysisTop and offers a few additional features as well as containing the configuration file for AnalysisTop. This step is run centrally for multiple analyses on the grid (see section 5.3.2) as the datasets are too big for local processing. The output (SingletopNTuples) is then transferred to the local site in Bonn.

The biggest background of events at the LHC consists of pure hadronic signatures originating from pile-up or pure hadronic interactions. To exclude those events, a signature which contains at least one electron or muon was chosen. These objects are well discriminated against jets. Consequently, the triggers which require at least one electron or muon are allowed to have large trigger rates (see section 5.3). Therefore, this analysis keeps events which pass any of these triggers. All other events are neglected.

To further limit the number of stored events, a number of reconstructed objects of leptonic origin are required to store the events. These include electrons, muons or τ_{had}. The event has to hold a combination of two or three of these objects. Additionally, at least one reconstructed electron or muon has to have at least 27 GeV (which is needed to pass the trigger threshold).

The conditions of data taking vary (see for example section 5.1). To control this any events which are stored need to be found in a list of runs with high data quality and no known issues [132]. To make sure the simulated datasets mimic the conditions of data taking the simulation is subsequently reweighted. This means that the simulation weight (see section 4.1) of each simulated event is multiplied with an additional weight. This weight is higher for run conditions (in terms of $\langle\mu\rangle$ or μ) which were underrepresented in the simulation with respect to data and lower for conditions which were overrepresented in the simulation [177].

To reduce the size of the output files, only objects which are needed for the analysis and pass basic quality criteria are stored. This leads to omitting all photon and large radius jet related information. Electrons, muons, τ_{had}, small radius jets and $E_{\text{T}}^{\text{miss}}$ are stored. The required working points can be found in tables 6.1 to 6.5. Note that for every objects two sets of requirements are given called 'loose' and 'tight'. In a first step only the 'loose' requirements have to be fulfilled, the application of the 'tight' working points is described in section 6.2.

WP	loose	tight
p_{T}	10 GeV	s
$|\eta|$	$0<|\eta|<1.37, 1.52<|\eta|<2.47$	s
Identification	LooseAndBLayerLH	TightLH
Isolation	-	PLIVTight

Table 6.1: Requirements imposed on electrons passing the loose and tight selection. If no requirement is applied this is marked with an '-', if the tight selection does not differ from the loose selection this is marked with an 's'.

WP	loose	tight
p_{T}	10 GeV	s
$|\eta|$	<2.5	s
Identification	Loose	Medium
Isolation	Loose	PLIVTight

Table 6.2: Requirements imposed on muons passing the loose and tight selection. If no requirement is applied this is marked with an '-', if the tight selection does not differ from the loose selection this is marked with an 's'.

WP	loose	tight
p_{T}	20 GeV	s
$|\eta|$	$0<|\eta|<1.37, 1.52<|\eta|<2.5$	s
Identification	RNNLoose	RNNMedium
electron BDT	Loose	s

Table 6.3: Requirements imposed on τ_{had} passing the loose and tight selection. If no requirement is applied this is marked with an '-', if the tight selection does not differ from the loose selection this is marked with an 's'.

Another way AnalysisTop manages the size of its outputfiles is by narrowing down the variables stored in its output files. For objects only the reconstructed information for both given WPs is stored by default. To allow more flexibility, additional WP information is stored via SingleTopAnalysis. The framework also

WP	loose	tight
p_T	20 GeV	s
$|\eta|$	<4.5	s
fJVT	tight	s

Table 6.4: Requirements imposed on jets passing the loose and tight selection. If no requirement is applied this is marked with an '-', if the tight selection does not differ from the loose selection this is marked with an 's'.

Objects	tight
WP	tight
min jet p_T	30 GeV
use fJVT	yes

Table 6.5: Requirements imposed on E_T^{miss}. There is no differentiation between loose and tight.

stores 'truth'-information for reconstructed objects in simulated samples. This refers to object information before detector simulation and subsequent reconstruction. With this information the accuracy of the reconstruction can be tested.

One example for this is the variable 'tau_true_pdg'. This can be used to assess whether a reconstructed τ_{had} originated from a hadronically decaying τ in the simulation or if the simulated object was something else (for example an electron) and the reconstruction put a τ_{had} in its place. To achieve this, the set of simulated objects is scanned for a τ in the vicinity of the reconstructed τ_{had} first, from highest to lowest p_T. If it is closer than $\Delta R = 0.2$ the truth τ is believed to be the origin of the reconstructed τ_{had}. Then the set of simulated muons is scanned. Only if any μ is close enough (again $\Delta R = 0.2$) and it has a p_T higher than the previously matched τ, it is taken as the new 'best origin'. Subsequently, this is done for truth e and jets. This procedure ensures to find the best match to a 'hard' object (meaning an object that originated from the hard scatter) while giving priority to τ to be matched to τ_{had}. An identifying number (PDG ID) of the matched object is stored in the variable 'tau_true_pdg'. Other information on this object like the p_T and origin of the truth object is stored, too. Similar procedures are applied for reconstructed electrons and muons.

By storing the information about truth objects related to reconstructed objects, studies on reconstruction performance are facilitated. However, any information about the relationship of the truth particles and the particles they originated from (for example a b originating from a t or H) is lost. To avoid this the 'truth' information of an event can be stored. This information is highly dependent on the process, for example a scan for a H and its decay products is lost work in a tZq sample. Therefore, truth information is only stored for processes the analysis is interested in and only for the respective samples. For this thesis, only the tHq information is relevant and is described in the following. Objects in the tHq final state (see figure 3.2) are a (light) quark, a H, a t and the second b from gluon splitting. For these (and all further described) objects p_T, mass, η, ϕ and if applicable PDG ID are stored. However, for the example of a light quark it becomes apparent that it will radiate off FSR which will alter its properties (especially p_T). To keep all necessary information, both the properties before any FSR (right out of the Feynman diagram) and the information after FSR (last available information before detector simulation) is stored. The distinction is done for all other hadronic objects, too. If the objects decay further, the properties 'after FSR' will be determined right before the decay. One of these decaying particles is the t (it can come as a t or a $\bar{t}$ but

there is no need to differentiate at this point). It decays into a *b* and a *W* where the latter will decay further into two distinct decay products. These are either two quarks or a charged lepton and a neutrino, which are all stored. In case the lepton is a τ it will decay further. However, the information needed from the τ decay just concerns the visible decay products. These are defined as all τ decay products, except the neutrinos. A similar distinction is needed for the decay products of the *H*. It will decay into two decay products, mostly two *b*, two τ or two bosons (*W* or *Z*). In case the decay products are τ, again the visible decay products are stored separately. In case the *H* decays in a *Z* or *W* they will again decay into two decay products. Both can be a τ, whose visible decay products will again be stored separately. All this information becomes crucial in the truth analysis described in section 5.4.

6.2 tHqLoop

To locally process the output of SingleTopAnalysis, a local processing step is utilized to perform the final alterations on the samples. A common framework 'TopLoop' gives the utilities of reading in SingleTopN-Tuples, looping over all events and creating outputs that can be used for the fit. In the output files every value is filled once for each event (called flat NTuples), as opposed to once for each object. Built upon TopLoop the more specialized 'tHqLoop' performs several more channel specific tasks during the loop. This includes the application of tighter object selections, selecting events which are more signal-like and a full event reconstruction.

6.2.1 Cutflow

Task one that has to be tackled is to define a signal channel. This refers to a set of events which are selected for further analysis, often based on a number of reconstructed objects. Some analyses consist of several channels which all aim at enriching the signal with respect to non-signal (background) samples, but are different enough that splitting them is beneficial in further processing the samples; for example, because the background composition is different. In this thesis specifically channels with τ_{had} are considered. As discussed in section 3.1 in the *tHq* process the most likely origin of those is the *H* decaying into two τ. These can decay either both hadronically, resulting in two τ_{had}, or one of them can decay hadronically resulting in one τ_{had}. If both τ decay leptonically there is no τ_{had} originating from the *H*. On the other hand the *W* from *t* can decay into a τ_{had}. However, the probability for this is lower than for a decay into *e* or μ (and significantly lower than purely hadronic decay modes). As described above, requiring *e* or μ (also referred to as 'light leptons') significantly reduces the background and one of them is required to trigger the event. This leads us to three scenarios in terms of leptons which make a possible channel:

- **1 *e* or μ + 1 τ_{had}:** This event selection is just enough to trigger the event and contain a τ_{had} which potentially could originate from the *H*. However, background processes with a larger cross-section like $t\bar{t}$ and *Z*+jets show the exact same signature. Additionally, the rate of falsely identified τ_{had} is relatively high (in comparison to *e* or μ), which allows a fraction of even more background processes like *W*+jets to enter this channel. For these reasons this channel is not considered.

- **2 *e* or μ + 1 τ_{had}:** In this event selection the contribution of background processes with one light lepton can be neglected. Also, the contribution of processes with two light leptons is limited, as an additional fake τ_{had} is required. Any of the two light leptons could have triggered the event. For convenience this channel can be split further:

- **2 *e* or *μ* OS + 1 *τ*_{had}:** In background samples like $t\bar{t}$ and Z+jets the two light leptons have different (opposite) sign. Requiring this, enhances the background. However, this channel still contains a significant fraction of signal events and is thus considered for the analysis. Additionally, events with the *H* decaying into two light leptons and the *t* decaying into a τ_{had} is included here. This enhances the absolute number of signal slightly.

- **2 *e* or *μ* SS + 1 *τ*_{had}:** The benefactor of dividing the $2e/\mu + 1\tau_{\text{had}}$ channel by light lepton charges is the channel with two light leptons carrying the same sign. High cross-section background samples like $t\bar{t}$ and Z+jets enter this channel rarely.

- **1 *e* or *μ* + 2 *τ*_{had}:** This channel maximizes the number of τ that can be expected from a *tHq* decay while requiring one light lepton which has triggered the event. Without falsely reconstructed τ_{had} this signature is rare for any background processes. However, one or both τ_{had} can be falsely reconstructed. Therefore, the channel is contaminated with processes like $t\bar{t}$, Z+jets and even W+jets.

The next object selection considers the number of jets. From the Feynman diagram one can see that at least one non-*b*-tagged jet and one *b*-tagged jet can be expected. Looking at the distribution of jets in the *tHq* sample (figure 6.1) this expectation is met. An event is required to contain at least two jets. At higher jet multiplicity the number of *tHq* events drops. Therefore, at most six jets are allowed in a single event.
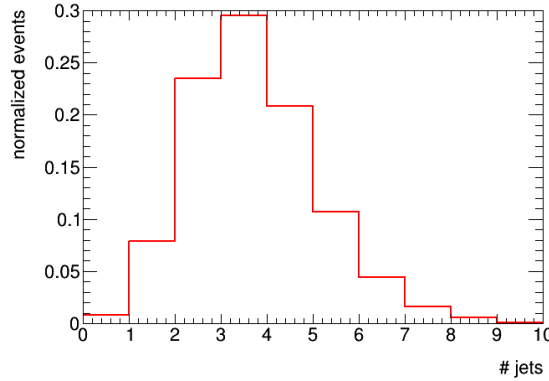
Figure 6.1: Number of jets in the *tHq* sample for all channels combined. All lepton related requirements are applied. The sample is normalized to a total of one event.

Another discriminating variable is the number of *b*-tagged jets. In the *tHq* process at least one *b*-tagged jet is expected. Therefore, this is the main discriminating variable with respect to Z+jets and W+jets events, which are not expected to contain *b*-tagged jets if the boson decayed leptonically. On the other hand $t\bar{t}$ events have a higher *b*-tagged-jet multiplicity. A comparison of the number of *b*-tagged jets in the *tHq* and $t\bar{t}$ sample are shown in figure 6.2. While at higher numbers of *b*-tagged jets per event the $t\bar{t}$ sample is dominating, just allowing one of those jets a high percentage of *tHq* events would be cut. Therefore, the number of *b*-tagged jets is required to be 1 or 2. The *b*-tagging requirements can be found in table 6.6.

To ensure the best quality of the objects every object has to pass the 'tight' selection described in section 6.1. However, in order to study the impact of falsely identified leptons (see section 6.2.4) 'loose not tight' selections are applied. This refers to an object which passes the loose selection, but fails the tight selection. As a consequence of this, the loose selection of one channel might overlap with the tight
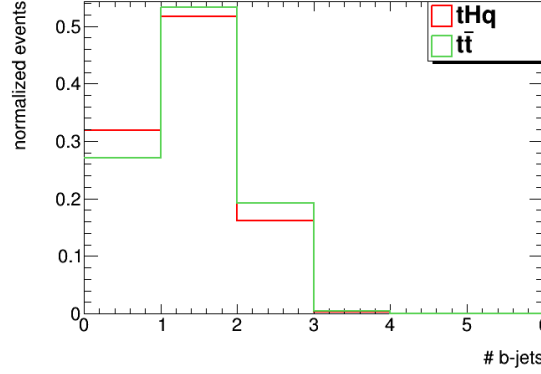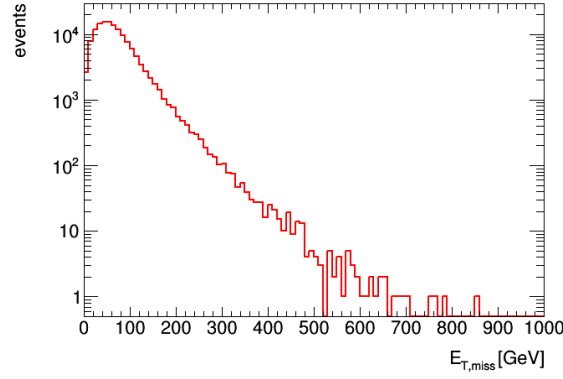
Figure 6.2: Number of jets in the *tHq* sample for all channels combined. All lepton related requirements as well as the requirements on the total number of jets are applied. Both samples are normalized to a total of one event.

WP	loose	tight
p_T	20 GeV	s
$|\eta|$	<2.5	s
btag	Dl1r	s
efficiency	85%	70%

Table 6.6: Requirements imposed on *b*-tagged jets passing the loose and tight selection. Any jet that is not passing the requirements is considered a light jet. If no requirement is applied this is marked with an '-', if the tight selection does not differ from the loose selection this is marked with an 's'.

selection of another channel.[1] To avoid this, an event which contains more τ_{had} than required is vetoed, taking into account loose object selections. The minimum number of required objects are vetoed based on the tight selection. This combination ensures the exact event selection on tight objects plus vetoing additional loose not tight τ_{had}.

Another potentially discriminating variable is the measured E_T^{miss}. As it is mostly originating from neutrinos, it is expected to be sizeable in *tHq* and $t\bar{t}$ events. Whereas, the E_T^{miss} from Z+jets events might approach 0. As figure 6.3 shows, by cutting on a E_T^{miss} >5 GeV the number of processed *tHq* (and $t\bar{t}$) events is almost unchanged whereas the number of Z+jets events is reduced by about 10%. For technical reasons also a maximum allowed E_T^{miss} has to be set. This is set to 800 GeV to keep as many *tHq* events as possible (compare the E_T^{miss} distribution in figure 6.4).

The effect of the above described cuts can be studied in figure 6.5 comparing the fractional impact on the most important samples. More cutflows for more samples can be found in appendix C.

6.2.2 Reconstruction

Since the structure of *tHq* events is known, within this sample it should be possible to reconstruct all event objects (like the *H*) and their properties (like masses). Actually, any event can be treated as if it was a *tHq* event. For events that have a similar structure but different objects (for example *tZq*) these properties

---

[1] This mostly applies to the $2e/\mu$SS + $1\tau_{\text{had}}$ channel and the channel containing just 2 same sign *e* or *μ* which are not described in this thesis.
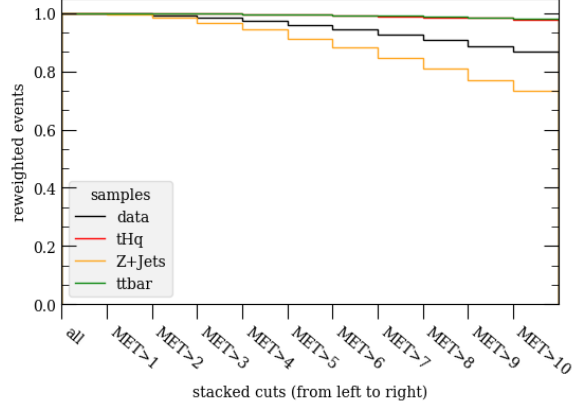
Figure 6.3: Comparison of E_T^{miss} cuts in the most important samples for all channels combined. Here, E_T^{miss} is called 'MET' and given in terms of GeV. All lepton related requirements as well as the requirements on the total number of jets are applied. All samples are reweighted to one event in the first bin.

Figure 6.4: E_T^{miss} in the *tHq* sample for all channels combined. All lepton related requirements as well as the requirements on the total number of jets are applied. With a cut on 800 GeV only one event is lost.
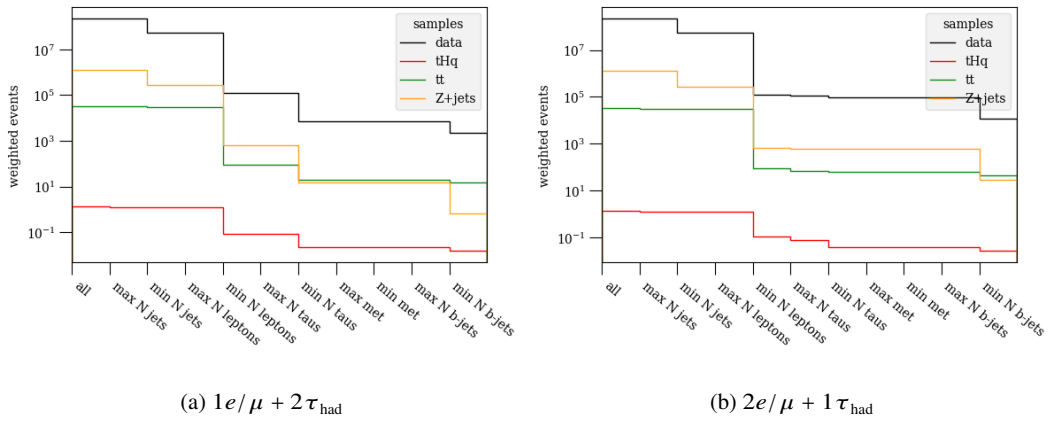
(a) $1e/\mu + 2\tau_{\text{had}}$

(b) $2e/\mu + 1\tau_{\text{had}}$

Figure 6.5: Cutflow on the cuts described in section 6.2.1. The most important samples for the $1e/\mu + 2\tau_{\text{had}}$ channel (a) and the combined $2e/\mu + 1\tau_{\text{had}}$ channels (b) are shown. All samples are weighted to their respective cross-section.

69

(especially the Higgs mass) are expected to be discriminating between the true processes. For events with a different structure (most other backgrounds) these variables are expected to be nonsensical and thus also discriminating. The overall structure of event reconstruction is the same in all three channels, the differences between the channels are discussed within the respective sections.

6.2.2.1 Object association

In a first step, the reconstructed objects need to be associated to either the H, the t or the spectator jet (j^{spect}). The jet association is similar in all channels and has been influenced by the reconstruction in [178]: the leading (highest p_T) b-tagged jet is associated with the t ($j^t_{b\text{-tag}}$, see figure 6.6). The jet which has the largest invariant mass together with $j^t_{b\text{-tag}}$ (which hints at them emerging in opposite direction) is associated with j^{spect}.

$$j^t_{b\text{-tag}} = \max(p_T(b)), b \in (b\text{-tagged jets}) \tag{6.1}$$

$$j^{\text{spect}} = \max\left(\left(j^t_{b\text{-tag}} + j\right)^2\right), j \in (\text{jets}) \tag{6.2}$$

$$m_{b,j^{\text{spect}}} = \text{mass}\left(j^t_{b\text{-tag}} + j^{\text{spect}}\right) \tag{6.3}$$

In the $2e/\mu + 1\tau_{\text{had}}$ channels non-b-tagged jets are preferred as the j^{spect}.
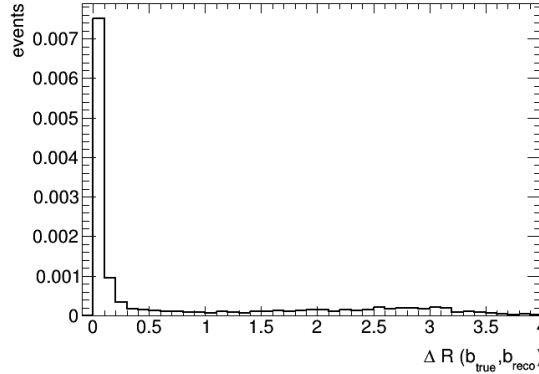
Figure 6.6: Distance between truth b-jet from t and the reconstructed b-tagged jet leading in p_T in the $1e/\mu + 2\tau_{\text{had}}$ channel of the tHq sample.

The next step in the object association are the leptons. The three leptons are expected to originate from the H (two) and the t (one). How the association is done varies largely by channel.

Lepton association in the 1 e/μ + 2 τ_{had} channel   In the $1e/\mu + 2\tau_{\text{had}}$ channel two τ_{had} and one light lepton have to be associated to the H and t. In table 6.7 the expected origin of leptons in the $1e/\mu + 2\tau_{\text{had}}$ channel is shown. The numbers stem from multiplying the different branching ratios with the theoretical cross-section of 73 fb. Neither the $j^t_{b\text{-tag}}$ decay, nor any detector efficiencies are taken into account. Therefore, just the expected percentages can directly be transferred to reconstructed events. The most probable train of events would be the H decaying into τ and those decaying hadronically. This would lead to the association of both τ_{had} to the H and assuming the t decayed into a light lepton. In table 6.7 one can see that (from purely theoretical considerations) this is correct in 86% of the expected events.

		$H \to \tau\tau$	$H \to WW$	$H \to ZZ$
$1e/\mu + 2\tau_{\text{had}}$ events	86.7	76.8	6.0	4.4
$e + \tau_{\text{had}} + \tau_{\text{had}}$	44.6	39.4	3.0	2.2
$\mu + \tau_{\text{had}} + \tau_{\text{had}}$	42.1	36.9	3.0	2.2
% correct assoc.	86.1 %	88.6 %	50.0 %	93.8 %

Table 6.7: Higgs decay origin of $1e/\mu + 2\tau_{\text{had}}$ events in *tHq* from theoretical considerations.

Lepton association in the 2 e/μ OS + 1 τ_{had} channel   In both $2e/\mu + 1\tau_{\text{had}}$ channels just one τ_{had} needs to be associated. Its highest probability origin is the H (see above) so the τ_{had} is always associated with the H. This is favoring the events where H decayed into two τ with opposite signs. Hence, the other lepton associated with the H is the direct decay product of the other τ. Therefore, the sign of that light lepton is opposite to the τ_{had} in the event. As both light leptons in this channel have different signs this association is unique. In table 6.8 it is shown that this association is correct in about 86% of the $H \to \tau\tau$ decays, but just in 72% of all expected events.

		$H \to \tau\tau$	$H \to WW$	$H \to ZZ$
$2e/\mu$OS + $1\tau_{\text{had}}$ events	66.5	42.5	20.7	3.3
$e + e + \tau_{\text{had}}$	17.5	10.8	5.2	1.5
$\mu + \mu + \tau_{\text{had}}$	16.8	10.4	5.1	1.2
$e + \mu + \tau_{\text{had}}$	32.1	21.3	10.3	0.5
% correct assoc.	72.2 %	86.4 %	50.0 %	28.9 %

Table 6.8: Higgs decay origin of $2e/\mu$OS + $1\tau_{\text{had}}$ events in *tHq* from theoretical calculation.

Lepton association in the 2 e/μ SS + 1 τ_{had} channel   As described above the τ_{had} in this channel is associated with the H. However, the association of the light leptons via charge cannot work (as light lepton charges are the same). Different approaches have been tried: assuming the leading light lepton originated from t [179], cutting on m_{reco}^{H} and m_{reco}^{t} (defined in section 6.2.2.2 and section 6.2.2.3) and finally a BDT. As the association depends on reconstruction level variables the accuracy cannot be investigated on a branching ratio level. Instead, the truth variables described in section 6.1 are used to match reconstructed and truth leptons if their distance is small: $\Delta R \leq 0.01$. The events where this is possible (86%) build the baseline for accuracy comparisons between the different methods. The first assumption only results in an accuracy of about 61%. The more advanced cut technique achieves an 84% accuracy. In comparison, the BDT achieves even 88% accuracy on reconstructable events. Therefore, the BDT method is used in this thesis and described in more detail.

The variables utilized by the BDT are m_{vis}^{H} (both options, as at this stage it is not decided which is the correct association) and ΔR as well as $\Delta \eta$ between the different objects (light leptons, τ_{had} and b-tagged-jets). More variables have been tested, but in order to avoid instabilities just the most influential are used. The target of the BDT is to separate events in which the leading light lepton originated from the t (see hypothesis one) from the events where it originated from the H. In contrast to other MVA applications, a cut on the output does not determine whether the event is signal-like or not. It just determines which set of variables are kept and thus how well the association and therefore the reconstruction is working. This

determines how meaningful the output variables are in the signal sample. In table 6.9 the percentages of events that are correctly classified via the BDT are shown, under the assumption that all $2e/\mu\text{SS} + 1\tau_{\text{had}}$ events contribute equally.

	Total	correct assignment	wrong assignment
leading e/μ from top	56.3 %	85.6 %	14.4 %
subleading e/μ from top	42.6 %	74.4 %	28.2 %
total		79.9	20.1 %

Table 6.9: Percentages of SS events which are correctly assigned after p_T ordering and the BDT-based assignment. All events are weighted equally in this table. Therefore, the numbers do not correspond to the percentages in the text which utilize the individual event weights.

6.2.2.2 Top and E_T^{miss} reconstruction

The next step in the reconstruction is to define the t properties. The known variables come from the reconstructable decay products of the top – the reconstructed four-momenta of the $\text{j}_{b\text{-tag}}^{t}$ and the lepton from the W-decay. The four-momentum of the neutrino(s) of the W decay is not reconstructable. In case the W decayed into a τ_{had} two neutrinos are created from the initial W-decay and the subsequent τ decay respectively. The unknown momenta are treated as one object in the reconstruction. Additionally, the mass of both W and t are known quantities.

One goal can be to reconstruct the full neutrino four-momentum. As the neutrino mass is known to be practically zero, three variables have to be calculated. In case of a $W \to \tau$ decay at least two neutrinos are emitted. Therefore, this assumption does not hold, but the difference to zero can be neglected. Such, additionally to the two known mass variables an additional assumption has to be made. One possible assumption could be to linearly correlate $E_\text{z}^{\text{miss}}|_{\text{reco}}^{t}$ with $p_{\text{z,reco}}^{\ell(t)}$. Due to the quadratic nature of the equations (quartic after combination) this approach has the downside of not producing one result for each event, but instead zero, two or four results which can be real or imaginary. This leads to some events not being able to be reconstructed and a struggle to pick the correct result in most cases [179].[2]

After $\vec{E}_\text{T}^{\text{miss}}|_{\text{reco}}^{t}$ has been reconstructed, this information can be used to draw some conclusions on the H system:

$$E_\text{T}^{\text{miss}} = \vec{E}_\text{T}^{\text{miss}}|_{\text{reco}}^{t} + \vec{E}_\text{T}^{\text{miss}}|_{\text{reco}}^{H} + \text{res} \tag{6.4}$$

$$\tag{6.5}$$

The residual term consists of E_T^{miss} mis-measurements, such as any additional unmeasured objects (for example ν in the $\text{j}_{b\text{-tag}}^{t}$ decay) and energy lost in the passive material of the detector (see section 5.4.8). Assuming this contribution is small one can use

$$\vec{E}_\text{T}^{\text{miss}}|_{\text{reco}}^{H} \approx \vec{E}_\text{T}^{\text{miss}} - \vec{E}_\text{T}^{\text{miss}}|_{\text{reco}}^{t} \tag{6.6}$$

which can in turn be used to reconstruct the H system (see section 6.2.2.3). This assumption and the modified E_T^{miss} are used for all following reconstruction steps.

---

[2] In the best configuration about 19% of the simulated events have zero solutions and only 45% of events are reconstructed to have a real solution.

Since E_T^{miss} is just a two-dimensional variable, the used parts of the last equation are

$$\left(E_x^{miss}\right)_{reco}^H \approx E_x^{miss} - E_x^{miss}\big|_{reco}^t \tag{6.7}$$

$$\left(E_y^{miss}\right)_{reco}^H \approx E_y^{miss} - E_y^{miss}\big|_{reco}^t \tag{6.8}$$

which only needs $E_x^{miss}\big|_{reco}^t$ and $E_y^{miss}\big|_{reco}^t$. This two-dimensional vector can be calculated from the W and t mass, or derived from truth distributions in simulation. The latter approach is used as it does not make use of quadratic equations and allows the (incomplete) t and W-mass to be plotted as a crosscheck. The used equations

$$p_T^{miss}\big|_{reco}^t = \frac{\alpha}{p_{T,reco}^{\ell(t)}} \tag{6.9}$$

$$\phi^{miss}\big|_{reco}^t = \phi_{reco}^{\ell(t)} \pm \frac{\pi}{2} \tag{6.10}$$

are derived from figure 6.7(a) and figure 6.7(b). α is fit to $1\,615.18\,\text{GeV}^2$ and the angle is derived in a way that the system of $\ell(t)$, $j_{b\text{-tag}}^t$ and $\vec{E}_T^{miss}\big|_{reco}^t$ is balanced in the xy-plane [180]. The resulting m_{reco}^t can be found in figure 6.8.

(a) e/μ from t vs. ν from t

(b) p_T e/μ from t vs. ν from t
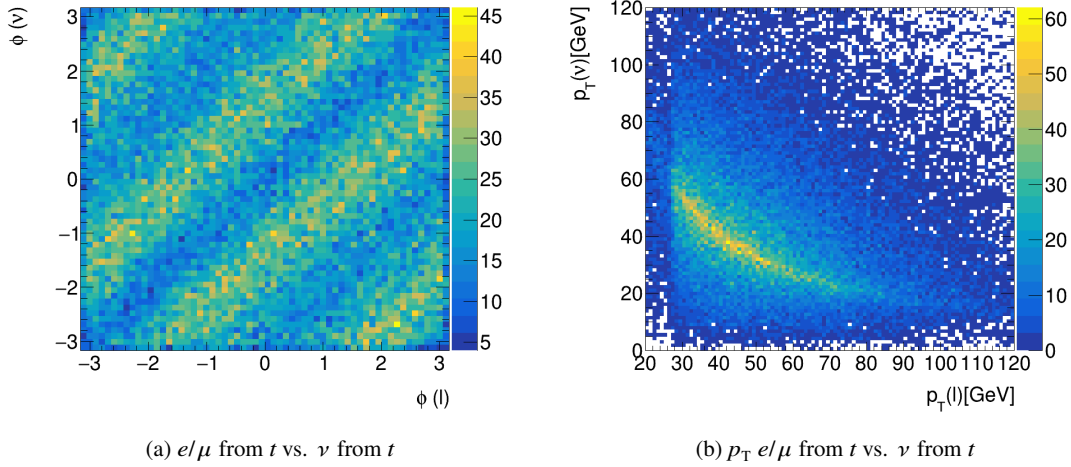
Figure 6.7: Comparison of the true t decay products in ϕ and p_T in all three channels. Events that pass any of the three selections when reconstructed are included equally.

6.2.2.3 H mass reconstruction

The last step in the reconstruction is to reconstruct the Higgs system. Especially m_{reco}^H is expected to be a discriminating variable against tZq which is not discriminatable via other variables. The easiest way to derive this property is m_{vis}^H which is completely independent of any neutrino-reconstruction:

$$m_{vis}^H = m\left(l_{1,reco}^H + l_{2,reco}^H\right) \tag{6.11}$$

where $l_{1,reco}^H$ and $l_{2,reco}^H$ denote both reconstructed leptons originating from H, which can be either light leptons or τ_{had}. This variable has a low uncertainty as only few assumptions have to be made. However, the
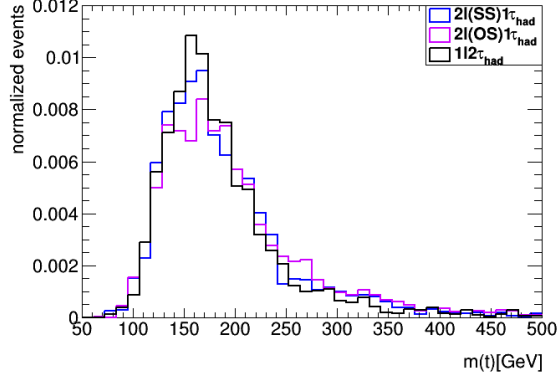
Figure 6.8: Reconstructed t mass in all three channels. Each distribution is normalized to one.

mass is missing a component and therefore is expected to be lower than the known m_{truth}^H. The distributions of m_{vis}^H in the three channels are shown in figure 6.7(b). Notably, the distribution from $1e/\mu + 2\tau_{\text{had}}$ events is smaller than the distributions from $2e/\mu + 1\tau_{\text{had}}$ events due to the lower number of ν.

(a) m_{vis}^H $1e/\mu + 2\tau_{\text{had}}$

(b) m_{vis}^H $2e/\mu + 1\tau_{\text{had}}$
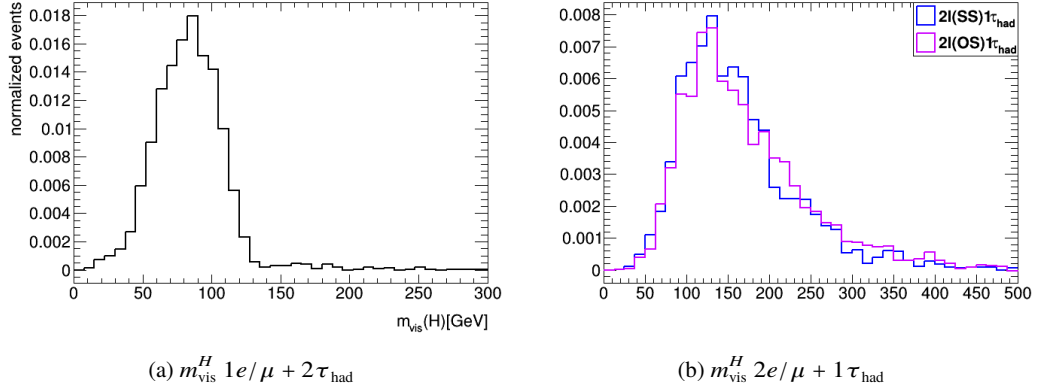
Figure 6.9: m_{vis}^H for the three channels. Note that the range for the $1e/\mu + 2\tau_{\text{had}}$ distribution is smaller.

Other ways to reconstruct the H mass have been developed for the $H \rightarrow \tau\tau$ analysis [181] and can be applied with the assumption stated in equation (6.8). However, this leads to an additional uncertainty on top of the reconstruction method. This is especially visible in case of the 'transverse mass' the most straightforward reconstruction method when taking into account $E_{\text{T}}^{\text{miss}}$:

$$m_{\text{T}}^H = m\left(l_{1,\text{reco}}^H(x,y) + l_{2,\text{reco}}^H(x,y) + \vec{E}_{\text{T}}^{\text{miss}}|_{\text{reco}}^H(x,y)\right) \tag{6.12}$$

This variable was first developed in UA1 ([182]) and works best in a clean environment with not too much boost and just one neutrino. In figure 6.10 one can see that m_{T}^H is significantly lower than the simulated mass m_{truth}^H (which corresponds to the literature value of the H mass) as the z-component of both leptons was neglected. Additionally, the spread is broad due to the same effect and the mis-measurement of $E_{\text{T}}^{\text{miss}}$.

A more involved way to take into account the neutrinos is the 'collinear approximation' [183, 184]. In this, it is approximated that the neutrinos from τ decays are collinear to the visible products from the
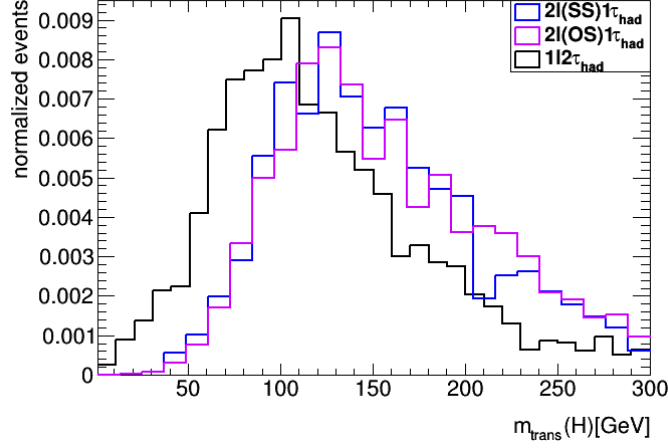
Figure 6.10: m_T^H for the three channels. All distributions are normalized to one.

same decay. This not only allows to take into account the *z*-component of the visible objects, but also to estimate the *z*-component of neutrino momenta. However, this assumption only holds for a boosted *H* system. This is not true for *tHq* events as can be seen in figure 6.11. Therefore, the collinear approximation is not compatible with the other methods.
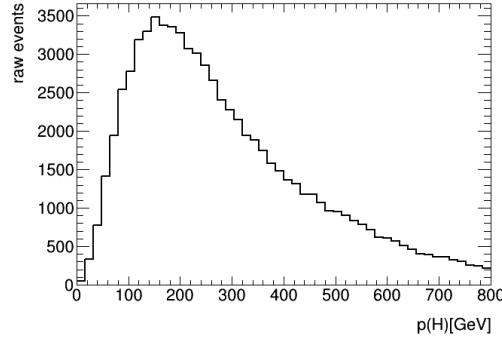
Figure 6.11: Momentum of the simulated *H*. All events from the three channels are included with weight one.

The most refined concept of Higgs mass reconstruction used in this thesis is the MissingMassCalculator (MMC). It is not searching for an analytical correct solution, but the most likely solution. More details about this can be found in [181, 184]. It has been originally developed for the $H \to \tau\tau$ analysis [185]. The MMC scans the phase space of *H* and τ_had related parameters to find the most likely solution for each event. Figure 6.12 shows the *H*-mass reconstructed with the MMC for *tHq* events (m_MMC^H). It achieves both, a small *H* mass resolution and an accurate position of the *H*-mass peak. The scanned phase space is expected to look different for *tHq* than for the pure $H \to \tau\tau$ events. A study was made to use events in a more similar phase space (*tZq*). The results can be seen in figure 6.13. Unfortunately, the low statistics in the used sample does not result in a more accurate description. Therefore, the default configuration is used.

The plots in figure 6.14 show the comparisons of the best discussed *H*-mass reconstruction techniques for the *tHq* sample for the different channels. While the m_vis^H offers a surprisingly narrow distribution, the
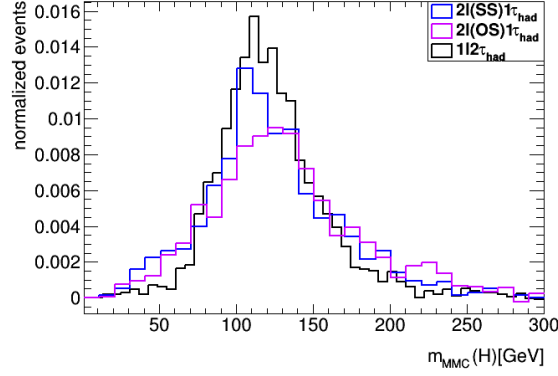
Figure 6.12: m_{MMC}^H for the three channels. All distributions are normalized to one.
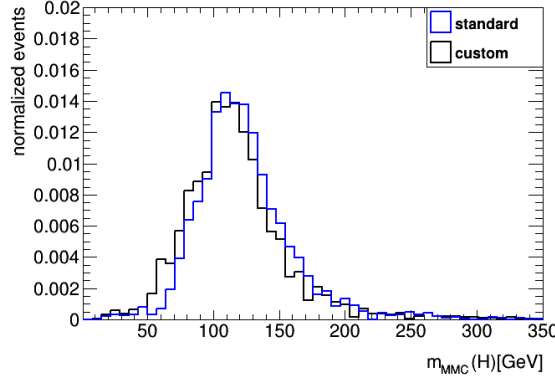
Figure 6.13: Comparison of m_{MMC}^H calculated with a standard and a custom made configuration file of the MMC for the $1e/\mu + 2\tau_{\mathrm{had}}$ channel. Both distributions are normalized to one.

output estimate is too low. The m_{T}^H has a broader distribution with a wrong peak position. m_{MMC}^H offers the best combination of a narrow distribution and an accurate peak position.

6.2.3 Channel or sample specific cuts

After the reconstruction more cuts are implemented to ensure the best event quality. These cuts may just apply to specific channels and can rely on the reconstruction.

- **Trigger match**: As described in section 5.3 an e or μ trigger is required to fire for each kept event. It is possible to get the information about the object that has been reconstructed in the trigger-level analysis. This object can then be matched in ΔR with the fully reconstructed e or μ. In the $2e/\mu + 1\tau_{\mathrm{had}}$ channel any event which did not keep the fully reconstructed triggered object is removed. In the $1e/\mu + 2\tau_{\mathrm{had}}$ channel the events which would fail this cut are kept, but the numbers are in the single digits.

- *tHq* **events in** $t\ell\ell q$ **samples:** The $t\ell\ell q$ samples (described in section 3.3) can by definition contain *tHq* events. With the kept truth information it is possible to veto those events.

(a) $1e/\mu + 2\tau_{had}$

(b) $1e/\mu + 2\tau_{had}$

(c) $1e/\mu + 2\tau_{had}$
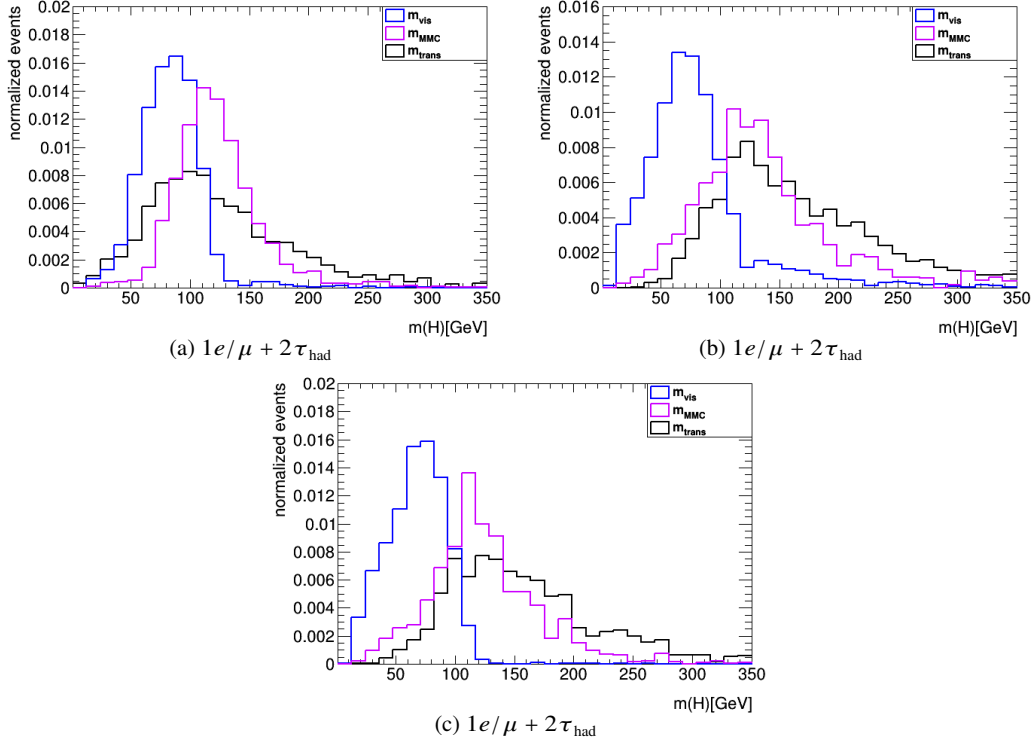
Figure 6.14: Comparisons of *H* mass reconstruction techniques for all three channels. All distributions are normalized to one.

- **Leading light lepton p_T:** The leading *e* or μ is required to pass 27 GeV as it is supposed to trigger the event. In case another lepton triggered the event this lepton is also required to pass this cut.

- **Subleading light lepton p_T:** At lower p_T the rate of falsely reconstructed leptons rises. To mitigate that effect a p_T cut on the subleading light lepton in the $2e/\mu + 1\tau_{had}$ channel is performed. This requires 14 GeV if the subleading light lepton has a lower p_T than the τ_{had}. (In case the subleading light lepton has a higher p_T than the τ_{had}, it automatically has a $p_T >20$ GeV as this is the minimum p_T of any τ_{had}.) This cut was optimized to ensure the highest S/B in the $2e/\mu SS + 1\tau_{had}$ channel and was subsequently adopted for the $2e/\mu OS + 1\tau_{had}$ channel.

- **1 e/μ + 2 τ_{had} :** To enable the reconstruction in the $1e/\mu + 2\tau_{had}$ channel the events are required to contain one light lepton and two τ_{had} which carry an opposite sign.

- **Higgs mass:** The reconstructed Higgs mass in all channels can be constrained to be in a range around the known *H* mass in the $1e/\mu + 2\tau_{had}$ channel. As this information is additionally processed by the MVA methods, the range is set to $\pm 1\,000$ GeV to minimize the effect of this cut.

- **$2e/\mu + 1\tau_{had}$:** To enable the reconstruction in the $2e/\mu + 1\tau_{had}$ channel processed event has to contain one τ_{had}, two light leptons, at least one *b*-tagged jet and one additional jet.

In figure 6.15 the effect of these cuts can be viewed for the most important samples, the effect for all other samples can be found in appendix C.
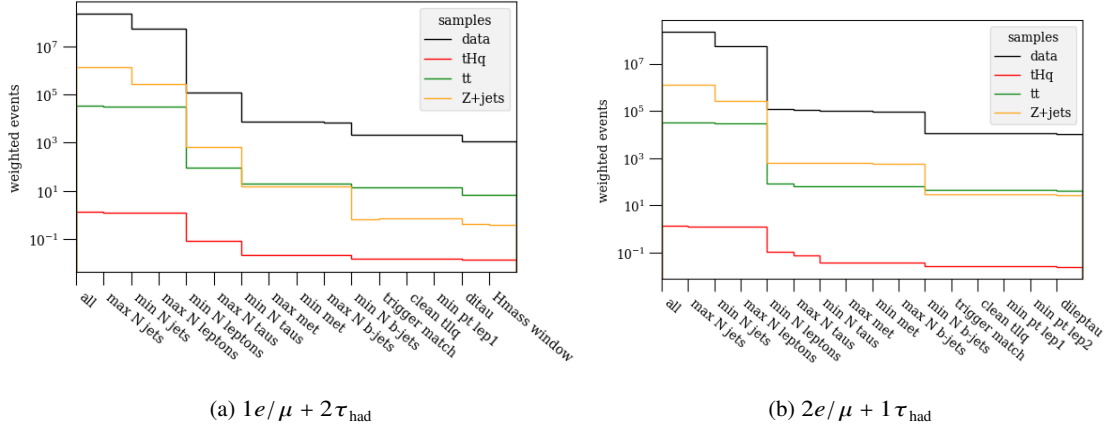
(a) $1e/\mu + 2\tau_{\text{had}}$

(b) $2e/\mu + 1\tau_{\text{had}}$

Figure 6.15: Cutflow on the cuts described in section 6.2.1 and section 6.2.3. The most important samples for the $1e/\mu + 2\tau_{\text{had}}$ channel ((a)) and the combined $2e/\mu + 1\tau_{\text{had}}$ channels ((b)) are shown. All samples are weighted to their respective cross-section.

6.2.4 Fake estimation

This thesis relies on interpreting reconstructed objects as originating from the primary interaction. However, as described in section 5.4.3 objects can be falsely reconstructed as being prompt while they were either originating from a later decay (for example non-prompt electrons) or never present (for example jets originating from light leptons misinterpreted as τ_{had}). For the purpose of this thesis both are addressed as 'fakes'. To estimate the abundance of fakes, the reconstructed objects in simulated datasets are matched with the simulated truth objects via ΔR. Tables 6.10 and 6.11 give an overview of the estimated fake (including non-prompt) contributions in the used regions. A split of the fake contributions by sample can be found in figure 6.17 (the full tables are in appendix D). It is obvious that the major backgrounds in the $1e/\mu + 2\tau_{\text{had}}$ and $2e/\mu\text{OS} + 1\tau_{\text{had}}$ channels are dominated by τ_{had} fakes, which mostly originate from quarks and gluons. Several challenges emerge from this: First, we cannot necessarily go back to first principles when reconstructing the background processes. For example the two t cannot be reconstructed in the $t\bar{t}$ sample, as one τ_{had} is expected to originate from a not further specified jet. Another challenge is that the MVA used to separate signal from background events would benefit most from being able to recognize fakes. However, each sample has different origins of fakes and the object internal variables are both not available at MVA level and already used in the reconstruction. Therefore, cutting on this could lead to sabotaging the reconstruction efficiency. The biggest challenge is that the estimation of fakes using simulated datasets is not completely accurate.

The two approaches used in this thesis to correct for the estimation of fakes are the scaling method and a Template Fit (TF). Both methods rely on identifying the fake events in simulation. This is done by identifying the truth matched objects described in section 6.1. In a second step the number of events in data is compared to the sum of the simulated datasets. From the difference a multiplicative weight is calculated and applied to simulated fake events. If this is done in the region used for the further analysis, effects that are present in data but not in simulated datasets might be lost. This includes amplified or not present signal. Therefore, this step of the analysis has to be performed in a different region. The most popular choice for an alternative region is to require the reconstructed leptons passing the loose WP but not the tight one ('LooseNotTight'). This ensures that the other objects (like jets) stay the same and thus relevant processes

	No fakes	τ_{had} gluon fake	τ_{had} quark fake	τ_{had} other fake
backgrounds in $2e/\mu$SS + $1\tau_{\text{had}}$	81.3 ± 1.6	9.2 ± 0.6	23.0 ± 1.1	2.8 ± 0.4
backgrounds in $2e/\mu$OS + $1\tau_{\text{had}}$	472 ± 6	$1\,567 \pm 27$	$5\,640 \pm 70$	$2\,240 \pm 50$

	e/μ fake	e charge flip
backgrounds in $2e/\mu$SS + $1\tau_{\text{had}}$	19.6 ± 1.6	33.0 ± 3.2
backgrounds in $2e/\mu$OS + $1\tau_{\text{had}}$	28.3 ± 1.9	340 ± 13

Table 6.10: Fake origins in the $2e/\mu$SS + $1\tau_{\text{had}}$ and $2e/\mu$OS + $1\tau_{\text{had}}$ channels. All backgrounds are evaluated together. One event might fall into multiple categories. A more detailed table can be found in appendix D. The full weights before fake correction are applied.

