Universität Bonn

Physikalisches Institut

Search for top-quark-pair associated Higgs production with tau leptons using the ATLAS detector at the LHC

David Hohn

The search for Higgs boson production in association with a pair of top quarks in multilepton final states is presented. Seven final states defined by the number and flavour of charged leptons are analysed in the data set corresponding to an integrated luminosity of 36.1 fb^{-1} recorded at 13 TeV with the ATLAS detector. The result is an observed excess of events over the expected backgrounds from Standard Model processes with a significance of 4.1 standard deviations (expected 2.8). The measured cross section of the $t\bar{t}H$ process $\sigma_{t\bar{t}H} = 790 \pm 150(\text{stat.})\,^{+170}_{-150}(\text{syst.})\,\text{fb} = 790\,^{+230}_{-210}\,\text{fb}$ is compatible with the expected one of $\sigma^{\text{SM}}_{t\bar{t}H} = 507^{+35}_{-50}\,\text{fb}$.
Particular focus is put on the two final states with two electrons and/or muons (of same or opposite electric charges) and one hadronically decaying tau lepton which contribute significantly to the result.

Physikalisches Institut der
Universität Bonn
Nussallee 12
D-53115 Bonn

BONN-IR-2019-001
CERN-THESIS-2018-367
Februar 2019
ISSN-0172-8741

Search for top-quark-pair associated Higgs production with tau leptons using the ATLAS detector at the LHC

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

von
David Hohn
aus
Hilden

Bonn, Juni 2018

Dieser Forschungsbericht wurde als Dissertation von der Mathematisch-Naturwissenschaftlichen Fakultät der Universität Bonn angenommen und ist auf dem Hochschulschriftenserver der ULB Bonn `http://hss.ulb.uni-bonn.de/diss_online` elektronisch publiziert.

1. Gutachter:	Prof. Dr. Norbert Wermes
2. Gutachter:	Prof. Dr. Ian Brock
Tag der Promotion:	3.9.2018
Erscheinungsjahr:	2019

Contents

CHAPTER 1

Introduction

It is a truth universally acknowledged, that a chiral gauge theory describing massive particles, must be in want of a broken symmetry. Whether the statement is true or not, the LHC and its experiments, ATLAS and CMS, have spectacularly demonstrated that a new boson exists in 2012. Since the discovery it has been confirmed by a multitude of measurements e.g., of mass, spin, CP invariance, decay width, cross sections (total, fiducial and differential) and couplings, that the boson behaves consistently with a theory of spontaneously broken electroweak symmetry as suggested by P. Higgs, F. Englert et al. [1–3]. And thus the Standard Model (SM) does have a mechanism to explain the masses of its gauge bosons as well as its fermions. This discovery was the most momentous event in recent particle physics history as evidenced by the bestowal of the Nobel prize in physics on P. Higgs and F. Englert [4]

> "*for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles, and which recently was confirmed through the discovery of the predicted fundamental particle, by the ATLAS and CMS experiments at CERN's Large Hadron Collider*"

in the very next year.

As the boson was discovered in bosonic (W, Z, γ) decays, another milestone in confirming that the observed boson is the predicted Higgs boson was the measurement of its coupling to fermions. Only in such, so called Yukawa couplings, do the fermions acquire mass. This was demonstrated by ATLAS and CMS for the heaviest lepton, the tau lepton, and the heavy b quark. These particles are the prime candidates for analysis because the coupling strength is proportional to the fermion's mass. Evidence for the coupling to the heaviest particle in the SM, the top quark, for which the Yukawa coupling is expected to be strongest, was elusive until 2017. This thesis presents the analysis that provided the first evidence for top-Higgs Yukawa coupling in multilepton final states at ATLAS [5].

While the top-Yukawa coupling is the largest of its type in the SM, measuring the direct tree-level interaction between the Higgs boson and top quarks is difficult. The mass of the Higgs boson is small, such that the decay to top quarks is extremely suppressed. So the only available avenue to observe it, is to provide a top quark for the Higgs to couple to in its production. The $t\bar{t}H$ process, while rare from an inclusive point of view, is the most common Higgs production mechanism involving tree-level top-Higgs couplings with a cross section of $507\,^{+35}_{-50}$ fb for an LHC energy of $\sqrt{s} = 13$ TeV. Observing a process with such a small cross section requires sophisticated analysis techniques and as large a signal acceptance as possible. The latter requirement is motivation for looking for the $t\bar{t}H$ process in multilepton final states which come about via $H \to WW^*, \tau\tau, ZZ^*$ decays. By specifically considering tau leptons in hadronic decays one can enhance the fraction of $H \to \tau\tau$ events which contain only fermionic Higgs couplings.

However, at the present experimental and statistical sensitivity no separation between the Higgs decays is possible. The sophisticated techniques are applied to perform background estimation and reduction.

The multilepton signatures are chosen in a way that reduces contributions from SM backgrounds. Other ways to search for $t\bar{t}H$ are briefly discussed as the general strategy to find $t\bar{t}H$ is to join forces of and combine as many analyses as possible. The channel with final state $2\ell(\mathrm{OS})1\tau_{\mathrm{had}}$ is analysed for the first time.

The work described in this thesis is embedded in the publication by the ATLAS collaboration which reports evidence for $t\bar{t}H$ production [5]. The personal contributions of the author include analysis of events with two light leptons and one hadronic tau lepton, namely the $2\ell(\mathrm{OS})1\tau_{\mathrm{had}}$ and $2\ell(\mathrm{SS})1\tau_{\mathrm{had}}$ final states, as well as the statistical combination of all channels. Therefore, particular focus is put on these topics while also describing the context in which the work was performed where it is relevant and conducive to the understanding of $t\bar{t}H$ as a whole.

The thesis is structured as follows. The theoretical background required for the understanding of the importance of the $t\bar{t}H$ process and the multilepton signatures is described in Chapter 2. The experimental setup used to record data in form of the LHC collider and ATLAS detector is described in Chapter 3. The following Chapters 4 to 6 describe the $t\bar{t}H$ analyses in general and the multilepton one in particular with special focus on the channels with two electrons or muons and one hadronically decaying tau lepton. Chapters 7 to 9 present the results and how they are interpreted statistically. Chapter 10 concludes.

CHAPTER 2

Theoretical Background

Particle physics is the branch of physics that studies the elementary particles and fundamental forces of nature. The theoretical models used to describe the understanding of nature have undergone many iterations. Physicists have come from bewilderment at the seeming lack of structure in the particle "zoo" to a consistent theory of quantum fields. The following sections describe the latter.

2.1 Matter and forces

All known matter in the universe can be described as consisting of elementary particles with half-integer spin, called fermions. Fermions are distinguished into leptons and quarks based on their interaction with each other. Furthermore, fermions can be arranged in a pattern of three generations. Each subsequent generation of fermions has larger masses than the previous one but otherwise identical properties.

Leptons can be electrically charged like the electron (e) as well as electrically neutral like the neutrino (ν). The second generation lepton is called muon (μ) while the third one is called tau lepton, tauon or simply and succinctly tau (τ), each also accompanied by their neutrinos.

On the other hand, all quarks are electrically charged where the up quark has a fractional charge of $\frac{2}{3}$ and the down quark of $-\frac{1}{3}$. Their later generation equivalents are charm, strange, top and bottom. A summary of the fermions and world averages of measurements of their masses is shown in Tab. 2.1. In the table the neutrino masses are indicated to be approximately zero. In fact their masses are non-zero as evidenced by the existence of neutrino oscillations, but are negligible for the context of $t\bar{t}H$ searches. Hence neutrinos are treated as massless in the remainder of this text.

1st generation		2nd generation		3rd generation		
mass [MeV]		mass [MeV]		mass [GeV]		charge [e]
ν_e	~ 0	ν_μ	~ 0	ν_τ	~ 0	0
e	0.511	μ	106	τ	1.78	-1
u	$2.2^{+0.5}_{-0.4}$	c	$1\,275^{+25}_{-35}$	t	173.0 ± 0.4	$\frac{2}{3}$
d	$4.7^{+0.5}_{-0.3}$	s	95^{+9}_{-3}	b	$4.18^{+0.04}_{-0.03}$	$-\frac{1}{3}$

Table 2.1: Leptons and quarks arranged by generations. Note the change of mass unit in the third generation. The unit of charge is the elementary positive charge. Uncertainties smaller than the shown significant digits are omitted [6].

Quarks also carry an additional colour charge which can take the three values called red, green and blue. However, quarks have never been observed in a free state. Instead they always form bound states of two or three quarks called hadrons. Hadrons are composed of quarks in such a way that the colour charges cancel, i.e. they are colourless. As a consequence the colour charges cannot be observed directly. Another consequence of the number of quarks inside hadrons being two or three is that hadrons have integer electric charge and the fractional charges of quarks are never observed freely, either.

Finally each fermion has an antifermion partner which has identical properties but opposite charge numbers. Fermions are always produced in pairs with one fermion and one antifermion. Because of the larger masses, the later generations decay into earlier generations. Only particles of the first generation can form stable matter.

The four fundamental forces in physics are the gravitational, weak, electromagnetic and strong forces. All but gravitation can at present be explained by the exchange of mediating particles between fermions. The mediators are called bosons and have spin 1. Electro-magnetism is mediated by the massless photon (γ) which couples to the electric charge. So all quarks and the charged leptons are affected by the electromagnetic interaction. The range of the force is infinite and becomes weaker at larger distances.

Massive charged $W^{\pm}$ and neutral Z boson carry the weak force. Due to their large masses the force has a limited range. The weak force affects all fermions.

The strong force is mediated by eight gluons (g). Gluons are massless and carry colour charges. The eight gluons differ only in their colour states which are different combinations and mixtures of the three colour charges, red, green and blue. This self-charge of gluons causes self-interactions which generate the peculiar effects of confinement and asymptotic freedom. These are explained in the following sections.

Some properties of the force-carrying bosons are summarised in Tab. 2.2.

Mediator	mass [GeV]	electric charge [e]	force
γ	0	0	electromagnetic
$W^{\pm}$	80.379 ± 0.012	± 1	weak
Z	91.1876 ± 0.0021	0	weak
g	0	0	strong

Table 2.2: Masses and charges of the mediating particles of the electroweak and strong forces [6].

2.2 Standard Model

The theoretical model that describes most of particle physics is called the Standard Model (SM) [7–9]. It obtained its name by its exceptional success in predicting phenomena that were later observed and the excellent precision and accuracy of the agreement with experimental measurements. It is a quantum field theory (QFT) in which matter is represented by fermion fields with half-integer spin. The physical particles are quantisations of those fields and the interactions between particles are described by gauge fields which arise from local gauge symmetries. The following sections briefly describe how the gauge principle is applied to the Lagrangian formalism of QFT and is based on citations [10–12].

A Lagrangian describes dynamic physical systems in generalised coordinates. It is defined as the subtraction of the kinetic (T) and potential energy (V), $\mathcal{L} = T - V$. The time integral of $\mathcal{L}$ is called the action which is used in the variational principle of least action. This principle states that the system takes

the path between states of the integrated time which minimises the action. From the principle, the time evolution of the system can be derived.

In particle physics one can derive field equations and Feynman rules from the Lagrangian. These rules are used in calculating transition amplitudes and cross sections.

2.2.1 QED

A Lagrangian that gives the Dirac equation for fermions by using the principle of least action is

$$\mathcal{L}_{\text{Dirac}} = i\bar{\psi}\gamma_\mu\partial^\mu\psi - m\bar{\psi}\psi, \tag{2.1}$$

where ψ is the fermion field and γ^μ are the Dirac gamma matrices. The fermion field changes under local transformations of the symmetry of the abelian unitary U(1) group as

$$\psi(x) \to \psi'(x) = e^{i\alpha(x)}\psi(x), \tag{2.2}$$

where $\alpha(x)$ is a phase change that depends on the space-time coordinate x. In order for the Lagrangian to stay invariant under this transformation a new derivative needs to be introduced

$$\partial_\mu \to D_\mu = \partial_\mu + iqA_\mu(x), \tag{2.3}$$

where A_μ is a new vector field. It is also called gauge field because it was introduced to restore the gauge (phase) invariance. The new derivative D_μ is called gauge covariant derivative. At the same time the gauge field is required to transform as

$$A_\mu(x) \to A'_\mu(x) = A_\mu(x) - \frac{1}{q}\partial_\mu\alpha(x). \tag{2.4}$$

The new, under U(1) locally invariant, Lagrangian is thus

$$\mathcal{L}_{\text{Dirac,loc.inv.}} = i\bar{\psi}\gamma_\mu D^\mu\psi - m\bar{\psi}\psi \tag{2.5}$$

$$= i\bar{\psi}\gamma_\mu\partial^\mu\psi - m\bar{\psi}\psi - q\bar{\psi}\gamma_\mu\psi A^\mu. \tag{2.6}$$

The additional term compared to Eq. 2.1 represents an interaction between the fermion field and the new vector gauge field with a coupling strength q. Gauge field excitations can be interpreted as the photon which makes the interaction the electromagnetic one between charged fermions and the photon. By adding the free term of the gauge field to the Lagrangian one obtains the one of quantum electrodynamics

$$\mathcal{L}_{\text{QED}} = i\bar{\psi}\gamma_\mu\partial^\mu\psi - m\bar{\psi}\psi - q\bar{\psi}\gamma_\mu\psi A^\mu - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{2.7}$$

where $F_{\mu\nu}$ is the electromagnetic field strength tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \tag{2.8}$$

and the coupling strength q becomes the electric charge. It should be pointed out that a hypothetical mass term of the gauge field of the form $A_\mu A^\mu$ is not invariant and is therefore forbidden. This matches with the physical property of the photon which is massless.

Here the simple requirement that $\mathcal{L}$ be invariant under local transformations demands that the photon vector field A_μ exists and that it interacts with the fermions.

2.2.2 Electro-weak extension

The description of QED through the U(1) symmetry can be extended to describe electromagnetic and weak interactions by considering an additional symmetry from the special unitary group SU(2). The gauge transformations are taken from the group $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$, where the subscript L indicates that SU(2) acts only on left-handed fields and Y is the weak hypercharge. Here it is necessary to distinguish between left- and right-handed chiral fermion fields. They can be written as

$$\psi = \frac{1}{2}(1-\gamma^5)\psi + \frac{1}{2}(1+\gamma^5)\psi = \psi_L + \psi_R, \tag{2.9}$$

where $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$ is the chirality operator.

The known particles group into left-handed doublets and right-handed singlets as

$$\psi_L = \begin{pmatrix}\nu_e\\ e\end{pmatrix}_L, \quad \begin{pmatrix}\nu_\mu\\ \mu\end{pmatrix}_L, \quad \begin{pmatrix}\nu_\tau\\ \tau\end{pmatrix}_L, \tag{2.10}$$

$$\psi_L = \begin{pmatrix}u\\ d\end{pmatrix}_L, \quad \begin{pmatrix}c\\ s\end{pmatrix}_L, \quad \begin{pmatrix}t\\ b\end{pmatrix}_L, \tag{2.11}$$

$$\psi_R = e_R, \qquad \mu_R, \qquad \tau_R, \tag{2.12}$$

$$\psi_R = u_R, \qquad c_R, \qquad t_R, \tag{2.13}$$

$$\psi_R = d_R, \qquad s_R, \qquad b_R. \tag{2.14}$$

Local transformations of the fields are then

$$\psi_L(x) \to \psi'_L(x) = e^{i\beta(x)Y + i\alpha(x)\frac{1}{2}\tau}\,\psi_L(x) \tag{2.15}$$

$$\psi_R(x) \to \psi'_R(x) = e^{i\beta(x)Y}\qquad \psi_R(x), \tag{2.16}$$

where τ are the three Pauli matrices and $\alpha(x)$ is thus a phase rotation with three components. One can see that the right-handed fields transform only with U(1) as before, with weak hypercharge instead of electric charge, but the left-handed one with the full $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$.

Satisfying local gauge invariance requires the introduction of four gauge fields, three $W_\mu^{1,2,3}$ from $\mathrm{SU}(2)_L$ and one B_μ from $\mathrm{U}(1)_Y$. They transform as

$$\vec{W}_\mu(x) \to \vec{W}'_\mu(x) = \vec{W}_\mu(x) - \frac{1}{g}\partial_\mu\vec{\alpha}(x) - \vec{\alpha}(x) \times \vec{W}_\mu(x) \tag{2.17}$$

$$B_\mu(x) \to B'_\mu(x) \;= B_\mu(x) - \frac{1}{g'}\partial_\mu\beta(x), \tag{2.18}$$

with the covariant derivative

$$\partial_\mu \to D_\mu = \partial_\mu + i\frac{g}{2}\vec{\tau}\cdot\vec{W}_\mu(x) + i\frac{g'}{2}YB_\mu(x), \tag{2.19}$$

in which g and g' are the coupling strengths associated with the $\mathrm{SU}(2)_L$ and $\mathrm{U}(1)_Y$ groups, respectively, and the cross product $\times$ is used as a shorthand notation for the anticommuting SU(2) structure constants.

The locally invariant Lagrangian with added free terms of the gauge fields is thus

$$\mathcal{L}_{\mathrm{EW}} = \mathcal{L}_{\psi}^{\mathrm{free}} + \mathcal{L}^{\mathrm{interaction}} + \mathcal{L}_{\mathrm{gauge}}^{\mathrm{free}} \tag{2.20}$$

$$\mathcal{L}_{\psi}^{\mathrm{free}} = i\overline{\psi}\gamma_{\mu}\partial^{\mu}\psi - m\overline{\psi}\psi \tag{2.21}$$

$$\mathcal{L}^{\mathrm{interaction}} = -\frac{g^{'}}{2}Y\overline{\psi}\gamma^{\mu}B_{\mu}\psi - \frac{g}{2}\overline{\psi}_{L}\gamma^{\mu}\vec{\tau}\cdot\vec{W}_{\mu}\psi_{L} \tag{2.22}$$

$$\mathcal{L}_{\mathrm{gauge}}^{\mathrm{free}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}. \tag{2.23}$$

The field strength tensors are defined as

$$\vec{F}_{\mu\nu} = \partial_{\mu}\vec{W}_{\nu} - \partial_{\nu}\vec{W}_{\mu} + g\vec{W}_{\mu} \times \vec{W}_{\nu} \tag{2.24}$$

$$B_{\mu\nu} = \partial_{\mu}B_{\nu} - \partial_{\nu}B_{\mu}, \tag{2.25}$$

where the non-abelian nature of $\mathrm{SU}(2)_L$ and the additional term in the tensor implies cubic and quartic self interactions of the W_{μ} gauge fields. The W_{μ} fields couple to left-handed fermions only, while B_{μ} still couples to both left- and right-handed.

The W_{μ} and B_{μ} fields can be rotated so that they correspond to the physical $W^{\pm}$, Z and A(photon) fields

$$W_{\mu}^{\pm} = \frac{1}{\sqrt{2}}(W_{\mu}^{1} \mp W_{\mu}^{2}) \tag{2.26}$$

$$A_{\mu} = \quad B_{\mu}\cos\theta_{W} + W_{\mu}^{3}\sin\theta_{W} \tag{2.27}$$

$$Z_{\mu} = -B_{\mu}\sin\theta_{W} + W_{\mu}^{3}\cos\theta_{W}. \tag{2.28}$$

Here θ_W is the weak mixing angle and relates to the g and $g^{'}$ couplings as $\tan\theta_W = \frac{g^{'}}{g}$. The rotated fields represent the physical charged $W^{\pm}$ bosons which couple charged leptons to their neutrinos and up-type quarks to down-type quarks, the Z boson which couples to any left-handed fermions and the photon which couples to any electrically charged fermions.

The electroweak quantum numbers isospin I_3 (its third component) and hypercharge Y are related to the electric charge Q by

$$Q = I_3 + \frac{1}{2}Y, \tag{2.29}$$

such that the electromagnetic U(1) symmetry is maintained. The values of these numbers for all relevant particles are summarised in Tab. 2.3.

As is the case for QED, massive gauge fields are forbidden in the Lagrangian. However, in this case it is a contradiction with the knowledge that the W and Z bosons are massive. Furthermore the mass term of the fermion fields is also forbidden because it mixes left- and right-handed fields which is in disagreement with physical massive matter particles. The solution to both problems is the introduction of a complex scalar field which breaks the symmetry.

2.2.3 Symmetry breaking

By adding a complex scalar field with a particular potential the electroweak symmetry can be broken which allows the existence of mass terms for the gauge bosons. In the process a new massive scalar boson emerges, the Higgs boson. This is called the Higgs mechanism, first described almost simultaneously by Higgs, Englert, Brout, Guralnik, Hagen and Kibble [1–3].

Particle	I_3	Y	Q
$\begin{pmatrix}\nu_{e\,L}\\ e_L\end{pmatrix}$	$+\frac{1}{2}$ $-\frac{1}{2}$	-1 -1	0 -1
(e_R)	0	-2	-1
$\begin{pmatrix}u_L\\ d_L\end{pmatrix}$	$+\frac{1}{2}$ $-\frac{1}{2}$	$+\frac{1}{3}$ $+\frac{1}{3}$	$+\frac{2}{3}$ $-\frac{1}{3}$
(u_R)	0	$+\frac{4}{3}$	$+\frac{2}{3}$
(d_R)	0	$-\frac{2}{3}$	$-\frac{1}{3}$
$W^{\pm}$	±1	0	±1
Z, γ	0	0	0
H	$-\frac{1}{2}$	1	0

Table 2.3: Summary of the electroweak quantum numbers, third component of the weak isospin I_3 and weak hypercharge Y, which are related to the electric charge by $Q = I_3 + \frac{Y}{2}$, for all relevant particles. Subscript L and R stand for left- and right-handed fermions. The second and third generation fermions are exactly similar.

An isospin doublet of complex scalar fields $\Phi = \begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix}$ can be added to the Lagrangian in a generic form as

$$\mathcal{L}_{\text{scalar}} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi), \tag{2.30}$$

with the same covariant derivative as Eq. 2.19 and a particular choice of potential

$$V(\Phi) = \mu^2(\Phi^\dagger \Phi) + \lambda(\Phi^\dagger \Phi)^2. \tag{2.31}$$

The potential is symmetric around zero and when $\mu^2 > 0$ also has a minimum at zero. When $\mu^2 < 0$ the lowest energy state is degenerate and describes a circle as illustrated in Fig. 2.1 for a single complex field. These minima are located at $\phi_0^2 = \frac{-\mu^2}{2\lambda} = \frac{v^2}{2}$ in which $v \equiv \sqrt{\frac{-\mu^2}{\lambda}}$ is called the vacuum expectation value.

A Lagrangian should be considered in the proximity of a minimum of the potential. So by choosing one specific minimum from the degenerate states the symmetry is spontaneously broken. The conventional choice of minimum is $\Phi_0 = \frac{v}{\sqrt{2}}\begin{pmatrix}0\\ 1\end{pmatrix}$ for which one can rewrite Φ to be centred around it in terms of shifted fields, $\eta_{1,2,3}$ and h, as

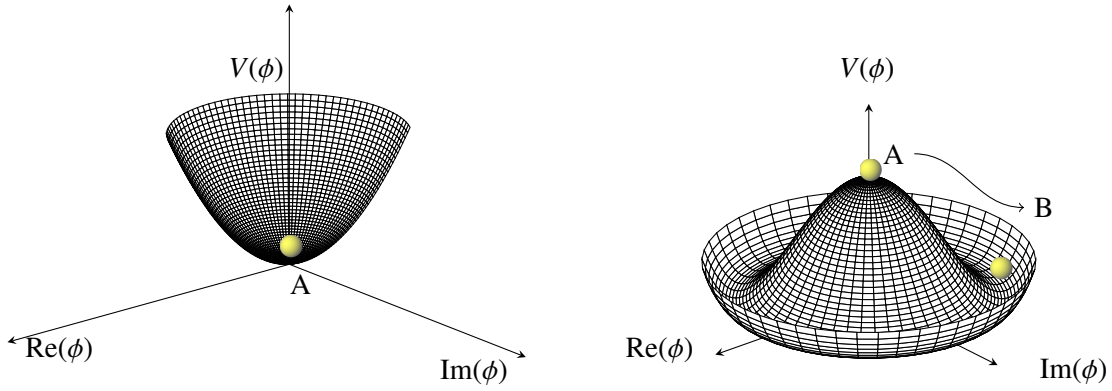
$$\Phi(x) = \begin{pmatrix}\eta_1(x) + i\eta_2(x)\\ \frac{1}{\sqrt{2}}(v + h(x)) - i\eta_3(x)\end{pmatrix}. \tag{2.32}$$

By making an appropriate choice of gauge transformation for Φ (unitary gauge)

$$\Phi(x) \to \Phi'(x) = e^{-i\frac{1}{2v}\vec{\tau}\cdot\vec{\eta}(x)}\Phi(x) = \frac{v + h(x)}{\sqrt{2}}\begin{pmatrix}0\\ 1\end{pmatrix}, \tag{2.33}$$

the η fields are made to be hidden from the Lagrangian and the associated degrees of freedom reappear as mass terms for the massive weak gauge bosons. The h field corresponds to a massive scalar boson H, the Higgs boson.

The expanded scalar addition to the Lagrangian in terms of the h field and the rotated electroweak



(a) $\mu^2 > 0$ (b) $\mu^2 < 0$

Figure 2.1: Illustration of the potential of a complex scalar field of type $V(\phi) = \mu^2\phi^2 + \lambda\phi^4$ with $\lambda > 0$ for two different values of μ^2. In both cases the potential is symmetric around the origin. In (a) $\mu^2 > 0$ and the ground state is also the origin labelled A. In (b) $\mu^2 < 0$ and the ground state is a degenerate circle where one arbitrary ground state is labelled B and breaks the electroweak symmetry.

fields is

$$\mathcal{L}_{\text{scalar}} = \frac{1}{2}\partial_\mu h\,\partial^\mu h + \frac{g^2}{4}(v+h)^2\left(W_\mu^+ W^{-\mu} + \frac{1}{2\cos^2\theta_W}Z_\mu Z^\mu\right) - \lambda v^2 h^2 - \lambda v h^3 - \frac{\lambda}{4}h^4 \tag{2.34}$$

Here one can see terms quadratic in the W and Z fields which represent mass terms. By comparing to known mass relations for gauge fields one can recognise the vector boson masses as

$$m_W = \frac{gv}{2} \qquad \text{and} \qquad m_Z = \frac{gv}{2\cos\theta_W} = \frac{m_W}{\cos\theta_W}, \tag{2.35}$$

as well as the scalar Higgs boson mass $m_H = 2\lambda v^2$. As is required by the existence of the massless photon there is no mass term for the A_μ field. Furthermore there are terms which imply interactions between the vector and scalar fields in vertices like hVV and $hhVV$ with coupling strength proportional to m_V^2. Lastly, one can recognise cubic and quartic self-interactions of the h field. The value of v can be deduced from the boson mass relations to be roughly 246 GeV. The parameter λ is unconstrained which means that no prediction for the mass of the Higgs boson was possible before its discovery.

2.2.4 Fermion masses

The fermion masses can also be generated in this theory with the scalar Higgs field. It was proposed by Weinberg [9] that a coupling between the fermion fields and the scalar fields can explain lepton masses. These "Yukawa" couplings are extensible to quarks as well. The addition to the Lagrangian due to the Yukawa terms is

$$\mathcal{L}_{\text{Yukawa}} = -y_f\left(\overline{\psi}_L\Phi\psi_R + \overline{\psi}_R\tilde{\Phi}\psi_L\right) \tag{2.36}$$

$$= -\frac{y_f}{\sqrt{2}}\,v\,\overline{\psi}\psi - \frac{y_f}{\sqrt{2}}\,h\,\overline{\psi}\psi, \tag{2.37}$$

where $\tilde{\Phi} = -i\tau_2\Phi^* = -\frac{1}{\sqrt{2}}\begin{pmatrix} v+h \\ 0 \end{pmatrix}$. One can identify a mass term for fermions where the Yukawa coupling y_f is rearranged as the fermion mass as

$$m_f = \frac{y_f v}{\sqrt{2}}. \tag{2.38}$$

The interaction strength in the second term between the Higgs field and fermions is therefore proportional to the fermion mass or Yukawa coupling.

When expanding the spinors in Eq. 2.37 for quarks of all generations, one obtains mass terms that mix the generations which cannot be interpreted as physical particles with a mass. However, it is possible to rotate the definition of quark eigenstates such that the mixed-generation terms are hidden in the rotation. The rotated states are called mass eigenstates because they have well-defined masses, while the other ones are called interaction eigenstates of the weak interaction. The rotation is parametrised with a unitary matrix V_{CKM}, named after Cabibbo, Kobayashi and Maskawa [13, 14]. It connects both types of eigenstates by

$$u_i^I = u_j \tag{2.39}$$

and

$$d_i^I = V_{\mathrm{CKM}} d_j \tag{2.40}$$

$$\begin{pmatrix} d^I \\ s^I \\ b^I \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{2.41}$$

where the superscript I indicates the interaction basis and i, j are generation indices. The up-type quarks are conventionally set equal in both descriptions and only the down-type quarks rotate. A consequence of the rotation is that the flavour and generation of quarks expressed in mass eigenstates can change in interactions with a $W^{\pm}$. The transitions between flavours $i \to j$ are described by the individual elements of V_{CKM} and are proportional to $|V_{ij}|^2$. Its structure is almost diagonal with considerable mixing between the first and second generation, but almost no mixing with the third generation because the top quark decays almost exclusively to a bottom quark. A representation that illustrates the structure well is in Wolfenstein parameters [15], where the values of all parameters are $0 < A, \lambda, \eta, \rho < 1$ and higher powers of the parameters become smaller

$$V_{\mathrm{CKM}} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4). \tag{2.42}$$

The quark mixing is also the theoretical explanation that mesons of mixed-flavour quarks are unstable and for CP violation in neutral meson mixing and meson decays.

2.2.5 QCD

The interactions of quantum chromodynamics are inferred from symmetries of the special unitary SU(3) group by the same principle of local gauge invariance. The symmetry of the group generates eight vector fields which correspond to the QCD bosons, the gluons. There are three types of conserved charges. To form neutral states a charge can cancel its anticharge or all three kinds of charges can be added. In

analogy to additive colour mixing the charges are called red, green and blue and explain the "chromo" part of the theory's name. Gluons carry mixtures of colour and anticolour, while quarks carry only one colour charge.

A peculiar property of QCD is that colour charges are confined in colour-neutral states of bound states of quarks. A quark and an antiquark can form mesons in which the quarks' colours cancel, e.g. pions. Baryons are bound states of three quarks or three antiquarks where the colour sum is also neutral, e.g. protons and neutrons. Confinement is due to antiscreening of the colour charge at large distances by the gluons which themselves carry colour. An instructive way to imagine this is the Lund string model which says that when two quarks are pulled apart there is a gluon flux tube between them that acts like a spring. When the distance increases the potential energy of the gluon string increases until it is large enough to create a new quark-antiquark pair from the vacuum. This can take place until the quark energies are low enough to form bound hadrons. As a result, energetic quarks and gluons are not observable directly but only as a spray of hadrons in the same general direction, called jets.

On the other hand, at short distances quarks behave almost freely because the effective interaction strength of the strong force decreases with distance. This behaviour of QCD is known as asymptotic freedom.

2.2.6 Calculating observable cross sections

From the Lagrangian of the SM, rules for calculating transition amplitudes between initial and final states can be derived, the Feynman rules. They are visually depicted by diagrams in which particles are represented by lines or "propagators". Vertices where lines come together represent interactions. Following the rules each propagator and vertex adds a factor to the transition amplitude. The correspondence to the Lagrangian can be made in that vertex factors are derived from the interaction terms and propagator factors from the free terms in $\mathcal{L}$ [12]. In the diagrams fermions are represented by solid straight lines, electroweak vector bosons by sinusoidal ("wavy") lines, gluons by cycloid ("springy") lines and Higgs bosons by dashed straight lines. In this thesis the Feynman diagrams are drawn with time running from left to right, so initial states are on the left and final ones on the right.

Renormalisation

A cross section is then computed as the absolute square of the transition amplitudes summed over all possible Feynman diagrams and integrated over all available phase space. Higher orders of the perturbation series for the amplitude are called loop corrections because they appear as closed loops in the Feynman diagrams. As the particles appearing in the loop are entirely internal to the transition their momenta has to be integrated over all values. However the vertex and propagator factors are such that the momentum integral is divergent for large momenta and the higher-order cross section would be infinite.

The solution is to regularise, e.g. by a cutoff, the divergent integral into a finite and a divergent part. If the theory is renormalisable, which the SM is, the divergent part can be absorbed inside the coupling strengths of the interactions that the loop modifies (e.g. charge, mass).

This can be understood by the fact that the coupling strength in the no-loop calculation are not physically measurable, they are the bare couplings and divergences in terms of unphysical parameters are not meaningful. Effective couplings are obtained by regularisation and renormalisation. In this process the couplings gain a dependence on an arbitrary momentum scale μ_R, called renormalisation scale, and are said to be running. The requirement that any observable should be independent of this scale gives rise to renormalisation group equations (RGE) which determines how a coupling evolves with the scale. A coupling, e.g. α_s, can be measured for a given scale μ and be evolved to any desired value of μ' by the

RGE. In Fig. 2.2 the evolution of the strong coupling constant α_s is shown with several measurements at different momentum transfers which agree well with each other. Here one can see that the coupling becomes smaller for larger momentum transfers which corresponds to the aforementioned asymptotic freedom. Conversely, the growing coupling with small momentum transfers can be seen and diverges quickly for momentum values below 1 GeV. This divergence of the perturbative description is called the Landau pole and requires non-perturbative descriptions in this regime. This long-distance behaviour of α_s is another explanation for colour confinement.

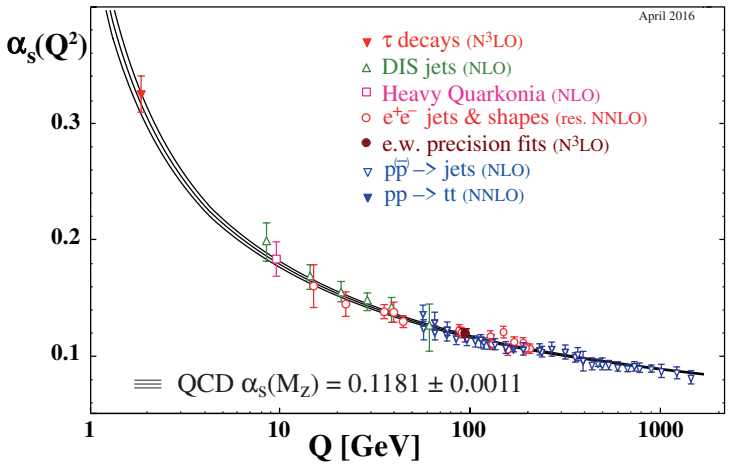


Figure 2.2: Running of the strong coupling constant α_s with momentum scale Q [6][1].

In calculations with all orders there is no observable dependence on μ_R, but at fixed order there is. At first it seems problematic that the predicted cross section (or any observable) depends on a completely arbitrary scale. However, this can be used to estimate theoretical uncertainties due to missing orders in the calculation by varying μ_R.

Factorisation

The factorisation theorem states that short distance (high momentum) and long distance (low momentum) behaviour in a scattering process do not affect each other and can be factorised [16]. In practice this means that a cross section can be computed perturbatively in terms of incoming quarks and gluons, while their non-perturbative behaviour inside the proton is described by parton[2] density functions (PDF). The total cross section of a certain process in proton–proton collisions can therefore be written as

$$\sigma = \sum_{i,j} \int_0^1 \mathrm{d}x_1 \int_0^1 \mathrm{d}x_2 \, f_i(x_1, \mu_F) \, f_j(x_2, \mu_F) \, \hat{\sigma}_{ij}\left(x_1 x_2 s, \alpha_s(\mu_R), \mu_F\right) \tag{2.43}$$

where $\hat{\sigma}_{ij}$ is the perturbative partonic cross section with the initial state partons labelled i and j, $f_i(x, \mu_F)$ is the distribution function of parton i inside the proton at the momentum scale μ_F, $x_{1,2}$ are the momenta

[1] "Quantum Chromodynamics" review
[2] A parton is thus any quark or gluon.

fractions of the incoming partons and $\sqrt{s}$ is the centre of mass energy of the pp collision. Here the factorisation scale μ_F represents the threshold between the perturbative calculation of the cross section and the non-perturbative measurement of parton densities in form of PDFs. A parton distribution function $f_i(x, \mu)$ encodes the probability density of finding a parton i with a longitudinal momentum fraction x inside the proton at a momentum scale μ. An example of a set of PDFs is shown in Fig. 2.3 at a low and high momentum scale. One can see two peaks near $x \sim 0.2$ for the valence up and down quarks which are the main constituents of the proton and give it its quantum numbers. At low values of x the sea quarks and especially gluons are the dominant contributions. Sea quarks and gluons arise from quantum fluctuations like gluon radiation and $q\bar{q}$ pair creation. The importance of the gluon PDF becomes even greater when the proton is probed at larger momenta where they can carry over half of the proton momentum. The PDFs at low and high momentum scales can be imagined as low and high resolution "pictures" of the proton structure.

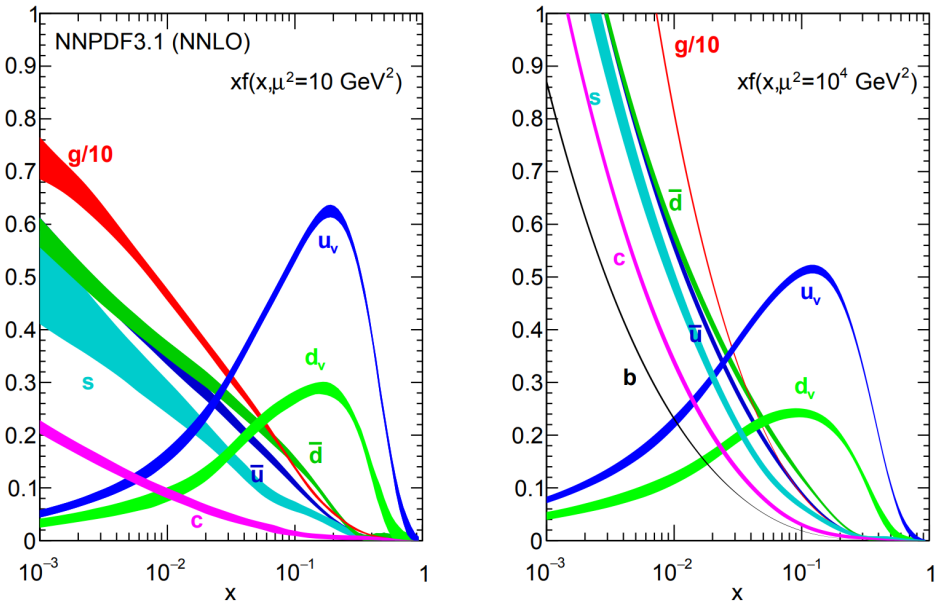


Figure 2.3: Example of proton parton distribution functions: NNPDF3.1 NNLO PDFs at momentum scales (left) $\mu^2 = 10\,\mathrm{GeV}^2$ and (right) $\mu^2 = 10^4\mathrm{GeV}^2$ [17].

The factorisation scale is arbitrary and as is the case for renormalisation any observable should not depend on the choice of μ_F. The RGEs that can be derived from this requirement are called DGLAP[3] equations which allow the evolution of PDFs from a given scale μ to another one μ'. First measurements of PDFs were performed in deep inelastic scattering of electrons and protons, e.g. in ep collisions of HERA. Nowadays, PDFs sets are global analyses of a large variety of additional QCD measurements from the hadron colliders, Tevatron and LHC. Different PDF sets exist that are published by several collaborations. Again variation of μ_F is used as a way to estimate the theoretical uncertainty. Typically both scales, renormalisation and factorisation, referred to as QCD scales, are set equal and varied coherently.

The cross section calculations that are computed by using the Standard Model as described are in excellent agreement with experimental measurements. At the LHC alone, the evidence is overwhelming that the SM makes reliable predictions and is valid. Fig. 2.4 shows a multitude of production cross section

[3] after Dokshitzer, Gribov, Lipatov, Altarelli and Parisi

measurements performed by ATLAS at the LHC which are compared to the theoretical SM predictions and shows how well theory and experiments agree. The figure also illustrates the relative frequency of the processes. One should note that jet production is several orders of magnitude more likely than weak boson production. In fact the most frequent process with high momentum transfers at a hadron collider is multiple jet production (called multijet). It is thus possible to use leptonic weak boson decays to reject jet backgrounds. The mere presence of an isolated lepton is a powerful tool not only for background reduction but also data acquisition.

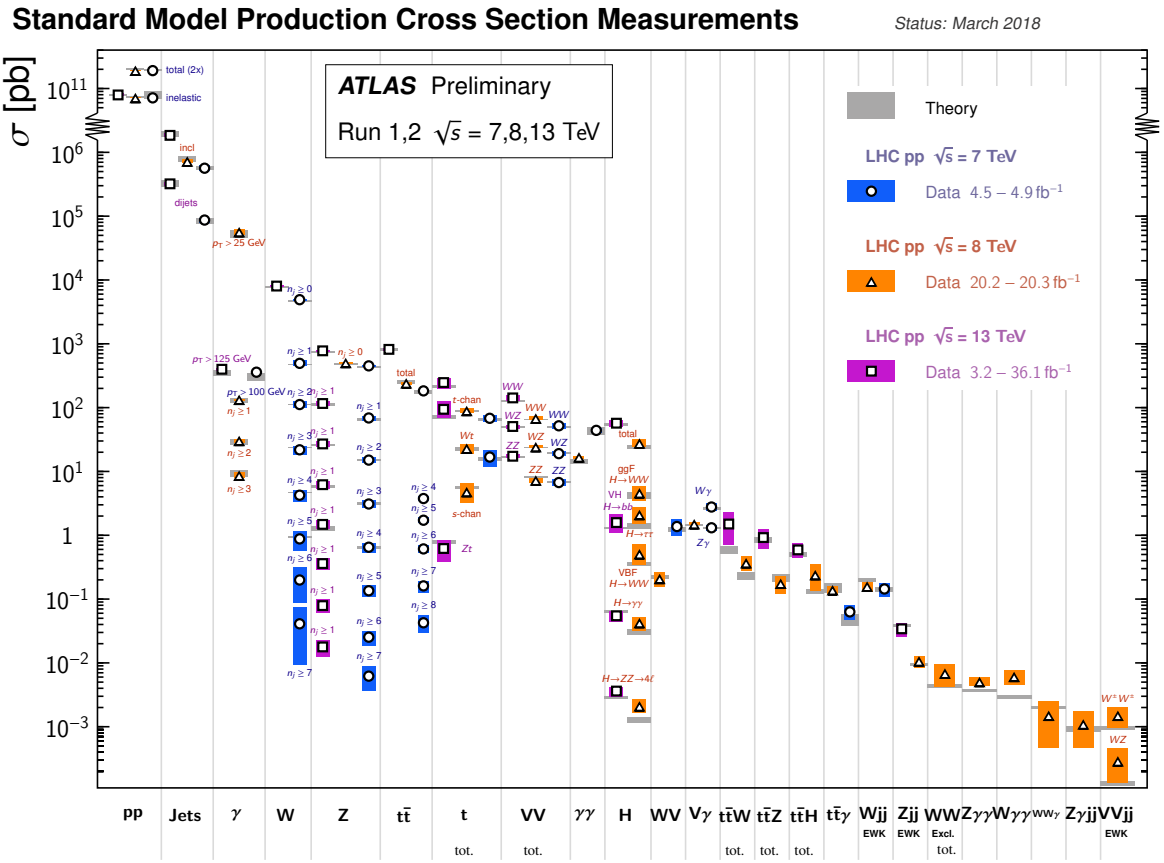


Figure 2.4: Summary of ATLAS cross section measurements of Standard Model processes [18]. The theoretical predictions and uncertainties are shown as grey bands, and the measurements in different markers and colours for different centre-of-mass energies of the LHC.

2.3 Simulating proton collisions

In modern particle physics it is useful to simulate various processes by numerical methods. The simulation uses Monte Carlo methods to generate "events" according to the probability densities given by the available phase space and transition amplitudes of the given processes. An event consists of a description of all final state particles in terms of their momenta. The number of particles in hadron

collisions is large[4] due to the hadronisation of partons and accordingly give each event many degrees of freedom for their momenta. In this high-dimensional phase space integration, the random sampling of numbers using MC methods is the most efficient solution [21]. In order to confidently achieve convergence of the integrals as well as sufficient population of events in all regions of phase space many events are generated e.g. routinely several hundreds of thousands or millions. Due to sampling failures and resampling of random numbers from the appropriate probability densities all events are weighted with appropriate event weights that maintain the differential and inclusive cross section normalisations. Simulated sets of events for a given process are called "Monte Carlo samples", or simply "MC" in particle physics jargon. This emphasises the numerical nature of the prediction which can suffer from low statistics.

As with the inclusive cross section calculation, MC generation factorises the proton structure from the interaction with high momentum exchange (hard scatter HS) using PDFs and matrix elements (ME). The matrix elements correspond conceptually to the short distance and high scale partonic cross section in Eq. 2.43.

The long distance evolution of the partons that are produced by the HS are described by parton shower (PS) algorithms. The PS evolves the partons from the high scale to the low scale of approximately 1 GeV. It does this by iterative parton branchings of the type $q \to qg$, $g \to q\bar{q}$ and $g \to gg$ governed by splitting functions that determine the momentum fraction of the radiated parton. Similarly to proton PDFs the splitting functions are evolved between scales by RGEs and in fact the PS can be conceptually imagined like a reverse PDF. Due to the iterative nature of the algorithm the splittings are self-similar at all stages and create the fractal-like structure of jets. The probability of the branchings is divergent for collinear and soft emissions. Collinear means that the angle between the original and the radiated parton is zero while soft means that the radiated parton carries no momentum away from the original one. Both types of divergences, sometimes jointly called infrared, are not physically observable. Therefore an infrared cutoff is introduced that ensures that a splitting is resolvable, i.e. has a large enough angle or momentum to be observable. The cutoff provides a stopping criterion for the PS algorithm where lower scale splittings are less likely to be resolvable and eventually stop [22]. Parton branching can also occur before the hard scattering interaction and depending on whether it occurs before or after, it is called initial or final state radiation (ISR and FSR).

To combine hard, large angle radiations from the ME and mostly soft, collinear radiation from the PS, matching and merging procedures have to be used. They also avoid that phase space regions are considered multiple times or not at all [23].

Hadronisation is the process of combining the showered partons into observable colour neutral hadrons. This step of the event generation is not calculated perturbatively but instead relies on phenomenological models of QCD. Two commonly used models exist, the string model and the cluster model [6][5].

The string model is implemented in the Pythia [24, 25] event generation program and is based on the concept of linear confinement. It assumes a potential between quarks that increases linearly with distance like a spring with tension of approximately 1 GeV/fm. Colour-connected pairs of quarks and antiquarks are interpreted as such elastic *strings*. During time evolution the strings break when the potential energy can create a new quark-antiquark pair which form the endpoints of new strings. Eventually, the strings are decayed directly into two hadrons with appropriate quark content.

The cluster model is based on the concept of preconfinement and is implemented in the event generators Herwig [26] and Sherpa [27]. Preconfinement implies that colour-singlet (colour neutral) systems of

[4] half a dozen charged particles per unit of pseudorapidity in events without high-p_{T} activity [19] and 8 to 20 charged particles per high-p_{T} jet [20]

[5] "Monte Carlo event generators" review

partons have a universal invariant mass distribution at low scales. *Clusters* are intermediate objects created from quark-antiquark pairs. Any gluons that are left over from the PS are split forcibly into such pairs. Clusters with invariant mass above a cutoff are split into two new ones and are then interpreted as excited mesons which decay to hadron pairs. Low-mass clusters may also decay into single hadrons.

The characteristic structure of gluon jets which have a softer hadron spectrum with larger multiplicities compared to quark jets can be understood intuitively in both models. In the string model, a gluon spans two coloured strings which create more breaks and in the cluster model, the $g \to q\bar{q}$ splittings are enhanced as the first step of the algorithm.

Both hadronisation models have parameters such as, the fractions of momentum transfers and choice of quark flavours and other quantum numbers in newly created quarks, that are fitted to experimental data. The fitted parameters depend on the cutoff scale and the specific implementation of the PS. So several sets of best-fit parameters ("tunes") exist for the different event generators.

Finally, all unstable particles produced in the described processes are decayed to stable ones or those that have lifetimes long and momentum/mass ratios large enough to be stable in the reference frame of a detector, such as electrons, muons, neutrinos, photons, kaons and charged pions. The stable or metastable particles are what is observable to the detectors.

It is observed that collision events with a hard scattering have an elevated level of particle production. This pedestal effect occurs even spacially away from the hard particles and is called underlying event (UE). One explanation for, and way to simulate, the effect is multiple parton interaction (MPI) of the remaining partons in the protons that participated in the hard scattering event. The selection of events with hard scattering causes a bias for such higher levels of UE.

On the other hand, one can consider minimum bias events without the presence of hard interactions. This kind of event consists of elastic as well as inelastic and diffractive interactions of protons. When they happen simultaneously to a hard interaction they are called pileup events. Pileup can be simulated by MC and "added" to the HS event. It is also possible to record minimum bias data with the respective detectors for which MC is being produced and overlay those data events on the MC HS event. The latter is not yet routinely used by high energy physics experiments.

The main aspects and components of MC event simulation are sketched in Fig. 2.5. The figure is also a schematic representation of the factorisation theorem of PDFs and partonic cross sections.

2.4 Higgs boson physics

The Higgs boson has a rich and diverse variety of experimentally observable signatures that are determined by its production and decay modes. It interacts with all massive particles and even the massless ones via loops. So it is not surprising that a particle that touches every other known particle provides a rich field of study. The most important properties of the Higgs boson are described in the next sections.

2.4.1 Production modes

The major production modes of the Higgs boson at the LHC in order of abundance are gluon-gluon fusion (ggF), massive vector boson fusion (VBF), associated vector boson production (VH) and associated heavy quark pair production ($t\bar{t}H$ and $b\bar{b}H$). The inclusive cross sections of these processes are shown in Fig. 2.6 where also the order of the perturbative expansion of the calculations and their uncertainties are indicated.

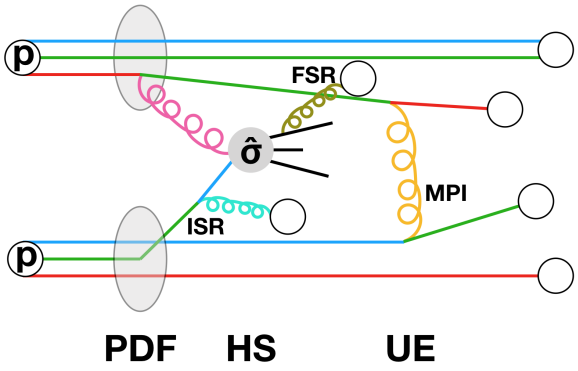


Figure 2.5: Schematic drawing of a proton–proton collision as it relates to the factorisation theorem and Monte Carlo event generation. The partonic cross section $\hat{\sigma}$ corresponds to the hard scattering (HS) interaction and factorises from the parton distribution function (PDF) of the incoming protons (p). Initial and final state radiation (ISR/FSR) can be considered part of the parton shower (PS) of MC generators with only one exemplary splitting each. The underlying event (UE) refers to any other increased activity in the presence of a HS event e.g. by multiple parton interactions (MPI). White circles indicate observable colour neutral hadrons.

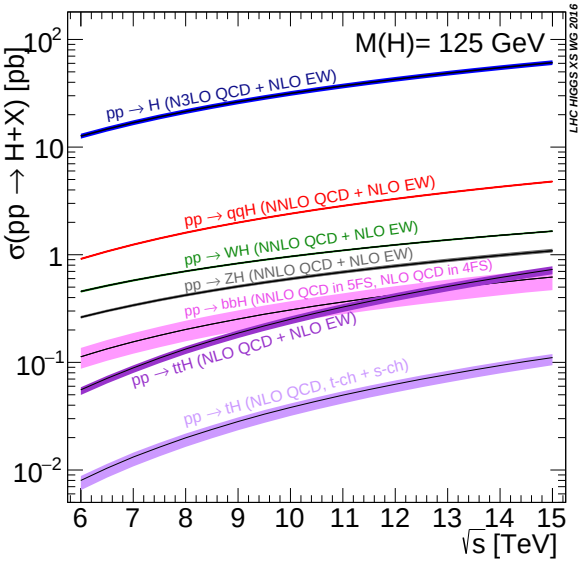


Figure 2.6: Cross sections of Higgs boson production modes as a function of the pp collision centre-of-mass energy [28]. The bands indicate the theoretical uncertainties.

Gluon-gluon fusion

Gluon-gluon fusion (ggF) is the most likely production mechanism of the Higgs boson at the LHC. Since the gluons are massless, the interaction with the Higgs boson proceeds via heavy quark loops, to which both of the bosons couple. Thus the lowest and leading order (LO) diagram already includes a loop (see Fig. 2.7(a)). The total ggF cross section has been computed at NNNLO[6] in perturbative expansions of QCD with NLO electroweak (EWK) corrections. The convergence of the computations with additional orders is slow. The ratios between higher and lower order calculations are called k-factors e.g.,

$$k_{\mathrm{NLO}} = \frac{\sigma_{\mathrm{NLO}}}{\sigma_{\mathrm{LO}}}, \tag{2.44}$$

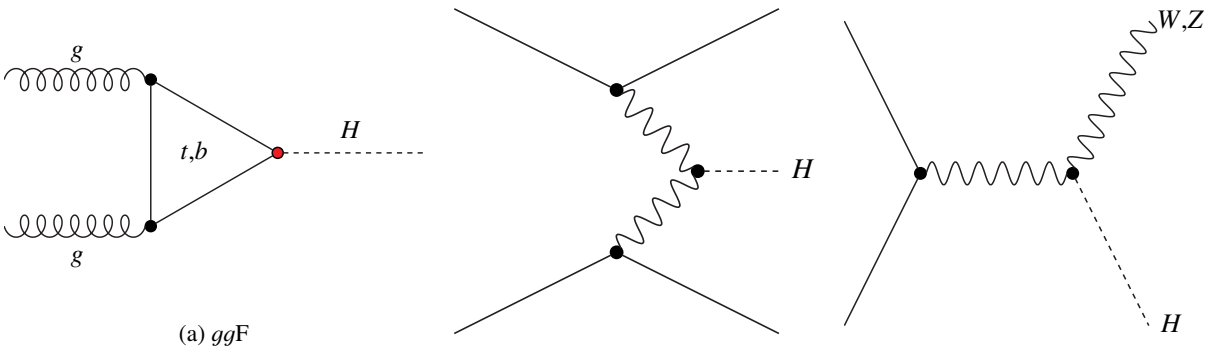
and are expected to converge to one when sufficient orders are considered in the calculations. In case of ggF the NLO and NNLO k-factors are large at approximately 2 and 1.4, respectively. The uncertainties due to QCD scale variations are 5-7 % and due to PDF and α_s are 3.2 % [29].

Vector boson fusion and vector boson associated production

Weak vector boson fusion (VBF) is the second most likely Higgs production mode. It has a characteristic signature of two additional quarks that fragment and hadronise into two jets which have large rapidities and a large rapidity gap between them. Additionally, at leading order there is no colour flow between them, so QCD radiation in the gap is suppressed. An example of a Feynman diagram for the VBF process is shown in Fig. 2.7(b).

Higgs production in association with a weak vector boson (VH) is also called Higgsstrahlung. The Feynman diagram of VH shown in Fig. 2.7(c) resembles Drell-Yan production of $W^{\pm}$ and Z bosons with the radiation of a Higgs boson. The accompanying vector boson also provides a unique signature which can be exploited by reconstructing or tagging its leptonic decays.

The cross sections of both VBF and VH have been computed at NNLO in QCD with NLO EWK corrections [29].



(a) ggF (b) VBF (c) VH

Figure 2.7: Examples of Feynman diagrams for Higgs production in gluon-gluon fusion, massive vector boson fusion and massive vector boson associated modes.

[6] each N is spoken as “next-to-” and means that one additional perturbative expansion term is included

Top quark pair associated production

Higgs production in association with a heavy quark pair is a rare process. Due to the direct coupling between the quarks and the Higgs the cross section is proportional to the quark mass which makes the top quark the primary participant in this production process. For top quarks this process is the main way to determine the top-Higgs coupling as direct $H \to t\bar{t}$ decays are kinematically forbidden. Several examples of Feynman diagrams for the $t\bar{t}H$ process are shown in Fig. 2.8. The $t\bar{t}H$ production mode is the main focus of this thesis.

The cross section of $t\bar{t}H$ production has been computed at NLO in QCD and electroweak couplings to be 507 fb for the LHC operating at $\sqrt{s} = 13\,\text{TeV}$ [29–33]. Both QCD scales are set to the same value $\mu_0 = \mu_R = \mu_F = m_t + m_H/2$, where the Higgs mass and top quark masses are taken as 125.09 GeV and 172.5 GeV. The associated uncertainties are derived by varying the scales independently by factors 1/2 and 2 while also keeping the ratio μ_R/μ_F between 1/2 and 2, for a total of four variations. The combined scale uncertainty is $^{+5.8}_{-9.2}\%$. The chosen value for the scales is motivated by the fact that the uncertainties from these variations are smallest at this value. Uncertainties from the PDF and strong coupling α_s are 3.6 %.

The NLO QCD correction increases the cross section by 25 %. The NLO EWK contributions spoil the simple proportionality between the cross section at the top-Yukawa coupling because they introduce Higgs to vector boson and Higgs self-couplings rather than top-Higgs couplings. Fortunately the size of the correction is small at less than 2 % and thus is smaller than the scale uncertainties.

All theoretical uncertainties and corrections of the cross section that are described above are summarised in Fig. 2.9 which shows $\sigma_{t\bar{t}H}$ as a function of m_H in the vicinity of the known measured $m_H = 125.09\,\text{GeV}$.

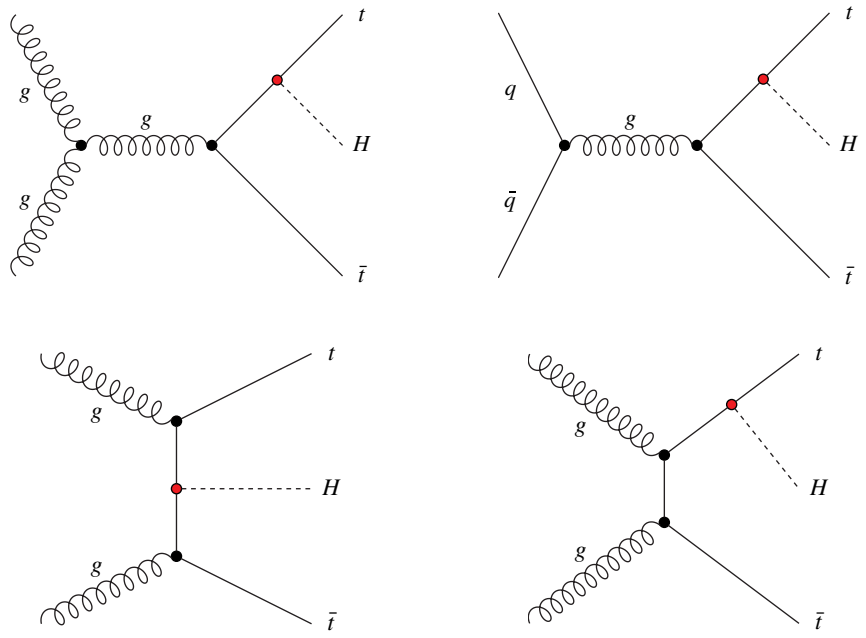


Figure 2.8: Examples of Feynman diagrams of $t\bar{t}H$ production at LO.

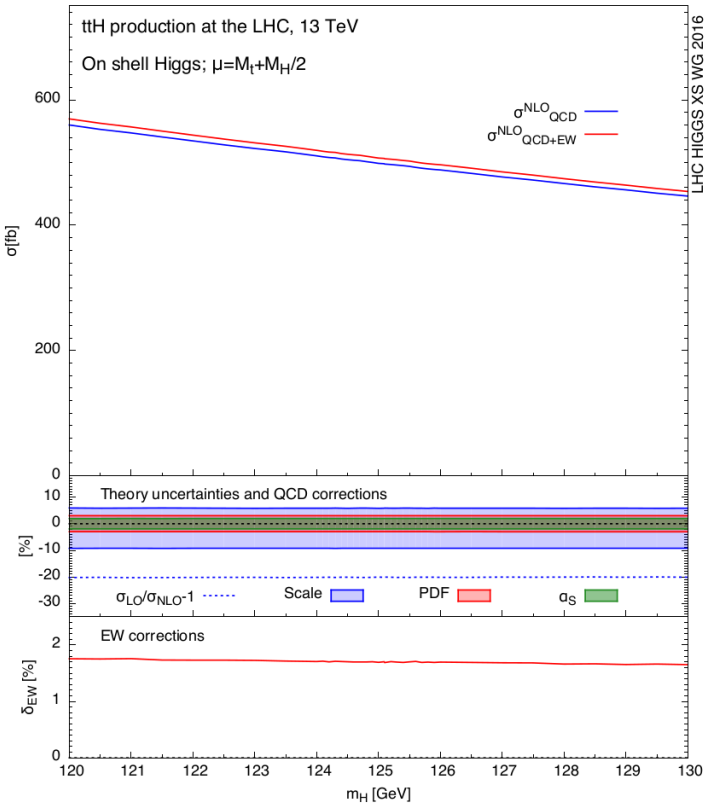


Figure 2.9: Predicted $t\bar{t}H$ cross section as a function of m_H for pp collisions at $\sqrt{s} = 13\,\text{TeV}$ [29]. The upper panel shows the cross section with and without the NLO EWK corrections in fb. The middle panel shows the size of the scale, PDF and α_s uncertainties with blue, red and green bands, as well as the size of the NLO QCD correction as dotted line in %. The bottom panel shows the size of the NLO EWK correction in %.

2.4.2 Decay modes

The Higgs boson decays to all massive particles directly and to massless particles via loops of massive particles. As the coupling is proportional to the particle mass, the branching ratios to heavy particles are large if the decay is kinematically allowed. The Feynman diagrams for the different three kinds of decays, to fermions, to vector bosons and via loops, are shown in Fig. 2.10.

In order of prevalence the Higgs decay modes are $b\bar{b}$, WW^*, gg, $\tau\bar{\tau}$, $c\bar{c}$, ZZ^*, $\gamma\gamma$, $Z\gamma$, $\mu\bar{\mu}$ and more. Of these $Z\gamma$ and $\mu\bar{\mu}$ are experimentally still inaccessible due to small branching fractions, while gg and $c\bar{c}$ are to date out of reach of measurement due to overwhelming jet background. The values of branching fractions of the decays that are relevant for $t\bar{t}H$ searches are shown in Fig. 2.11 for a Higgs mass of $m_H = 125\,\text{GeV}$. Relative uncertainties on these predicted branching fractions are between 1-2% in size and arise from missing orders in the calculations as well as propagated uncertainties of the input parameters (quark masses and α_s) [29]. The parametric uncertainties introduce correlations between the decay modes. There is also the trivial correlation that all fractions have to add to unity.

The Higgs boson was first discovered in the ZZ^*, $\gamma\gamma$ and WW^* decays with subsequent charged leptonic decays of the weak bosons by the two large competing experiments at the LHC, ATLAS [34] and CMS [35]. Despite the small branching fractions of $ZZ^* \to 4\ell$ and $\gamma\gamma$ these channels proved to be sensitive by reconstructing the decay mother with the good energy/momentum resolution of photon, electron and muon signatures in the detectors.

Establishing experimental evidence for the fermionic Higgs decays was more challenging due in large parts to the hadronic nature of the final states. Both the $b\bar{b}$ and $\tau\bar{\tau}$ decays present themselves with jets in the final state, the b-quarks by taggable b–jets and the τ leptons mostly by reconstructable hadronic decays. To reduce the ever-present jet background, these two fermionic analyses relied on the Higgs production modes with additional taggable signatures VH and VBF. So the challenge in comparison to the bosonic decays was twofold, in the smaller accessible production cross section and increased hadronic backgrounds. Nevertheless strong evidence[7] for the $b\bar{b}$ decay and a significant observation of the $\tau\bar{\tau}$ decay was found by ATLAS [36, 37] and CMS [38, 39].

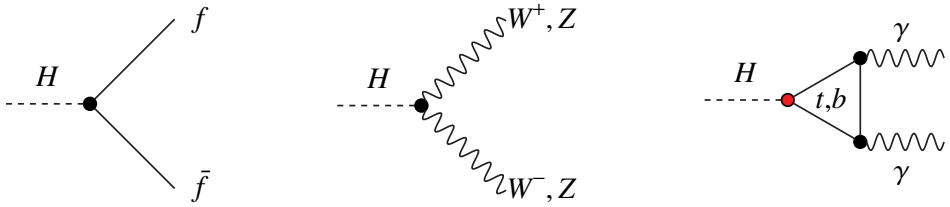


Figure 2.10: Examples of Feynman diagrams for Higgs decay into a fermion pair, a massive vector boson pair and a photon pair. The $H \to \gamma\gamma$ diagram also exists with a W loop.

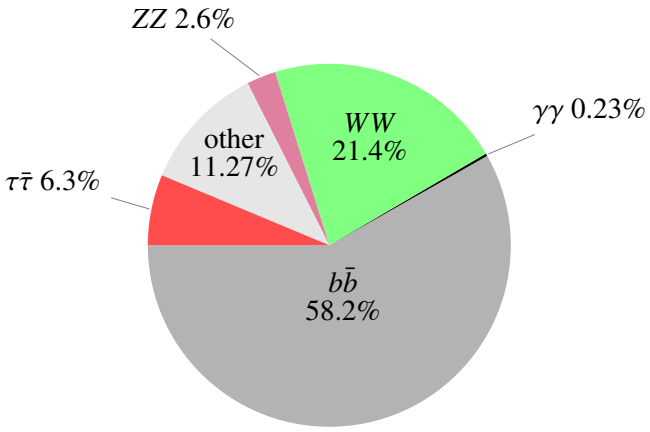


Figure 2.11: Fractions of decay modes of a Higgs boson with $m_H = 125\,\text{GeV}$ [29]. Those relevant to the $t\bar{t}H$ searches are labelled by name, the rest are combined in "other" and include the remaining fermionic decays and the loop-induced bosonic gg and $Z\gamma$ decays. Relative uncertainties are less than 2 % in all named cases and omitted for readability.

2.4.3 Yukawa couplings and other properties

The experimental results that establish the existence of the $H \to b\bar{b}$ and $H \to \tau\bar{\tau}$ decays simultaneously also provide a measurement of the b- and τ-Yukawa couplings. The results are consistent with the expectations derived from the Standard Model. Searches for decays into fermions exist for second generation quarks and leptons, specifically $H \to c\bar{c}$ and $H \to \mu\bar{\mu}$, e.g. by ATLAS [40, 41]. However, the results are not statistically significant due to small statistics and large backgrounds. Improvements to these analyses and more data can eventually lead to evidence for these Higgs decays if they behave as expected.

[7] Conventionally "evidence" refers to a significance of 3σ and "observation" to 5σ. See Sec. 7.2 for a definition of significance.

Direct measurements of the top-Yukawa coupling are accessible by measuring the $t\bar{t}H$ cross section. However, it is also possible to infer knowledge about it indirectly. All the Feynman diagrams shown in this section have vertices marked in red colour where a top-Yukawa coupling can appear. That is obviously the case in all $t\bar{t}H$ diagrams (Fig. 2.8), but also in the loop induced $gg \to H$ (Fig. 2.7(a)) and $H \to \gamma\gamma$ diagrams (rightmost in Fig. 2.10) as the two major examples where a top loop is an important contribution to the total amplitude.

The loop contributions can be parametrised in the κ-framework which scales Higgs production cross sections and decay widths based on the participating particles [42]. It assumes the narrow width approximation so that production and decay factorise, allowing for scaling each separately. For example, κ_t, the top coupling modifier, changes the $t\bar{t}H$ cross section as

$$\sigma_{t\bar{t}H} = \kappa_t^2 \cdot \sigma_{t\bar{t}H}^{\mathrm{SM}}, \tag{2.45}$$

so that a value of $\kappa_t = 1$ recovers the SM prediction. This convention is for all other κ as well.

Fits of κ-values were performed to all Higgs measurements performed by ATLAS and CMS with data from the first period of data taking of LHC (Run 1) [42]. The gluon and photon loops can be resolved in terms of top, bottom and W coupling modifiers as

$$\sigma_{gg\mathrm{F}} = \sigma_{gg\mathrm{F}}^{\mathrm{SM}} \cdot \left(1.06 \cdot \kappa_t^2 + 0.01 \cdot \kappa_b^2 - 0.07 \cdot \kappa_t\kappa_b\right) \tag{2.46}$$

$$\Gamma_{\gamma\gamma} = \Gamma_{\gamma\gamma}^{\mathrm{SM}} \cdot \left(1.59 \cdot \kappa_W^2 + 0.07 \cdot \kappa_t^2 - 0.66 \cdot \kappa_W\kappa_t\right), \tag{2.47}$$

where top-bottom and top-W interference account for the negative mixed terms. The above equations clearly assume that the loops only contain the known SM particles but no potentially existing new particles predicted by theoretical models beyond the Standard Model (BSM). This assumption is reasonable as the best-fit modifier values κ_g and κ_γ that treat the ggF and $H \to \gamma\gamma$ loops as effective interactions are found to be compatible with the expectation in the SM. For the fit of the effective modifiers fewer assumptions about BSM contributions are necessary.

The best-fit value of $\kappa_t = 0.87 \pm 0.15$ is compatible with the expectation of 1, also shown in Fig. 2.12. One might argue that this fact is not surprising given the observed measurements of ggF and $H \to \gamma\gamma$ interactions and the strong assumptions made in the fit. The relative uncertainty of the results of approximately 17 % can be of more interest. For a significant improvement of this uncertainty a significant measurement of the $t\bar{t}H$ cross section is necessary.

Another property of the Higgs boson that has been measured to very high precision is its mass. A value of

$$m_H = 125.09 \pm 0.24\ \mathrm{GeV} \tag{2.48}$$

is achieved by combining measurements that reconstruct Higgs bosons in $H \to \gamma\gamma$ and $H \to 4\ell$[8] events recorded by ATLAS and CMS in Run 1 of LHC [43]. The good energy/momentum resolution of reconstructed photons, electrons and muons allows for the high precision in the mass measurement.

The spin and parity quantum numbers of the Higgs boson were examined in bosonic decays (WW^*, ZZ^* and $\gamma\gamma$) [44, 45] and in the bosonic production mode VBF tagged by $\tau\bar{\tau}$ decays [46]. The analysis of decays tested and rejected several alternative hypothesis in which the Higgs boson has spin and parity quantum numbers different from the ones predicted in the SM, scalar with even parity. The analysis of the VBF production tested for admixtures of pseudoscalar components to the observed Higgs boson and also found compatibility with the SM prediction within its confidence interval.

The expected zero charge of the Higgs boson is trivially confirmed by several observations of its decay

[8] $\ell = e, \mu$

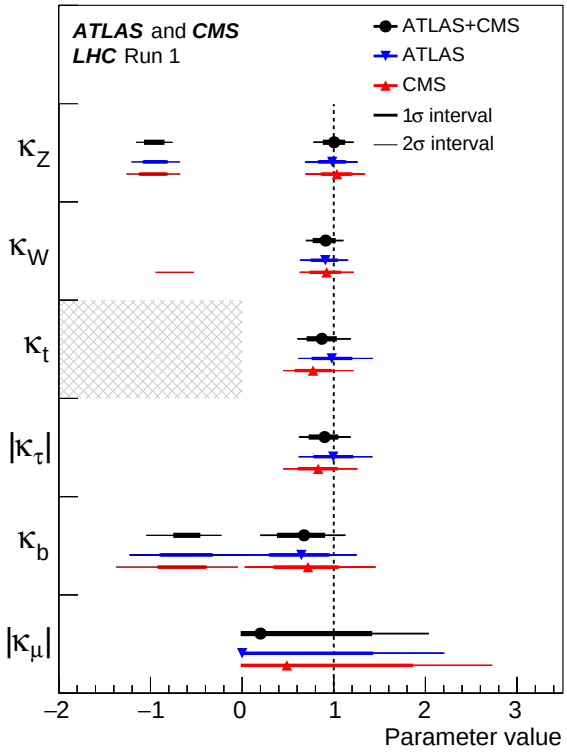


Figure 2.12: Higgs coupling modifiers defined in the κ-framework [42]. Results shown are the combination of all Higgs analyses performed by ATLAS and CMS on the full dataset of the first data-taking period (Run 1) of the LHC. The grey hatched area indicates an arbitrary choice of sign for κ_t for which only the relative sign to κ_b and κ_W are meaningful.

into neutral final states.

In summary, all properties of the Higgs boson that are measured and for which predictions exist in the SM are consistent with those predictions at the current experimental sensitivities.

2.5 Top quark physics

The existence of the top quark was experimentally discovered in collision events of the Tevatron recorded and analysed by the CDF [47] and D0 [48] collaborations in 1995. Its existence was predicted and expected since the 1970s. The explanation for CP violation in quark mixing relies on a complex phase in the quark mass matrix, V_{CKM}. The minimal size of the matrix for which such a phase is possible is 3-dimensional, which are interpreted as quark generations. The discovery of a fifth quark, the bottom quark, in its quarkonium resonance state [49] and the τ lepton [50] opened the third generation of fermions. This solidified the expectation that the top quark should exist to complete the isospin doublet of the b-quark.

The current world average value of the top quark mass that combines measurements from the Tevatron and LHC [6] is

$$m_t = 173.0 \pm 0.4 \text{ GeV}, \tag{2.49}$$

where the component of the uncertainty that is attributable to systematic sources is larger than the

statistical one. Due to this large mass, several orders of magnitude larger compared to all other fermions, the experimental signatures of the top quark are different than the other quarks'. The mass is above the threshold for the decay $t \to Wb$ with an on-shell W boson. As a consequence the top quark's lifetime is too short to form hadronic bound states and thus the weak decay is the dominant and practically only decay mode.

The dominant production mechanism of top quarks at the LHC is pair production $gg \to t\bar{t}$ with approximately 10 % contributions from $q\bar{q} \to t\bar{t}$ [6][9]. The inclusive cross section in pp collisions of $\sqrt{s} = 13$ TeV is predicted with NNLO terms in QCD to be

$$\sigma_{t\bar{t}} = 831.8^{+19.8}_{-29.2} \pm 35.1 \text{ pb} \tag{2.50}$$

with uncertainties arising from QCD scales and PDFs [51]. Decays of top-quark pair events are categorised by the decay of the W boson from each of the $t \to Wb$ decays. Depending on the number of $W \to \ell\nu$ an $W \to qq'$ decays, top-quark pair decays are called dileptonic, lepton+jets and allhadronic. The branching fractions of these categories are shown in Fig. 2.13, where one can see that the dileptonic mode is the least and the other two approximately equally likely. The characteristic experimental signatures of $t\bar{t}$ events are 0 to 2 isolated leptons and missing momentum, 0 to 4 jets initiated by the W decays, and 2 b–jets initiated by the b–quarks.

Combining the $t\bar{t}$ decay branching fractions with the H boson decay branching fractions yields the experimental signatures of $t\bar{t}H$ events. The categorisation of $t\bar{t}H$ events is described in Chapter 4.

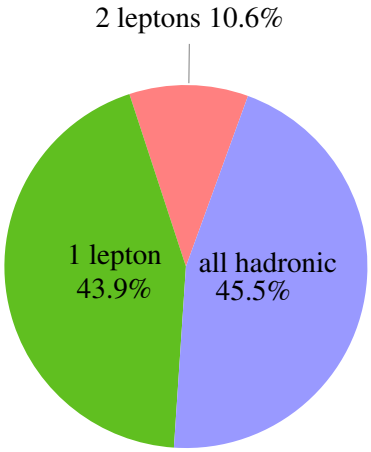


Figure 2.13: Fractions of top-quark pair decay modes [6]. Each top quark decays to Wb and the W decays are indicated in the figure. Relative uncertainties are less than 1 % in all cases and omitted for readability.

2.6 Tau lepton physics

Tau leptons have a mass of

$$m_\tau = 1776.86 \pm 0.12 \text{ MeV}, \tag{2.51}$$

making them the most massive leptons in the SM [6]. Like its older sister, the muon, the tau lepton only decays weakly to other less massive leptons, but due to its relatively large mass also to hadrons. An example of a Feynman diagram of the decay is shown in Fig. 2.14. In the figure ℓ stands for electron or

[9] "The Top Quark" review

muon. The tau neutrino conserves the tau lepton number and is not directly observable. Considering that the two quarks can appear in three colour combinations, there are five diagrams that contribute to the decay process. Naively, one would expect a decay to e and μ in a fraction of $2/5$ of times and a decay to hadrons in $3/5$ of times [52]. This intuition matches well with the measured branching fractions shown in Fig. 2.15. When recognising that the decay diagrams are just W boson decays with the additional neutrino, the naive expectation also matches the known W branching ratios when corrected for the kinematically allowed decay products.

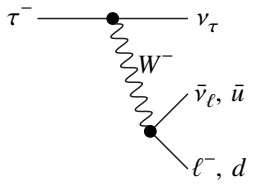


Figure 2.14: Example of Feynman diagram of tau lepton decay.

There are several different hadronic[10] decay modes of a tau lepton. The most common ones are listed in Tab. 2.4. The two broad characteristic categories are the decays to only one charged hadron or three charged hadrons through intermediate light scalar or vector mesons (η, ρ, ω,...). Decays to more than three exist but are rare. As an example for a tau decay Fig. 2.16 is showing the decay to three charged and one neutral pion. A description of how the occurrence of a tau decay is inferred from the hadron signatures is given in Sec. 3.4.7. With an average proper decay length of $c\tau = 87.03\,\mu\text{m}$ [6], tau leptons can travel measurable fractions of millimetres in detectors which is one of the experimental signatures used to identify its decays.

Decays to strange mesons (kaons) are rare because the transition amplitude in the corresponding $W \to us$ decay is proportional to the CKM matrix element [53]

$$|V_{us}| = 0.22508^{+0.00030}_{-0.00028}, \tag{2.52}$$

and the probability for it to occur is suppressed by the square of its value. Inversely, measuring the branching ratios of decay modes involving kaons allows to infer knowledge about the strange quark.

Measurements of the hadronic spectra allow for precision measurements in perturbative QCD [54]. One example is the determination of the strong coupling constant at the scale of the tau lepton mass $\alpha_s(m_\tau)$ as already seen in Fig. 2.2. The relevant measurements were performed at lepton colliders where there are little to no hadronic backgrounds.

At a hadron collider such as the LHC, precision QCD measurements of and with hadronic tau lepton decays are hindered by the necessity to distinguish them from quark/gluon jets. At such machines, the main merit of the tau lepton is its large mass. It is used mainly in searches for processes with couplings proportional to mass. Indeed the first evidence for the existence of a Higgs interaction of Yukawa type, i.e. with fermions, was found with tau leptons [55, 56].

Furthermore, tau leptons provide a third lepton flavour that opens search channels for processes that might violate the conservation of the lepton quantum numbers which is not foreseen in the SM. Or, as is the case in the $t\bar{t}H$ analysis, it simply increases the acceptance to a rare process.

The fact that the tau lepton is so much heavier than the other leptons which were originally named after the greek word for small or light, often leads to the use of the oxymoron heavy lepton to signify the tau lepton and the tautology light lepton to signify electrons and muons. Nevertheless, the latter phrase is

[10] Semileptonic decay would be more accurate as there is also the tau neutrino among the daughter particles. However, the neutrino is always present. Thus hadronic decay is used to refer to a decay to hadrons.

used in this thesis as it is useful to refer to electrons and muons with a single word when the difference is irrelevant.

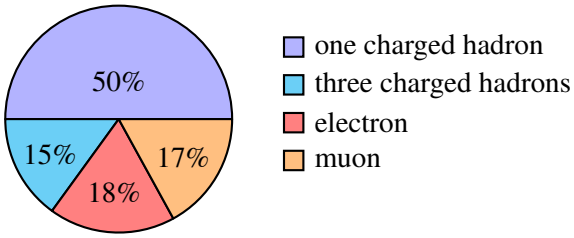


Figure 2.15: Dominant tau lepton decay modes [6]. Relative uncertainties are less than 1 % in all cases and omitted for readability.

Decay mode	Branching fraction in %
h^-	11.51 ± 0.05
$h^-\pi^0$	25.93 ± 0.09
$h^- \geq 2\pi^0$	10.81 ± 0.09
$h^-h^-h^+$	9.80 ± 0.05
$h^-h^-h^+ \geq 1\pi^0$	5.29 ± 0.05

Table 2.4: Branching fractions of major hadronic decay modes of the tau lepton [6]. $h^\pm$ signifies a charged hadron which is a pion in most cases. Tau neutrinos are present in all decays and omitted from the table.

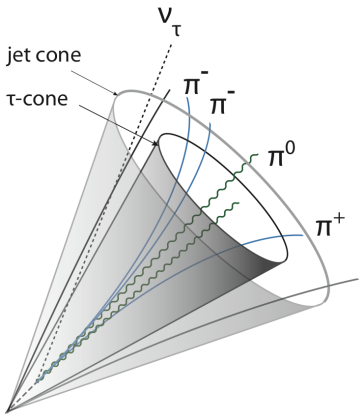


Figure 2.16: Drawing of a tau decay to three charged and one additional neutral pion [57]. The cones are instrumental in its reconstruction. All pions and the neutrino originate from the displaced vertex where the tau decayed. The neutral pion is observable by the two photons it decays to. Other tracks not related to the tau decay are background.

CHAPTER 3

ATLAS detector at the LHC

3.1 LHC

The large hadron collider (LHC) is a particle accelerator and storage ring for protons and heavier ions (Xe, Au, Pb) [58]. It is located at CERN near Geneva, Switzerland, in a tunnel about 100 m underground and has a circumference of 27 km, making it the largest machine in the world. Protons are accelerated in bunches by a chain of smaller accelerators at CERN. In the LHC they are further accelerated by RF cavities up to an energy of 6.5 TeV. Several hundred bunches with spacings of at least 25 ns make up beams of protons that circulate around LHC in both directions. The beams are kept on the circular path by superconducting dipole magnets and are made to cross in four points around the ring where the resulting proton-proton collision can have a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$[1]. There was a period of data taking from 2010–2012 with a lower energy of 7–8 TeV and there are plans for raising it to 14 TeV.