	τ_{had} gluon fake	τ_{had} quark fake	τ_{had} other fake
backgrounds in $1e/\mu$ + $2\tau_{\text{had}}$	88 ± 4	433 ± 12	93 ± 15

	τ_{had} 2 gluon fake	τ_{had} 2 quark fake	τ_{had} 2 other fake
backgrounds in $1e/\mu$ + $2\tau_{\text{had}}$	143 ± 8	531 ± 12	199 ± 14

	No fakes	e/μ fake	e charge flip
backgrounds in $1e/\mu$ + $2\tau_{\text{had}}$	101.8 ± 1.7	1.2 ± 0.4	21.3 ± 1.9

Table 6.11: Fake origins in the $1e/\mu$ + $2\tau_{\text{had}}$ channel. All backgrounds are evaluated together. One event might fall into multiple categories. A more detailed table can be found in appendix D. The full weights before fake correction are applied.
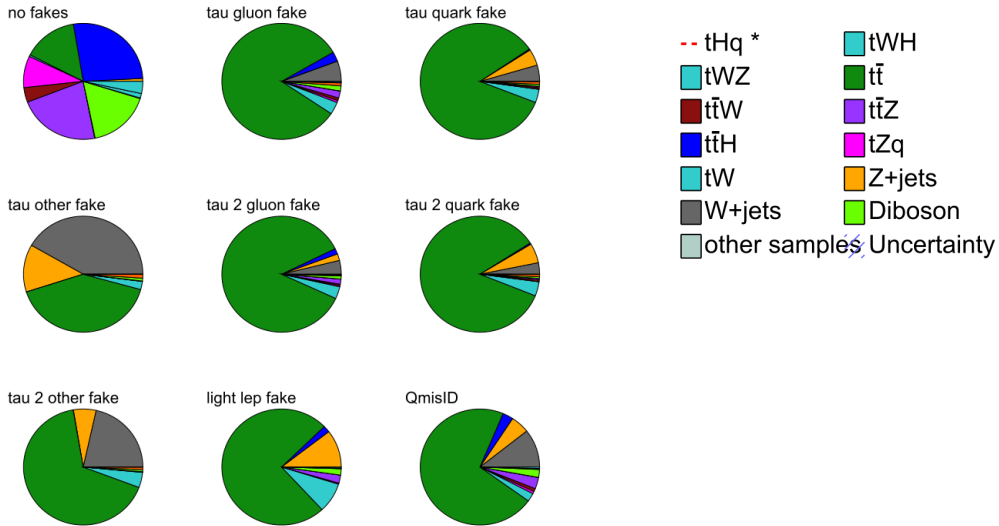
Figure 6.16: Fake origins in the different samples of the $1e/\mu$ + $2\tau_{\text{had}}$ channel. One event might fall into multiple categories. The full weights before fake correction are applied. The color code given here is commonly used across multiple plots. As *tHq* is the signal process it is not shown in these piecharts.

are similar. Additionally, in this region the number of events including fakes are enhanced, which reduces the statistical uncertainty of the estimation.

(a) $2e/\mu\text{SS} + 1\tau_{\text{had}}$　　　　　(b) $2e/\mu\text{OS} + 1\tau_{\text{had}}$
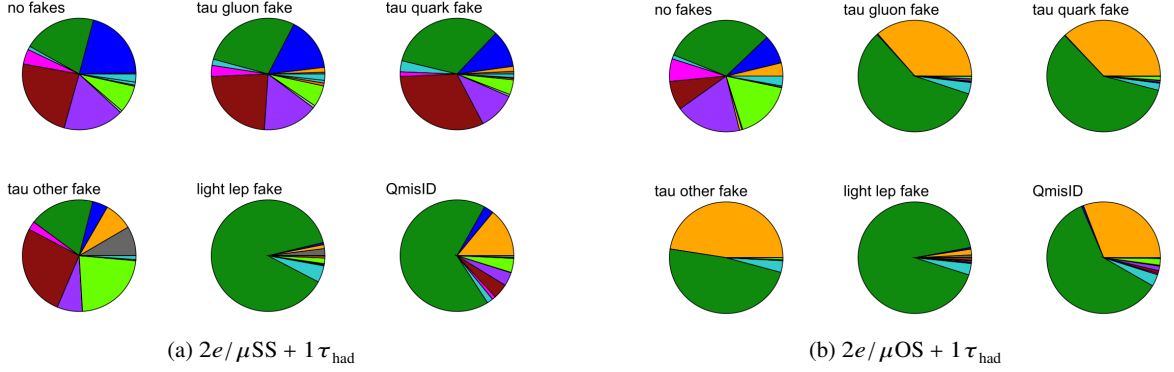
Figure 6.17: Fake origins in the different samples in the $2e/\mu + 1\tau_{\text{had}}$ channels. One event might fall into multiple categories. The full weights before fake correction are applied. The color code presented in figure 6.16 is used.

In the easiest implementation of the scaling method the number of simulated fake events in one CR is scaled by a factor f to match the number of data events.

$$N_{\text{data,CR}} = N_{\text{sim,true,CR}} + fN_{\text{sim,fake,CR}} \tag{6.13}$$

If the fake sources in the SR vary, several CRs are required to create several scaling factors. One of the reasons to use several CRs and thus different factors in this analysis is the prongness of τ_{had}. As described in section 5.4.6 1-prong and 3-prong τ_{had} are used in the analysis. The sources of misidentification are slightly different and also differently common for 1-prong and 3-prong τ_{had}. To take this into account, the LooseNotTight CR is split into two regions, one containing 1-prong τ_{had} and the other containing 3-prong τ_{had}. In each of these regions one f is calculated. For each simulated event in the SR it needs to be checked what prongness the containing fake τ_{had} possesses, and the according factor is multiplied with this event.

This strategy is expanded in the $1e/\mu + 2\tau_{\text{had}}$ channel in the following way: the CR containing loose leptons is split into regions where one of the three leptons (light lepton, leading τ_{had} and subleading τ_{had}) is loose while the other objects pass the tight criteria. In the case of a loose τ_{had}, the region is further split by prongness of the loose τ_{had} and number of b-tagged jets in the event. The latter influences the fake probability by altering the sources of backgrounds containing fakes (especially $t\bar{t}$). The CR including one loose light lepton is not split further. This is not necessary since the probability of a fake light lepton is lower in comparison to τ_{had} (see table 6.11). Additionally, the available statistics of the simulated datasets is not high enough to benefit from a split with respect to the rise in statistical uncertainty. To evaluate the impact of that statistical uncertainty the Poissonian uncertainty is calculated and used for the systematic variation. The application of this method in the $1e/\mu + 2\tau_{\text{had}}$ region is shown in figure 6.18. While the simulation before the fake estimation was underestimating the total number of events, afterwards it is overestimated. This suggests an overestimation of fakes.

A more involved method to estimate the realistic fake contribution is a Template Fit. As the name suggests it makes use of the previously discussed method of a binned profile likelihood fit (section 4.3). The advantage of that approach is that two different contributions can be fit simultaneously, in this case different fake sources. As described before, the main sources of τ_{had} fakes are jets. These jets may originate from quarks or gluons which impacts the jet structure, especially the tracks associated with it. As this influences the misreconstruction of a jet as a τ_{had} these two samples are fit and thus estimated simultaneously. The fit variable needs to be sensitive to the quark/gluon origin of the fake τ_{had}. Therefore,
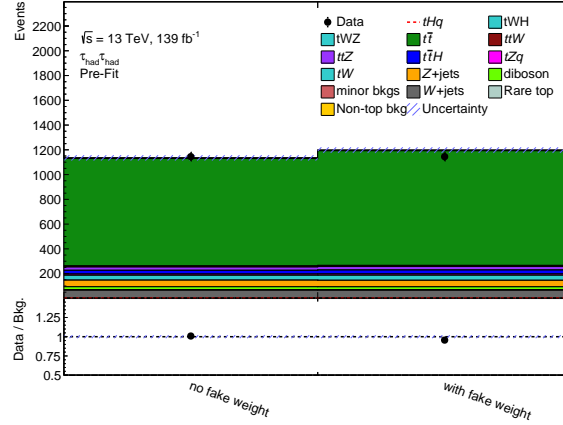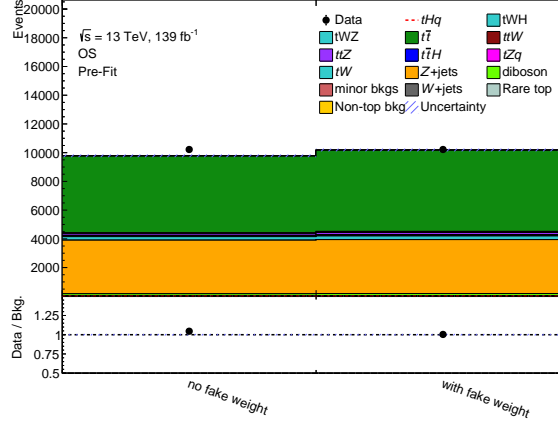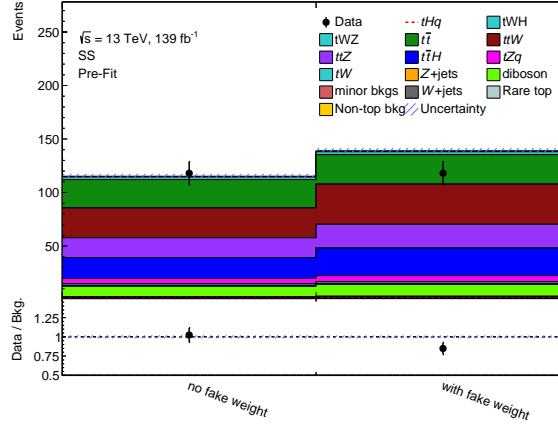
Figure 6.18: Application of the fake estimation in the $1e/\mu + 2\tau_{\mathrm{had}}$ channel.

a variable is needed to discriminate between those two samples. For events with only one b-tagged jet the τ_{had} property 'JetTrackWidth' is used (it measures the distance between the tracks associated with the τ_{had} object). For events with two b-tagged jets, instead a gradient BDT proved to be more discriminating. The implementation, available as LightGBM [186], is used with τ_{had} specific variables (including the JetTrackWidth) as input as well as information about the light leptons and two leading light jets. The split of events is given by the truth matched τ_{had} origin (see above) which is expanded to also identify the type of hadrons matched with τ_{had}. The dominating samples are q-jet (quark originating jet) and g-jet (gluon originating jet), but smaller contributions are given by 'unknown' (truth matching failed), 'tau' (matched to a simulated τ) and 'electron' (matched to a simulated electron). As the $2e/\mu\mathrm{OS} + 1\tau_{\mathrm{had}}$ channel has higher statistics than the other channel the fake estimation is not only split by the number of b-tagged jets and prongness, but also by $p_{\mathrm{T}}^{\tau_{\mathrm{had}}}$ (as this channel just includes one τ_{had} no separate estimation has to be performed for the subleading τ_{had}). q-jet and g-jet contributions behave differently at low $p_{\mathrm{T}}^{\tau_{\mathrm{had}}}$ which is to be expected as the jet modelling heavily depends on jet p_{T} and also most jets are created at low p_{T}. Uncertainties here come from two sources: the statistical uncertainty (estimated from the counting method described above) and the shape uncertainty estimated by the fit. Both are propagated independently through the following steps. The application of the resulting weights is shown in figure 6.19

The third channel $2e/\mu\mathrm{SS} + 1\tau_{\mathrm{had}}$ again suffers from low statistics. Therefore, a Template Fit is not feasible and the scaling method is used instead. Other than for the $1e/\mu + 2\tau_{\mathrm{had}}$ channel a split in $p_{\mathrm{T}}^{\tau_{\mathrm{had}}}$ is preferred over a split in the b-tagged jet multiplicity, as the background contribution does not significantly vary with the latter, but a doubling of the estimated fake contribution is observed at low $p_{\mathrm{T}}^{\tau_{\mathrm{had}}}$. The application of the fake estimation in this channel is shown in figure 6.20

For all channels the simulation is scaled up after applying the fake channels. However, the agreement is not perfect after the correction which is a result of the shortcomings of the methods: more variables than can be explored have an influence on the jet modelling and the τ_{had} reconstruction, misreconstruction from the LooseNotTight region cannot directly be transferred to the tight region and statistics limits the meaningfulness of the fake estimation. Overall, if it would be straightforward to understand the underlying processes both the simulation and reconstruction would have been improved to a point that these corrections would become (largely) obsolete.

Figure 6.19: Application of the fake estimation in the $2e/\mu\text{OS} + 1\tau_{\text{had}}$ channel.

Figure 6.20: Application of the fake estimation in the $2e/\mu\text{OS} + 1\tau_{\text{had}}$ channel.

6.2.5 Yields

The number of events in the three channels, after the application of all previously described cuts, weights and the fake estimation are called pre-selection yields. They build the baseline for the further analysis with more involved statistical methods. The complete yields including all samples are shown in table 6.12. Some samples have a negligible contribution for certain channels. To simplify the analysis and avoid high uncertainties from low statistics of these samples, they are combined to 'minor backgrounds' (marked with 'x'). This category is filled with any sample which has a lower yield than the signal in the relevant preselection region. However, tWH is exempt from this rule, as it is so similar to tHq and expected to show similar behavior in case of a BSM scenario. Additionally, Z+jets and W+jets in $2e/\mu\text{SS} + 1\tau_{\text{had}}$ are not combined to allow for keeping track of their fake and charge-flip properties.

	$2e/\mu\text{OS} + 1\tau_\text{had}$	$2e/\mu\text{SS} + 1\tau_\text{had}$	$1e/\mu + 2\tau_\text{had}$
tHq	1.91 ± 0.05	1.26 ± 0.04	1.92 ± 0.04
tWH	2.42 ± 0.10	0.98 ± 0.06	1.47 ± 0.06
tWZ	18.87 ± 0.09	2.711 ± 0.034	3.92 ± 0.04
$t\bar{t}$	5 648 ± 29	27.6 ± 2.0	926 ± 12
$t\bar{t}W$	70.0 ± 0.9	34.4 ± 0.7	8.54 ± 0.27
$t\bar{t}Z$	137.8 ± 1.2	22.1 ± 0.5	28.2 ± 0.6
$t\bar{t}H$	60.67 ± 0.29	25.87 ± 0.23	32.15 ± 0.21
tZq	39.9 ± 0.4	5.66 ± 0.16	9.84 ± 0.18
tW	255 ± 6	1.6 ± 0.4	40.9 ± 2.4
Z+jets	3 750 ± 90	0.88 ± 0.27	52 ± 12
diboson	177.1 ± 2.4	11.1 ± 0.5	22.6 ± 0.8
triboson	x 0.98 ± 0.07	x 0.35 ± 0.05	x 0.180 ± 0.026
t-channel	x 0.30 ± 0.15	x 0.18 ± 0.09	4.7 ± 0.5
s-channel	x 0.00 ± 0.00	x 0.00 ± 0.00	x 0.39 ± 0.11
W+jets	x 0.28 ± 0.15	0.68 ± 0.26	66 ± 14
multitop	x 1.68 ± 0.04	x 0.791 ± 0.029	x 0.204 ± 0.013
other Higgs	10.6 ± 2.4	x 1.0 ± 0.8	x 0.00 ± 0.00
total background	10 140 ± 90	130.3 ± 2.4	1 187 ± 22
data	10 221	118	1 145

Table 6.12: Yields of the three different channels after all cuts. Fake estimation weights are applied. Sample-region combinations marked with 'x' are combined per channel in the further analysis.

6.3 Neural Network application

To get a handle on the signal events and discriminate them from the background events NNs are used. The setup is similar for all three channels, a categorical NN is build targeting three types of processes: *tHq*, *tZq* and all other backgrounds. Each of those three samples is scaled to make up a third of the population, to ensure the different cross-sections do not bias the training. The benefit of separating *tZq* from other backgrounds is that *tZq* is discriminated against specifically. If it was a part of the overall background distribution it would not be highly weighted and likely ignored in the training. Since it resembles the signal distributions (apart from the reconstructed Higgs mass) it would likely resemble the signal. This is a problem since the statistics is ten times higher than the signal statistics (see table 6.12). Additionally, forcing the NN to separate *tZq* and *tHq* lays the focus on the reconstructed Higgs mass. This is expected to be a strong variable to separate *tHq* events from other processes. Another general advantage of having a categorical setting arises in the fit: having a three-dimensional output gives more flexibility in designing regions and choosing a sensitive variable to be fit. In that the output targeting *tZq* does not have to be used explicitly. Since all three outputs add up to one it simply allows for the additional degree of freedom for the output targeting other backgrounds.

Each NN is optimized for the specific input set using a combination of an automated procedure and a human guided grid search to find the optimal hyperparameters. The variables given to the NN are chosen to both include reconstruction-level object based information and variables reconstructed to assemble *tHq* features (see section 5.4). A full list of the used variables can be found in table E.1. Table 6.13 shows some

of the hyperparamets. Apart from information about the differing objects (subleading τ_{had} and lepton from top) the $1e/\mu + 2\tau_{\text{had}}$ channel uses more information about jets and the fully reconstructed top as it depends more on the jet faking τ_{had} and the top reconstruction is more reliable (as just one light lepton is available).

	$1e/\mu + 2\tau_{\text{had}}$	$2e/\mu\text{OS} + 1\tau_{\text{had}}$	$2e/\mu\text{SS} + 1\tau_{\text{had}}$
hidden layers	3	6	6
nodes per layer	112	120	100
dropout	0.2	0.65	0.25
training epochs	50	150	170

Table 6.13: Final hyperparameters of the NNs for all three channels.

6.3.1 NN performance

To ensure stability of the trained NN against fluctuations in the data, several quality standards are taken into account. First of all, a 10-fold-cross-validation is performed, meaning the training is done and validated on ten subsamples to ensure the training is stable against small fluctuations in the sample. The relevant information about the used NN performance can be seen in the signal to background separation.

Figure 6.21 shows the signal to background separation of the $1e/\mu + 2\tau_{\text{had}}$ NN. The performance in terms of AUC is highest for the *tHq* separation (80%), but significantly lower (63% and 61%) for the other two targets. Since the main interest of the analysis is in separating *tHq* and not separating the other two targets from each other, this is still a good result.

(a) background target
(b) tZq target
(c) tHq target
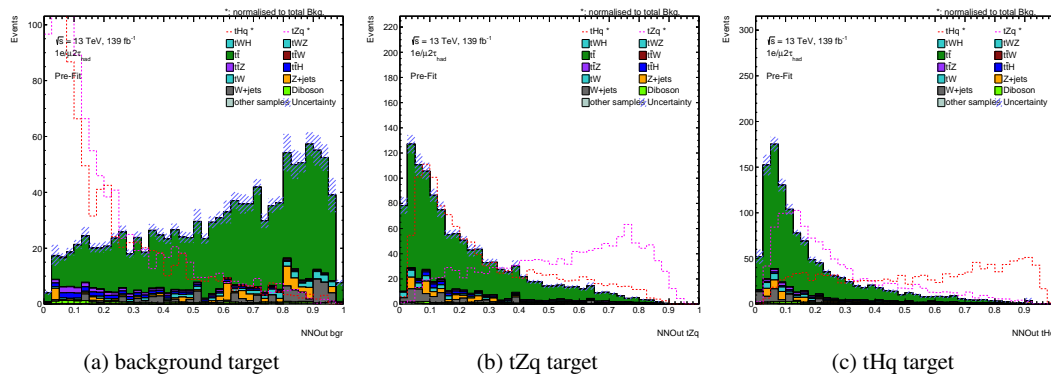
Figure 6.21: The three NN outputs in the $1e/\mu + 2\tau_{\text{had}}$ channel. *tHq* and *tZq* are overlayed and scaled to the same integral as the sum of all (other) backgrounds to show the shape differences.

The network trained on $2e/\mu\text{OS} + 1\tau_{\text{had}}$ events shows a similar behavior. The AUCs are 87%, 73% and 75% for the *tHq*, *tZq* and background target. But looking at the signal to background separation in figure 6.22 the high background ratio cannot be overcome just using the NN information.

When training the network on $2e/\mu\text{SS} + 1\tau_{\text{had}}$ events, the result is not as satisfactory as for the other channels. The AUCs achieved are just 69%, 69% and 81%. In figure 6.23 it becomes clear that the signal separation is not comparable to the other channels. The reason for this is unclear. One possible cause could be that the input features and dominating background samples differ so much that the optimal

Figure 6.22: The three NN outputs in the $2e/\mu$OS + $1\tau_{\mathrm{had}}$ channel. *tHq* and *tZq* are overlayed and scaled to the same integral as the sum of all (other) backgrounds to show the shape differences.
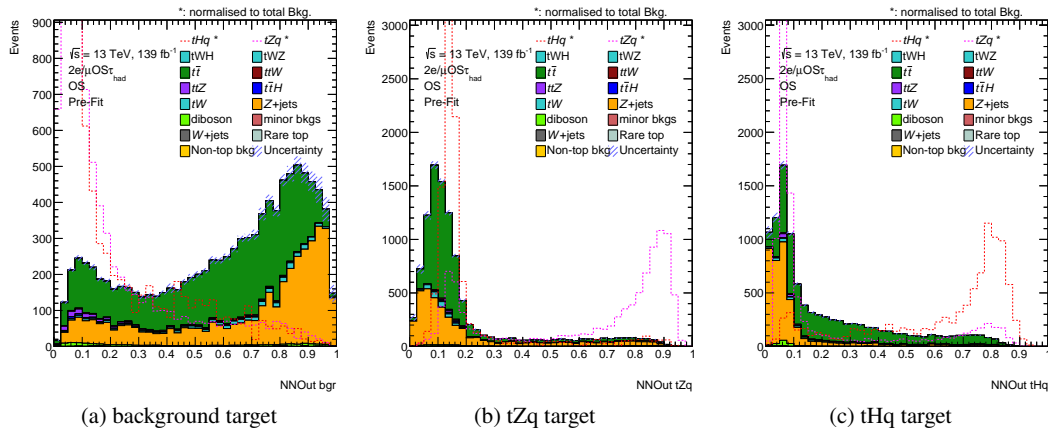
hyperparameters could not be found with the used optimization technique. Broader scans (for example with different targets) did not result in any significant improvement. Another possible cause could be that crucial information needed to separate the samples is not transferred to the network. This option is disfavored as the input features include all basic object information. A third cause could be some sort of bug in the input samples or network application. While a hunt for these is ongoing nothing could be found.

One cross-check is to apply the network trained on $2e/\mu$OS + $1\tau_{\mathrm{had}}$ events to the $2e/\mu$SS + $1\tau_{\mathrm{had}}$ events. It is found that this network is better at classifying $2e/\mu$SS + $1\tau_{\mathrm{had}}$ events than the network trained on these events. Figure 6.24 shows the resulting signal separation. This cross-check disfavors the second possible cause described above. As the result applying the network trained on $2e/\mu$OS + $1\tau_{\mathrm{had}}$ events on the $2e/\mu$SS + $1\tau_{\mathrm{had}}$ events is the best available, moving forward this network score is used.

Figure 6.23: The three outputs of the NN trained on $2e/\mu$SS + $1\tau_{\mathrm{had}}$ events applied in the $2e/\mu$SS + $1\tau_{\mathrm{had}}$ channel. *tHq* and *tZq* are overlayed and scaled to the same integral as the sum of all (other) backgrounds to show the shape differences.

(a) background target

(b) tZq target

(c) tHq target

Figure 6.24: The three outputs of the NN trained on $2e/\mu$OS + $1\tau_{\text{had}}$ events applied in the $2e/\mu$SS + $1\tau_{\text{had}}$ channel. *tHq* and *tZq* are overlayed and scaled to the same integral as the sum of all (other) backgrounds to show the shape differences.

Setting Limits on the *tHq* signal strength

But there's no sense crying over every mistake
You just keep on trying till you run out of cake
And the science gets done [...]

<div align="right">

– Valve, Jonathan W. Coulton, *Still Alive*

</div>

The final step in the analysis is a complete statistical analysis of the three different channels. This chapter is organized as follows: first, the fit setup of all three channels is described in section 7.1, section 7.2 and section 7.3. In section 7.4 the systematic uncertainties are introduced and the impact on the three fits is discussed. Section 7.5 discusses the results obtained from fits using only simulation. Some data is introduced for cross-checks in section 7.6 while the final results with data are discussed in section 7.7.

7.1 1 e/μ + 2 τ_{had} fit setup

This section first introduces the decision process for the fit setup on the example of the $1e/\mu + 2\tau_{\text{had}}$ channel. Subsequent sections show the analogous results for the other two channels. For the decision process on the analysis startegy it is important to not look at the regions of data which are expected to contain signal. This concept is called blinding and is an important step at preventing scientists from biasing their results by following statistical fluctuations they see in data. Therefore, if not stated otherwise, the decision process is completely based on simulation which is used to build Asimov data (see also section 4.3).

To obtain the best information about the *tHq* events in the $1e/\mu + 2\tau_{\text{had}}$ selection, first a signal region is selected. The basis of this are the preselection events which altogether achieve an S/B of 0.16%. On the one hand, finding a subset with a higher S/B is advantageous to help the fit. On the other hand statistical uncertainties will rise on a small subset of events, while informative value will shrink. Therefore, the arbitrary lower bound of 1 signal event in the SR is chosen. In total the goal is not to maximize S/B, but to significantly improve it with respect to the preselection region.

The SR definition can be refined with the NN outputs described already in section 6.3.1. From all three outputs the NN_{tZq} is not showing a separation between *tHq* and the background events. Therefore, only NN_{tHq} and NN_{bgr} are considered for the signal region definition. In figure 7.1 the results between possible cuts on both variables are compared. As one can see, cuts on NN_{tHq} achieve higher S/B in total, but all of these cuts define SRs including less than one signal event. For any given number of signal events bigger than one, a cut on NN_{bgr} is advantageous. The distribution is less steep for more than 1.4 signal events.

However, no plateau can be found. In order to create a signal enriched SR a cut at NN_{bgr} <0.45 is chosen where the S/B is tripled with respect to the preselection S/B. At this point 1.71 signal events are in the SR, while the majority of background is removed. The resulting yields can be seen in table 7.1.
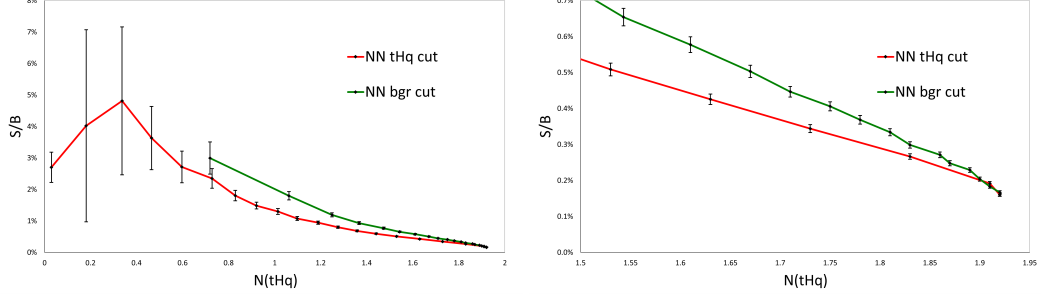
Figure 7.1: S/B against number of *tHq* events in the SR for different cuts on NN_{tHq} and NN_{bgr} in the $1e/\mu + 2\tau_{\mathrm{had}}$ channel.

	SR	
tHq	1.71	$\pm$ 0.04
tWH	1.07	$\pm$ 0.05
tWZ	2.698	$\pm$ 0.030
$t\bar{t}$	276	$\pm$ 7
tW	9.9	$\pm$ 1.2
t-channel	0.86	$\pm$ 0.23
$t\bar{t}W$	3.38	$\pm$ 0.17
$t\bar{t}Z$	19.5	$\pm$ 0.5
$t\bar{t}H$	21.69	$\pm$ 0.18
tZq	8.62	$\pm$ 0.17
Z+jets	11	$\pm$ 5
W+jets	15	$\pm$ 5
diboson	12.9	$\pm$ 0.5
other samples	0.22	$\pm$ 0.04
total background	383	$\pm$ 9

Table 7.1: Yields in the $1e/\mu + 2\tau_{\mathrm{had}}$ SR

In a second step, the bins for the SR are defined. The used variable should be sensitive to shape differences between signal and background events. NN_{tHq} possesses that property (see figure 6.21(c)). An advantageous consequence of the cut on NN_{bgr} is the availability of the full range of NN_{tHq} to differentiate out the different shapes of signal and background event distributions as shown in figure 7.2.

To define the exact boundaries of the bins, an automatic binning function is used [187]. This method can be steered by two variables which control the total number of bins and the focus on the signal or background samples (for CRs what the algorithm takes as the 'signal' sample can be redefined). In the two limits the sample which is not focussed appears flat, which means the number of events is the same across all bins. Given that the variable used peaks on the right for signal (and right for background), the
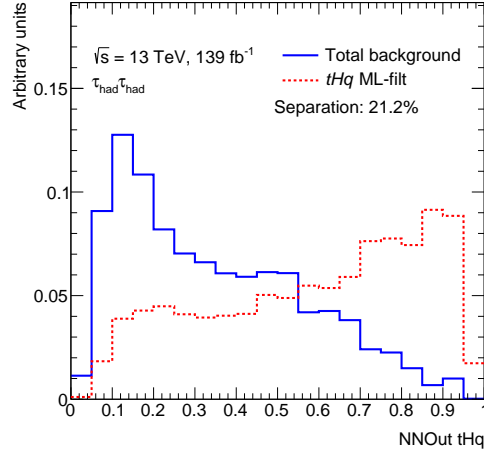
Figure 7.2: Shape comparison of *tHq* versus total background on NN_{tHq} in the $1e/\mu + 2\tau_{\text{had}}$ channel SR.

background (signal) sample will then peak on the left (right) side of the distribution. The most common setting weights both samples equally which creates two separate peaks for signal and background samples. Three examples with different focus are shown in figure 7.3.

(a) Focus on signal   (b) Equal focus   (c) Focus on background
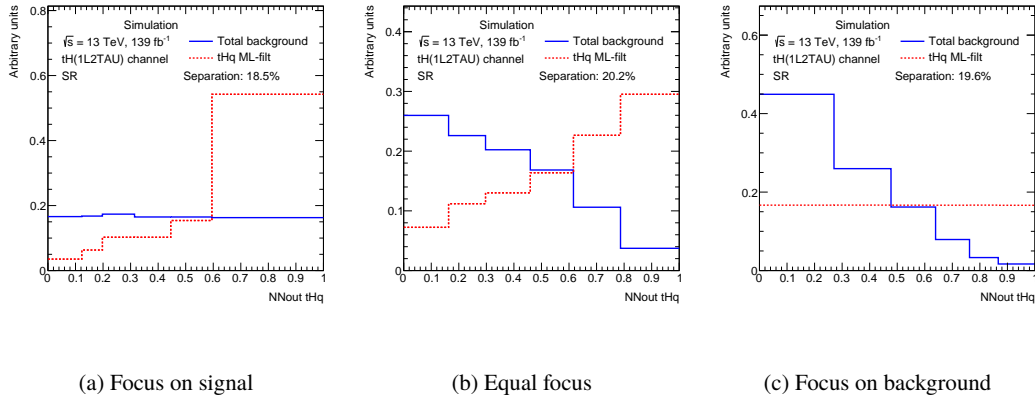
Figure 7.3: Examples for the automatic binning function. All three binnings were performed on the same distribution (SR in the $1e/\mu + 2\tau_{\text{had}}$ channel). Each of the three binnings requires six bins. From left to right the binning function focuses on signal, both distributions equally and background.

The chosen binning in the SR is optimized by the minimal $\Delta\mu(tHq)$ in a fit with only simulated events (Asimov fit) and only the SR. The lower boundaries for statistics per bin are given in terms of the uncertainty on the γ which only depend on the available number of simulated events in this bin. $\Delta\gamma$ per bin should be at maximum $\pm 20\%$. Additionally, each bin should contain at least 10 events. When increasing the number of bins a plateau in $\Delta\mu(tHq)$ is reached at about 7 bins. A subsequent scan of weighting signal versus background shows a minimum at focussing on the signal. Figure 7.4(a) shows the resulting shapes in signal and background with separated background samples (signal is shown as a separate distribution and scaled to the total number of backgrounds, data is not shown). The resulting $\Delta\gamma$ can be seen in figure 7.4(b).

(a) $1e/\mu + 2\tau_{\mathrm{had}}$ SR

(b) γ associated with the $1e/\mu + 2\tau_{\mathrm{had}}$ SR
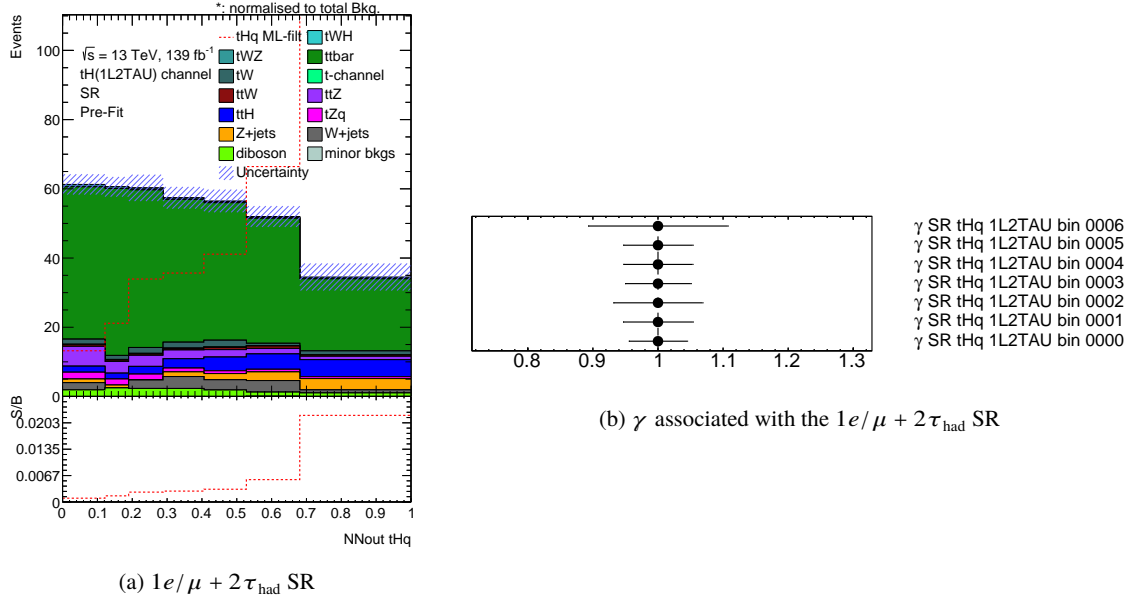
Figure 7.4: SR in the $1e/\mu + 2\tau_{\mathrm{had}}$ channel and the resulting γ. The γ are associated to the bins from left to right and numbered starting with 0.

7.1.1 Treatment of reducible backgrounds

From the yields (table 7.1) it becomes clear that the reducible backgrounds $t\bar{t}$, Z+jets and W+jets are the dominating backgrounds. Section 6.2.4 introduced the used method to estimate the impact of events including fake leptons. However, there are a number of caveats which limit the accuracy expectation for that estimation. First, the estimation is done on a different region than the fit. Second, the estimation is done inclusively over all samples, which may result in a slightly wrong result for specific samples. Third, systematic uncertainties are not taken into account in the fake estimation. Especially uncertainties on the modelling of dominating reducible backgrounds (namely $t\bar{t}$) are expected to have an impact on the fake estimation. Therefore, treatment of the dominating reducible backgrounds in the fit has to be evaluated carefully.

There are three methods to incorporate some degree of uncertainty into the fit: shape or normalization impact from a Nuisance Parameter (can also appear combined) and norm factors. As described in section 4.3 the differentiation which parameters are determined using a NP or a NF is solely physics motivated. However, a NF is seen as something that is estimated in a fit with no (or little) prior knowledge. In contrast, a NP is seen as a parameter which has an impact on the fit, but was determined beforehand. Therefore, to determine whether the described backgrounds should be included as a NP or NF, the impact of the different configurations is tested. In case the impact of a NF is higher, the NF is picking up a variation that was not described by the NP. In this case it seems to be the safe option to incorporate NFs for the dominating reducible backgrounds.

The impact of three different configurations for the $t\bar{t}$ background is tested. In all configurations $t\bar{t}$ modelling uncertainties are reduced to the comparison of different generators, which are expected to have the highest impact. Other NPs are not taken into account to receive a clearer picture. The first configuration shown in figure 7.5 assumes that the previous fake estimation was perfect. Therefore, the

overall normalization of the $t\bar{t}$ sample is not changeable during the fit. In this configuration modelling uncertainties are just allowed to change the shape of the sample. The second configuration shown figure 7.6 treats $t\bar{t}$ as any other sample. The sample after fake estimation is taken as the baseline and shape and normalization NPs are included. The third configuration shown in figure 7.7 associates a NF to $t\bar{t}$. It is allowed to vary between 0 and 2 times the nominal value. As this NF determines the normalization of the sample, the modelling uncertainties just impact the shapes. From the comparison of the three approaches, it becomes visible that there is a small difference in the result depending on the treatment. This hints at additional information being present in the fit. Allowing a NF to be determined in the fit allows for the most information to be picked up from the fit. Therefore, the dominating reducible backgrounds are associated with NFs in this thesis as described in strategy three.[1]
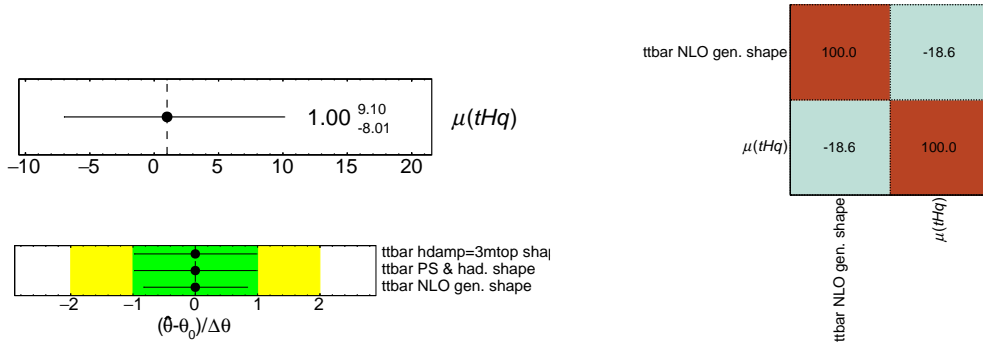
Figure 7.5: Fit results from an Asimov fit performed with strategy one. The only degrees of freedom for $t\bar{t}$ are the shapes of the systematic variations. Not all NPs appear in the correlation matrix due to pruning (see section 7.4).
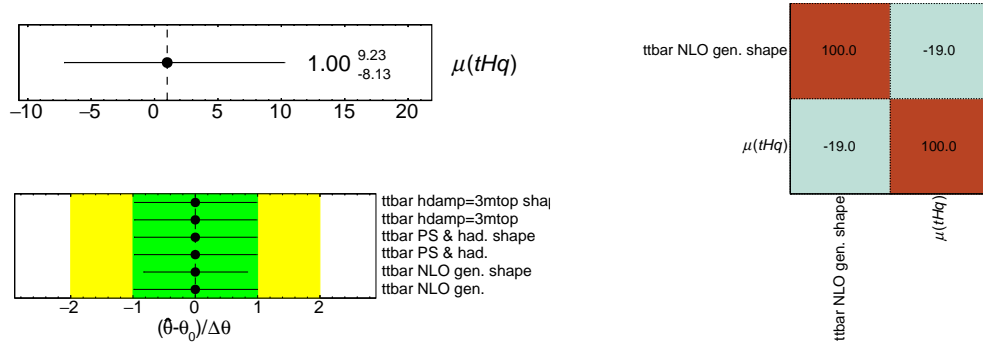
Figure 7.6: Fit results from an Asimov fit performed with strategy two. The only degrees of freedom for $t\bar{t}$ are the systematic variations. Not all NPs appear in the correlation matrix due to pruning (see section 7.4).

---

[1] This does not mean the fake estimation done before was not useful or should be disregarded. Besides studying the fake composition, the samples after correction are a better starting point for the fit. This correction is done on an even-by-event basis, taking into account information from event reconstruction (associated origin of leptons) as well as p_{T} and b-tagged jet multiplicity. Not all of this information is available at fit level anymore, or could be associated with a different scale factor.
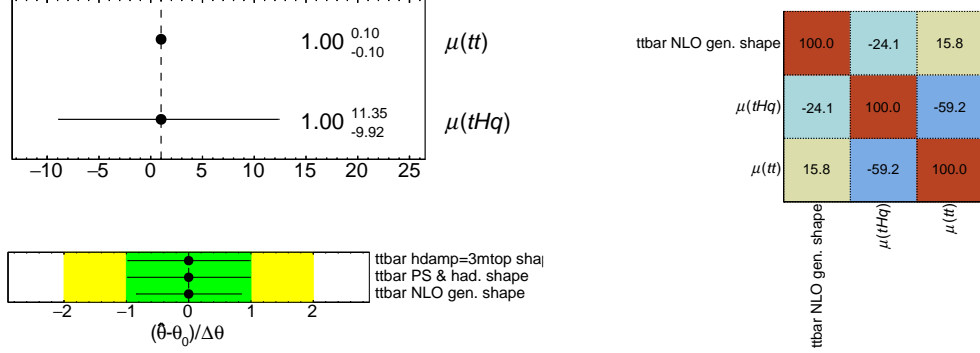
Figure 7.7: Fit results from an Asimov fit performed with strategy three. The degrees of freedom for $t\bar{t}$ being a NF and the shape of the systematic variations. This fit is evaluated in the SR. Not all NPs appear in the correlation matrix due to pruning (see section 7.4).
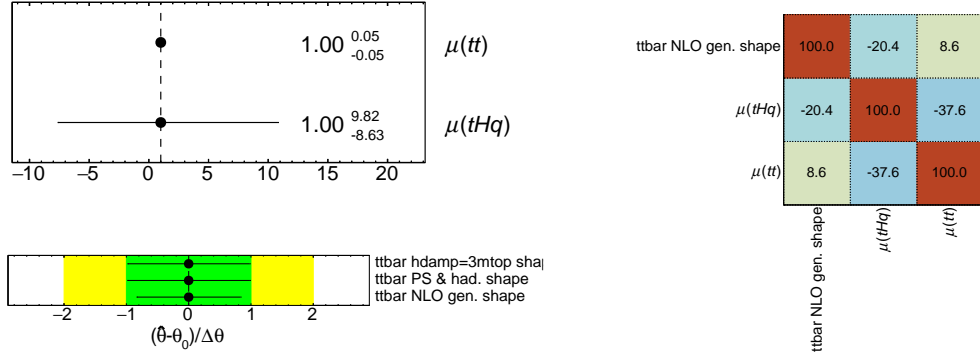
Figure 7.8: Fit results from an Asimov fit performed with strategy three and an additional CR. The degrees of freedom for $t\bar{t}$ are a NF and the shape of the systematic variations. This fit is evaluated in the SR and the $t\bar{t}$ CR. Not all NPs appear in the correlation matrix due to pruning (see section 7.4).

7.1.1.1 Control regions

As described above additional NFs are measured in the fit. Therefore, a CR is selected to allow a proper estimation which can be forwarded to the SR fit. CRs need to be orthogonal to the SR and between each other. Additionally, they are ideally dominated by the background they are associated with. If this is not possible, a region which is enriched in that background is sufficient, as long as the dominating backgrounds are well estimated. A last condition is to be comparable to the SR. Only in that case, the estimation of the NF in the CR is transferrable to the SR.

The obvious region which fulfills the above stated conditions is the preselection region with an inverted SR cut: $NN_{\text{bgr}} > 0.45$. Table 7.2 shows the sample composition in that region. The reducible backgrounds are dominating, especially $t\bar{t}$. A fit comparable to the previous studies including a $t\bar{t}$ NF measured in this CR can be seen in figure 7.8. The reduced $\Delta\mu(t\bar{t})$ and the reduced correlation between $\mu(tHq)$ and $\mu(t\bar{t})$ show the information gathered in this region.

However, what about Z+jets and W+jets? On the one hand, it is possible to argue that all three samples are reducible and therefore one NF is enough to estimate the impact of all fakes on the SR. On the other

hand, the three samples differ in the expected number of fake leptons (two for W+jets, one for the others) and b-tagged-jet multiplicity (two for $t\bar{t}$, none for the others). The latter already impacts the fake estimation described in section 6.2.4. In case the estimated NFs for the samples are different, it is better to associate different NFs to them. Unfortunately the generated number of events in the simulated W+jets samples leads to statistical fluctuations in the sample. Therefore, one can say the 'statistics' is not sufficient to estimate a separate NF. The effect of Additionally, a sufficient CR which favors this sample and not Z+jets does not exist. Therefore, Z+jets and W+jets are always estimated together. The associated NF is called $\mu(\mathrm{V}+\mathrm{jets})$. An anti-correlation between $\mu(t\bar{t})$ and $\mu(\mathrm{V}+\mathrm{jets})$ is expected as enhancing one background needs to be countered by lowering the other to result in the same sum of events. Therefore, two separate NFs are used for the fit and a separate CR is used. This CR should enable a decorrelation of both NFs and enhance the statistics of both Z+jets and W+jets. The CR used in this thesis is defined separately from the preselection: instead of requiring 1 or 2 jets b-tagged at 70% efficiency, one jet passing the 85% efficiency cut, but not the 70% cut is required. All other requirements stay the same. This region is enhanced in events with a spuriously b-tagged jet, namely Z+jets and W+jets. The sample composition of this region can be seen in table 7.2. For this region similar studies as for the preselection region are done. They can be found in appendix F.

Both CRs can be separated into different bins. Different configurations have been tested in terms of the used variable and binning options for both regions separately. Dividing the $t\bar{t}$ CR does not reduce any of the $\Delta\mu$ or the correlations significantly – unless about 40 bins are used, which would reduce the statistics per bin too much. The optimal configuration for the V+jets CR is two bins, on basis of $m_{b,\mathrm{j}^{\mathrm{spect}}}$ (see section 6.2.2.1). This variable was designed to find the t-channel structure of the signal. Z+jets and W+jets events neither have the $\mathrm{j}^{\mathrm{spect}}$ nor are they expected to contain b-quarks. Instead, the jets are expected to recoil against the boson. The resulting shape of $m_{b,\mathrm{j}^{\mathrm{spect}}}$ therefore differs from both the signal and $t\bar{t}$.

7.1.2 Overall fit setup

To summarize the decision process, three regions with the following differentiating characteristics are created:

- Signal region:
    - 1 or 2 jets b-tagged at 70% efficiency
    - NN_{bgr} <0.45
- Control region $t\bar{t}$:
    - 1 or 2 jets b-tagged at 70% efficiency
    - NN_{bgr} >0.45
- Control region V+jets:
    - 1 jet b-tagged at 85% efficiency but failing 70% efficiency

The yields in the three regions are shown in table 7.2. Figure 7.9 shows the fit distribution with the chosen binnings for all three regions. Figure 7.10 shows the associated $\Delta\gamma$ which are all within the set constraint.

The central characteristics of the statistics-only Asimov fit in this channel are shown in figure 7.11. The expected uncertainty on $\mu(tHq)$ is estimated with Minos (see section 4.3):

$$\Delta\mu(tHq)(\mathrm{stat.}) = {}^{+7.5}_{-7.5} \tag{7.1}$$

(a) SR

(b) CR $t\bar{t}$
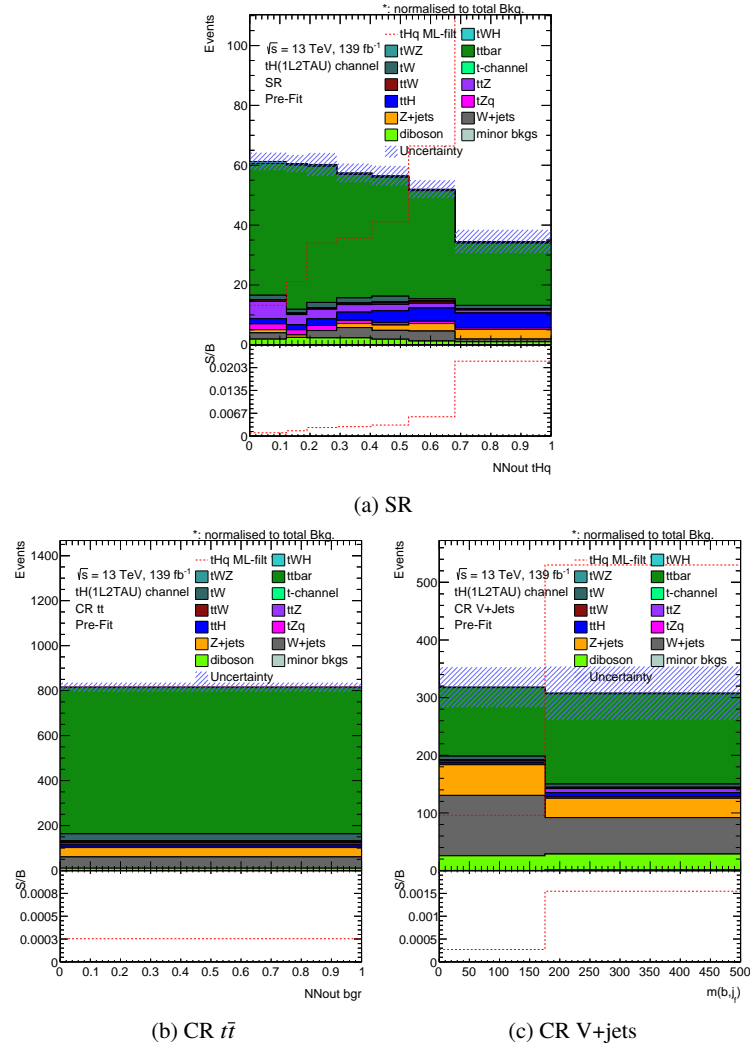
(c) CR V+jets

Figure 7.9: The three fit distributions in the $1e/\mu + 2\tau_{\mathrm{had}}$ channel.
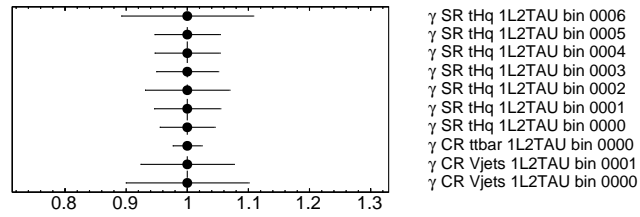
Figure 7.10: $\Delta\gamma$ of all bins used in the $1e/\mu + 2\tau_{\mathrm{had}}$ channel. The bins per region are numerated beginning with 0.

	SR	CR $t\bar{t}$	CR V+jets
tHq	1.71 $\pm$ 0.04	0.207 $\pm$ 0.013	0.563 $\pm$ 0.021
tWH	1.07 $\pm$ 0.05	0.399 $\pm$ 0.033	0.51 $\pm$ 0.04
tWZ	2.698 $\pm$ 0.030	1.224 $\pm$ 0.020	1.426 $\pm$ 0.022
$t\bar{t}$	276 $\pm$ 7	650 $\pm$ 10	275 $\pm$ 6
tW	9.9 $\pm$ 1.2	31.0 $\pm$ 2.1	11.8 $\pm$ 1.2
t-channel	0.86 $\pm$ 0.23	3.8 $\pm$ 0.4	1.39 $\pm$ 0.27
$t\bar{t}W$	3.38 $\pm$ 0.17	5.90 $\pm$ 0.23	2.63 $\pm$ 0.15
$t\bar{t}Z$	19.5 $\pm$ 0.5	8.65 $\pm$ 0.33	9.61 $\pm$ 0.33
$t\bar{t}H$	21.69 $\pm$ 0.18	10.47 $\pm$ 0.12	11.20 $\pm$ 0.13
tZq	8.62 $\pm$ 0.17	1.23 $\pm$ 0.07	3.01 $\pm$ 0.10
Z+jets	11 $\pm$ 5	42 $\pm$ 11	87 $\pm$ 21
W+jets	15 $\pm$ 5	50 $\pm$ 13	170 $\pm$ 50
diboson	12.9 $\pm$ 0.5	9.8 $\pm$ 0.7	52.9 $\pm$ 1.4
other samples	0.22 $\pm$ 0.04	0.55 $\pm$ 0.10	1.4 $\pm$ 0.7
total background	383 $\pm$ 9	815 $\pm$ 20	630 $\pm$ 60

Table 7.2: Yields of the $1e/\mu + 2\tau_{\text{had}}$ SR and CRs.
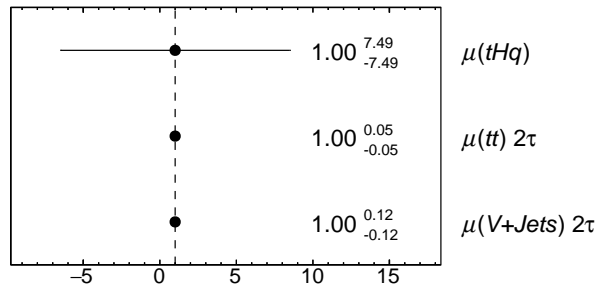
Figure 7.11: Fit results from an Asimov fit performed on all regions, but only taking into account the Asimov data and no NPs. '2τ' marks the NFs to be estimated in the $1e/\mu + 2\tau_{\text{had}}$ channel.

The uncertainties on $\mu(t\bar{t})$ and $\mu(\text{V+jets})$ are estimated with Migrad and are 5% and 12% respectively.

When taking into account the statistics of the simulated samples, the expected result is already less precise (see figure 7.12(a)):

$$\Delta\mu(tHq)(\gamma\text{-factors}) = ^{+9.6}_{-8.4} \tag{7.2}$$

The additional uncertainty shows a non-negligible of the $\Delta\gamma$. This likely stems from the fact that the fit is background dominated and some of our leading samples are limited in statistics. The correlations from this fit are shown in figure 7.12(b). While $\mu(t\bar{t})$ and $\mu(\text{V+jets})$ are highly correlated (according to the expectation), $\mu(tHq)$ is just moderately correlated with both.

	$\mu(V+Jets)$ 2τ	$\mu(tt)$ 2τ	$\mu(tHq)$
$\mu(V+Jets)$ 2τ	100.0	-63.5	-1.6
$\mu(tt)$ 2τ	-63.5	100.0	-27.6
$\mu(tHq)$	-1.6	-27.6	100.0

(a)   (b)
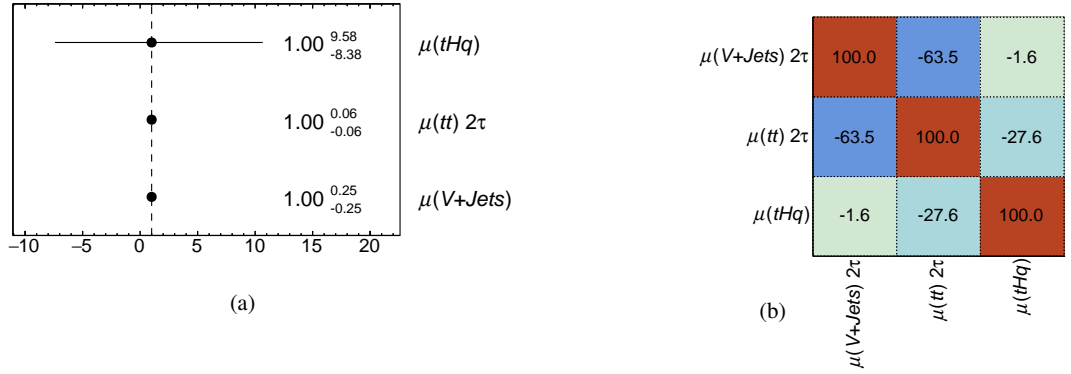
Figure 7.12: Fit results from an Asimov fit performed on all regions of the $1e/\mu + 2\tau_{\text{had}}$ channel, but only taking into account the Asimov data and the statistical uncertainty of the simulated samples.

7.2 2 e/μ OS + 1 τ_{had} fit setup

The $2e/\mu\text{OS} + 1\tau_{\text{had}}$ channel is similar to the $1e/\mu + 2\tau_{\text{had}}$ channel. It is dominated by the reducible backgrounds $t\bar{t}$ and Z+jets and contains between one and two signal events (see table 6.12). However, the amount of background is almost ten times more than in the $1e/\mu + 2\tau_{\text{had}}$ region. The S/B in the preselection region is only 0.01%. Therefore, the expected $\Delta\mu(tHq)$ from this channel is higher, but the strategy is still similar.

First, for a definition of the SR the cuts on NN_{bgr} and NN_{tHq} are compared in figure 7.13. Cuts on NN_{tHq} outperform cuts on NN_{bgr} in most of the possible points. A similar cut to the $1e/\mu + 2\tau_{\text{had}}$ channel was chosen: at around 1.7 events (in the SR) the S/B is tripled. In this range cuts on NN_{bgr} perform similar to cuts on NN_{tHq}, which is fortunate for the later use of NN_{tHq}. The chosen cut is done at $NN_{\text{bgr}} < 0.5$ where 1.7 signal events are left in the SR and only a third of the background. The resulting yields can be seen in table 7.3.

To handle the dominating reducible backgrounds $t\bar{t}$ and Z+jets, a similar strategy to the $1e/\mu + 2\tau_{\text{had}}$ channel is implemented. The baseline for the CRs is the inverted NN_{bgr} cut: $NN_{\text{bgr}} > 0.5$. In contrast to the $1e/\mu + 2\tau_{\text{had}}$ channel W+jets events do not play a role. Therefore, the NFs in this channel are $\mu(t\bar{t})$ targeting $t\bar{t}$ and $\mu(Z + \text{jets})$ targeting Z+jets. Additionally, the Z+jets sample already has a 36% contribution in the preselection region. This makes it easier to find a CR for the Z+jets sample. The
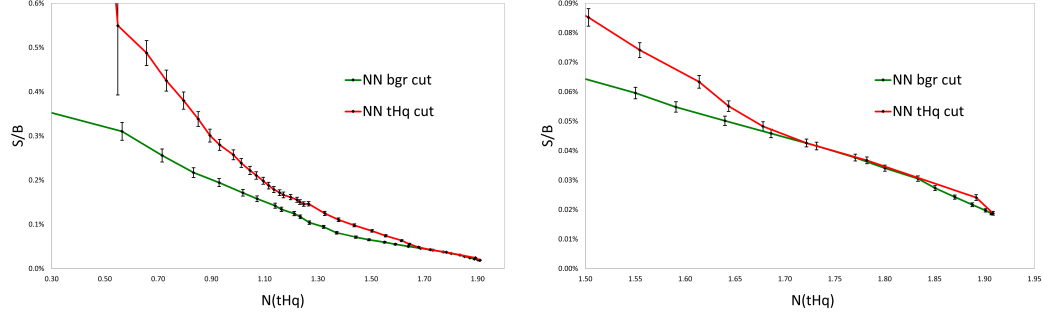
Figure 7.13: Possible cuts on NN_{bgr} and NN_{tHq} in the 2*e*/*µ*OS + 1τ_{had} channel. For each cut the left signal events in the SR is shown against S/B.

.

	SR		CR $t\bar{t}$		CR Z+jets	
tHq	1.7	$\pm$ 0.4	0.045	$\pm$ 0.014	0.18	$\pm$ 0.05
tWH	1.6	$\pm$ 0.4	0.20	$\pm$ 0.08	0.65	$\pm$ 0.15
tWZ	15	$\pm$ 8	0.7	$\pm$ 0.4	3.1	$\pm$ 1.7
$t\bar{t}$	2 020	$\pm$ 100	1 030	$\pm$ 50	2 600	$\pm$ 70
$t\bar{t}W$	37	$\pm$ 12	16	$\pm$ 5	24	$\pm$ 7
$t\bar{t}Z$	107	$\pm$ 31	11.8	$\pm$ 3.5	19	$\pm$ 5
$t\bar{t}H$	41	$\pm$ 12	8.2	$\pm$ 2.2	11.1	$\pm$ 3.0
tZq	36	$\pm$ 7	0.73	$\pm$ 0.19	3.2	$\pm$ 1.0
tW	76	$\pm$ 17	31	$\pm$ 6	148	$\pm$ 24
Z+jets	930	$\pm$ 120	181	$\pm$ 22	2 640	$\pm$ 310
diboson	94	$\pm$ 10	7.3	$\pm$ 1.0	75	$\pm$ 8
minor bkgs	5.2	$\pm$ 3.2	1.6	$\pm$ 1.1	7	$\pm$ 5
total background	3 360	$\pm$ 220	1 290	$\pm$ 60	5 500	$\pm$ 400

Table 7.3: Yields of 2*e*/*µ*OS + 1τ_{had} regions

described CR is simply divided by the number of *b*-tagged jets. The region where two of them are required is dominated by the $t\bar{t}$ sample. In the region with just one *b*-tagged jet Z+jets is close to 50%. All sample yields can be seen in table 7.3. The region definitions are summarized in table 7.4.

	Z+jets CR	$t\bar{t}$ CR	*tHq* SR
Number of *b*-tagged jets	1	2	1-2
NN_{bgr} cut	> 0.5	> 0.5	< 0.5

Table 7.4: All region definitions in the $2e/\mu\mathrm{OS} + 1\tau_{\mathrm{had}}$ channel.

The binning studies are performed independently in all three regions. Unsurprisingly, the best differentiating variable for the SR is NN_{tHq}. The automatic binning performs best focussing on the background samples (see above), which is explainable by the abundance of background events. However, the best performing binning creates 13 bins, some of them barely passing the threshold of ten events per bin. Manually, the bins are merged, resulting in only six bins. The number of events per bin allow the fit to be more stable, while the loss in $\Delta\mu(tHq)$ is less than 1%. The best performing variable for the CRs is $E_{\mathrm{T}}^{\mathrm{miss}}$. Its use decorrelates the $t\bar{t}$ and Z+jets NFs best. This is not surprising as neutrinos are mostly expected in $t\bar{t}$ samples. The resulting binned plots of all regions can be found in figure 7.14, the γ factors in figure 7.15.

From the fit only taking into account the Asimov data $\Delta\mu(tHq)$ is estimated to be:

$$\Delta\mu(tHq)(\mathrm{stat.}) = {}^{+15.7}_{-15.7} \tag{7.3}$$

When also taking into account the statistical uncertainties of the simulated samples the result is not impacted a lot:

$$\Delta\mu(tHq)(\gamma\text{-factors}) = {}^{+17.3}_{-17.3} \tag{7.4}$$

All results can be seen in figure 7.17. No high $\Delta\gamma$ are observed and as expected the correlation between the two background NFs is high (above 40%).

(a) SR

(b) CR $t\bar{t}$

(c) CR *Z*+Jets
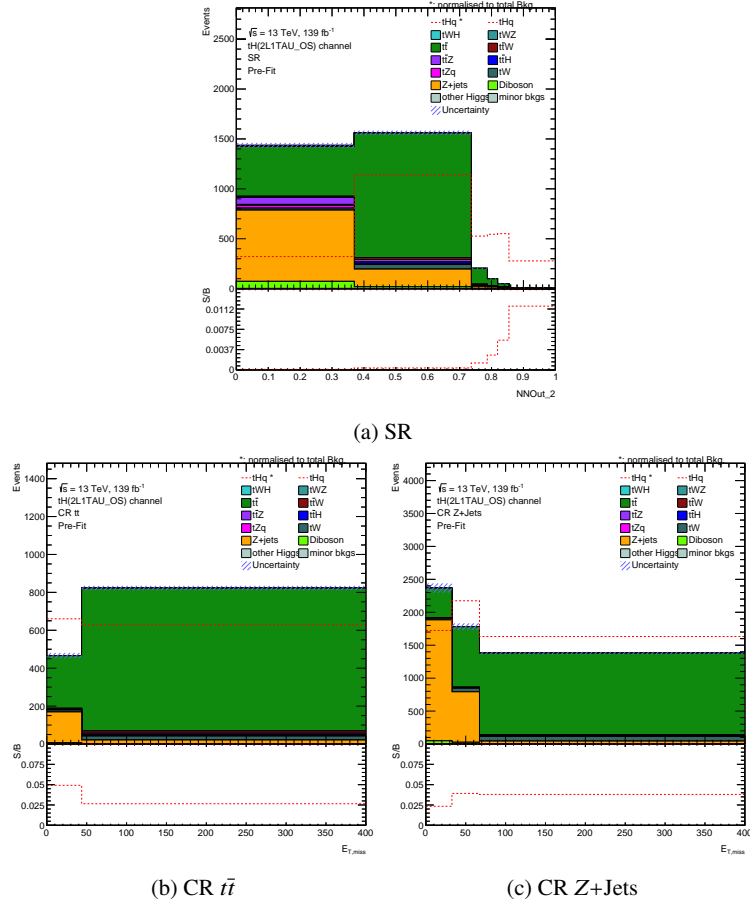
Figure 7.14: The three fit distributions in the $2e/\mu$OS + $1\tau_{\text{had}}$ channel.
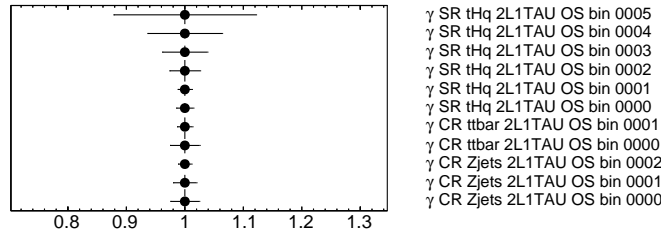
Figure 7.15: Gammas of all bins used in the $2e/\mu$OS + $1\tau_{\text{had}}$ channel.
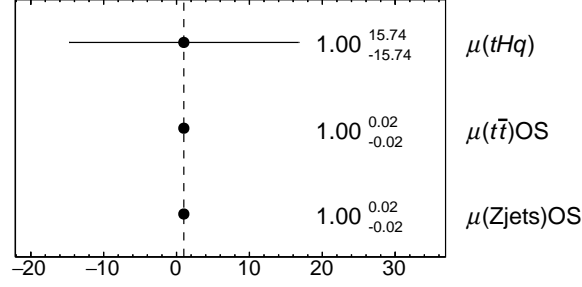
Figure 7.16: Fit result of the $2e/\mu$OS + $1\tau_{\text{had}}$ channel taking into account only the Asimov data and no NPs. The background NFs have an attached 'OS' to not confuse them with the $1e/\mu$ + $2\tau_{\text{had}}$ NFs.
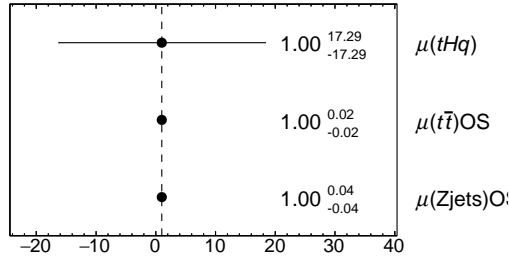
Figure 7.17: Fit results from an Asimov fit performed on all regions of the $2e/\mu$OS + $1\tau_{\text{had}}$ channel, but only taking into account the expected data from Asimov and statistical uncertainties of the simulated samples.

7.3 2 e/μ SS + 1 τ_{had} fit setup

Samples in the $2e/\mu$SS + $1\tau_{\text{had}}$ channel are differently distributed than in the other two described channels. The only reducible background which is dominant is the $t\bar{t}$ sample. Otherwise, irreducible samples are dominating, especially $t\bar{t}Z$, $t\bar{t}W$ and $t\bar{t}H$. This is mostly a result of the same-sign requirement which reduces the reducible backgrounds $t\bar{t}$ and Z+jets. Therefore, the overall background statistics is lower. The S/B in the preselection region is the highest of all described channels with 0.97%. As described in section 6.3.1 the performance of the dedicated NN is not as strong as for the other channels. Therefore, the NN trained on $2e/\mu$OS + $1\tau_{\text{had}}$ events is used. Figure 7.18 shows the possible SR definition cuts on NN_{tHq}. As described above at least one expected signal event in the SR is required. This exact number is achieved with a cut on NN_{tHq} >0.4. At this margin roughly half of the background events are still in the SR, which is worse than for the other channels. However, due to the relatively clean preselection region the S/B in the signal region is still the best with 1.36%.

Only one CR is defined with the inverse cut NN_{tHq} <0.4. Splitting this CR according to the number of *b* tagged jets is not helpful as all dominating backgrounds are expected to contain two *b*-tagged jets per event. Additionally, the statistics in the CR is not high, which would only be diminished by a further
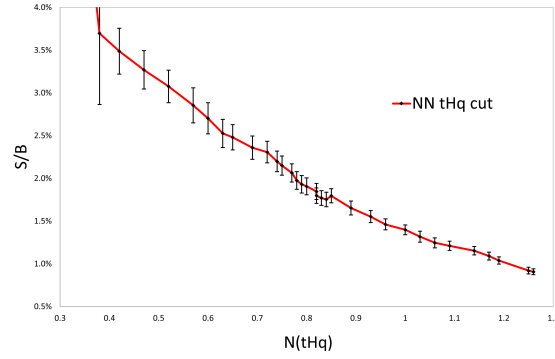
Figure 7.18: Possible cuts on NN_{tHq} in the $2e/\mu$SS + 1τ_{had} channel. For each cut the left signal events in the SR is shown against S/B.

.

splitting. No additional NFs are associated with the one defined CR. Instead, the CR is used to evaluate the NPs and decorrelate the $t\bar{t}$ sample from the irreducible backgrounds. This is achieved by binning in the transverse component of the sum of the momenta from all objects used in the analysis H_T. However, another consequence of the low expected statistics is the low number of bins, which can be created and still pass the minimal criteria for statistics per bin. This only allows for a low number of bins and favors a signal-focussed binning (see figure 7.3). The resulting binned distributions are shown in figure 7.19, the associated $\Delta\gamma$ in section 7.3.

	SR	CR
tHq	1.00 $\pm$ 0.04	0.261 $\pm$ 0.019
tWH	0.62 $\pm$ 0.05	0.36 $\pm$ 0.04
tWZ	1.516 $\pm$ 0.025	1.195 $\pm$ 0.022
$t\bar{t}$	14.7 $\pm$ 1.5	12.9 $\pm$ 1.3
$t\bar{t}W$	14.7 $\pm$ 0.4	22.9 $\pm$ 0.5
$t\bar{t}Z$	13.0 $\pm$ 0.4	9.08 $\pm$ 0.29
$t\bar{t}H$	15.52 $\pm$ 0.18	10.35 $\pm$ 0.14
tZq	4.06 $\pm$ 0.13	1.60 $\pm$ 0.08
tW	1.1 $\pm$ 0.4	0.51 $\pm$ 0.26
Z+jets	0.24 $\pm$ 0.20	0.64 $\pm$ 0.19
W+jets	0.46 $\pm$ 0.29	0.22 $\pm$ 0.19
diboson	6.2 $\pm$ 0.4	4.87 $\pm$ 0.33
other samples	1.6 $\pm$ 0.8	0.72 $\pm$ 0.09
total background	73.7 $\pm$ 1.9	65.3 $\pm$ 1.6

Table 7.5: Yields of $2e/\mu$SS + 1τ_{had} regions

(a) SR

(b) CR
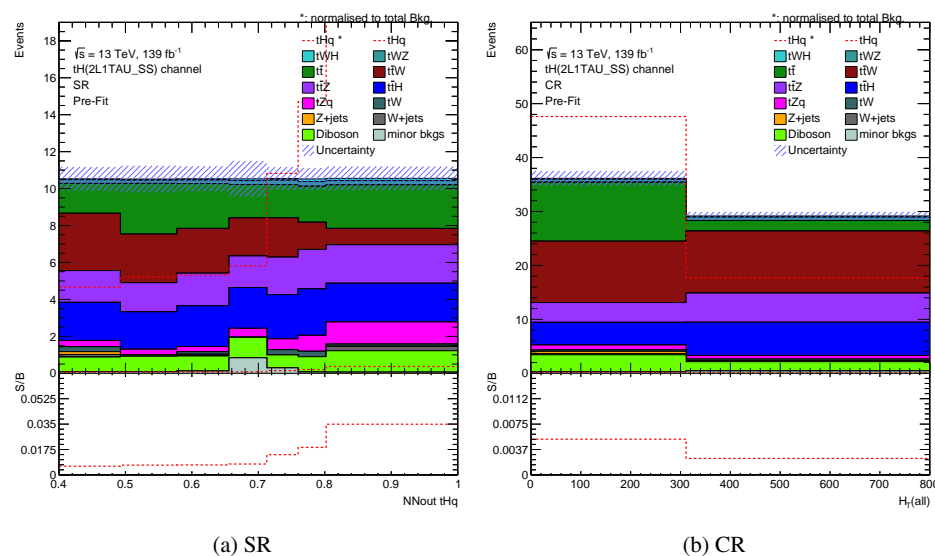
Figure 7.19: The two binned distributions used in the $2e/\mu\mathrm{SS} + 1\tau_{\mathrm{had}}$ channel fit.

(a) Taking into account only the Asimov data.

(b) Taking into account the Asimov data and $\Delta\gamma$.
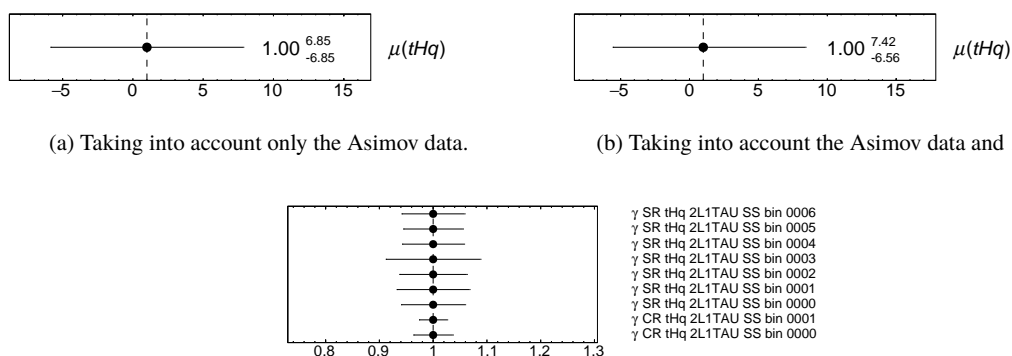
Figure 7.20: Fit results from an Asimov fit in the $2e/\mu\mathrm{SS} + 1\tau_{\mathrm{had}}$ channel performed on all regions, but only taking into account statistical information.

The result of the fit not taking into account any systematic uncertainties gives

$$\Delta\mu(tHq)(\mathrm{stat.}) \quad = {}^{+6.9}_{-6.9} \tag{7.5}$$

$$\Delta\mu(tHq)(\gamma\text{-factors}) = {}^{+7.4}_{-6.6} \tag{7.6}$$

which can also be seen in figure 7.20.

7.4 Systematic uncertainties in the fit

The topic of systematic uncertainties has been touched several times already, for example in section 4.3 and section 7.1.1. This section gives a more organized overview of the treatment of the different systematic

uncertainties.

The most important differentiation is the source of the uncertainty. One source of uncertainties are simulation and theoretical knowledge about the different processes. A complete list of all systematic uncertainties related to the different samples can be found in table B.7 and table B.8. Depending on the sample this can be just a cross-section uncertainty, which is determined by theory. For other samples the evaluation is done more thoroughly. In this case several independent uncertainties are related to one sample. Additional to the cross-section uncertainty, the samples might be produced with different generators or with different settings for one generator. Common variations are done on the initial state radiation, final state radiation as well as the refactorization and renormalization scale. The last two sources are treated differently from the others, as different scales have to be taken into account. To keep the number of NPs entering the fit small, they are combined per sample. For that, both scales are evaluated together in seven different settings (additional to the default setting) and the envelope of all variations is used for the final fit. As $t\bar{t}$ is one of the dominating samples in all three channels the uncertainties associated with this sample are explained as an example:

- **Cross-section:** overall variation of $\pm 6\%$
- **Parton shower and hadronization:** comparing the Pythia to a Herwig sample (both samples are generated with AFII)
- **NLO generator:** comparing a Powheg and a Madgraph sample (both samples are generated with Herwig and AFII)
- **Renormalization and refactorization scale:** comparing seven different configurations with respect to the standard configuration ($\mu_R - \mu_F = $ 0.5-0.5; 0.5-1; 1-0.5; 1-1; 1-2; 2-1; 2-2) and taking the envelope
- **hdamp:** damping of the singular shower contributions in Powheg [59]; varied to three times the top mass, this is double the default scale (both samples are generated with AFII)
- **ISR:** Initial State Radiation variation
- **FSR:** Final State Radiation variation

For all samples associated with a NF, only the shape aspect of a NP is taken into account. An example of this is shown in figure 7.21 (figure 7.21(b) to figure 7.21(c)). This also means that the theoretical cross-section uncertainty of for example $t\bar{t}$ is completely omitted. Samples which are subsumed under 'other samples' are varied together by an overall up and down variation of 50%.

A second source of systematic uncertainties are originating in the object reconstruction. Complete lists are given in tables B.1 to B.6. Some context to the origin of the different uncertainties is given in the respective part of section 5.4. The statistical uncertainty on the fake estimation is also described in table B.5. Two additional systematic uncertainties are attached to the luminosity ($\pm 1.7\%$ [126]) and the pile-up structure (see section 6.1). All these uncertainties impact all samples. However, some uncertainties are found to be different for samples which underwent the full Geant4 chain and samples which utilized AFII. In this case the evaluation is also done separately.

As described in section 4.3 the NPs are allowed to vary within a Gaussian constraint given by previous experiments. The comparison is done between different samples. These are generated for example by different generators, varying the cross-section, changing weights associated to different objects or even the whole reprocessing of the event under different reconstruction parameters. In case just a variation in one direction is available, the impact is symmetrized into an up and a down variation. An example of this is shown in figure 7.21(b) in comparison to figure 7.21(a).

In total (before the envelope) there are about 270 NPs (for each channel). Each of them is evaluated for all bins in all regions, most also for all samples, separately. To mitigate the impact of statistic fluctuations on the result, all NPs which have a shape component are smoothed. In that, bins with large fluctuations are merged, then the histogram is smoothed according to [188]. An example for smoothing is shown in figure 7.21 in comparison of figure 7.21(c) to figure 7.21(d).
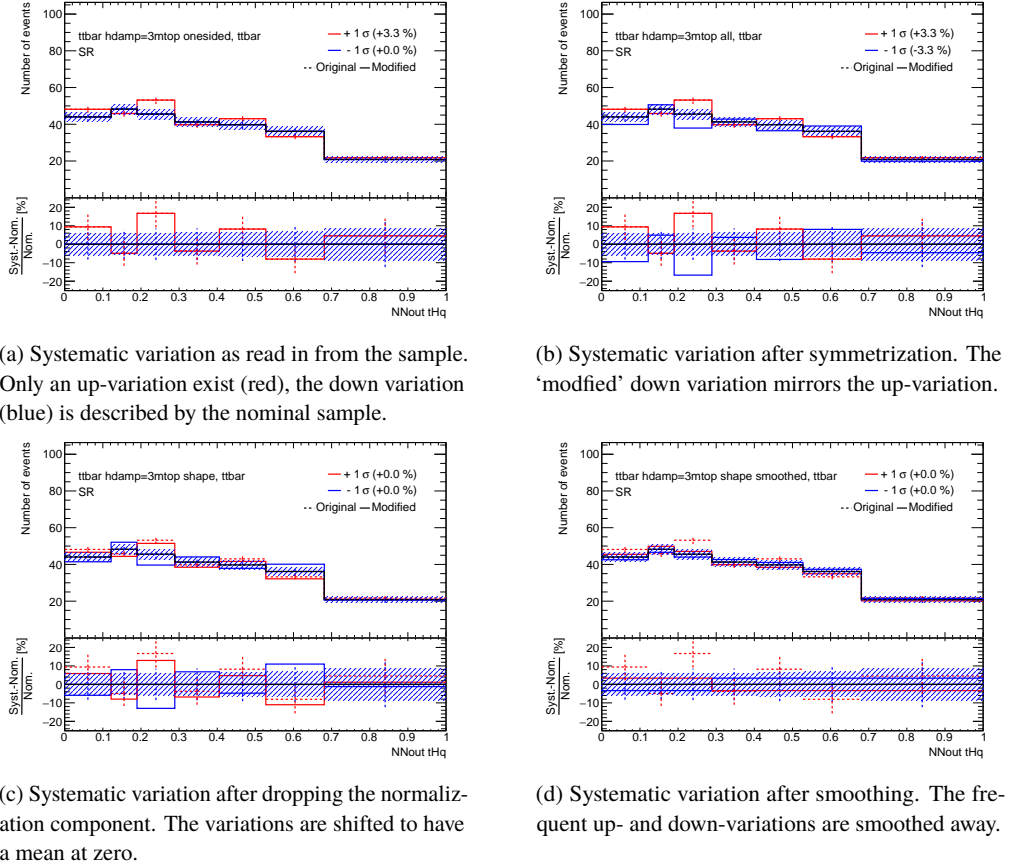

(a) Systematic variation as read in from the sample. Only an up-variation exist (red), the down variation (blue) is described by the nominal sample.

(b) Systematic variation after symmetrization. The 'modfied' down variation mirrors the up-variation.

(c) Systematic variation after dropping the normalization component. The variations are shifted to have a mean at zero.

(d) Systematic variation after smoothing. The frequent up- and down-variations are smoothed away.

Figure 7.21: Different stages of processing NPs before the fit, shown on the example of the $t\bar{t}$ hdamp NP in the $1e/\mu + 2\tau_{\text{had}}$ SR. (a) shows the original variation, (b) shows the systematic variation after symmetrization, additionally in (c) the normalization component is dropped and (d) shows the final used variation after smoothing. The red crosses depict the original up-variations with uncertainties, the blue crosses the down variation (in this example all on the nominal line). The respective continuous lines depict the systematic variation after all modifications.

Some of the multitude of variation-sample-channel combinations do not impact the fit. For the sake of stability those NPs are omitted. This process is called pruning. The shape and normalization impact are pruned separately if the impact in this specific variation-sample-channel combination is smaller than 1% and 0.5% respectively. The complete plots which indicates which NPs have been pruned are collected in appendix G. Most of the systematic uncertainties are pruned in normalization and shape for at least a few samples. Some of them even for all combinations. Note that NPs which are not used after pruning will not appear in any further plots.

7.5 Asimov fit results

This section shows the results of fits which use Asimov data in all regions, including the *tHq* sample with its SM expectation value. This setup is the most comparable to the final fit with data and is used to evaluate the impact of the different NPs.

7.5.1 Asimov fit result in the 1 e/μ + 2 τ_{had} channel

When incorporating the systematic uncertainties as NPs into the fit, the first step is to evaluate the constraint evoked by the fit on every NP. As described in section 4.3, the expectation is that most NPs are found to be fully constrained by the given uncertainty. Just in cases where the fit is able to find additional information on that NP in the fit dataset the fit should constrain the NP further. This can be the case because the fit dataset is more specific than the dataset that has been evaluated before or because the fit is sensitive to the variation invoked by that NP.

Figure 7.22 shows the constraints on all NPs which survived the pruning. Only the NP associated with $t\bar{t}$ generator variation (NLO generator shape, shown in figure 7.22(e)) is constrained more than a few percent. This uncertainty is expected to impact the fit as $t\bar{t}$ is the dominating background. In figure 7.23 the shape variation of that alternative sample is shown. Other NPs which are slightly constrained are other $t\bar{t}$ related NPs and jet related NPs. As the phase space contains multiple jets, several variables are dependent on the jet and jets are the primary source of fake τ_{had}, this is not surprising.

The correlations between the NPs and NFs are shown in figure 7.24. Only NPs which are correlated with at least one other NP or NF more than 15% are shown. Apart from the correlation of the NFs (see section 7.1.1.1) most correlations are between the jet-related NPs and the background-related NFs. As discussed above this is expected, mostly due to the impact of jets on fakes. The POI has the highest correlation with the leading jet energy resolution uncertainty, probably due to the same effect. One correlation is found between the PS variation in the $t\bar{t}$ and tW sample. Since both samples interfere (see section 2.2.2) this is also not worrisome.

This leads to the first results with all NPs. Figure 7.25 shows the evaluated $\Delta\mu(tHq)$, $\Delta\mu(t\bar{t})$ and $\Delta\mu(\text{V+jets})$ from the Asimov fit including all NPs. The uncertainties of the background NFs are doubled ($\Delta\mu(t\bar{t})$) and tripled ($\Delta\mu(\text{V+jets})$) in comparison to the fit without all NPs. Fortunately the impact on $\Delta\mu(tHq)$ is not as large:

$$\Delta\mu(tHq)(\text{Asimov}) = {}^{+11.3}_{-10.1} \tag{7.7}$$

Still, in comparison to the result incorporating only the expected statistics in equation (7.1) $\Delta\mu(tHq)$ is about doubled. Since this fit is including more uncertainties, this is expected. Nevertheless, it is interesting to study the source of uncertainty.

7.5.1.1 NP impact

One way of studying the source of high $\Delta\mu$ is to remove groups of NPs from the fit and study the fit result. Table 7.6 shows $\Delta\mu(tHq)$ (evaluated with Migrad) in those cases. First, the simulation statistics effect is removed from the fit. Therefore, the fit result becomes more sensitive and the impact can be calculated by quadratic subtraction. The same is done with all groups of NPs shown in figure 7.22. This approach cannot correctly determine the correlations between the NPs of the different groups, but it can give an idea about the influence each group of NPs has. From table 7.6 it becomes clear that while the Asimov data has

(a) experimental

(b) JES/JER

(c) flavor-tagging

(d) theory

(e) theory $t\bar{t}$
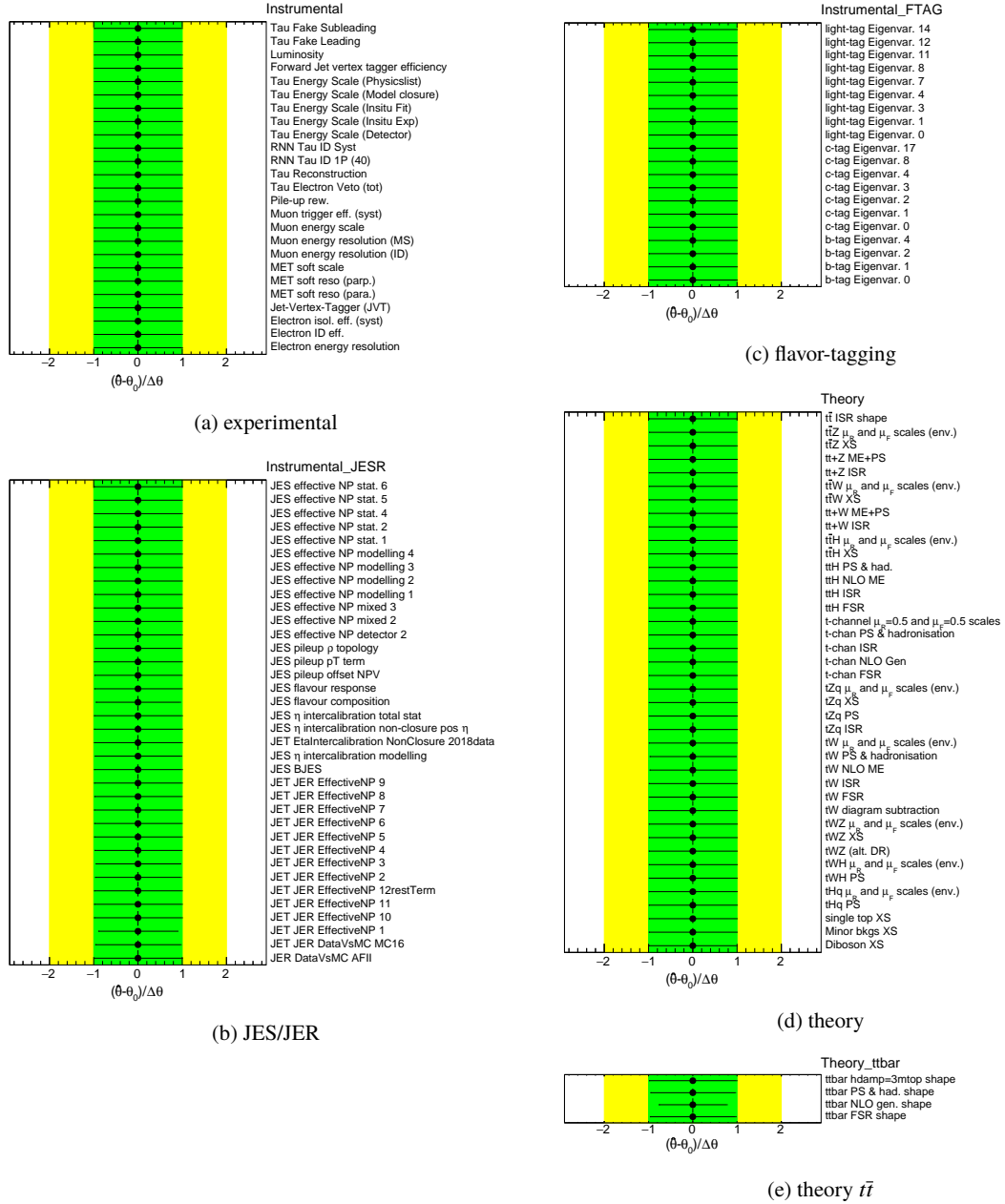
Figure 7.22: Constraints of all NPs in the $1e/\mu + 2\tau_{\mathrm{had}}$ region after the Asimov fit. The NPs are roughly grouped by their origin.

a large impact, the limited simulated data statistics is also a big factor. All other NP groups have a smaller impact.

The same procedure can be done for individual NPs as well. Figure 7.26 shows the NPs which have the biggest impact on the fit result. This plot is a bit more complex as it contains three pieces of information for each NP: the constraint, the pre- and post-fit impact. The constraint of the NPs is the same as shown in
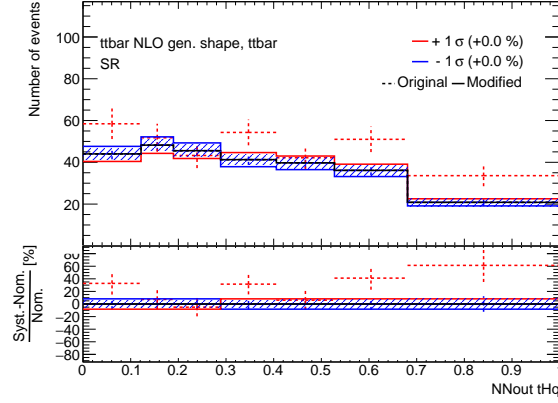
Figure 7.23: Variation of the $t\bar{t}$ NLO Generator NP in the $1e/\mu + 2\tau_{\text{had}}$ SR.
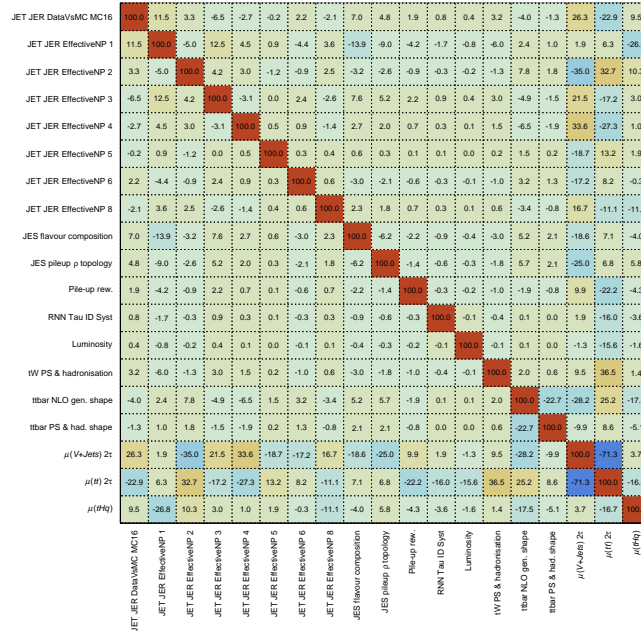
Figure 7.24: Correlations between NFs and NPs in the $1e/\mu + 2\tau_{\text{had}}$ channel. Only NPs which have at least one correlation above 15% are included.

	Full fit	stat.	γ-factors	all other NPs
impact	10.68	7.5	5.25	7.92

	experimental	flavor-tagging	JES/JER	theory	theory $t\bar{t}$	NFs
impact	1.55	0.27	3.92	1.94	2.57	2.14

Table 7.6: Impact of different groups of NPs on the $1e/\mu + 2\tau_{\text{had}}$ fit on Asimov data.
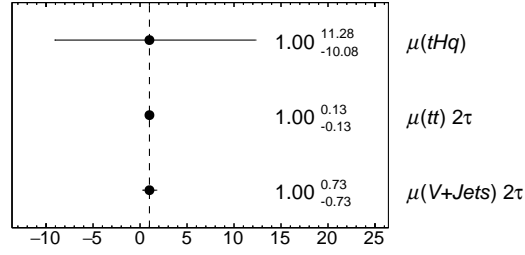
Figure 7.25: Fit result for the Asimov fit in the $1e/\mu + 2\tau_{\text{had}}$ channel, taking into account all NPs.

figure 7.22 and figure 7.4(b) and depicted by the black bar. The scale for this is given by the lower x-axis. The pre- and post-fit impact is associated with the upper x-axis scale, which shows the impact on $\Delta\mu(tHq)$. For the pre-fit impact the NP is fixed to the nominal value and a fit is performed. This result is compared to the full fit and the impact is depicted in the empty blue box. For the post-fit impact the NP is instead constrained to the post-fit estimate. As this fit uses Asimov data, the pre- and post-fit impact only differs if the fit was able to constrain the NP. The NPs are then ordered by their influence on their final result and the 20 most influential NPs are shown in the ranking plot. In this case the ranking plot (figure 7.26) shows a break down of the results in table 7.6. The limited simulated statistics of events in the most sensitive SR bin is the dominating uncertainty. Furthermore, from the jet related NPs the leading jet energy resulution NP (NP1) is dominating the fit. The leading $t\bar{t}$ related uncertainty is related to the NLO generator and has already been discussed above.
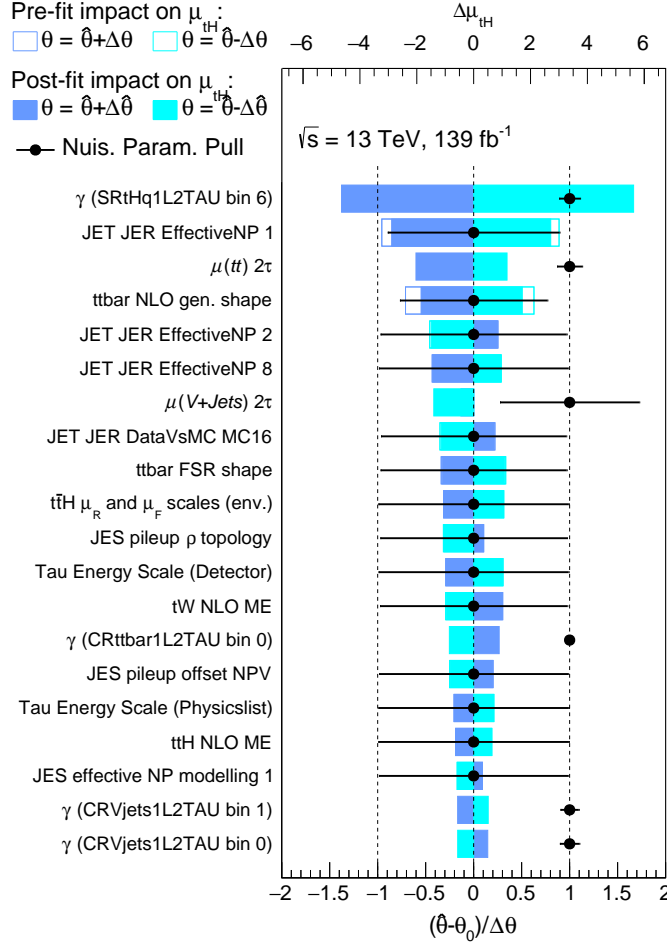
Figure 7.26: Ranking of the 20 most impactful NPs on the Asimov fit in the $1e/\mu + 2\tau_{\mathrm{had}}$ channel.

7.5.1.2 Inverted coupling and limit

The fit described above focuses on measuring the *tHq* process. However, as figure 7.25 shows how big $\Delta\mu(tHq)$ is in case the SM holds. But what if the SM has to be altered? As described in section 3.1 in case k_t is altered the expected *tHq* cross-section rises. The biggest effect can be seen if $k_t = -k_W$, this scenario is called 'inverted coupling'. Therefore, *tHq* and *tWH* samples are tested which are simulated under the inverted coupling hypothesis. These samples are inserted into the same fit procedure as described above and the inverted-coupling $\mu(tHq)$ is measured. The result of the fit is shown in figure 7.27. It becomes apparent that in this case $\mu(tHq)$ can be measured with a higher precision:

$$\Delta\mu(tHq)(\text{Asimov,ic}) = (\text{stat})^{+1.4}_{-1.4}(\text{tot})^{+2.4}_{-2.4} \tag{7.8}$$

Both the fit only taking into expected data and the fit including all NPs profit from the increased signal sample cross-section.

Another approach is to estimate how much the fit is able to limit the *tHq* cross-section. The procedure
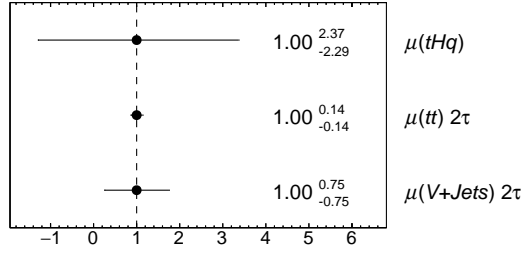
Figure 7.27: Fit result for the Asimov fit in the $1e/\mu + 2\tau_{\text{had}}$ channel under the inverted coupling hypothesis, taking into account all NPs.

for this has already been described in section 4.4. The expected limit from the $1e/\mu + 2\tau_{\text{had}}$ channel is:

$$\text{limit(exp.)} = 23.5 \left(\pm 1\sigma^{34.5}_{16.9} \pm 2\sigma^{49.9}_{12.6} \right) \mu(tHq)(\text{SM}) \tag{7.9}$$

The variations describe the 1σ and 2σ deviations of that limit in both directions. More limits and significances are shown in appendix G. In case the signal sample is described by the inverted coupling, the limit falls to:

$$\text{limit(exp.)} = 4.6 \left(\pm 1\sigma^{6.6}_{9.2} \pm 2\sigma^{3.3}_{2.5} \right) \mu(tHq)(\text{IC}) \tag{7.10}$$

This shows the presented fit setup is sensitive to the inverted coupling *tHq* sample. More information from fits under the inverted coupling hypothesis are shown in appendix J.

7.5.2 Asimov fit result in the 2 e/μ OS + 1 τ_{had} channel

The same studies as above have also been performed in the $2e/\mu\text{OS} + 1\tau_{\text{had}}$ channel. The results concerning the NP constraints and correlations can be found in figures 7.28 and 7.29. More NP plots are shown in appendix G Overall, the effects look similar to the $1e/\mu + 2\tau_{\text{had}}$ channel: $t\bar{t}$ and jet related NPs are correlated with the background-related NFs and partially constrained. Additional small constraints are visible in the b-tagging related NPs. This can be due to the CR definitions, which depend on the b-tagging. $\mu(tHq)$ has the highest correlation with the $t\bar{t}$ FSR NP. Figure 7.30 shows the $t\bar{t}$ variation due to this sample change in the SR. It becomes visible that NN_{tHq} deems the FSR variation of $t\bar{t}$ as more signal-like. Therefore, an about 20% correlation of this NP is totally reasonable.

The $2e/\mu\text{OS} + 1\tau_{\text{had}}$ fit result with all NPs applied are shown in figure 7.31. In comparison to the $1e/\mu + 2\tau_{\text{had}}$ channel $\Delta\mu(tHq)$ is higher, but in comparison to the fit without NPs the change is in the same range:

$$\Delta\mu(tHq)(\text{Asimov}) = {}^{+22.3}_{-22.3} \tag{7.11}$$

$\Delta\mu(t\bar{t})$ and $\Delta\mu(Z + \text{jets})$ are 4% and 10% respectively. Both numbers are significantly lower than for the comparable uncertainties in $1e/\mu + 2\tau_{\text{had}}$. This is explainable by the fact that the backgrounds have a factor of 10 more statistics in this channel. Additionally, there is no need to simultaneously estimate W+jets, which is another sample with a low number of simulated events in these channels.
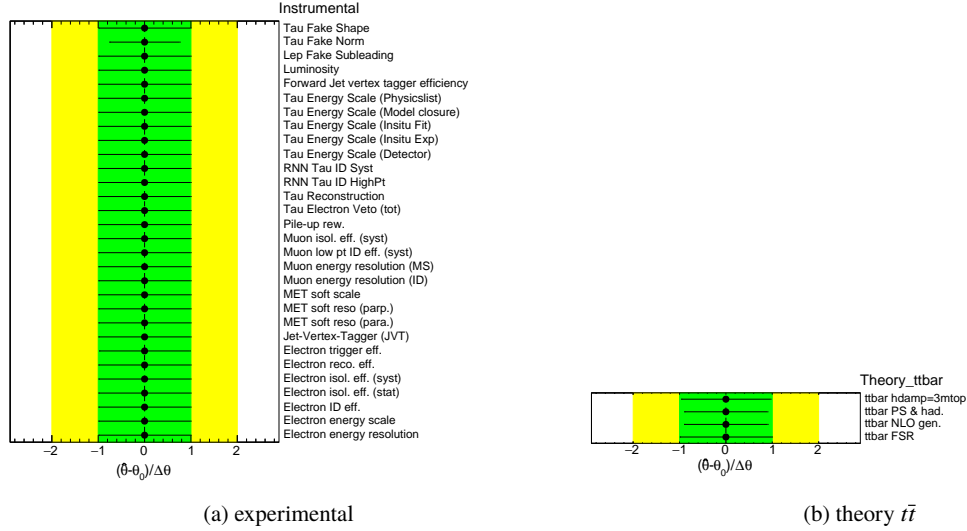
(a) JES/JER

(b) theory $t\bar{t}$
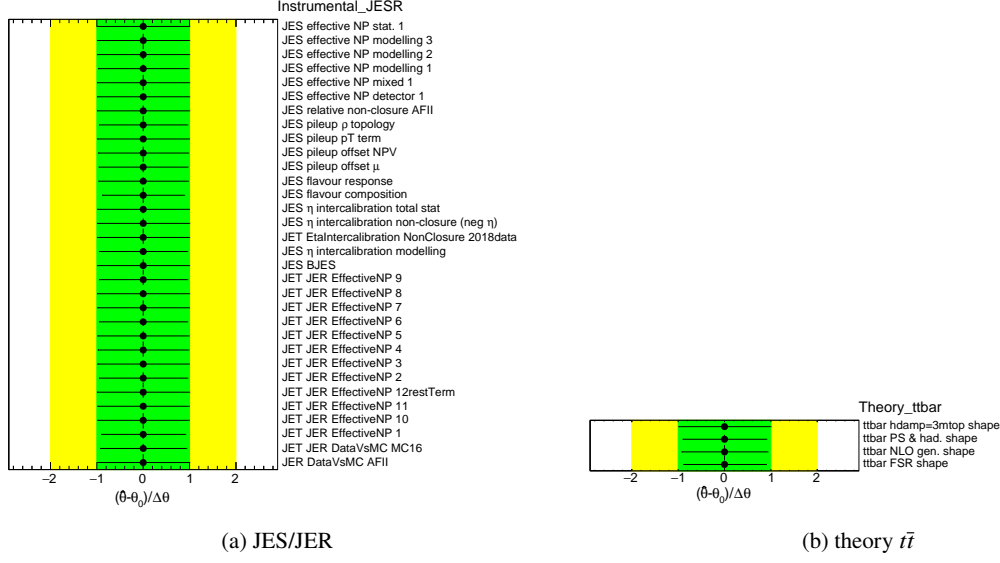
Figure 7.28: Some constraints of the NPs in the $2e/\mu$OS + $1\tau_{\text{had}}$ region after the Asimov fit. The NPs are roughly grouped by their origin. All constraints are shown in appendix G.

The impact of the groups of NP are shown in table 7.7, the ranking of individual NPs in figure 7.32. In comparison to the $1e/\mu$ + $2\tau_{\text{had}}$ channel (table 7.6 and figure 7.26) the influences of all NP groups are enhanced. However, this is a combined effect of the direct impact and the correlations. Regarding the individual NP rankings, the $t\bar{t}$ FSR NP has the highest rank. This is expected due to the above described shape effect. Other highly ranked NPs are mainly related to jets, the limited simulated statistics and other $t\bar{t}$ related NPs.