The four detectors that measure the collisions in each interaction point are: ATLAS [59], CMS [60], LHCb [61] and ALICE [62]. The two general purpose detectors, ATLAS and CMS, are designed to discover and measure the Higgs boson and search for new physics. The delivered luminosity to these experiments is as high as possible. LHCb is specialised in measuring B-hadron decays, while ALICE is focused on measuring the heavy ion collision data. Figure 3.1 shows a drawing of the LHC, the chain of injecting accelerators and the four interaction points with the corresponding detectors.

The beam intensities are expressed as a luminosity L which acts as a proportionality factor between the cross section of a process σ and its corresponding interaction rate $\dot{N}$,

$$\dot{N} = L \cdot \sigma. \tag{3.1}$$

To summarise a period of data taking one can quote the time-integrated luminosity which can be interpreted as a total amount of interactions. To achieve the highest possible luminosities the proton beams are focused by multi-pole magnets when they are made to cross inside ATLAS. A consequence of it, is that in each bunch crossing more than one proton–proton collision occurs. One makes a distinction between collisions with large momenta transfers which likely include physical processes of interest and other collisions called “pileup”. The number of interactions follows a Poisson distribution with mean

$$\langle \mu \rangle = \frac{L \cdot \sigma_{\text{inelastic}}}{f_{\text{LHC}}}, \tag{3.2}$$

where f_{LHC} is the revolution frequency of the beams in the LHC. Figure 3.2 shows that during 2015–2016

[1] $s = (p_1 + p_2)^2$, p are the 4-momenta of the incoming protons, s is known as a Mandelstam variable.

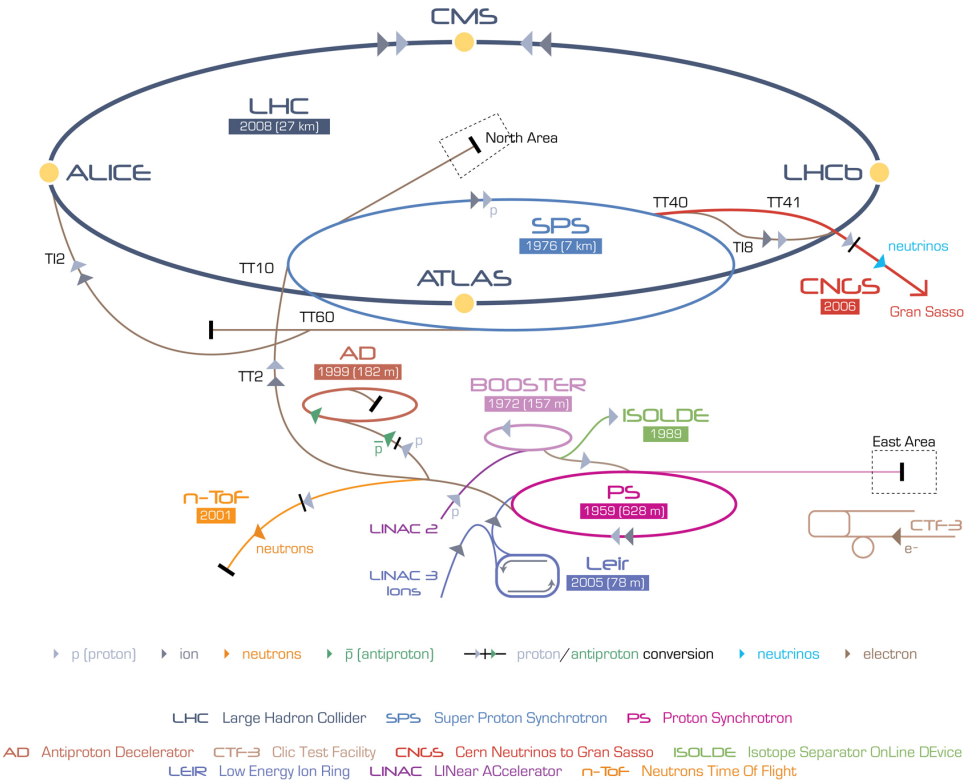


Figure 3.1: Drawing of the LHC, its four interaction points with detectors, and its injection chain. [63]

the mean number of interactions per bunch crossing in ATLAS was 24. Additionally, even collisions from preceding and subsequent crossings can contribute to pileup in sub-detectors that have signal lengths and readout times longer than the 25 ns between bunches.

Beams that are not made to cross in any interaction point, can circulate within the LHC almost indefinitely. However, when they do cross the colliding protons are removed from the beams and their intensities decay exponentially. In the years 2015–2016, the half life of crossing beams was about 10 h, which in ideal conditions lead to a useful time of roughly 15 h of collisions before the LHC had to be refilled with replenished beams.

3.2 ATLAS

ATLAS is a general purpose detector that can measure particles in almost the full solid angle [59]. The detector has a cylindrical shape of 44 m in length and 25 m in diameter and is located at the interaction point (IP) of the LHC that is closest to the CERN Meyrin site. Its main components are a tracking detector that measures the trajectories and momenta of charged particles, calorimeters that measure the energies of particles that interact electromagnetically and strongly, tracking detectors that measure muon trajectories and momenta, as well as magnets that bend the trajectories of charged particles. With these ATLAS can measure all types of particles that can be produced in high energy collisions except weakly interacting, electrically neutral neutrinos.

The coordinate system that is used throughout has its origin in the middle of the detector at the nominal IP. The z–axis coincides with the beam line and is positive in the southern direction. The positive x–axis points to the centre of the LHC ring and the positive y–axis is upwards. Due to the cylindrical symmetry

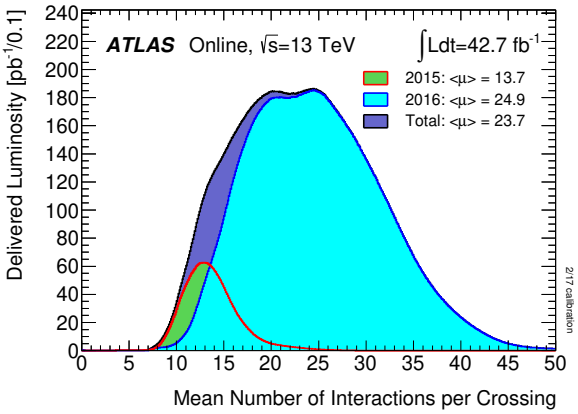


Figure 3.2: Mean number of interactions per crossing for the 2015–2016 pp collision data delivered to ATLAS at 13 TeV centre-of-mass energy [64].

of the detector it is convenient to use spherical coordinates. The azimuthal angle ϕ is measured from the x–axis clockwise around the z–axis. The polar angle θ is measured from the positive z–axis and is transformed to pseudorapidity

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right) \tag{3.3}$$

due to the property that a difference in pseudorapidity is invariant under Lorentz transformations along z. Thus the distance in $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ is also invariant under such transformation. In this coordinate system a particle's 4-momentum that originates in the IP is expressed as $(p_{\mathrm{T}}, \eta, \phi, E)$, where p_{T} is the momentum component in the plane perpendicular to z and E is the particle's energy.

Most detector components are divided into a central part with low values of η called *barrel* and parts at each end of the cylinder with large values of η called *end caps*. A drawing of ATLAS that contains labels for the major components is shown in Fig. 3.3. The components are described in the following sections.

Inner tracking detectors

The inner detector (ID) is located closest to the beam pipe and the collisions. Its function is to precisely measure the trajectories of charged particles within $|\eta| < 2.5$. It uses semiconducting sensors which are segmented rectangularly (pixels) and into strips (SCT) that detect electron-hole pairs, as well as a gaseous detector (TRT) that detects ionisation and transition radiation.

The **pixel** detector is the first sub-detector that a particle encounters. It consists of four cylindrical layers in the barrel part and three disks in the end caps and fits entirely inside a volume of $\pi \cdot (15\,\mathrm{cm})^2 \cdot 130\,\mathrm{cm}$. The innermost layer is located at a distance of only 3.3 cm to the IP and was installed as an upgrade in 2013–2014 after the first period of LHC data taking. The reduced distance to the IP and the increased number of measurement points for trajectories have improved the resolution of reconstructed vertices in the second period of LHC data taking. The spacial resolution of the pixel detector is 10 μm in the r–ϕ direction and 115 μm in the other direction (z for barrel, r for disks) [59].

The SemiConducting Tracker (**SCT**) is segmented into 12 cm long strips with a distance of 80 μm between strips. Each module is assembled from two layers of strips at an angle of 2.3 degrees to allow measurement of a two-dimensional space point. The SCT is assembled in four cylindrical layers between radii 30 cm $< r <$ 51 cm from the IP and nine disks on both sides extending to $|z| = 2.7$ m. The intrinsic accuracy of the SCT is 17 μm in r-ϕ and 580 μm in z (r) for barrel (disk) modules.

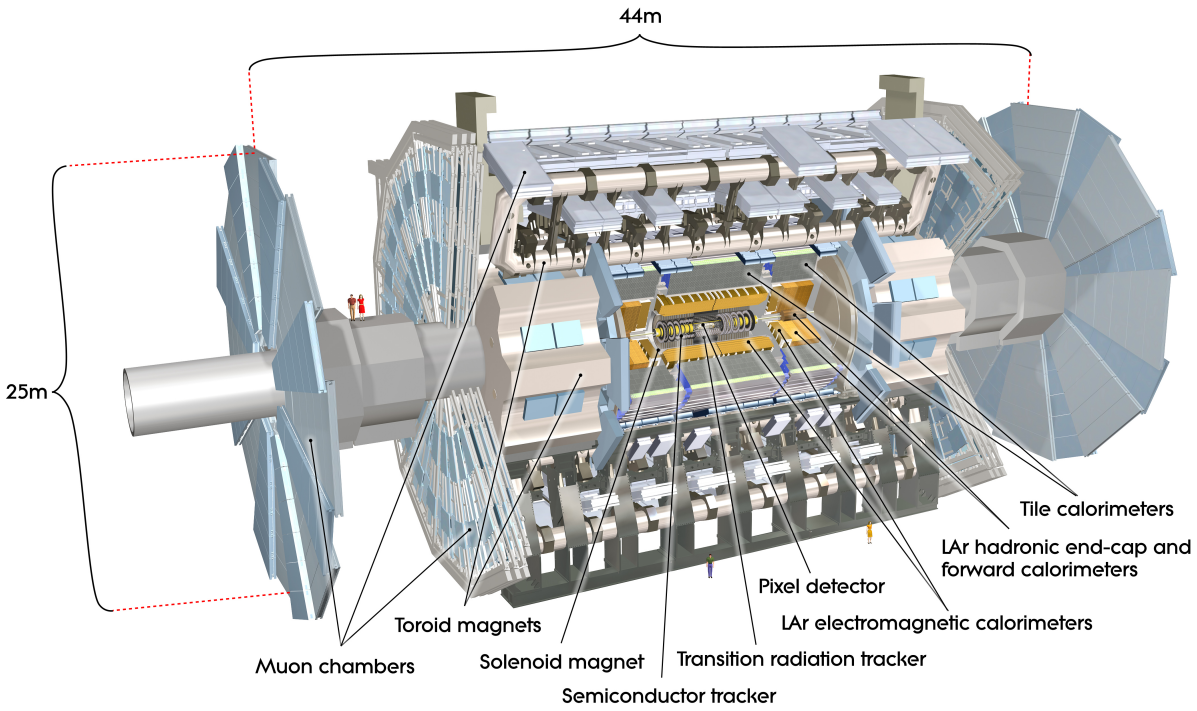


Figure 3.3: Drawing of the ATLAS detector [65]. Cutaways reveal the inner parts of the detector. The main components used to measure particles are labelled. Four human figures are drawn to help appreciate the scale of the size.

Both the pixel and SCT detectors rely on measuring electron-hole pairs that are created by traversing charged particles.

The Transition Radiation Tracker (**TRT**) provides measurement of trajectories in a range of $|\eta| < 2$. It is made of tubes filled with a xenon, CO_2 and oxygen gas mixture that have anode wires in their centre. The operating principle is the collection of charges that are created through ionisation by passing charged particles. The electric potential between cathode (tube wall) and anode cause the ions and electrons to drift in opposite directions. While drifting the electrons can create secondary ionisation. Close to the wire which has a small diameter the radial electric field is strong and creates an avalanche of secondary electrons. The number of electrons is amplified by four orders of magnitudes by this process. The position of where the charged particle passed is calculated from the measured drift time of electrons and the known drift velocity. Thereby the position is constrained to a circle around the wire with accuracy of 130 µm.

In between the tubes are fibres or foils made of polypropylene which provide boundaries of different dielectric properties between them and the gas in the ID volume that cause transition radiation (TR) when traversed by charged particles. The TR photons are absorbed by the xenon atoms that subsequently ionise and create larger signals than a minimum-ionising particle. Each tube is read out with two thresholds, where the higher threshold indicates presence of TR. The amount of TR generated is proportional to the Lorentz factor γ of the traversing particle and allows for discrimination of electrons and pions. Due to their small mass, electrons have a large γ factor and produce more TR than pions.

The tubes have a diameter of 4 mm and length of 144 cm. In the barrel they are aligned parallel to the

beam line and thus only measure space points in the r–ϕ plane. In the end cap they are arranged radially (like spokes) in wheels.

The measurements of space points depends on precise knowledge of the position of the detector and relative alignment of all components. It is required to have the same order of accuracy as the intrinsic resolution of the modules. Knowledge of alignment and position is calibrated regularly with collision and cosmic ray data.

The entire ID is situated in a magnetic field of up to 2 T that is generated by a solenoid magnet. The field bends the trajectories of charged particle in the r–ϕ plane. From the curvature of the reconstructed trajectories one can calculate the particle's momentum.

Calorimeters

The calorimeters in ATLAS measure the energy of incident particles in a range $|\eta| < 4.9$ and are situated around the ID. There are two kinds of calorimeters: electromagnetic (EM), that measures electromagnetic cascades from electrons and photons, and hadronic, that measure hadronic showers from jets. Both calorimeters are sampling i.e., have alternating layers of absorbing and detecting material.

The EM calorimeter (ECAL) consists of lead absorber plates in an accordion structure that also act as electrodes with liquid argon (**LAr**) in between. Readout electrodes are between the absorber plates. The barrel reaches up to $|\eta| < 1.475$ which is supplemented by end caps in the range $1.375 < |\eta| < 3.2$. The readout is segmented in ϕ–η as well as longitudinally. There are three layers in the longitudinal direction with higher ϕ–η granularity for smaller radii as can be seen in Fig. 3.4. Additionally, there is a pre-sampler which is a layer of liquid argon in front of the accordion structure. It can measure whether a particle has already started forming a shower inside the ID volume.

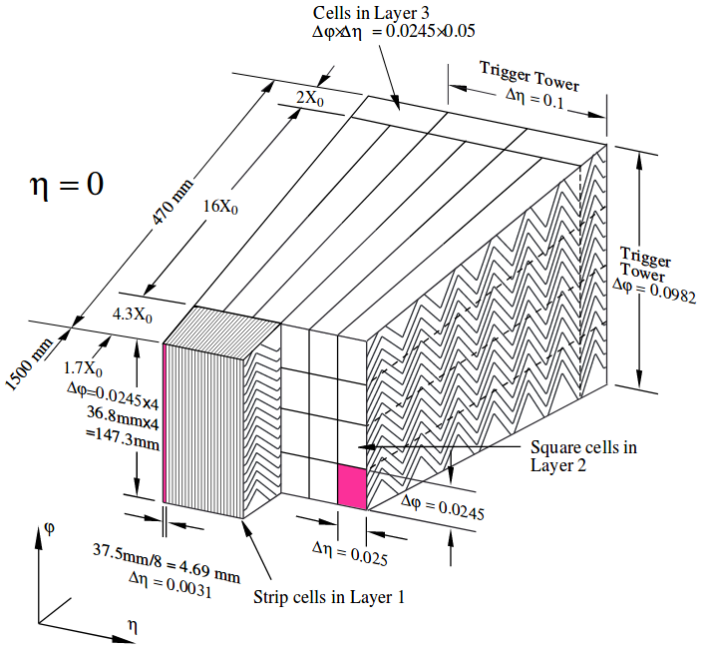


Figure 3.4: Sketch of a barrel module of the electromagnetic calorimeter. The segmentation is also shown.

The hadronic calorimeter (HCAL) uses steel as absorber and scintillating **tiles** as active medium in the range $|\eta| < 1.7$ and LAr with copper absorbers in the end caps ($1.5 < |\eta| < 3.2$). In the forward regions

$3.1 < |\eta| < 4.9$ there are forward calorimeters (**FCal**) made of LAr with copper and tungsten absorbers.

The fractional energy resolution for electrons in the LAr calorimeter is $\frac{\sigma(E)}{E} = \frac{10\,\%}{\sqrt{E(\mathrm{GeV})}} \oplus 0.2\,\%$. The tile calorimeter has a resolution of $\frac{\sigma(E)}{E} = \frac{56\,\%}{\sqrt{E(\mathrm{GeV})}} \oplus 6\,\%$ for pions [59].

The material depth of the LAr calorimeter is at least 22 radiation lengths X_0 and of the tile calorimeter approximately 7.4 interaction lengths λ. This means that most particles are contained inside the calorimeters and do not reach the muon detectors downstream.

Muon detectors

The outermost parts of ATLAS are the muon detectors that can measure the trajectories of muons within $|\eta| < 2.7$ and trigger the recording of events with muons within $|\eta| < 2.4$, jointly called muon spectrometer (MS). There are four different designs based on gaseous drift chamber technology. Two are designed to provide precision measurement of the trajectories in two coordinates, while the other two are dedicated to have a fast response time for triggering and provide one supplemental coordinate measurement. The muon trajectory bending is effected by a magnetic field generated by toroid magnets. Muon momenta up to several TeV can be measured. The toroid magnets also give ATLAS a distinctive appearance on drawings and assembly pictures. The magnetic field created by the toroids is perpendicular to the solenoid field in the ID. This provides two distinct bending planes for muons.

Data acquisition

During data taking there can be up to 40 million pp bunch crossings occurring inside ATLAS per second. It is not possible to record all of these potential collision events due to bandwidth limitation of writing data to storage media as well as limitations of reading signals from the detector. Furthermore, most of the events have low momenta transfers which are of little interest. Therefore it is crucial to have a system that triggers the recording of interesting events.

The triggering of events happens in two stages. A first level (L1) reduces the rate of events to 100 kHz and its rate capacity dictated by how fast the over 80 million channels of ATLAS that can be read out. The L1 trigger uses large energy deposits in the calorimeters and coincidences in muon chambers to make these fast decisions. The data of events that pass the L1 decision is kept in temporary buffers that allows a longer time for the next level to make a decision. The high level trigger (HLT) uses simplified algorithms to reconstruct particles within 200 ms on average and records events at a rate of 1 kHz.

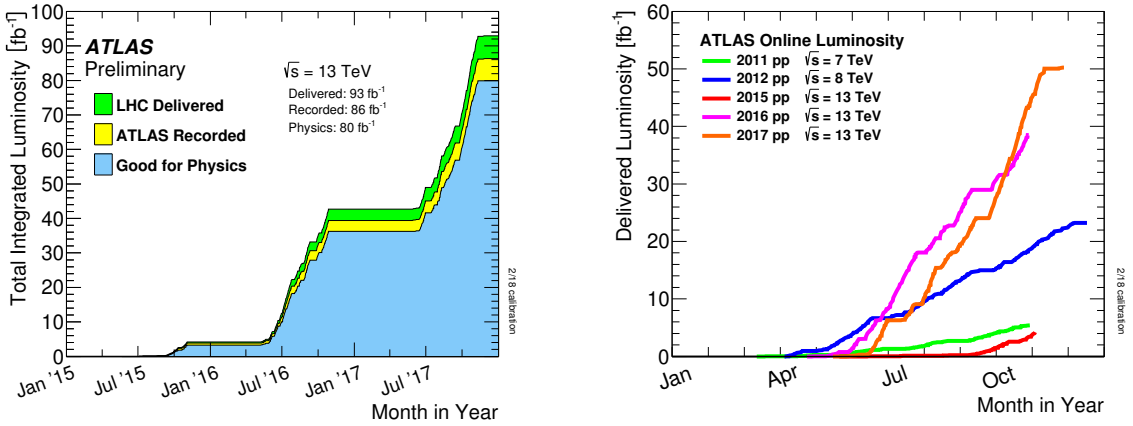
Time is the most precious resource at every stage of the trigger system. To save time the simplified event and object reconstruction in the HLT is performed only for the parts of the detector that had a positive L1 decision, so called regions of interest (ROI). The reconstruction algorithms yield objects that may have been caused by particles in the detector such as electrons, muons, jets and hadronic taus [66, 67]. The rate of occurrence of such particles and positive trigger decisions is inversely proportional to the momentum/energy of those particles. Therefore a minimum threshold on the observed transverse momentum or energy is applied to the trigger decisions to keep the rate within the limits that are possible for the detector and the data processing facilities. Furthermore, it is possible to apply quality requirements such as isolation to the trigger objects.

3.3 Recorded and simulated data

The pp collision data that is analysed in this thesis was recorded during the second period of LHC data taking in the years 2015 and 2016. The triggers used are sensitive to the presence of one and two

electrons and muons. The p_{T} thresholds of the single-electron (single-muon) triggers are 24 (20) GeV in 2015 and 26 GeV for both in 2016. In 2015 (2016) the trigger threshold is 18+8 (22+8) GeV for dimuon and 12+12 (17+17) GeV for dielectron triggers. The electron+muon trigger has thresholds of 17+14 GeV in both years. The efficiencies of the triggers are between 82 % for muon triggers and 99 % for electron triggers [68]. The muon efficiencies are lower because of smaller geometric acceptance due to non-instrumented regions where there are support structures.

The data is only considered for physics data analysis when all detector component were operating adequately and the LHC beam conditions were stable. The total data recorded in 2015–2016 that is of good quality corresponds to an integrated luminosity of 36.1 fb^{-1}. Figure 3.5(a) shows the cumulative total integrated luminosity delivered to and recorded by ATLAS, and of good quality for analysis in the years 2015–2017. Both the data-taking and data-quality efficiencies are approximately 90 %. Figure 3.5(b) shows the delivered luminosity for all years of pp data taking so far.



(a) (b)

Figure 3.5: Luminosity at ATLAS. (a) Cumulative luminosity versus time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams for pp collisions at 13 TeV centre-of-mass energy in 2015–2017 [69]. (b) Cumulative luminosity delivered to ATLAS versus day during stable beams and for high energy pp collisions [70].

The creation of simulated samples of signal and background events exploits the factorisation of perturbative hard-scattering matrix element computation (ME), parton distribution functions (PDF) and parton showering (PS). Events are generated using Monte Carlo methods (MC). The programs that were used are listed in Tab. 3.1. "Tune" refers to a set of parameters of the parton showering program that is optimised with data. The "PDF" column lists what is used in the ME generation. The PDFs used by the parton shower programs are NNPDF 2.3 LO when using the A14 tune or CTEQ6L1 for samples that use the UE-EE-5 or Perugia2012 tunes. Alternative samples that vary the generator and parton shower program are used to estimate uncertainties for $t\bar{t}H$ and $t\bar{t}V$.

All MC samples use a detailed ATLAS detector simulation [71] that is based on the GEANT4 [72] program. For a few small background and alternative samples a faster simulation [73] based on parametrised calorimeter responses is used. Pileup is modelled by generating simulated low-momentum QCD events and laying them over the simulated hard-scatter event. The pileup events are generated with PYTHIA 8 [24, 25] using the A2 [74] tune and the MSTW2008LO parton distribution function [75]. Since MC is often generated before data taking the pileup configuration that is used is a best guess. Therefore the MC is reweighted, once data is taken, such that the mean number of collisions per bunch crossing matches the one in data.

The MC samples of $t\bar{t}H$, $t\bar{t}V$, VV and $t\bar{t}$ are described in detail in Refs. [76–78].

Process	Event generator	ME order	PS	Tune	PDF
$t\bar{t}H$	MG5_aMC [79]	NLO	Pythia 8 [25]	A14 [80]	NNPDF 3.0 NLO [81]
alternative	MG5_aMC	NLO	Herwig++ [26]	UE-EE-5 [82]	CT10 [83]
$tHqb$	MG5_aMC	LO	Pythia 8	A14	CT10
tHW	MG5_aMC	NLO	Herwig++	UE-EE-5	CT10
$t\bar{t}W$	MG5_aMC	NLO	Pythia 8	A14	NNPDF 3.0 NLO
alternative	Sherpa 2.1.1 [27, 84–87]	LO multileg	Sherpa	default	NNPDF 3.0 NLO
$t\bar{t}(Z/\gamma^* \to ll)$	MG5_aMC	NLO	Pythia 8	A14	NNPDF 3.0 NLO
alternative	Sherpa 2.1.1	LO multileg	Sherpa	default	NNPDF 3.0 NLO
tZ	MG5_aMC	LO	Pythia 6 [24]	Perugia2012 [88]	CTEQ6L1 [89, 90]
tWZ	MG5_aMC	NLO	Pythia 8	A14	NNPDF 2.3 LO [91]
$t\bar{t}t$, $t\bar{t}t\bar{t}$	MG5_aMC	LO	Pythia 8	A14	NNPDF 2.3 LO
$t\bar{t}W^+W^-$	MG5_aMC	LO	Pythia 8	A14	NNPDF 2.3 LO
$t\bar{t}$	Powheg-BOX v2 [92]	NLO	Pythia 8	A14	NNPDF 3.0 NLO
$t\bar{t}\gamma$	MG5_aMC	LO	Pythia 8	A14	NNPDF 2.3 LO
single top	Powheg-BOX v1 [93–95]	NLO	Pythia 6	Perugia2012	CT10
VV, VVV	Sherpa 2.1.1	MEPS NLO	Sherpa	default	CT10
$Z \to l^+l^-$	Sherpa 2.2.1	MEPS NLO	Sherpa	default	NNPDF 3.0 NLO

Table 3.1: List of programs and configurations used to generate signal and background samples. Explanations are given in the text. "MEPS" refers to Sherpas matching of matrix element to parton shower.

3.4 Reconstruction of particles and particle-like observables

$t\bar{t}H$ events can contain particles of the kind electron, muon and tau lepton[2]. Objects that have a 4-momentum and can be treated like particles are (b–)jets, tracks and decayed tau leptons. The process of combining detector signatures to represent physical particles is called particle reconstruction. Figure 3.6 shows a sketch of the ATLAS detector and how particles are reconstructed from the signatures they left in it.

Particle identification is the process of enhancing the purity of the true target particle among the reconstructed particles. The term "object" is used to refer to reconstructed electrons, muons, tau leptons and (b–)jets jointly.

The reconstructed objects represent physical particles or observables. The correspondence between them is listed in Tab. 3.2. Electron and muon objects have both high efficiency and purity, so the mapping can be considered to be 1–to–1. For b–jets and τ_{had} there is a trade off between efficiency and purity. Both reconstruction and identification are described in the following.

3.4.1 Tracks and vertices

Trajectories of charged particles (tracks) in the ID are reconstructed from the measured space points (hits) in the pixel, SCT and TRT detectors. First, a pattern recognition algorithm is used to create candidate tracks from multiple hits. In the solenoidal magnet field the tracks follow a helical path which

[2] Their anti particles are implied.

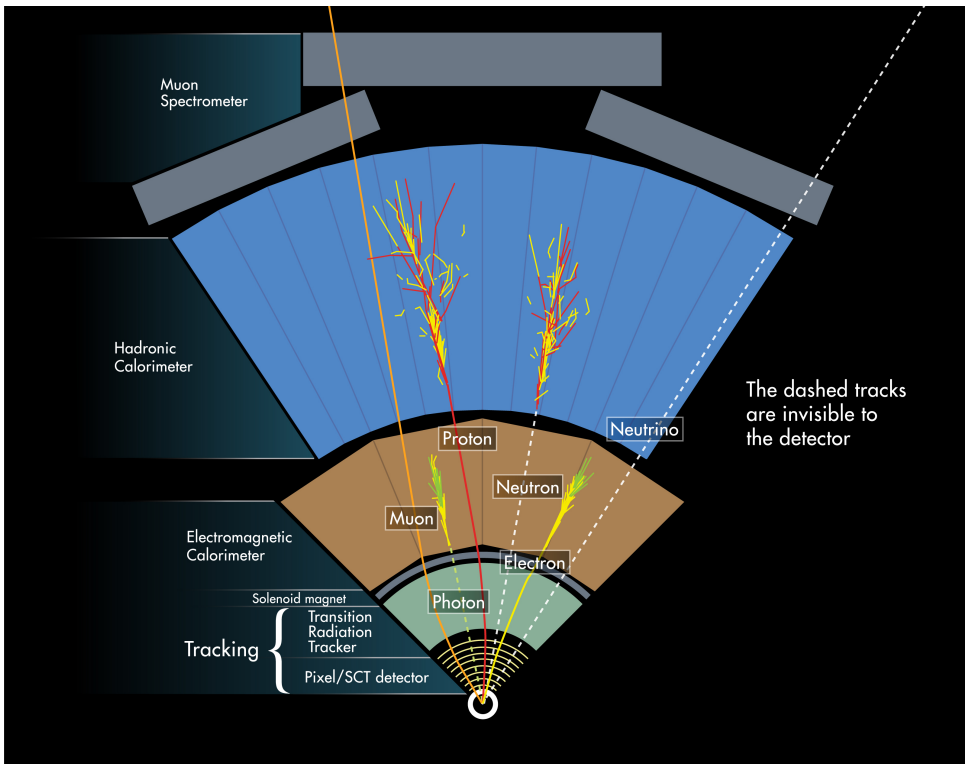


Figure 3.6: Drawing of how particles are reconstructed from their signatures in different parts of ATLAS. [96]

Reconstructed object	Physical equivalent	Efficiency	Purity
track	charged particle	high	high
vertex	position of interaction with outgoing charged particles	high	high
electron	electron	high	high
muon	muon	high	high
jet	quark- or gluon-initiated jet	high	high
b–tagged jet	b–quark initiated jet	medium	medium
τ_{had}	visible part of τ decay to 1 or 3 charged pions	low	medium

Table 3.2: Correspondence of reconstructed objects to physical particles or observables. Typical efficiencies and purities are given qualitatively. Efficiencies are defined as the ratio of number of reconstructed (and identified where applicable) objects to the total true number of the targeted particle/observable. Purity is the ratio of number of reconstructed objects that are truly the targeted particle/observable to total number of reconstructed objects.

is described by five parameters. The track parameters are then estimated by a fit to the found hits. Finally the tracks need to satisfy quality requirements on the number of hits to reduce those tracks that are badly measured or combinatorial fakes [97]. Tracks that point to a common origin are used to reconstruct vertices [98]. The vertex with the largest sum of squared transverse momenta of the associated tracks with $p_{\mathrm{T}} > 400\,\mathrm{MeV}$ is called the primary vertex (PV) of the event.

3.4.2 Electrons

Individual calorimeter cells are grouped into topological clusters by an algorithm that implicitly performs noise suppression [99]. Electrons are reconstructed from clusters of cells in the electromagnetic calorimeter that can be matched to a track in the inner detector [100, 101]. Their minimum p_{T} is 15 GeV and have to be within $|\eta_{\mathrm{cluster}}| < 2.47$. Electrons with clusters within $1.37 < |\eta_{\mathrm{cluster}}| < 1.52$ are excluded because it is the region in the detector where the transition between the barrel and end cap electromagnetic calorimeters occurs. In the transition region there is a relatively large amount of inactive material which leads to worse energy measurement performance compared to regions which consist of one homogeneous calorimeter system.

To reduce contributions from jets, photon conversions and heavy hadron decays a likelihood discriminator that uses shower shape and tracking information is used. The efficiency to reconstruct and identify an electron is greater than 98 % in $Z \to e\bar{e}$ events.

3.4.3 Muons

Muons are reconstructed from tracks in the inner detector and (partial) tracks in the muon spectrometer [102]. The p_{T} is required to be greater than 15 GeV and have $|\eta| < 2.5$. To reduce muons from heavy hadron decays which would have a distinctive kink between ID and MS, identification criteria are applied based on track quality and compatibility of the tracks measured in ID and MS. The efficiency to reconstruct and identify prompt muons is greater than 99 % in $Z \to \mu\bar{\mu}$ and greater than 98 % in simulated $t\bar{t}$ events.

3.4.4 Common selections for electrons and muons

Both electrons and muons are selected with a loose isolation requirement [101, 102] and are required to be compatible with the primary vertex.

Isolation is defined by two observables based on tracking and calorimeter information. Tracking isolation is the sum of transverse momenta of all tracks with $p_{\mathrm{T}} > 1\,\mathrm{GeV}$ from the primary vertex in a cone of $\Delta R < 10\,\mathrm{GeV}/p_{\mathrm{T}}(\ell)$ excluding the one of the electron or muon itself. At low p_{T} the cone has a maximum size of 0.2 and 0.3 for electrons and muons, respectively. Calorimeter isolation is the sum of transverse energies of the clusters in a cone of $\Delta R < 0.3$ around the electron or muon. Clusters that belong to the electron and ones that are within $\Delta R < 0.1$ of the muon are excluded from the sum. The loose requirement uses both types of isolation to select 99% of electrons and muons in $Z \to \ell\bar{\ell}$ events.

The compatibility with originating from the PV is ensured by requiring the longitudinal impact parameter to be $|z_0 \sin\theta| < 0.5\,\mathrm{mm}$ and the transverse impact parameter significance to be $|d_0|/\sigma_{d_0} < 5(3)$ for electrons (muons).

Furthermore, reconstructed electrons and muons are required to be matched to the trigger signature that triggered the event.

3.4.5 Jets

Jets are reconstructed from topological clusters [103] with the anti-k_t algorithm using a cone size of $\Delta R = 0.4$ [104]. Their p_{T} is required to be larger than 25 GeV and they have to be central ($|\eta| < 2.5$).

The ATLAS calorimeters do not respond equally to electromagnetic and hadronic showers. Furthermore part of the energy in hadronic showers can be invisible due to hadron decays to neutrinos, recoil and delayed processes. Therefore the energy of jets is calibrated to the hadronic scale with MC. Corrections to reduce the impact of pileup are also applied. Finally a calibration using p_{T} balances of dijet, multijet, jet+photon and jet+Z events in data is performed [105].

A multivariate discriminant that uses the fraction of tracks inside the cone that are compatible with the primary vertex reduces jets from pileup processes with $p_{\mathrm{T}} < 60$ GeV and $|\eta| < 2.4$ with an efficiency of 92 % [106].

3.4.6 Jet flavour tagging

Jets can be initiated by quarks and gluons. Furthermore hadronic decays of tau leptons are also reconstructed as jets. These different kinds of jets have unique properties that are used to tag them. In particular for $t\bar{t}H$ events, tagging b–jets is useful to infer the presence of top quarks (which decay almost exclusively to b quarks).

A multivariate discriminant is used to tag jets that were initiated by a b quark. The discriminant is developed on MC where it is possible to give a truth label to jets. In a process called "truth-matching", hadrons from the history record of the MC program are matched to the reconstructed jet. The flavour of the highest energy hadron found in a cone of $\Delta R < 0.3$ gives the jet a label of b, c or "light" if no flavoured hadron is found. Jets matched to a hadronic tau lepton decay are also labelled accordingly.

The b–jets are tagged by exploiting the fact that they contain B-hadrons which have a measurable lifetime before decaying. The decay products of the B-hadrons originate from a vertex that is displaced from the PV. The distance of the secondary vertices and the properties of the tracks from these vertices are discriminating observables. Equivalently the tracks that are associated to the b–jet have larger impact parameters[3] in 2D and 3D to the primary vertex than light or gluon jets. The tagging algorithm uses templates of discriminating track and vertex observables created for b–jets, c–jets and light jets to give the likelihood that a given jet is a b–jet. Finally, the likelihoods are combined with a reconstruction of the $B \to D$ hadron decay chain where possible into a single discriminant [107, 108]. If a jet has a b–tagging discriminant value above a certain threshold it is considered b–tagged. The working point that is used has an average efficiency of tagging b–jets with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$ of 70 % in $t\bar{t}$ events.

The efficiencies to b–tag light/gluon jets, c–jets and hadronic taus expressed as rejection[4] factors are 380, 12, 55, respectively [108, 109]. Corrections are applied to MC events such that the efficiencies are compatible with data. The corrections to the b–jet efficiency are compatible with unity, while the corrections to the other efficiencies can be as large as 20 %. The τ_{had} efficiencies are not measured in data and so their corrections are assumed to be equal to the c–jet efficiency corrections as the τ_{had} contributions can usually be neglected. To account for the assumption the uncertainty of the τ_{had} efficiency correction is made larger.

[3] distance of closest approach of the track to a vertex
[4] Rejection is the reciprocal of efficiency.

3.4.7 Hadronically decaying tau leptons

Tau leptons have a significant lifetime but decay before the first layer of the pixel detector in most cases. The dominant decay modes are to charged hadrons (see Fig. 2.15 and Tab. 2.4). Such hadronic decays are reconstructed as jets and can be distinguished from quark/gluon jets with tracking and calorimeter observables [110]. Figure 3.7 shows an example of a tau decay to one neutral and one charged pion, as well as a quark/gluon jet that might mimic a tau decay. All decay products are typically in a narrow cone around the tau lepton direction. True and reconstructed hadronically decaying tau leptons are called τ_{had}.

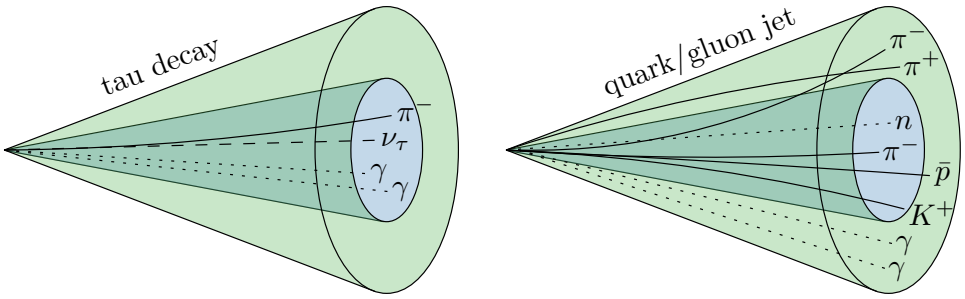


Figure 3.7: Drawing of an exemplary (left) 1-prong hadronic tau decay and (right) quark/gluon jet [111]. Two cone sizes with respect to the jet axis are shown. The outer cone is used to reconstruct the jet. The inner cone (τ-cone) is used to define many observables for the identification.

Tau lepton decays to the lighter charged leptons are not distinguishable from prompt electrons or muons. Thus they are reconstructed and identified as electrons or muons. The notation for these is τ_{lep}. These leptons are only identifiable as tau decays products when the presence of a tau lepton can be inferred from correlations in the events e.g. when there are tau lepton pairs.

Reconstruction The τ_{had} candidates start as jets reconstructed with the anti-k_T algorithm with size 0.4. Unlike the quark/gluon jets described above, the tau candidates are initially calibrated with the "local hadronic calibration" (LC) [112]. The LC method reweights the energy of calorimeter cells to compensate for the different response of the calorimeter to electromagnetic and hadronic showers. The weighting uses the energy density of cells which is correlated to the shower shapes. EM showers are narrower than hadronic showers and have a higher density.

The momentum of the τ_{had} candidate is set to the total energy of clusters inside an inner cone of $\Delta R < 0.2$ in the jet. For the purpose of defining the observables used in the identification, the clusters in the last layer of the ECAL are counted as hadronic.

Vertex and track selection Out of all reconstructed vertices the vertex with the largest sum of p_{T}^2 of tracks associated to the jet is assigned to the τ_{had} candidate. This may be different from the vertex that maximises the sum for all tracks (called PV above) and is called tau vertex (TV).

Tracks in the core region ($\Delta R < 0.2$) are selected with $p_{\text{T}} > 1\,\text{GeV}$ and requirements on the number of hits in the ID. The tracks also have to be compatible with the tau vertex.

In correspondence to the tau lepton decay to one or three charged hadrons, τ_{had} candidates are defined as 1-prong when there is one track in the core region and 3-prong when there are three tracks in the core region.

All above selections are approximately 70 % efficient for true hadronic taus with $p_{\text{T}} > 20\,\text{GeV}$.

Secondary vertices are reconstructed from tracks in the core region for τ_{had} candidates with multiple such tracks.

Identification The discrimination with respect to quark/gluon jets is provided by a multivariate discriminant that combines track impact parameters, secondary vertices and shower shape observables. The discriminant is a boosted decision tree (BDT)[5] and is trained separately for 1-prong and 3-prong τ_{had} candidates on $Z/\gamma^* \to \tau\tau$ as signal and dijet events as background. The following list describes the observables that are the inputs to the BDT [110].

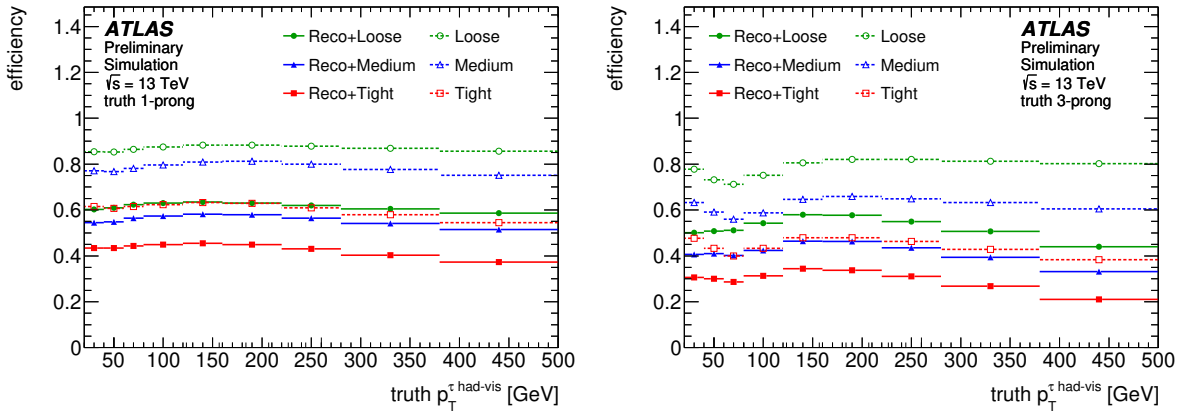
1. *central energy fraction*, f_{cent}: fraction of energy deposited in a region $\Delta R < 0.1$ to all energy within $\Delta R < 0.2$ of the τ_{had} candidate at the EM scale.
2. *leading track momentum fraction*, $f^{-1}_{\text{leadtrack}}$: fraction of calorimeter energy at EM scale to the transverse momentum of the highest-p_{T} track in the core region.
3. *track radius*, $R^{0.2}_{\text{track}}$: p_{T}–weighted ΔR distance between tracks in the core region to τ_{had} direction.
4. *track impact parameter significance*, $|S_{\text{leadtrack}}|$: absolute value of impact parameter of track in core region divided by its uncertainty. Only used for 1-prong τ_{had} candidates.
5. *fraction of track p_T in the isolation region*, $f^{\text{track}}_{\text{iso}}$: fraction of scalar p_{T} sum of tracks in isolation region ($0.2 < \Delta R < 0.4$) to p_{T} sum of all tracks. Only used for 1-prong τ_{had} candidates.
6. *maximum ΔR distance*, ΔR_{Max}: maximum ΔR between any track in core region and τ_{had} direction. Only used for 3-prong τ_{had} candidates.
7. *transverse flight path significance*, $S_{\text{T}}^{\text{flight}}$: transverse distance between secondary vertex and tau vertex divided by its uncertainty. Only used for 3-prong τ_{had} candidates.
8. *track mass*, m_{track}: invariant mass of vector sum of all tracks, assuming all tracks are pions. Only used for 3-prong τ_{had} candidates.
9. *fraction of EM energy from charged pions*, $f_{\text{EM}}^{track-HAD}$: fraction of core track momenta minus hadronic cluster energy to electromagnetic cluster energy. Clusters are LC calibrated.

$$\frac{\sum_{\text{tracks}}^{\Delta R<0.2} p - \sum_{\text{HAD clusters}} E}{\sum_{\text{EM clusters}} E}$$

10. *fraction of EM energy to track momentum*, $f^{\text{EM}}_{\text{track}}$: fraction of electromagnetic cluster energy to core track momenta. Clusters are LC calibrated.
11. *track-plus-EM mass*, $m_{\text{EM+track}}$: invariant mass of vector sum of tracks and EM clusters in core region. Each EM cluster is assumed to have zero mass.
12. *fraction of EM-plus-track to transverse momentum*, $p_{\text{T}}^{\text{EM+track}}/p_{\text{T}}$: fraction of p_{T} of vector sum of tracks and two most energetic EM clusters to p_{T} of vector sum of clusters.

All observables were reweighted to minimise the correlation with pileup. Working points of τ_{had} identification are defined by placing requirements on the identification discriminant. The cut thresholds are p_{T} dependent to give an approximately constant efficiency. Figure 3.8 shows the combined reconstruction and identification efficiency. For 1(3)-prong candidates it is 55 %(40 %) for the used "Medium" working point. The rejection factor against quark/gluon jets in a dijet MC sample is approximately 50(100) for 1(3)-prong candidates.

[5] A description of BDT is in Sec. 5.4.



(a) 1-prong (b) 3-prong

Figure 3.8: Efficiency for τ_{had} identification (with respect to reconstructed τ_{had}) (open symbols) and combined reconstruction and identification (full symbols) as a function of true p_{T} for (a) 1-prong and (b) 3-prong τ_{had} candidates [110].

Corrections are applied to MC so that the efficiencies match those measured in $Z \to \tau\bar{\tau}$ events [113]. The size of the uncertainties on the corrections are 5 %.

Rejection of muon, electron, *b*–jet and pileup backgrounds Other backgrounds that are selected by the τ_{had} reconstruction and identification are electrons, muons, *b*–jets and pileup jets.

Predominantly 1-prong τ_{had} candidates suffer from electron backgrounds where the electron mimics the signature of a charged pion. A BDT is trained to reject this background. Discriminating observables to distinguish electrons from pions are shower shapes in the calorimeters and hits in the TRT where electrons generate transition radiation while pions do not.

Muon background is rejected by requiring the τ_{had} not to geometrically coincide with a reconstructed muon with $p_{\text{T}} > 2\,\text{GeV}$.

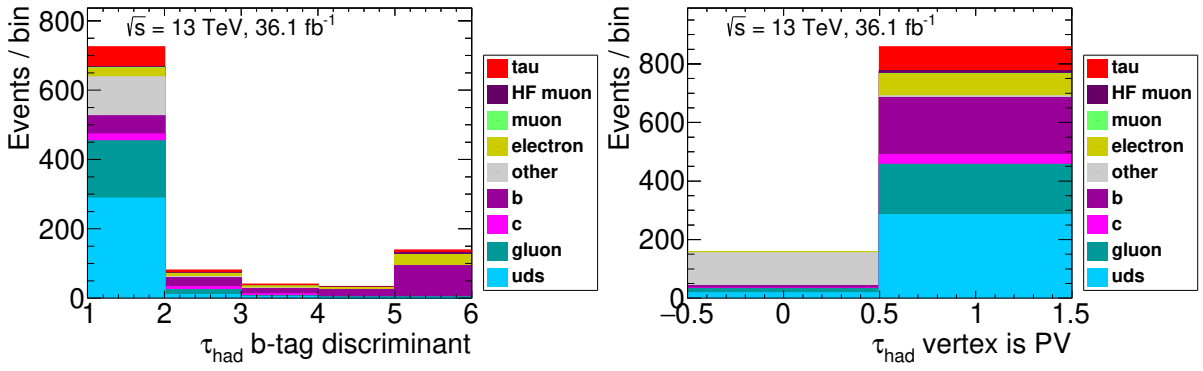
Backgrounds from *b*– and pileup jets are not yet treated in a standard way in ATLAS. Therefore, the standard τ_{had} definition is modified in this thesis.

The jet flavour labelling described above is extended to also match electrons, muons, quarks and gluons. True electron, muon and tau leptons are matched within a distance $\Delta R < 0.3$. The "light" label is split into quark (uds) and gluon by matching true quarks and gluons in a cone of $\Delta R < 0.3$. The order of precedence for matching is electrons/muons, B- and D-hadrons, then quarks and gluons. A jet that can not be matched to anything within the cone is assumed to originate from pileup interactions (called "other" in Fig. 3.9).

b–jet fakes are reduced by requiring that the τ_{had} does not overlap with a *b*–tagged jet. This cut is illustrated in Fig. 3.9(a) which shows the flavour label of the tau jet as a function of the *b*–tagging discriminant of the jet that overlaps with the τ_{had} (before said cut). Jets with a discriminant in the last two bins are considered *b*–tagged. The figure shows that rejecting the events in those bins removes *b*–jet fakes but does not significantly affect true hadronic τ–lepton decays that are selected.

A τ_{had} candidate can also be faked by pileup jets. A pileup jet originates from one of the secondary collisions during a bunch crossing and consists mainly of pions. Such pions are not compatible with originating from the primary vertex. Therefore pileup jet fakes are reduced by requiring the vertex that is associated to the τ_{had} to be the PV. Figure 3.9(b) shows that it is effective in removing only pileup jets

while leaving true hadronic τ–lepton decays unaffected.



(a) values of 4 – 6 are rejected

(b) value 0 is rejected

Figure 3.9: Distributions of true compositions of (a) b–tagging discriminant of the jet matched to the τ_{had} and (b) equality of τ_{had} vertex with primary vertex. "other" are unmatched jets and assumed to be from pileup. "HF muon" are muons matched to a heavy flavour hadron decay. Event selection is the one described in Tab. 5.1.

Energy calibration The τ_{had} energy is corrected with a boosted regression tree (BRT) to the energy of the generated hadrons in $Z \to \tau\bar{\tau}$ MC [113]. The BRT is a multivariate algorithm which provides a correction for a target observable based on multiple input observables. The target in this case is the true τ_{had} energy and the inputs are shower shape observables (similar to those used for the LC calibration) and reconstructed hadron momenta inside the τ_{had}. To improve the energy resolution at low values another dedicated algorithm reconstructs neutral pions in the τ_{had} [114]. Neutral pions predominantly decay to two collimated photons. Those photons can be seen as two local maxima in the first layer of the ECAL. The reconstructed neutral energy is combined with the charged momentum measured by the ID. At low τ_{had} p_{T} the BRT improves the p_{T} resolution by a factor 2. Figure 3.10 shows the p_{T} resolution with the BRT correction and with the calorimeter-only calibration. The resolution is measured in $Z \to \tau\bar{\tau}$ events with an uncertainty of 1.2(3.0) % for 1(3)-prong τ_{had} candidates.

3.4.8 Missing energy

Energy that escapes the detector via neutrinos cannot be reconstructed directly. To infer the magnitude of this missing energy all reconstructed objects and all tracks that are associated with the primary vertex (soft) but not to any of the reconstructed objects are added vectorially as

$$-E_{\text{T}}^{\text{miss}} = \sum^{\text{electrons}} E_{\text{T}} + \sum^{\text{muons}} E_{\text{T}} + \sum^{\tau_{\text{had}}} E_{\text{T}} + \sum^{\text{jets}} E_{\text{T}} + \sum^{\text{soft}} E_{\text{T}}, \tag{3.4}$$

where E_{T} is a 4-momentum with a zero longitudinal component.

The energy should be conserved when summed, so any net energy is an indicator that something is undetected. This is also a reason why it is important that ATLAS covers as much of the solid angle as possible. The missing energy is unconstrained in the direction along z because the momenta of the interacting partons is not known. However, the net transverse momentum in any collision event must be zero and so the negative direction of the vector sum gives the ϕ-direction of the missing energy [115, 116].

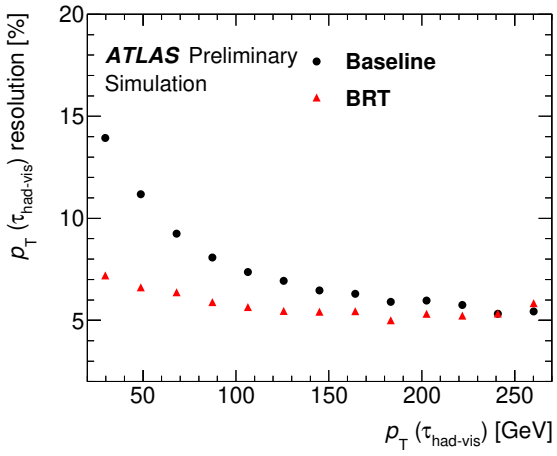


Figure 3.10: Resolution of τ_{had} energy calibration with regression correction and only calorimeter based (baseline) [113].

3.4.9 Overlap removal

The reconstructed objects described above can be geometrically overlapping with each other. This happens when a physical particle is duplicated as several reconstructed objects or when several physical particles are actually nearby each other. An example for duplication is an electron that is reconstructed as an electron as well as a jet. Both the reconstructed electron and jet were caused by the same physical particle. An example for nearby objects is a muon inside a b–jet created by a hadron decay. Overlap removal thus resolves ambiguities and affects the isolation of particles.

The overlap removal algorithm is following the sequence of decisions in Tab. 3.3 to determine which objects will not be considered in the analysis i.e. removed. All requirements use fixed ΔR distances except for overlapping jets and muons. The cone for the jet-muon step starts at $\Delta R = 0.4$ for muons with $p_T < 28\,\text{GeV}$ but shrinks for muons with higher p_T. This allows more muons to survive at high p_T and increases signal acceptance. The variable cone size is illustrated in Fig. 3.11.

Keep	Remove	Critical distance in ΔR
muon	electron	0.1
electron	electron	0.1
electron	jet	0.3
jet	muon	$\min(0.4, 0.04 + \frac{10\,\text{GeV}}{p_T(\mu)})$
electron	τ_{had}	0.2
muon	τ_{had}	0.2
b–tagged jet	τ_{had}	0.3
τ_{had}	jet	0.3

Table 3.3: Algorithm to resolve ambiguities of overlapping reconstructed objects. The type of object in the "Remove" column is removed when it is closer than the critical distance to the type of object in the "Keep" column. For same type objects the one with lower p_T is removed.

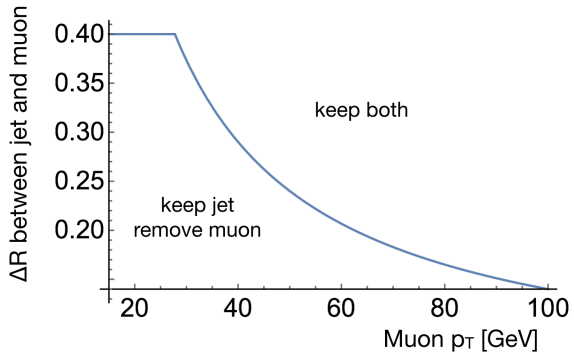
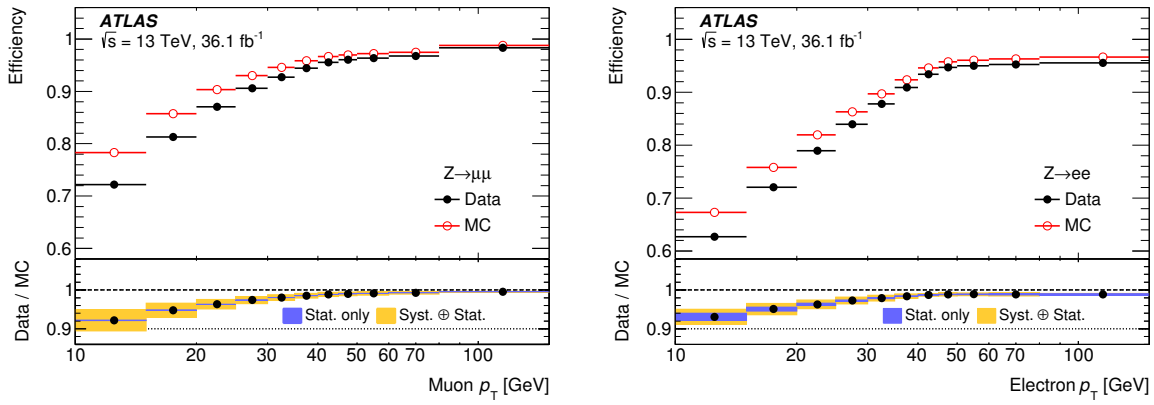


Figure 3.11: Illustration of the variable cone size of the jet-muon overlap removal step. The critical distance between a jet and a muon is $\min(0.4, 0.04 + \frac{10\,\text{GeV}}{p_\text{T}(\mu)})$. If the distance between any jet-muon pair is smaller than the critical distance, the muon is removed.

3.4.10 Prompt electron and muon tagging

Prompt electrons and muons are defined as direct decay products of heavy bosons or tau leptons. Electrons and muons that originate from the decay of a B- or D-hadron decay are called non-prompt. Additionally it is possible that photons in the jet shower convert to an electron pair asymmetrically. Such leptons really belong to the jet. However it is possible that the lepton is reconstructed and identified and then has priority or is further than the critical distance in the overlap removal procedure. The result is the loss of a real b–jet which is replaced by an additional lepton. This is a significant background to the $t\bar{t}H$ analyses with multiple leptons.

Therefore a multivariate discriminant was developed that tags the promptness of electrons and muons [5]. The discriminant observables in the tagger are derived from b–tagging algorithms. The light lepton is matched to a jet that is reconstructed from tracks rather than clusters ("track jet"). The b–tagging impact parameter likelihood discriminants together with isolation observables, number of tracks in the jet, angular distance between lepton and jet and lepton p_T to jet p_T ratio are inputs to a BDT training with prompt and non-prompt leptons in simulated $t\bar{t}$ events. A requirement on the output of the tagging discriminant defines a working point that rejects non-prompt leptons. The efficiency to select prompt electrons and muons is measured in $Z \to \ell\bar{\ell}$ events and is compared to MC in Fig. 3.12. It is 96 % (98 %) for electrons (muons) with medium p_T greater than roughly 45 GeV, but considerably lower at low p_T with 60 % (70 %). The rejection factor of non-prompt leptons is approximately 20. This factor is approximately two times better than what can be achieved with conventional isolation definitions that rely on calorimeter or track cones.



(a) muons (b) electrons

Figure 3.12: Efficiency of the prompt lepton tagging as function of p_{T} for (a) muons and (b) electrons. Red open (black solid) points represent $Z \to \ell\bar{\ell}$ MC (data). The lower panels show the ratio between data and MC with blue (yellow) shading for statistical (full) uncertainties [5].

CHAPTER 4

Search for top associated Higgs production

The production of $t\bar{t}H$ is a rare process. Its cross section is roughly 1 % of the total Higgs production. To analyse as many $t\bar{t}H$ events as possible an inclusive search strategy is adopted which consists of combining multiple analyses which target all the H decays that are experimentally accessible.

4.1 Multiple final states

There are four analyses which are currently sensitive to $t\bar{t}H$ production with the ATLAS detector. Each is optimised to specific H decays with distinct challenges. Aside from the selection of the H decay products, the analyses exploit the unique signature of the associated top-quark pair which almost always decays to two b quarks which initiate b–jets which can be tagged and 0–2 leptons.

$t\bar{t}H$ with $H \to b\bar{b}$ The H decay to a bottom quark pair has the largest branching fraction. Such a decay in a $t\bar{t}H$ event means that there are four b–jets which can be reconstructed as b–tagged jets. In principle, all $t\bar{t}$ pair decay modes are accessible experimentally and have been analysed in Run 1 of the LHC. In practice, the most challenging is the fully hadronic final state with at least 8 jets at leading order [117, 118]. The challenges lie both in data acquisition (trigger) and background reduction and prediction (multijet).

Thus, the most sensitive analyses of the $H \to b\bar{b}$ final state require at least one top quark to decay to a light lepton. This makes data acquisition comparatively trivial by using lepton triggers [119].

The dominant background in the analysis is $t\bar{t}$ production with additional b–quarks from NLO QCD radiation and gluon splitting. While the inclusive $t\bar{t}$ cross section is predicted and measured with high precision, the cross section of this particular phase space is not [120].

$t\bar{t}H$ with $H \to \gamma\gamma$ The signature of a Higgs decay to two photons can be selected by the resonant invariant H mass. The production in association with a top-quark pair is experimentally accessible in both the hadronic and leptonic $t\bar{t}$ decays by tagging the presence of a b–jet. With the branching fraction of $H \to \gamma\gamma$ decays ($2.270 \cdot 10^{-3}$ [29]) and the typical efficiency and acceptance of 5 % ([121]) only approximately 2 events are expected to be selected in the data collected in 2015–2016.

$t\bar{t}H$ with $H \to ZZ^* \to 4\ell$ The decay of a H boson to 4 leptons via a pair of Z bosons provides an exceedingly pure signature with very little background. The quadruplet of leptons allows a full and precise reconstruction of the H boson kinematics. In this search only light leptons, e and μ, are considered

for their high reconstruction and identification efficiencies and high energy/momentum resolutions. After selecting the lepton quadruplet mass to be close to the H mass, backgrounds are very small. The high purity also allows to search for $t\bar{t}H$ in both hadronic and leptonic $t\bar{t}$ decays. However, the decay is rare with BR($H \to 4\ell$)=$1.251 \cdot 10^{-4}$ where $\ell = e, \mu$ [29], so only approximately 2 events are expected to have been produced and less than 1 to be selected during 2015–2016 [122].

Both searches via $H \to \gamma\gamma$ and $H \to 4\ell$ decays are therefore statistically limited in the 36.1 fb^{-1} dataset.

$t\bar{t}H$ with $H \to WW^*$, $H \to \tau\bar{\tau}$ and other $H \to ZZ^*$ All other H decays that are experimentally available for analysis are studied together in the "multilepton" analysis. The name refers to the presence of many charged leptons in the final state, in this case including tau leptons. Unlike the three previously described analyses, no single specific H decay is targeted here. Instead a combination of various $H \to WW^*$, $H \to \tau\bar{\tau}$ and $H \to ZZ^*$ decays are selected.

The targeted and selected decay modes of the four $t\bar{t}H$ searches are summarised in Tab. 4.1. The multilepton analysis is described in detail in the following.

Analysis name	$t\bar{t}$ decay	H decay
ttHbb	1–2ℓ	$b\bar{b}$
ttHML "multilepton"	1–2ℓ	(WW^*, $\tau\bar{\tau}$, ZZ^*) →1–4ℓ
ttHZZ	0–2ℓ	$ZZ^* \to 4\ell$
ttH$\gamma\gamma$	0–2ℓ	$\gamma\gamma$

Table 4.1: Summary of $t\bar{t}H$ decay modes targeted by different analyses. In general $\ell = e, \mu$, except in the multilepton analysis that explicitly considers hadronic tau decays and where ℓ stands for all charged leptons.

4.2 Final states with many leptons

The $t\bar{t}H$ multilepton analysis is characterised by final states with many charged leptons, many jets and b–jets. It defines several analysis categories ("channels") by counting the number and charges of leptons in the final state. In total there are 7 channels with 2–4 selected charged leptons. Electrons and muons are counted separately from tau leptons because of the fundamentally different way of their reconstruction and identification which have different background contributions. In the context of the multilepton analysis, electrons and muons are both called light leptons or simply leptons where it is unambiguous. Tau leptons are identified only by their hadronic decays and are therefore called "hadronic taus" or τ_{had}, symbolically. The categorisation of channels is sketched in Fig. 4.1. On top of the lepton counting described in the following, the general selection requirements in all channels are a minimum number of jets (2-4) and a minimum number of b–tags among those jets.

The channels are the following:

2ℓ(SS)$0\tau_{\text{had}}$ selects two light leptons which have electric charges of the same sign (SS). Any hadronic taus are vetoed. The main backgrounds are $t\bar{t}W$, $t\bar{t}Z$, and non-prompt leptons. Two multivariate discriminants are trained to separate signal from $t\bar{t}V$ and signal from non-prompt background. Both discriminants are combined linearly to create a single final observable.

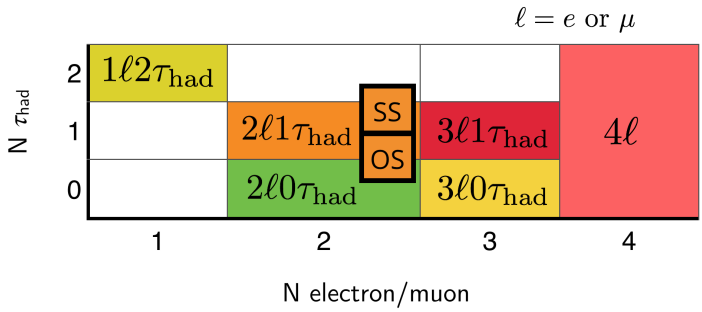


Figure 4.1: Diagram of how analysis channels are defined in the ttHML analysis by counting leptons.

$3\ell 0\tau_{had}$ selects three light leptons which have an absolute sum of electric charges equal to 1. Any hadronic taus are vetoed. The dominant backgrounds are the same as for $2\ell(\text{SS})0\tau_{had}$ above and diboson. Four discriminants are trained each with background $t\bar{t}W$, $t\bar{t}Z$, $t\bar{t}$ and diboson. Five orthogonal regions are defined using the discriminants that are enriched in each background forming control regions. The region enriched in $t\bar{t}H$ is consequently the signal region.

4ℓ selects four light leptons with zero charge sum. No explicit requirement on hadronic taus is made, however an analogous requirement of zero hadronic taus would have negligible impact. This channel is distinguished from the $H \to 4\ell$ analysis by removing events with $m_{4\ell}$ close to m_H. The largest fraction of signal events selected are with the $H \to WW^*$ decay. Therefore there are no lepton flavour correlations between the selected leptons and any combination of electrons and muons is possible. The requirement of total zero charge sum implies that there are two pairs each of zero charge sum. These lepton pairs of opposite-sign (OS) charge are classified by their flavour. There can be 0, 2 or 4 OS pairs with the same flavour (SF). Two analysis regions are defined by the presence or absence of such SFOS lepton pairs, called Z enriched and Z depleted. The dominant background in the enriched region is $t\bar{t}Z$. A discriminant is trained against it to reduce its contribution in the Z enriched region with a cut.

$1\ell 2\tau_{had}$ selects one light lepton and two hadronically decaying tau leptons which are reconstructed with opposite charge. The dominant background is fake hadronic taus from $t\bar{t}$ events. A discriminant is trained to distinguish $t\bar{t}H$ and $t\bar{t}$.

$2\ell(\text{SS})1\tau_{had}$ selects two light leptons with the same charge and one τ_{had}. The main backgrounds are non-prompt leptons, fake τ_{had} and $t\bar{t}V$. A BDT is trained to separate $t\bar{t}H$ from all other backgrounds in a loose selection region.

$2\ell(\text{OS})1\tau_{had}$ selects two light leptons with opposite charge and one τ_{had}. The dominant background is fake τ_{had} in $t\bar{t}$ events. A BDT is trained to distinguish $t\bar{t}H$ from $t\bar{t}$. The absolute sum of charges of electrons, muons and τ_{had} is 1.

$3\ell 1\tau_{had}$ selects three light leptons and one τ_{had}. The dominant background is $t\bar{t}Z$ with and without a fake τ_{had}.

The channels that veto or do not explicitly select hadronic taus ($2\ell(\text{SS})0\tau_{had}$, $3\ell 0\tau_{had}$, 4ℓ) select $t\bar{t}H$ events with mainly H decays into W boson pairs. Smaller contributions between 10-20 % from $H \to \tau\bar{\tau}$ decays and $H \to ZZ^*$ (only for $3\ell 0\tau_{had}$ and 4ℓ) are also present.

The $t\bar{t}H$ events in the channels that select at least one hadronic tau consist of at least 60 % $H \to \tau\bar{\tau}$ decays. The rest is made up of $H \to WW^*$ decays. In the presence of two hadronic taus the fraction of $H \to \tau\bar{\tau}$ events is increased to over 95 %. The composition of $t\bar{t}H$ events in all signal regions is shown in Fig. 4.2. The acceptance and efficiency to select $t\bar{t}H$ events in each channel is given in Tab. 4.2. These numbers multiplied by the expected $t\bar{t}H$ cross section and integrated luminosity give how many $t\bar{t}H$ events are expected in the analysis. They are influenced by the multiplicity of selected objects as well as the efficiencies to select those objects i.e., they are expected to be larger for channels with electrons or muons than for channels with hadronic taus.

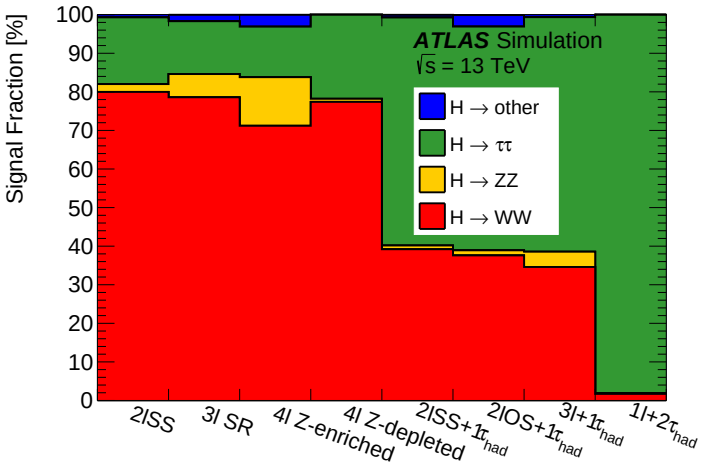


Figure 4.2: Fractions of $t\bar{t}H$ events with major Higgs decay modes selected by ML channels [5].

	$2\ell(\text{SS})0\tau_{\text{had}}$	$3\ell 0\tau_{\text{had}}$	4ℓ	$1\ell 2\tau_{\text{had}}$	$2\ell(\text{SS})1\tau_{\text{had}}$	$2\ell(\text{OS})1\tau_{\text{had}}$	$3\ell 1\tau_{\text{had}}$	Total
$A \times \epsilon\ [10^{-4}]$	23	13	0.6+0.1	2.3	1.7	7.8	0.8	50

Table 4.2: Acceptance times efficiency ($A \times \epsilon$) for $t\bar{t}H$ signal in each analysis channel. This includes Higgs boson and top quark branching fractions, detector acceptance, and reconstruction and selection efficiency, and is computed relative to inclusive $t\bar{t}H$ production considering all Higgs boson and top decays. In the 4ℓ channel, the two numbers correspond to the Z enriched and the Z depleted categories [5].

A charged lepton which is visible to the detector and reconstructable for analysis is accompanied by a neutrino which is undetected when produced in processes involving charged W bosons. This production mechanism is frequent in both the $t\bar{t}H$ signal as well as the backgrounds. In case of tau-lepton decays to electrons or muons there are even two neutrinos. The reconstructed missing energy only provides indirect information about the vector sum of these undetected particles. All individual kinematic information is hidden from and inaccessible to the analysis. This makes a full event reconstruction with assignment of leptons to their mother particles difficult. There are studies, e.g. for the $2\ell(\text{SS})1\tau_{\text{had}}$ channel, which show that event reconstruction with high efficiency is possible only if the neutrino momenta were known, or inversely, decomposition of the missing energy is possible if the assignment of observed particles to their mothers were known [123].

For this reason, almost all channels in the analysis use multivariate classifiers in the form of BDTs as a brute force replacement for full event reconstruction.

4.2.1 Backgrounds

The charges and numbers of light leptons are chosen such that they have to originate from both the H and the $t\bar{t}$ system in $t\bar{t}H$ events. At first order, there is no possibility to select leptons that are decay products of solely one of the charge-correlated systems. This property of the selection is chosen not only because it is characteristic for $t\bar{t}H$ events but also because it reduces backgrounds.

A distinction can be made between reducible and irreducible backgrounds. The latter consists of processes that can have the same final state in terms of lepton counting as $t\bar{t}H$ at leading order. Reducible backgrounds contribute through secondary processes like non-prompt leptons or are selected because of fake hadronic taus. In the multilepton analysis the irreducible backgrounds are estimated with MC, while the reducible are estimated with data-driven methods. The motivation to use these methods rather than MC is the fact that the secondary processes that make up the reducible backgrounds are rare. It is thus inefficient to generate a sufficient number of MC events which can populate the particular phase space that is selected. Furthermore, these backgrounds arise in the context of jet fragmentation, hadronisation, hadron decays and detector simulation. The correct description of these processes relies heavily on tuning of free parameters of the MC generator programs. It may thus be the case that MC does not describe these backgrounds well when they are scrutinised this closely.

Non-prompt electrons and muons in $t\bar{t}$ events

The $t\bar{t}$ process is counted as a reducible background in all signal regions. Examples of Feynman diagrams of $t\bar{t}$ are shown in Fig. 4.3.

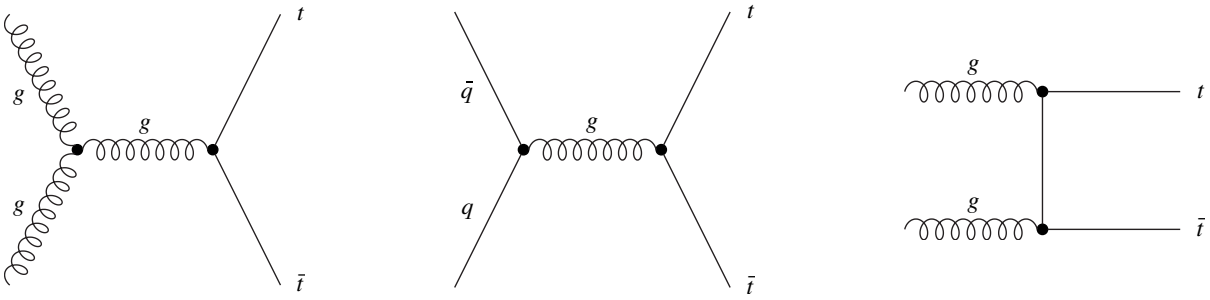
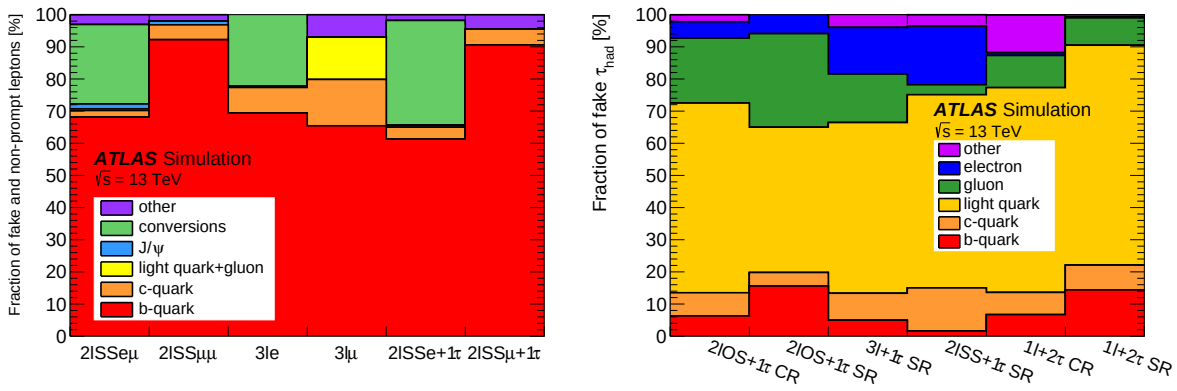


Figure 4.3: Examples of Feynman diagrams of the $t\bar{t}$ process.

Non-prompt electrons and muons originating from secondary processes like meson decays and photon conversions that occur inside jets are the main source of same-sign light leptons in the $t\bar{t}$ process and account for almost half of the background in the $2\ell(\mathrm{SS})0\tau_{\mathrm{had}}$ channel and a quarter in $3\ell0\tau_{\mathrm{had}}$. To contribute to the $3\ell0\tau_{\mathrm{had}}$ channel, a non-prompt lepton is selected in addition to the two prompt ones from a dileptonic $t\bar{t}$ decay. The composition of secondary processes that contribute to the background is shown in Fig. 4.4.

In case of additional muons, the selected muon needs to be sufficiently separated from the jet that produced it to survive the overlap removal procedure (see Sec. 3.4.9). So two reconstructed objects remain in the event: the jet and the non-prompt muon. In case of electrons, the jet is removed when it is too close to the electron. So only one object remains from the jet: the non-prompt electron. A consequence is that events with muons tend to have more jets than events with electrons for background processes. This tendency is visible in Fig. 4.5, where the q mis-id[1] background is only relevant for events

[1] q mis-id refers to the incorrect assignment of charge to reconstructed electrons. See next section.



(a) Non-prompt light leptons

(b) Fake hadronically decaying tau leptons

Figure 4.4: True composition from simulation of (a) non-prompt light leptons and (b) fake hadronically decaying tau leptons in ML channels [5].

with electrons and one can infer that bins with fewer jets have a large contribution from events that contain electrons.

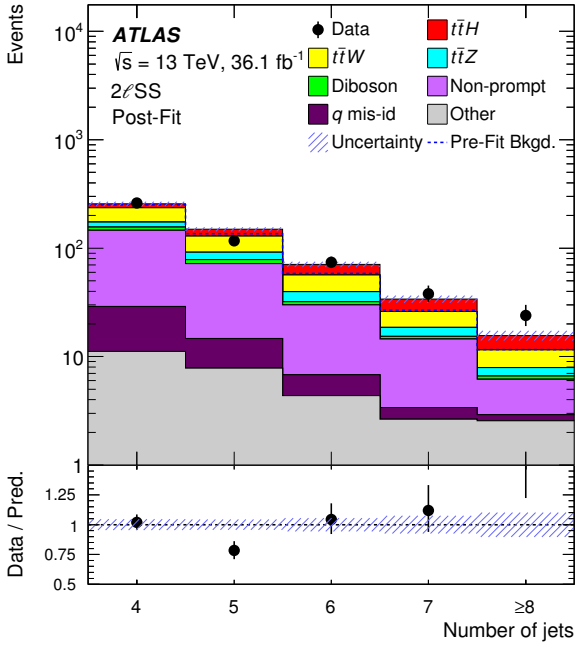


Figure 4.5: Number of jets in the $2\ell(\mathrm{SS})0\tau_{\mathrm{had}}$ channel [5]. Bins with larger contributions from the q mis-id background contain more electron events. Events with electrons are more likely to have fewer jets rather than more.

The contributions from non-prompt light leptons are estimated with a data-driven method, the so-called matrix method [5]. It works by estimating the rate of selecting prompt and non-prompt leptons between lepton definitions that are loose-but-not-tight or tight. The loose definition is naturally enriched in non-prompt contributions which allows to select side-bands close to the signal regions with large statistics. The rates are measured in $t\bar{t}$ events with two light leptons of same charge for the non-prompt

and opposite charge for the prompt leptons and lower jet multiplicity than the signal regions. The estimated non-prompt background is then the events in the side-bands reweighted by the appropriate combination of rates. This procedure assumes that the rates are compatible between regions. The rates are parametrised as functions of lepton p_{T}, the number of b–tagged jets and the distance between lepton and nearest jet. The total uncertainty of the estimate is 20-30 %.

The non-prompt background in the $2\ell(\mathrm{SS})1\tau_{\mathrm{had}}$ channel is estimated with a simpler method due to low statistics in the potential side-bands. The fake factor method that is used is similar to the matrix method. The difference is that prompt leptons are estimated from MC instead of data. The method is described in Chapter 6.

Misassignment of reconstructed electron charge in $t\bar{t}$ events

Electrons have a small probability to be reconstructed with the wrong charge. This can happen by two distinct processes. One is bremsstrahlung of an electron, where the photon subsequently converts into an electron pair ($e^{\pm} \to e^{\pm}\gamma^{*} \to e^{\pm}e^{+}e^{-}$). It is possible that the wrong track is associated with the EM clusters, and thereby the charge is misassigned. The probability of this process depends on the amount of material that is traversed and therefore increases with larger absolute pseudorapidity.

The other source of charge misassignment is due to too small curvature of the track. The direction of the curvature of the track determines the charge of the traversing particle. When the curvature is small the direction can be misassigned. This happens for electrons with high p_{T} which produce less curved tracks. The combined effect of both sources is that the rate of misassignment increases for larger values of both $|\eta|$ and p_{T}.

This background is called "q mis-id" in the plots and only relevant for the $2\ell(\mathrm{SS})0\tau_{\mathrm{had}}$ channel and to a lesser extent in $2\ell(\mathrm{SS})1\tau_{\mathrm{had}}$. This is because in the $2\ell(\mathrm{SS})0\tau_{\mathrm{had}}$ channel a process with large cross section ($t\bar{t}$) is promoted to be selected by the charge misassignment. In events with three leptons there is no process with much larger cross section that could be promoted in such a way. Here $2\ell(\mathrm{SS})1\tau_{\mathrm{had}}$ is a slight exception because $t\bar{t}$ events with a jet faking a hadronic tau can contribute via this process.

The rate of assigning the wrong charge to tightly identified electrons is measured with $Z \to ee$ events. Electrons pairs with an invariant mass within 10 GeV of the Z mass are selected and the number of events with same-charge pairs N^{SS} ($e^{\pm}e^{\pm}$) is compared to the number of events with all pairs N. The relation between these numbers can be expressed with the charge misidentification rate ε as

$$N^{SS} = 2\varepsilon(1-\varepsilon)N. \tag{4.1}$$

The rate is estimated from these events in bins of p_{T} and $|\eta|$ using a likelihood fit. As expected, the rates are lower for central, low-p_{T} (≈ 15 GeV) electrons at $5 \cdot 10^{-5}$ than for high-p_{T}, forward ($|\eta| > 2$) electrons at 10^{-2} (see Fig. 4.6).

The contribution of charge misassignment is finally estimated by reweighting events with the $2\ell(\mathrm{SS})0\tau_{\mathrm{had}}$ or $2\ell(\mathrm{SS})1\tau_{\mathrm{had}}$ selection but with an inverted electron charge cut.

The uncertainties on the rates are approximately 30 % and arise mainly from a validity test performed by comparing the m_Z distribution of $Z \to ee$ events measured with same-charge electron pairs to the predicted events with same-charge pairs.

Misassignment of muon charge is negligible due to the additional track measurement in the muon spectrometer. This provides a longer track and lever arm to estimate the track parameters. Another factor is that the solenoid and toroid magnetic fields are perpendicular to each other. So both central and forward muons experience a significant magnetic field component which gives a curvature to the track.

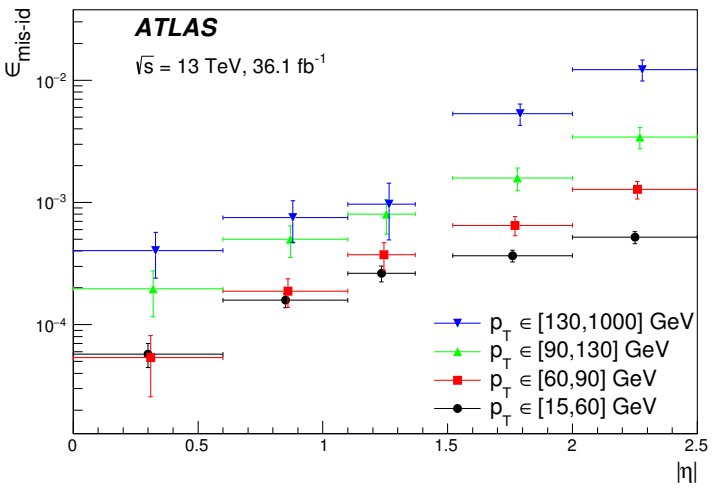


Figure 4.6: Charge misassignment rate of selected electrons as function of $|\eta|$ for several p_{T} ranges. Error bars include statistical and systematic uncertainties [5].