	Full fit	stat.	γ-factors	all other NPs
impact	22.3	15.7	7.6	16.6

	experimental	flavor-tagging	JES/JER	theory	theory $t\bar{t}$	NFs
impact	3.1	0.9	6.8	3.9	11.7	4.1

Table 7.7: Impact of different groups of NPs on the $2e/\mu$OS + $1\tau_{\text{had}}$ channel fit.
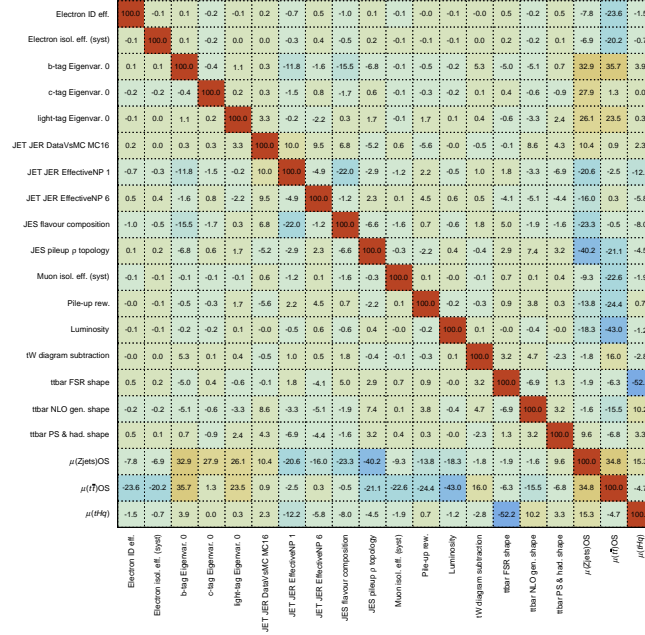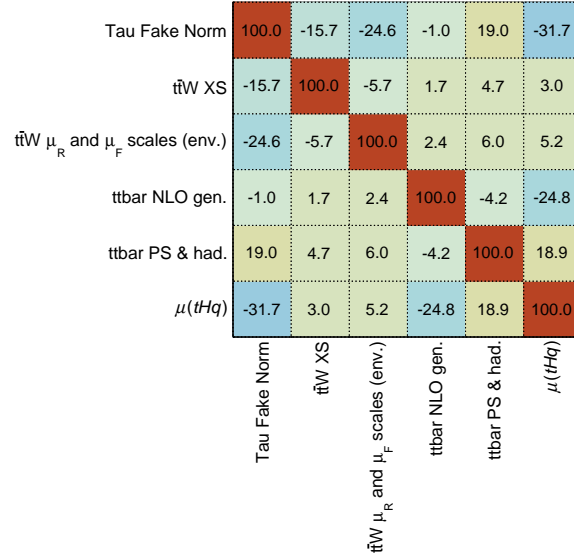
Figure 7.29: Correlations between NFs and NPs in the $2e/\mu$OS + $1\tau_{\text{had}}$ channel. Only NPs which have at least one correlation above 15% are included.
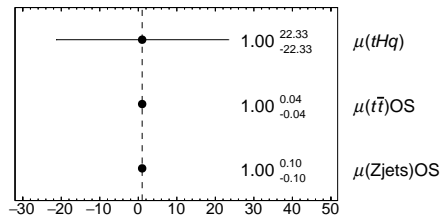
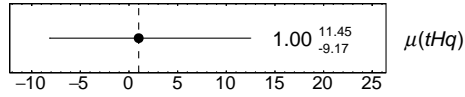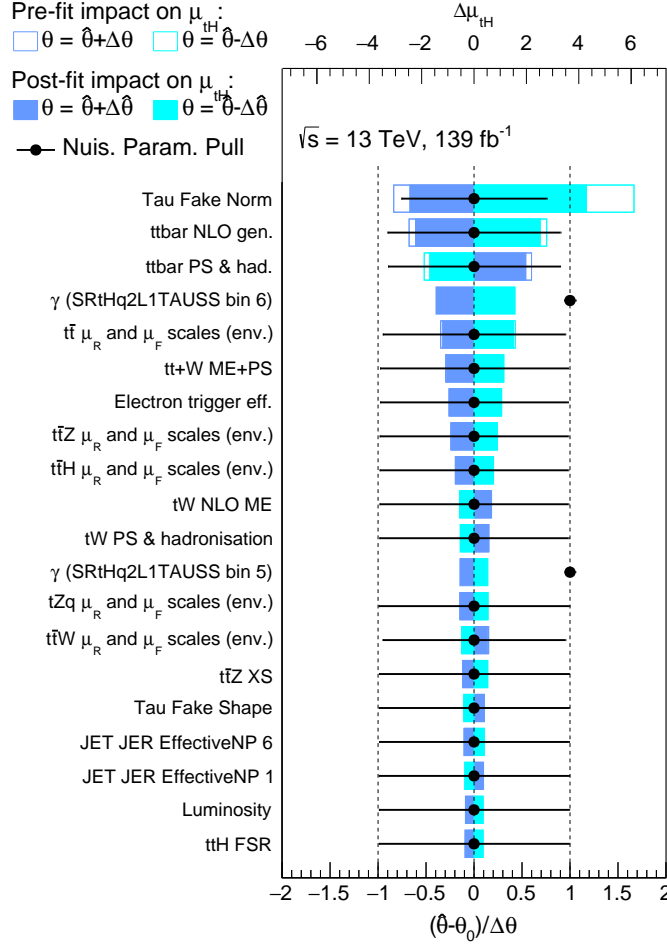Figure 7.30: Fit result for the Asimov fit in the $2e/\mu$OS + $1\tau_{\text{had}}$ channel, taking into account all NPs.
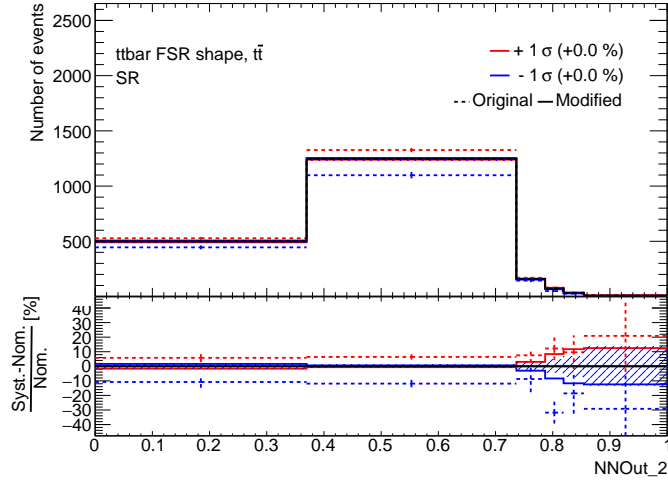
Figure 7.31: Variation associated with the $t\bar{t}$ FSR NP in the $2e/\mu$OS + $1\tau_{had}$ SR.
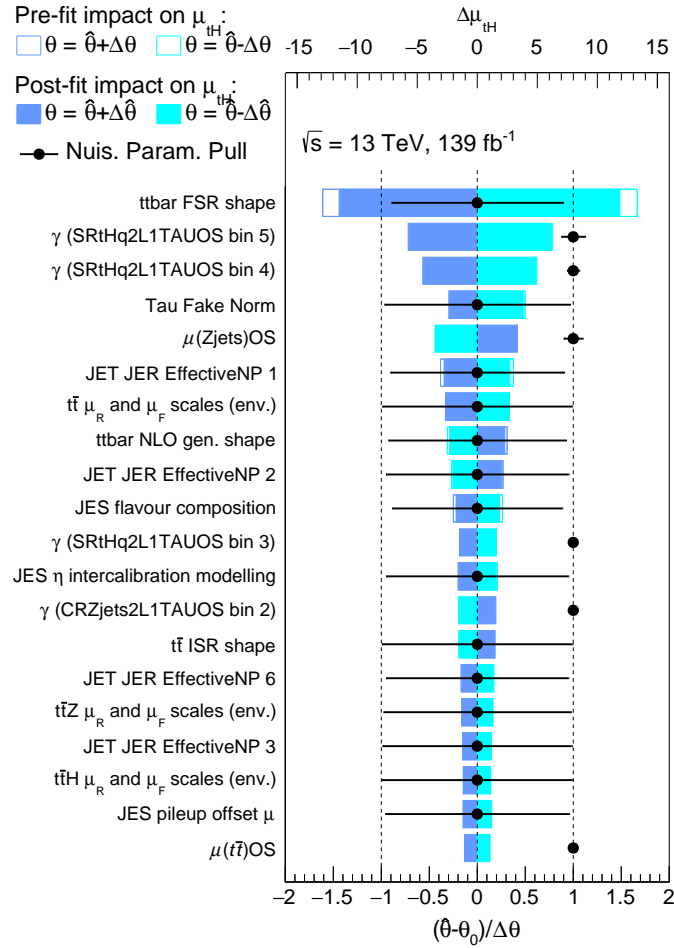
Figure 7.32: Ranking of the 20 most impactful NPs in the Asimov fit in the $2e/\mu$OS + $1\tau_{had}$ channel.

7.5.3 Asimov fit result in the 2 *e*/*μ* **SS + 1** τ_{had} **channel**

Since the $2e/\mu \text{SS} + 1\tau_{\text{had}}$ channel depends less on the reducible backgrounds it is expected that the full Asimov fit is less influenced by their variations. This is validated by figure 7.33 and figure 7.34. Noticeable is the overall smaller correlation matrix, which means fewer NPs have at least one 15% correlation with another NP or $\mu(tHq)$. The NPs which were present in figure 7.24 and figure 7.29, but are missing here were mostly jet related NPs which were correlated with the background related NFs. As no additional NFs are included in this channel, their absence is expected. However, $t\bar{t}$ related NPs are still highly correlated with $\mu(tHq)$ and also constrained. Additionally, $t\bar{t}W$ related NPs show a similar behavior. This is unique to this channel, as this is the only channel where $t\bar{t}W$ is one of the leading backgrounds. Another NP is highly correlated and constrained: 'Tau Fake Norm'. This NP originates in the fake estimation and describes the impact of the possible normalization variations on the fake estimate (similar to the shape and normalization influence of NPs described for this fit). The constraint hints at the fact that it is possible to learn more information on the correct fake estimation from this fit. Therefore, the uncertainty on the fake estimation was too big. This might originate in the low statistics present in the $2e/\mu\text{SS} + 1\tau_{\text{had}}$ fake estimation.

(a) experimental   (b) theory $t\bar{t}$

Figure 7.33: Some constraints of all NPs in the $2e/\mu\text{OS} + 1\tau_{\text{had}}$ region after the Asimov fit. The NPs are roughly grouped by their origin. All Constraints can be found in appendix G.

The result of the Asimov fit including all NPs is:

$$\Delta\mu(tHq)(\text{Asimov}) = {}^{+11.5}_{-9.2} \tag{7.12}$$

This result is comparable to the $1e/\mu + 2\tau_{\text{had}}$ channel result. Therefore, also the impact of the NP groups in table 7.8 can be compared to that fit. In comparison, the $\Delta\gamma$ impact is small. However, the experimental NPs (which include the fake estimation) and the impact from theory are enhanced, resulting in a comparable result. This is confirmed by the individual NP rankings: the above described 'Tau Fake Norm' NP is the highest ranked NP, followed by theory related NPs.

Figure 7.34: Correlations between NFs and NPs in the $2e/\mu\text{SS} + 1\tau_{\text{had}}$ channel. Only NPs which have at least one correlation above 15% are included.

Figure 7.35: Fit result for the Asimov fit in the $2e/\mu\text{SS} + 1\tau_{\text{had}}$ channel, taking into account all NPs.

	Full fit	stat.	γ-factors	all other NPs
impact	10.3	6.9	1.7	7.3

	experimental	flavor-tagging	JES/JER	theory	theory $t\bar{t}$
impact	3.7	0.2	0.8	2.9	3.1

Table 7.8: Impact of different groups of NPs on the $2e/\mu\text{SS} + 1\tau_{\text{had}}$ channel fit.

Figure 7.36: Ranking of the 20 most impactful NPs in the Asimov fit in the $2e/\mu SS + 1\tau_{had}$ channel.

7.5.4 Asimov fit result of the channel combination

Since the results from all three channels are known, the results can also be combined. The statistical power of the combination is expected to be higher than for the single channels. It is driven by the most sensitive result. In the fit of all three channels together, only $\mu(tHq)$ is varied across all regions. $\mu(V+jets)$ and $\mu(Z+jets)$ are just used for the $1e/\mu + 2\tau_{had}$ and $2e/\mu OS + 1\tau_{had}$ regions respectively. $\mu(t\bar{t})$ is defined twice: once for the $1e/\mu + 2\tau_{had}$ events ('2τ') and once for the $2e/\mu OS + 1\tau_{had}$ events ('OS'). The NFs are not combined as they predominantly describe events with fake contributions, which are different for the different channels. The NPs which are comparable are varied across all regions together. However, NPs which are used only by shape in one channel and with both shape and normalization in another are not combined. The NP comparisons can be found in appendix G.

Figure 7.37 depicts the expected sensitivity due to the expected data statistics (stat) and the fit including all NPs (tot) for each individual channel and the combined fit. The combined Asimov fit result is found to be:

$$\Delta\mu(tHq)(\text{Asimov}) = (\text{stat})^{+4.6}_{-4.6}(\text{tot})^{+7.4}_{-6.7} \qquad (7.13)$$

This meets the expectations formulated above. As the other NFs are just evaluated in one of the channels, the expected uncertainty in the combined fit almost equals the result from one channel (see appendix G).
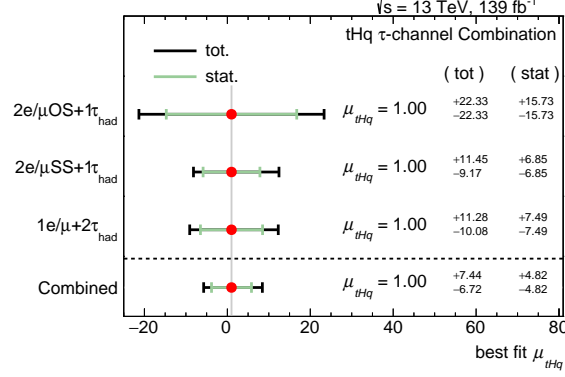
Figure 7.37: Comparison of the Asimov fit results in all three channels as well as the combined fit result.

Similar to section 7.5.1.2 the expected limit can be calculated for the combined fit:

$$\text{limit(exp.)} = 15.47 \left(\pm 1\sigma^{22.35}_{16.94} \pm 2\sigma^{31.76}_{8.30}\right) \; \mu(tHq)(\text{SM}) \tag{7.14}$$

Additionally, the expectations in the inverted coupling scenario are set:

$$\Delta\mu(tHq)(\text{Asimov,ic}) = (\text{stat})^{+1.0}_{-1.0}(\text{tot})^{+1.81}_{-1.67} \tag{7.15}$$

Both values are slightly improved with respect to the $1e/\mu + 2\tau_{\text{had}}$ channel expectation, but the addition of more channels juust slightly improves the expected uncertainty.

7.6 Partial unblinding

After the expected results based on purely simulated samples have been presented, the measured data is taken into account. In a first step, data is only used in regions or bins in which little signal is expected. With these studies the impact of these regions on the background estimates can be determined. Additionally, the impact on NPs can be seen. In this section only the most interesting studies are shown, more cross-checks can be found in appendix H.

7.6.1 CR assumptions

To validate the assumptions made in section 7.1.1 on the possible behavior of fake-driven NFs a series of fits is performed with data in the CRs. Only this way it can be tested what happens with the NFs and if they change their behaviour. However, the signal is not important for the crosschecks. Therefore, in the SR Asimov data is used and $\mu(tHq)$ is not shown. To be comparable with figure 7.8 the 'shape only' NPs associated with $t\bar{t}$ are used in the fit. However, for simplicity the other NPs are not included. Figure 7.38 shows the impact of a combined NF across all three sample in the described CR. This can be compared to a fit with the same CR but two NFs in figure 7.39. In comparison figure 7.40 shows the impact of an additional CR for $\mu(V+\text{jets})$. All figures include the correlation between the NFs. It is visible that the

additional NFs are highly anti-correlated amongst each other. However, the correlation with $\mu(tHq)$ is decreased. The additional CR further decorrelates the NFs. Furthermore, the NFs are pulled differently. Especially $\mu(V+\text{jets})$ is pulled down more than $\mu(t\bar{t})$ and the uncertainty on $\mu(V+\text{jets})$ is higher.

To crosscheck the effect of separating $\mu(Z+\text{jets})$ and $\mu(W+\text{jets})$ this fit is done with both CRs and shown in figure 7.41. The uncertainties on the NFs are even more enlarged and $\mu(Z+\text{jets})$ is even pulled to zero. This is clearly an unphysical solution and hints at too little statistics in these samples to result in a correct estimation. This supports the decision to treat both samples together.
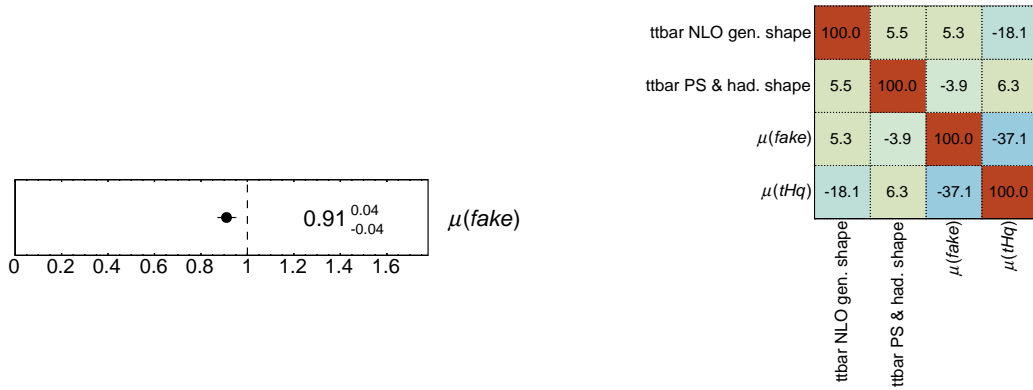
Figure 7.38: Fit results from an Asimov fit in the SR and data in the CR performed with one additional NF covering $t\bar{t}$, Z+jets and W+jets.
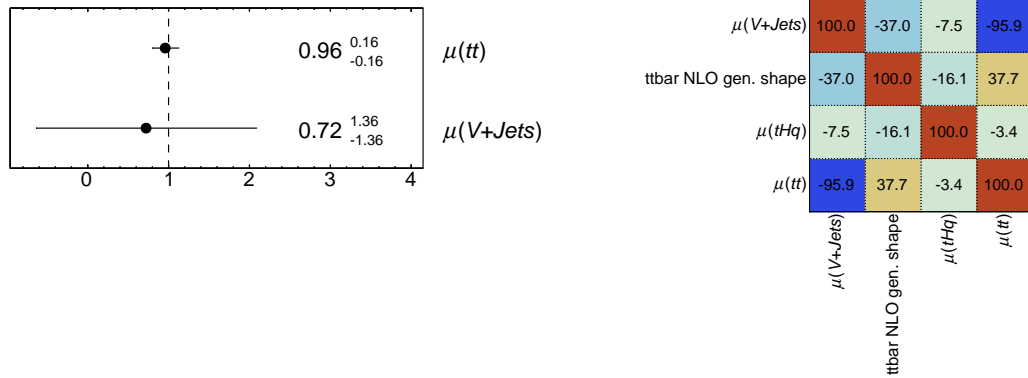
Figure 7.39: Fit results from an Asimov fit in the SR and data in the CR performed with one additional NF covering $t\bar{t}$ and another covering Z+jets and W+jets.

7.6.2 NP studies

Another study that can be done with partial unblinded data is the impact on the NPs. This is the first time the value of a NP after the fit can deviate from the nominal value. Any unexpected deviation from the
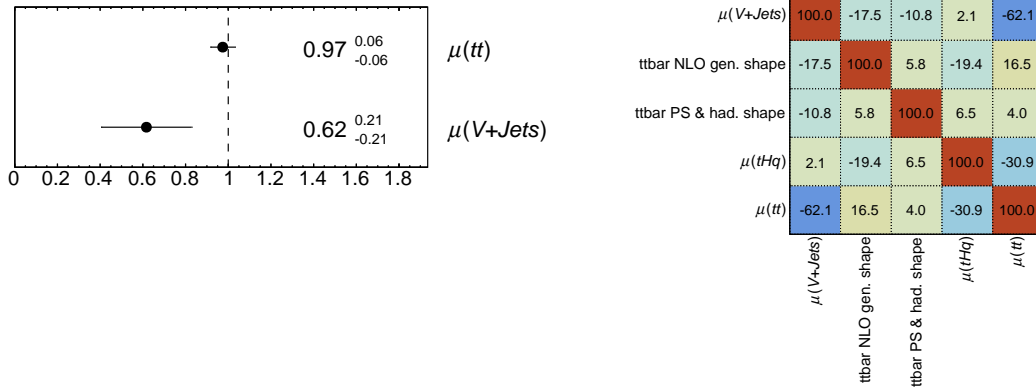
Figure 7.40: Fit results from an Asimov fit in the SR and data in both CRs performed with one additional NF covering $t\bar{t}$ and another covering Z+jets and W+jets.
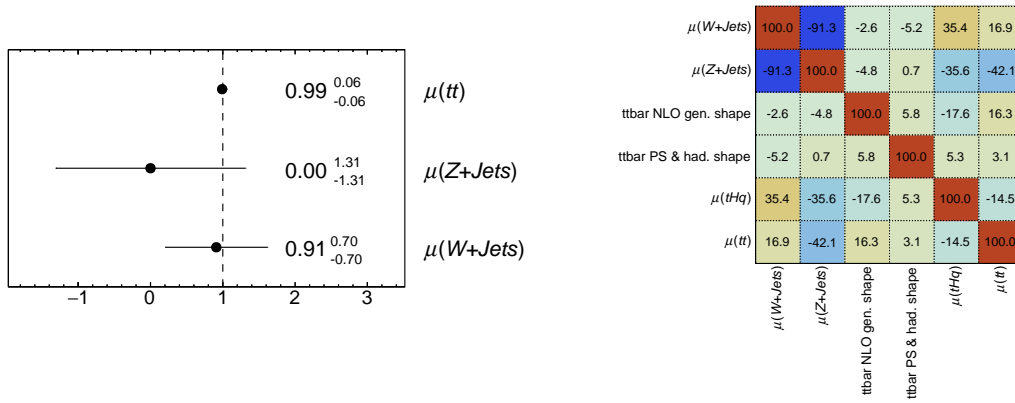
Figure 7.41: Fit results from an Asimov fit in the SR and data in the CRs performed with three separate NFs for $t\bar{t}$, Z+jets and W+jets.

central value can hint at a methodical error and needs to be investigated. Any deviation which exceeds 1σ is often deemed as problematic.

The highest ranked NPs described in section 7.5.1, section 7.5.2 and section 7.5.3 are shown in this section. For all channels the $t\bar{t}$ theory related uncertainties have a large impact on the final fit result. Figure 7.42 shows the best fit result of these NPs in three different fits: control regions only, data in CRs only and only insensitive bins. The first fit only takes into account the CRs, but uses data in them. The second fit additionally takes into account the SR, but uses Asimov data in it. Both of these fits are performed under the background-only hypothesis as the studied regions are expected to have minimal signal contribution. Therefore, $\mu(tHq)$ is fixed to zero. The third fit uses all bins in which the signal is not significant – as the analysis in principle is sensitive to the inverted coupling samples this includes all bins which have a contribution of less than 5% of tHq +tWH under the inverted coupling hypothesis. The 'CR only' fit only shows a small pull. However, the fit has a significantly lower sensitivity. When Asimov data in the SR is added, the pull becomes more visible but is still small. A stronger pull is observed when

fitting on all bins insensitive to the signal. None of the pulls is worryingly strong.

(a) CR only

(b) CR data + SR Asimov

(c) insensitive bins
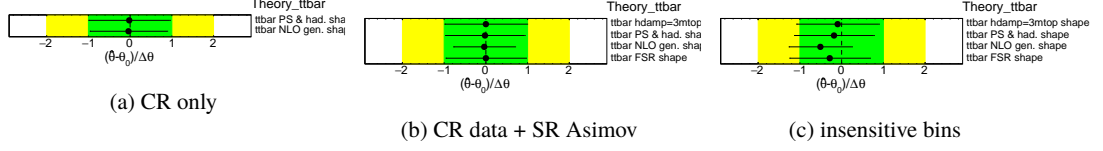
Figure 7.42: NP studies on three steps of integrating data into the $1e/\mu + 2\tau_{\text{had}}$ fit.

The same study is presented for $2e/\mu\text{OS} + 1\tau_{\text{had}}$ in figure 7.43. The corresponding $2e/\mu\text{SS} + 1\tau_{\text{had}}$ studies only include the CR studies and are shown in figure 7.44. In contrast to the other channels all bins in all regions pass the sensitivity cut. Therefore, the study on insensitive bins is not performed. The pulls in both channels are insignificant.

(a) CR only

(b) CR data + SR Asimov

(c) insensitive bins
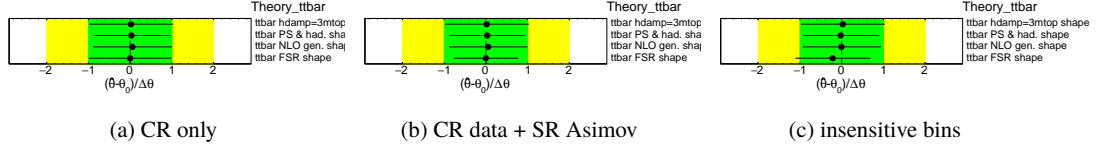
Figure 7.43: NP studies on three steps of integrating data into the $2e/\mu\text{OS} + 1\tau_{\text{had}}$ fit.

(a) CR only

(b) CR data + SR Asimov
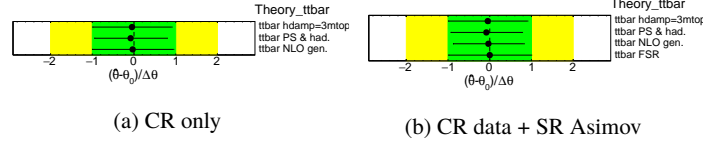
Figure 7.44: NP studies on three steps of integrating data into the $2e/\mu\text{SS} + 1\tau_{\text{had}}$ fit.

7.7 Results with data

This section includes the final fit results from each individual channel and the combination of all channels. In these the all available data is used. The NPs after the fit are further studied in appendix I.

7.7.1 Results in the 1 e/μ + 2 τ_{had} channel

Figure 7.45 shows the binned distributions in the $1e/\mu + 2\tau_{\text{had}}$ channel before and after the fit. After the fit the simulation matches the data better, especially in the CRs. In the SRs the data is fluctuating due to the low statistics. Especially the most sensitive bin has a discrepancy with the simulated datasets if *tHq* is not included. According to these observations the different samples are scaled by the fit:

$$\mu(t\bar{t}) \quad = \quad 0.99^{+\ 0.11}_{-\ 0.11} \tag{7.16}$$

$$\mu(\text{V+jets}) = \quad 0.57^{+\ 0.48}_{-\ 0.48} \tag{7.17}$$

$$\mu(tHq) \quad = 13.13^{+11.51}_{-10.07} \tag{7.18}$$

It becomes apparent that $\mu(t\bar{t})$ is almost at the nominal value, while $\mu(\text{V+jets})$ is scaled down. However, $\Delta\mu(\text{V+jets})$ is so high that it is fully compatible with 1. The signal $\mu(tHq)$ is scaled up, but still compatible with 1 at 1.3σ. This is not a significant derivation. Looking at the data statistics effect only (no NPs)

$$\mu(t\bar{t})_{\text{stat}} \quad = \quad 0.99^{+0.04}_{-0.04} \tag{7.19}$$

$$\mu(\text{V+jets})_{\text{stat}} = \quad 0.57^{+0.09}_{-0.09} \tag{7.20}$$

$$\mu(tHq)_{\text{stat}} \quad = 13.13^{+8.09}_{-7.48} \tag{7.21}$$

it is crucial to improve the data statistics to improve the final result. If this result confirms the tendency towards a higher than SM contribution of tHq events is impossible to state at this point.
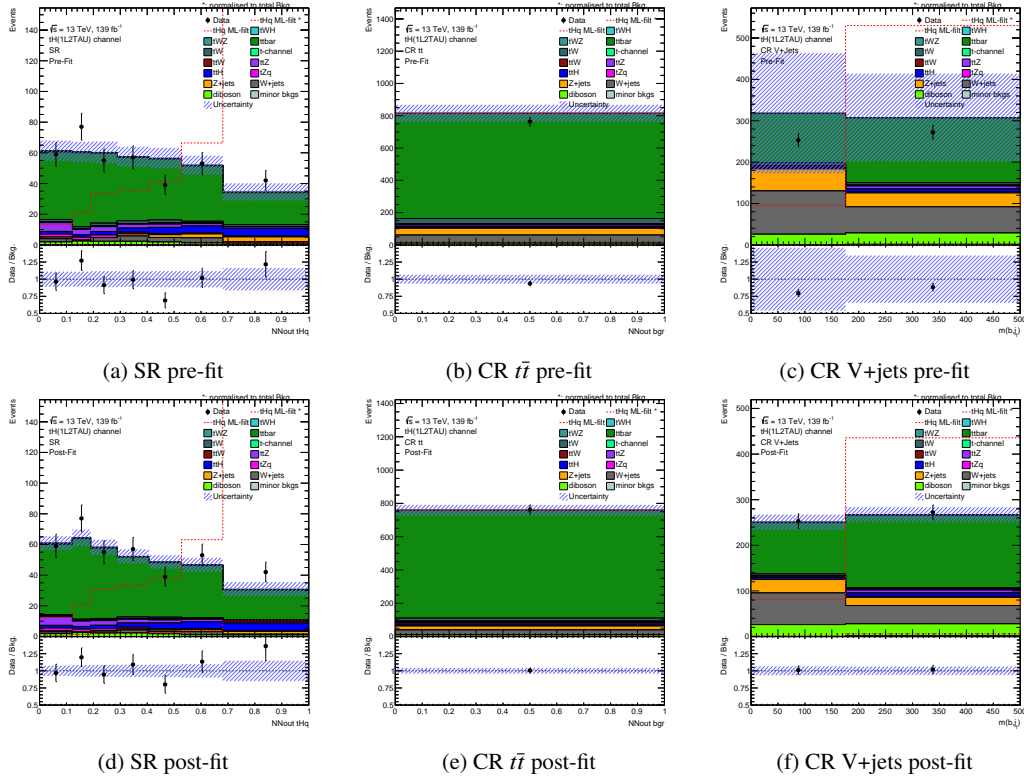
Figure 7.45: Simulation to data agreement in the $1e/\mu + 2\tau_{\text{had}}$ channel before and after the fit to data.

(a) SR pre-fit (b) CR $t\bar{t}$ pre-fit (c) CR V+jets pre-fit

(d) SR post-fit (e) CR $t\bar{t}$ post-fit (f) CR V+jets post-fit

To look further into possible improvements of the $1e/\mu + 2\tau_{\text{had}}$ fit the impact of NP groups is evaluated. Table 7.9 shows the impact of different NPs on the fit result. The findings are similar to the Asimov fit result table 7.6. Big improvements are expected from an increment in statistics of the simulated samples and a better $t\bar{t}$ description. Not to be ignored is the impact NPs have on $\mu(\text{V+jets})$. Therefore, an improvement of the NPs (namely the limited statistics of the simulated samples) might also indirectly benefit the $\mu(tHq)$ estimate.

	Full fit	stat.	γ-factors	all other NPs
impact	10.79	7.78	5.55	7.90

	experimental	flavor-tagging	JES/JER	theory	theory $t\bar{t}$	NFs
impact	2.02	0.40	3.18	1.92	2.88	2.13

Table 7.9: Impact of different groups of NPs on the $1e/\mu + 2\tau_{\text{had}}$ channel fit.

7.7.2 Results in the 2 e/μ OS + 1 τ_{had} channel

The result of the $2e/\mu$OS $+ 1\tau_{\text{had}}$ channel is expected to be even less sensitive than the result from the $1e/\mu + 2\tau_{\text{had}}$ channel. Figure 7.46 shows the agreement of simulation to data before and after the fit. In this channel the data statistics is not a big issue in most bins. The agreement before the fit was already good. After the fit the small deviations are gone. In contrast to the $1e/\mu + 2\tau_{\text{had}}$ channel, the data shows a small downward fluctuation in the most sensitive bins. However, this is fully compatible within 1σ. The NFs after the fit with and without NPs are scaled as follows:

$$\mu(Z + \text{jets}) = 1.04 \pm (\text{stat})^{+\ 0.02}_{-\ 0.02}(\text{syst})^{+\ 0.10}_{-\ 0.10} \tag{7.22}$$

$$\mu(t\bar{t}) = 1.01 \pm (\text{stat})^{+\ 0.02}_{-\ 0.02}(\text{syst})^{+\ 0.04}_{-\ 0.04} \tag{7.23}$$

$$\mu(tHq) = -21.12 \pm (\text{stat})^{+15.47}_{-14.14}(\text{syst})^{+21.84}_{-22.30} \tag{7.24}$$

Note that the uncertainties with NPs include the effects from data. While $\mu(t\bar{t})$ and $\mu(V+\text{jets})$ are hardly scaled and fully compatible with 1, $\mu(tHq)$ is scaled negatively by a factor of 21. However, the uncertainty covers that scaling. Again, an improvement of data statistics would facilitate a better understanding of the *tHq* scale.

7.7.3 Results in the 2 e/μ SS + 1 τ_{had} channel

This is the most sensitive channel from the S/B expectations. However, this comes with low data statistics. This can also be seen in the plots including data before and after the fit in figure 7.47. The data in the SR fluctuates with respect to the expectation due to the low statistics. After the fit both match a bit better, but the fit has limited possibilities to scale the backgrounds independently of each other. This can especially be seen in the CR. One of the NPs which is significantly varied is 'Tau Fake Norm' as shown in figure 7.48. It is pulled down to the point where it is barely compatible with 1 within the post-fit uncertainty estimation. The final scale of $\mu(tHq)$ is found to be

$$\mu(tHq) = -2.5 \pm (\text{stat})^{+8.0}_{-7.0}(\text{syst})^{+12.0}_{-10.0} \tag{7.25}$$

This includes a slight downscale with respect to the standard model value, but a large enough $\Delta\mu(tHq)$ to cover the SM value at 1σ and the inverted coupling value at 2σ. Again, improving the data statistics would yield the biggest improvement. Additional improvements are expected mostly from the fake estimation as was shown in figure 7.36.

(a) SR pre-fit (b) CR $t\bar{t}$ pre-fit (c) CR Z+jets pre-fit

(d) SR post-fit (e) CR $t\bar{t}$ post-fit (f) CR Z+jets post-fit

Figure 7.46: Simulation to data agreement in the $2e/\mu$OS + $1\tau_{\mathrm{had}}$ channel before and after the fit to data.

(a) SR pre-fit

(b) CR post-fit

(c) CR pre-fit

(d) SR post-fit
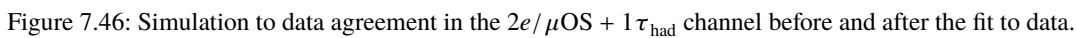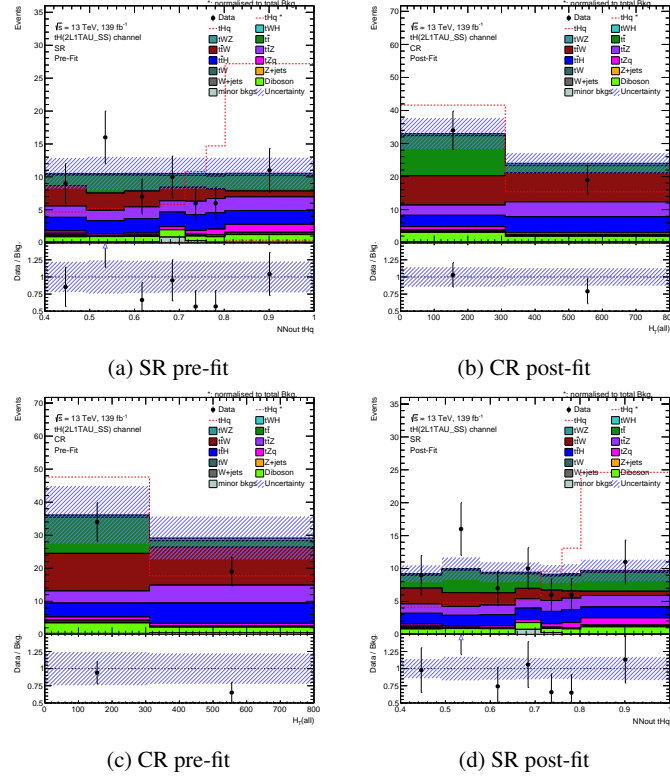
Figure 7.47: Simulation to data agreement in the $2e/\mu$SS + $1\tau_{\text{had}}$ channel before and after the fit to data.
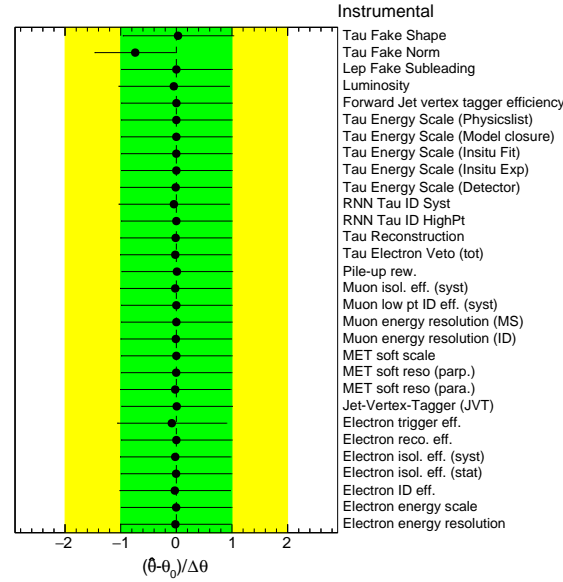
Figure 7.48: Pull of the experimental NPs in the $2e/\mu$SS + $1\tau_{\text{had}}$ fit with data.

7.7.4 Results in the combination of all channels

The final result in this thesis is given by a combined fit of all channels with data. As described in section 7.5.4 the comparable NPs are estimated in all regions together, while the background NFs are kept separate. Figure 7.49 shows a comparison of the individual channel results and the fit result from the combined fit.

$$\mu(tHq)_{\text{comb,stat}} = 9.3 \pm {}^{+5.5}_{-5.2} \tag{7.26}$$

$$\mu(tHq)_{\text{comb,full}} = 5.4 \pm {}^{+7.8}_{-7.1} \tag{7.27}$$

The $1e/\mu + 2\tau_{\text{had}}$ result has the biggest impact on the combination fit, especially in the fit without NPs. The final result with NPs prefers a value higher than the SM expectation. All the while it is fully compatible with the SM expectation at 0.76σ. Expected and observed significances per channel are shown in appendix G and appendix I.
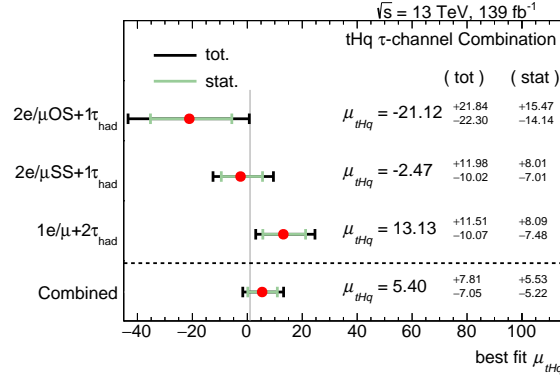
Figure 7.49: $\mu(tHq)$ results from the individual channel fits and the combination.

Clearly, a further improvement of the analysis would lead to fascinating results. The impact the different groups of NPs have on the final result is shown in table 7.10. After the separation of the γ-factors the other groups of NPs have a similar impact on the final result as the fit using only statistical effects.

	Full fit	stat.	γ-factors	all other NPs
impact	7.43	5.38	2.88	5.22

	experimental	flavor-tagging	JES/JER	theory	theory $t\bar{t}$	NFs
impact	1.86	0.32	1.80	1.93	2.05	1.08

Table 7.10: Impact of different groups of NPs on the combination fit

When ranking the NPs in figure 7.50 it becomes apparent that the NPs which were most important for the individual channels also dominate the final result. The three most important NPs are: the number of simulated events in the most significant bin of the $1e/\mu + 2\tau_{\text{had}}$ channel, the $2e/\mu\text{SS} + 1\tau_{\text{had}}$ τ_{had} fake estimation and the $t\bar{t}$ FSR variation (from the $2e/\mu\text{OS} + 1\tau_{\text{had}}$ and $1e/\mu + 2\tau_{\text{had}}$ channel). The latter two are also pulled more than 1σ in the final fit. Therefore, it is quite possible that a future improvement of the analysis could already improve from focussing on those NPs.
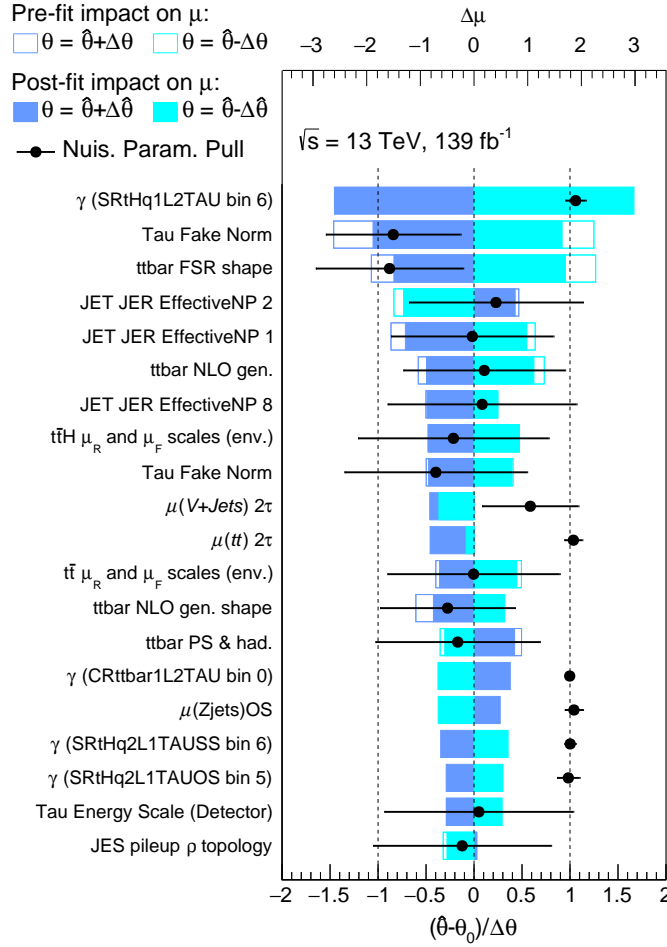
Figure 7.50: Ranking of the 20 most impactful NPs on the data fit in the combination of all channels.

To study the final fit result, the likelihood curve of the final result is studied. The most interesting study estimates the $\mu(tHq)$ likelihood together with the other likelihood of the other NFs. Figure 7.51 shows the 2-dimensional likelihood map of $\mu(tHq)$ and $\mu(t\bar{t})(2\tau)$. The bin with the lowest $-\ln(L)$ is white. The other bins are color-coded as to the difference in the $-\ln(L)$ with respect to that bin. The dotted lines mark the 1,2 and 3σ and boundaries. The plot shows that both NFs are not much correlated with each other and $\mu(t\bar{t})$ is a lot better constrained. The SM value ($\mu(tHq)$ and $\mu(t\bar{t})$) is well included in 1σ. Similar studies with the other NFs are shown in figure 7.52. While the two NFs from the $2e/\mu$OS $+$ $1\tau_{\text{had}}$ fit behave similar to the already discussed NF, μ(V+jets) estimated from the $1e/\mu$ $+$ $2\tau_{\text{had}}$ fit is well less constrained. The plotted range shows a plateau spanning almost the complete plotted range of μ(V+jets). The SM value at (1,1) is included in this plateau. Only little correlation with $\mu(tHq)$ can be observed towards higher values (compare also to the correlation matrix in figure I.6). This hints at the fact that the low sensitivity does not spuriously influence the final result for $\mu(tHq)$. Nevertheless, an improvement on the sensitivity for μ(V+jets) might be beneficial for future studies.

Since the fit result has not concluded in a discovery, the limit the final result can set on $\mu(tHq)$ is calculated. Figure 7.53 compares the limit calculations of all three channels and the combined limit. The
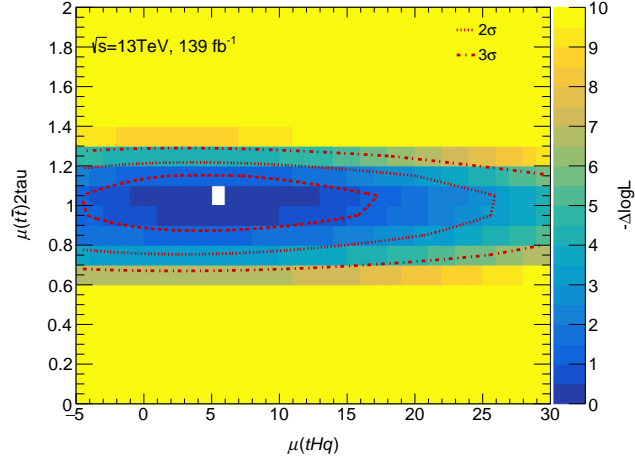
Figure 7.51: 2-dimensional likelihood scan of $\mu(tHq)$ versus $\mu(t\bar{t})$ (from $1e/\mu + 2\tau_{\text{had}}$) in the fit combinating all three channels.
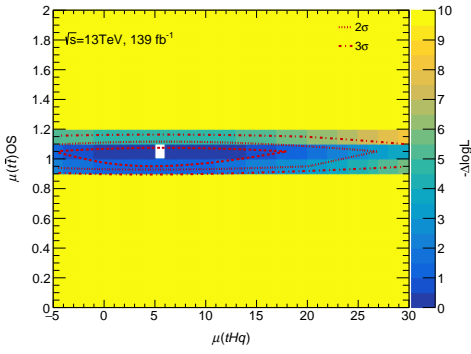
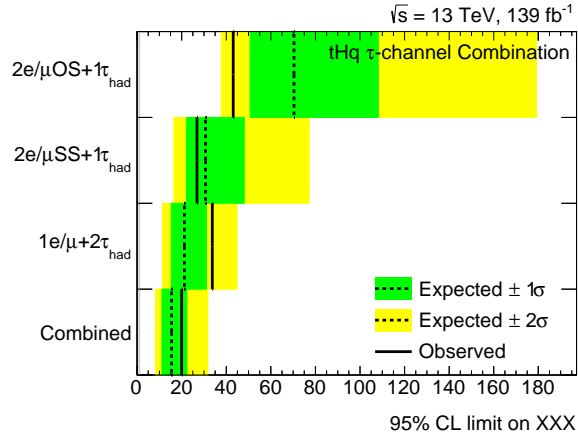(a) $\mu(tHq)$ vs. $\mu(t\bar{t})$ (from $2e/\mu\text{OS} + 1\tau_{\text{had}}$)
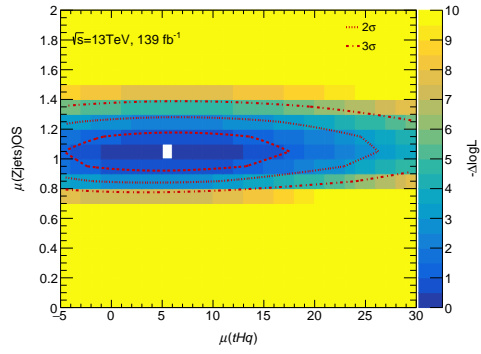
(b) $\mu(tHq)$ vs. $\mu(Z + \text{jets})$ (from $2e/\mu\text{OS} + 1\tau_{\text{had}}$)
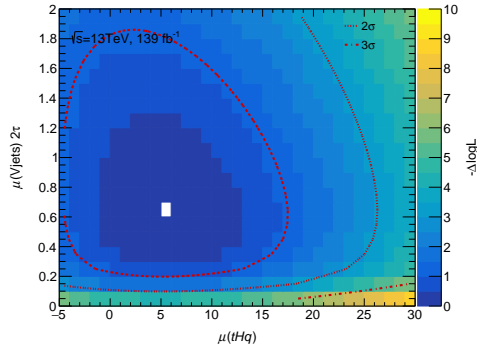
(c) $\mu(tHq)$ vs. $\mu(V+\text{jets})$ (from $1e/\mu + 2\tau_{\text{had}}$)

Figure 7.52: 2-dimensional likelihood scan of $\mu(tHq)$ versus the other NFs in the fit combinating all three channels.

Figure 7.53: Limits on $\mu(tHq)$ from the individual channel fits and the combination.

$2e/\mu$OS + $1\tau_{\text{had}}$ limit calculation is not completely comparable with the other results as the calculation only converges if signal is injected in the calculation. The combination results taking into account only the two most sensitive channels can be found in appendix K. The combined limit from all three channels together is 20 times the SM expectation. This is well inside the 1σ limit of the SM expectation presented in equation (7.14). Additionally, this result is comparable with the previous searches for *tHq* presented in section 3.5. Expected and observed limits per channel are shown in appendix G and appendix I.

To round up, the results on the basis of the SM expectation are compared to the results under the inverted coupling hypothesis. Figure 7.54 show the $\mu(tHq)$ estimates under this hypothesis. More information from fits under the inverted coupling hypothesis are shown in appendix J. All three channels as well as the combination disfavor this scenario. In the combination

$$\mu(tHq)(\text{ic}) = -4.45 \pm (\text{stat})^{+0.90}_{-0.85}(\text{syst})^{+1.49}_{-1.46} \tag{7.28}$$

the uncertainties match the expected uncertainties presented in equation (7.15). The central value deviates -2.8σ from the expected value. This disfavors the inverted coupling hypothesis.

7.7.5 Discussion of the results

The results with data presented in this section match the expectation of the SM hypothesis presented in section 7.5. One channel ($1e/\mu + 2\tau_{\text{had}}$) shows an upward fluctuation, but this is neither significant, nor is it reflected in the other channels.

To learn more about the *tHq* process the uncertainties of the combined result would need to be smaller. About half of the impact on the uncertainty stems from the low number of observed data events. Incorporating events from Run 3 of the LHC will be an opportunity to improve that part of the analysis for future studies. The rest of the impact stems from different analysis, simulation and reconstruction related sources. The impact of the simulated number of events for certain samples is unfortunate. An improvement is expected from incorporating newer samples which have been simulated with more statistics. Sadly, they were not available in time for this thesis. An alternative approach to mitigate the impact of this effect would be to combine more sensitive bins with less sensitive bins. This would increase the available number
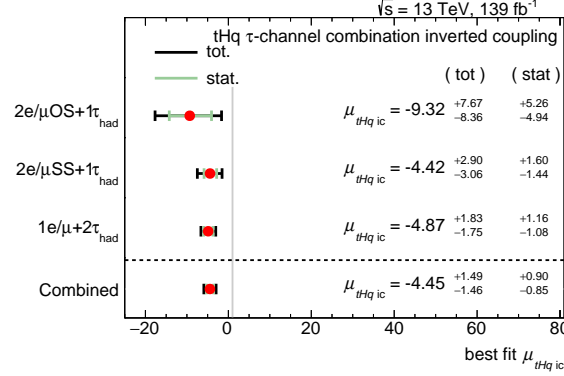
Figure 7.54: Results of $\mu(tHq)$ under the inverted coupling hypothesis for the individual channel fits and the combination.

of simulated events, but decrease the sensitivity of the final result. Therefore, this path was not pursued. Another choice which has a negative impact on the final result is the fake estimation in the $2e/\mu\text{SS} + 1\tau_{\text{had}}$ channel. The number of events which are associated with background processes in this channel is low. This is also true for the CR which is used to determine the fakes. Subsequently, the uncertainty on the fake estimation increases and in this case the statistical fluctuation of the result lead to an estimate which did not match the SR. Alternative methods to estimate the fakes in this region (for example estimating the fakes in a different region with more events) might be suited better. A third effect is the signal-like behavior the $t\bar{t}$ FSR variation has in the NN output in the $2e/\mu\text{OS} + 1\tau_{\text{had}}$ channel. This is not a wrong behavior of the NN, but unfortunate. Training the NN on different input variables might have resulted in a different behavior. All three of the described effects have already been seen and discussed before data was introduced in the regions sensitive to signal.

The combination of the three channels with τ_{had} in this thesis resulted in a limit on the tHq cross-section of 20 times the SM cross-section (15.5 expected). This is not as good as some of the previously published results (section 3.5). However, this thesis focuses on three challenging channels. Results including these channels have never been published by ATLAS before and especially the expected limit shows remarkable compatibility with the published results by ATLAS.

The other research question of this thesis was the alternative 'inverted coupling' hypothesis. The result under this hypothesis is disfavored. However, the used statistical method to evaluate the factor for the alternative cross-section is not ideal to evaluate this hypothesis, since an observed negative cross-section is not physical (as there is no expected negative interference with another process). Still, the observed deviation of -2.8σ from the expected result under this hypothesis is close to being an evidence (3σ) against this hypothesis. The fact that all channels independently prefer a cross-section lower than the expected one in this hypothesis further solidifies that this scenario is disfavored. To receive more information on the relative coupling of the H from the tHq process a scan of different coupling scenarios would be the next step.