$t\bar{t}W$ and $t\bar{t}Z$

Events with a top-quark pair and additional W– or Z–boson radiation are called $t\bar{t}W$ and $t\bar{t}Z$, or $t\bar{t}V$ jointly. An example of a Feynman diagram of the $t\bar{t}W$ process is shown in Fig. 4.7. In all signal regions $t\bar{t}V$ are irreducible backgrounds.

As $t\bar{t}W$ is a charged final state it is only produced via quark initial states. Furthermore, the cross section of $t\bar{t}W^+$ is larger than $t\bar{t}W^-$ because of the relative momentum fractions of up– and down–quarks in the colliding protons. The diagrams for $t\bar{t}Z$ are similar to the ones of $t\bar{t}W$ and $t\bar{t}H$ with a Z replacing the respective bosons.

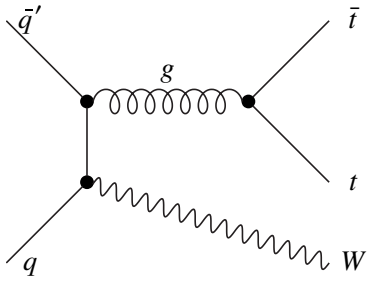
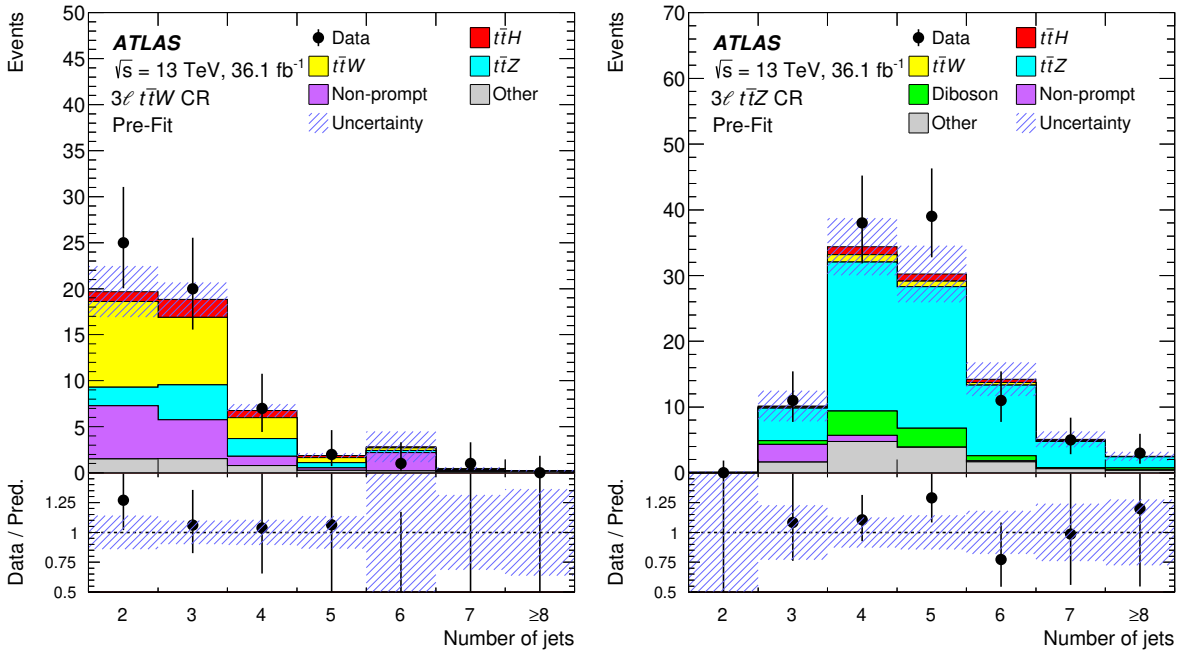


Figure 4.7: Example of Feynman diagram of the $t\bar{t}W$ process.

There are two dedicated control regions for $t\bar{t}V$ defined for events with three light leptons. The definition of CR is performed with the BDTs that are trained and discriminate against each $t\bar{t}V$ process. The number of jets in these regions is plotted in Fig. 4.8, which shows good agreement between the data and the prediction for $t\bar{t}V$, both in terms of total events and as a function of the number of jets.

Diboson

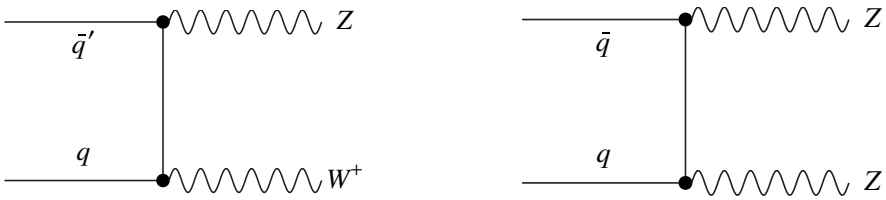
Events with two weak bosons are another irreducible background. Specifically, WZ and ZZ where the bosons decay leptonically can result in final states with the same number of leptons as $t\bar{t}H$. In order to be



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

Figure 4.8: Control regions for the (a) $t\bar{t}W$ and (b) $t\bar{t}Z$ processes in events with three light leptons. The shown observable is the number of jets. $t\bar{t}V$ is normalised to its theoretical prediction. [5]

selected, additional QCD radiation is necessary in diboson events and so requirements on the number of jets and b–tagged jets are effective in reducing it. Two examples for WZ and ZZ production Feynman diagrams are given in Fig. 4.9.



(a) WZ (b) ZZ

Figure 4.9: Examples of Feynman diagrams of (a) WZ production and (b) ZZ production.

Hadronically decaying tau leptons faked by other signatures

The reconstruction and identification of quark or gluon initiated jets as τ_{had} is called fake because it was not initiated by a true tau lepton decay. Such fake hadronic taus can be selected in any event that contains jets and can mimic the final state of the signal. For example, a dileptonic $t\bar{t}$ event where one lepton is light (electron or muon) and the other is a tau lepton that decayed hadronically is selected with an additional fake τ_{had} in the $1\ell 2\tau_{\text{had}}$ channels. In that case the background belongs to the reducible category.

Another example is $t\bar{t}Z$ in the $3\ell 1\tau_{\text{had}}$ channel. It can be selected both with and without a fake τ_{had} and depending on the $t\bar{t}$ and Z decay is attributed to the irreducible or reducible background.

Other sources of fake τ_{had} are electrons and muons mimicking the signature of 1–prong τ_{had}. These electrons and muons fail their respective identification and quality requirements as otherwise they would have had precedence in the overlap removal procedure. Electrons leave a calorimeter signature which can have similar shower shapes to a charged pion of a τ_{had}. Muons have to undergo unusually large energy loss in the calorimeters to be able to fake a τ_{had}. The relative frequency of the light lepton fakes compared to the jet fakes is low because real leptons are rarer than jets. The fractional contributions of the different types of fakes in all channels with hadronic taus are shown in Fig. 4.4(b).

The background from fake τ_{had} is estimated with information from data in all channels. The methods used in $2\ell(\text{OS})1\tau_{\text{had}}$ and $2\ell(\text{SS})1\tau_{\text{had}}$ are described in Chapters 5 and 6 and also briefly in the following.

In the $2\ell(\text{OS})1\tau_{\text{had}}$ channel a data-driven method that reweights side-band data events is used. The fake τ_{had} background in $2\ell(\text{SS})1\tau_{\text{had}}$ and $3\ell1\tau_{\text{had}}$ is using truth-matched MC which contains fake τ_{had} and is rescaled with a factor derived from the $2\ell(\text{OS})1\tau_{\text{had}}$ estimate. So the estimates in all three channels which select one τ_{had}, are related.

In the $1\ell2\tau_{\text{had}}$ channel the fake τ_{had} estimate is exploiting the fact that there is a pair of τ_{had} which is selected with opposite charges. The fake τ_{had} consists dominantly of $t\bar{t}$ events with one or two fake τ_{had} and a prompt light lepton. The fake rate from jets does not depend on the charge of the τ_{had} and thus a fake hadronic tau originating from a jet has no charge correlation to a real τ_{had} in the event. This property is used by inverting the charge cut on the τ_{had} pair such that they have same-sign charge. The SS events do not contain any significant $t\bar{t}H$ signal and are the estimate of the fake τ_{had} background in this channel.

Other backgrounds

Other backgrounds only have small contributions in the signal regions. They include mainly rare processes with top quarks like single top, tZ, tWZ, $t\bar{t}WW$, triple and quadruple top-quark production and tH. The latter is counted as background despite being sensitive to the top-Yukawa coupling because it presents itself differently. The present analysis is optimised for $t\bar{t}H$ and tH only contributes negligibly. Many of these backgrounds are irreducible but have small predicted cross section, either inclusively or in the phase space selected by the analysis.

Processes without top-quarks that also contribute are weak triple boson production and Drell-Yan Z+jets. These contribute with additional QCD radiation to produce enough jets to be selected. Furthermore, Z+jets always requires at least one non-prompt lepton or fake τ_{had} in order to contribute to any signal region.

Background summary

The relative contributions of the backgrounds to all signal regions are shown in Fig. 4.10. One can see considerable differences in the signal-to-background ratios and most channels include discriminants that further enhance the ratio.

In general the strategy of selecting a large number of leptons or unlikely charge combinations works well for light leptons, while channels with τ_{had} suffer from the fake background and rely on larger expected statistics to recover significance. The statistical significances ($S/\sqrt{B}$) and purities (S/B) of all signal regions are shown in Fig. 4.11.

4.3 Current state of $t\bar{t}H$ analyses

Several searches for $t\bar{t}H$ have been performed by both ATLAS and CMS in the decay channels described above. All results are summarised in Tab. 4.3. The combination of all $t\bar{t}H$ analyses performed in Run 1

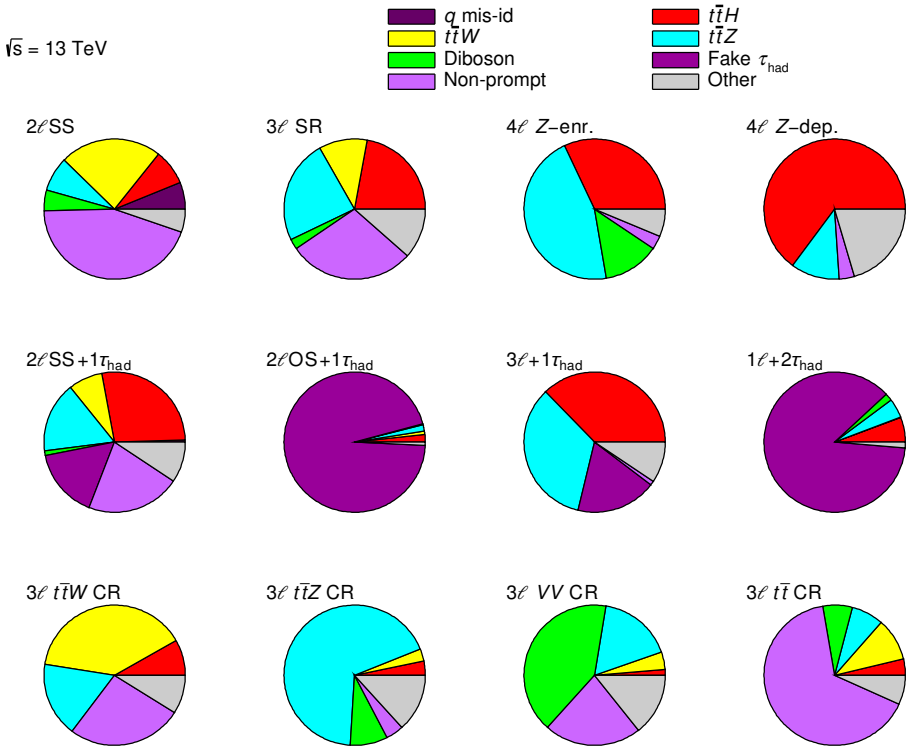


Figure 4.10: Fractions of processes selected in all signal and control regions of the multilepton analysis.

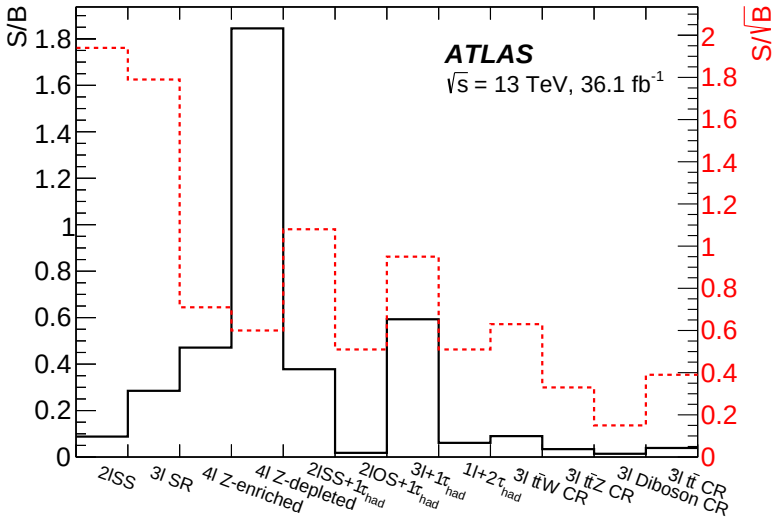


Figure 4.11: Purity (S/B) and statistical significance ($S/\sqrt{B}$) of ML channels [5].

by ATLAS and CMS yielded a mild observed excess over the background-only ($\mu = 0$) as well as the SM ($\mu = 1$) expectation [42]. This excess ensured that the interest in measuring the $t\bar{t}H$ process in Run 2 was further increased. Both ATLAS and CMS observed significant evidence for an excess over the background-only hypothesis by analysing data recorded in 2015 and 2016 of about $36\,\text{fb}^{-1}$ [5, 124]. A significant observation was achieved almost concurrently also by both collaborations by different strategies. CMS combined the $36\,\text{fb}^{-1}$ data with the Run 1 result [125]. ATLAS did the same but also added approximately $44\,\text{fb}^{-1}$ of data recorded in 2017 to the analyses in the ZZ and $\gamma\gamma$ channels [126].

The author of this thesis made major contributions to the multilepton analysis in Run 1 [127] that entered the combined result [42]. The work described in this thesis entered the Run 2 multilepton result [5] that was crucial to both the evidence and observation.

Analysis	ATLAS		CMS	
	$\mu_{t\bar{t}H}$	Obs. (exp.) significance	$\mu_{t\bar{t}H}$	Obs. (exp.) significance
Run 1 combination [42]	$2.3\,^{+0.7}_{-0.6}$	4.4 (2.0)	$2.3\,^{+0.7}_{-0.6}$	4.4 (2.0)
ttHbb [119, 128]	0.8 ± 0.6	1.4 (1.6)	0.72 ± 0.45	1.6 (2.2)
ttHML [5, 124]	$1.6\,^{+0.5}_{-0.4}$	4.1 (2.8)	$1.23\,^{+0.45}_{-0.43}$	3.2 (2.8)
ttHZZ [122, 129]	< 7.1 (95% CL)	–	$0.00\,^{+1.18}_{-0.00}$	–
tt$H\gamma\gamma$ [121, 130]	0.5 ± 0.6	1.0 (1.8)	$2.2\,^{+0.9}_{-0.8}$	3.3 (1.5)
2015–2016 combination [5, 131]	1.2 ± 0.3	4.2 (3.8)	$1.18\,^{+0.31}_{-0.27}$	–
Run 1+2015–2016 combination [125]			$1.26\,^{+0.31}_{-0.26}$	5.2 (4.2)
2015–2017 combination [126]	$1.32\,^{+0.28}_{-0.26}$	5.8 (4.9)		
Run 1+2015–2017 combination [126]	–	6.3 (5.1)		

Table 4.3: Summary of $t\bar{t}H$ analyses by ATLAS and CMS. "Run 1 combination" refers to results by both ATLAS and CMS with data of the years 2011–2012 and numbers are repeated in both experiment columns for ease of reading. "2015–2016 combination" refers to the combination of all analyses (ttHbb, ttHML, ttHZZ, tt$H\gamma\gamma$) of each experiment separately. The 2015–2017 combination by ATLAS only updated the ZZ and $\gamma\gamma$ channels with 2017 data. The first column in each experiment shows the best-fit signal strengths and limits at 95% confidence level, where applicable. The second column shows the observed and expected significance with respect to the background-only hypothesis. Significances marked with a dash are not given by the cited works and not meaningful for null observations/limits.

4.4 Dilepton+τ_{had} analysis

Both the $2\ell(\text{SS})1\tau_{\text{had}}$ and $2\ell(\text{OS})1\tau_{\text{had}}$ channels are conceptually similar to the $3\ell0\tau_{\text{had}}$ one in that they select three objects reconstructed from charged leptons. The substitution of a τ_{had} for one of the light leptons, however opens the door to much increased backgrounds from jets faking the τ_{had}. The $2\ell(\text{SS})1\tau_{\text{had}}$ channel partially remedies this issue by adopting the same charge cut on the light leptons that works well in $2\ell(\text{SS})0\tau_{\text{had}}$, it is thus a hybrid between the $2\ell(\text{SS})0\tau_{\text{had}}$ and $3\ell0\tau_{\text{had}}$ channels.

Unfortunately, the consequence of the same sign charge cut is reduced acceptance for $t\bar{t}H$ events. When considering the two light leptons and the τ_{had} in both channels as generic charged objects, there

are eight different permutations of selecting three charged objects. Six of these have an absolute sum of charges equal to one, as is additionally required in the event selection. Only two of the six permutations have same-sign light leptons, while four permutations have opposite-sign light leptons. The additional same sign charge cut is simply more stringent than the inclusive requirement over all three objects. These permutations are illustrated in Tab. 4.4.

Examples of Feynman diagrams of $t\bar{t}H$ events that can be selected in $2\ell(\mathrm{SS})1\tau_{\mathrm{had}}$ are shown in Fig. 4.12. One notes how the light leptons are, one each, originating from the H system and the $t\bar{t}$ system. The $2\ell(\mathrm{OS})1\tau_{\mathrm{had}}$ channel can select events with the same diagrams with appropriate rearranging of leptonic and hadronic W decays. Additionally, it can select events with diagrams shown in Fig. 4.13. Here it should be noted that both light leptons originate from a single system, either the H or the $t\bar{t}$ one. The difference in signal acceptance between both channels can be seen in Tab. 4.2.

Historically, the $2\ell(\mathrm{SS})1\tau_{\mathrm{had}}$ channel is the more successful one of the two. It was part of the Run 1 multilepton result [127] and all subsequent ones. This is despite the fact that feasibility studies for conditions similar to the contemporary ones concluded that *"this channel can only be used as a corroborative channel for determination of the top Yukawa coupling, but not as a main discovery channel"* [132]. However, this statement is at present true for all $t\bar{t}H$ multilepton channels.

On the other hand, the $2\ell(\mathrm{OS})1\tau_{\mathrm{had}}$ final state is for the first time analysed in the 2015–2016 multilepton analysis by ATLAS [5] and through this thesis.

$\ell = e, \mu$	Electric charge	SS	SS	OS	OS	OS	OS
$t \to W^+ \to \ell^+$	+1	×			×		×
$\bar{t} \to W^- \to \ell^-$	−1		×		×	×	
$H \nearrow W^+, \tau^+ \to \ell^+$	+1	×		×		×	
$H \searrow W^-, \tau^- \to \ell^-$	−1		×	×			×
Charge sum	0	+2	−2	0	0	0	0

Table 4.4: Combinations of selecting pairs of charged particles with same-sign (SS) and opposite-sign (OS) electric charge in a simplified $t\bar{t}H$ event. Each row stands for the detectable leptonic decay products (e, μ) of the H boson and the top quarks. The H decay proceeds either via a $H \to WW^*$ or $H \to \tau\bar{\tau}$ decay. The columns represent possible combinations of creating the selected SS or OS pairs. Crossed entries represent reconstructed, identified and selected leptons, while blank entries can be absence of lepton due to a hadronic W decay, leptons outside detector acceptance or not satisfying selection criteria.

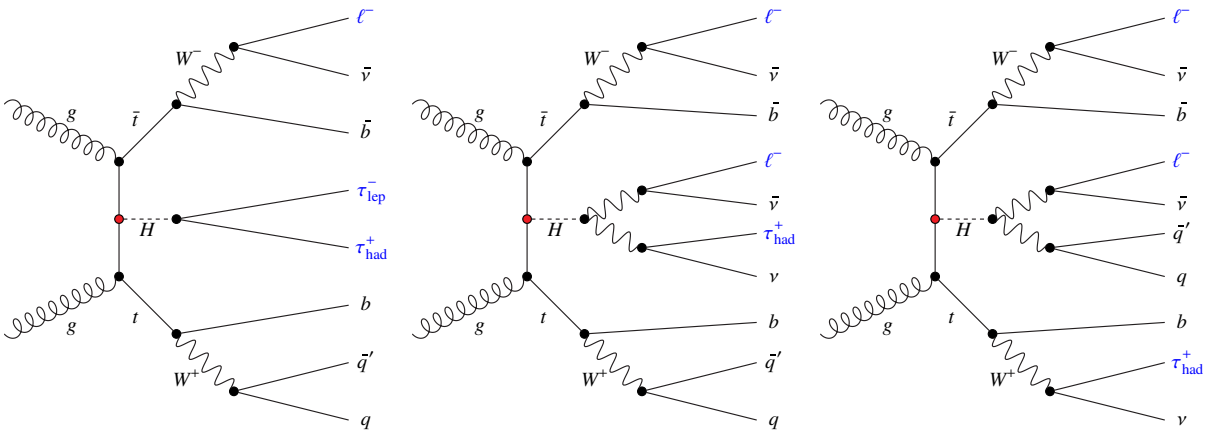


Figure 4.12: Examples of Feynman diagrams of the $t\bar{t}H$ process selected in the $2\ell(\mathrm{SS})1\tau_{\mathrm{had}}$ channel. ℓ and τ_{lep} stand for electron or muon. Observables marked blue are counted to define and name the analysis categories.

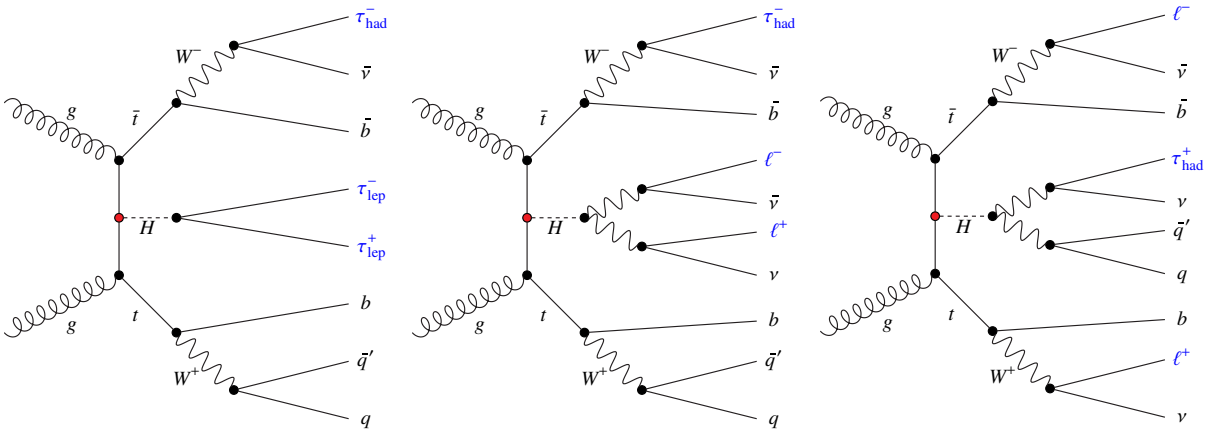


Figure 4.13: Examples of additional Feynman diagrams of the $t\bar{t}H$ process selected in the $2\ell(\mathrm{OS})1\tau_{\mathrm{had}}$ channel. ℓ and τ_{lep} stand for electron or muon. Observables marked blue are counted to define and name the analysis categories.

CHAPTER 5

Analysis of $t\bar{t}H \rightarrow 2\ell(\mathrm{OS})1\tau_{\mathrm{had}}$ final state

The final state with $2\ell(\mathrm{OS})1\tau_{\mathrm{had}}$ is a more likely signature of the $t\bar{t}H$ process than $2\ell(\mathrm{SS})1\tau_{\mathrm{had}}$ as described in the preceding chapter in Sec. 4.4. This means that the $t\bar{t}H$ branching ratio to the $2\ell(\mathrm{OS})1\tau_{\mathrm{had}}$ final state and consequently the number of selected signal events is larger when compared to the $2\ell(\mathrm{SS})1\tau_{\mathrm{had}}$ final state. The disadvantage is that it comes with much larger background from neutral resonances like Z+jets, and $t\bar{t}$ that can decay into a pair of charged leptons which have opposite charges. However, the large number of selected events also allows to perform robust estimation and validation of the background predictions.

5.1 Object selection

Reconstructed electrons and muons only have to satisfy basic quality and isolation requirements in this channel to maximise the number of selected events. Opposite charge light leptons are so abundant that contributions from fake or non-prompt leptons are negligible relative to $t\bar{t}$. Therefore there is no merit to having a tighter selection requirement.

The τ_{had} identification criteria are optimised for the special conditions in this channel. The abundance of jets both non-tagged and b–tagged is the motivation for the τ_{had} selections, in particular the ones rejecting b–tagged and pileup jets, described in Sec. 3.4.7.

5.2 Event selection

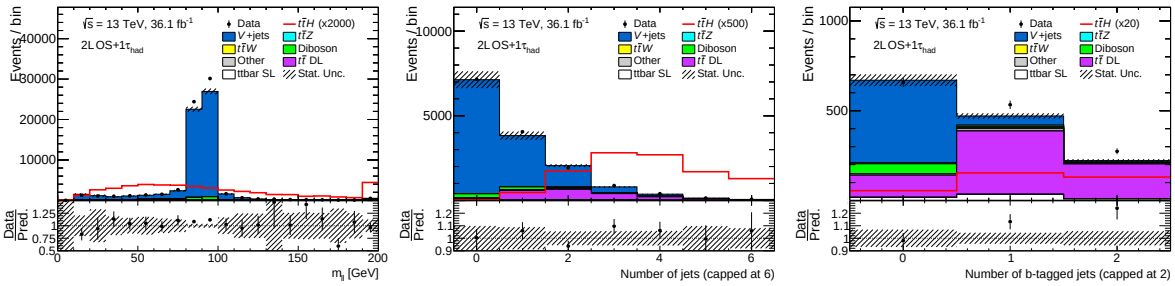
A final state with two light leptons of opposite charge can be produced in many different ways at a hadron collider. The most abundant sources of such events are resonances such as the Z boson, and top quark pair production.

The resonant background can be efficiently rejected by making requirements on the invariant mass, $m_{\ell\ell}$, of the light lepton pair. All events with a light lepton pair of same flavour that have an invariant mass close to the Z mass ($|m_{\ell\ell} - 10\,\mathrm{GeV}| < 91.2\,\mathrm{GeV}$) or below 12 GeV are removed.

To further reduce $t\bar{t}$ and Z+jets, the number of jets is required to be at least three. At least one of the jets has to be b–tagged. These selection cuts are summarised in Tab. 5.1 and are called signal region (SR) in the following. Figure 5.1 shows the distributions of each observable used to define the SR before the cut on each observable. Subfigure (a) starts with only the requirements on the number of leptons and each subsequent subfigure adds the requirement on the previous observable.

Light leptons	τ_{had}	Jets
=2	=1	≥ 3
opposite charge	$p_{\text{T}} > 25\,\text{GeV}$	≥ 1 b–tagged
$p_{\text{T}}(\ell_1) > 25\,\text{GeV}$		
$p_{\text{T}}(\ell_2) > 15\,\text{GeV}$		
$\lvert m_{ee/\mu\mu} - m_Z \rvert > 10\,\text{GeV}$		
$m_{ee/\mu\mu} > 12\,\text{GeV}$		

Table 5.1: Selection criteria in the $2\ell(\text{OS})1\tau_{\text{had}}$ channel.



(a) $m_{\ell\ell}$ to be cut not near m_Z (b) n. of jets to be cut ≥ 3 (c) n. of b–tagged jets to be cut ≥ 1

Figure 5.1: Distributions of (a) $m_{\ell\ell}$, (b) number of jets and (c) number of b–tagged jets in $2\ell(\text{OS})1\tau_{\text{had}}$ before the cuts on these observables. Each subsequent plot includes the cut on the previous observable.

5.3 Fake τ_{had} estimate

The background is dominated by processes with two real light leptons and one fake τ_{had}. The corresponding physical processes that contribute are $t\bar{t}$ with 90% and Z+jets with 10%. The origin of the τ_{had} fakes are jets initiated by quarks, gluons as well as jets from pileup interactions. These jets can have similar detector signatures to real hadronic tau decays. Such signatures with few tracks in a narrow cone are however unlikely for jets and the probability for them to occur is not well modelled in MC. This can be seen in the two rightmost bins of Fig. 5.1(c), where MC predicts approximately 20 % fewer events than are observed in data. In order to have a more accurate prediction of the background a data-driven method is employed.

5.3.1 Data-driven estimate

The fake factor (FF) method is used to estimate the background with a fake hadronic tau. It uses data events in a side-band region that are reweighted with appropriate factors to describe the background in the signal region. The reweighting factors are determined from data in control regions. Many analyses in ATLAS and CMS successfully use this method for estimation of fake τ_{had} backgrounds (e.g. Refs. [133, 134]).

The side-band is defined by reversing the requirement on the τ_{had} identification discriminant. Such side-band τ_{had} replace the identified τ_{had} of the event selection. The events that are selected by this modified τ_{had} definition are enriched in fake τ_{had}.

To derive the necessary reweighting factors (in the following fake factors (FF)) a control region is

defined by changing the requirement on the number of selected jets. This control region (in the following extraction region of FF (ER)) also has the accompanying side-band. The FF is defined as the number of events with the nominal τ_{had} definition divided by the number of events with the side-band τ_{had} definition. It is instructive to think of the two regions in terms of their place in the fraction as numerator and denominator. Figure 5.2 shows a sketch of all regions and how they relate to each other. The FF is calculated as

$$\frac{\text{number of events with } \tau_{\text{had}} \text{ that passes identification criteria}}{\text{number of events with } \tau_{\text{had}} \text{ that fails identification criteria}} = \frac{N_A}{N_B} = \text{FF}_{AB} \equiv \text{FF}_{CD} = \frac{N_C}{N_D}. \quad (5.1)$$

The number of estimated background events in the signal region is thus

$$N_A = \text{FF} \cdot N_B. \quad (5.2)$$

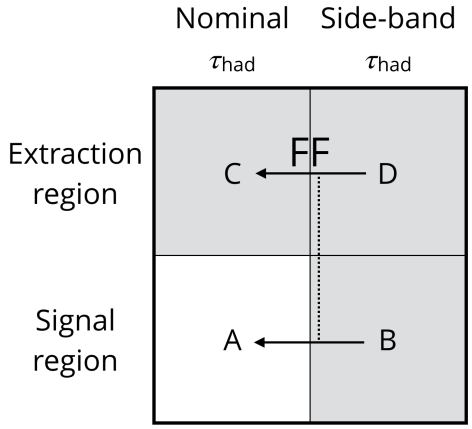


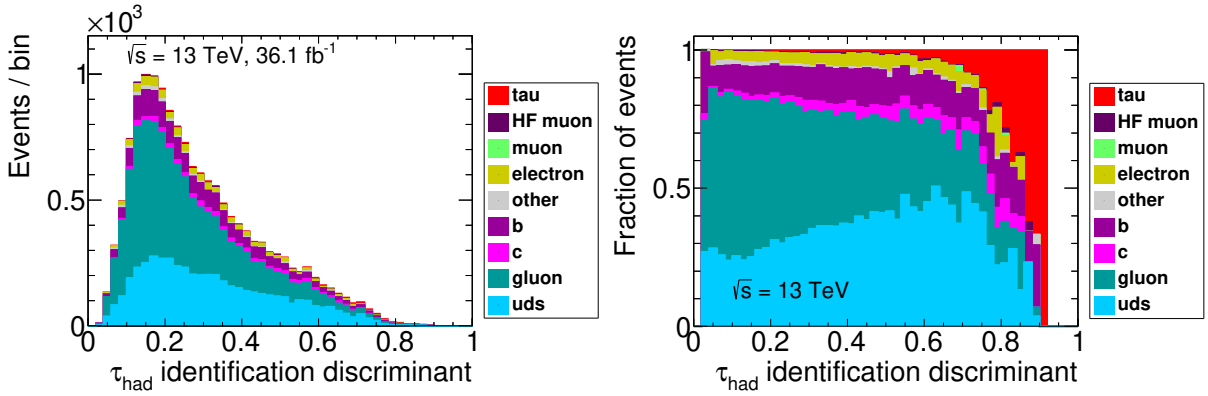
Figure 5.2: Sketch of FF method. The arrows represent the ratio of events that define the fake factor FF. The dotted line represents that the FF is extrapolated from the extraction region to the signal region.

Assumptions

The assumption that is made in this estimation technique is that the fake factors of the signal and extraction regions are the same. For this to hold the following should be true:

1. the type of τ_{had} fake is similar across all four regions,
2. contributions from real τ_{had} are negligible in all four regions or if this is not the case,
3. contributions from real τ_{had} can be subtracted using truth-matched MC.

Figure 5.3 shows the true (MC) composition of the reconstructed τ_{had} candidates as a function of their identification discriminant in the SR-like event selection without the τ_{had} identification requirement which is equivalent to the sum of regions A and B. Low values of the discriminant mean that the reconstructed τ_{had} is unlikely to be a true τ_{had}, while high values mean the opposite. The threshold for the nominal identification requirement is 0.6-0.7, depending on the p_{T} and number of tracks of the τ_{had} candidate. One can see that the composition changes significantly between low and high values, as expected. In order to satisfy assumption 1. from above an additional requirement of >0.35 is placed on

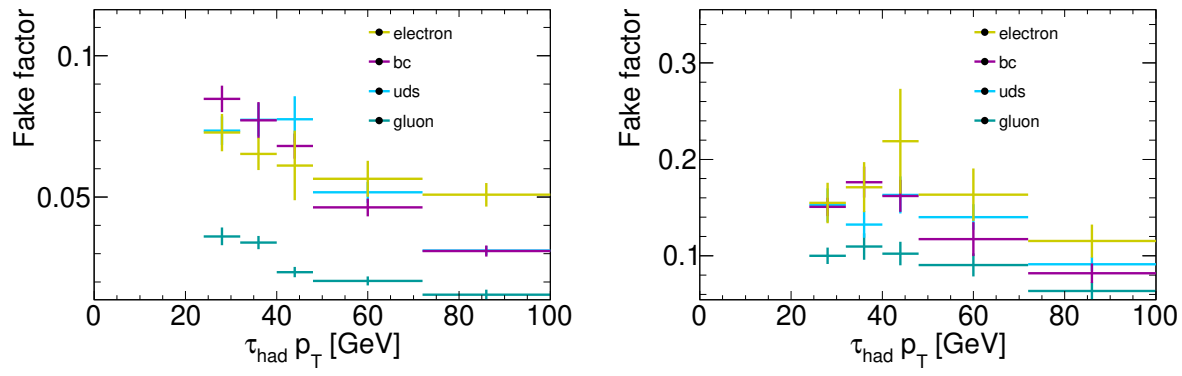


(a) normalised to expected number of events at 36.1 fb^{-1}

(b) normalised arbitrarily per bin

Figure 5.3: True composition of reconstructed τ_{had} candidates as function of identification score in the $2\ell(\text{OS})1\tau_{\text{had}}$ SR-like selection (without τ_{had} identification requirement and equivalent to the sum of regions A and B). The left subfigure shows each component normalised to the expected number of events at 36.1 fb^{-1}. Therefore the true "tau" component is barely visible as the plot is dominated by fake τ_{had}. The right subfigure shows the same distribution where each bin is normalised independently to show the relative contributions of each component. Here one can see that the nominal identification requirement of 0.6-0.7 selects bins with large fractions of "tau" events. The truth flavour labelling is done in the same way as for Fig. 3.9.

reverse-identified τ_{had} candidates. This ensures that the fraction of quark- to gluon-initiated jets selected as τ_{had} is comparable between nominal and side-band τ_{had} definitions. Furthermore the cut selects jets that have a detector signature more similar to a hadronic tau decay. Figure 5.4 shows the fake factor with and without the additional requirement on the tau identification discriminant. Without the requirement there are significant differences between the fake factors for gluon- and quark-initiated jets, whereas with the requirement the fake factors are similar. The remaining differences in composition thus have less impact on the fake factors. The difference in the absolute values of FF in the two cases (with and without additional requirement) do not affect the final estimated number of events as it is compensated by the corresponding change in the definition of region B.



(a) without additional requirement

(b) with additional requirement

Figure 5.4: Fake factors (a) without and (b) with additional identification requirement on reverse identified τ_{had} candidates as described in the text. Different values of FF between (a) and (b) are compensated by the corresponding different definition of region B.

It is not possible to define regions of phase space that are solely comprised of fake τ_{had} candidates. Therefore the contribution from real τ_{had} is subtracted in all regions. Equation 5.1 is modified such that $N_X = N_X^{\text{data}} - N_X^{\text{MC real } \tau_{\text{had}}}$ for each region X={B,C,D}. Those events that are truth-matched to real τ_{had} in the SR side-band are shown in Fig. 5.5 weighted by the FF. They are the contribution that is subtracted in the SR. The number of events to be subtracted is less than 13 which is less than 2 % of the total fake τ_{had} estimate.

The $t\bar{t}H$ signal present in the SR side-band is subtracted whether the τ_{had} is truth-matched or not i.e., all of the $t\bar{t}H$ signal. It means that all $t\bar{t}H$ events can be interpreted as signal, rather than counting the part with fake τ_{had} as part of the background. More subtracted $t\bar{t}H$ in the side-band would result in less estimated background from fake τ_{had} and possibly an overestimate of the signal strength. This makes the fake estimate dependent on the signal strength. To account for this effect the amount of subtracted $t\bar{t}H$ events in the side-band is varied coherently with the amount of $t\bar{t}H$ events in the SR for the statistical interpretation (see Sec. 7.3.2). The equivalent effect of this is a reduction of the selected $t\bar{t}H$ events in the signal region of 25 %.

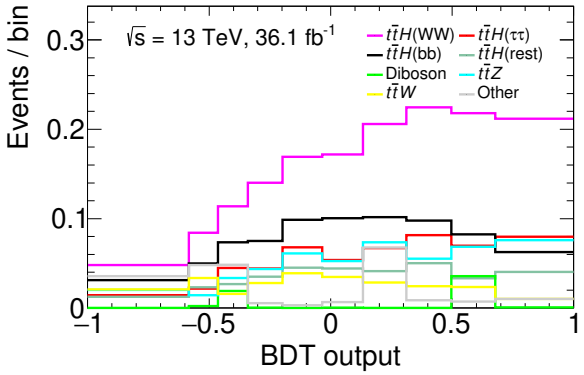


Figure 5.5: FF-weighted backgrounds with true τ_{had} and $t\bar{t}H$ in the SR side-band region. These events are subtracted in the SR with nominal τ_{had} definition.

Choice of extraction region

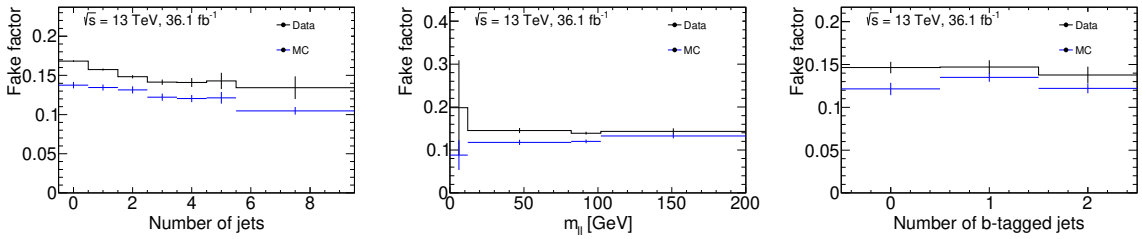
To compute the FF, appropriate regions have to be chosen that are orthogonal to the SR. A region is appropriate when the FF that is extracted from it, is compatible with the FF of the SR. The SR is defined by cuts on only three observables, $m_{\ell\ell}$ and number of (b–tagged) jets, so to define new regions that are not overlapping with the SR, one has to modify the requirements on those. Figure 5.6 shows the FF as a function of $m_{\ell\ell}$ and number of (b–tagged) jets. Each bin of these plots define regions of phase space that could potentially be used as FF extraction region. A constant value of FF in these plots would indicate appropriateness of the corresponding event selection for use in the FF extraction.

Figure 5.6(a) contains the loosest selection with only requirements on the number and charges of leptons. Specifically there is no cut on the number of jets, so that FF(N_{jet}) can be defined for all multiplicities. One can see higher values of the FF for low jet multiplicities. However, at 2 and more jets the FF becomes constant. Therefore, events with any number of jets above 2 can be used to extract the FF.

Figure 5.6(b) narrows the selection to events with at least 3 jets. The FF is plotted as a function of the invariant mass of the light lepton pair. It is constant near the Z mass as well as away from it. So it is possible to change the cut on $m_{\ell\ell}$ to create an alternative region.

Figure 5.6(c) further includes the Z–veto cut and shows the FF as function of the number of b–tagged jets. Again the FF is constant. It is concluded that any b–tagged jet multiplicity is valid for the fake estimate.

Here discrepancies between data and MC are of no importance in this consideration and are, in fact, another argument for using the data-driven method.



(a) only lepton selection (b) NJet ≥ 3 (c) NJet ≥ 3, Z–veto

Figure 5.6: Fake factors as function of (a) number of jets, (b) invariant mass of light lepton pair, (c) number of b–tagged jets in data (black) and MC (blue) events. The subcaptions describe the cuts that are applied to the events and are further described in the text.

Using the previous conclusions three extraction regions are defined. The nominal ER is defined by the same selection as the SR except that no b–tagged jets are allowed in the event. The processes that are selected are 90% Z+jets and 10% $t\bar{t}$. Two additional ERs are defined, called "Z enriched" and "$t\bar{t}$ enriched". One that inverts the Z mass veto of the light lepton pair and also vetoes b–tagged jets, thereby further increasing the fraction of Z+jets events to close to 100%. The other ER changes the SR selection by requiring exactly 2 jets. By keeping the b-tag requirement it is pure in $t\bar{t}$ events and comparable in composition to the SR.

These definitions are summarised in Tab. 5.2. The distributions of p_{T} of the τ_{had} and the background compositions in all ERs can be seen in Appendix A.

Region	Number of jets	Number of b–tagged jets	$\lvert m_{ee/\mu\mu} - m_Z \rvert$ in GeV
Signal	≥ 3	≥ 1	>10
Extraction	≥ 3	$= 0$	>10
Z+jets enriched	≥ 3	$= 0$	<10
$t\bar{t}$ enriched	$= 2$	≥ 1	>10

Table 5.2: Definitions of signal and fake factor extraction regions.

Parametrisation of fake factors

Rather than using one inclusive value of FF in Eq. 5.1 one can compute it as a function of an observable. The full equation then becomes

$$\frac{N_A^{\mathrm{data}}(p_{\mathrm{T}}) - N_A^{\mathrm{MC\ real}\tau_{\mathrm{had}}}(p_{\mathrm{T}})}{N_B^{\mathrm{data}}(p_{\mathrm{T}}) - N_B^{\mathrm{MC\ real}\ \tau_{\mathrm{had}}}(p_{\mathrm{T}})} = \frac{N_C^{\mathrm{data}}(p_{\mathrm{T}}) - N_C^{\mathrm{MC\ real}\ \tau_{\mathrm{had}}}(p_{\mathrm{T}})}{N_D^{\mathrm{data}}(p_{\mathrm{T}}) - N_D^{\mathrm{MC\ real}\ \tau_{\mathrm{had}}}(p_{\mathrm{T}})} = \mathrm{FF}(p_{\mathrm{T}}), \tag{5.3}$$

where p_{T} is the transverse momentum of the τ_{had}, to illustrate one possible parametrisation. Properties of the reconstructed τ_{had} are obvious choices for this parametrisation. Apart from the p_{T}, also η and the number of tracks of the τ_{had}, and the average number of bunch crossings in the event are considered. Figure 5.7(a) shows that the FF as a function of η of theτ_{had} is constant and does not need to be further considered. However, p_{T} and number of tracks of the τ_{had}, as well as, the average number of bunch crossings of the event in Figs. 5.7(b) to 5.7(d) are correlated to the FF. The features that can be seen are that side-band τ_{had} candidates with low transverse momentum and one track are more likely to be identified as nominal τ_{had} candidates. In the case of number of tracks, the effect is partly given by design as the identification working points have different fake rejection efficiencies for 1- and 3-prong τ_{had} candidates. Furthermore, events with more pileup activity have a higher probability of having a side-band τ_{had} candidate that is also identified.

When using a FF as a function of an observable that is similar between regions (i.e. the ratio of FFs is constant and equal to one), the estimated number of events is the same as when using an unparametrised FF. This is demonstrated in Tab. 5.3, which shows the predicted number of events for unparametrised FF, as well as parametrised by p_{T} and number of tracks of the τ_{had}, and the average number of bunch crossings. All predictions are compatible with each other and any choice of parametrisation would be valid.

Nevertheless, it is chosen to use p_{T} of the τ_{had} as the parameter of the FF. The reason is that FF (p_{T}) corrects the shape of the p_{T} distribution of the side-band to accurately describe the distribution in the SR. It also means that it is a deliberate choice to ignore the shape of FF (number of tracks) and FF (average number of bunch crossings). The consequence is that those distributions are not well described in the SR. It is permissible to do so because neither observable is used for defining analysis regions or otherwise.

Parametrisation of FF	Predicted fake yields in SR
None	736 ± 35
p_{T} of τ_{had}	746 ± 36
Number of tracks of τ_{had}	726 ± 37
Average number of bunch crossings	733 ± 35

Table 5.3: Predicted yields of fake τ_{had} background in the SR using FFs that are unparametrised, parametrised as function of τ_{had} transverse momentum and number of tracks, and the average number of bunch crossings in an event.

Validity of the method

The validity of the method can be demonstrated by calculating FF (p_{T}) using MC and also applying it to MC in the SR side-band. It provides a description of the fake background in the SR that used only events from control regions. This description should then be compatible with the direct MC prediction in the SR. Any mismodelling of the fakes in MC are present on both sides of this comparison and should therefore cancel. Figure 5.8(a) shows this test. In the plots "FF MC" refers to SR side-band MC events weighted by the FF determined in ER MC events, while "MC" refers to the direct prediction of MC events in the SR. Subfigures (b) and (c) are the result of performing the test separately for τ_{had} fakes from gluon- and quark-initiated jets. The fact that both distributions match well to each other indicates that the test is passed successfully and gives confidence that the method will give an accurate description of the fakes in data.

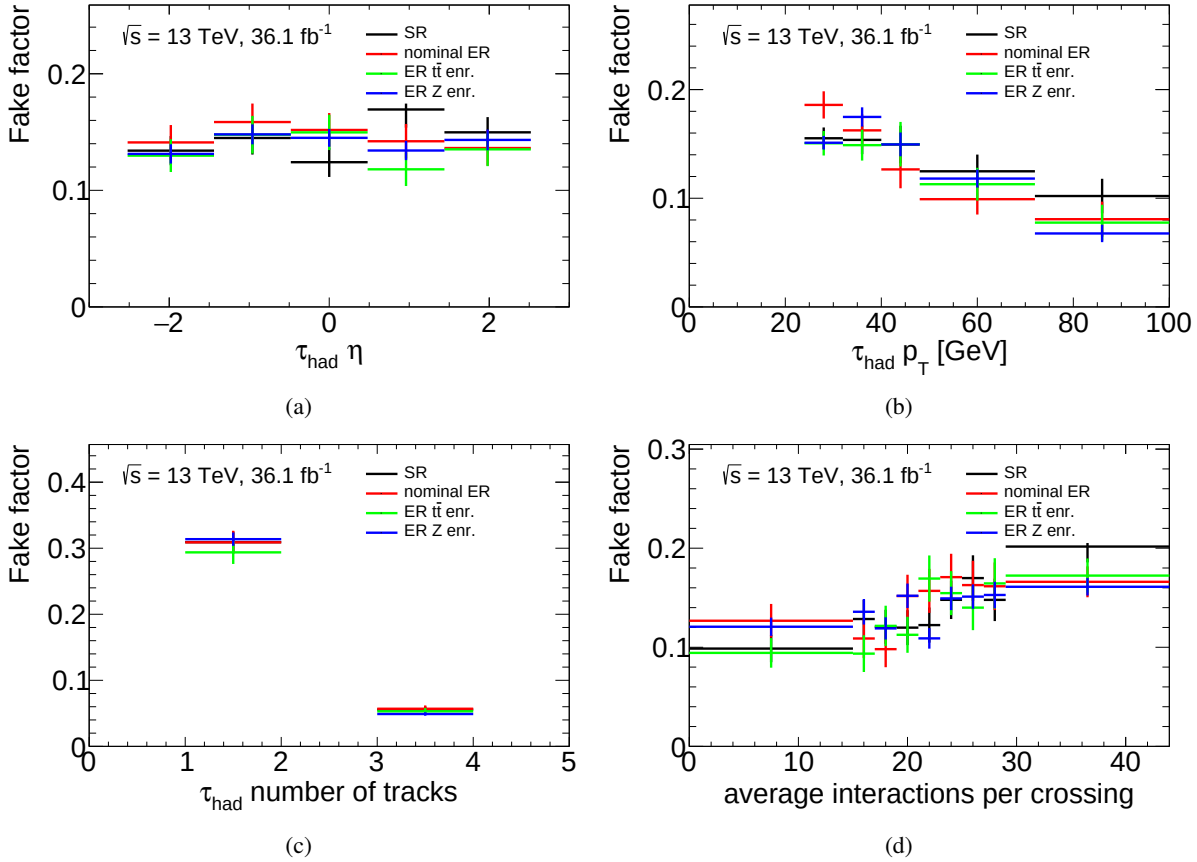


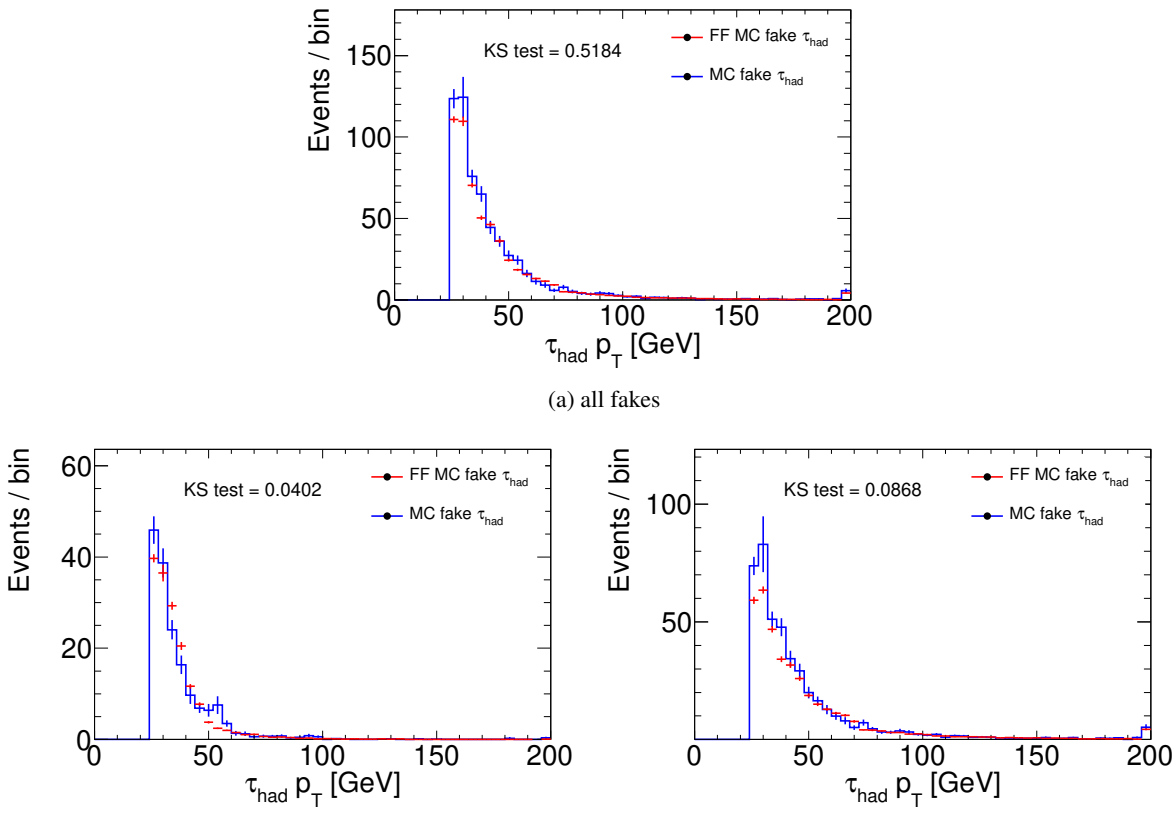
Figure 5.7: Fake factors as function of (a) pseudorapidity, (b) transverse momentum, (c) number of tracks of the τ_{had} and (d) average number of interactions per bunch crossing in data events in the SR, the nominal ER and the two alternative ERs.

5.3.2 Uncertainties

The statistical uncertainty of the fake estimate is propagated from the two event counts in the ER and the event count in the SR side-band that enter the calculation. The largest contribution comes from the smallest event count, which is the number of events in the ER with nominal τ_{had} selection. Table 5.4 shows the expected and observed number of events in all regions. In the row "weighted side-band" the estimated number of events from fake τ_{had} backgrounds is $N_{\text{Data}} - N_{\text{real}\,\tau_{had}} - N_{t\bar{t}H} = 746 \pm 36$. Thus the statistical uncertainty is 4.9 %.

Systematic uncertainties in the fake τ_{had} estimate can arise from

1. the choice of FF extraction region,
2. the difference of fake composition between SR and ER,
3. the choice of observable(s) for the parametrisation as well as their binning,
4. the choice of the reverse identification requirement in the τ_{had} selection,
5. the assumption that real τ_{had} contribution can be subtracted using MC,
6. and $t\bar{t}H$ contamination in the control regions.



(a) all fakes

(b) gluon fakes

(c) quark fakes

Figure 5.8: Validity test of fake factor method. FF MC denotes that the method is applied to MC to predict fakes in the SR, MC denotes the direct prediction of fakes in the SR. The test is performed for (a) all fakes, (b) gluon fakes and (c) quark fakes.

Event selection	τ_{had} selection	$N_{t\bar{t}H}$	$N_{\text{real }\tau_{\text{had}}}$	$N_{\text{fake }\tau_{\text{had}}}$	N_{Data}	FF
SR	nominal	14.2 ± 0.3	70.8 ± 3.5	621 ± 17	807	0.144 ± 0.006
	side-band	25.1 ± 0.4	68.0 ± 4.3	4 674 ± 40	5 116	
	weighted side-band	3.49 ± 0.17	9.3 ± 0.8	683 ± 32	759 ± 37	
ER	nominal	2.75 ± 0.1	56.4 ± 3.6	613 ± 33	657	0.146 ± 0.007
	side-band	4.27 ± 0.17	46.8 ± 3.9	5 030 ± 120	4 134	
ER Z+jets enr.	nominal	0.28 ± 0.05	80.3 ± 5.0	1 885 ± 69	1 806	0.141 ± 0.004
	side-band	0.48 ± 0.06	43.3 ± 3.2	15 012 ± 226	12 311	
ER $t\bar{t}$ enr.	nominal	2.56 ± 0.09	62.7 ± 3.9	486 ± 16	608	0.136 ± 0.007
	side-band	3.11 ± 0.10	74.9 ± 3.9	3 636 ± 51	4 059	

Table 5.4: Event yields of $t\bar{t}H$, real τ_{had} and fake τ_{had} MC and data in signal region and its side-band, as well as the control regions used to determine the fake factor (FF). The FFs in their column are unparametrised. The weighted side-band row, however, is weighted with FF(p_T) from the ER. Quoted uncertainties are of statistical nature and omitted on unweighted event counts.

All systematic uncertainties of the fake estimate are summarised in Tab. 5.5 and are discussed in the following.

The effect of 1. is estimated by having two alternative extraction regions. Differences in estimated numbers of events are considered as uncertainties. The ER(Z enr.) estimated number of events is 4.4 % smaller than the nominal estimate, while the ER($t\bar{t}$ enr.) estimate is 8.2 % less. Both alternative ERs also have different fake compositions to the ER. Compositions of the fake τ_{had} are shown for all regions in Fig. 5.9. In particular the $t\bar{t}$ enriched ER has more b-jet fakes. Moreover, the minimum identification requirement even for the reversed τ_{had} identification ensures not only that gluon- and quark fake fractions but also that their FFs are similar in all regions. So the effect of 2. is implicitly included by varying the ER.

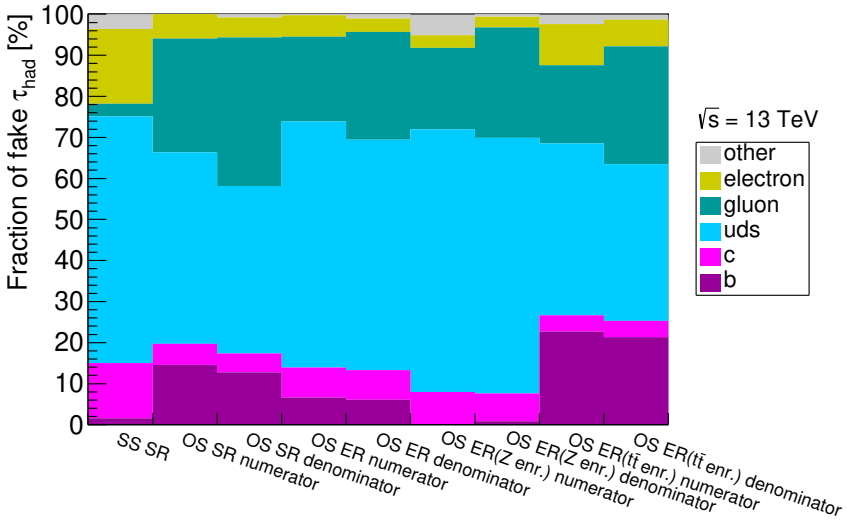


Figure 5.9: Fractional composition of the fake τ_{had} in the $2\ell(\text{SS})1\tau_{\text{had}}$ and $2\ell(\text{OS})1\tau_{\text{had}}$ channels. "numerator" and "denominator" refer to the τ_{had} definition in the FF method.

The choice of p_{T} as functional parameter of the FF ensures good modelling of the p_{T} of the τ_{had}. The binning of the FF is chosen as fine as the available statistics allows it. It was shown that the parameter does not significantly affect the prediction (see Tab. 5.3).

Changing the reverse identification requirement will first and foremost change the size of the side-band regions. Therefore it is not possible to disentangle statistical effects from real systematic effects. Both 3. and 4. are not explicitly considered as a source of systematic uncertainty.

Contributions from real τ_{had} and $t\bar{t}H$ are subtracted in all control regions. The number of events subtracted in this way are varied by ±20 % for real τ_{had} backgrounds. The number of subtracted $t\bar{t}H$ events in the SR side-band are varied coherently with the signal strength parameter of the statistical interpretation. To illustrate the size of the effect it is shown in the table with variations of factors 0 and 2. The average effects of the subtractions are smaller than the statistical uncertainty.

5.4 Suppression of $t\bar{t}$ background using BDT

A boosted decision tree (BDT) is a machine learning technique that is widely used in high energy particle physics to classify events. In this thesis it is used to enhance the separation between signal and

Uncertainty	FF estimated number of events	Δ(nominal) [%]
Nominal ER	746 ± 36	—
$t\bar{t}$ enriched ER	685 ± 35	−8.2
Z+jets enriched ER	714 ± 22	−4.4
Scale real τ_{had} subtraction by 0.8	760 ± 36	1.9
Scale real τ_{had} subtraction by 1.2	732 ± 36	−1.9
Scale $t\bar{t}H$ subtraction by 0	752 ± 36	0.8
Scale $t\bar{t}H$ subtraction by 2	740 ± 36	−0.8

Table 5.5: Effects of systematic uncertainties of the fake τ_{had} estimate on the predicted number of events from fake backgrounds. The third column is the difference to the nominal estimate in %.

background events. As signal and background serve $t\bar{t}H$ and $t\bar{t}$ MC, respectively.

A simple decision tree is a binary tree in which each node represents the splitting of events based on an observable X_1 at the threshold value t_1. The left panel of Fig. 5.10 shows a sketch of such a simple decision tree that uses two observables. The right panel shows how that tree partitions the 2D space spanned by the observables. Each X is chosen from the set of all observables that are input to the algorithm.

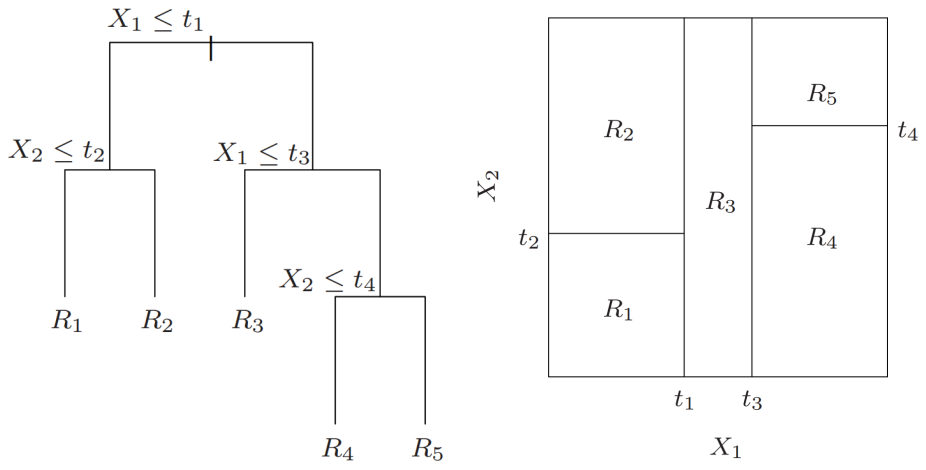


Figure 5.10: Sketch of a decision tree and the partitions that it creates in the 2D-space of 2 observables [135]. The output nodes $R_{1...4}$ are labelled signal or background by the majority of events in the node.

Both the choice of observable and threshold value is optimised. The purity p is the ratio of signal events to all events in a node. A purity of $p = 0.5$ indicates a node that is equally mixed with signal and background. Ideal nodes have purities close to 1 or 0, since nodes that are pure in background are equally valuable as nodes pure in signal to separate the two classes. A separation estimator called "Gini index" is defined as $p \cdot (1 - p)$, which is symmetric for these two cases and is used as the figure of merit to be optimised [136].

The splitting is performed until a stopping criterium is reached, either a maximum number of splits of 4 or a minimum number of events in any node of at least 5 % of the total events. The last node (leaf)

in each branch of the tree is labelled to be signal- or background-like by the majority of the contained events. The power of a BDT is that it combines many (here 100) simple decision trees. This is achieved by a process called boosting. Events that end up in the wrong leaves of the tree (i.e. signal events in a node labelled as background and vice versa) are given a larger weight (boosted) for the next iteration of tree-splitting optimisation. This optimisation of cut thresholds and the boosting is known jointly as "training" of the BDT. The final output of all trees for a given event, called "BDT output", is the sum of all tree decisions weighted by the normalised boosting weight. The background (signal) label is mapped to -1 (1) for the output. So the weighted sum is a continuous distribution between -1 and 1, where a value in between the two indicate degrees of likeness to be signal or background.

The training of a BDT has many parameters that can be tuned to optimise its behaviour. The stopping criteria for the tree splitting are influenced by limited MC statistics. More splittings can be more sensitive to differences in signal and background, while fewer splittings are less susceptible to statistical fluctuations. Therefore the criteria are chosen to have as many splits as possible with the given statistics.

The BDT can become sensitive to statistical fluctuations that are present in the particular MC events used. This is called overtraining and results in different responses of the BDT classifier to data that was seen during the training and unseen data. To evaluate whether overtraining is present one splits the training events into at least two parts deterministically, typically by even or odd event numbers. Only one part of the events is used for training. The other unseen part can then be used to make sure that the BDT responses are the same between the trained and unseen events.

Unfortunately this procedure halves the available event statistics for training and evaluation. To recover the full statistics for evaluation purpose it is possible to train two BDTs on both the even and odd samples, and evaluate each even event with the odd BDT and vice versa.

To also increase event statistics for the training one can split into more than two parts. In this thesis, all training events are split four-ways. Therefore there are four BDTs that have been trained on ¾ of all events and evaluated on the remaining ¼. The splitting is performed on the event number which is a unique integer identifier of each event. The number is divided by four and the remainder r is used to make four independent subsets. Table 5.6 shows a sketch of this procedure.

	$r = 0$	$r = 1$	$r = 2$	$r = 3$
BDT1	Training	Evaluation	Evaluation	Evaluation
BDT2	Evaluation	Training	Evaluation	Evaluation
BDT3	Evaluation	Evaluation	Training	Evaluation
BDT4	Evaluation	Evaluation	Evaluation	Training

Table 5.6: Sketch of the 4-fold cross-evaluation of the BDT. r is the remainder after dividing the event number by 4.

The parameters with the largest impact on the BDT performance are the set of input observables $\{X_1, ..., X_N\}$. The observables are chosen based on their discriminating power between signal and background. The goal for the choice of the set of observables is to include as much information about the event as possible, while minimising the number of observables. The reason to keep the number of observables low is simplicity. Every observable should be verified to have a good agreement between data and MC. Additionally it is easier to understand the effect of systematic uncertainties when there are fewer observables.

The number of possible observables in a collision event is large. In the absence of full event reconstruction there are no single, simple observables that have good discrimination, but many that have low

or moderate discrimination. The BDT can then concentrate many low performance observables into one high performance one.

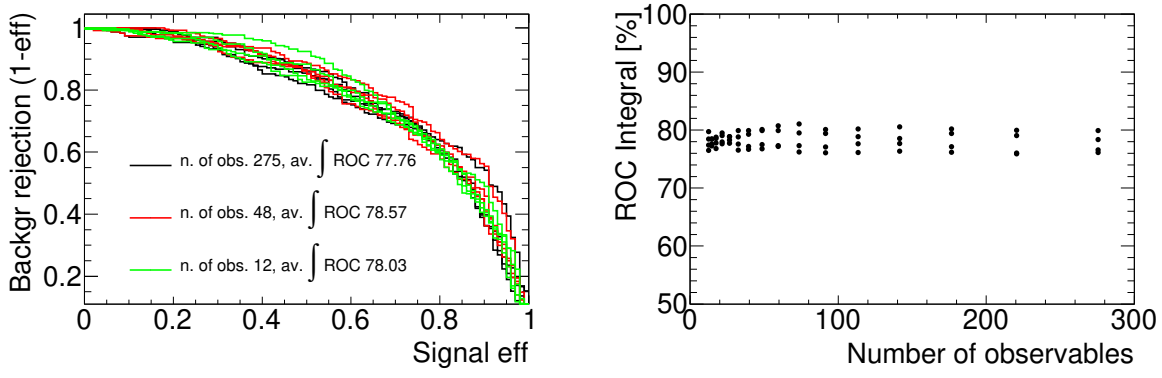
5.4.1 Optimisation of observables

The set of observables that are input to the BDT are determined algorithmically. An initial list of observables is generated by computing ΔR, minimal ΔR, $\Delta\phi$, $\Delta\eta$, k_T [1], invariant mass and scalar p_{T} sum of all combinations of the following reconstructed objects: τ_{had}, both light leptons, first three leading jets in p_{T}, as well as the last jet, first two jets that are b–tagged and $E_{\mathrm{T}}^{\mathrm{miss}}$. The p_{T} of all above objects are also added. This results in a list of 275 potential observables to serve as input to the BDT. The goal of the algorithm that chooses the final set is to select the smallest number of observables with the best performance. Since these criteria are subjective, an arbitrary choice is made to reduce the set of observable to one dozen.

The algorithm trains a BDT with the full set of observables and discards the least important 20%. The figure of merit to determine the importance of an observable is the ranking obtained by counting how often the observable was used in a tree splitting weighted by the gain in separation achieved in the child nodes [136]. The discarding is repeated until a dozen observables are identified.

A ROC curve is a graphical summary of the performance of a classifier and shows the background rejection (1− efficiency) vs. signal efficiency [137]. A better classifier therefore has a curve closer to the top right corner. Each point on the ROC curve is a specific cut on the classifier output. By moving along the line from the left to right, the classifier output cut threshold changes from rejecting all background (but not leaving any signal either) to rejecting no background (but keeping all events). In a search analysis signal events are rare so a cut threshold with high signal efficiency is usually chosen. That means that only the right side of the ROC curve is relevant for estimating the performance of a classifier.

Figure 5.11(a) shows the ROC curves for three iterations (first, last and one intermediate). In each iteration there are four curves for each of the 4–fold BDTs. When ROC curves have similar shapes, as is the case here, their integral can be used as a figure of merit to compare them. To ascertain that the final set of observables is performing well the area under the ROC curve (ROC integral) is plotted in Fig. 5.11(b) for each iteration of the optimisation. It can be seen that no performance is lost by reducing the set of observables from 275 to a dozen.



(a) ROC curve

(b) ROC integral

Figure 5.11: ROC curves for the first, last, and an intermediate iteration and ROC integrals for all iterations of the optimisation of BDT observables. For each iteration the four folds of the training are shown.

[1] $k_T = \min(p_{\mathrm{T}}^a, p_{\mathrm{T}}^b) \cdot \Delta R(a,b)$ with a maximum of 400 GeV

The optimisation results in the set of observables in Tab. 5.7. Two changes are made to the optimised set of a dozen observables. First, the transverse momentum of the leading b–tagged jet is removed due to strong correlation with leading jet p_{T}. Second, the number of jets and b–tagged jets is added. Due to their integer nature which is different to the other observables that have continuous values they are not included in the optimisation algorithm. Therefore the final number is 13 observables.

The range of momentum-based and jet-counting observables is restricted for the purpose of training the BDT. The restrictions are motivated by aspects of the training algorithm as well as considerations of data/MC agreement.

During training, the BDT algorithm scans the full range of observable values for optimal cut thresholds. With larger ranges the search space and so the distance between considered cut thresholds becomes larger. The momentum-based observables have long continuously falling distributions. The allowed values for these are restricted to essentially remove the tail. All events are still used with a maximum rather than the actual value. Capping of the allowed range prevents too coarse searches of the cut thresholds. Furthermore, little information about $t\bar{t}H$ is expected to be found at large values of p_{T}.

The allowed number of (b–tagged) jets are capped at 6 (2) for a different reason. Large multiplicities of jets are caused by higher order QCD radiation. In MC, which is used to train the BDT, these are partially approximated by parton shower algorithms that often may not model data well. This potential region of mismodelling is avoided by capping the values.

All BDT inputs and the output can be seen in Figs. 5.12 to 5.15. All distributions show good agreement of the prediction to data. It is also determined with Kolmogorov-Smirnov tests between BDT outputs obtained from seen and unseen data that no overtraining is present in this BDT.

	Observables	Abbreviation used in figures
Lepton properties	Invariant mass of light lepton pair	$m_{\ell\ell}$
	Sum p_{T} of light leptons	
	p_{T} of τ_{had}	τ_{had} p_{T}
Jet properties	Leading jet p_{T}	
	Sum p_{T} of jets	
	Sum p_{T} of b–tagged jets	
	Number of jets	
	Number of b-tagged jets	
Angular distances	Smallest ΔR distance between a light lepton and a jet	min ΔR(lj)
	Smallest ΔR distance between a light lepton and a b–tagged jet	min ΔR(lb)
	Smallest ΔR distance between a non-tagged jet and a b–tagged jet	min ΔR(jb)
	ΔR distance between the leading light lepton and the τ_{had}	$\Delta R(\mathrm{l}_0, \tau_{\mathrm{had}})$
	ΔR distance between the sub-leading light lepton and the τ_{had}	$\Delta R(\mathrm{l}_1, \tau_{\mathrm{had}})$

Table 5.7: Chosen set of observables for BDT.

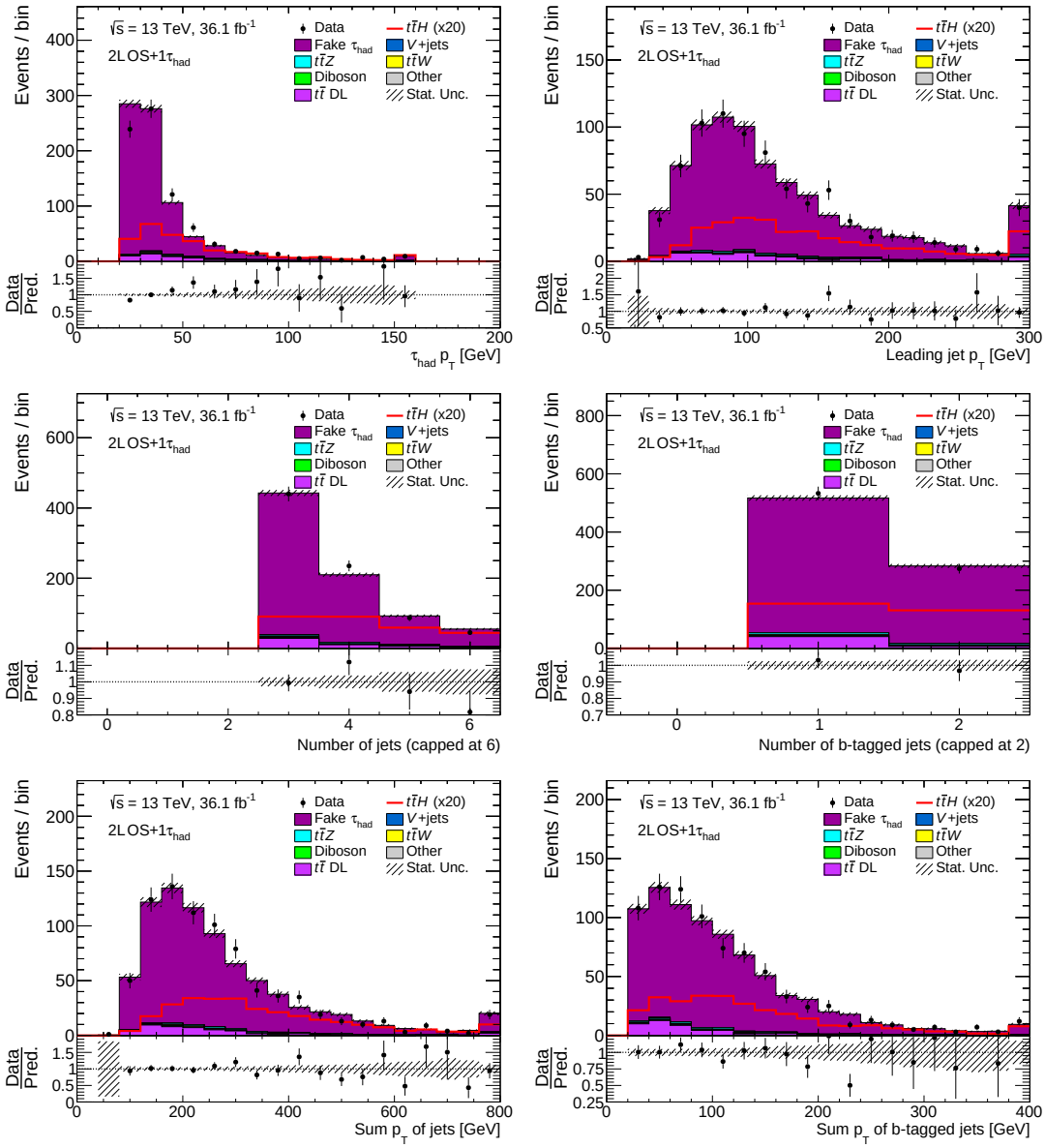


Figure 5.12: BDT input observables.
Top row: Leading tau p_{T}. Leading jet p_{T}.
Middle row: Number of jets. Number of b–tagged jets.
Bottom row: Scalar sum of jet p_{T}. Scalar sum of b–tagged jet p_{T}.

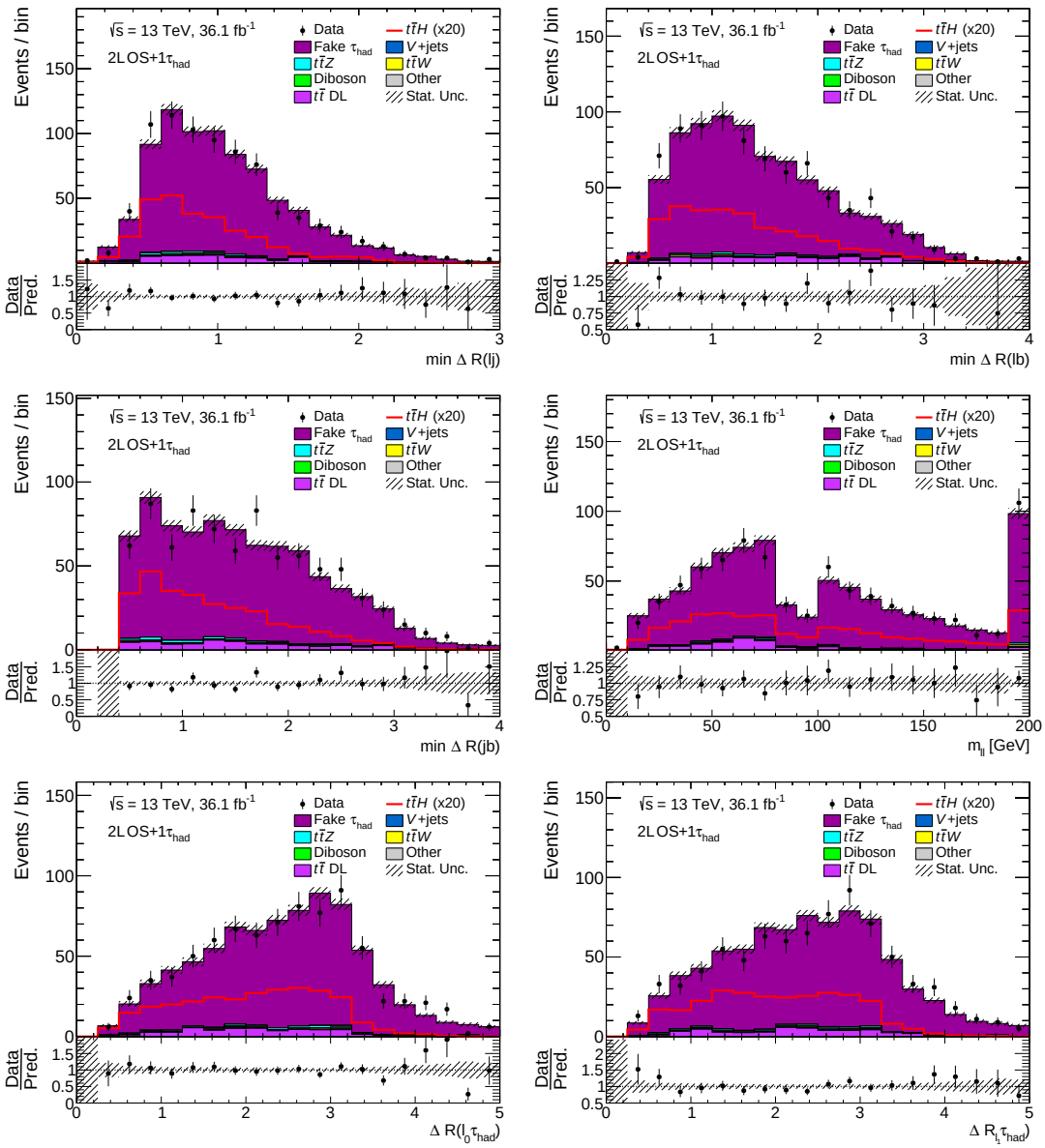


Figure 5.13: BDT input observables.
Top row: Minimum ΔR between a lepton and a jet. Minimum ΔR between a lepton and a b–tagged jet.
Middle row: Minimum ΔR between a jet and a b–tagged jet. Invariant mass of the dilepton system. Here the Z–veto is clearly visible.
Bottom row: ΔR between leading lepton and tau. ΔR between subleading lepton and tau.