Summary

This thesis presents the first study to find events of the rare *tHq* process with ATLAS data in channels with hadronically decaying τ. A baseline framework is developed to select as many signal events as possible in three channels: $1e/\mu + 2\tau_{\text{had}}$, $2e/\mu\text{OS} + 1\tau_{\text{had}}$ and $2e/\mu\text{SS} + 1\tau_{\text{had}}$. Within this framework a full reconstruction of *tHq* events is performed which leads to a *H* mass reconstruction. A binned profile likelihood fit with Asimov data and real data is done for the three channels separately and in the combination of all three channels. The expected and observed SM result of the combined fit is:

$$\mu(tHq)_{\text{expected}} = 1 \quad \pm^{+7.4}_{-6.7} \tag{8.1}$$

$$\mu(tHq)_{\text{observed}} = 5.4 \pm^{+7.8}_{-7.1} \tag{8.2}$$

This result is fully compatible with the SM. A limit on the *tHq* cross-section is set to 20 times the SM expectation (15.5 expected).

The BSM scenario of an inverted *H* to *t* coupling with respect to the *H* to *W* coupling is disfavored at the 2.8σ level. A significant improvement of the result is expected when more data is available. However, additional improvements on the result are expected if the size of the simulated datasets is increased. Other fields of improvements are a better fake estimation and estimation of uncertainties related to $t\bar{t}$ modelling.

The presented result is compatible with the previous results. No study on *tHq* has been published quoting the exact same final states as used in this thesis that the result could be directly compared to. The enlarged uncertainty with respect to the published results (additionally) including other channels is expected, since three challenging final states are targeted.

MC samples

Breakdown of MC samples used. The samples marked as nominal are used as the nominal sample for the associated to a process, the others are used as a variation. The following columns describe which generator has been used for the associated generation steps. 'aMC' denotes the NLO version of Madgraph. The last column describes any notable specifications to distinguish nominal and systematic samples.

process	nominal	DSID	ME	shower	detector	
tHq	x	346799	aMC	Pythia	AFII	
		508775	aMC	Herwig	AFII	
		508774	aMC	Pythia	AFII	$y_t = -1$
tWH	x	508776	aMC	Pythia	AFII	DR1
		512275	aMC	Herwig	AFII	DR1
		508777	aMC	Pythia	AFII	DR1, $y_t = -1$
tWZ	x	412118	aMC	Pythia	Geant4	DR1
		410408	aMC	Pythia	Geant4	minDR1
$t\bar{t}$	x	410470, 410471	Powheg	Pythia	Geant4	[189, 190]
		410470	Powheg	Pythia	AFII	
		411233,411234	Powheg	Herwig	AFII	
		412116,412117	aMC	Herwig	AFII	
		410480,410482	Powheg	Pythia	AFII	hdamp
tW	x	410646,410647	Powheg	Pythia	Geant4	DR1, [191]
		410654,410655	Powheg	Pythia	Geant4	DS
		410646,410647	Powheg	Pythia	AFII	DR1
		412002	aMC	Pythia	AFII	
		600727,600728	Powheg	Herwig	AFII	
t-channel	x	410658,410659	Powheg	Pythia	Geant4	
		600017,600018	Powheg	Herwig	AFII	[35, 192]
		410658,410659	Powheg	Pythia	AFII	
		500289	aMC	Herwig	AFII	

Table A.1: Used samples and generators used to create them.

process	nominal	DSID	ME	shower	detector	
$t\bar{t}W$	x	410155,412123	aMC	Pythia	Geant4	
		700168,700205	Sherpa	Sherpa	Geant4	
		410376,410377	aMC	Pythia	Geant4	ISR variation
$t\bar{t}Z$	x	[1]	aMC	Pythia	Geant4	
		410142,410143	Sherpa	Sherpa	Geant4	
		[2]	aMC	Pythia	Geant4	ISR
$t\bar{t}H$	x	346343,346344,346345	Powheg	Pythia	Geant4	[193]
		346343,346344,346345	Powheg	Pythia	AFII	
		346443,346444,346445	aMc	Pythia	AFII	
		346346,346347,346348	Powheg	Herwig	AFII	
tZq	x	512059	aMC	Pythia	Geant4	
		512059	aMC	Pythia	AFII	
		512130	aMC	Herwig	AFII	
		512169,512168	aMC	Pythia	AFII	ISR
Z+jets	x	[3]	Sherpa	Sherpa	Geant4	
W+jets	x	[4]	Sherpa	Sherpa	Geant4	
Diboson	x	[5]	Sherpa	Sherpa	Geant4	
Triboson	x	[6]	Sherpa	Sherpa	Geant4	
ggf H	x	342282,342283	Powheg	Pythia	Geant4	[194, 195]
VBF H	x	342282,342283	Powheg	Pythia	Geant4	[196, 197]
WH	x	342284	Pythia	Pythia	Geant4	
ZH	x	342285	Pythia	Pythia	Geant4	
s-channel	x	410644,410645	Powheg	Pythia	Geant4	[192]
3top	x	304014	Madgraph	Pythia	Geant4	LO
4top	x	412043	aMC	Pythia	AFII	

---

[1] 410156,410157,410218,410219,410220,410276,410277,410278

[2] 410370,410372,410374,410371,410373,410375

[3] 364114,364115,364116,364117,364118,364119,364120,364121,364122,364123,364124,364125,364126,364127,364100,364101,364102, 364103,364104,364105,364106,364107,364108,364109,364110,364111,364112,364113,364128,364129,364130,364131,364132,364133, 364134,364135,364136,364137,364138,364139,364140,364141,364204,364205,364206,364207,364208,364209,364198,364199,364200, 364201,364202,364203,364210,364211,364212,364213,364214,364215

[4] 364170,364171,364172,364173,364174,364175,364176,364177,364178,364179,364180,364181,364182,364183,364156,364157,364158, 364159,364160,364161,364162,364163,364164,364165,364166,364167,364168,364169,364184,364185,364186,364187,364188,364189, 364190,364191,364192,364193,364194,364195,364196,364197

[5] 364250,364253,364254,364255,364283,364284,364286,364287,364288,364289,364290,363355,363356,363357,363358,363359,363360, 363489

[6] 364242,364243,364244,364245,364246,364247,364248,364249

Systematics

Name	variation	type
ATLAS_JVT	JVT	weight
ATLAS_fJVT	fJVT	weight
FTAG_B[00-44]	b-tagging uncertainties on b-tagged jets	weight
FTAG_C[00-19]	tagging uncertainties on c-tagged jets	weight
FTAG_LIGHT[00-19]	b-tagging uncertainties on not-tagged jets	weight
JER_DataVsMC_MC16	JER data-simulation comparison	Pseudodata
JER_EffectiveNP_[1-12]	JER	Pseudodata
BJES_Response	b-tagged jet energy scale	tree
SingleParticle_HighPt	JES high p_T	tree
PunchThrough_[MC16/AFII]	JES punch-through from EMCal to HCal	tree
Pileup_OffsetMu	μ impact on JES	tree
Pileup_OffsetNPV	NPV impact on JES	tree
Pileup_PtTerm	p_T impact on JES	tree
Pileup_RhoTopology	ρ impact on JES	tree
Flavor_Response	JES from different hadron types	tree
Flavor_Composition	different composition of hadron types on JES	tree
EtaIntercalibration_Modelling	JES modelling variation	tree
EtaIntercalibration_TotalStat	JES stat component	tree
EtaIntercalibration_NonClosure_2018data	JES 2018 non-closure	tree
EtaIntercalibration_NonClosure_highE	JES high energy non-closure	tree
EtaIntercalibration_NonClosure_negEta	JES non-closure from $-\eta$	tree
EtaIntercalibration_NonClosure_posEta	JES non-closure from $+\eta$	tree
RelativeNonClosure_AFII	JES non-closure from AFII	tree
EffectiveNP_Detector[1-2]	detector related variation of JES	tree
EffectiveNP_Mixed[1-3]	mixed related variation of JES	tree
EffectiveNP_Statistical[1-6]	statistics related variation of JES	tree

Table B.1: List of Jet related systematics.

Name	variation	type
ID	e identification	weight
Isol_Syst	e isolation systematic influence	weight
Isol_Stat	e isolation statistics influence	weight
Reco	e reconstruction	weight
Trigger	e trigger	weight
EG_RESOLUTION	e resolution	tree
EG_SCALE_AFII	e scale for AFII samples	tree
EG_SCALE	e scale	tree

Table B.2: List of electron related systematics.

Name	variation	type
ID_SYST	μ identification systematics influence	weight
ID_SYST_LOWPT	μ identification at low p_T systematics influence	weight
Isol_SYST	μ isolation systematics influence	weight
Trigger_SYST	μ trigger systematics influence	weight
TTVA_SYST	μ TTVA systematics influence	weight
ID_STAT	μ identification statistics influence	weight
ID_STAT_LOWPT	μ identification at low p_T statistics influence	weight
Isol_STAT	μ isolation statistics influence	weight
Trigger_STAT	μ trigger statistics influence	weight
TTVA_STAT	μ TTVA statistics influence	weight
ID	μ energy resolution identification influence	tree
MS	μ energy resolution μ scale influence	tree
SAGITTA_RESBIAS	μ sagitta residual bias	tree
SAGITTA_RHO	μ sagitta ρ topology	tree
SCALE	μ energy scale	tree

Table B.3: List of muon related systematics.

Name	variation	type
TRUEHADTAU_EFF_ELEOLR_TOTAL	electron overlap removal	weight
TRUEELEC-TRON_EFF_ELEBDT_STAT	electron rejection BDT statistics influence	weight
TRUEELEC-TRON_EFF_ELEBDT_SYST	electron rejection BDT systematics influence	weight
TRUEHADTAU_EFF_RECO_TOTAL	τ_{had} reconstruction	weight
TRUEHADTAU_EFF_RN-NID_1PRONGSTATSYSTPT2025	RNN for 1-prong τ_{had} with p_T =20 to 25 Gev	weight
TRUEHADTAU_EFF_RN-NID_1PRONGSTATSYSTPT2530	RNN for 1-prong τ_{had} with p_T =25 to 30 Gev	weight
TRUEHADTAU_EFF_RN-NID_1PRONGSTATSYSTPT3040	RNN for 1-prong τ_{had} with p_T =30 to 40 Gev	weight
TRUEHADTAU_EFF_RN-NID_1PRONGSTATSYSTPTGE40	RNN for 1-prong τ_{had} with p_T >40 Gev	weight
TRUEHADTAU_EFF_RN-NID_3PRONGSTATSYSTPT2025	RNN for 3-prong τ_{had} with p_T =20 to 25 Gev	weight
TRUEHADTAU_EFF_RN-NID_3PRONGSTATSYSTPT2530	RNN for 3-prong τ_{had} with p_T =25 to 30 Gev	weight
TRUEHADTAU_EFF_RN-NID_3PRONGSTATSYSTPT3040	RNN for 3-prong τ_{had} with p_T =30 to 40 Gev	weight
TRUEHADTAU_EFF_RN-NID_3PRONGSTATSYSTPTGE40	RNN for 3-prong τ_{had} with p_T >40 Gev	weight
TRUEHADTAU_EFF_RNNID_HIGHPT	RNN for τ_{had} with high p_T	weight
TRUEHADTAU_EFF_RNNID_SYST	RNN systematics influence	weight
TRUEHADTAU_SME_TES_DE-TECTOR	TES detector effects	tree
TRUEHADTAU_SME_TES_INSITU-EXP	insitu TES experimental effects	tree
TRUEHADTAU_SME_TES_INSITUFIT	insitu TES fit impact	tree
TRUEHADTAU_SME_TES_MODEL_CLOSURE	TES closure	tree
TRUEHADTAU_SME_TES_PHYSIC-SLIST	TES physicslist influence	tree

Table B.4: List of τ_{had} related systematics.

Name	variation	type
FAKE_DITAU_SIMJET_1L2TAU	subleading τ_{had} fake estimation in $1e/\mu + 2\tau_{\text{had}}$	weight
FAKE_DITAU_JETSIM_1L2TAU	leading τ_{had} fake estimation in $1e/\mu + 2\tau_{\text{had}}$	weight
FAKE_LEP_SYS_1L2TAU	light lepton fake estimation in $1e/\mu + 2\tau_{\text{had}}$	weight
FAKE_DILEP_SIMJET_2L1TAU_OS	subleading light lepton fake estimation in $2e/\mu\text{OS} + 1\tau_{\text{had}}$	weight
FAKE_DILEP_JETSIM_2L1TAU_OS	leading light lepton fake estimation in $2e/\mu\text{OS} + 1\tau_{\text{had}}$	weight
FAKE_TAU_NORM_2L1TAU_OS	τ_{had} fake estimation normalization influence in $2e/\mu\text{OS} + 1\tau_{\text{had}}$	weight
FAKE_TAU_SHAPE_2L1TAU_OS	τ_{had} fake estimation shape influence in $2e/\mu\text{OS} + 1\tau_{\text{had}}$	weight
FAKE_DILEP_SIMJET_2L1TAU_SS	subleading light lepton fake estimation in $2e/\mu\text{SS} + 1\tau_{\text{had}}$	weight
FAKE_DILEP_JETSIM_2L1TAU_SS	subleading light lepton fake estimation in $2e/\mu\text{SS} + 1\tau_{\text{had}}$	weight
FAKE_TAU_NORM_2L1TAU_SS	τ_{had} fake estimation normalization influence in $2e/\mu\text{SS} + 1\tau_{\text{had}}$	weight
FAKE_TAU_SHAPE_2L1TAU_SS	τ_{had} fake estimation normalization influence in $2e/\mu\text{SS} + 1\tau_{\text{had}}$	weight

Table B.5: List of fake estimation related systematics.

Name	variation	type	onesided
MET_SoftTrk_Scale	$E_{\text{T}}^{\text{miss}}$ scale of soft componenet	tree	
MET_SoftTrk_ResoPara	$E_{\text{T}}^{\text{miss}}$ resolution of soft component parallel to the final $E_{\text{T}}^{\text{miss}}$	tree	yes
MET_SoftTrk_ResoPerp	$E_{\text{T}}^{\text{miss}}$ resolution of soft component perpendicular to the final $E_{\text{T}}^{\text{miss}}$	tree	yes
PRW_DATASF	pile-up-rewighting	weight	
lumi	luminosity	norm	

Table B.6: List of $E_{\text{T}}^{\text{miss}}$ and other ATLAS related systematics.

Name	variation	type	onesided	comparison
ttH_XS	$t\bar{t}H$ cross-section	norm		+7% -10%
ttW_XS	$t\bar{t}W$ cross-section	norm		+13% -12%
ttZ_XS	$t\bar{t}Z$ cross-section	norm		+10% -12%
singleTop_XS	tW,t-channel cross-section	norm		±5%
tWZ_XS	tWZ cross-section	norm		±50%
tZq_XS	tZq cross-section	norm		+7.8% -8%
Diboson_XS	diboson cross-section	norm		±24.5%
Zjets_XS	Z+jets cross-section	norm		±35%
Wjets_XS	W+jets cross-section	norm		±40%
other_XS	minor samples cross-section	norm		±50%
tHq_muR_muF_[1-7]	tHq renormalization and refactorization scale – envelope	weight	yes	
tWH_muR_muF_[1-7]	tWH renormalization and refactorization scale – envelope	weight	yes	
tWZ_muR_muF_[1-7]	tWZ renormalization and refactorization scale – envelope	weight	yes	
ttH_muR_muF_[1-7]	$t\bar{t}H$ renormalization and refactorization scale – envelope	weight	yes	
tZq_muR_muF_[1-7]	tZq renormalization and refactorization scale – envelope	weight	yes	
tW_muR_muF_[1-7]	tW renormalization and refactorization scale – envelope	weight	yes	
ttW_muR_muF_[1-7]	$t\bar{t}W$ renormalization and refactorization scale – envelope	weight	yes	
ttZ_muR_muF_[1-7]	$t\bar{t}Z$ renormalization and refactorization scale – envelope	weight	yes	
ttbar_muR_muF_[1-7]	$t\bar{t}$ renormalization and refactorization scale – envelope	weight	yes	
t-chan_muR_muF_[1-7]	t-channel renormalization and refactorization scale – envelope	weight	yes	
ttH_ISR	$t\bar{t}H$ ISR	weight		
ttH_FSR	$t\bar{t}H$ FSR	weight		
tW_ISR	tW ISR	weight		
tW_FSR	tW FSR	weight		
tchan_ISR	t-channel ISR	weight		
tchan_FSR	t-channel FSR	weight		
ttbar_ISR_shape	$t\bar{t}$ ISR	weight		
ttbar_FSR_shape	$t\bar{t}$ FSR	weight		

Table B.7: List of MC sample related systematics.

Name	variation	onesided	comparison
tHq_PS	*tWH* parton showering	yes	508775
tWH_PS	*tWH* parton showering	yes	512275
tWZ_DR	*tWZ-tt̄Z* interference	yes	410408
ttH_ME	*tt̄H* matrix element	yes	346443,346444,346445 (AFII vs. AFII)
ttH_PS	*tt̄H* parton showering	yes	346346,346347,46348 (AFII vs. AFII)
tZq_ISR	*tZq* ISR		up 512169, down 512168 (AFII vs. AFII)
tZq_PS	*tZq* parton showering	yes	512130 (AFII vs. AFII)
tW_DS	*tW-tt̄* interference	yes	410654,410655 (AFII vs. AFII)
tW_ME	*tW* matrix element	yes	412002 (AFII vs. AFII)
tW_PS		yes	600727,600728 (AFII vs. AFII)
tchan_Gen	t-channel generator	yes	500289 vs 600017,600018 (AFII vs. AFII)
tchan_PS	t-channel parton showering	yes	600017,600018 vs 410658,410659
ttW_ME_PS	*tt̄W* matrix element and parton showering	yes	700168,700205
ttW_ISR			up 410376, down 410377
ttZ_ME_PS		yes	410142,410143
ttZ_ISR	*tt̄Z* ISR		up 410370,410372,410374; down 410371,410373,410375
ttbar_PS(_shape)	*tt̄* parton showering	yes	411233,411234 (AFII vs. AFII)
ttbar_Gen(_shape)	*tt̄* generator	yes	412116,412117 vs 411233,411234
ttbar_hdamp3mtop(_shape)	*tt̄* hdamp		410480 (AFII vs. AFII)

Table B.8: List of MC sample related systematics which are estimated with alternative samples. If not stated otherwise, the samples are compared to the nominal sample (or the AFII sample with the same DSID). All DSIDs can be compared to the lists in appendix A.

Cutflows

1 e/μ + 2 τ_{had} cutflows
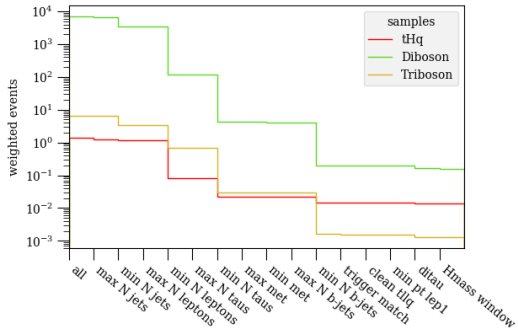
Complete cutflows of *tHq* in comparison with triboson and diboson samples.

Complete cutflows of *tHq* in comparison with s-channel single top production and 3/4-top samples.

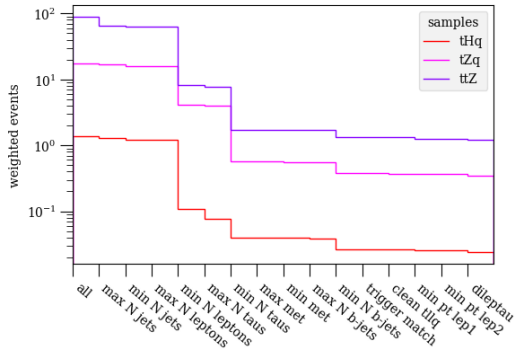Complete cutflows of *tHq* in comparison with $t\bar{t}$ and t-channel single top production samples.

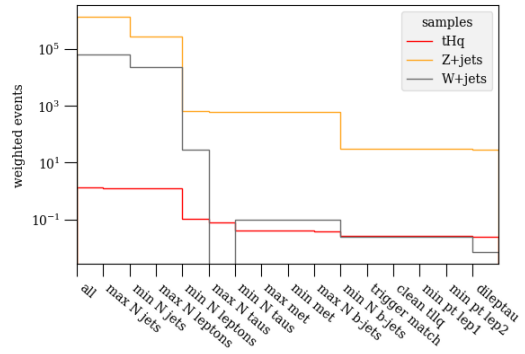Complete cutflows of *tHq* in comparison with $t\bar{t}W$ and $t\bar{t}H$ samples.

Complete cutflows of *tHq* in comparison with *tW* and *tWZ* samples.

Complete cutflows of *tHq* in comparison with *tWH* and other *H* containing samples.

Complete cutflows of *tHq* in comparison with *tZq* and *t̄Z* samples.

Complete cutflows of *tHq* in comparison with Z+jets and *W*+jets samples.

2 *e*/μ + 1 τ_had cutflows

Complete cutflows of *tHq* in comparison with triboson and diboson samples.

Complete cutflows of *tHq* in comparison with s-channel single top production and 3/4-top samples.
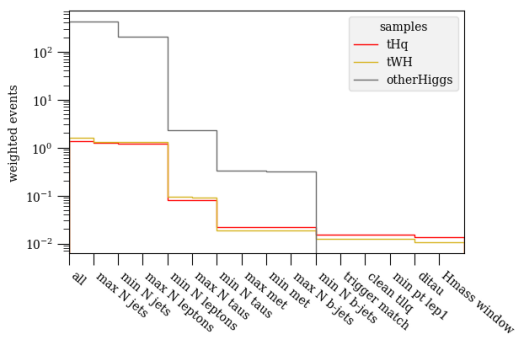
Complete cutflows of *tHq* in comparison with $t\bar{t}$ and t-channel single top production samples.
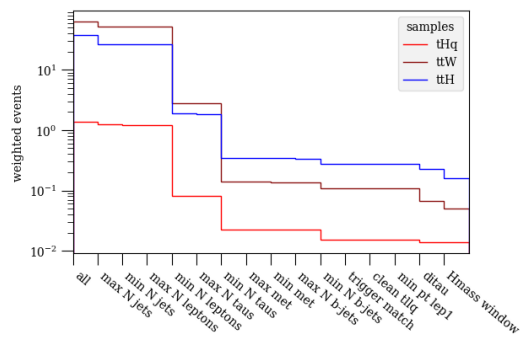
Complete cutflows of *tHq* in comparison with $t\bar{t}W$ and $t\bar{t}H$ samples.
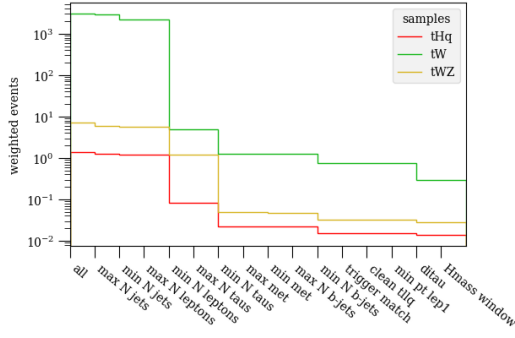
Complete cutflows of *tHq* in comparison with *tW* and *tWZ* samples.

Complete cutflows of *tHq* in comparison with *tWH* and other *H* containing samples.
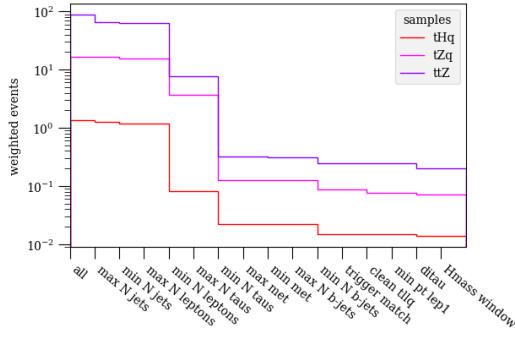
Complete cutflows of *tHq* in comparison with *tZq* and $t\bar{t}Z$ samples.

Complete cutflows of *tHq* in comparison with Z+jets and W+jets samples.

Fakes

	no fakes		tau gluon fake		tau quark fake		tau other fake	
tHq	0.831	± 0.025	0.050	± 0.007	0.084	± 0.008	0.0081	± 0.0025
tWH	0.72	± 0.04	0.069	± 0.013	0.047	± 0.013	0.000	± 0.005
tWZ	1.880	± 0.025	0.166	± 0.008	0.233	± 0.009	0.0308	± 0.0032
$t\bar{t}$	16.9	± 1.5	2.6	± 0.6	7.6	± 1.0	0.51	± 0.24
$t\bar{t}W$	19.2	± 0.4	2.13	± 0.14	7.32	± 0.24	0.73	± 0.07
$t\bar{t}Z$	14.00	± 0.33	1.44	± 0.10	2.50	± 0.11	0.198	± 0.033
$t\bar{t}H$	16.81	± 0.14	1.42	± 0.04	2.47	± 0.06	0.120	± 0.012
tZq	3.53	± 0.10	0.294	± 0.034	0.31	± 0.04	0.068	± 0.019
tW	0.83	± 0.32	0.16	± 0.12	0.74	± 0.30	0.0	± 0.0
Z+jets	0.22	± 0.17	0.13	± 0.05	0.33	± 0.09	0.23	± 0.19
diboson	6.43	± 0.27	0.57	± 0.10	0.99	± 0.11	0.62	± 0.27
triboson	0.27	± 0.04	0.017	± 0.009	0.006	± 0.006	0.0	± 0.0
t-channel	0.0	± 0.0	0.06	± 0.06	0.09	± 0.05	0.009	± 0.009
s-channel	0.0	± 0.0	0.0	± 0.0	0.0	± 0.0	0.0	± 0.0
W+jets	0.0	± 0.0	0.05	± 0.05	0.18	± 0.11	0.23	± 0.12
multitop	0.479	± 0.018	0.066	± 0.007	0.130	± 0.009	0.0010	± 0.0016
other Higgs	0.0	± 0.0	0.0	± 0.0	0.0	± 0.0	0.0	± 0.0
Total background	81.3	± 1.6	9.2	± 0.6	23.0	± 1.1	2.8	± 0.4

	light lep fake		QmisID	
tHq	0.0060	± 0.0015	0.045	± 0.006
tWH	0.0007	± 0.0007	0.041	± 0.009
tWZ	0.0136	± 0.0025	0.157	± 0.007
$t\bar{t}$	17.3	± 1.5	22.1	± 1.7
$t\bar{t}W$	0.033	± 0.017	1.43	± 0.10
$t\bar{t}Z$	0.053	± 0.028	1.30	± 0.10
$t\bar{t}H$	0.117	± 0.012	0.912	± 0.032
tZq	0.034	± 0.010	0.342	± 0.032
tW	0.95	± 0.34	0.56	± 0.27
Z+jets	0.23	± 0.17	4.7	± 2.7
diboson	0.32	± 0.26	1.35	± 0.15
triboson	0.0	± 0.0	0.012	± 0.006
t-channel	0.12	± 0.07	0.04	± 0.04
s-channel	0.0	± 0.0	0.0	± 0.0
W+jets	0.40	± 0.16	0.0	± 0.0
multitop	0.0020	± 0.0017	0.027	± 0.004
other Higgs	0.0	± 0.0	0.0	± 0.0
Total background	19.6	± 1.6	33.0	± 3.2

Table D.1: Contributions of different fake origins in the individual samples in the $2e/\mu$SS + $1\tau_{\text{had}}$ channel. One event might fall into multiple categories. The full weights before fake correction are applied.

	no fakes		tau gluon fake		tau quark fake		tau other fake	
tHq	1.103	± 0.028	0.086	± 0.008	0.225	± 0.013	0.016	± 0.004
tWH	1.62	± 0.06	0.143	± 0.020	0.284	± 0.028	0.020	± 0.007
tWZ	13.44	± 0.06	1.091	± 0.019	2.538	± 0.029	0.173	± 0.008
$t\bar{t}$	149	± 4	913	± 11	3 321	± 21	1 082	± 12
$t\bar{t}W$	39.9	± 0.6	3.85	± 0.18	14.46	± 0.34	1.28	± 0.10
$t\bar{t}Z$	88.7	± 0.7	9.59	± 0.28	26.5	± 0.6	1.67	± 0.16
$t\bar{t}H$	39.68	± 0.18	3.74	± 0.05	7.81	± 0.08	0.606	± 0.020
tZq	30.94	± 0.31	2.48	± 0.09	1.31	± 0.07	0.358	± 0.033
tW	5.6	± 0.9	52.1	± 2.6	117	± 4	75.9	± 3.1
Z+jets	18	± 4	570	± 25	2 080	± 70	1 060	± 50
diboson	79.8	± 1.0	10.9	± 0.6	64.6	± 1.8	13.0	± 0.8
triboson	0.75	± 0.05	0.039	± 0.012	0.027	± 0.010	0.013	± 0.007
t-channel	0.06	± 0.06	0.0	± 0.0	0.16	± 0.10	0.10	± 0.07
s-channel	0.0	± 0.0	0.0	± 0.0	0.0	± 0.0	0.0	± 0.0
W+jets	0.0	± 0.0	0.0	± 0.0	0.16	± 0.10	0.07	± 0.07
Rare top	1.070	± 0.027	0.124	± 0.010	0.224	± 0.012	0.0042	± 0.0021
other Higgs	3.2	± 1.3	0.7	± 0.6	4.6	± 1.6	2.3	± 1.1
Total background	472	± 6	1 567	± 27	5 640	± 70	2 240	± 50

	light lep fake		QmisID	
tHq	0.0037	± 0.0015	0.066	± 0.007
tWH	0.0033	± 0.0029	0.052	± 0.015
tWZ	0.0163	± 0.0020	0.612	± 0.013
$t\bar{t}$	26.1	± 1.9	205	± 5
$t\bar{t}W$	0.063	± 0.020	2.84	± 0.15
$t\bar{t}Z$	0.121	± 0.030	4.62	± 0.19
$t\bar{t}H$	0.136	± 0.013	1.90	± 0.04
tZq	0.034	± 0.010	1.30	± 0.06
tW	0.94	± 0.34	11.7	± 1.2
Z+jets	0.48	± 0.35	105	± 11
diboson	0.069	± 0.022	6.6	± 0.5
triboson	0.004	± 0.004	0.031	± 0.010
t-channel	0.16	± 0.10	0.06	± 0.06
s-channel	0.0	± 0.0	0.0	± 0.0
W+jets	0.23	± 0.12	0.0	± 0.0
Rare top	0.0002	± 0.0006	0.056	± 0.006
other Higgs	0.0	± 0.0	0.7	± 0.5
Total background	28.3	± 1.9	340	± 13

Table D.2: Contributions of different fake origins in the individual samples in the $2e/\mu\text{OS} + 1\tau_{\text{had}}$ channel. One event might fall into multiple categories. The full weights before fake correction are applied.

	no fakes		tau gluon fake		tau quark fake		tau other fake	
tHq	1.693	± 0.035	0.111	± 0.009	0.040	± 0.006	0.0027	± 0.0009
tWH	1.28	± 0.06	0.121	± 0.016	0.061	± 0.012	0.0005	± 0.0005
tWZ	3.435	± 0.033	0.245	± 0.009	0.104	± 0.006	0.0025	± 0.0007
$t\bar{t}$	15.1	± 1.4	72.5	± 3.1	367	± 7	37.9	± 2.3
$t\bar{t}W$	4.17	± 0.18	0.65	± 0.07	1.86	± 0.12	0.058	± 0.019
$t\bar{t}Z$	22.7	± 0.4	1.72	± 0.13	1.52	± 0.18	0.046	± 0.019
$t\bar{t}H$	27.31	± 0.19	2.35	± 0.06	1.32	± 0.04	0.024	± 0.006
tZq	8.98	± 0.17	0.51	± 0.04	0.151	± 0.023	0.008	± 0.006
tW	0.47	± 0.24	3.0	± 0.6	16.1	± 1.4	2.1	± 0.5
Z+jets	0.8	± 0.4	0.0	± 2.2	20	± 7	12	± 7
diboson	17.3	± 0.5	1.24	± 0.12	2.3	± 0.5	0.72	± 0.34
triboson	0.166	± 0.024	0.011	± 0.006	0.0025	± 0.0016	0.008	± 0.008
t-channel	0.0	± 0.0	0.70	± 0.19	3.2	± 0.4	0.94	± 0.23
s-channel	0.0	± 0.0	0.024	± 0.024	0.28	± 0.09	0.08	± 0.04
W+jets	0.0	± 0.0	5.0	± 1.6	20	± 6	39	± 12
Rare top	0.116	± 0.009	0.017	± 0.004	0.024	± 0.004	0.00015	± 0.00015
other Higgs	0.0	± 0.0	0.0	± 0.0	0.0	± 0.0	0.0	± 0.0
Total background	101.8	± 1.7	88	± 4	433	± 12	93	± 15

	tau 2 gluon fake		tau 2 quark fake		tau 2 other fake		light lep fake		QmisID	
tHq	0.095	± 0.009	0.070	± 0.007	0.0086	± 0.0025	0.0012	± 0.0006	0.035	± 0.005
tWH	0.081	± 0.016	0.075	± 0.015	0.007	± 0.004	0.0022	± 0.0022	0.036	± 0.008
tWZ	0.219	± 0.009	0.163	± 0.007	0.0130	± 0.0020	0.0035	± 0.0009	0.084	± 0.005
$t\bar{t}$	123	± 4	451	± 8	132	± 4	0.93	± 0.35	15.2	± 1.4
$t\bar{t}W$	0.54	± 0.07	2.36	± 0.14	0.23	± 0.04	0.0	± 0.0	0.18	± 0.04
$t\bar{t}Z$	2.03	± 0.14	2.23	± 0.20	0.10	± 0.06	0.028	± 0.012	0.68	± 0.08
$t\bar{t}H$	2.22	± 0.05	1.97	± 0.05	0.148	± 0.013	0.023	± 0.006	0.617	± 0.028
tZq	0.49	± 0.04	0.187	± 0.025	0.045	± 0.011	0.004	± 0.004	0.142	± 0.023
tW	4.8	± 0.8	21.1	± 1.7	8.4	± 1.0	0.10	± 0.10	0.50	± 0.24
Z+jets	3	± 4	29	± 7	12	± 7	0.13	± 0.14	1.1	± 0.8
diboson	1.20	± 0.18	2.7	± 0.5	1.1	± 0.4	0.022	± 0.021	0.47	± 0.11
triboson	0.018	± 0.008	0.0009	± 0.0008	0.008	± 0.008	0.0	± 0.0	0.0042	± 0.0023
t-channel	0.34	± 0.12	3.1	± 0.4	1.37	± 0.28	0.0	± 0.0	0.009	± 0.009
s-channel	0.05	± 0.04	0.25	± 0.08	0.07	± 0.04	0.0	± 0.0	0.012	± 0.012
W+jets	5	± 5	17	± 5	43	± 12	0.0	± 0.0	2.3	± 0.9
Rare top	0.015	± 0.004	0.048	± 0.005	0.0019	± 0.0013	0.0	± 0.0	0.0078	± 0.0018
other Higgs	0.0	± 0.0	0.0	± 0.0	0.0	± 0.0	0.0	± 0.0	0.0	± 0.0
Total background	143	± 8	531	± 12	199	± 14	1.2	± 0.4	21.3	± 1.9

Table D.3: Contributions of different fake origins in the individual samples in the $1e/\mu + 2\tau_{\text{had}}$ channel. One event might fall into multiple categories. The full weights before fake correction are applied.

More information on the NNs

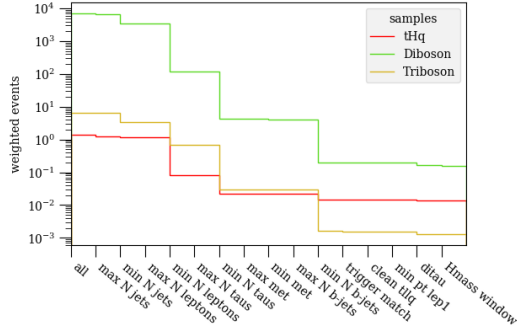| variable	explanation	$1e/\mu + 2\tau_{\text{had}}$	$2e/\mu\text{OS} + 1\tau_{\text{had}}$	$2e/\mu\text{SS} + 1\tau_{\text{had}}$
E_jet1	leading jet energy	x		
E_jet2	subleading jet energy	x		
E_lep1	leading $e/\mu/\tau_{\text{had}}$ energy	x		
HT_all	H_T	x	x	x
HtransMass	m_{T}^{H}	x		
HvisEta	η_{vis}^{H}	x	x	x
HvisMass	m_{vis}^{H}	x		
HvisPt	$p_{\text{T,vis}}^{H}$	x	x	x
MMC_out_1	m_{MMC}^{H}	x	x	x
M_b_jf	$m_{b,\text{j}^{\text{spect}}}$		x	x
TvisEta	η_{vis}^{t}	x	x	x
TvisMass	m_{vis}^{t}	x	x	x
TvisPhi	ϕ_{vis}^{t}	x		
TvisPt	$p_{\text{T,vis}}^{t}$		x	x
deltaPhiTau	$\Delta\phi\left(\tau_1, tau_2\right)$	x	x	x
deltaRTau	$\Delta R\left(\tau_1, tau_2\right)$	x	x	x
eta_b	$\eta(\text{j}_{b\text{-tag}}^{t})$	x	x	x
eta_jet1	$\eta(\text{leading jet})$	x	x	x
eta_jet2	$\eta(\text{subleading jet})$	x		
eta_jf	$\eta(\text{j}^{\text{spect}})$	x	x	x
eta_lep1	$\eta(\text{leading } e/\mu/\tau_{\text{had}})$	x		
had_tau_1_E	$E(\text{leading } \tau_{\text{had}})$	x		
had_tau_1_eta	$\eta(\text{leading } \tau_{\text{had}})$	x	x	x
had_tau_1_phi	$\phi(\text{leading } \tau_{\text{had}})$	x		
had_tau_1_m	$m(\text{leading } \tau_{\text{had}})$		x	x
had_tau_1_pt	$p_{\text{T}}(\text{leading } \tau_{\text{had}})$	x	x	x
had_tau_1_px	$p_x(\text{leading } \tau_{\text{had}})$	x		
had_tau_1_py	$p_y(\text{leading } \tau_{\text{had}})$	x		
had_tau_1_pz	$p_z(\text{leading } \tau_{\text{had}})$	x		

had_tau_2_E	$E(\text{subleading } \tau_{\text{had}})$	x		
had_tau_2_eta	$\eta(\text{subleading } \tau_{\text{had}})$	x		
had_tau_2_phi	$\phi(\text{subleading } \tau_{\text{had}})$	x		
had_tau_2_pt	$p_{\text{T}}(\text{subleading } \tau_{\text{had}})$	x		
had_tau_2_px	$p_x(\text{subleading } \tau_{\text{had}})$	x		
had_tau_2_py	$p_y(\text{subleading } \tau_{\text{had}})$	x		
had_tau_2_pz	$p_z(\text{subleading } \tau_{\text{had}})$	x		
lep_Higgs_eta	$\eta(e/\mu(H))$		x	x
lep_Higgs_phi	$\phi(e/\mu(H))$		x	x
lep_Higgs_pt	$p_{\text{T}}(e/\mu(H))$		x	x
lep_Top_E	$E(e/\mu(t))$	x		
lep_Top_eta	$\eta(e/\mu(p_{\text{T}}))$	x	x	x
lep_Top_phi	$\phi(e/\mu(p_{\text{T}}))$	x	x	x
lep_Top_pt	$p_{\text{T}}(e/\mu(p_{\text{T}}))$	x	x	x
lep_Top_px	$p_x(e/\mu(p_{\text{T}}))$	x		
lep_Top_py	$p_y(e/\mu(p_{\text{T}}))$	x		
lep_Top_pz	$p_z(e/\mu(p_{\text{T}}))$	x		
lep_top_higgs_mass	$m(e/\mu(p_{\text{T}}), e/\mu(H))$		x	
E_spectjet	$E(\text{j}^{\text{spect}})$	x		
met	$E_{\text{T}}^{\text{miss}}$	x	x	x
met_x	$E_{\text{x}}^{\text{miss}}$	x		
met_y	$E_{\text{y}}^{\text{miss}}$	x		
phi_met	$\phi(E_{\text{T}}^{\text{miss}})$		x	x
sumet	$\sum p_{\text{T}}$		x	x
phi_b	$\phi(\text{j}_{b\text{-tag}}^{t})$	x	x	x
phi_jet1	$\phi(\text{leading jet})$		x	x
phi_jf	$\phi(\text{j}^{\text{spect}})$	x	x	x
phi_lep1	$\phi(\text{leading } e/\mu/\tau_{\text{had}})$	x		
pt_b	$p_{\text{T}}(\text{j}_{b\text{-tag}}^{t})$	x	x	x
pt_jet1	$p_{\text{T}}(\text{leading jet})$		x	x
pt_jet2	$p_{\text{T}}(\text{subleading jet})$	x		
pt_jf	$p_{\text{T}}(\text{j}^{\text{spect}})$	x	x	x
pt_lep1	$p_{\text{T}}(\text{leading } e/\mu/\tau_{\text{had}})$	x		
top_mt	$m_T(t)$	x		
top_phi	$\phi(t)$	x		
top_pt	$p_{\text{T}}(t)$	x		

Table E.1: Variables used in the NNs for the tree different channels with explanation. The variables are sorted by type and not relevance for the networks.

1 e/μ + 2 τ_{had} with one loose but not tight b-tagged jet

Cutflows

Complete cutflows of *tHq* in comparison with triboson and diboson samples.

Complete cutflows of *tHq* in comparison with s-channel single top production and 3/4-top samples.

Complete cutflows of *tHq* in comparison with $t\bar{t}$ and t-channel single top production samples.
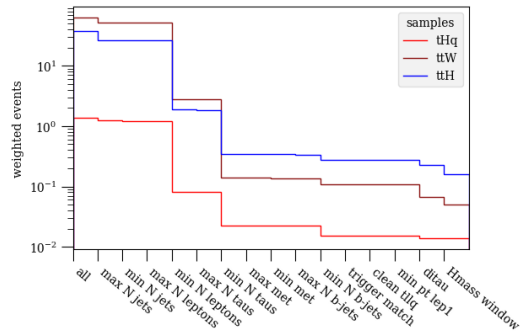
Complete cutflows of *tHq* in comparison with $t\bar{t}W$ and $t\bar{t}H$ samples.
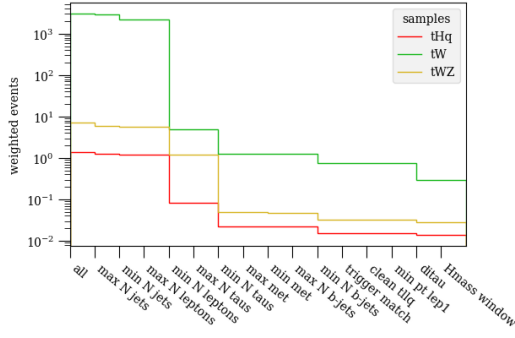
Complete cutflows of *tHq* in comparison with *tW* and *tWZ* samples.

Complete cutflows of *tHq* in comparison with *tWH* and other *H* containing samples.

Complete cutflows of *tHq* in comparison with *tZq* and *t̄Z* samples.
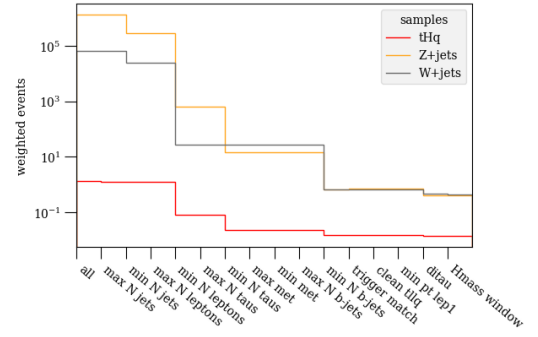
Complete cutflows of *tHq* in comparison with Z+jets and W+jets samples.

Fakes

In the $1e/\mu + 2\tau_{\text{had}}$ looseNotTight b-tagged jet region the fake weights derived from the $1e/\mu + 2\tau_{\text{had}}$ region with one tight b-tagged jet are applied.

	no fakes		tau gluon fake		tau quark fake		tau other fake	
tHq	0.519	± 0.020	0.032	± 0.005	0.010	± 0.004	0.0010	± 0.0007
tWH	0.471	± 0.034	0.043	± 0.010	0.027	± 0.008	0.0005	± 0.0005
tWZ	1.272	± 0.021	0.089	± 0.006	0.040	± 0.004	0.0015	± 0.0005
$t\bar{t}$	4.2	± 0.8	23.0	± 1.8	116	± 4	13.3	± 1.5
$t\bar{t}W$	1.27	± 0.10	0.096	± 0.035	0.44	± 0.06	0.022	± 0.011
$t\bar{t}Z$	8.03	± 0.27	0.73	± 0.08	0.58	± 0.11	0.015	± 0.014
$t\bar{t}H$	9.73	± 0.11	0.897	± 0.035	0.524	± 0.027	0.009	± 0.004
tZq	2.79	± 0.09	0.184	± 0.025	0.036	± 0.013	0.0036	± 0.0028
tW	0.13	± 0.10	1.5	± 0.5	4.2	± 0.7	1.00	± 0.33
Z+jets	0.9	± 0.4	0	± 15	23	± 6	24	± 10
diboson	39.3	± 0.8	2.5	± 0.4	5.3	± 0.8	2.5	± 0.7
triboson	0.257	± 0.024	0.021	± 0.010	0.006	± 0.005	0.0010	± 0.0010
t-channel	0.0	± 0.0	0.24	± 0.11	0.98	± 0.24	0.18	± 0.08
s-channel	0.0	± 0.0	0.0	± 0.0	0.034	± 0.026	0.042	± 0.030
W+jets	0.0	± 0.0	13	± 9	24	± 26	130	± 50
multi top	0.069	± 0.007	0.0113	± 0.0027	0.0165	± 0.0029	0.0	± 0.0
other Higgs	1.0	± 0.7	0.6	± 0.6	0.0	± 0.0	0.0	± 0.0
Total background	69.3	± 1.4	43	± 9	175	± 27	170	± 50

	tau 2 gluon fake		tau 2 quark fake		tau 2 other fake		light lep fake		QmisID	
tHq	0.036	± 0.005	0.020	± 0.004	0.0027	± 0.0016	0.0010	± 0.0005	0.0083	± 0.0027
tWH	0.024	± 0.010	0.018	± 0.008	0.0035	± 0.0027	0.0	± 0.0	0.006	± 0.004
tWZ	0.083	± 0.005	0.071	± 0.005	0.0054	± 0.0013	0.0013	± 0.0007	0.0249	± 0.0034
$t\bar{t}$	38.3	± 2.3	142	± 4	41.8	± 2.4	0.14	± 0.14	2.8	± 0.6
$t\bar{t}W$	0.17	± 0.04	0.79	± 0.08	0.098	± 0.023	0.0	± 0.0	0.039	± 0.021
$t\bar{t}Z$	0.58	± 0.08	0.69	± 0.13	0.094	± 0.034	0.004	± 0.006	0.23	± 0.05
$t\bar{t}H$	0.794	± 0.032	0.742	± 0.032	0.057	± 0.007	0.010	± 0.004	0.216	± 0.017
tZq	0.141	± 0.025	0.066	± 0.015	0.016	± 0.007	0.0055	± 0.0030	0.036	± 0.013
tW	1.7	± 0.5	5.4	± 0.8	3.3	± 0.7	0.0	± 0.0	0.0	± 0.0
Z+jets	11	± 4	21	± 16	24	± 10	0.6	± 0.4	9	± 4
diboson	4.3	± 0.5	5.0	± 0.7	3.4	± 0.8	0.07	± 0.05	1.29	± 0.16
triboson	0.014	± 0.006	0.015	± 0.007	0.0019	± 0.0012	0.00022	± 0.00016	0.0073	± 0.0029
t-channel	0.18	± 0.10	0.89	± 0.21	0.33	± 0.15	0.0	± 0.0	0.0	± 0.0
s-channel	0.0009	± 0.0009	0.034	± 0.027	0.041	± 0.030	0.0	± 0.0	0.0	± 0.0
W+jets	0	± 22	54	± 12	120	± 50	0.0	± 0.0	7	± 4
multi top	0.0066	± 0.0026	0.027	± 0.004	0.0002	± 0.0009	0.00029	± 0.00029	0.0025	± 0.0011
other Higgs	0.0	± 0.0	0.0	± 0.0	0.0	± 0.0	0.0	± 0.0	0.0	± 0.0
Total background	57	± 5	230	± 21	190	± 50	0.8	± 0.4	21	± 6

Table F.1: Contributions of different fake origins in the individual samples in the $1e/\mu + 2\tau_{\text{had}}$ looseNotTight *b*-tagged jet region. One event might fall into multiple categories. The full weights before fake correction are applied.
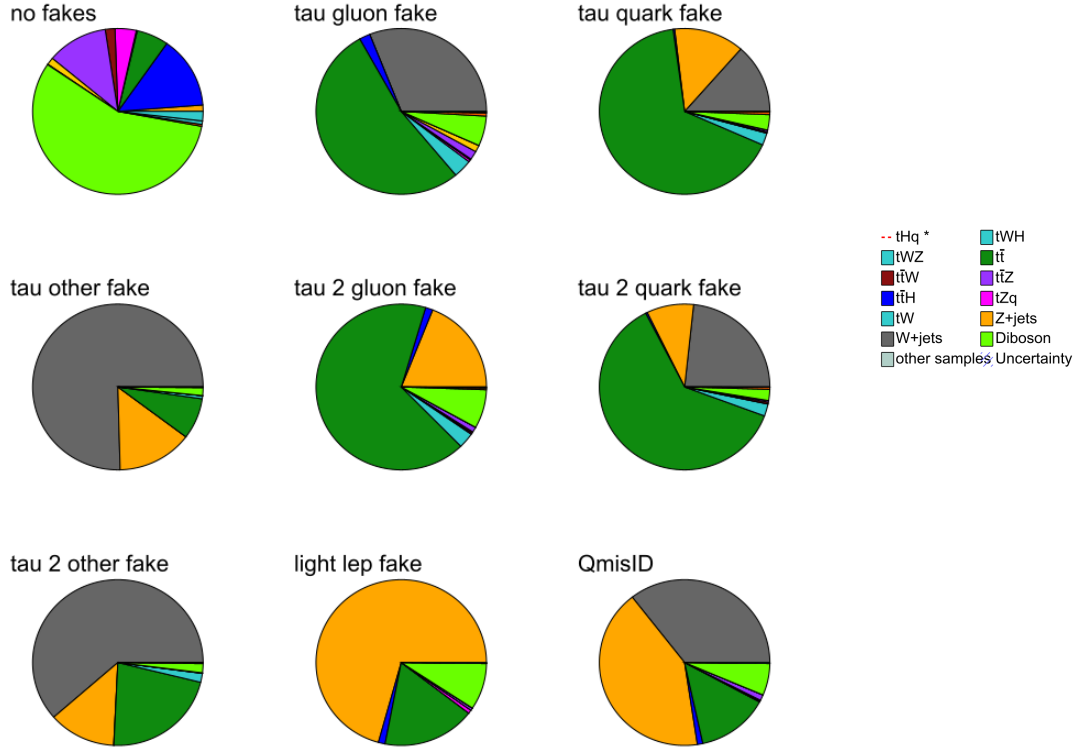
Figure F.5: Fake origins in the different samples of the $1e/\mu + 2\tau_{\text{had}}$ looseNotTight b-tagged jet region. One event might fall into multiple of the categories (except for 'no fakes') The full weights before fake correction are applied. As *tHq* is the signal process it is not shown in these piecharts.
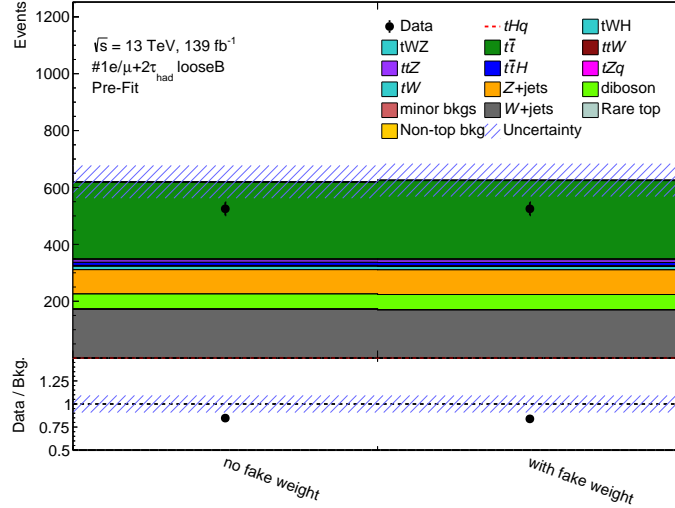

Figure F.6: Application of the fake estimation in the $1e/\mu + 2\tau_{\text{had}}$ looseNotTight b-tagged jet region.