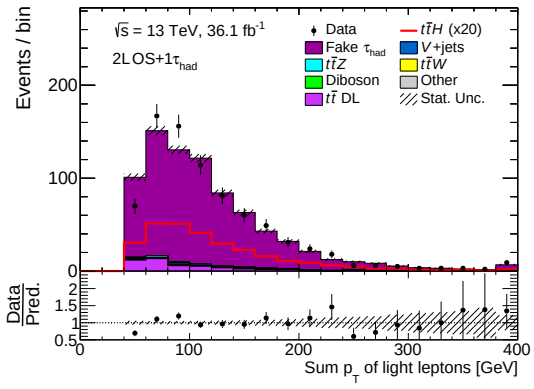


Figure 5.14: BDT input observable. Scalar sum of lepton p_{T}

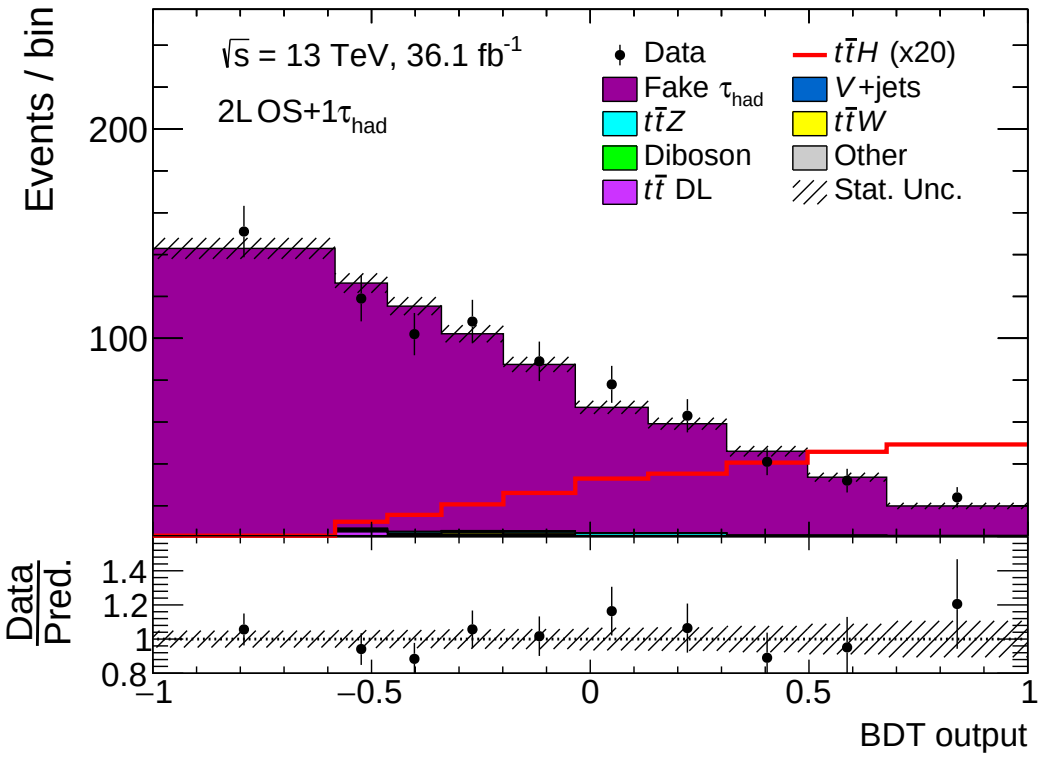


Figure 5.15: BDT output in $2\ell(\mathrm{OS})1\tau_{\mathrm{had}}$

5.5 Orthogonality with other $t\bar{t}H$ analyses

Because this channel is analysed for the first time, concerns about overlapping selections with the $t\bar{t}H(H \to b\bar{b})$ dilepton channel exist. The $H \to b\bar{b}$ selection also contains two light leptons of opposite charge. Major differences exist in the jet selections. In $t\bar{t}H$ events with a $H \to b\bar{b}$ decay, four b–jets are present and these are also tagged in the analysis. This contrasts starkly with the b-tagging requirement of at least one in the $2\ell(\mathrm{OS})1\tau_{\mathrm{had}}$ selection.

The overlap of the event selections is found to be negligible. However, to be completely orthogonal, the $H \to b\bar{b}$ analysis vetoes events with any reconstructed τ_{had}. The effect on the signal yield is less than 1 ‰.

CHAPTER 6

Analysis of $t\bar{t}H \rightarrow 2\ell(\text{SS})1\tau_{\text{had}}$ final state

In this channel, final states with two light leptons of the same charge and one τ_{had} are analysed. By inverting the charge cut on the light leptons with respect to the channel described in the previous chapter the backgrounds and requirements on electron and muon identification are fundamentally different. The main advantage of this channel is that due to smaller expected background, this channel has a high signal purity of approximately 0.4 (see Fig. 4.11 and Fig. 6.1). However, the total number of expected events (about 10) is smaller.

6.1 Object selection

The selection criteria for electrons, muons and τ_{had} are identical to the ones used in $2\ell(\text{OS})1\tau_{\text{had}}$, except that identification and isolation requirements on electrons and muons are additionally tightened to reduce contributions from processes with non-prompt leptons. Rejection of non-prompt electrons and muons is achieved by using the dedicated discriminant that was described in Sec. 3.4.10.

For electrons, an additional multivariate classifier is used to reject candidates that were reconstructed with the wrong charge. It has an efficiency of 95 % for correct charge assignment and a rejection factor of approximately 17 for wrong charge assignment.

6.2 Event selection

The event selection is similar to $2\ell(\text{OS})1\tau_{\text{had}}$. The differences are that the light lepton charges have the same sign and the number of jets is at least four. The Z veto is kept only for events with two electrons. $Z \rightarrow ee$ events can be selected when an electron charge is misreconstructed. These cuts constitute the signal region and are summarised in Tab. 6.1.

6.3 Backgrounds

The reducible backgrounds in this channel are non-prompt electrons and muons mainly from $t\bar{t}$ events with one prompt W to lepton decay. The processes $t\bar{t}Z$, $t\bar{t}W$ and diboson contribute with 2 prompt leptons and can have a real or fake τ_{had}. The background composition is shown in Fig. 6.1.

Other backgrounds are processes that are rare in the selected phase space and include tZ, tW, tWZ, $t\bar{t}WW$, $t\bar{t}t$, tH, radiative $t \rightarrow Wb\ell\ell$ decays in $t\bar{t}$ and triboson production.

Light leptons	τ_{had}	Jets
=2	=1	**≥ 4**
same charge		≥ 1 *b*-tagged
pass PLI requirement		
$\mathbf{\lvert m_{ee} - m_Z \rvert > 10\,GeV}$		

Table 6.1: Selection criteria in the $2\ell(SS)1\tau_{had}$ channel. Differences to the selection in $2\ell(OS)1\tau_{had}$ are in bold face. "PLI" refers to the prompt lepton tagger which is called PromptLeptonIsolation in ATLAS jargon.

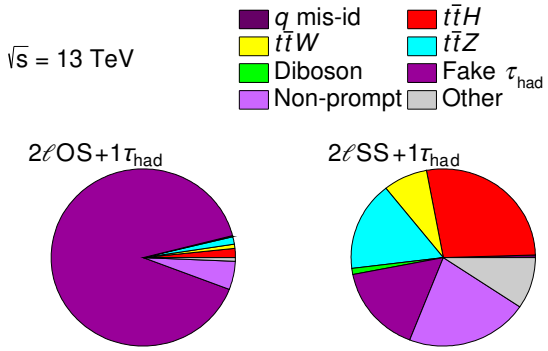


Figure 6.1: Fractional contributions from all predicted processes in $2\ell(OS)1\tau_{had}$ and $2\ell(SS)1\tau_{had}$ channels.

The background from misassigning the reconstructed electron charges is estimated as described in Sec. 4.2.1 which yields 0.05 ± 0.02 events. As the method inverts the charge cut on the light lepton pair to apply reweighting factors, there is overlap with events of the $2\ell(OS)1\tau_{had}$ channel. However it is small due to the tighter light lepton and event selections. The impact of the overlap is ignored as the estimated contribution is also small in the present channel.

6.4 Non-prompt electron and muon estimate

The background with non-prompt leptons is estimated analogously to the fake τ_{had} estimate in the $2\ell(OS)1\tau_{had}$ channel with a fake factor method. The fake factors are estimated separately for electrons and muons as functions of p_T and are shown in Fig. 6.2. The regions of side-band light leptons are defined by reversing identification as well as isolation cuts of one of the light leptons in the event. In most cases this is the subleading one. The extraction regions are defined by modifying the SR to use events with only two or three jets and also allowing events with no τ_{had}. The uncertainty on the estimate is limited by the number of events in the SR side-band and the non-closure found in the MC validity test. Its total size is 55 % [5].

6.5 Fake τ_{had} estimate

Events with τ_{had} faked by jets are also a main contribution to the background in this channel. If the fake τ_{had} is accompanied by a non-prompt lepton its contribution to the background is included in the non-prompt light lepton estimate, since the light lepton SR side-band also contains the τ_{had} selection.

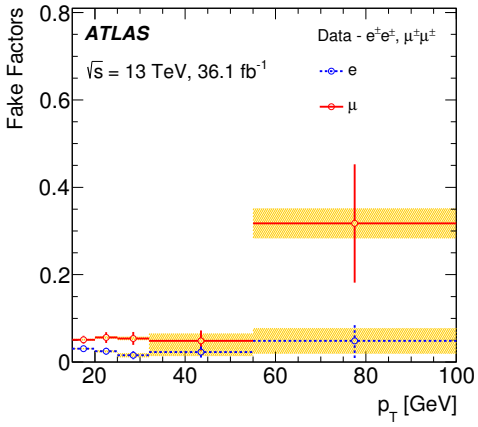


Figure 6.2: Fake factors for electrons and muons as function of p_{T} determined in $2\ell(\mathrm{SS})1\tau_{\mathrm{had}}$ ERs. Error bars are statistical and the shaded area indicates the systematic uncertainty [5].

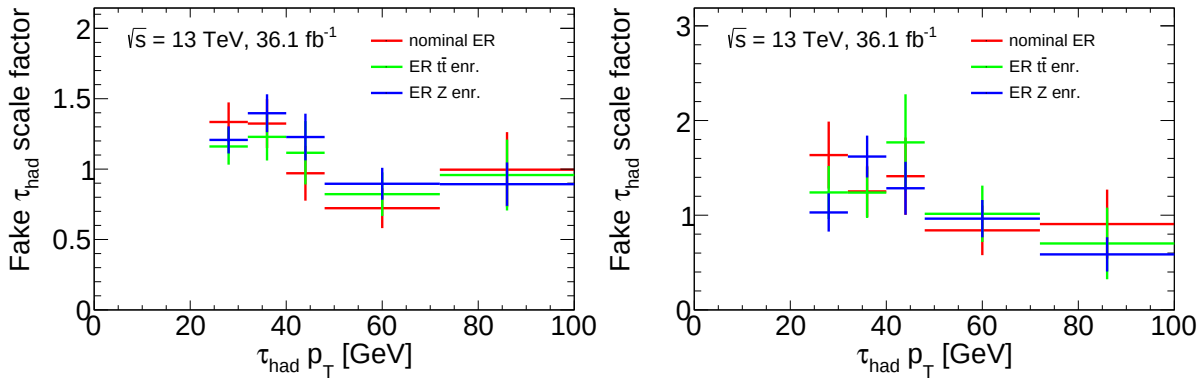
What is not estimated, are events with two prompt SS light leptons and a fake τ_{had}. Such events are selected in processes like $t\bar{t}W$ and $t\bar{t}Z$.

The same arguments as for $2\ell(\mathrm{OS})1\tau_{\mathrm{had}}$ hold for advantages of estimating this background by using a data-driven method. Unfortunately, the fake factor method as described previously cannot be applied here. The fundamental assumption that the side-band regions are (nearly) signal free is not satisfied. In the analysis there is another channel with the final state of two same sign light leptons but vetoing any additional τ_{had}. By reverting the identification of the τ_{had} there is a large overlap with that signal region. The MC subtraction as performed in $2\ell(\mathrm{OS})1\tau_{\mathrm{had}}$ is thus problematic here due to the much larger signal contamination in the side-band. The introduced correlation between the expected signal and the estimated background would be large. Furthermore, there would be events selected in both the $2\ell(\mathrm{SS})0\tau_{\mathrm{had}}$ SR and $2\ell(\mathrm{SS})1\tau_{\mathrm{had}}$ CR, which introduces statistical correlation that would have to be taken into account in the statistical interpretation. For simplicity, these problems are avoided by not using a fake factor method.

Instead, background with a fake τ_{had} is estimated by MC. The knowledge gained from performing the FF method in $2\ell(\mathrm{OS})1\tau_{\mathrm{had}}$ is used by scaling the MC. A scale factor is calculated as the ratio of fake τ_{had} background estimated with the FF method to the background estimated by MC in the $2\ell(\mathrm{OS})1\tau_{\mathrm{had}}$ channel. It is assumed to be applicable here because all involved MC in both channels use the same parton shower and hadronisation algorithm that create the reconstructed jets. So the frequency of jet-to-τ_{had} fakes should be the same independent of the rest of the event topology. The scale factor is 1.20 ± 0.07 and the relative differences to the SFs obtained from the varied extraction regions are defined as uncertainties (see Tab. 6.2). The scale factor is consistent with what was observed qualitatively in the previous chapter. A correlation between the SF and τ_{had} p_{T} is not observed within uncertainties as shown in Fig. 6.3.

ER used in FF method	Scale factor
Nominal ER	1.20 ± 0.07
$t\bar{t}$ enriched ER	1.10 ± 0.06
Z enriched ER	1.15 ± 0.05

Table 6.2: Scale factors for fake τ_{had} backgrounds. Defined as the ratio of fake τ_{had} background estimate with the FF method to the truth-matched (to fake τ_{had}) MC estimate in the $2\ell(\text{OS})1\tau_{\text{had}}$ channel.



(a) 1-prong (b) 3-prong

Figure 6.3: Ratio of FF estimate to MC estimate of fake τ_{had} background for (a) 1-prong and (b) 3-prong τ_{had} in the $2\ell(\text{OS})1\tau_{\text{had}}$ SR. The ratio is used as a scale factor.

CHAPTER 7

Statistical interpretation

The statistical interpretation of the results consists of two related parts: estimation of free parameters and explicit hypothesis testing. The parameter of interest is the signal strength which is a scaling factor for the expected number of signal events and is defined as the ratio of observed to predicted cross section as $\mu = \frac{\sigma_{\text{observed}}}{\sigma_{\text{predicted}}}$. It is allowed to be fitted to negative values but will be restricted to physically meaningful values for the hypothesis tests. All other parameters that are estimated are called nuisance parameters (NP).

The hypothesis to be tested implicitly is that the free parameters describe the data well. Explicitly this is interpreted as whether or not the data is compatible with the presence and absence of a $t\bar{t}H$ signal.

7.1 Estimating the $t\bar{t}H$ signal strength

The signal strength μ of $t\bar{t}H$ is estimated with a binned maximum likelihood (ML) fit. The likelihood function is defined as

$$L(\mu) = \prod_{i}^{\text{bins}} P(n_i|\mu \cdot s_i + b_i), \tag{7.1}$$

where $P(k|\nu)$ is the probability of observing k events when ν events are expected, n is the observed number of events in data, s and b are the expected number of signal and background events. The product runs over all bins in the analysis.

Systematic uncertainties are taken into account by introducing NPs $\vec{\theta}$ into the likelihood function as follows

$$L\left(\mu, \vec{\theta}\right) = \prod_{i}^{\text{bins}} P\left(n_i|\mu \cdot s_i(\vec{\theta}) + b_i(\vec{\theta})\right) \cdot \prod_{j}^{\text{systematics}} A(\theta_j|\tilde{\theta}_j). \tag{7.2}$$

The vector notation of θ indicates that there are multiple NPs. The effect of the NPs is that they can change the expected number of events and are additional parameters that need to estimated. However, they are only partially free parameters because $A(\theta|\tilde{\theta})$ is a gaussian constraint term that penalises the likelihood function when the NP θ is different from its nominal value $\tilde{\theta}$. The second product runs over all systematic uncertainties.

The nominal values and variances of the NPs are obtained from auxiliary measurements which are described in Sec. 7.3.

The estimated values, $\hat{\mu}$ and $\hat{\vec{\theta}}$, which are called maximum likelihood estimators or best-fit values, are obtained by minimising the negative logarithm of the likelihood function. Minimising the negative

(logarithm of) the likelihood is a technical choice[1] and equivalent to maximising the likelihood. The variance of the estimators are determined from the covariance matrix which is evaluated numerically as the inverse of the second derivatives of the log likelihood function using the observed data and the best-fit NP values $\hat{\vec{\theta}}$ [139, 140].

The technical implementation of the likelihood function and the probability distribution functions contained therein is done with the HistFactory [141] and RooFit [142] software packages.

7.2 Testing for discovery of $t\bar{t}H$

To test whether the $t\bar{t}H$ process is discovered, the null hypothesis H_0 is that the observed data is compatible with the expectation when only considering background processes (i.e. $\mu = 0$). The alternative hypothesis is that $\mu > 0$. The following test statistic is defined, where H_0 is represented by the numerator,

$$q_0 = \begin{cases} -2 \ln \dfrac{L(0, \hat{\hat{\vec{\theta}}})}{L(\hat{\mu}, \hat{\vec{\theta}})} & \text{if } \hat{\mu} \geq 0 \\ 0 & \text{if } \hat{\mu} < 0 \end{cases}, \tag{7.3}$$

where $\hat{\hat{\vec{\theta}}}$ are the ML estimators of the NPs under the condition that $\mu = 0$, i.e. no signal. Smaller values of q_0 imply a better compatibility with H_0. A downward fluctuation of data yielding $\hat{\mu} < 0$ is also incompatible with H_0 but is not physically meaningful. Therefore the statistic is defined to be zero in those cases.

To determine the p-value of the observed q_0, knowledge about the sampling distribution of the test statistic is required. The sampling distribution can be found by performing toy experiments in which for each experiment a pseudo-dataset is generated from the expectation. The toys and the resulting q-distribution $f(q_\mu|\mu')$ thus depend on the assumed true value of μ' as well as on the value of μ that is being tested. The sampling distributions are known asymptotically [143], so that the computationally expensive process of generating toys is not necessary. The p-value is then the integral from q_0 to the right of the sampling distribution.

A transformation of the p-value to a significance Z is made such that the one-sided integral of a standard[2] normal distribution from Z standard deviation to the right is equal to p. In particle physics it is customary to require a significance of 5σ or $p < 2.87 \cdot 10^{-7}$ to claim the discovery of a new particle. This is an expression of strong confidence in the Standard Model. Therefore extraordinary proof is required to reject a hypothesis that is derived from the SM to avoid falsely rejecting it. Even in cases where both the null and alternative hypotheses are derived from the SM, as is the case here in the search for $t\bar{t}H$, the conservative requirement of 5σ is maintained to prevent false or premature claims of discovery.

The expected significance of the analysis is obtained from the median of the q_0-distribution assuming a true value of $\mu' = 1$. In other words, the median is the significance of the incompatibility of the background-only hypothesis with hypothetical data if data contained the signal exactly as expected.

[1] The computer program MINUIT minimizes functions [138].
[2] zero mean, unit variance

7.3 Systematic uncertainties

Systematic uncertainties enter the ML fit as parameters of which there exists prior knowledge. These so-called constraints are obtained in auxiliary measurements e.g., jet resolution is measured with jet events, etc. Many of these measurements use the same ML fit formalism to estimate parameters. In principle it is possible to use the likelihood functions of all auxiliary measurements in a global analysis of all ATLAS data. However, this would be prohibitively expensive in terms of computation resources. Instead, the simplified formalism of nuisance parameters is used. All uncertainties of the auxiliary measurements are expressed as uncorrelated components represented by the NPs. Each NP has a nominal value and $\pm 1\sigma$ uncertainty values (up/down). Values of the NPs in between these three fixed values are interpolated and values beyond the up/down variations are extrapolated.

Sources of uncertainties that are external to ATLAS, like theoretically predicted cross sections, are also expressed using NPs for consistency.

Nuisance parameters can be grouped by their properties and origin into instrumental, theoretical, related to background estimation and simulation statistics.

Many of the following uncertainties are separated into independent uncorrelated components which correspond to the nuisance parameters in the ML fit. Table 7.1 shows all components and also indicates whether they affect all bins in the fit in a uniform manner (overall normalisation) or affect bins differentially (shape).

7.3.1 Instrumental uncertainties

Instrumental uncertainties are related to the instrument that is used to record the data i.e. the ATLAS detector. The performance of the detector is measured for all reconstructed and measured particles and observables. Differences of e.g. energy resolution or reconstruction efficiencies between MC and data are corrected in MC so that it matches data. The measurements are described in the following.

Electrons, muons and hadronically decaying tau leptons

Efficiencies to reconstruct, identify and select electrons [101], muons [102] and τ_{had} [110] are measured in data with Z boson decays to the respective charged leptons. The measurements use a tag-and-probe method which tags one leg of the Z decay and measures the other leg. The "tag" places very strict requirements on the reconstructed object. Together with the knowledge about the invariant mass of Z, a probe object can be selected without using the cuts for which the efficiency has to be measured. For electrons and muons, measurement with J/Ψ events complement the low p_{T} phase space. For the τ_{had} measurements the tag is actually the muon from a leptonic tau-lepton decay, thus using $Z \to \tau_{\mu}\tau_{\text{had}}$ events.

Energies and momenta measured by ATLAS were calibrated with test beam data and MC simulations. Any residual differences between data and MC are parametrised with a "scale" and a "resolution" parameter. The parameters are fitted in $Z \to \ell\ell$ data to the invariant mass distribution. The parameter names indicate the effect that they have on the distribution i.e., "scale" shifts the peak and "resolution" affects the width. The τ_{had} calibration only fits a "scale" as the hadronic resolution is too large to be sensitive to the Z width.

Jets

The jet energy calibration consists of several steps that correct for and reduce the impact of pileup effects. The non-compensating nature of the calorimeter is also corrected for. Any residual differences between

data and MC are adjusted by in-situ measurements [105]. The measurements rely on balancing the p_{T} of jets against other well measured objects. The references used are photons and Z bosons reconstructed by electrons an muons. Well measured central jets are also used to correct more forward jets.

The efficiency of the pileup rejection discriminant is measured in $Z \to \mu\mu$ + jet events where the reconstructed Z is tagging the recoiling jet that is probed [106].

The efficiencies to b–tag b–jets, c–jets and light jets are measured in data [107]. No measurement is available for the b–tagging efficiency of tau jets. Therefore the same efficiency corrections are applied to tau jets as to c–jets. The relevant uncertainty components are increased due to this assumption. This c-to-τ extrapolation affects this analysis because the τ_{had} definition is implicitly using b–tagging for vetoing.

Luminosity

The luminosity of the collisions produced by the crossing of LHC beams is measured by several luminometers in ATLAS [144]. The measurements are calibrated regularly by absolute luminosity measurements with van-der-Meer scans which displace the proton beams with respect to each other [145]. By changing the displacement the beam profile can be measured by the changing interaction rates.

The pileup reweighting is correlated to the luminosity measurement because the average number of interactions per bunch crossing are calculated from it.

Technically, there are two kinds of instrumental uncertainties. One type is related to momentum/energy calibration. Those are applied by modifying the nominal p_{T}[3] of objects. The effect is modified shapes of distributions of physical observables. However, there are also acceptance effects when the p_{T} migrates across an object or event selection requirement.

The other kind is related to efficiency corrections. These are applied as event weights that are multiplied to the original event weight. The effect can also be shape changes. However, there are no acceptance effects. The number of selected MC events is the same for all variations.

7.3.2 Uncertainties of data-driven background estimates

The uncertainties of the fake τ_{had} estimate are described in Sec. 5.3.2. In Tab. 7.1, "side-band statistics" refers to the signal region side-bands to which the fake factor weights are applied. Therefore they have as many components as there are bins in the fitted observables. "Fake factor statistical" refers to the statistical uncertainties of the fake factors that come from the number of events in the extraction regions and their side-bands. The number of components is equal to the number of bins which parametrise the fake factors.

The contamination of $t\bar{t}H$ in the SR side-bands are subtracted by modifying the Poisson term in the likelihood function in Eq. 7.2 to

$$P\Big(n_i|\mu \cdot s_i(\vec{\theta}) + b_i(\vec{\theta}) \underbrace{-\mu \cdot s_i^{\tau_{\mathrm{had}}\ \text{side-band}}(\vec{\theta})}_{\text{additional term}}\Big), \tag{7.4}$$

such that the subtracted events are completely anticorrelated to the signal events.

Uncertainties related to the non-prompt lepton estimate in $2\ell(\mathrm{SS})1\tau_{\mathrm{had}}$ behave analogously to the fake τ_{had} ones in the $2\ell(\mathrm{OS})1\tau_{\mathrm{had}}$ channel.

[3] or energy

7.3.3 Theoretical uncertainties

Theoretical uncertainties include both, effects on cross section calculations and MC event generation. All cross sections have uncertainties related to the QCD scales and PDFs/α_s used in the calculations. The cross sections and uncertainties are given in Appendix C. These uncertainties are evaluated by varying the factorisation and renormalisation scales used in the computation of the cross sections.

Uncertainties related to the event generation of $t\bar{t}H$, $t\bar{t}W$ and $t\bar{t}Z$ are estimated by comparing the nominal MC samples to the alternatives listed in Tab. 3.1. The A14 tune of the Pythia 8 parton shower provides an estimate of its uncertainty. For $t\bar{t}H$, the alternative sample uses the same matrix element generator but a different parton shower, Herwig++. For both $t\bar{t}V$, the alternative samples have different matrix element generators, Sherpa, with its built-in parton shower.

7.3.4 Pruning and smoothing

Nuisance parameters that have negligible impact on the result are removed. This is called "pruning". The criterion for removal is whether an NP changes the expected events in any sample by less than 1 %. It was determined that neither the uncertainty on the fitted signal strength nor the expected significance changes by this pruning. The number of NPs left after pruning is approximately 120 in each channel. When combining individual channels the total number of NPs can grow again to the pre-pruned number. However, the gain in fitting speed is largely maintained as the NPs do not affect all channels anymore.

Another procedure that increases the robustness of the fit is smoothing of shape uncertainties. Particularly, momentum smearing uncertainties can cause MC event migrations between bins which can create statistical fluctuations in a distribution. The smoothing reduces the fluctuations. It can also happen that a smoothed shape variation is in fact a constant offset and thus becomes a normalisation-only uncertainty (in the notation used in Tab. 7.1). It was also checked that smoothing does not significantly change the results.

A desirable effect of both pruning and smoothing is that the fitting time is reduced. This is particularly relevant for the combination with other $t\bar{t}H$ analyses where multiple likelihood functions are multiplied. While all analyses have many common uncertainties (mainly instrumental), each analysis also brings unique nuisance parameters to the combination. Therefore the combination always has a larger number of nuisance parameters and fitting time is longer. It is best to start with the minimal set of uncertainties that are relevant.

Uncertainties that do not provide both up and down variations, but only a nominal value and one variation are called 2-point systematics. These are transformed into NPs by symmetrising the variation around the nominal value. Examples of 2-point systematics are the variations of ER extraction region in the 2ℓ(OS)$1\tau_{\text{had}}$ fake τ_{had} estimate (see 1st row of Fig. D.2).

7.3.5 Pulls and constraints

For the nuisance parameters that are constrained by auxiliary measurements or other knowledge as described above, it is possible to compare the fitted values and their uncertainties to the nominal values i.e., what was "input" to the fit. A pull is the difference between the fitted and nominal values of an NP, while the difference in the uncertainties of the fitted and nominal values is called constraint (see e.g. the black points in Figs. 8.2 and 9.2). A priori, the presence of pulls or constraints is not problematic, but it can be indicative of an inadequate fit model which does not describe the data well. For example, a constraint of an instrumental NP like jet energy scale (JES) would indicate that this analysis can measure the JES better than the dedicated auxiliary analyses and might be unexpected.

Source of uncertainty	Type	Components
Instrumental		
Luminosity	N	1
Pileup reweighting	SN	1
Electron	SN	6
Muon	SN	15
τ_{had}	SN	10
Jet energy	SN	28
Jet vertex tagging	SN	1
Jet flavour tagging	SN	126
Data-driven background modelling		
Light lepton side-band statistics	N	1
Light lepton fake factor statistical	N	10
τ_{had} side-band statistics	SN	10
τ_{had} fake factor statistical	SN	5
τ_{had} fake factor ER variation	SN	2
τ_{had} fake factor real subtraction	SN	2
$t\bar{t}H$ modelling		
Cross section	N	2
Renormalisation and factorisation scales	S	3
Parton shower and hadronisation model	SN	1
Higgs boson branching fractions	N	4
Shower tune	SN	1
$t\bar{t}W$ modelling		
Cross section	N	2
Renormalisation and factorisation scales	S	3
Matrix-element MC event generator	SN	1
Shower tune	SN	1
$t\bar{t}Z$ modelling		
Cross section	N	2
Renormalisation and factorisation scales	S	3
Matrix-element MC event generator	SN	1
Shower tune	SN	1
Other background modelling		
Cross section	N	15
Shower tune	SN	1
Total		259

Table 7.1: Systematic uncertainties in the ML fit as nuisance parameters. “N” indicates that the uncertainty acts like a normalisation-only effect. “S” means that the uncertainty changes the shape of the fitted observables. “SN” means the uncertainty changes both the normalisation and shape.

7.4 Blinding

All ATLAS analyses are performed "blind" to reduce analysers' bias. There can be different criteria for blinding an analysis. In ATLAS it is often possible to remove the data points in observable distributions where the signal to background ratio expected from MC is larger than a threshold. This strategy was applied to the ttHML analysis with a threshold of 15 %.

In practice a blind analysis means that the optimisation of selections and design of background estimates are performed in a blind way [146]. Only once the analysers are confident in all aspects of the analysis, the data is unblinded. In ATLAS there is a customary procedure, where a blinded analysis is presented to colleagues that where not involved in its design. The colleagues then give the approval to unblind if they are equally satisfied with the design of the analysis. Afterwards no changes are permitted to be made.

This procedure ensures that no analyser is unconsciously biasing the result. Any unexpected outcome is only observable after unblinding. The result is thus protected from any expectations to confirm or reject the hypothesis.

Another aspect of a blind analysis is the validation of the fit model. For the purpose of a blind fit an Asimov dataset is used [143]. It is a pseudo dataset that is generated from the MC prediction. The Asimov data is exactly equal to the prediction. When used to perform the ML fit, there should be a null result where all parameters are fitted to be exactly their prior values. The potential non-integer nature of the Asimov dataset is not a problem because the factorial term of the Poisson PDF becomes a constant offset in the logarithm of the likelihood. The constant term does not affect the minimisation and cancels in the ratio of the hypothesis test, and is therefore dropped.

CHAPTER 8

Results of dilepton+τ_{had} channels

This chapter presents the results obtained from the two previously described dilepton+τ_{had} channels, $2\ell(\text{SS})1\tau_{\text{had}}$ and $2\ell(\text{OS})1\tau_{\text{had}}$.

The discriminant observables that are used to build the templates for the maximum likelihood fit are the BDT output with 10 bins in the $2\ell(\text{OS})1\tau_{\text{had}}$ channel and the event yield (1 bin) in the $2\ell(\text{SS})1\tau_{\text{had}}$ channel. Fake τ_{had} and non-prompt light lepton backgrounds are estimated with data-driven methods in OS and SS, respectively. All other backgrounds are estimated from MC and normalised to predicted cross sections and the luminosity of the pp collisions. The truth-matched fake τ_{had} background in SS is scaled with a factor derived from the data-driven estimate of OS.

The results of the maximum likelihood fit are shown in Tab. 8.1. The observed best-fit signal strength of $t\bar{t}H$ is

$$\hat{\mu} = 2.5^{+1.4}_{-1.1} = 2.5^{+1.0}_{-1.0}(\text{stat.})^{+0.9}_{-0.6}(\text{syst.}). \tag{8.1}$$

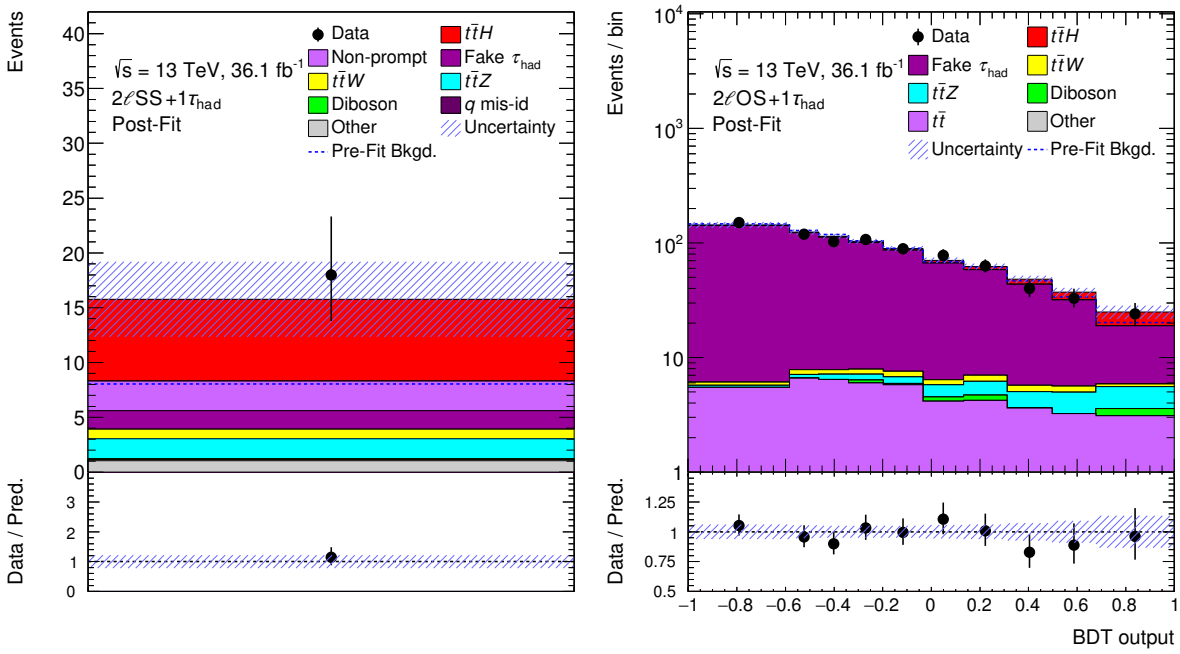
Uncertainties on the signal strengths are split into their statistical and systematic components. This is done fixing all nuisance parameters to their best-fit values which gives the statistical $\Delta\mu$. Subtracting quadratically the $\Delta\mu(\text{stat.})$ from the $\Delta\mu$ obtained in the full fit gives the systematic contribution. The fitted signal strengths of each individual channel are obtained in a combined fit with one signal strength per channel. It means that all nuisance parameters are correlated.

Figure 8.1 shows the fitted observables with the observed data events and the best-fit predicted background and $t\bar{t}H$ signal.

8.1 Significance of excess

The upward deviation of $\mu_{t\bar{t}H}$ from the expected value of 1 is due to a larger than expected number of events in the $2\ell(\text{SS})1\tau_{\text{had}}$ channel of 18 events. The significance of the excess with respect to the background-only hypothesis is 2.3σ. Since the best-fit value is much larger than the expectation, the significance of the excess with respect to the signal-plus-background hypothesis is also computed. This is done by changing the conditional maximum likelihood in the numerator of the test statistic Eq. 7.3 to maximise under the $\mu = 1$ assumption. This significance i.e., that the excess is incompatible with SM $t\bar{t}H$ production, is 1.7σ.

All event yields of both channels are shown in Tab. 8.2 before the fit and with the best-fit values. The reduction of the uncertainty on the fake τ_{had} background in the $2\ell(\text{OS})1\tau_{\text{had}}$ channel from ±80.6 pre-fit to ±28.1 is not a constraint. It is rather due to the fact that the uncertainties in the pre-fit columns are not considering correlations while the post-fit uncertainties do. The two fake τ_{had} ER variation nuisance



(a) (b)

Figure 8.1: Post-Fit plots of the discriminants used in the combined fit of (a) $2\ell(\mathrm{SS})1\tau_{\mathrm{had}}$ and (b) $2\ell(\mathrm{OS})1\tau_{\mathrm{had}}$. The signal is scaled to the observed best-fit signal strength. The expected (pre-fit) background is overlayed as a blue dashed line and is almost identical to the post-fit background. "Non-prompt" and "Fake τ_{had}" refer to the data-driven estimates of these backgrounds.

Channel	Signal strength $\hat{\mu}$		Significance in σ		
			$\mu = 0$		$\mu = 1$
	Expected	Observed	Expected	Observed	Observed
$2\ell(\mathrm{OS})1\tau_{\mathrm{had}}$	$1.0\,^{+1.6}_{-1.5}(\mathrm{stat.})\,^{+1.5}_{-1.3}(\mathrm{syst.})$	$0.5\,^{+1.6}_{-1.5}(\mathrm{stat.})\,^{+1.5}_{-1.1}(\mathrm{syst.})$	0.5	0.3	—
$2\ell(\mathrm{SS})1\tau_{\mathrm{had}}$	$1.0\,^{+1.2}_{-1.0}(\mathrm{stat.})\,^{+0.7}_{-0.5}(\mathrm{syst.})$	$3.3\,^{+1.5}_{-1.3}(\mathrm{stat.})\,^{+1.0}_{-0.7}(\mathrm{syst.})$	0.8	2.5	1.7
Combination	$1.0\,^{+1.0}_{-0.8}(\mathrm{stat.})\,^{+0.7}_{-0.5}(\mathrm{syst.})$	$2.5\,^{+1.0}_{-1.0}(\mathrm{stat.})\,^{+0.9}_{-0.6}(\mathrm{syst.})$	1.0	2.3	1.3

Table 8.1: Expected and observed best-fit signal strengths and significances with respect to the background-only ($\mu = 0$) hypothesis. Significances of excesses over the signal-plus-background ($\mu = 1$) hypothesis are also given. In the latter case the expected column is omitted as it is equal to zero.

parameters are highly anticorrelated, as not only are the variations of the total expected background events in the same direction (see Tab. 5.5), but also the shape effect on the BDT output are similar (see Fig. D.2). The anticorrelation is present both pre- and post-fit and can be seen in Appendix E.

	$2\ell(\mathrm{SS})1\tau_{\mathrm{had}}$	$2\ell(\mathrm{OS})1\tau_{\mathrm{had}}$	$2\ell(\mathrm{SS})1\tau_{\mathrm{had}}$	$2\ell(\mathrm{OS})1\tau_{\mathrm{had}}$
	Pre-fit yields		Post-fit yields	
Fake τ_{had}	1.66 ± 0.39	759.7 ± 80.6	1.69 ± 0.37	732.7 ± 28.1
$t\bar{t}$ in FF side-band with real τ_{had}	—	-6.8 ± 1.5	—	-6.8 ± 1.5
$t\bar{t}W$ in FF side-band with real τ_{had}	—	-0.3 ± 0.1	—	-0.3 ± 0.1
$t\bar{t}Z$ in FF side-band with real τ_{had}	—	-0.5 ± 0.1	—	-0.5 ± 0.1
$t\bar{t}H$ in FF side-band	—	-3.3 ± 0.3	—	-7.7 ± 4.2
Non-prompt	2.43 ± 1.40	—	2.71 ± 1.52	—
q mis-id	0.05 ± 0.02	—	0.05 ± 0.02	—
$t\bar{t}$	—	44.2 ± 7.5	—	43.0 ± 7.2
$t\bar{t}Z$	1.83 ± 0.92	11.4 ± 1.9	1.82 ± 0.90	11.0 ± 1.8
$t\bar{t}W$	0.88 ± 0.24	6.5 ± 1.3	0.87 ± 0.23	6.3 ± 1.3
Diboson	0.12 ± 0.18	2.0 ± 1.3	0.13 ± 0.18	1.9 ± 1.2
Other	1.06 ± 0.24	5.8 ± 1.5	1.05 ± 0.23	5.7 ± 1.4
Total background	8.03 ± 1.76	818.7 ± 81.0	8.31 ± 1.84	785.3 ± 29.5
$t\bar{t}H\,(H \to WW)$	0.70 ± 0.11	3.0 ± 0.6	1.72 ± 0.82	7.3 ± 3.6
$t\bar{t}H\,(H \to \tau\tau)$	1.79 ± 0.27	7.9 ± 1.3	4.38 ± 2.08	18.6 ± 9.3
$t\bar{t}H\,(H \to ZZ)$	0.02 ± 0.02	0.2 ± 0.0	0.04 ± 0.05	0.9 ± 0.2
$t\bar{t}H\,(H \to \mathrm{Others})$	0.01 ± 0.01	0.2 ± 0.1	0.01 ± 0.02	0.4 ± 0.3
$t\bar{t}H\,(H \to WW)$ fake τ_{had}	0.45 ± 0.15	2.0 ± 0.4	1.15 ± 0.55	4.9 ± 2.3
$t\bar{t}H\,(H \to \tau\tau)$ fake τ_{had}	0.03 ± 0.05	0.3 ± 0.1	0.08 ± 0.13	0.8 ± 0.3
$t\bar{t}H\,(H \to ZZ)$ fake τ_{had}	0.01 ± 0.01	<0.01	0.03 ± 0.04	<0.01
$t\bar{t}H\,(H \to \mathrm{Others})$ fake τ_{had}	0.01 ± 0.01	0.2 ± 0.1	0.03 ± 0.03	0.5 ± 0.3
Total $t\bar{t}H$	3.03 ± 0.34	13.8 ± 1.5	7.45 ± 2.31	33.5 ± 10.3
Total	11.06 ± 2.00	832.5 ± 82.1	15.76 ± 3.43	818.8 ± 30.4
Data	18	807	18	807

Table 8.2: Event yields of $2\ell(\mathrm{SS})1\tau_{\mathrm{had}}$ and $2\ell(\mathrm{OS})1\tau_{\mathrm{had}}$ channels before and after the fit. Negative entries indicate the subtracted contributions of the fake τ_{had} estimate. Entries that are not applicable to a channel are marked "—".

8.2 Impact of systematic uncertainties

The uncertainty of the measured signal strength is dominated by the statistical component in the $2\ell(\mathrm{SS})1\tau_{\mathrm{had}}$ channel, while in $2\ell(\mathrm{OS})1\tau_{\mathrm{had}}$ the dominant uncertainty is of systematic nature and re-

lated to the fake τ_{had} background. The impact of individual nuisance parameter is evaluated by auxiliary ML fits in which the NP values are fixed to their up and down variations. The difference $\Delta\mu$ of the best-fit signal strength to the unconditional fit corresponds to the impact of the NP. The impacts are presented in Fig. 8.2, sorted by size. Both the pre-fit and post-fit impacts are evaluated, but no significant differences are observed in data.

The NPs with the largest impacts are related to the data-driven background estimates. Instrumental uncertainties related to jet energy calibration, τ_{had} reconstruction, identification and calibration, and flavour tagging are important because they relate to the objects that are used in the event selection.

The NP labelled "Flavor tagging c-jet/τ_{had}" is ranked highly because of the requirement of the τ_{had} to not be *b*–tagged. This NP is related to the efficiency correction of tagging *c*–jets. Due to the lack of a dedicated efficiency measurements of *b*–tagging jets spawned by hadronic tau decays, the NP is assigned also to the correction for τ_{had} and its uncertainty is inflated by 20 %. This uncertainty could be reduced by a dedicated measurement as well as a dedicated tagger for rejecting *b*–jets faking hadronic taus.