Yields

	$1e/\mu + 2\tau_{\text{had}}$ looseB
tHq	0.554 ± 0.020
tWH	0.50 ± 0.04
tWZ	1.399 ± 0.022
$t\bar{t}$	269 ± 6
$t\bar{t}W$	2.43 ± 0.14
$t\bar{t}Z$	9.32 ± 0.31
$t\bar{t}H$	10.91 ± 0.12
tZq	2.97 ± 0.10
tW	11.9 ± 1.2
Z+jets	86 ± 21
diboson	52.4 ± 1.4
triboson	0.284 ± 0.026
t-channel	1.40 ± 0.28
s-channel	0.08 ± 0.04
W+jets	170 ± 50
multitop	0.107 ± 0.008
other Higgs	1.0 ± 0.7
Total background	620 ± 60
Data	525

Table F.2: Yields before fake weight application.

	$1e/\mu$ + $2\tau_{\text{had}}$ looseB
tHq	0.563 ± 0.021
tWH	0.51 ± 0.04
tWZ	1.426 ± 0.022
$t\bar{t}$	275 ± 6
$t\bar{t}W$	2.45 ± 0.14
$t\bar{t}Z$	9.61 ± 0.33
$t\bar{t}H$	11.20 ± 0.13
tZq	3.01 ± 0.10
tW	11.8 ± 1.2
Z+jets	87 ± 21
diboson	52.9 ± 1.4
triboson	0.287 ± 0.026
t-channel	1.39 ± 0.27
s-channel	0.07 ± 0.04
W+jets	170 ± 50
multitop	0.116 ± 0.009
other Higgs	1.0 ± 0.7
Total background	620 ± 60
Data	525

Table F.3: Yields after fake weight application.

Further Asimov fit studies

Expected limits and significance

The expected limits as explained in section 7.5.1.2 are also calculated for the other channels. They are shown together with the expected significances (see section 4.4):

1 e/μ + 2 τ_{had} :

$$\text{limit(exp.)} = 23.5 \ \left(\pm 1\sigma^{34.5}_{16.9} \ \pm 2\sigma^{49.9}_{12.6}\right) \ \mu(tHq)(\text{SM}) \tag{G.1}$$

$$\text{significance(exp.)} = 0.11 \tag{G.2}$$

2 e/μ OS + 1 τ_{had} :

$$\text{limit(exp.)} = 70.4 \ \left(\pm 1\sigma^{108.2}_{50.7} \ \pm 2\sigma^{179.2}_{37.8}\right) \ \mu(tHq)(\text{SM}) \tag{G.3}$$

$$\text{significance(exp.)} = 0.05 \tag{G.4}$$

This result is not fully comparable to the other results. For the limit in the $2e/\mu\text{OS} + 1\tau_{\text{had}}$ channel a signal has to be injected, otherwise no limit is found.

2 e/μ SS + 1 τ_{had} :

$$\text{limit(exp.)} = 30.7 \ \left(\pm 1\sigma^{48.1}_{22.1} \ \pm 2\sigma^{77.2}_{16.5}\right) \ \mu(tHq)(\text{SM}) \tag{G.5}$$

$$\text{significance(exp.)} = 0.09 \tag{G.6}$$

Combination:

$$\text{limit(exp.)} = 15.5 \ \left(\pm 1\sigma^{22.4}_{11.1} \ \pm 2\sigma^{31.8}_{18.3}\right) \ \mu(tHq)(\text{SM}) \tag{G.7}$$

$$\text{significance(exp.)} = 0.14 \tag{G.8}$$

Pruning plots of all channels
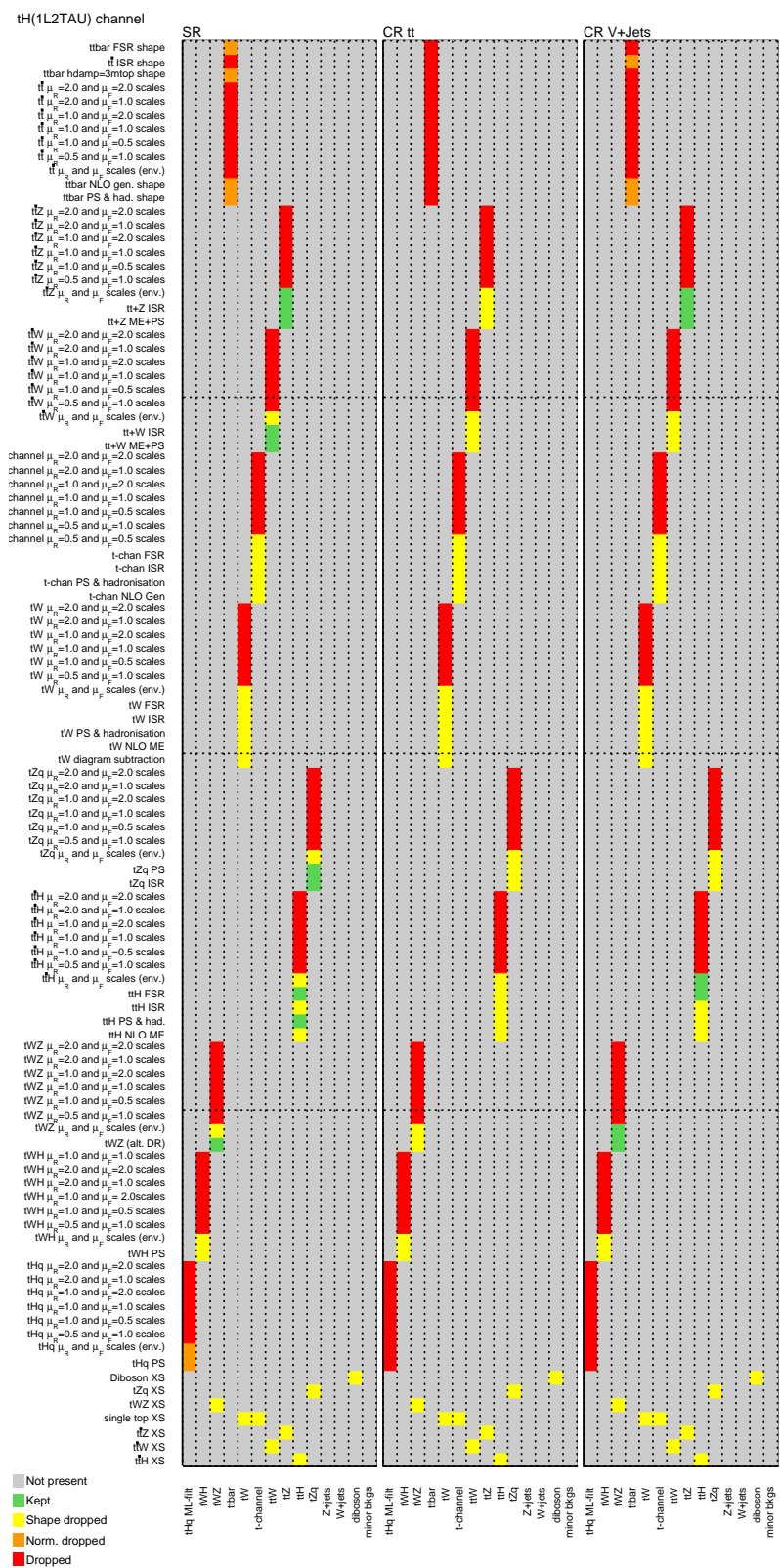
Figure G.1: Pruning of the different theory related NPs in the $1e/\mu + 2\tau_{\text{had}}$ channel. Grey combinations are not present, green ones are kept. Red combinations are completely dropped. For orange combinations only the shape component is kept, while for yellow ones only the normalization is kept.
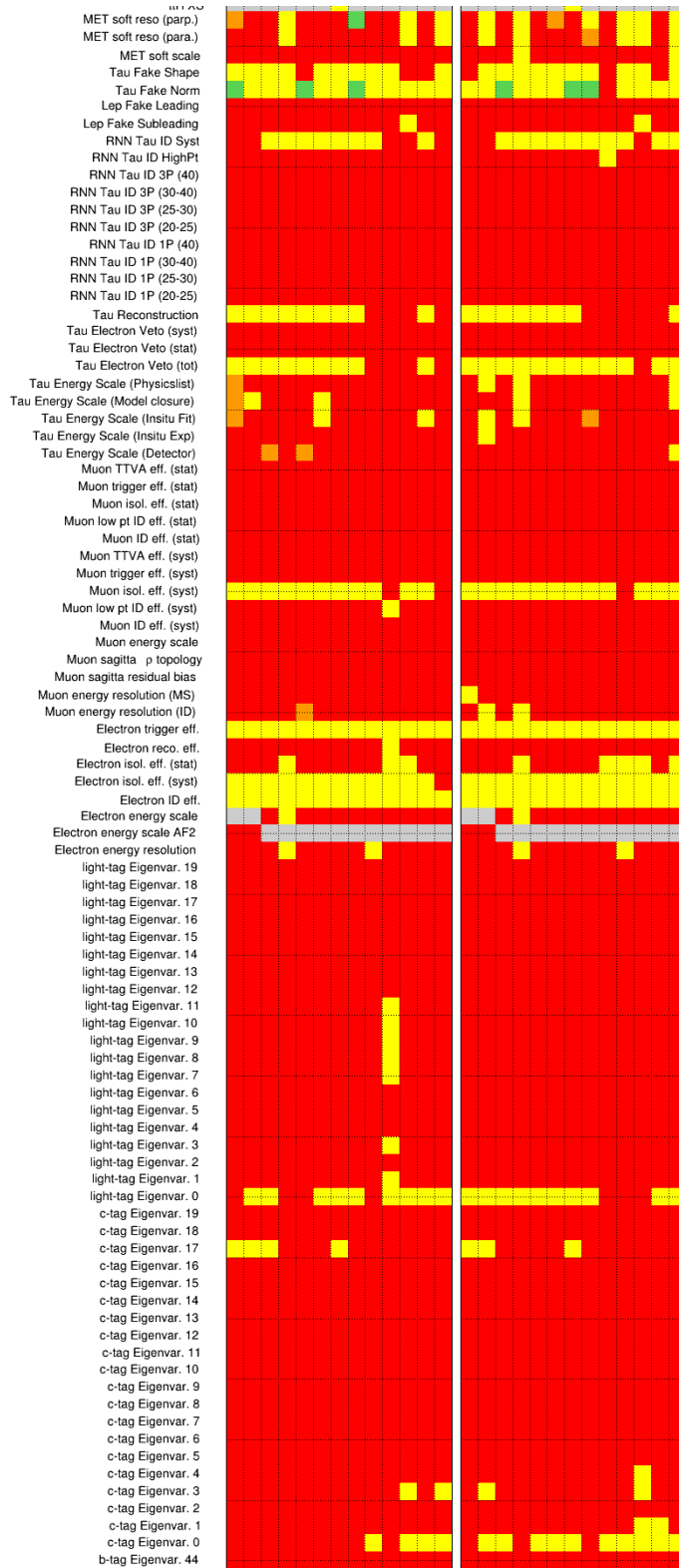
Figure G.2: Pruning of different reconstruction related NPs in the $1e/\mu + 2\tau_{\text{had}}$ channel. Grey combinations are not present, green ones are kept. Red combinations are completely dropped. For orange combinations only the shape component is kept, while for yellow ones only the normalization is kept.
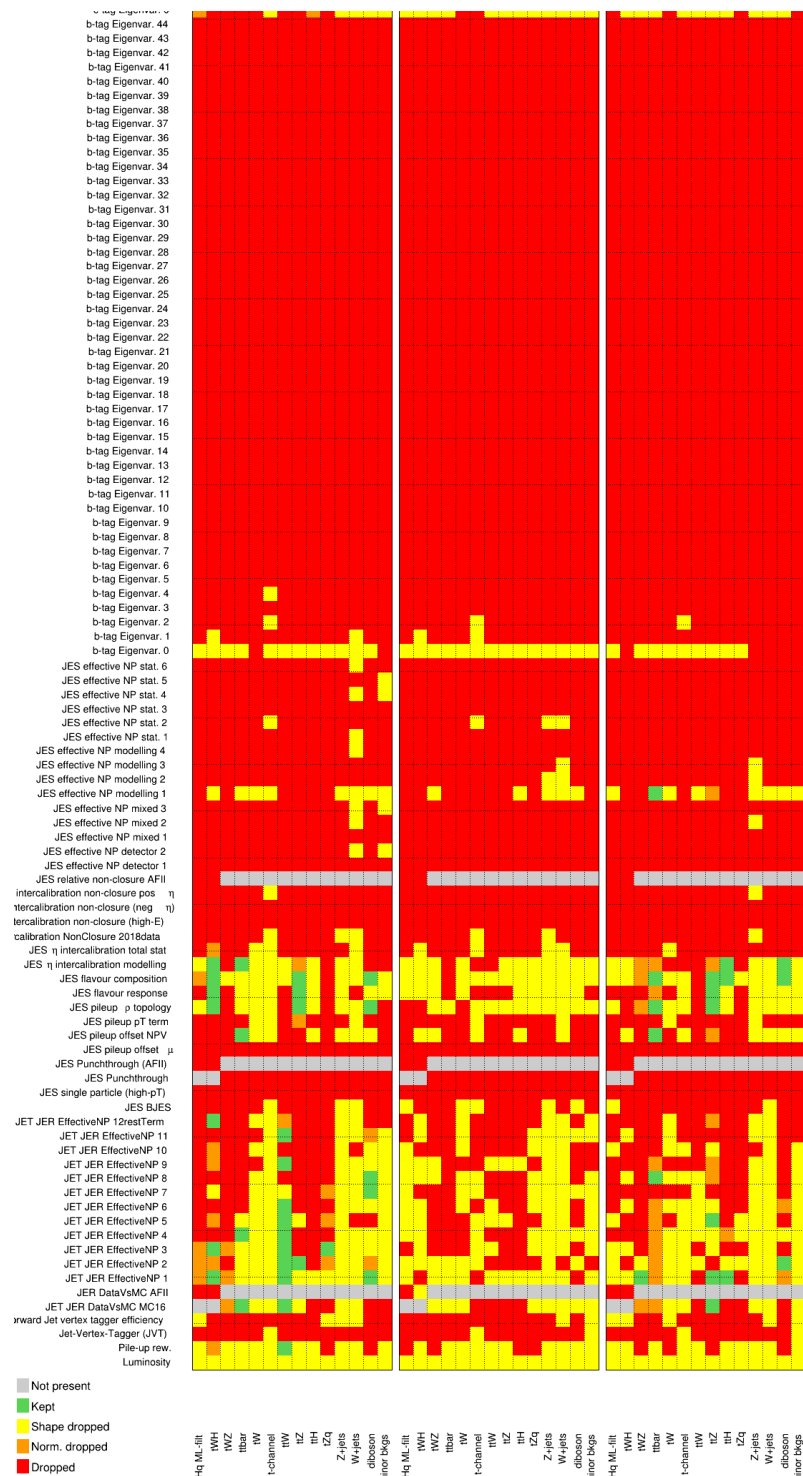
Figure G.3: Pruning of the last NPs in the $1e/\mu + 2\tau_{\mathrm{had}}$ channel. Grey combinations are not present, green ones are kept. Red combinations are completely dropped. For orange combinations only the shape component is kept, while for yellow ones only the normalization is kept.
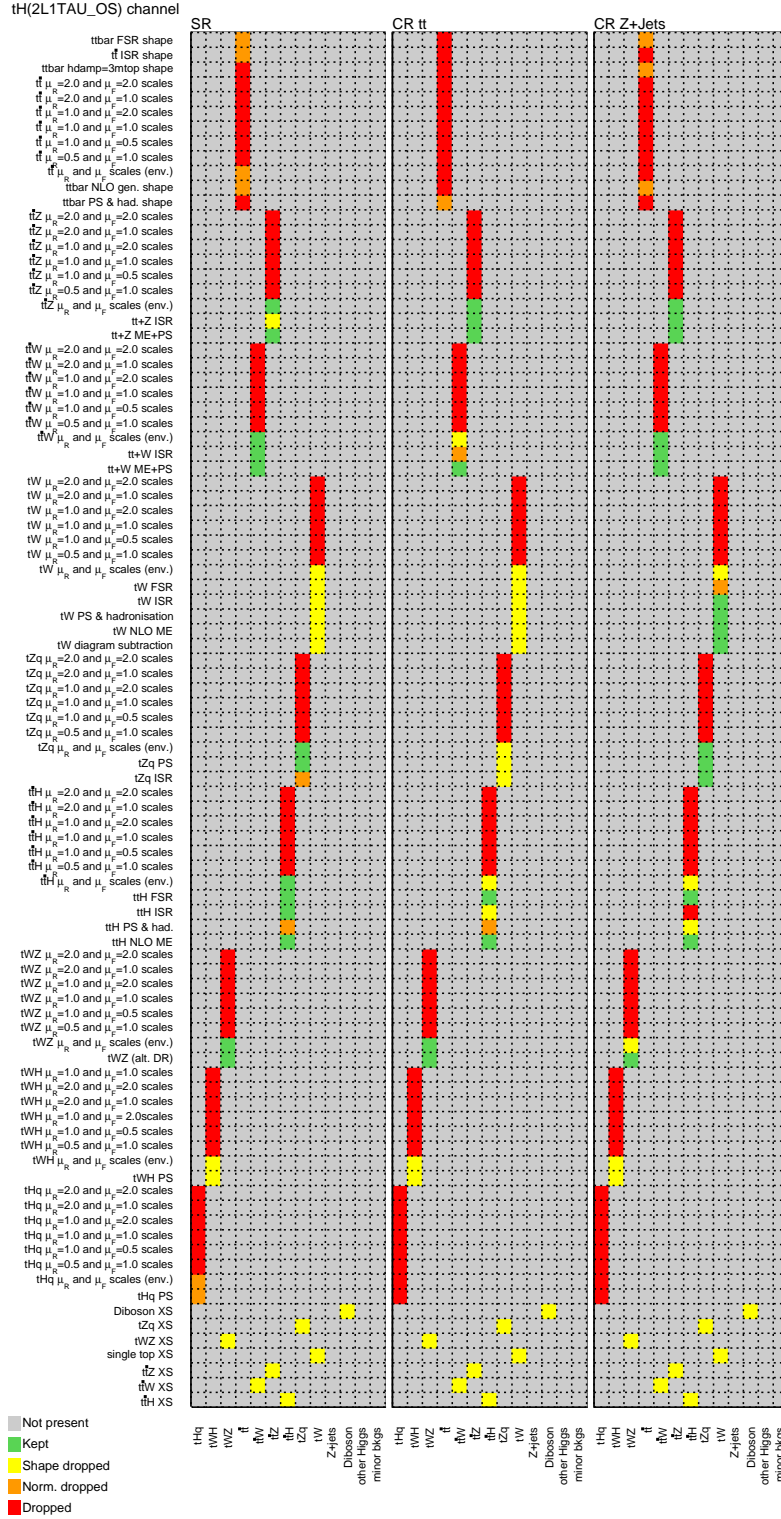
Figure G.4: Pruning of the different theory related NPs in the $2e/\mu\mathrm{OS} + 1\tau_{\mathrm{had}}$ channel. Grey combinations are not present, green ones are kept. Red combinations are completely dropped. For orange combinations only the shape component is kept, while for yellow ones only the normalization is kept.
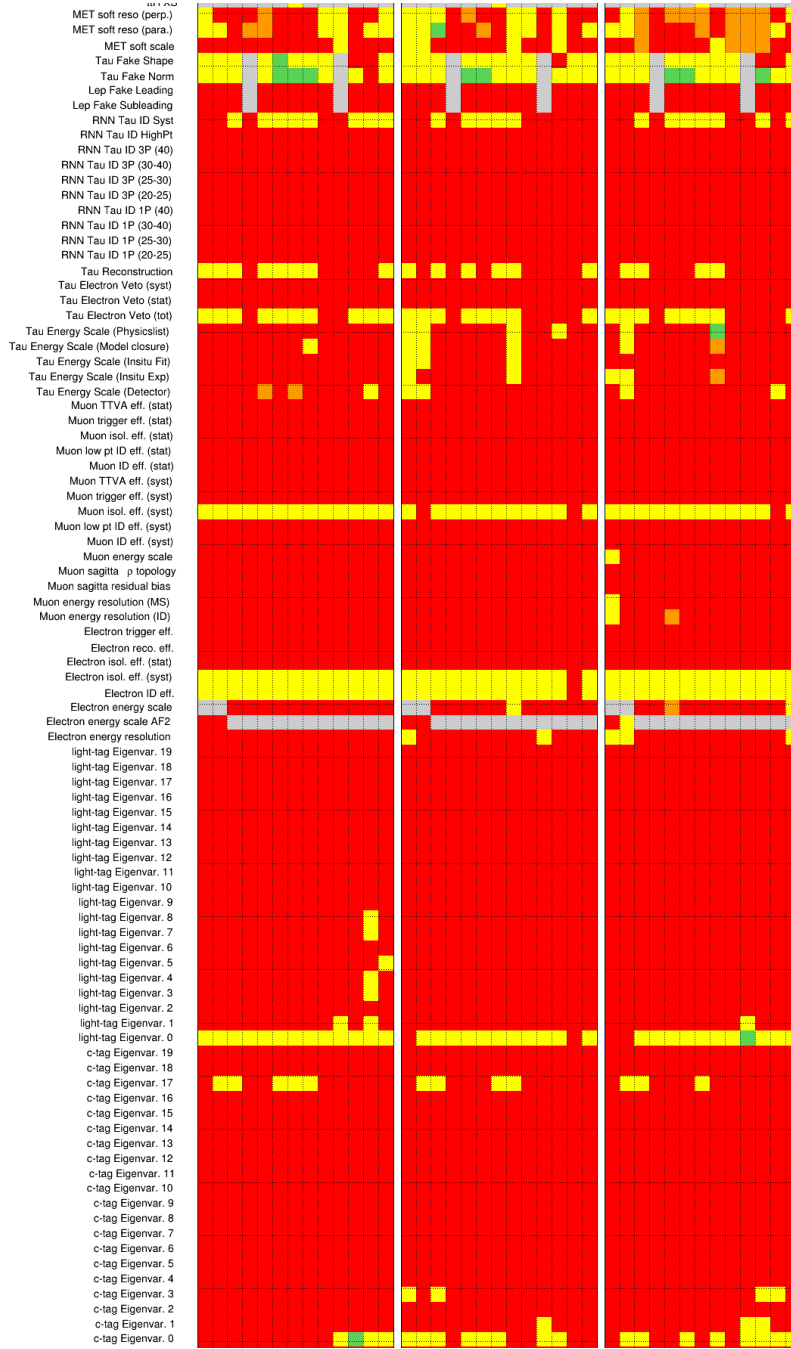
Figure G.5: Pruning of different reconstruction related NPs in the $2e/\mu$OS $+ 1\tau_{\mathrm{had}}$ channel. Grey combinations are not present, green ones are kept. Red combinations are completely dropped. For orange combinations only the shape component is kept, while for yellow ones only the normalization is kept.

Figure G.6: Pruning of the last NPs in the $2e/\mu$OS + $1\tau_{\text{had}}$ channel. Grey combinations are not present, green ones are kept. Red combinations are completely dropped. For orange combinations only the shape component is kept, while for yellow ones only the normalization is kept.
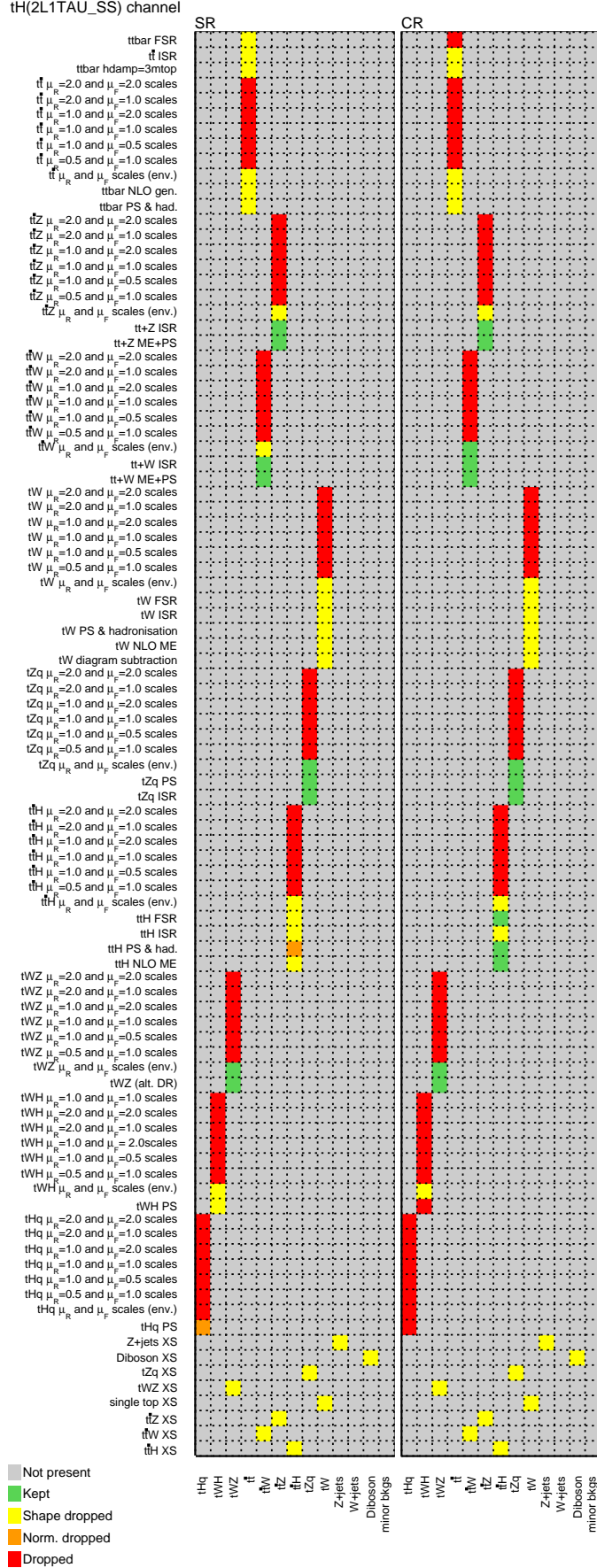
Figure G.7: Pruning of the different theory related NPs in the $2e/\mu\mathrm{SS} + 1\tau_{\mathrm{had}}$ channel. Grey combinations are not present, green ones are kept. Red combinations are completely dropped. For orange combinations only the shape component is kept, while for yellow ones only the normalization is kept.

Figure G.8: Pruning of different reconstruction related NPs in the $2e/\mu\text{SS} + 1\tau_{\text{had}}$ channel. Grey combinations are not present, green ones are kept. Red combinations are completely dropped. For orange combinations only the shape component is kept, while for yellow ones only the normalization is kept.

Not present

NP plots

(a) experimental

(b) JES/JER

(c) flavor-tagging

(d) theory

(e) theory $t\bar{t}$
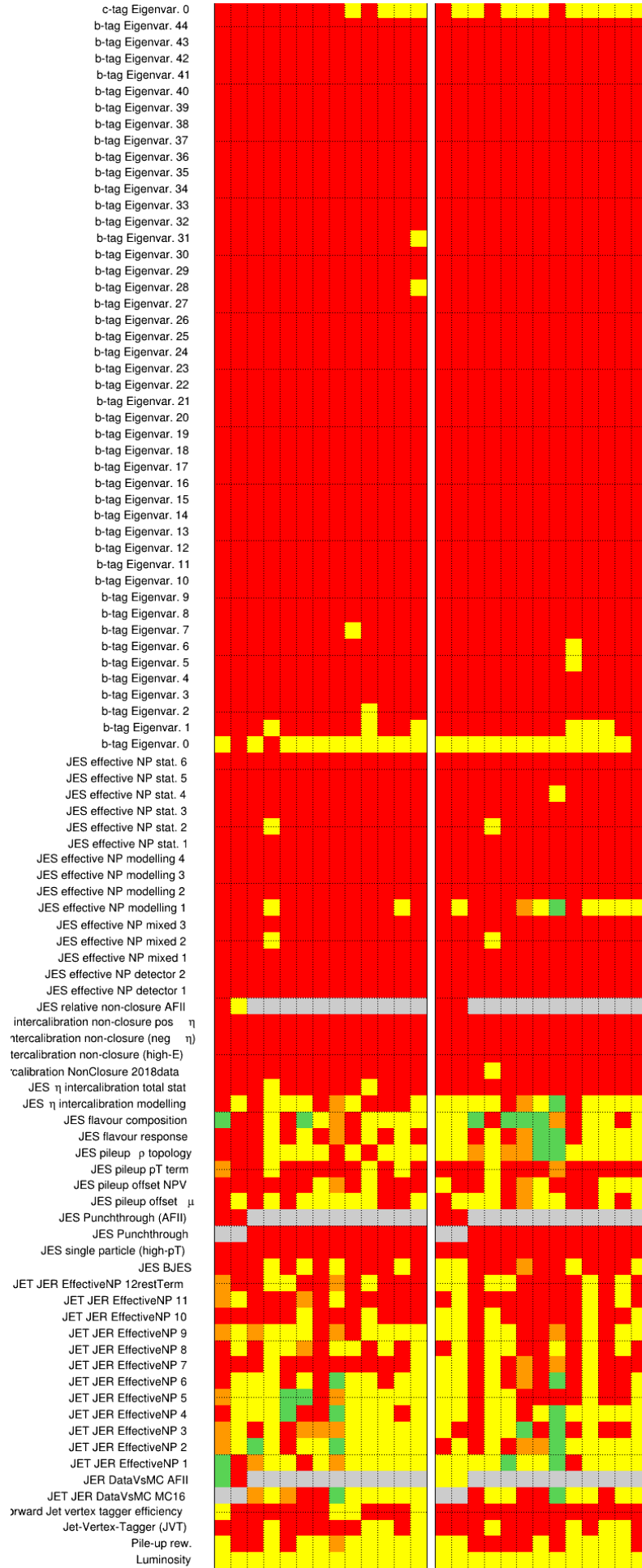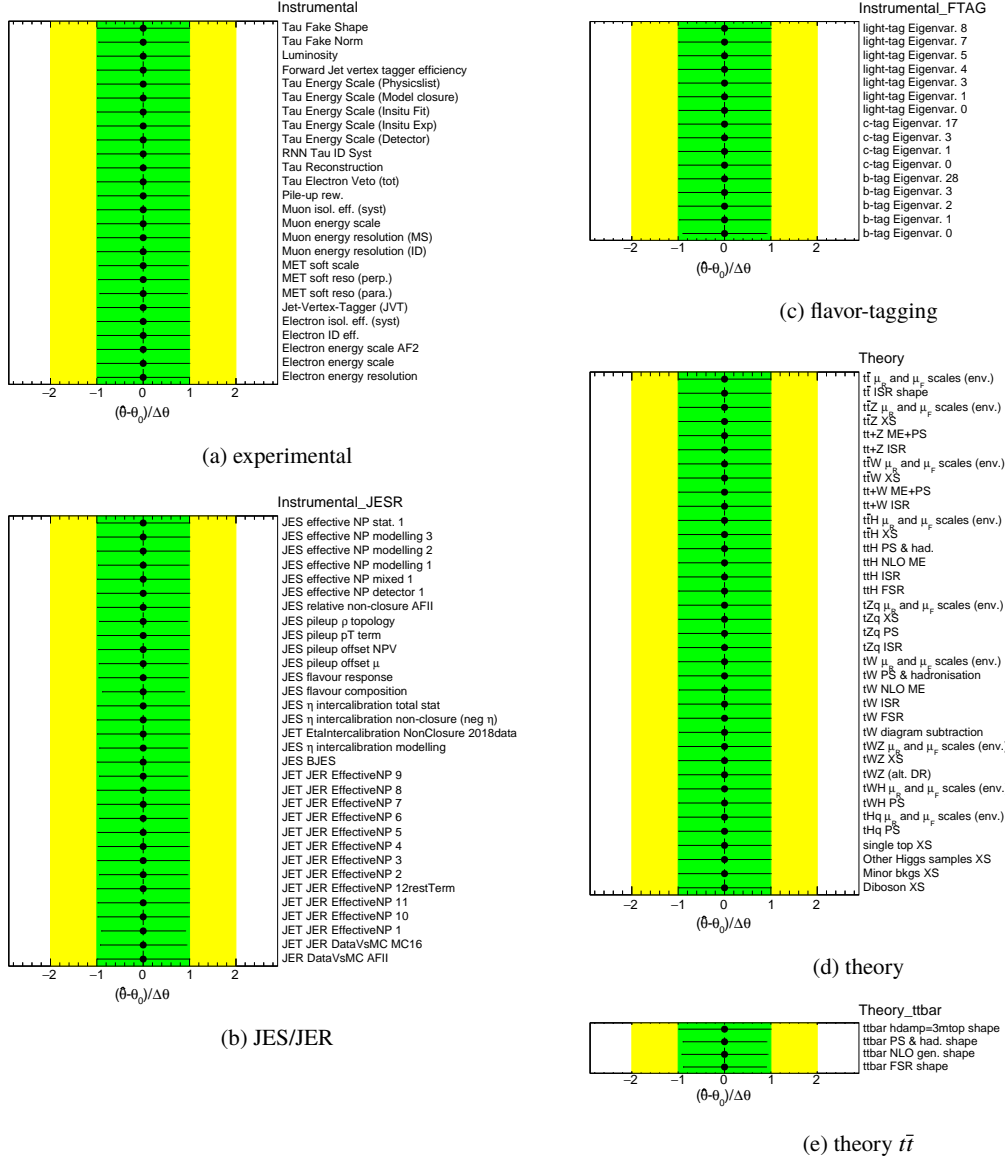
Figure G.10: Constraints of all NPs in the $2e/\mu\mathrm{OS} + 1\tau_{\mathrm{had}}$ region after the Asimov fit. The NPs are roughly grouped by their origin.

(a) experimental

(b) JES/JER

(c) flavor-tagging

(d) theory

(e) theory $t\bar{t}$
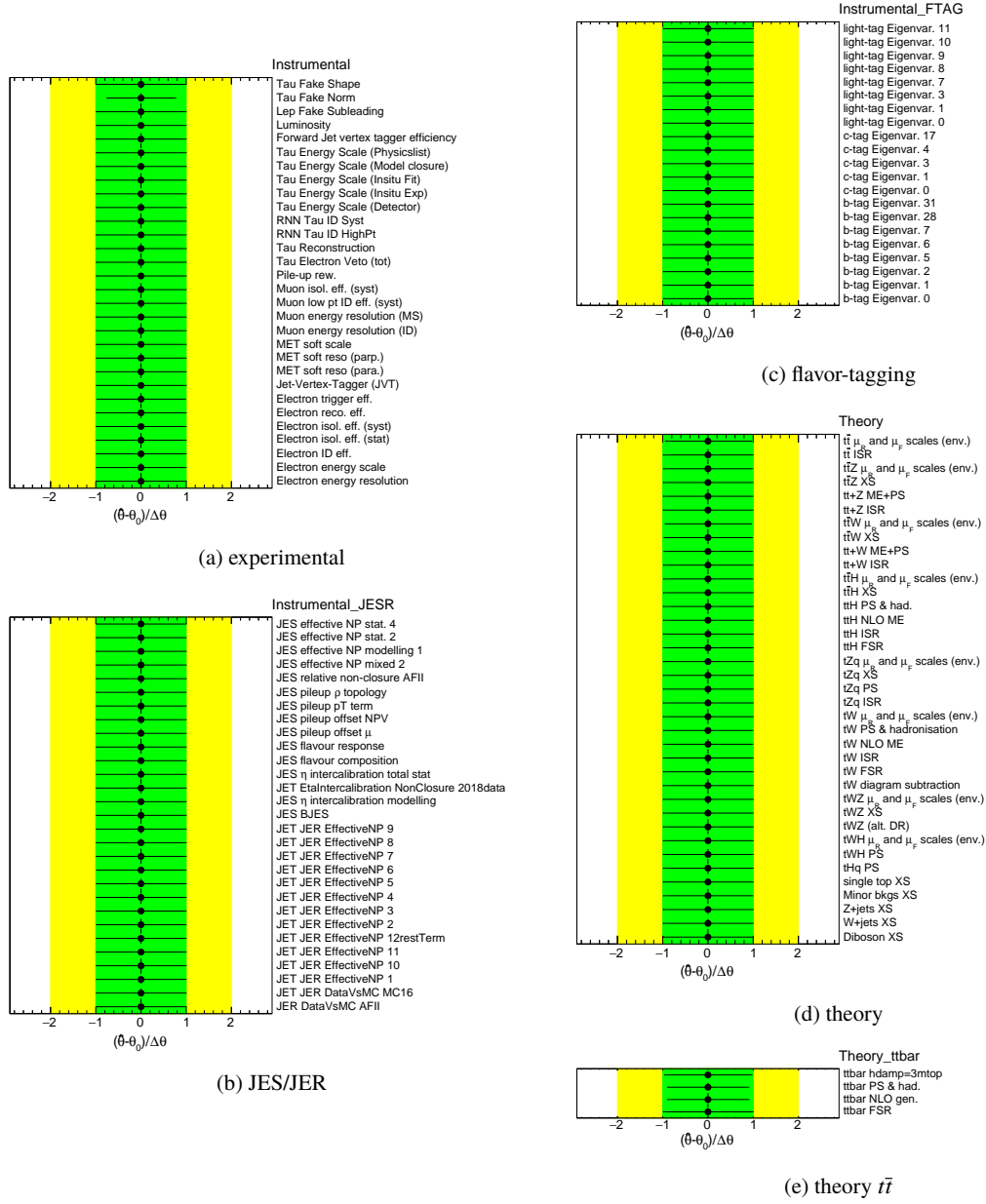
Figure G.11: Constraints of all NPs in the $2e/\mu\mathrm{OS} + 1\,\tau_{\mathrm{had}}$ region after the Asimov fit. The NPs are roughly grouped by their origin.

Combination

	2e/μOS+1τ_had	2e/μSS+1τ_had	1e/μ+2τ_had	Combine
$\mu(t\bar{t})$OS	$1.00^{+0.04}_{-0.04}$	$-10.00^{+0.00}_{-0.00}$	$-10.00^{+0.00}_{-0.00}$	$1.00^{+0.04}_{-0.04}$
μ(Zjets)OS	$1.00^{+0.10}_{-0.10}$	$-10.00^{+0.00}_{-0.00}$	$-10.00^{+0.00}_{-0.00}$	$1.00^{+0.09}_{-0.09}$
$\mu(tt)$ 2τ	$-10.00^{+0.00}_{-0.00}$	$-10.00^{+0.00}_{-0.00}$	$1.00^{+0.13}_{-0.13}$	$1.00^{+0.12}_{-0.12}$
$\mu(V+Jets)$ 2τ	$-10.00^{+0.00}_{-0.00}$	$-10.00^{+0.00}_{-0.00}$	$1.00^{+0.73}_{-0.73}$	$1.00^{+0.70}_{-0.70}$
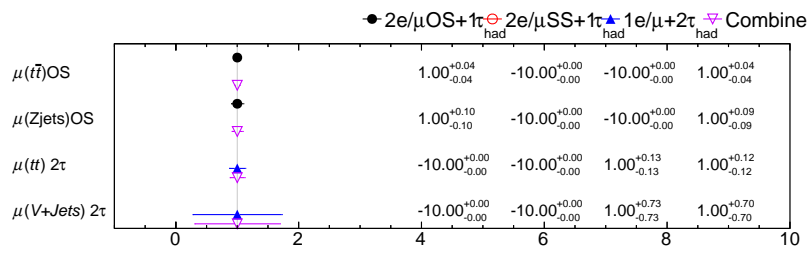
Figure G.12: NFs from the fits to Asimov data (except $\mu(tHq)$), before and after combination.

(a) experimental

(b) JES/JER

(c) flavor-tagging

(d) theory

(e) theory $t\bar{t}$
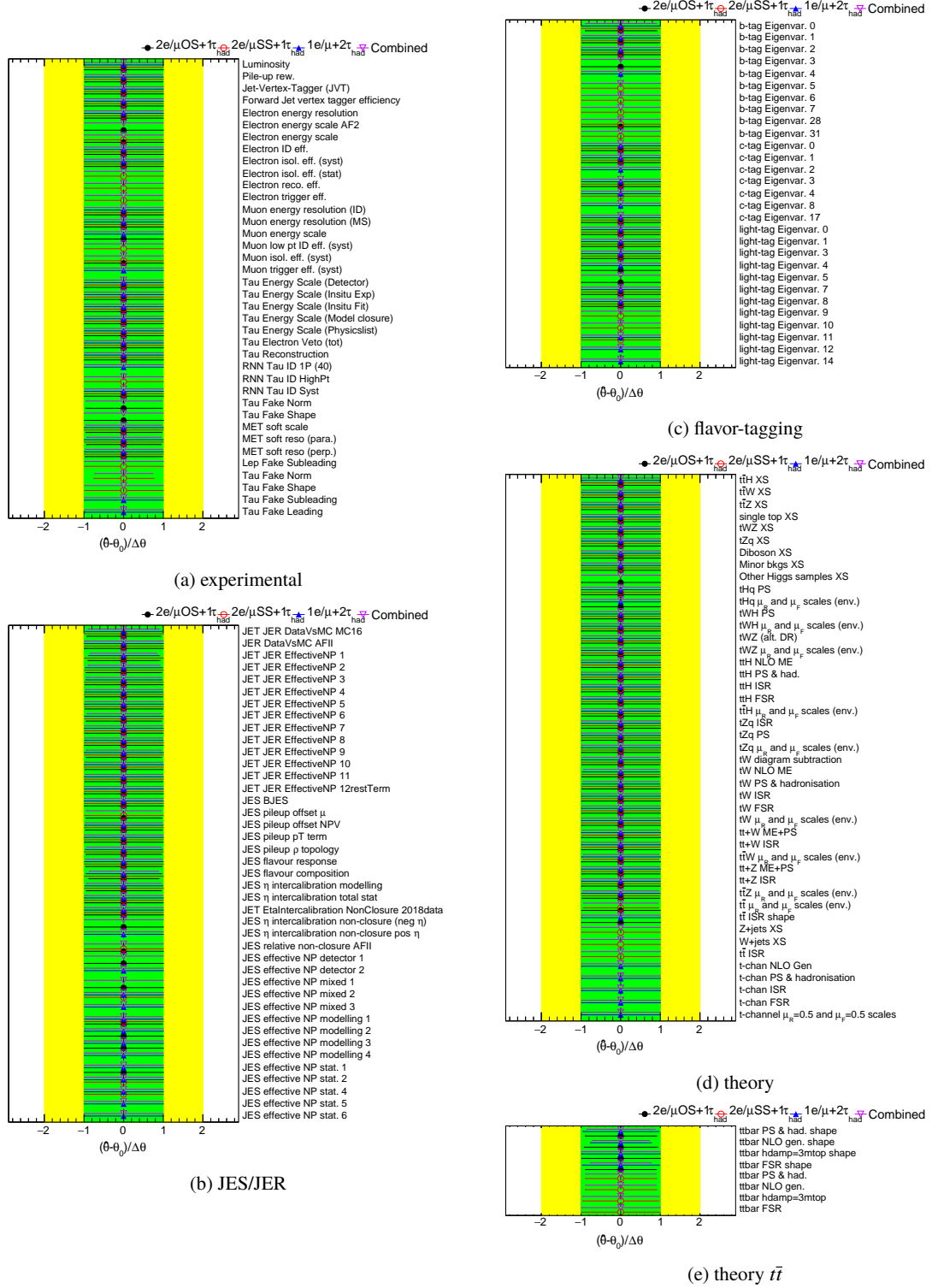
Figure G.13: Constraints of all NPs in the all channels after the Asimov fit. The NPs are roughly grouped by their origin.

Additional studies on the partial unblinded fits

Background fit in the CRs

The fits presented here are fitted under the background only hypothesis. Therefore, $\mu(tHq)$ is fixed to 0. Only information about the CRs is included in the fit. The fit is performed with data in these regions.

1 e/μ + 2 τ_{had}

Figure H.1: NFs determined from the fit in the $1e/\mu + 2\tau_{had}$ channel using the CRs only and fixing $\mu(tHq)$ to zero.

(a) CR $t\bar{t}$ pre-fit

(b) CR $t\bar{t}$ post-fit

(c) CR V+jets pre-fit

(d) CR V+jets post-fit

Figure H.2: Simulation to data agreement in the $1e/\mu + 2\tau_{\text{had}}$ channel before and after the fit to the background in the CRs.

Instrumental
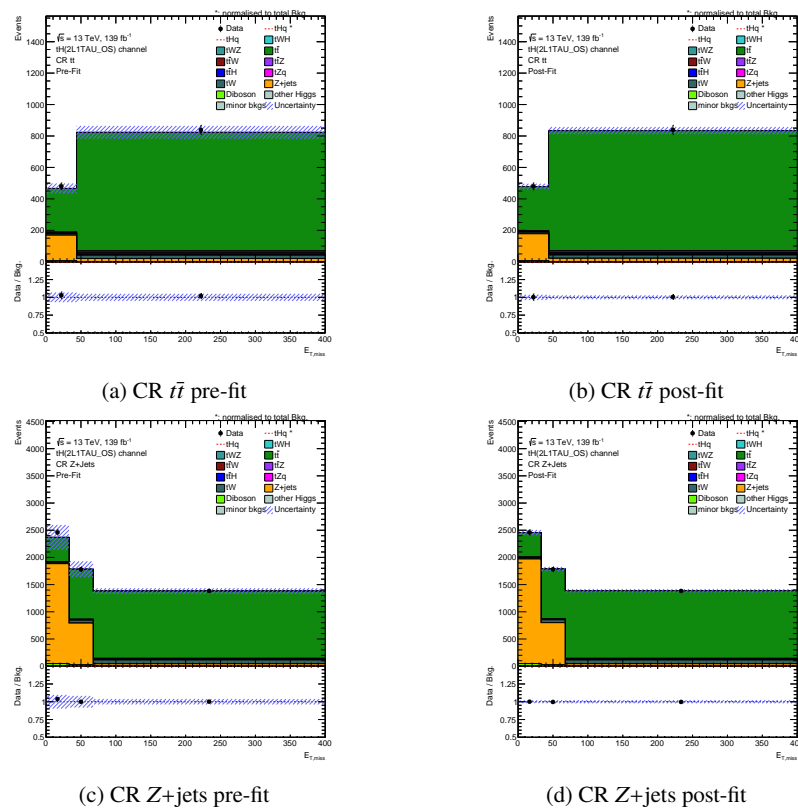
Tau Fake Subleading
Tau Fake Leading
Luminosity
Forward Jet vertex tagger efficiency
Tau Energy Scale (Physicslist)
Tau Energy Scale (Model closure)
Tau Energy Scale (Insitu Fit)
Tau Energy Scale (Insitu Exp)
Tau Energy Scale (Detector)
RNN Tau ID Syst
RNN Tau ID 1P (40)
Tau Reconstruction
Tau Electron Veto (tot)
Pile-up rew.
Muon trigger eff. (syst)
Muon energy scale
Muon energy resolution (MS)
Muon energy resolution (ID)
MET soft reso (parp.)
MET soft reso (para.)
Jet-Vertex-Tagger (JVT)
Electron isol. eff. (syst)
Electron ID eff.
Electron energy resolution

$(\theta-\theta_0)/\Delta\theta$

(a) experimental

Instrumental_JESR

JES effective NP stat. 2
JES effective NP modelling 3
JES effective NP modelling 2
JES effective NP modelling 1
JES effective NP mixed 2
JES pileup ρ topology
JES pileup pT term
JES pileup offset NPV
JES flavour response
JES flavour composition
JES η intercalibration total stat
JES η intercalibration non-closure pos η
JET EtaIntercalibration NonClosure 2018data
JES η intercalibration modelling
JES BJES
JET JER EffectiveNP 9
JET JER EffectiveNP 8
JET JER EffectiveNP 7
JET JER EffectiveNP 6
JET JER EffectiveNP 5
JET JER EffectiveNP 4
JET JER EffectiveNP 3
JET JER EffectiveNP 2
JET JER EffectiveNP 12restTerm
JET JER EffectiveNP 11
JET JER EffectiveNP 10
JET JER EffectiveNP 1
JET JER DataVsMC MC16
JER DataVsMC AFII

$(\theta-\theta_0)/\Delta\theta$

(b) JES/JER

Instrumental_FTAG

light-tag Eigenvar. 14
light-tag Eigenvar. 12
light-tag Eigenvar. 11
light-tag Eigenvar. 8
light-tag Eigenvar. 7
light-tag Eigenvar. 4
light-tag Eigenvar. 3
light-tag Eigenvar. 1
light-tag Eigenvar. 0
c-tag Eigenvar. 17
c-tag Eigenvar. 8
c-tag Eigenvar. 4
c-tag Eigenvar. 3
c-tag Eigenvar. 2
c-tag Eigenvar. 1
c-tag Eigenvar. 0
b-tag Eigenvar. 2
b-tag Eigenvar. 1
b-tag Eigenvar. 0

$(\theta-\theta_0)/\Delta\theta$

(c) flavor-tagging

Theory

$t\bar{t}$ ISR shape
$t\bar{t}Z$ μ_R and μ_F scales (env.)
$t\bar{t}Z$ XS
tt+Z ME+PS
tt+Z ISR
$t\bar{t}W$ μ_R and μ_F scales (env.)
$t\bar{t}W$ XS
tt+W ME+PS
tt+W ISR
$t\bar{t}H$ μ_R and μ_F scales (env.)
$t\bar{t}H$ XS
ttH PS & had.
ttH NLO ME
ttH ISR
ttH FSR
t-channel μ_R=0.5 and μ_F=0.5 scales
t-chan PS & hadronisation
t-chan ISR
t-chan NLO Gen
t-chan FSR
tZq μ_R and μ_F scales (env.)
tZq XS
tZq PS
tZq ISR
tW μ_R and μ_F scales (env.)
tW PS & hadronisation
tW NLO ME
tW ISR
tW FSR
tW diagram subtraction
tWZ μ_R and μ_F scales (env.)
tWZ XS
tWZ (alt. DR)
tWH μ_R and μ_F scales (env.)
tWH PS
single top XS
Minor bkgs XS
Diboson XS

$(\theta-\theta_0)/\Delta\theta$

(d) theory

Theory_ttbar

ttbar PS & had. shape
ttbar NLO gen. shape

$(\theta-\theta_0)/\Delta\theta$

(e) theory $t\bar{t}$

Figure H.3: Constraints of all NPs in the $1e/\mu + 2\tau_{\text{had}}$ region after the background fit in the CRs on data. The NPs are roughly grouped by their origin.

2 e/μ OS + 1 τ_{had}

Figure H.4: NFs determined from the fitin the $2e/\mu\text{OS} + 1\tau_{\text{had}}$ channel using the CRs only and fixing $\mu(tHq)$ to zero.

(a) CR $t\bar{t}$ pre-fit

(b) CR $t\bar{t}$ post-fit

(c) CR Z+jets pre-fit

(d) CR Z+jets post-fit

Figure H.5: Simulation to data agreement in the $2e/\mu\text{OS} + 1\tau_{\text{had}}$ channel before and after the fit to the background in the CRs.
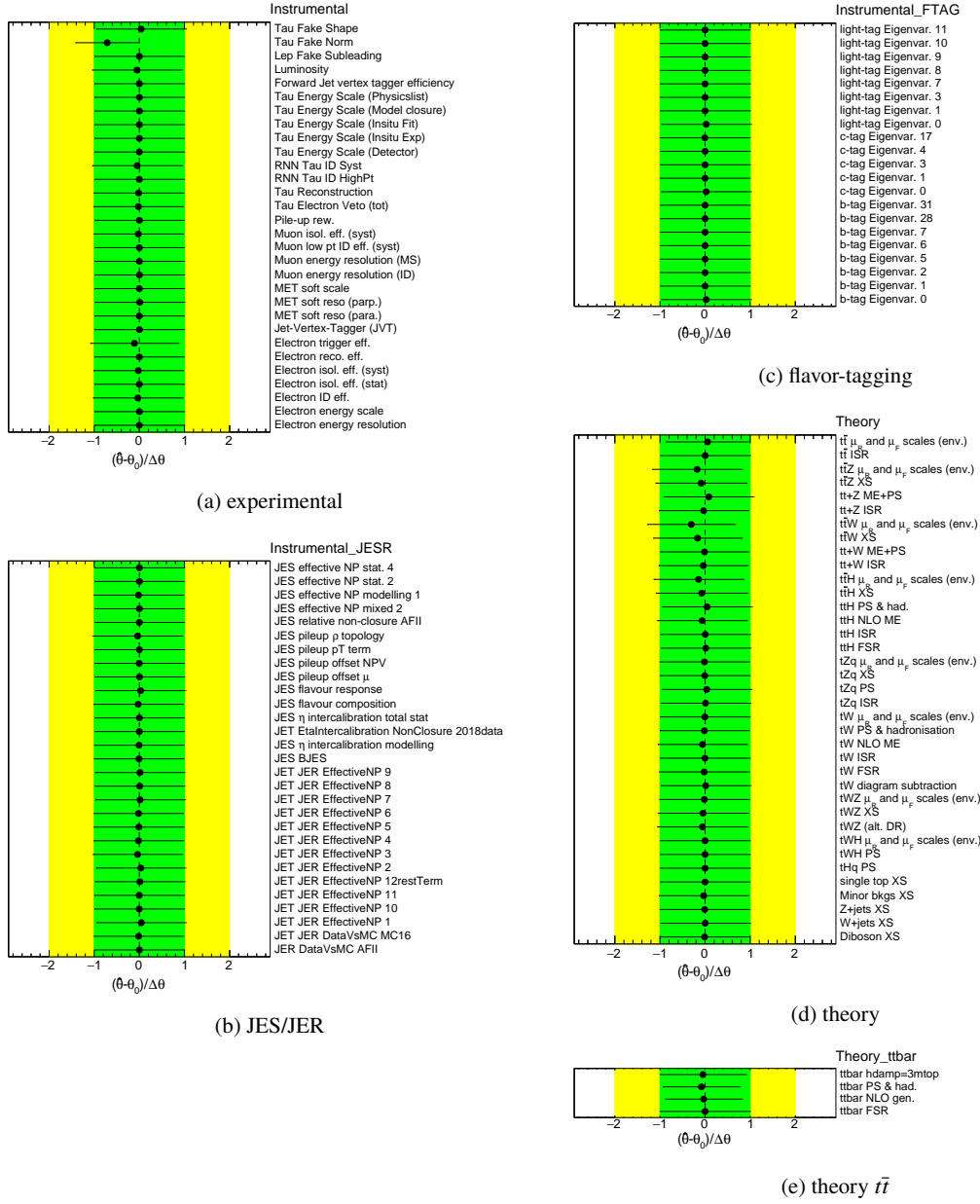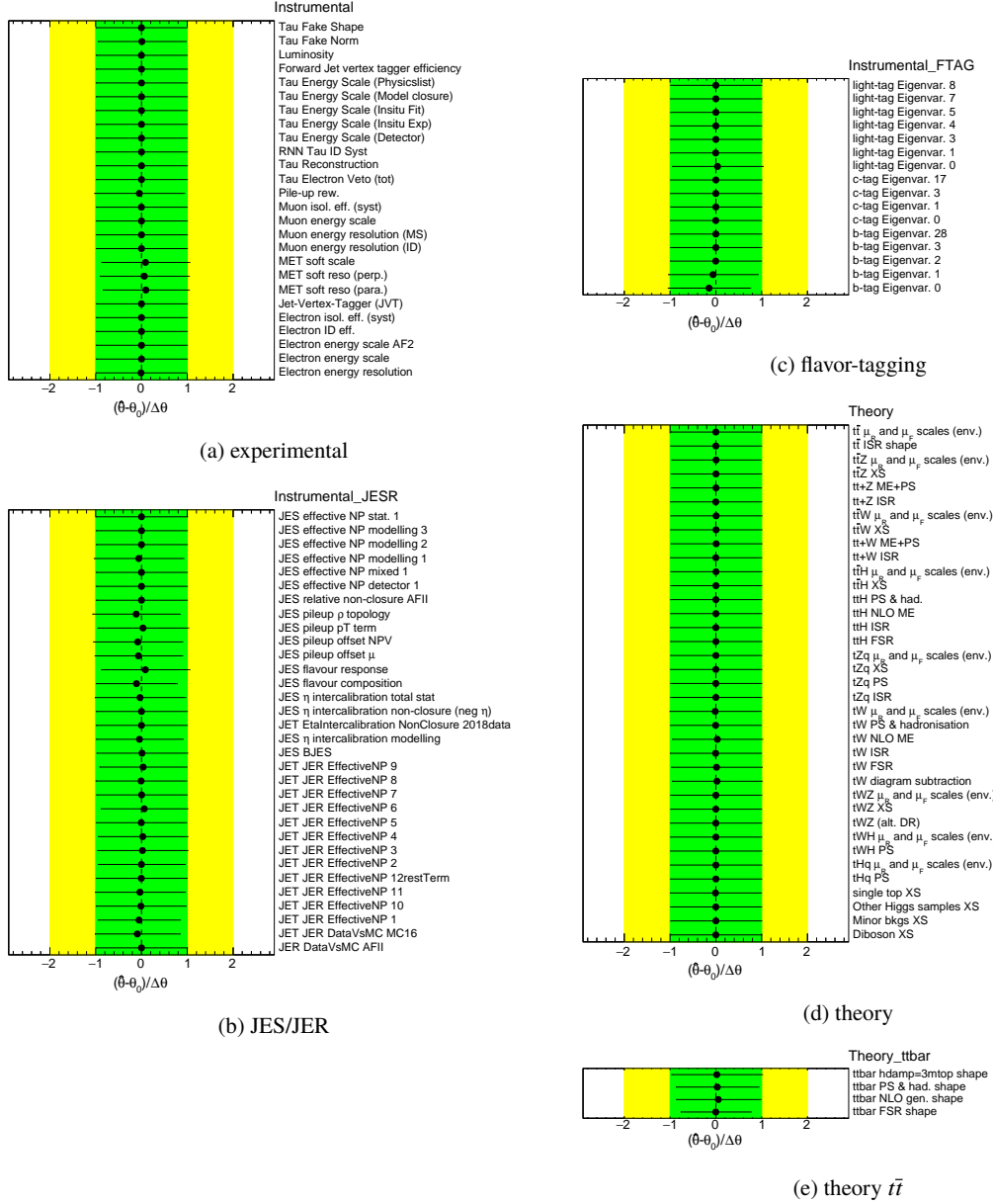
Instrumental

Tau Fake Shape
Tau Fake Norm
Luminosity
Forward Jet vertex tagger efficiency
Tau Energy Scale (Physicslist)
Tau Energy Scale (Model closure)
Tau Energy Scale (Insitu Fit)
Tau Energy Scale (Insitu Exp)
Tau Energy Scale (Detector)
RNN Tau ID Syst
Tau Reconstruction
Tau Electron Veto (tot)
Pile-up rew.
Muon isol. eff. (syst)
Muon energy scale
Muon energy resolution (MS)
Muon energy resolution (ID)
MET soft scale
MET soft reso (perp.)
MET soft reso (para.)
Jet-Vertex-Tagger (JVT)
Electron isol. eff. (syst)
Electron ID eff.
Electron energy scale AF2
Electron energy scale
Electron energy resolution

(a) experimental

Instrumental_FTAG

light-tag Eigenvar. 1
light-tag Eigenvar. 0
c-tag Eigenvar. 17
c-tag Eigenvar. 3
c-tag Eigenvar. 1
c-tag Eigenvar. 0
b-tag Eigenvar. 28
b-tag Eigenvar. 3
b-tag Eigenvar. 2
b-tag Eigenvar. 1
b-tag Eigenvar. 0

(c) flavor-tagging

Instrumental_JESR

JES effective NP stat. 1
JES effective NP modelling 3
JES effective NP modelling 2
JES effective NP modelling 1
JES effective NP mixed 1
JES effective NP detector 1
JES relative non-closure AFII
JES pileup ρ topology
JES pileup pT term
JES pileup offset NPV
JES pileup offset μ
JES flavour response
JES flavour composition
JES η intercalibration total stat
JES η intercalibration non-closure (neg η)
JET EtaIntercalibration NonClosure 2018data
JES η intercalibration modelling
JES BJES
JET JER EffectiveNP 9
JET JER EffectiveNP 8
JET JER EffectiveNP 7
JET JER EffectiveNP 6
JET JER EffectiveNP 5
JET JER EffectiveNP 4
JET JER EffectiveNP 3
JET JER EffectiveNP 2
JET JER EffectiveNP 12restTerm
JET JER EffectiveNP 11
JET JER EffectiveNP 10
JET JER EffectiveNP 1
JET JER DataVsMC MC16
JER DataVsMC AFII

(b) JES/JER

Theory

$t\bar{t}Z$ μ_R and μ_F scales (env.)
$t\bar{t}Z$ XS
tt+Z ME+PS
tt+Z ISR
$t\bar{t}W$ μ_R and μ_F scales (env.)
$t\bar{t}W$ XS
tt+W ME+PS
tt+W ISR
$t\bar{t}H$ μ_R and μ_F scales (env.)
$t\bar{t}H$ XS
ttH PS & had.
ttH NLO ME
ttH ISR
ttH FSR
tZq μ_R and μ_F scales (env.)
tZq XS
tZq PS
tZq ISR
tW μ_R and μ_F scales (env.)
tW PS & hadronisation
tW NLO ME
tW ISR
tW FSR
tW diagram subtraction
tWZ μ_R and μ_F scales (env.)
tWZ XS
tWZ (alt. DR)
tWH μ_R and μ_F scales (env.)
tWH PS
single top XS
Other Higgs samples XS
Minor bkgs XS
Diboson XS

(d) theory

Theory_ttbar

ttbar hdamp=3mtop shape
ttbar PS & had. shape
ttbar NLO gen. shape
ttbar FSR shape

(e) theory $t\bar{t}$

Figure H.6: Constraints of all NPs in the $2e/\mu$OS + $1\tau_{\text{had}}$ region after the background fit in the CRs on data. The NPs are roughly grouped by their origin.

2 e/μ SS + 1 τ_{had}

(a) CR pre-fit

(b) CR post-fit

Figure H.7: Simulation to data agreement in the $2e/\mu$SS + $1\tau_{\text{had}}$ channel before and after the fit to the background in the CRs.

(a) experimental

(b) JES/JER

(c) flavor-tagging

(d) theory

(e) theory $t\bar{t}$

Figure H.8: Constraints of all NPs in the $2e/\mu$SS $+ 1\tau_{\text{had}}$ region after the background fit in the CRs on data. The NPs are roughly grouped by their origin.

Background fit with data in the CRs and simulation in the SRs

The fits presented here are fitted under the background only hypothesis. Therefore, $\mu(tHq)$ is fixed to 0. Data is used in the CRs, Asimov data in the SR.

1 e/μ + 2 τ_{had}
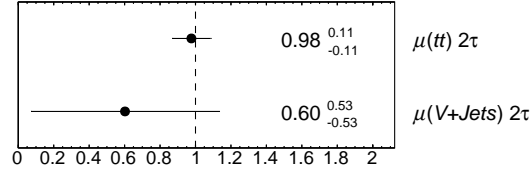
Figure H.9: NFs determined from the fit in the $1e/\mu + 2\tau_{\text{had}}$ channel using data in the CRs and fixing $\mu(tHq)$ to zero.

(a) CR $t\bar{t}$ pre-fit

(b) CR $t\bar{t}$ post-fit

(c) CR V+jets pre-fit

(d) CR V+jets post-fit
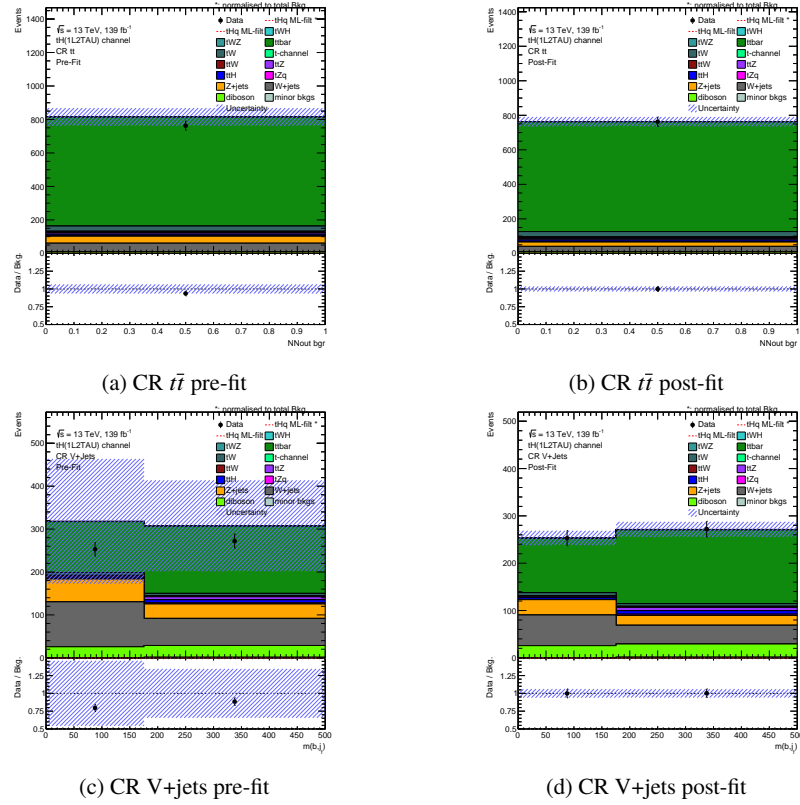
Figure H.10: Simulation to data agreement in the $1e/\mu + 2\tau_{\text{had}}$ channel before and after the fit to the background with data in the CRs.
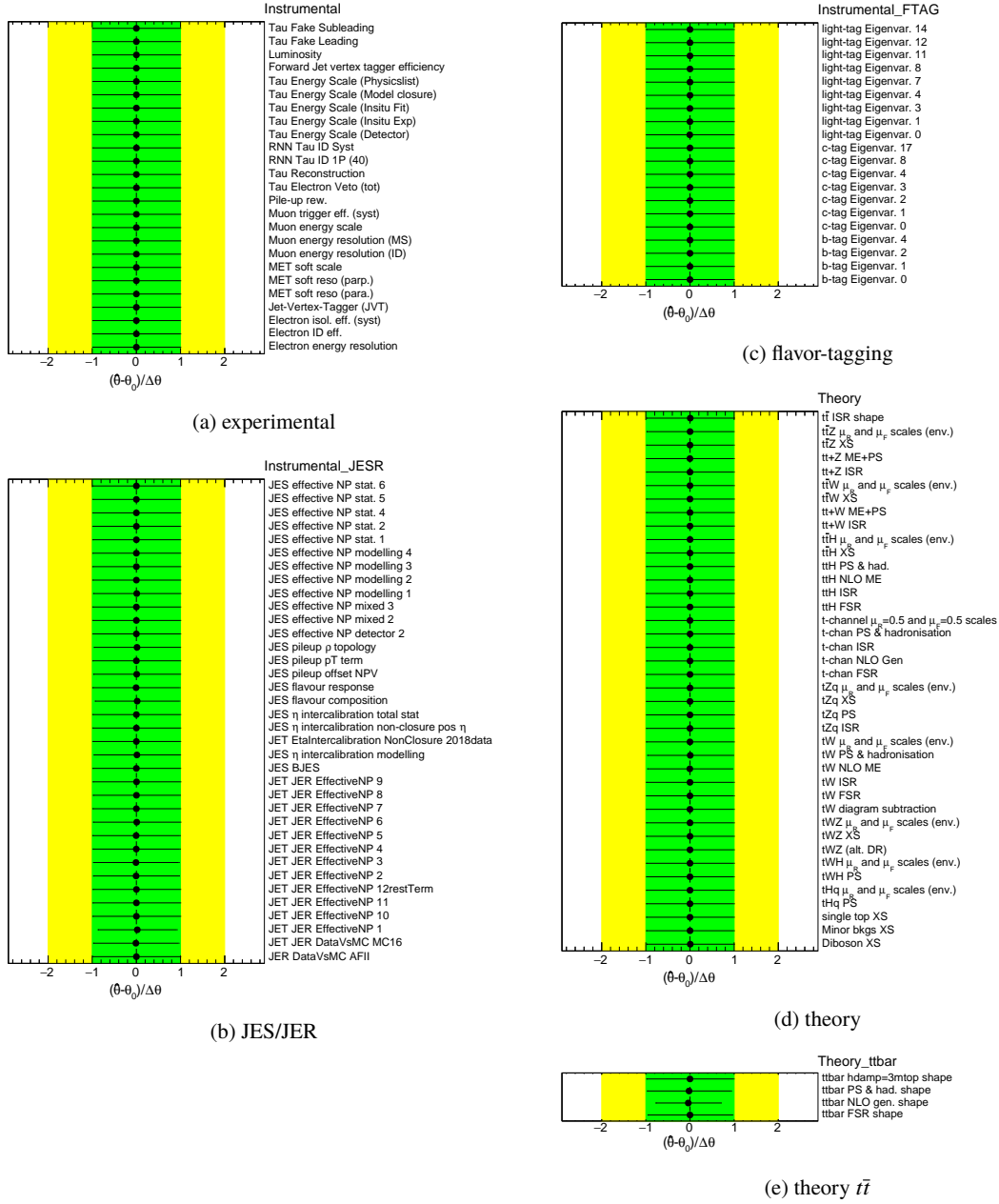
Figure H.11: Constraints of all NPs in the $1e/\mu + 2\tau_{\text{had}}$ region after the background fit on data in the CRs. The NPs are roughly grouped by their origin.

2 e/μ OS + 1 τ_{had}

179
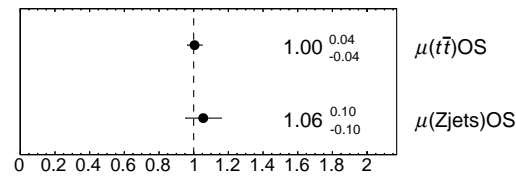
Figure H.12: NFs determined from the fit in the $2e/\mu\mathrm{OS} + 1\tau_{\mathrm{had}}$ channel using data in the CRs and fixing $\mu(tHq)$ to zero.

(a) CR $t\bar{t}$ pre-fit

(b) CR $t\bar{t}$ post-fit

(c) CR Z+jets pre-fit

(d) CR Z+jets post-fit

Figure H.13: Simulation to data agreement in the $2e/\mu\mathrm{OS} + 1\tau_{\mathrm{had}}$ channel before and after the fit to the background with data in the CRs.

(a) experimental

(b) JES/JER

(c) flavor-tagging

(d) theory

(e) theory $t\bar{t}$
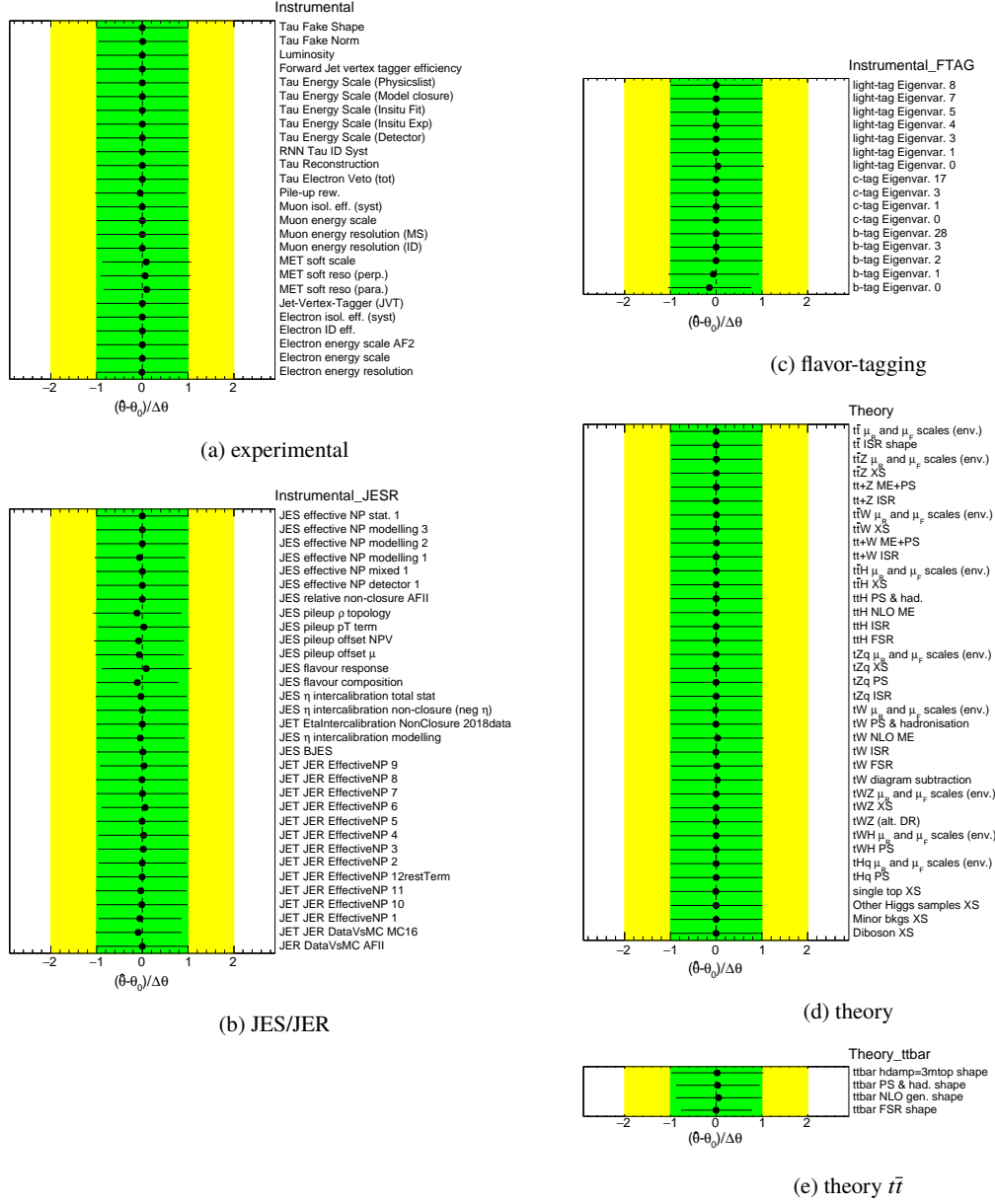
Figure H.14: Constraints of all NPs in the $2e/\mu$OS + $1\tau_{\text{had}}$ region after the background fit using data in the CRs. The NPs are roughly grouped by their origin.

2 e/μ SS + 1 τ_{had}

(a) CR pre-fit

(b) CR post-fit
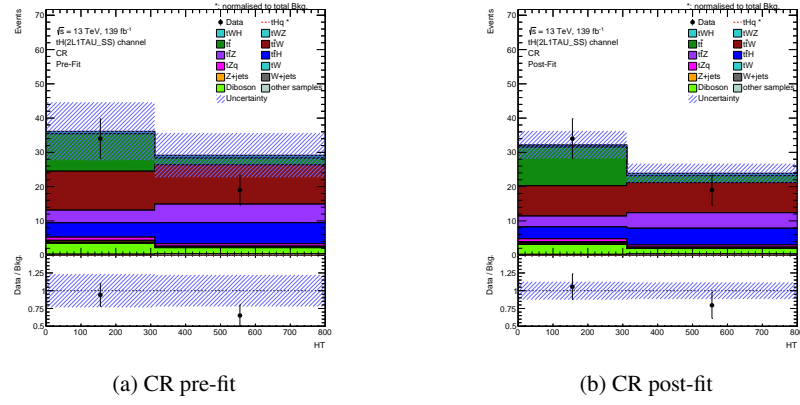
Figure H.15: Simulation to data agreement in the $2e/\mu$SS + $1\tau_{\text{had}}$ channel before and after the fit to the background with data in the CRs.

Instrumental

Tau Fake Shape
Tau Fake Norm
Lep Fake Subleading
Luminosity
Forward Jet vertex tagger efficiency
Tau Energy Scale (Physicslist)
Tau Energy Scale (Model closure)
Tau Energy Scale (Insitu Fit)
Tau Energy Scale (Insitu Exp)
Tau Energy Scale (Detector)
RNN Tau ID Syst
RNN Tau ID HighPt
Tau Reconstruction
Tau Electron Veto (tot)
Pile-up rew.
Muon isol. eff. (syst)
Muon low pt ID eff. (syst)
Muon energy resolution (MS)
Muon energy resolution (ID)
MET soft scale
MET soft reso (parp.)
MET soft reso (para.)
Jet-Vertex-Tagger (JVT)
Electron trigger eff.
Electron reco. eff.
Electron isol. eff. (syst)
Electron isol. eff. (stat)
Electron ID eff.
Electron energy scale
Electron energy resolution

(a) experimental

(b) JES/JER

(c) flavor-tagging

(d) theory

(e) theory $t\bar{t}$

Figure H.16: Constraints of all NPs in the $2e/\mu\mathrm{SS} + 1\tau_{\mathrm{had}}$ region after the background fit using data in the CRs. The NPs are roughly grouped by their origin.

Background fit with data in insensitive bins

These fits are performed similar to the final fit, but only in bins which are expected to not contain any signal. This is determined by a S/B per bin of 5% with respect to the inverted coupling hypothesis $tHq + tWH$. This strategy blinds the highest bins in the $1e/\mu + 2\tau_{\mathrm{had}}$ and $2e/\mu\mathrm{OS} + 1\tau_{\mathrm{had}}$ fit. In the $2e/\mu\mathrm{SS} + 1\tau_{\mathrm{had}}$ region this strategy blinds too many bins even in the CR. Therefore, this strategy is not performed in this

region.

1 e/μ **+ 2** τ_{had}

Figure H.17: NFs determined from the fit in the $1e/\mu + 2\tau_{\text{had}}$ channel using only insensitive bins.

Instrumental

Tau Fake Subleading
Tau Fake Leading
Luminosity
Forward Jet vertex tagger efficiency
Tau Energy Scale (Physicslist)
Tau Energy Scale (Model closure)
Tau Energy Scale (Insitu Fit)
Tau Energy Scale (Insitu Exp)
Tau Energy Scale (Detector)
RNN Tau ID Syst
RNN Tau ID 1P (40)
Tau Reconstruction
Tau Electron Veto (tot)
Pile-up rew.
Muon trigger eff. (syst)
Muon energy scale
Muon energy resolution (MS)
Muon energy resolution (ID)
MET soft scale
MET soft reso (parp.)
MET soft reso (para.)
Jet-Vertex-Tagger (JVT)
Electron isol. eff. (syst)
Electron ID eff.
Electron energy resolution

$(\theta - \theta_0)/\Delta\theta$

(a) experimental

Instrumental_JESR

JES effective NP stat. 6
JES effective NP stat. 5
JES effective NP stat. 4
JES effective NP stat. 2
JES effective NP stat. 1
JES effective NP modelling 4
JES effective NP modelling 3
JES effective NP modelling 2
JES effective NP modelling 1
JES effective NP mixed 3
JES effective NP mixed 2
JES effective NP detector 2
JES pileup ρ topology
JES pileup pT term
JES pileup offset NPV
JES flavour response
JES flavour composition
JES η intercalibration total stat
JES η intercalibration non-closure pos η
JET EtaIntercalibration NonClosure 2018data
JES η intercalibration modelling
JES BJES
JET JER EffectiveNP 9
JET JER EffectiveNP 8
JET JER EffectiveNP 7
JET JER EffectiveNP 6
JET JER EffectiveNP 5
JET JER EffectiveNP 4
JET JER EffectiveNP 3
JET JER EffectiveNP 2
JET JER EffectiveNP 12restTerm
JET JER EffectiveNP 11
JET JER EffectiveNP 10
JET JER EffectiveNP 1
JET JER DataVsMC MC16
JER DataVsMC AFII

$(\theta - \theta_0)/\Delta\theta$

(b) JES/JER

Instrumental_FTAG

light-tag Eigenvar. 14
light-tag Eigenvar. 12
light-tag Eigenvar. 11
light-tag Eigenvar. 8
light-tag Eigenvar. 7
light-tag Eigenvar. 4
light-tag Eigenvar. 3
light-tag Eigenvar. 1
light-tag Eigenvar. 0
c-tag Eigenvar. 17
c-tag Eigenvar. 8
c-tag Eigenvar. 4
c-tag Eigenvar. 3
c-tag Eigenvar. 2
c-tag Eigenvar. 1
c-tag Eigenvar. 0
b-tag Eigenvar. 4
b-tag Eigenvar. 2
b-tag Eigenvar. 1
b-tag Eigenvar. 0

$(\theta - \theta_0)/\Delta\theta$

(c) flavor-tagging

Theory

$t\bar{t}$ ISR shape
$t\bar{t}Z$ μ_R and μ_F scales (env.)
$t\bar{t}Z$ XS
tt+Z ME+PS
tt+Z ISR
$t\bar{t}W$ μ_R and μ_F scales (env.)
$t\bar{t}W$ XS
tt+W ME+PS
tt+W ISR
$t\bar{t}H$ μ_R and μ_F scales (env.)
ttH XS
ttH PS & had.
ttH NLO ME
ttH ISR
ttH FSR
t-channel μ_R=0.5 and μ_F=0.5 scales
t-chan PS & hadronisation
t-chan ISR
t-chan NLO Gen
t-chan FSR
tZq μ_R and μ_F scales (env.)
tZq XS
tZq PS
tZq ISR
tW μ_R and μ_F scales (env.)
tW PS & hadronisation
tW NLO ME
tW ISR
tW FSR
tW diagram subtraction
tWZ μ_R and μ_F scales (env.)
tWZ XS
tWZ (alt. DR)
tWH μ_R and μ_F scales (env.)
tWH PS
tHq μ_R and μ_F scales (env.)
tHq PS
single top XS
Minor bkgs XS
Diboson XS

$(\theta - \theta_0)/\Delta\theta$

(d) theory

Theory_ttbar

ttbar hdamp=3mtop shape
ttbar PS & had. shape
ttbar NLO gen. shape
ttbar FSR shape

$(\theta - \theta_0)/\Delta\theta$

(e) theory $t\bar{t}$

Figure H.18: Constraints of all NPs in the $1e/\mu + 2\tau_{had}$ region after the fit using only insensitive bins. The NPs are roughly grouped by their origin.

2 e/μ OS + 1 τ_{had}

Figure H.19: NFs determined from the fit in the $2e/\mu\mathrm{OS} + 1\tau_{\mathrm{had}}$ channel using only insensitive bins.

(a) experimental

(b) JES/JER

(c) flavor-tagging

(d) theory

(e) theory $t\bar{t}$

Figure H.20: Constraints of all NPs in the $2e/\mu\mathrm{OS} + 1\tau_{\mathrm{had}}$ region after the fit using only insensitive bins. The NPs are roughly grouped by their origin.

Additional studies on the fully unblinded fits

1 e/μ **+ 2** τ_{had}
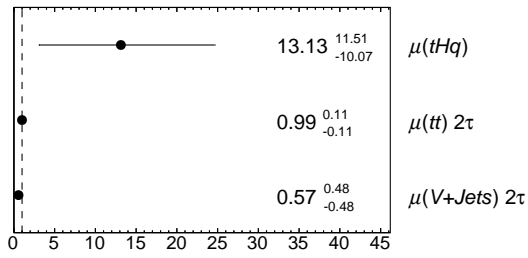
Figure I.1: NFs determined from the fit in the $1e/\mu + 2\tau_{\text{had}}$ channel using data.

(a) experimental

(b) JES/JER

(c) flavor-tagging

(d) theory

(e) theory $t\bar{t}$
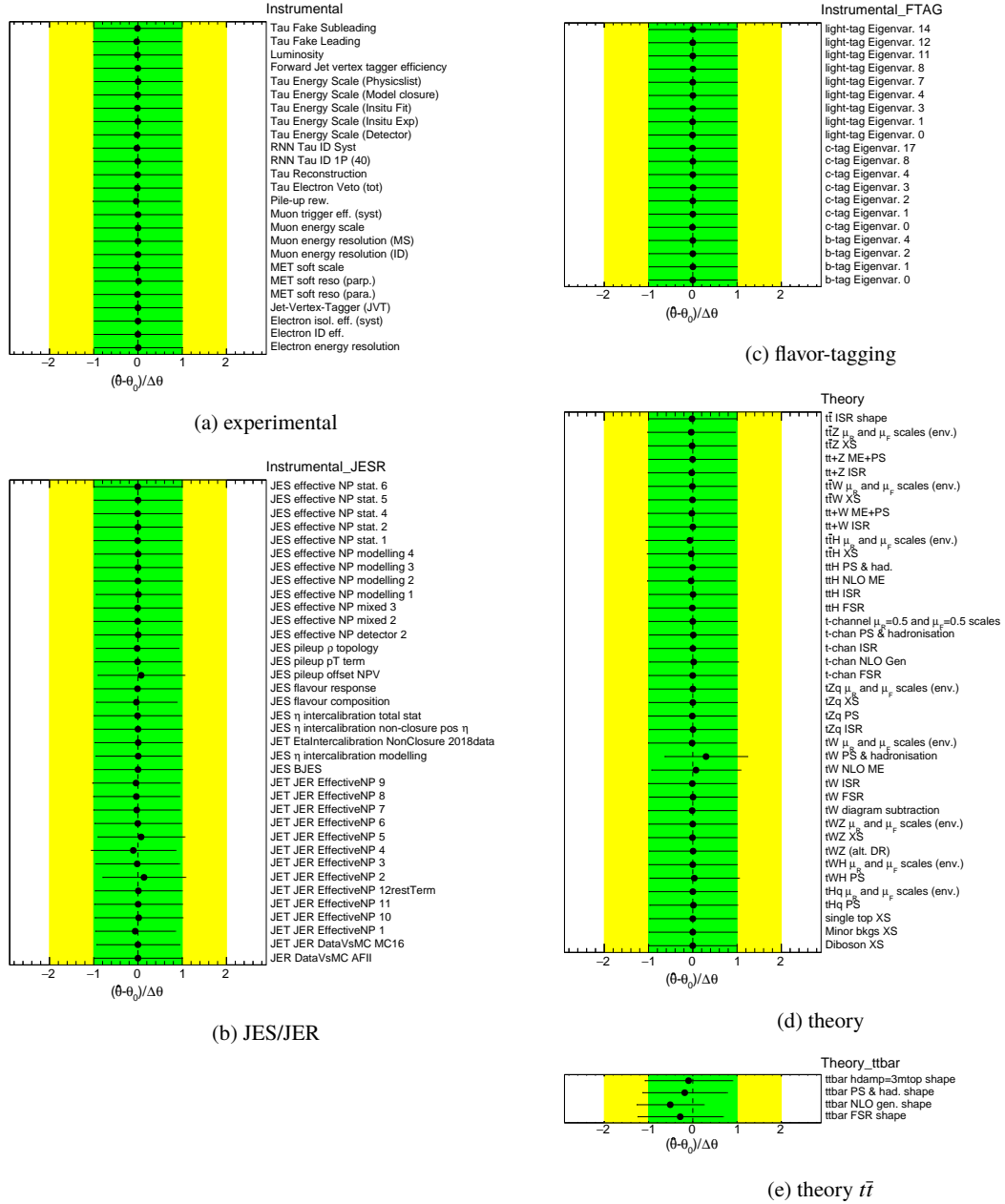
Figure I.2: Constraints of all NPs in the $1e/\mu + 2\tau_{\text{had}}$ region after the fit on data. The NPs are roughly grouped by their origin.

2 e/μ OS + 1 τ_{had}
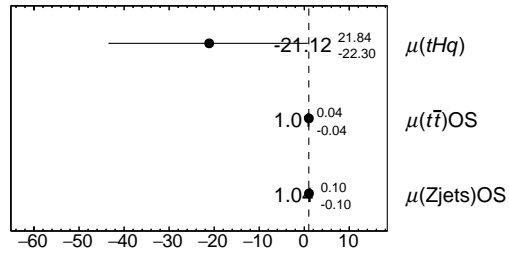
Figure I.3: NFs determined from the fit in the $2e/\mu\text{OS} + 1\tau_{\text{had}}$ channel using data.

(a) experimental

(b) JES/JER

(c) flavor-tagging

(d) theory

(e) theory $t\bar{t}$
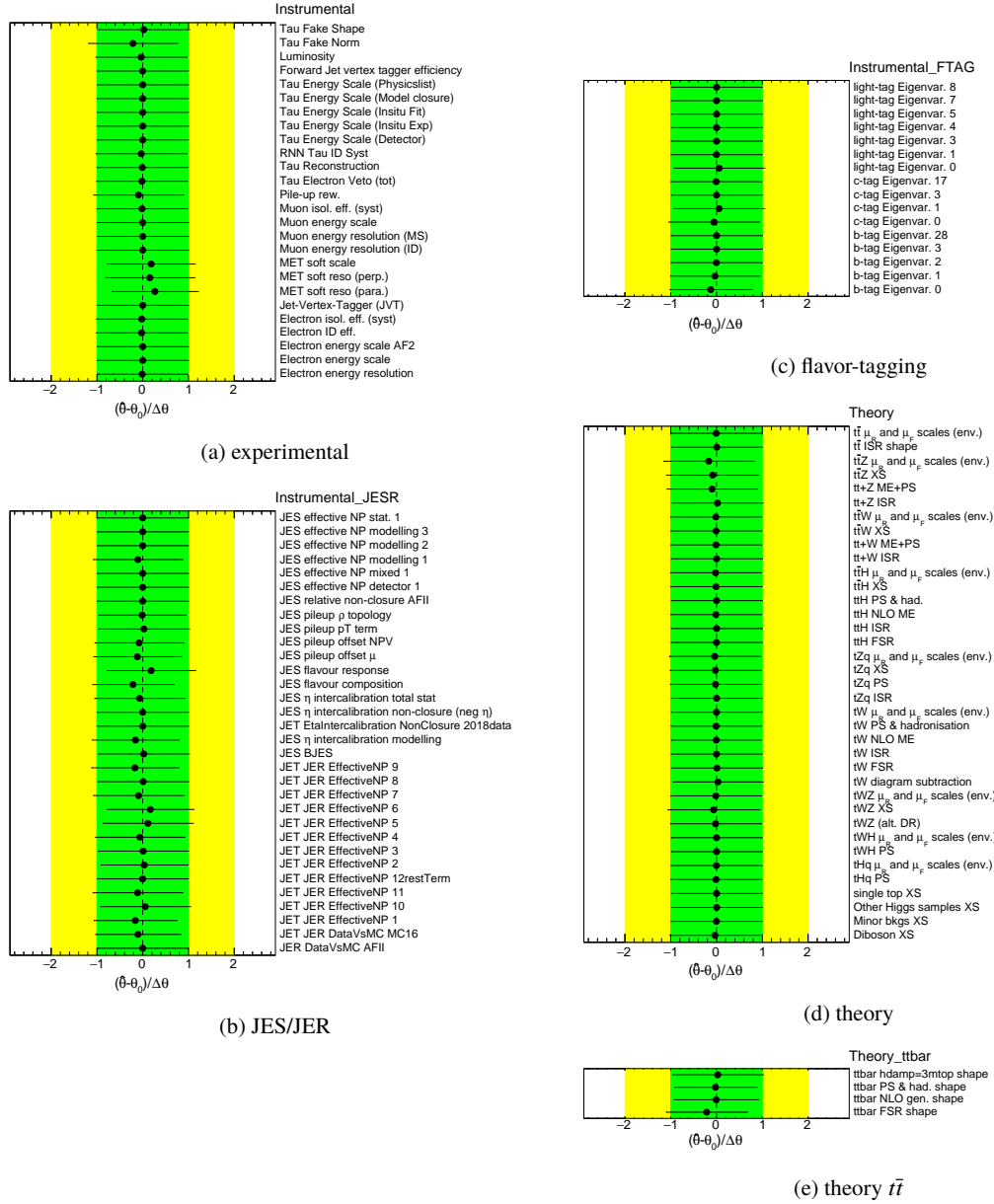
Figure I.4: Constraints of all NPs in the $2e/\mu\mathrm{OS} + 1\tau_{\mathrm{had}}$ region after the fit on data. The NPs are roughly grouped by their origin.

2 e/μ SS + 1 τ_{had}

(a) experimental

(b) JES/JER

(c) flavor-tagging

(d) theory

(e) theory $t\bar{t}$
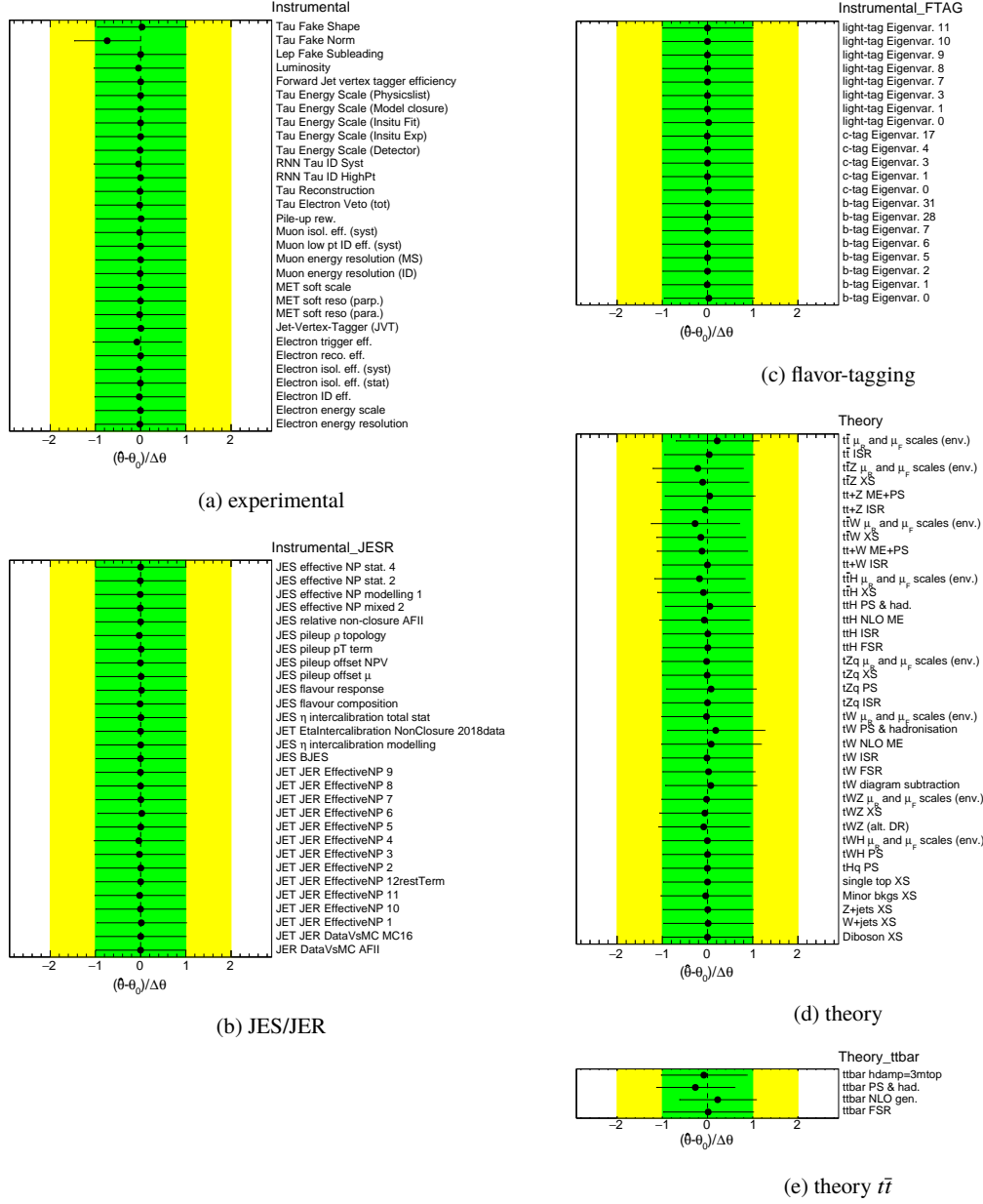
Figure I.5: Constraints of all NPs in the $2e/\mu$SS + $1\tau_{\text{had}}$ region after the data fit. The NPs are roughly grouped by their origin.

Combination

Figure I.6: Correlation matrix from all NPs in the combination fit with data. The correlations are comparable to the correlation in the individual fits. In total the highest correlations are shown between the NPs and the NFs associated with the reducible backgrounds. The highest correlations of $\mu(tHq)$ is observed with the τ_{had} fake related NPs and $t\bar{t}$ FSR (see section 7.5.2).
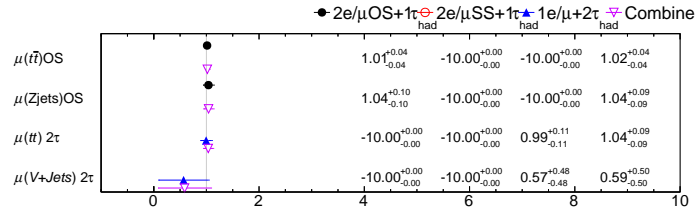
Figure I.7: NFs after the combination of all channels with data. Since the NFs are not combined between the channels the results are comparable with the results from the individual channel fits.

Limits and significance

Observed Limits and significance per channel

	limit	significance
$1e/\mu + 2\tau_{\text{had}}$	33.7	1.33
$2e/\mu \text{OS} + 1\tau_{\text{had}}$	43.1	-0.97
$2e/\mu \text{SS} + 1\tau_{\text{had}}$	26.9	-0.23
combination	20.0	0.76

Further inverted coupling studies

Asimov data

Under the inverted coupling scenario the cross-section of the *tHq* process is enhanced. For these fits the *tHq* and *tWH* samples are exchanged for samples which were generated under the inverted coupling hypothesis. The expected significance is 0.58σ.

(a) $1e/\mu + 2\tau_{\text{had}}$ no NPs

(b) $1e/\mu + 2\tau_{\text{had}}$ with NPs

(c) $2e/\mu\text{OS} + 1\tau_{\text{had}}$ no NPs

(d) $2e/\mu\text{OS} + 1\tau_{\text{had}}$ with NPs

(e) $2e/\mu\text{SS} + 1\tau_{\text{had}}$ no NPs

(f) $2e/\mu\text{SS} + 1\tau_{\text{had}}$ with NPs
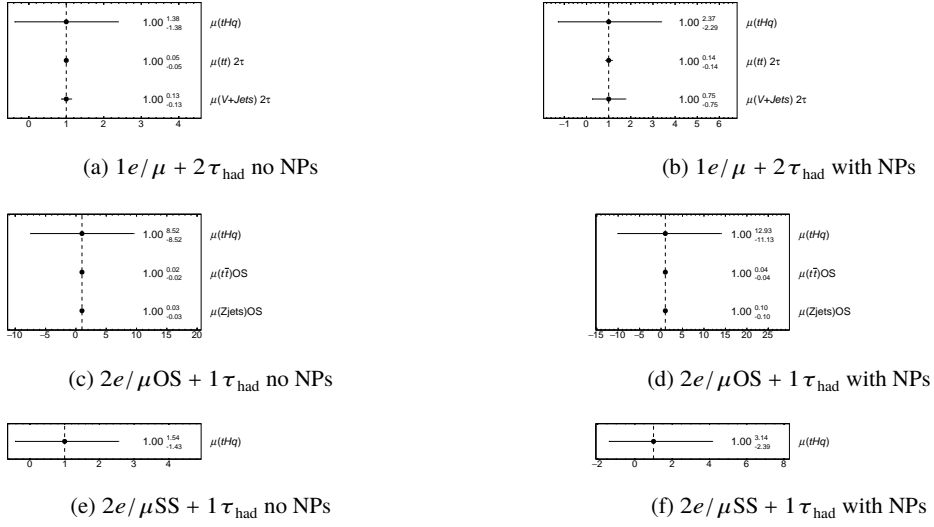
Figure J.1: NFs determined from the fit in all channels using Asimov data under the inverted coupling hypothesis. Both the fit results from a fit without NPs (left) and with NPs (right) are shown. $\Delta\mu(tHq)$ is smaller in comparison to the fit under the SM hypothesis. This can especially be seen in the left plots.

(a) experimental

(b) JES/JER

(c) flavor-tagging

(d) theory

(e) theory $t\bar{t}$
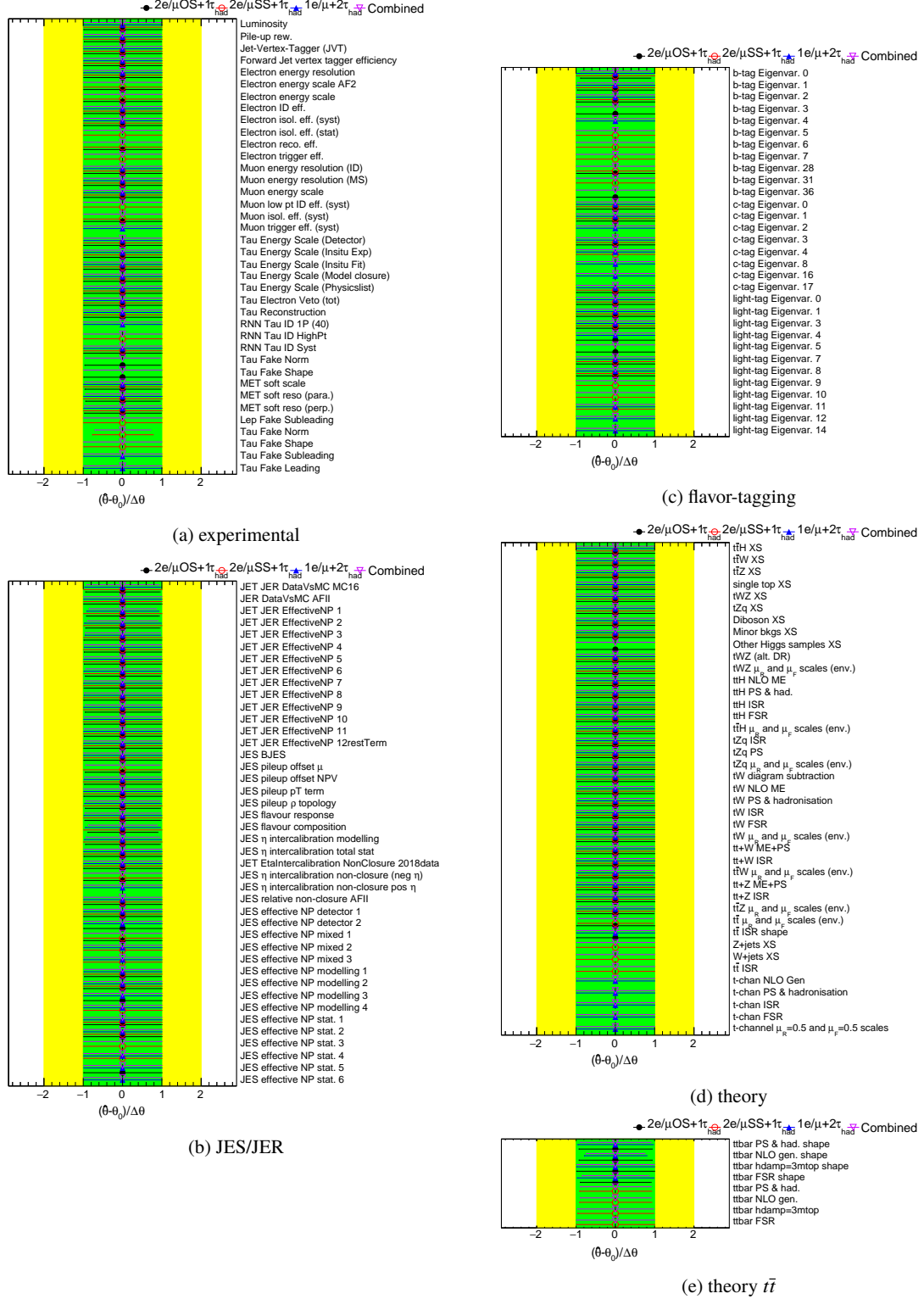
Figure J.2: Comparison of the constraints of all NPs in the combined fit on Asimov data under the inverted coupling hypothesis. The NPs are roughly grouped by their origin.

Data
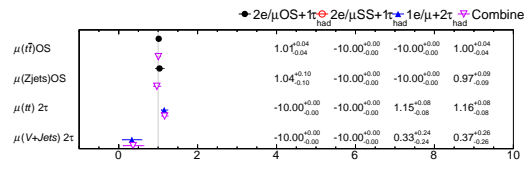
Figure J.3: NFs estimated from the combination of the fits on data under the inverted coupling hypothesis.

(a) experimental

(b) JES/JER

(c) flavor-tagging

(d) theory

(e) theory $t\bar{t}$
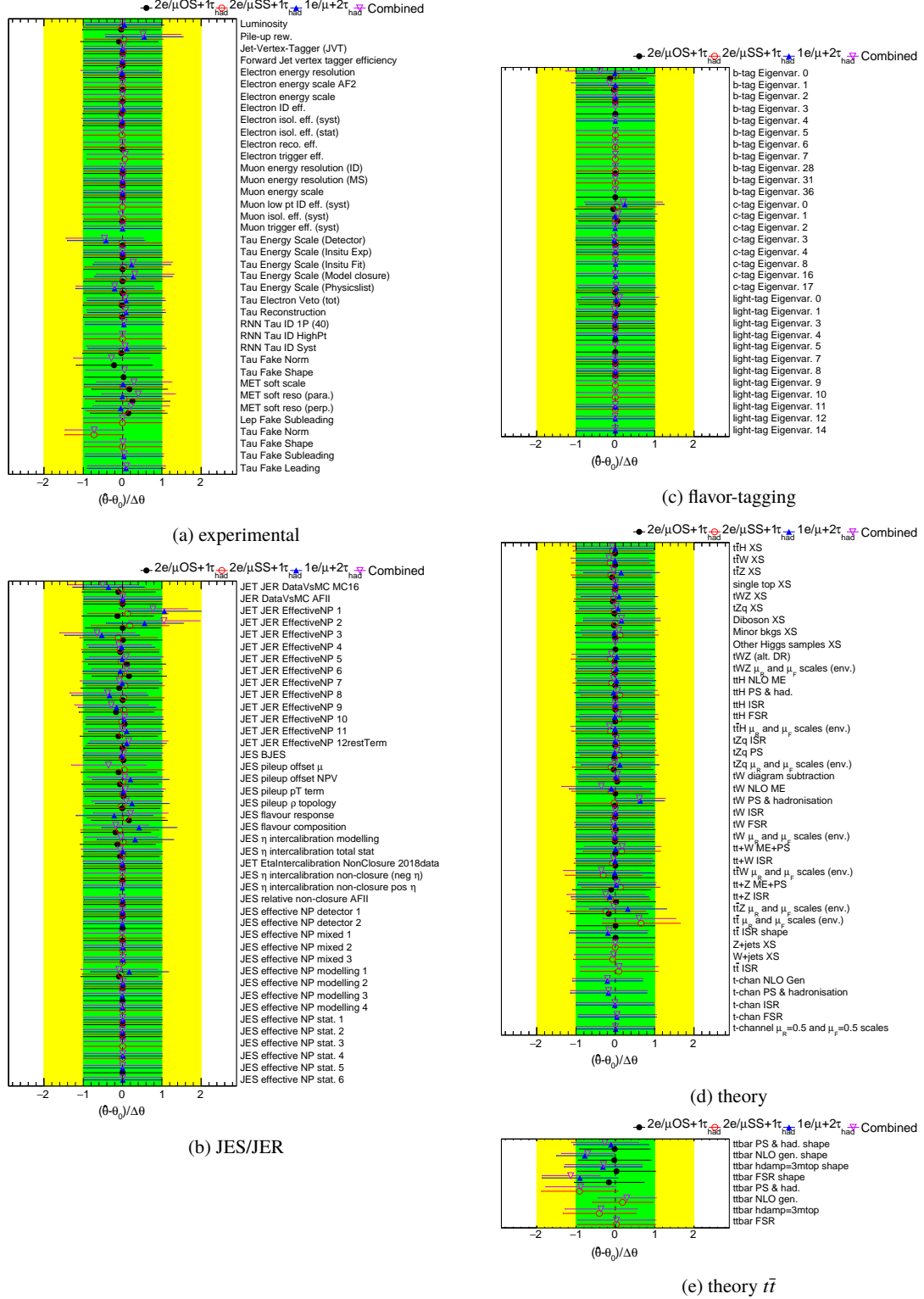
Figure J.4: Comparison of the constraints of all NPs in the combined fit on data under the inverted coupling hypothesis. The NPs are roughly grouped by their origin.

Combination of the two most sensitive fits

The $2e/\mu\text{OS} + 1\tau_{\text{had}}$ channel is not as sensitive as the other two channels. Therefore, the question arises if combining the $1e/\mu + 2\tau_{\text{had}}$ and $2e/\mu\text{SS} + 1\tau_{\text{had}}$ channel only leads to similar results. This would simplify the analysis and might prevent complications due to the $2e/\mu\text{OS} + 1\tau_{\text{had}}$ channel. In this section the combined fit on Asimov data and data in the $1e/\mu + 2\tau_{\text{had}}$ and $2e/\mu\text{SS} + 1\tau_{\text{had}}$ channel is shown and compared to the results in section 7.7.4.
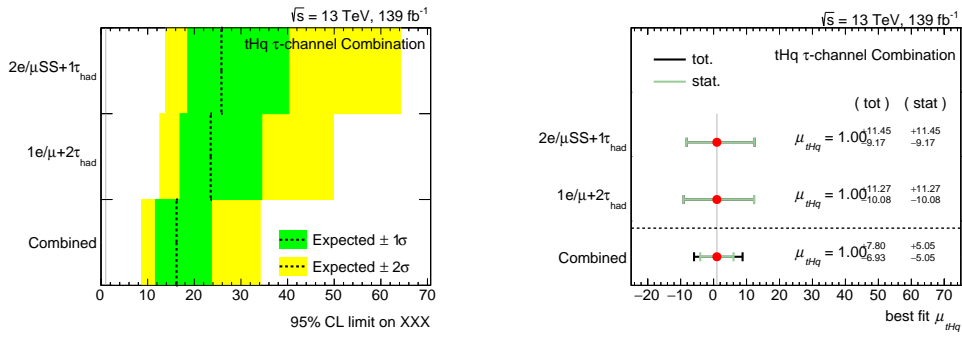
Figure K.1: Combination of the two most sensitive channels in the fit on Asimov data.

The results of the combination of only the $1e/\mu + 2\tau_{\text{had}}$ and $2e/\mu\text{SS} + 1\tau_{\text{had}}$ channel are comparable to the combination of all channels. But, the sensitivity is compromised. For completeness the result with data is shown, too.
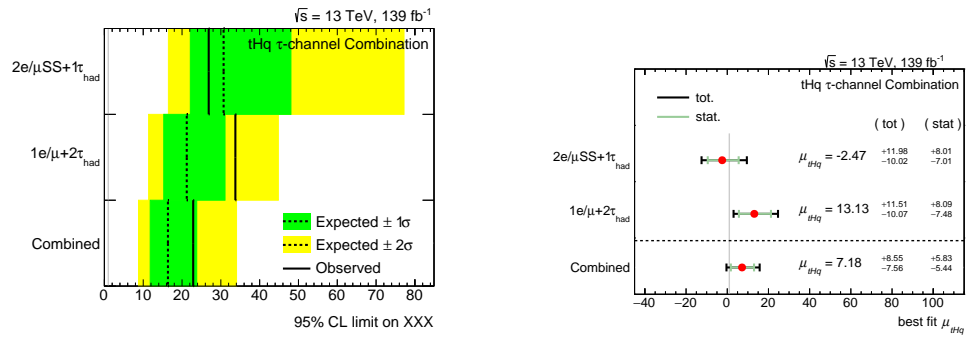
Figure K.2: Combination of the two most sensitive channels in the fit on data.

Bibliography

[1] S. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579; A. Salam, "Weak and Electromagnetic Interactions", *Elementary particle theory. Relativistic groups and analyticity. Proceedings of the Eighth Nobel Symposium*, ed. by N. Svartholm, Stockholm: Almquist & Wiksell, 1968 367; S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264, cit. on p. 3.

[2] G. 't Hooft and M. J. G. Veltman, *Regularization and Renormalization of Gauge Fields*, Nucl. Phys. B **44** (1972) 189 (cit. on p. 3).

[3] R. L. Workman et al., *Review of Particle Physics*, PTEP **2022** (2022) 083C01 (cit. on pp. 4, 5, 10–12, 14, 19).

[4] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Reading, USA: Addison-Wesley (1995) 842 p, Westview Press, 1995 (cit. on p. 5).

[5] T. Ohl, *Feynman Diagrams For Pedestrians*, 2019, URL: https://www.maria-laach.tp.nt.uni-siegen.de/downloads/files/2019/Ohl-2019.pdf (cit. on p. 5).

[6] L. Edelhäuser and A. K. Knochel, *Tutorium Quantenfeldtheorie, Was sie schon immer über QFT wissen wollten, aber bisher nicht zu fragen wagten*, ger, Lehrbuch, Berlin ; Heidelberg: Springer Spektrum, 2016, ISBN: 978-3-642-37675-7 (cit. on p. 5).

[7] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321 (cit. on p. 6).

[8] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132 (cit. on p. 6).

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

[10] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585 (cit. on p. 6).

[11] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. **145** (1966) 1156 (cit. on p. 6).

[12] T. W. B. Kibble, *Symmetry breaking in nonAbelian gauge theories*, Phys. Rev. **155** (1967) 1554 (cit. on p. 6).

[13] M. Pohl, *Particles, fields, space-time: from Thomson's electron to Higgs' boson*, Boca Raton: CRC Press, 2020, ISBN: 9780367347239, URL: https://bib-pubdb1.desy.de/record/448525 (cit. on p. 11).

[14]   ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv: `1207.7214 [hep-ex]` (cit. on p. 11).

[15]   CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30, arXiv: `1207.7235 [hep-ex]` (cit. on p. 11).

[16]   A. Nisati and V. Sharma, eds., *Discovery of the Higgs Boson*, Hackensack: World Scientific, 2017, ISBN: 978-981-4425-44-5, 978-981-4425-87-2 (cit. on p. 11).

[17]   ATLAS Collaboration, *A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery*, Nature **607** (2022) 52, arXiv: `2207.00092 [hep-ex]` (cit. on p. 11).

[18]   CMS Collaboration, *A portrait of the Higgs boson by the CMS experiment ten years after the discovery*, Nature **607** (2022) 60, arXiv: `2207.00043 [hep-ex]` (cit. on p. 11).

[19]   D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, (2016), arXiv: `1610.07922 [hep-ph]` (cit. on pp. 11, 12, 19).

[20]   X. Ji, *QCD Analysis of the Mass Structure of the Nucleon*, Phys. Rev. Lett. **74** (7 1995) 1071 (cit. on p. 12).

[21]   Y.-B. Yang et al., *Proton Mass Decomposition from the QCD Energy Momentum Tensor*, Phys. Rev. Lett. **121** (21 2018) 212001 (cit. on p. 12).

[22]   C. Lorcé, *On the hadron mass decomposition*, The European Physical Journal C **78** (2018) (cit. on p. 12).

[23]   J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43** (2016) 023001, arXiv: `1510.03865 [hep-ph]` (cit. on p. 13).

[24]   ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL: `https://cds.cern.ch/record/1966419` (cit. on p. 13).

[25]   S. Carrazza, S. Forte and J. Rojo, *Parton Distributions and Event Generators*, 2013, URL: `https://arxiv.org/abs/1311.5887` (cit. on p. 13).

[26]   R. D. Ball et al., *Parton distribution benchmarking with LHC data*, Journal of High Energy Physics **2013** (2013) (cit. on p. 13).

[27]   J. Rojo et al., *The PDF4LHC report on PDFs and LHC data: results from Run I and preparation for Run II*, Journal of Physics G: Nuclear and Particle Physics **42** (2015) 103103 (cit. on p. 13).

[28]   M. Tanabashi et al., *Review of Particle Physics*, Phys. Rev. D **98** (3 2018) 030001 (cit. on p. 13).

[29]   K. Kleinknecht and U. Uwer, "Symmetry Violations and Quark Flavour Physics", *Particle Physics Reference Library: Volume 1: Theory and Experiments*, ed. by H. Schopper, Cham: Springer International Publishing, 2020 519, ISBN: 978-3-030-38207-0, URL: `https://doi.org/10.1007/978-3-030-38207-0_9` (cit. on p. 13).

[30] E. Rabinovici, "Beyond the Standard Model",
Particle Physics Reference Library: Volume 1: Theory and Experiments, ed. by H. Schopper,
Cham: Springer International Publishing, 2020 455, ISBN: 978-3-030-38207-0,
URL: https://doi.org/10.1007/978-3-030-38207-0_8 (cit. on p. 14).

[31] B. Grzadkowski, M. Iskrzyński, M. Misiak and J. Rosiek,
Dimension-six terms in the Standard Model Lagrangian,
Journal of High Energy Physics **2010** (2010) (cit. on p. 14).

[32] N. Castro et al., *LHC EFT WG Report: Experimental Measurements and Observables*, 2022,
URL: https://arxiv.org/abs/2211.08353 (cit. on p. 14).

[33] J. J. Ethier et al.,
Combined SMEFT interpretation of Higgs, diboson, and top quark data from the LHC, 2021,
URL: https://arxiv.org/abs/2105.00006 (cit. on p. 14).

[34] J. Ellis, M. Madigan, K. Mimasu, V. Sanz and T. You,
Top, Higgs, diboson and electroweak fit to the Standard Model effective field theory,
Journal of High Energy Physics **2021** (2021) (cit. on p. 14).

[35] R. Frederix, E. Re and P. Torrielli, *Single-top t-channel hadroproduction in the four-flavour
scheme with POWHEG and aMC@NLO*, JHEP **09** (2012) 130,
arXiv: 1207.5391 [hep-ph] (cit. on pp. 16, 133).

[36] F. Demartin, F. Maltoni, K. Mawatari and M. Zaro,
Higgs production in association with a single top quark at the LHC,
The European physical journal. C, Particles and fields **75** (2015) (cit. on pp. 16, 18).

[37] D. Pagani, I. Tsinikos and E. Vryonidou,
NLO QCD+EW predictions for tHj and tZj production at the LHC, JHEP **08** (2020) 082,
arXiv: 2006.10086 [hep-ph] (cit. on pp. 16, 18, 19).

[38] E. L. Berger, J. Gao and H. X. Zhu, *Differential distributions for t-channel single top-quark
production and decay at next-to-next-to-leading order in QCD*,
Journal of High Energy Physics **2017** (2017) (cit. on p. 17).

[39] LHCTopWG, *NNLO+NNLL top-quark-pair cross sections*, 2015, URL:
https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO
(cit. on p. 18).

[40] G. Bordes and B. van Eijk, *On the associate production of a neutral intermediate-mass Higgs
boson with a single top quark at the LHC and SSC*, Physics Letters B **299** (1993) 315,
ISSN: 0370-2693 (cit. on p. 18).

[41] S. Biswas, E. Gabrielli and B. Mele,
Single top and Higgs associated production as a probe of the $Ht\bar{t}$ coupling sign at the LHC,
Journal of High Energy Physics **2013** (2013) (cit. on p. 18).

[42] M. Farina, C. Grojean, F. Maltoni, E. Salvioni and A. Thamm, *Lifting degeneracies in Higgs
couplings using single top production in association with a Higgs boson*, JHEP **05** (2013) 022,
arXiv: 1211.3736 (cit. on p. 18).

[43] C. Degrande, F. Maltoni, K. Mimasu, E. Vryonidou and C. Zhang,
Single-top associated production with a Z or H boson at the LHC: the SMEFT interpretation,
Journal of High Energy Physics **2018** (2018) (cit. on p. 18).

[44] S. Bhattacharya, S. Biswas, K. Pal and J. Wudka,
Associated production of Higgs and single top at the LHC in presence of the SMEFT operators,
2022, URL: https://arxiv.org/abs/2211.05450 (cit. on p. 18).

[45] CMS Collaboration, *Inclusive and differential cross section measurements of single top quark production in association with a Z boson in proton–proton collisions at $\sqrt{s}$ = 13 TeV*,
JHEP **02** (2021) 107, arXiv: 2111.02860 [hep-ex] (cit. on p. 19).

[46] ATLAS Collaboration, *Observation of the associated production of a top quark and a Z boson in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, JHEP **07** (2020) 124,
arXiv: 2002.07546 [hep-ex] (cit. on p. 19).

[47] S. Frixione, E. Laenen, P. Motylinski, C. White and B. R. Webber,
Single-top hadroproduction in association with a W boson, JHEP **07** (2008) 029,
arXiv: 0805.3067 [hep-ph] (cit. on p. 20).

[48] F. Demartin, B. Maier, F. Maltoni, K. Mawatari and M. Zaro,
tWH associated production at the LHC, Eur. Phys. J. C **77** (2017) 34,
arXiv: 1607.05862 [hep-ph] (cit. on p. 20).

[49] CMS Collaboration, *Search for the associated production of a Higgs boson with a single top quark in proton–proton collisions at $\sqrt{s}$ = 8 TeV*, JHEP **06** (2016) 177,
arXiv: 1509.08159 [hep-ex] (cit. on p. 24).

[50] CMS Collaboration, *Search for associated production of a Higgs boson and a single top quark in proton–proton collisions at $\sqrt{s}$ = 13 TeV*, Phys. Rev. D **99** (2019) 092005,
arXiv: 1811.09696 [hep-ex] (cit. on p. 24).

[51] CMS Collaboration,
Measurement of the Higgs boson production rate in association with top quarks in final states with electrons, muons, and hadronically decaying tau leptons at $\sqrt{s}$ = 13 TeV,
Eur. Phys. J. C **81** (2020) 378, arXiv: 2011.03652 [hep-ex] (cit. on p. 24).

[52] ATLAS Collaboration, *Search for H → γγ produced in association with top quarks and constraints on the Yukawa coupling between the top quark and the Higgs boson using data taken at 7 TeV and 8 TeV with the ATLAS detector*, Phys. Lett. B **740** (2015) 222,
arXiv: 1409.3122 [hep-ex] (cit. on p. 24).

[53] G. Aad and B. Abbott, *CP Properties of Higgs Boson Interactions with Top Quarks in the $t\bar{t}H$ and tH Processes Using H → γγ with the ATLAS Detector*,
Phys. Rev. Lett. **125** (6 2020) 061802,
URL: https://link.aps.org/doi/10.1103/PhysRevLett.125.061802
(cit. on p. 24).

[54] M. H. Kalos and P. A. Whitlock, *Monte Carlo methods*, eng, Weinheim: WILEY-VCH, 2008
XII, 203 S. ISBN: 3-527-40760-X, 978-3-527-40760-6 (cit. on p. 25).

[55] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv: `1405.0301 [hep-ph]` (cit. on p. 26).

[56] R. Frederix et al., *The automation of next-to-leading order electroweak calculations*, JHEP **07** (2018) 185, arXiv: `1804.10017 [hep-ph]` (cit. on p. 26), Erratum: R. Frederix et al., *Erratum to: The automation of next-to-leading order electroweak calculations*, Journal of High Energy Physics **2021** (2021).

[57] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv: `hep-ph/0409146` (cit. on p. 26).

[58] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv: `0709.2092 [hep-ph]` (cit. on p. 26).

[59] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, arXiv: `1002.2581 [hep-ph]` (cit. on pp. 26, 103).

[60] S. Frixione, E. Laenen, P. Motylinski and B. R. Webber, *Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations*, JHEP **04** (2007) 081, arXiv: `hep-ph/0702198` (cit. on p. 26).

[61] P. Artoisenet, R. Frederix, O. Mattelaer and R. Rietkerk, *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, JHEP **03** (2013) 015, arXiv: `1212.3460 [hep-ph]` (cit. on p. 26).

[62] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv: `1410.3012 [hep-ph]` (cit. on p. 26).

[63] B. Andersson, G. Gustafson, G. Ingelman and T. Sjöstrand, *Parton fragmentation and string dynamics*, Phys. Rept. **97** (1983) 31 (cit. on p. 26).

[64] T. Sjöstrand, *Jet fragmentation of multiparton configurations in a string framework*, Nucl. Phys. B **248** (1984) 469 (cit. on p. 26).

[65] ATLAS Collaboration, *Improvements in t$\bar{t}$ modelling using NLO+PS Monte Carlo generators for Run 2*, ATL-PHYS-PUB-2018-009, 2018, URL: `https://cds.cern.ch/record/2630327` (cit. on p. 26).

[66] ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling with Sherpa and MG5_aMC@NLO*, ATL-PHYS-PUB-2017-007, 2017, URL: `https://cds.cern.ch/record/2261938` (cit. on p. 26).

[67] ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling for Top2016*, ATL-PHYS-PUB-2016-020, 2016, URL: `https://cds.cern.ch/record/2216168` (cit. on p. 26).

[68] J. Bellm et al., *Herwig 7.1 Release Note*, (2017), arXiv: `1705.06919 [hep-ph]` (cit. on p. 26).

[69] S. Gieseke, F. Loshaj and P. Kirchgaesser,
      Soft and diffractive scattering with the cluster model in Herwig,
      The European Physical Journal C **77** (2016) (cit. on p. 26).

[70] E. Bothmann et al., *Event Generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034,
      arXiv: 1905.09127 [hep-ph] (cit. on p. 26).

[71] ATLAS Collaboration, *ATLAS simulation of boson plus jets processes in Run 2*,
      ATL-PHYS-PUB-2017-006, 2017, URL: https://cds.cern.ch/record/2261937
      (cit. on p. 26).

[72] ATLAS Collaboration, *Multi-Boson Simulation for* 13 *TeV ATLAS Analyses*,
      ATL-PHYS-PUB-2017-005, 2017, URL: https://cds.cern.ch/record/2261933
      (cit. on p. 26).

[73] GEANT4 Collaboration, S. Agostinelli et al., *Geant4 – a simulation toolkit*,
      Nucl. Instrum. Meth. A **506** (2003) 250 (cit. on p. 26).

[74] J. Allison et al., *Geant4 developments and applications*,
      IEEE Transactions on Nuclear Science **53** (2006) 270 (cit. on p. 26).

[75] J. Allison et al., *Recent developments in Geant4*,
      Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers,
      Detectors and Associated Equipment **835** (2016) 186 (cit. on p. 26).

[76] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823,
      arXiv: 1005.4568 [physics.ins-det] (cit. on p. 26).

[77] ATLAS Collaboration, *The simulation principle and performance of the ATLAS fast
      calorimeter simulation FastCaloSim*, ATL-PHYS-PUB-2010-013, 2010,
      URL: https://cds.cern.ch/record/1300517 (cit. on p. 26).

[78] W. Lukas, *Fast Simulation for ATLAS: Atlfast-II and ISF*,
      Journal of Physics: Conference Series **396** (2012) 022031 (cit. on p. 26).

[79] G. Bohm and G. Zech, *Einführung in Statistik und Messwertanalyse für Physiker*,
      Deutsches Elektronen-Synchrotron in der Helmholtz-Gemeinschaft, 2006,
      URL: http://www-library.desy.de/preparch/books/vstatmp.pdf
      (cit. on p. 27).

[80] J. H. Friedman, *Greedy function approximation: A gradient boosting machine.*,
      The Annals of Statistics **29** (2001) 1189 (cit. on p. 29).

[81] L.-G. Xia,
      Understanding the boosted decision tree methods with the weak-learner approximation, 2018,
      URL: https://arxiv.org/abs/1811.04822 (cit. on p. 29).

[82] F. Chollet et al., *Keras*, 2015, URL: https://github.com/fchollet/keras
      (cit. on p. 30).

[83] D. P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, 2014,
      URL: https://arxiv.org/abs/1412.6980 (cit. on p. 30).

[84] Cowan, Glen, Cranmer, Kyle, Gross, Eilam and Vitells, Ofer,
      Asymptotic formulae for likelihood-based tests of new physics, Eur. Phys. J. C **71** (2011) 1554
      (cit. on p. 31).

[85] F. a. James, *MINUIT: Function Minimization and Error Analysis Reference Manual*,
(1998), CERN Program Library Long Writeups,
URL: https://cds.cern.ch/record/2296388 (cit. on p. 32).

[86] K. Cranmer, *Practical Statistics for the LHC*, arXiv, 2015,
URL: https://arxiv.org/abs/1503.07622 (cit. on p. 32).

[87] M. Baak et al., *HistFitter software framework for statistical data analysis*,
The European Physical Journal C **75** (2015) (cit. on pp. 33, 34).

[88] A. L. Read, *Modified frequentist analysis of search results (the CL_s method)*, 2000,
URL: https://cds.cern.ch/record/451614 (cit. on p. 33).

[89] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, arXiv (2003) (cit. on p. 34).

[90] C. Grupen, *Astroparticle Physics*, Undergraduate Texts in Physics,
Springer International Publishing, 2020, ISBN: 9783030273392,
URL: https://books.google.de/books?id=gKbMDwAAQBAJ (cit. on p. 35).

[91] O. S. Brüning et al., *LHC Design Report, Volume 2*, CERN Yellow Reports: Monographs, 2004,
URL: https://cds.cern.ch/record/815187 (cit. on p. 36).

[92] O. S. Brüning et al., *LHC Design Report, Volume 1*, CERN Yellow Reports: Monographs,
Geneva: CERN, 2004, URL: https://cds.cern.ch/record/782076 (cit. on p. 36).

[93] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001 (cit. on p. 36).

[94] J. Billan et al., "Magnetic performance of the LEP bending magnets", 1989 1148,
URL: http://accelconf.web.cern.ch/p89/PDF/PAC1989_1148.PDF
(cit. on p. 36).

[95] J. Wenninger, *Operation and Configuration of the LHC in Run 2*, (2019),
URL: https://cds.cern.ch/record/2668326 (cit. on pp. 36, 37).

[96] K. Schindl, *The Injector Chain for the LHC; rev. version*, (1999) 7 p,
URL: http://cds.cern.ch/record/384396 (cit. on p. 37).

[97] R. Scrivens et al., *Overview of the status and developments on primary ion sources at CERN**,
(2011) 4 p, URL: https://cds.cern.ch/record/1382102 (cit. on p. 37).

[98] B. H. Wolf, *Handbook of ion sources*, Boca Raton, FL: CRC Press, 1995, ISBN: 978-0849325021,
URL: https://cds.cern.ch/record/306831 (cit. on p. 37).

[99] C. E. Hill, A. M. Lombardi, W. Pirkl, E. Tanke and M. Vretenar,
Performance of the CERN Linac 2 with a high intensity proton RFQ, (1994) 4 p,
URL: http://cds.cern.ch/record/276894 (cit. on p. 37).

[100] M. Weiss, *The RFQ2 complex: the future injector to CERN Linac 2*, (1992) 4 p,
URL: https://cds.cern.ch/record/238556 (cit. on p. 37).

[101] D. J. Warner, *Project study for a new 50 MeV linear accelerator for the C. P. S*, CERN, 1973,
URL: http://cds.cern.ch/record/414071 (cit. on p. 37).

[102] E. Boltezar et al., "THE NEW CERN 50-MEV LINAC.",
10th International Linear Accelerator Conference, 1980 S2,
URL: https://accelconf.web.cern.ch/l79/papers/s2-2.pdf
(cit. on p. 37).

[103] S. Ramberger et al., "Drift Tube Linac Design and Prototyping for the CERN Linac4", 2008 184,
URL: https://accelconf.web.cern.ch/LINAC08/papers/mop049.pdf
(cit. on p. 37).

[104] R. Cappi et al., *Beams in the PS complex during the LHC era; rev. version*, CERN, 1993,
URL: https://cds.cern.ch/record/252138 (cit. on p. 38).

[105] M. Benedikt, P. Collier, V. Mertens, J. Poole and K. Schindl, *LHC Design Report, Volume 3*,
CERN Yellow Reports: Monographs, 2004,
URL: https://cds.cern.ch/record/823808 (cit. on p. 38).

[106] H. Bartosik et al., *Injectors Beam Performance Evolution during Run 2*, (2019) 51,
URL: https://cds.cern.ch/record/2750276 (cit. on p. 38).

[107] S. Myers and H. E. Schopper, *Particle Physics Reference Library Volume 3: Accelerators and Colliders: Volume 3: Accelerators and Colliders*, 2020, ISBN: 978-3-030-34244-9 (cit. on p. 38).

[108] R. Schmidt,
"Machine Protection and Interlock Systems for Circular Machines - Example for LHC",
2014 Joint International Accelerator School: Beam Loss and Accelerator Protection: Newport Beach, CA, USA 5 - 14 Nov 2014, 2016 319, arXiv: 1608.03087 (cit. on pp. 38, 39).

[109] E. Mobs, *The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018*, (2018), General Photo,
URL: https://cds.cern.ch/record/2636343 (cit. on p. 39).

[110] ATLAS Collaboration,
ATLAS Detector and Physics Performance: Technical Design Report, Volume 1,
ATLAS-TDR-14; CERN-LHCC-99-014, 1999,
URL: https://cds.cern.ch/record/391176 (cit. on p. 41).

[111] CERN AC, "Layout of ATLAS. Dessin representant le detecteur ATLAS", 1998,
URL: https://cds.cern.ch/record/39038 (cit. on p. 41).

[112] ATLAS Collaboration, *ATLAS Run 2 Detector Status*, 2018,
URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/
ApprovedPlotsATLASDetector#ATLAS_Run_2_Detector_Status_AN1
(cit. on p. 42).

[113] A. L. Rosa and on behalf of ATLAS collaboration,
The ATLAS Insertable B-Layer: from construction to operation,
Journal of Instrumentation **11** (2016) C12036 (cit. on p. 41).

[114] G. Aad et al., *ATLAS pixel detector electronics and sensors*,
Journal of Instrumentation **3** (2008) P07007 (cit. on p. 41).

[115] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*,
ATLAS-TDR-19; CERN-LHCC-2010-013, 2010,
URL: https://cds.cern.ch/record/1291633 (cit. on p. 41).

[116] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*,
JINST **3** (2008) S08003 (cit. on pp. 42–44, 51).

[117] H. Kolanoski and N. Wermes, *Teilchendetektoren: Grundlagen und Anwendungen*,
Springer, 2016, ISBN: 978-3-662-45349-0, 978-3-662-45350-6 (cit. on pp. 42–44).

[118]   J. Pequenao, "Computer generated image of the ATLAS inner detector", 2008,
         URL: https://cds.cern.ch/record/1095926 (cit. on p. 43).

[119]   J. Pequenao, "Computer Generated image of the ATLAS calorimeter", 2008,
         URL: https://cds.cern.ch/record/1095927 (cit. on p. 45).

[120]   ATLAS Collaboration, *ATLAS Muon Spectrometer: Technical Design Report*,
         ATLAS-TDR-10; CERN-LHCC-97-022, CERN, 1997,
         URL: https://cds.cern.ch/record/331068 (cit. on p. 46).

[121]   ATLAS Collaboration, *Letter of Intent for the Phase-I Upgrade of the ATLAS Experiment*,
         CERN-LHCC-2011-012, LHCC-I-020, 2011,
         URL: http://cds.cern.ch/record/1402470 (cit. on p. 46).

[122]   J. Pequenao, "Computer generated image of the ATLAS Muons subsystem", 2008,
         URL: https://cds.cern.ch/record/1095929 (cit. on p. 46).

[123]   ATLAS Collaboration,
         ATLAS Magnet System: Magnet Project Technical Design Report, Volume 1,
         ATLAS-TDR-6; CERN-LHCC-97-018, 1997,
         URL: https://cds.cern.ch/record/338080 (cit. on p. 46).

[124]   A. M. Rodriguez Vera and J. Antunes Pequenao, "ATLAS Detector Magnet System",
         General Photo, 2021, URL: https://cds.cern.ch/record/2770604 (cit. on p. 47).

[125]   ATLAS Collaboration, *Luminosity determination for low-pileup datasets at $\sqrt{s}$ = 5 and 13 TeV
         using the ATLAS detector at the LHC*, ATLAS-CONF-2020-023, 2020,
         URL: https://cds.cern.ch/record/2725195 (cit. on p. 47).

[126]   ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s}$ = 13 TeV using the
         ATLAS detector at the LHC*, (2022), arXiv: 2212.09379 [hep-ex] (cit. on pp. 47, 103).

[127]   G. Avoni et al.,
         The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS,
         Journal of Instrumentation 13 (2018) P07017 (cit. on p. 47).

[128]   M. Trzebinski, *ATLAS Forward Proton Detectors*, 2017,
         URL: https://twiki.cern.ch/twiki/pub/Atlas/AFP_Figures/
         Trzebinski_AFP_CERN_seminar.pdf (cit. on p. 48).

[129]   ATLAS Collaboration, "ATLAS Forward Detectors", General Photo, 2018,
         URL: https://cds.cern.ch/record/2627582 (cit. on p. 48).

[130]   I. Bird et al., *Update of the Computing Models of the WLCG and the LHC Experiments*,
         Technical design report. LCG ; 2, 2014,
         URL: https://cds.cern.ch/record/1695401 (cit. on pp. 48, 50).

[131]   ATLAS Collaboration, *Operation of the ATLAS trigger system in Run 2*,
         JINST 15 (2020) P10004, arXiv: 2007.12539 [hep-ex] (cit. on pp. 48, 49).

[132]   ATLAS Collaboration,
         ATLAS data quality operations and performance for 2015–2018 data-taking,
         JINST 15 (2020) P04003, arXiv: 1911.04632 [physics.ins-det] (cit. on pp. 49, 64).

[133] A. Borga et al., *Evolution of the ReadOut System of the ATLAS experiment*, CERN, 2014,
URL: https://cds.cern.ch/record/1710776 (cit. on p. 49).

[134] G. Anders et al., *The upgrade of the ATLAS Level-1 Central Trigger Processor*,
Journal of Instrumentation **8** (2013) C01049 (cit. on p. 49).

[135] The ATLAS TDAQ Collaboration, *The ATLAS Data Acquisition and High Level Trigger system*,
Journal of Instrumentation **11** (2016) P06008 (cit. on p. 50).

[136] ATLAS Collaboration, *2015 start-up trigger menu and initial performance assessment of the ATLAS trigger using Run-2 data*, ATL-DAQ-PUB-2016-001, 2016,
URL: https://cds.cern.ch/record/2136007 (cit. on p. 50).

[137] ATLAS Collaboration, *Trigger Menu in 2016*, ATL-DAQ-PUB-2017-001, 2017,
URL: https://cds.cern.ch/record/2242069 (cit. on p. 50).

[138] ATLAS Collaboration, *Trigger Menu in 2017*, ATL-DAQ-PUB-2018-002, 2018,
URL: https://cds.cern.ch/record/2625986 (cit. on p. 50).

[139] ATLAS Collaboration, *Trigger Menu in 2018*, ATL-DAQ-PUB-2019-001, 2019,
URL: https://cds.cern.ch/record/2693402 (cit. on p. 50).

[140] *Preparing ATLAS Distributed Computing for LHC Run 2*, ATLAS Collaboration, 2014,
URL: https://cds.cern.ch/record/1669858 (cit. on p. 50).

[141] J. Elmsheuser and A. Di Girolamo, *Overview of the ATLAS distributed computing system*,
ATLAS Collaboration, 2018, URL: https://cds.cern.ch/record/2644515
(cit. on p. 50).

[142] ATLAS Collaboration, *ATLAS Computing: Technical Design Report*,
ATLAS-TDR-17; CERN-LHCC-2005-022, 2005,
URL: https://cds.cern.ch/record/837738 (cit. on p. 50).

[143] I. Bird, *Computing for the Large Hadron Collider*, Ann. Rev. Nucl. Part. Sci. **61** (2011) 99
(cit. on p. 50).

[144] M. Elsing, L. Goossens, A. Nairz and G. Negri,
The ATLAS Tier-0: Overview and operational experience,
Journal of Physics: Conference Series **219** (2010) 072011 (cit. on p. 50).

[145] ATLAS Collaboration,
ATLAS Detector and Physics Performance: Technical Design Report, Volume 2,
ATLAS-TDR-15; CERN-LHCC-99-015, 1999,
URL: https://cds.cern.ch/record/391177 (cit. on p. 51).

[146] ATLAS Collaboration, *Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2*, Eur. Phys. J. C **77** (2017) 673, arXiv: 1704.07983 [hep-ex]
(cit. on p. 51).

[147] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*,
Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **262** (1987) 444 (cit. on p. 51).

[148] T. Cornelissen et al., *The new ATLAS track reconstruction (NEWT)*, Journal of Physics: Conference Series **119** (2008) 032014, on behalf of the ATLAS Inner Detector software group
(cit. on p. 51).

[149] ATLAS Collaboration,
       A neural network clustering algorithm for the ATLAS silicon pixel detector,
       JINST **9** (2014) P09009, arXiv: 1406.7690 [hep-ex] (cit. on p. 51).

[150] ATLAS Collaboration,
       Early Inner Detector Tracking Performance in the 2015 Data at $\sqrt{s}$ = 13 TeV,
       ATL-PHYS-PUB-2015-051, 2015, URL: https://cds.cern.ch/record/2110140
       (cit. on p. 52).

[151] ATLAS Collaboration,
       Vertex Reconstruction Performance of the ATLAS Detector at $\sqrt{s}$ = 13 TeV,
       ATL-PHYS-PUB-2015-026, 2015, URL: https://cds.cern.ch/record/2037717
       (cit. on p. 52).

[152] F. Meloni, *Primary vertex reconstruction with the ATLAS detector*,
       Journal of Instrumentation **11** (2016) C12060, for the ATLAS Collaboration,
       URL: https://dx.doi.org/10.1088/1748-0221/11/12/C12060
       (cit. on p. 52).

[153] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in
       proton–proton collision data at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **76** (2016) 292,
       arXiv: 1603.05598 [hep-ex] (cit. on pp. 52, 56).

[154] W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*,
       ATL-LARG-PUB-2008-002, 2008, URL: https://cds.cern.ch/record/1099735
       (cit. on p. 53).

[155] ATLAS Collaboration,
       Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1,
       Eur. Phys. J. C **77** (2017) 490, arXiv: 1603.02934 [hep-ex] (cit. on p. 53).

[156] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS
       detector using the 2015–2017 LHC proton–proton collision data*, JINST **14** (2019) P12006,
       arXiv: 1908.00005 [hep-ex] (cit. on p. 53).

[157] ATLAS Collaboration,
       Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data,
       Eur. Phys. J. C **74** (2014) 3071, arXiv: 1407.5063 [hep-ex] (cit. on p. 53).

[158] ATLAS Collaboration, *Electron reconstruction and identification efficiency measurements with
       the ATLAS detector using the 2011 LHC proton–proton collision data*,
       Eur. Phys. J. C **74** (2014) 2941, arXiv: 1404.2240 [hep-ex] (cit. on p. 53).

[159] F. He, J. Liu, R. Ospanov and R. T. Roberts,
       Tagging non-prompt electrons and muons using lifetime and isolation information,
       ATLAS internal communication: CERN, 2018,
       URL: https://cds.cern.ch/record/2632152 (cit. on p. 55).

[160] ATLAS Collaboration,
       Jet reconstruction and performance using particle flow with the ATLAS Detector,
       Eur. Phys. J. C **77** (2017) 466, arXiv: 1703.10485 [hep-ex] (cit. on p. 57).

[161] M. Cacciari, G. P. Salam and G. Soyez, *FastJet user manual*, Eur. Phys. J. C **72** (2012) 1896, arXiv: 1111.6097 [hep-ph] (cit. on p. 57).

[162] M. Cacciari, "FastJet: A Code for fast k_t clustering, and more", *Proceedings, Deep inelastic scattering, 14th International Workshop, DIS 2006* (Tsukuba, Japan, 20th–24th Apr. 2006) 487, arXiv: hep-ph/0607071 [hep-ph] (cit. on p. 57).

[163] ATLAS Collaboration, *In situ calibration of large-radius jet energy and mass in* 13 *TeV proton–proton collisions with the ATLAS detector*, Eur. Phys. J. C **79** (2019) 135, arXiv: 1807.09477 [hep-ex] (cit. on p. 58).

[164] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at* $\sqrt{s}$ = 13 *TeV with the ATLAS detector*, Eur. Phys. J. C **81** (2020) 689, arXiv: 2007.02645 [hep-ex] (cit. on pp. 58, 59).

[165] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at* $\sqrt{s}$ = 8 *TeV using the ATLAS detector*, Eur. Phys. J. C **76** (2016) 581, arXiv: 1510.03823 [hep-ex] (cit. on p. 60).

[166] ATLAS Collaboration, *Forward jet vertex tagging using the particle flow algorithm*, ATL-PHYS-PUB-2019-026, 2019, URL: https://cds.cern.ch/record/2683100 (cit. on p. 60).

[167] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, ATL-PHYS-PUB-2017-013, 2017, URL: https://cds.cern.ch/record/2273281 (cit. on p. 60).

[168] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with tt̄ events in pp collisions at* $\sqrt{s}$ = 13 *TeV*, Eur. Phys. J. C **79** (2019) 970, arXiv: 1907.05120 [hep-ex] (cit. on p. 60).

[169] T. Barillari et al., *Local Hadronic Calibration*, ATLAS Collaboration, 2008, URL: https://cds.cern.ch/record/1112035 (cit. on p. 60).

[170] ATLAS Collaboration, *Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at* $\sqrt{s}$ = 8 *TeV*, Eur. Phys. J. C **75** (2015) 303, arXiv: 1412.7086 [hep-ex] (cit. on p. 60).

[171] ATLAS Collaboration, *Commissioning of the reconstruction of hadronic tau lepton decays in ATLAS using pp collisions at* $\sqrt{s}$ = 13 *TeV*, ATL-PHYS-PUB-2015-025, 2015, URL: https://cds.cern.ch/record/2037716 (cit. on p. 60).

[172] ATLAS Collaboration, *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at* $\sqrt{s}$ = 13 *TeV*, ATLAS-CONF-2017-029, 2017, URL: https://cds.cern.ch/record/2261772 (cit. on p. 61).

[173] ATLAS Collaboration, *Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment*, ATL-PHYS-PUB-2019-033, 2019, URL: https://cds.cern.ch/record/2688062 (cit. on p. 61).

[174] ATLAS Collaboration, *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*, ATL-PHYS-PUB-2015-045, 2015, URL: https://cds.cern.ch/record/2064383 (cit. on p. 61).

[175] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **78** (2018) 903, arXiv: 1802.08168 [hep-ex] (cit. on p. 62).

[176] ATLAS Collaboration, *E_T^{miss} performance in the ATLAS detector using 2015–2016 LHC pp collisions*, ATLAS-CONF-2018-023, 2018, URL: https://cds.cern.ch/record/2625233 (cit. on p. 62).

[177] W. Buttinger, *Using Event Weights to account for differences in Instantaneous Luminosity and Trigger Prescale in Monte Carlo and Data*, Internal communication: CERN, 2015, URL: https://cds.cern.ch/record/2014726 (cit. on p. 64).

[178] ATLAS Collaboration, *Measurement of the production cross-section of a single top quark in association with a Z boson in proton–proton collisions at 13 TeV with the ATLAS detector*, Phys. Lett. B **780** (2018) 557, arXiv: 1710.03659 [hep-ex] (cit. on p. 70).

[179] H. N. We, *Associated Production of a top quark and a Higgs Boson in pp collisions at 13 TeV using the ATLAS detector*, Masterthesis: Rheinische Friedrich-Wilhelms-Universität Bonn, 2020 (cit. on pp. 71, 72).

[180] M. Weiß, *Improvements on the mass reconstruction of single top quark associated Higgs production at ATLAS analyzing the decay channel $H \rightarrow \tau^+\tau^-$*, unpublished Materthesis (cit. on p. 73).

[181] M. Hübner, *Higgs Boson Production Cross-Section Measurements in the Di-Tau Final State at ATLAS*, PhD Thesis: Rheinische Friedrich-Wilhelms-Universität Bonn, 2022, URL: https://hdl.handle.net/20.500.11811/9734 (cit. on pp. 74, 75).

[182] C. Rubbia, "Physics Results of the UA1 Collaboration at the CERN Proton - Anti-Proton Collider", *11th International Conference on Neutrino Physics and Astrophysics*, UA1, 1984, URL: https://cds.cern.ch/record/155129/files/198411025.pdf (cit. on p. 74).

[183] R. Ellis, I. Hinchliffe, M. Soldate and J. Van Der Bij, *Higgs decay to $\tau^+\tau^-$ — A possible signature of intermediate mass Higgs bosons at high energy hadron colliders*, Nuclear Physics B **297** (1988) 221 (cit. on p. 74).

[184] A. Elagin, P. Murat, A. Pranko and A. Safonov, *A new mass reconstruction technique for resonances decaying to $\tau\tau$*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **654** (2011) 481 (cit. on pp. 74, 75).

[185] ATLAS Collaboration, *Search for the Standard Model Higgs boson in the $H \rightarrow \tau^+\tau^-$ decay mode in $\sqrt{s}$ = 7 TeV pp collisions with ATLAS*, JHEP **09** (2012) 070, arXiv: 1206.5971 [hep-ex] (cit. on p. 75).

[186] G. Ke et al., "LightGBM: A Highly Efficient Gradient Boosting Decision Tree",
Advances in Neural Information Processing Systems, ed. by I. Guyon et al., vol. 30,
Curran Associates, Inc., 2017, URL: https://proceedings.neurips.cc/paper/
2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf
(cit. on p. 81).

[187] T. P. Calvet, *Search for the production of a Higgs boson in association with top quarks and
decaying into a b-quark pair and b-jet identification with the ATLAS experiment at LHC*,
Presented 08 Nov 2017, PhD thesis: Aix-Marseille Université, 2017,
URL: https://cds.cern.ch/record/2296985 (cit. on p. 88).

[188] *Proceedings of the 1974 CERN School of Computing: Godøysund, Norway 11 - 24 Aug 1974.
3rd CERN School of Computing*, CERN, Geneva: CERN, 1974 294,
URL: https://cds.cern.ch/record/186223 (cit. on p. 104).

[189] S. Alioli, S.-O. Moch and P. Uwer,
Hadronic top-quark pair-production with one jet and parton showering, JHEP **01** (2012) 137,
arXiv: 1110.5251 [hep-ph] (cit. on p. 133).

[190] J. M. Campbell, R. K. Ellis, P. Nason and E. Re,
Top-Pair Production and Decay at NLO Matched with Parton Showers, JHEP **04** (2015) 114,
arXiv: 1412.1828 [hep-ph] (cit. on p. 133).

[191] E. Re,
Single-top Wt-channel production matched with parton showers using the POWHEG method,
Eur. Phys. J. C **71** (2011) 1547, arXiv: 1009.2450 [hep-ph] (cit. on p. 133).

[192] S. Alioli, P. Nason, C. Oleari and E. Re,
NLO single-top production matched with shower in POWHEG: s- and t-channel contributions,
JHEP **09** (2009) 111, arXiv: 0907.4076 [hep-ph] (cit. on pp. 133, 134).

[193] H. B. Hartanto, B. Jager, L. Reina and D. Wackeroth,
Higgs boson production in association with top quarks in the POWHEG BOX,
Phys. Rev. D **91** (2015) 094003, arXiv: 1501.04498 [hep-ph] (cit. on p. 134).

[194] S. Alioli, P. Nason, C. Oleari and E. Re,
NLO Higgs boson production via gluon fusion matched with shower in POWHEG,
JHEP **04** (2009) 002, arXiv: 0812.0578 [hep-ph] (cit. on p. 134).

[195] E. Bagnaschi, G. Degrassi, P. Slavich and A. Vicini,
Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM,
JHEP **02** (2012) 088, arXiv: 1111.2854 [hep-ph] (cit. on p. 134).

[196] P. Nason and C. Oleari,
NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG,
JHEP **02** (2010) 037, arXiv: 0911.5299 [hep-ph] (cit. on p. 134).

[197] J. M. Campbell et al., *NLO Higgs boson production plus one and two jets using the POWHEG
BOX, MadGraph4 and MCFM*, JHEP **07** (2012) 092, arXiv: 1202.5475 [hep-ph]
(cit. on p. 134).

List of Figures

List of Tables

Acknowledgements

I would like to say thank everyone who has suppported and encouraged me throughout the years.

Special thanks go towards Prof. Ian Brock who allowed me to work on my masterthesis and subsequently PHD thesis in his group. Thank you for all the encouragement and guidance. A big thank you also goes towards Dr. Regina Moles-Valls who guided me throughout my masterthesis and the first year of my PHD.

I would also like to thank other current and former members of the group: Thank you Federico Diaz Capriles for sharing the time in this group with me and the good company at schools, conferences and CERN visits. Thank you Christian Kirfel for countless discussions on everything from presentation styles to particle physics and of course for the machine learning contribution for this analysis. Thank you Florian Kirfel for joining the effort and the deep dive discussions and interest. Thank you Dr. Oleh Kivernyk for providing your view on physics, programming and some more and also the joined effort on this analysis. Thank you Nilima Akolkar for being my office mate, good discussions and support whenever I needed updates on your analysis.

Special thanks also go towards Han Na We, Thilo Scharnhorst and Mathias Weiß who worked with me on this analysis. Thanks for all the questions, answers and work. Of course I would also like to thank the 'senior PHD students' of the group Ozan Arslan, Anjishnu Bandyopadhyay, Irina Cioara and Rui Zhang who helped my a lot with long discussions on physics and guidance on the different parts of a PHD thesis. Big thanks go towards everyone I worked with during my time in the group: Richard Baumann, Marius Blaut, Nicolas Boeing, Chris Boever, Oliver But, Ellinor Eckstein, Anubhav Gupta, Kanhaiya Gupta, Alexander Johnston, Piet Nogga, Maxx Rahman, Marvin Schmitz, Chris Staude, David Stotko, Marius Stricker, Simon Thiele, Luka Vomberg and Dorothee Wohlleben. Thank you for discussions, questions and letting me be part of your scientific journey. Thanks for everyone in the group for good times in the office, game nights and social activities. Even though we could not always spend time in the offices talking to each other we did our best to keep the socializing alive even during hard times.

Of course, I also want to thank people outside of my group. Special thanks go towards Prof. Klaus Desch and the members of his group, especially Priv. Doz. Philip Bechtle and Dr. Christian Grefe for the good discussions and helpful comments. I also would like to thank everyone who joined me at my time at CERN and on conferences, especially Lara Schildgen, Dr. Matthias Hamer and Patrick Bauer.

I would also like to thank Nicole Felde and Andrea Fürstenberg for supporting our groups with everything from organization to paperwork and answering so many of my questions. This thesis would also not exist without the IT-support of the institute. Thanks to everyone and especially to Dr. Oliver Freyermuth for solving or at least explaining every possible and impossible problem concerning the computing infrastructure.

Big thanks also go to other members of the ATLAS collaboration which helped and supported me and sometimes urged me towards a better analysis. Special thanks go towards the members of the *tHq* analysis team. Thank you Carlos Escobar Ibanez, Valentina Vecchio and Nello Bruscino for organizing and building the foundations of the analysis together as well as answering quite some of my questions. I would also like

Curriculum Vitae

Personal Details

Name	Tanja Holm
Date of Birth	Troisdorf
Email	holm@physik.uni-bonn.de
Family status	Single

Education

2003–2012	Abitur, Albert-Einstein-Gymnasium, Sankt Augustin, Germany
2012–2015	BSc in Physics, Rheinische Friedrich-Wilhelms-Universität, Bonn, Germany.
2015–2018	MSc in Physics Rheinische Friedrich-Wilhelms-Universität, Bonn, Germany.
2018–2023	PhD in Physics, Rheinische Friedrich-Wilhelms-Universität, Bonn, Germany.

Professional Experience

2018–2023	Doctoral work at the University of Bonn, Germany.
2019	Fieldwork at CERN, Geneva, Switzerland.

Languages

German	Mother tongue
English	Fluent