The theoretical uncertainty on the $t\bar{t}H$ cross section propagates directly to the measured signal strength. These contributions can be made irrelevant when measuring the $t\bar{t}H$ cross section directly instead of normalising to the predicted value (see next chapter).

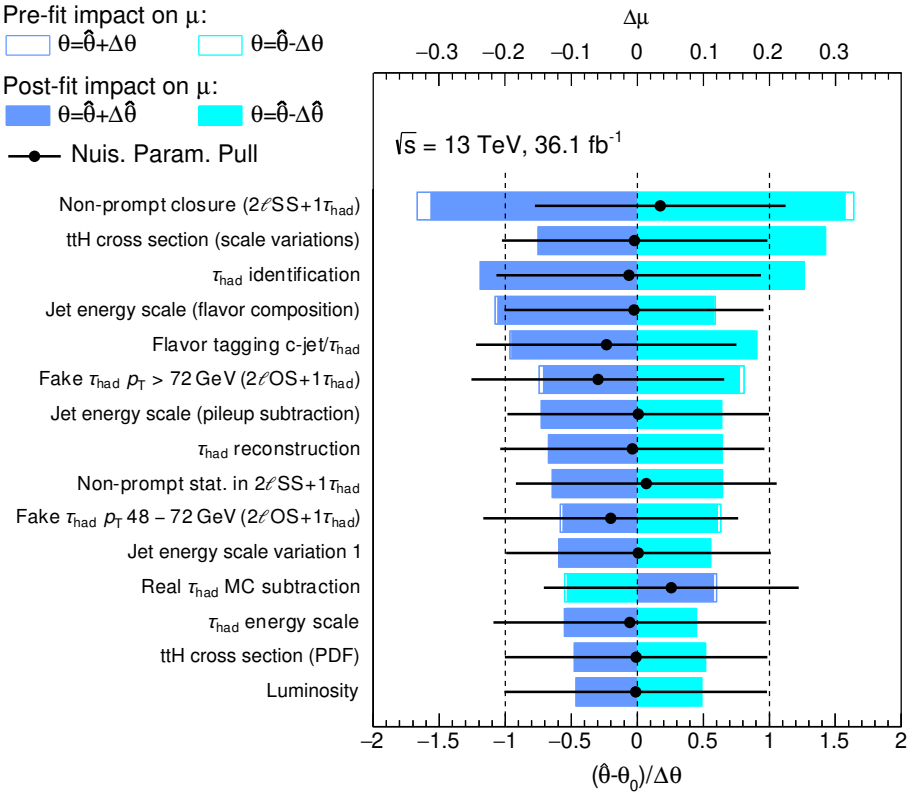


Figure 8.2: Impact of nuisance parameters on fitted signal strength uncertainty and best-fit values with uncertainties of nuisance parameters in the combined fit of dilepton+τ_{had} channels. The impact is calculated by fixing each NP to their up and down variations, both pre-fit (open bars) and post-fit (solid bars), and calculating the difference $\Delta\mu$ to the unconditional best-fit $\hat{\mu}$. The NPs are listed in order of decreasing size of their impact. The lower *x*–axis and black points show the best-fit value of the corresponding NP and its uncertainty relative to the pre-fit values and uncertainties.

CHAPTER 9

Results of all $t\bar{t}H$ multilepton channels

The ttHML analysis contains seven different analysis channels, of which two are described in detail in this thesis. The most significant channels are the $2\ell(\text{SS})0\tau_{\text{had}}$ and $3\ell0\tau_{\text{had}}$ ones with higher purity or larger event yield than the ones with hadronic taus. All following results are presented as they appear in the paper of Ref. [5], henceforth referred to as "the paper". This thesis describes a slightly different treatment of the channels than the paper. The differences are outlined below.

9.1 Differences to previous results

MC subtraction of $t\bar{t}$ background in FF method As described in Sec. 5.3.1, the backgrounds with true hadronic taus are subtracted in all control regions in the data-driven estimate of fake hadronic taus using truth-matched MC. Consequently, the $t\bar{t}$ background with a real hadronic tau is taken from MC in the signal region. These $t\bar{t}$ events contain leptonic decays of both top quarks where one of the leptons is a tau that decays hadronically. Therefore, one of the light leptons has to be non-prompt. In the paper this subtraction is not performed due to the assumption that the number of events with a real τ_{had} is negligible in selected $t\bar{t}$ events.

Concretely, the difference between the previously presented result in this thesis and the following result of the paper is that $t\bar{t}$ MC with true τ_{had} is included in this thesis while it is not in the paper. Therefore there is a difference in the total background yield prediction of 37 events which can be deduced from Tab. 8.2 by subtracting the entries containing "$t\bar{t}$" i.e., ($t\bar{t}$ in FF side-band with real τ_{had}) + ($t\bar{t}$) = 44 + (−7) = 37 fewer events in the paper.

Since, no subtraction is performed the events contributed in all control regions of the FF method and are therefore estimated as part of it. While the difference of 37 events is within the uncertainty of the fake estimate and does not invalidate either one of the treatments, more investigation might be desirable.

The expected significances in the $2\ell(\text{OS})1\tau_{\text{had}}$ channel is not affected by this difference. However, the observed significance is.

Discriminating observable in $2\ell(\text{SS})1\tau_{\text{had}}$ The discriminating observable used in the fit of the paper for the $2\ell(\text{SS})1\tau_{\text{had}}$ channel is a BDT that is trained on a loose event selection with $t\bar{t}H$ as signal and all other processes as background. In the loose selection the dominant selected process is $t\bar{t}$ with a non-prompt lepton. Accordingly, the BDT discriminates mainly against this process and less so against others like $t\bar{t}V$ that are more relevant in the tighter SR. The reason for the looser event selection for the BDT training, is a lack of sufficient MC statistics. The gain in expected significance from using the BDT distribution in two bins rather than the event count is approximately 30 %.

9.2 Combined fit model

Multivariate discriminants are used to create the templates that are fitted in five of the seven analysis channels. The 4ℓ channel also defines one of its signal regions with a cut on a BDT. The only channels that merely counts the number of events in a single bin is $3\ell 1\tau_{\text{had}}$. Additional control regions that enter the fit are defined with three light leptons that are enriched in $t\bar{t}W$, $t\bar{t}Z$, VV, and $t\bar{t}$, each with one bin. In total there are 32 bins in the maximum likelihood fit. The setup of discriminant observables and number of regions is summarised in Tab. 9.1.

	$2\ell(\text{SS})0\tau_{\text{had}}$	$3\ell 0\tau_{\text{had}}$	4ℓ	$1\ell 2\tau_{\text{had}}$	$2\ell(\text{SS})1\tau_{\text{had}}$	$2\ell(\text{OS})1\tau_{\text{had}}$	$3\ell 1\tau_{\text{had}}$
BDT trained against	Non-prompt and $t\bar{t}V$	$t\bar{t}$, $t\bar{t}W$, $t\bar{t}Z$, VV	$t\bar{t}Z$ / -	$t\bar{t}$	all	$t\bar{t}$	-
Discriminant	2×1D BDT shape	5D BDT shape	Event count	BDT shape	BDT shape	BDT shape	Event count
Number of bins	6	5	1 / 1	2	2	10	1
Control regions (1 bin each)	-	4	-	-	-	-	-

Table 9.1: Summary of combined fit model for all ttHML channels [5]. The two numbers in the 4ℓ column refer to the Z enriched and depleted categories.

9.3 Measured signal strength and significance

The best-fit signal strength of the combined fit of all channels is

$$\hat{\mu} = 1.56\,^{+0.49}_{-0.42}(\text{tot.}) = 1.56\,^{+0.30}_{-0.29}(\text{stat.})\,^{+0.39}_{-0.30}(\text{syst.}). \tag{9.1}$$

The upward deviation from the expectation that was already shown persists in the full combination. The significance of the excess with respect to the background-only hypothesis is 4.1σ. This constitutes evidence for the production of the $t\bar{t}H$ process. Since the expected significance is only 2.8σ, the compatibility with the SM hypothesis is also determined. The measured result is compatible with the SM expectation of $\mu = 1$ within 1.4σ.

Signal strengths for the individual channels are obtained from a fit with seven floating μ, one per channel. This setup means that all nuisance parameters are correlated among channels. This is particularly relevant for channels with 1 bin or a small number of bins. In such a situation all or most of the differences between the observed data and prediction can be absorbed in the signal strength that is freely floating, leaving the nuisance parameters unchanged. The combined fit allows best-fit values of the NPs to be propagated from channels with more statistical power to those with less.

All observed signal strengths are plotted in Fig. 9.1 sorted by the size of their uncertainty. All but $2\ell(\text{SS})1\tau_{\text{had}}$ are compatible with $\mu = 1$ within approximately one standard deviation. The channels 4ℓ and $3\ell 1\tau_{\text{had}}$ which expect a small number of total events are statistically limited. The other channels have similar contributions from statistical and systematic components of the uncertainty. Two channels, 4ℓ and $1\ell 2\tau_{\text{had}}$, observed a downward fluctuation below the expected number of background events. This is represented by a negative best-fit value of μ.

Both expected and observed signal strengths and significances are given in Tab. 9.2 for all channels, the full combination and partial combination of channels that do or do not select hadronic taus. The channels are sorted by presence of a τ_{had} in the event selection and the expected significance. One can

see that $2\ell(\text{SS})0\tau_{\text{had}}$ and $3\ell0\tau_{\text{had}}$ are the major contributors to the overall significance of the analysis. However, no single channel can be considered significant by itself. The strength of the ttHML analysis, which is also the overall strategy in the $t\bar{t}H$ search, lies in the combination of multiple channels.

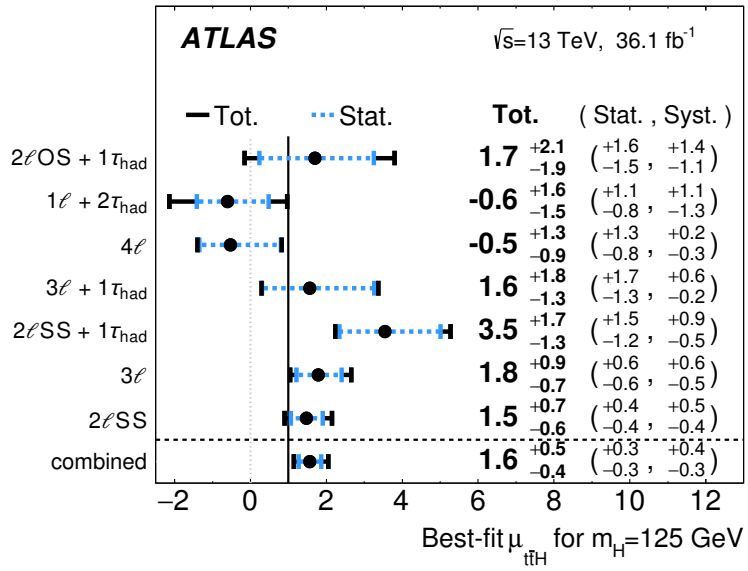


Figure 9.1: Best-fit signal strengths $\hat{\mu}_{t\bar{t}H}$ and their uncertainties for all ttHML channels and the combination. The individual μs are obtained in a combined fit with one μ per channel, and the combined with one common μ. The channels are sorted by decreasing size of uncertainty [5].

9.4 Impact of systematic uncertainties

The impact of systematic uncertainties is determined for single nuisance parameters and shown in Fig. 9.2. The highest ranking NP are those related to the most significant channels i.e., $2\ell(\text{SS})0\tau_{\text{had}}$ and $3\ell0\tau_{\text{had}}$, and are uncertainties on the theoretical prediction of $t\bar{t}H$ and $t\bar{t}V$ cross sections and instrumental jet uncertainties.

The data-driven estimates of non-prompt lepton background also have a large impact on the result but only appear in the ranking plot with one NP because the uncertainty is split into many components. The impact of groups of nuisance parameters is shown in Tab. 9.3. These impacts are calculated by performing a fit where the group of nuisance parameters is kept constant. The resulting $\Delta\mu$ is smaller than the total $\Delta\mu$ and the impact of the group is then defined as the quadratic difference of the two. The table is sorted by size of impact and shows the non-prompt background estimates in third place.

The second largest impact of a group of uncertainties is due to the jet calibration as seen in the ranking plot already. The jet calibration affects both the acceptance of events in the event selection as well as kinematic properties of the jets which enter the discriminating BDTs.

The largest impact on the results is caused by the uncertainty on the predicted $t\bar{t}H$ cross section. In the table, "ttH modelling (cross section)" only affects the predicted cross section i.e., the denominator of the signal strength $\mu = \frac{\sigma_{\text{observed}}}{\sigma_{\text{predicted}}}$. This uncertainty is irrelevant when only considering the direct measurement of the cross section without normalising to the prediction. This form of the result is given in the next section.

Channel	$\hat{\mu}$	$\Delta\hat{\mu}$(tot.)	$\Delta\hat{\mu}$(stat., syst.)	Significance ($\mu = 0$)	Significance ($\mu = 1$)
			Expected		
2ℓ(OS)$1\tau_{\text{had}}$	1.00	$^{+1.93}_{-1.76}$	$^{+1.48\ +1.24}_{-1.38\ -1.10}$	0.54	—
$1\ell 2\tau_{\text{had}}$	1.00	$^{+1.68}_{-1.45}$	$^{+1.14\ +1.23}_{-0.94\ -1.10}$	0.59	—
$3\ell 1\tau_{\text{had}}$	1.00	$^{+1.57}_{-1.11}$	$^{+1.51\ +0.48}_{-1.10\ -0.29}$	0.85	—
2ℓ(SS)$1\tau_{\text{had}}$	1.00	$^{+1.15}_{-0.90}$	$^{+1.05\ +0.48}_{-0.85\ -0.29}$	1.05	—
tau	1.00	$^{+0.82}_{-0.68}$	$^{+0.64\ +0.51}_{-0.58\ -0.36}$	1.53	—
4ℓ	1.00	$^{+1.81}_{-1.16}$	$^{+1.75\ +0.49}_{-1.15\ -0.18}$	0.81	—
$3\ell 0\tau_{\text{had}}$	1.00	$^{+0.74}_{-0.68}$	$^{+0.56\ +0.48}_{-0.52\ -0.43}$	1.48	—
2ℓ(SS)$0\tau_{\text{had}}$	1.00	$^{+0.60}_{-0.55}$	$^{+0.41\ +0.44}_{-0.39\ -0.38}$	1.83	—
notau	1.00	$^{+0.49}_{-0.45}$	$^{+0.33\ +0.36}_{-0.32\ -0.31}$	2.34	—
all	1.00	$^{+0.41}_{-0.38}$	$^{+0.29\ +0.29}_{-0.27\ -0.27}$	2.79	—
			Observed		
2ℓ(OS)$1\tau_{\text{had}}$	1.70	$^{+2.10}_{-1.85}$	$^{+1.56\ +1.41}_{-1.46\ -1.14}$	0.92	0.39
$1\ell 2\tau_{\text{had}}$	−0.61	$^{+1.57}_{-1.53}$	$^{+1.06\ +1.16}_{-0.84\ -1.28}$	—	—
$3\ell 1\tau_{\text{had}}$	1.57	$^{+1.80}_{-1.27}$	$^{+1.69\ +0.61}_{-1.26\ -0.20}$	1.29	0.44
2ℓ(SS)$1\tau_{\text{had}}$	3.54	$^{+1.73}_{-1.31}$	$^{+1.47\ +0.92}_{-1.25\ -0.53}$	3.42	2.24
tau	2.11	$^{+0.99}_{-0.79}$	$^{+0.71\ +0.70}_{-0.65\ -0.46}$	3.16	1.46
4ℓ	−0.53	$^{+1.35}_{-0.87}$	$^{+1.32\ +0.25}_{-0.82\ -0.29}$	—	—
$3\ell 0\tau_{\text{had}}$	1.79	$^{+0.87}_{-0.75}$	$^{+0.61\ +0.61}_{-0.58\ -0.48}$	2.43	0.99
2ℓ(SS)$0\tau_{\text{had}}$	1.47	$^{+0.67}_{-0.58}$	$^{+0.43\ +0.51}_{-0.41\ -0.41}$	2.64	0.86
notau	1.42	$^{+0.54}_{-0.48}$	$^{+0.34\ +0.42}_{-0.33\ -0.35}$	3.14	0.87
all	1.56	$^{+0.49}_{-0.42}$	$^{+0.30\ +0.39}_{-0.29\ -0.30}$	4.14	1.36

Table 9.2: Expected and observed best-fit signal strengths and significances for all channels in the ttHML analysis and combinations thereof. Individual signal strengths are obtained from a combined fit with one independent signal strength per channel. While the tau and notau sub-combinations are independent. Significances are given for both the background-only ($\mu = 0$) and the signal-plus-background hypothesis ($\mu = 1$)

The other entry, "$t\bar{t}H$ modelling (acceptance)" contain the uncertainty on Higgs branching fractions and the $t\bar{t}H$ MC parton shower and shower tune variations. It affects the number of selected events as well as the predicted cross section and therefore cannot be made irrelevant.

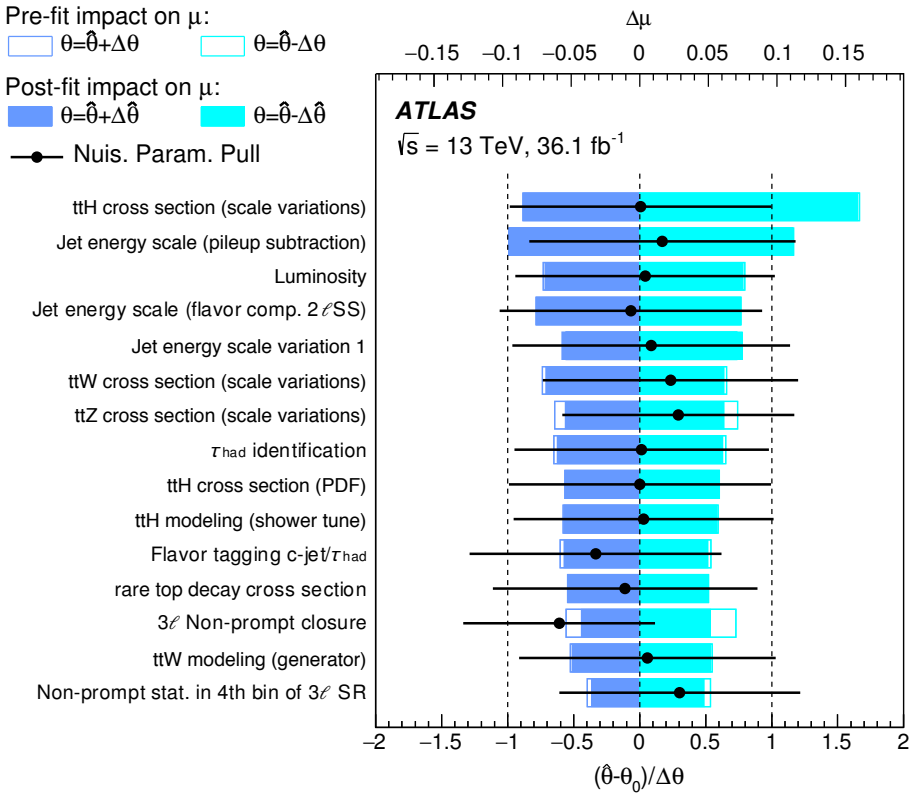


Figure 9.2: Impact of nuisance parameters on fitted signal strength uncertainty and best-fit values with uncertainties of nuisance parameters in the combined fit of all ttHML channels. The impact is calculated by fixing each NP to their up and down variations, both pre-fit (open bars) and post-fit (solid bars), and calculating the difference $\Delta\mu$ to the unconditional best-fit $\hat{\mu}$. The NPs are listed in order of decreasing size of their impact. The lower x–axis and black points show the best-fit value of the corresponding NP and its uncertainty relative to the pre-fit values and uncertainties [5].

9.5 Measuring $t\bar{t}H$ cross section

The inclusive cross section of the $t\bar{t}H$ process measured by this analysis is

$$\sigma_{t\bar{t}H} = 790 \pm 150(\text{stat.})\,^{+170}_{-150}(\text{syst.})\,\text{fb} = 790\,^{+230}_{-210}\,\text{fb}. \tag{9.2}$$

The relative uncertainty on the cross section is reduced compared to μ by the omission of the first entry in Tab. 9.3. The measurement is to be compared to the prediction of $\sigma^{\text{SM}}_{t\bar{t}H} = 507^{+35}_{-50}$ fb.

9.6 Visualisation of results

To visualise the results obtained from a fit with 32 bins it is of course straight-forward to look at the plots of all signal regions as in Figs. 9.3 and 9.4. However it is also convenient to make a single plot

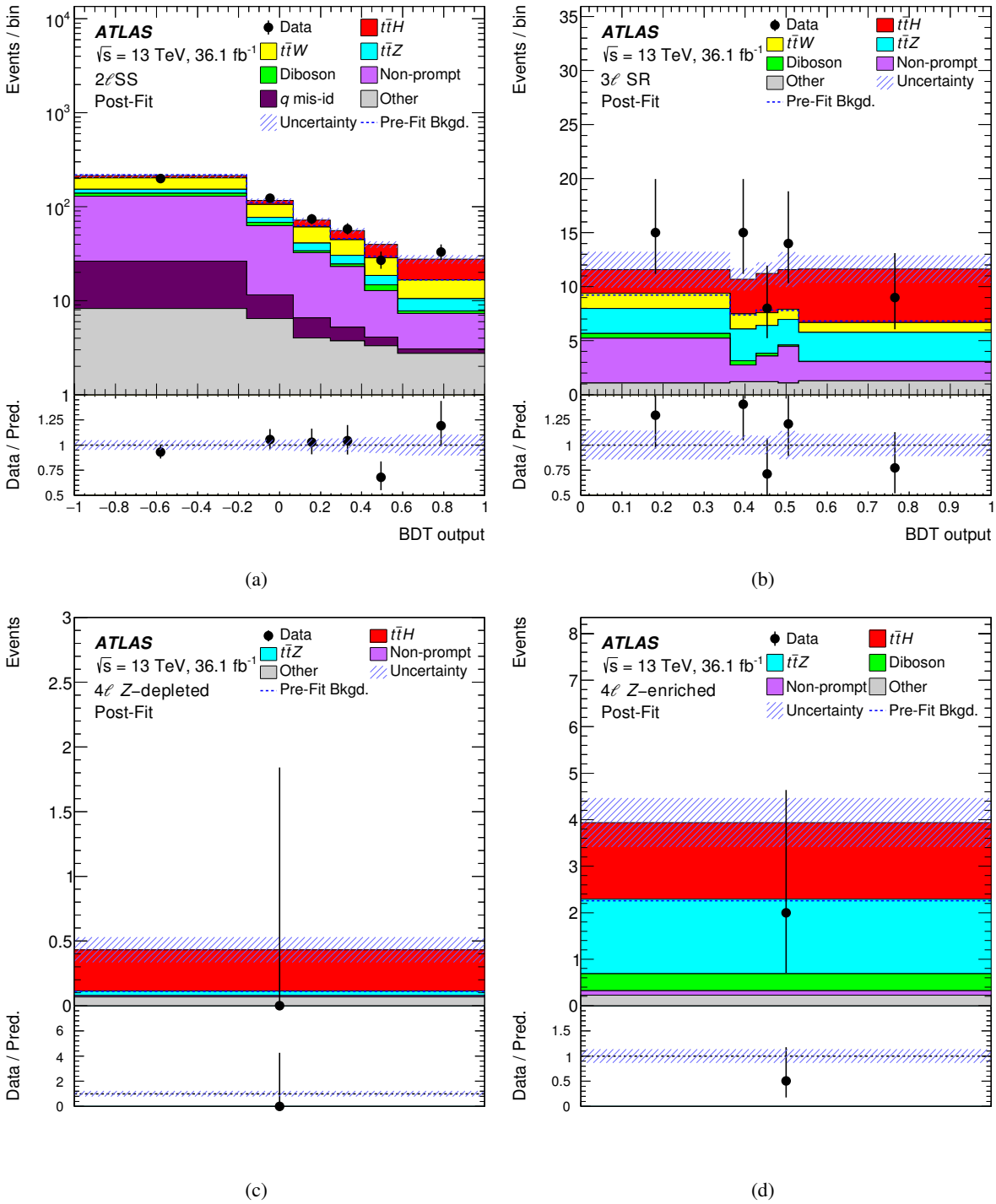
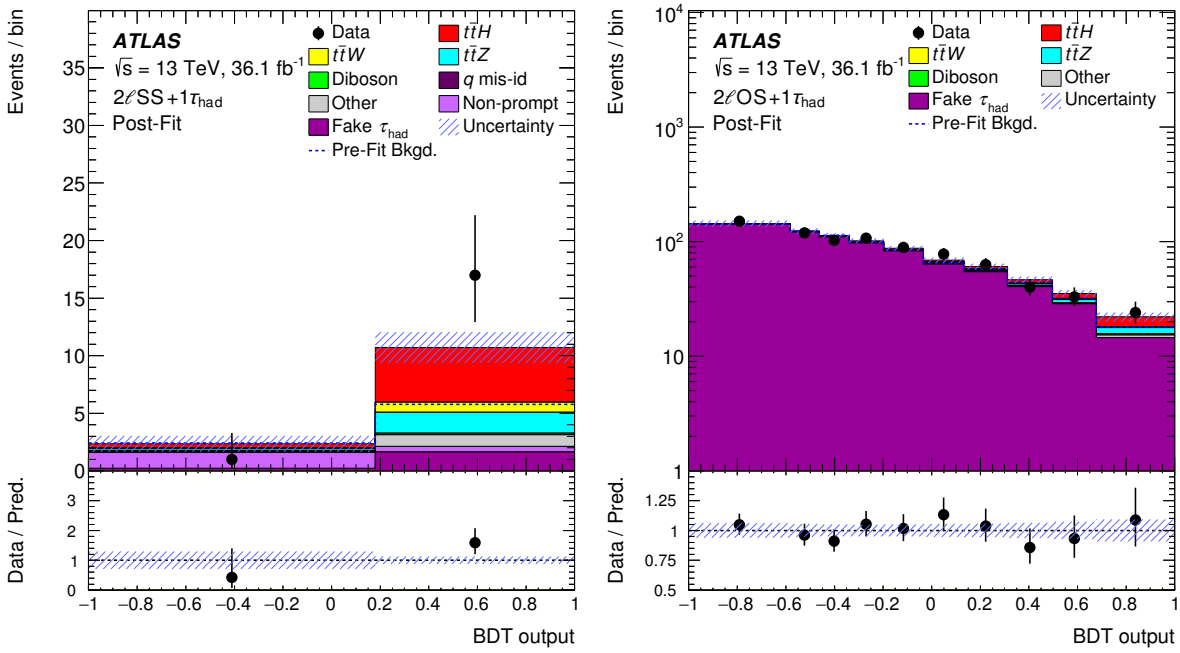
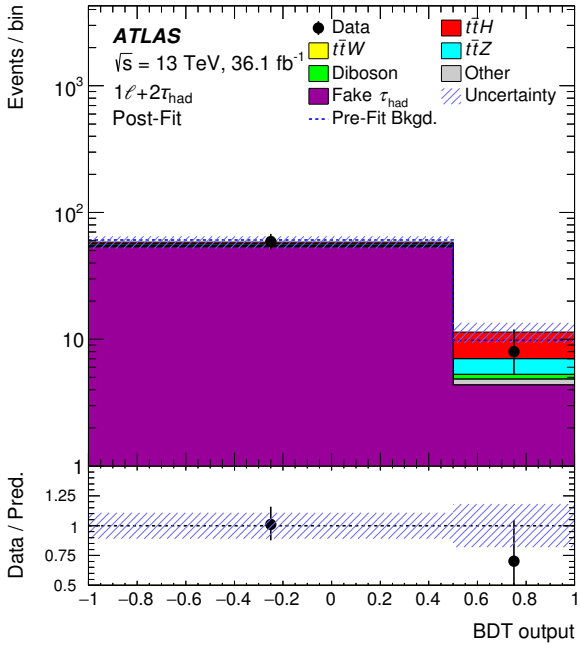


Figure 9.3: Post-Fit plots of the discriminants of the channels not selecting a τ_{had} used in the combined fit of all ttHML channels. The signal is scaled to the observed best-fit signal strength. The expected (pre-fit) background is overlayed as a blue dashed line and is almost identical to the post-fit background. "Non-prompt" refers to the data-driven estimates of this background [5].

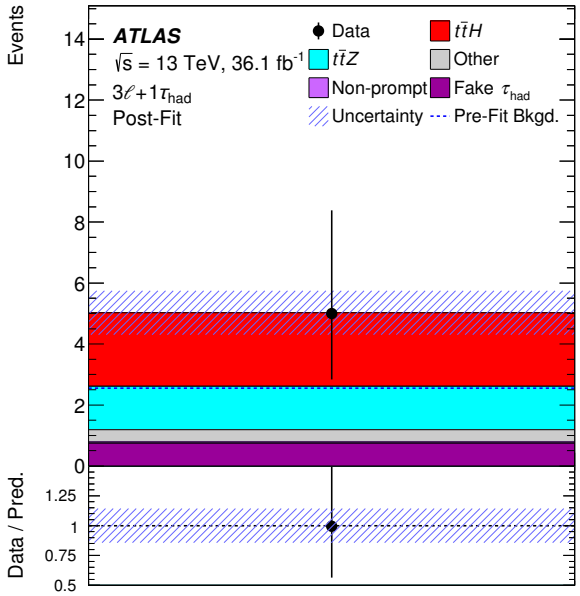


(a)

(b)



(c)



(d)

Figure 9.4: Post-Fit plots of the discriminants of the channels that do select a τ_{had} used in the combined fit of all ttHML channels. The signal is scaled to the observed best-fit signal strength. The expected (pre-fit) background is overlayed as a blue dashed line and is almost identical to the post-fit background. "Non-prompt" and "Fake τ_{had}" refer to the data-driven estimates of these backgrounds [5].

Uncertainty Source	$\Delta\mu$	
$t\bar{t}H$ modelling (cross section)	+0.20	−0.09
Jet energy scale and resolution	+0.18	−0.15
Non-prompt light-lepton estimates	+0.15	−0.13
Jet flavor tagging and τ_{had} identification	+0.11	−0.09
$t\bar{t}W$ modelling	+0.10	−0.09
$t\bar{t}Z$ modelling	+0.08	−0.07
Other background modelling	+0.08	−0.07
Luminosity	+0.08	−0.06
$t\bar{t}H$ modelling (acceptance)	+0.08	−0.04
Fake τ_{had} estimates	+0.07	−0.07
Other experimental uncertainties	+0.05	−0.04
Simulation sample size	+0.04	−0.04
Charge misassignment	+0.01	−0.01
Total systematic uncertainty	+0.39	−0.30

Table 9.3: Impacts of groups of systematic uncertainties on the best-fit signal strength $\hat{\mu}$ in the combined ttHML fit. Due to rounding and correlations the total uncertainty is different from the quadratic sum of decomposed uncertainties. The uncertainties are sorted by decreasing impact [5].

that condenses all the information and clearly shows the significant presence of the signal. Such a plot is shown in Fig. 9.5, which shows the bins of the analysis as a function of their $\log_{10}(S/B)$, where S is the expected number of $t\bar{t}H$ events ($\mu = 1$) and B is the number of background events in each bin. Bins are combined to give a continuously decreasing distribution. This way of plotting is essentially sorting all bins by their S/B, such that the bins with largest signal purity are on the right. The plot contains multiple overlayed lines for the expected and observed signal strengths as well as the expected background. The black dotted line that shows the best-fit background from a conditional fit with the background-only hypothesis ($\mu = 0$) is particularly interesting. This line can be interpreted as how much the background can be adjusted within its uncertainties to describe the data if there existed no signal. It corresponds exactly to the maximised likelihood function value in the numerator of the test statistic (Eq. 7.3). The lower panel shows the background-subtracted data normalised to the background uncertainty (also called "pulls"). It shows that the purest (rightmost) bins are also the most significant.

9.7 Fit cross-checks

Several checks are performed to evaluate the robustness and possible bias of the combined fit. All checks are concluded successfully.

9.7.1 Compatibility of channels

The observed signal strengths of the individual channels have relatively large differences. To determine whether the results from the two fits, one with seven independent μ and one with a single common μ, are

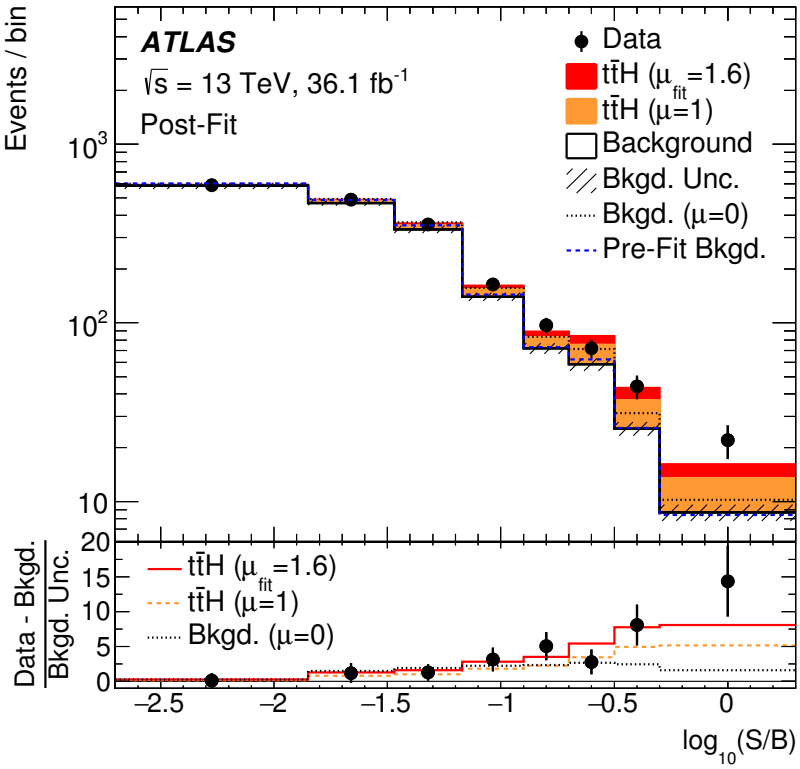


Figure 9.5: Summary plot of all ML channels. All bins of the analysis are sorted by S/B and combined as appropriate to plot $\log_{10}(S/B)$ where S is the expected signal yield and B the background yield from the unconditional fit. The $t\bar{t}H$ contribution is stacked on top of the background for values of $\mu = 1$ (orange) and best-fit $\hat{\mu} = 1.6$ (red). The pre-fit background is shown as a dashed blue line in the top panel. The background from the conditional fit with $\mu = 0$ is shown as a dotted, black line. The lower panel shows the background-subtracted data normalised to the background uncertainty ("pulls"). The range of the x–axis is defined by the lowest and largest expected S/B ratios of any bin [5].

compatible statistically, a χ^2-test is performed. The difference between values of the minimised negative log-likelihood functions are interpreted as the χ^2 with degrees of freedom of $7 - 1$. The probability that the signal strengths in all seven channels are compatible is 34 %.

9.7.2 Measuring $t\bar{t}V$ normalisation

The processes with a top-quark pair and a vector boson, $t\bar{t}W$ and $t\bar{t}Z$, are major backgrounds in most of the channels. Accordingly, the uncertainty on the theoretical cross sections that are used to predict the yield of these processes have a large impact on the significance of the analysis. The contribution from these uncertainties can be avoided when measuring the yield of $t\bar{t}V$ directly, and thus not relying on the cross section anymore for normalisation. Another goal of the cross check is to determine if there are systematic biases due to $t\bar{t}W$ backgrounds that affect the fitted signal strength.

A cross-check motivated by this idea is performed by adding free normalisation factors (NF) to the maximum likelihood fit that scales the number of events of $t\bar{t}W$ and $t\bar{t}Z$, together and separately one at a time. These free parameters replace the constrained parameters that are related to the $t\bar{t}V$ cross section uncertainties in these fits.

The best-fit values of the signal strength and normalisation factors are shown in Tab. 9.4. The fitted $t\bar{t}H$ signal strengths are compatible in all configurations. However, the significance to discover $t\bar{t}H$ is reduced. This means that there is not sufficient statistical power to measure the $t\bar{t}V$ normalisation better than the theoretical uncertainty on their cross sections. This is not surprising as the analysis is not designed to do that.

The fitted values of the $t\bar{t}V$ NFs are compatible with the prediction (NF=1) as well as with each other in all configurations. The uncertainty on the estimated NFs is between 20-30 %. The dedicated control regions for $t\bar{t}W$ and $t\bar{t}Z$ and discriminant power particularly in $2\ell(\mathrm{SS})0\tau_{\mathrm{had}}$ and $3\ell0\tau_{\mathrm{had}}$ are the dominant causes for this relatively small uncertainty. With more data is should be possible to measure simultaneously $t\bar{t}H$, $t\bar{t}W$ and $t\bar{t}Z$.

It is concluded that the $t\bar{t}V$ background is well understood and does not bias the result of the $t\bar{t}H$ analysis.

	NF	ΔNF(±tot.)	ΔNF(±stat. ± syst.)	Significance ($\mu = 0$)	
				Observed	Expected
$t\bar{t}H$	1.56	$^{+0.49}_{-0.42}$	$^{+0.30\ +0.39}_{-0.29\ -0.30}$	4.14	2.79
$t\bar{t}W$	const. 1				
$t\bar{t}Z$	const. 1				
$t\bar{t}H$	1.57	$^{+0.57}_{-0.50}$	$^{+0.43\ +0.37}_{-0.42\ -0.27}$	3.39	2.27
$t\bar{t}W$	0.92	$^{+0.32}_{-0.32}$	$^{+0.22\ +0.23}_{-0.21\ -0.24}$		
$t\bar{t}Z$	1.17	$^{+0.25}_{-0.22}$	$^{+0.16\ +0.19}_{-0.15\ -0.16}$		
$t\bar{t}H$	1.49	$^{+0.50}_{-0.44}$	$^{+0.34\ +0.36}_{-0.33\ -0.30}$	3.59	2.50
$t\bar{t}W$	const. 1				
$t\bar{t}Z$	1.16	$^{+0.25}_{-0.21}$	$^{+0.16\ +0.19}_{-0.15\ -0.16}$		
$t\bar{t}H$	1.63	$^{+0.56}_{-0.49}$	$^{+0.41\ +0.37}_{-0.39\ -0.29}$	3.70	2.44
$t\bar{t}W$	0.95	$^{+0.31}_{-0.32}$	$^{+0.22\ +0.23}_{-0.21\ -0.23}$		
$t\bar{t}Z$	const. 1				

Table 9.4: Fit cross-check with free normalisation factors (NF) of $t\bar{t}V$ processes. Each horizontal line delimits one fit configuration where $t\bar{t}V$ is normalised to the prediction (NF is constant and equal 1) or is fitted to the data (NF is estimated), separately for $t\bar{t}W$ and $t\bar{t}Z$.

CHAPTER 10

Conclusion

The analysis of LHC collision data recorded with the ATLAS detector at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ corresponding to $36.1\,\text{fb}^{-1}$ searching for the $t\bar{t}H$ process in multilepton final states is presented. The result is a measured total cross section of $t\bar{t}H$ of

$$\sigma_{t\bar{t}H} = 790 \pm 150(\text{stat.})\,^{+170}_{-150}(\text{syst.})\,\text{fb} = 790\,^{+230}_{-210}\,\text{fb}, \tag{10.1}$$

which compared to the prediction of $\sigma^{\text{SM}}_{t\bar{t}H} = 507^{+35}_{-50}\,\text{fb}$ establishes the first evidence for the existence of the $t\bar{t}H$ process at ATLAS. The observed significance of the measurement is 4.1σ, while the expected one is 2.8σ. The compatibility with the SM is 1.4σ.

The result also provides the first direct measurement of the top Yukawa coupling, which was previously known only through inference from loop interactions. The measured signal strength that is in good agreement with the SM expectation is more confirmation that the Higgs coupling to fermions is proportional to the mass.

The multilepton final state is categorised into 7 different analysis channels. All channels contribute to the combined result to achieve a significant measurement. The backgrounds in the analysis are distinguished into reducible and irreducible. Data-driven methods are developed and used to estimate the irreducible background of non-prompt electrons and muons, and fake hadronically decaying tau leptons. To reduce backgrounds multivariate discriminants are developed in all but one channels.

The channel with the final state $2\ell(\text{OS})1\tau_{\text{had}}$ is analysed for the first time. The sensitivity gained by its inclusion is small, however it is essential for validating the fake τ_{had} estimate in the more sensitive $2\ell(\text{SS})1\tau_{\text{had}}$ channel. In future, the $2\ell(\text{OS})1\tau_{\text{had}}$ channel might be considered a control region unless the background rejection can be increased.

Future developments of $t\bar{t}H$ analyses will become sensitive in single channels and separation of H decay modes can allow measurements of $t\bar{t}H$ with $H \to \tau\bar{\tau}$ to compete and compare with $t\bar{t}H$ with $H \to b\bar{b}$ in fully fermionic Higgs couplings.

Bibliography

[1] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508 (cit. on pp. 1, 7).

[2] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321 (cit. on pp. 1, 7).

[3] 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 pp. 1, 7).

[4] Nobel Media AB, *The Nobel Prize in Physics 2013*, URL: http://www.nobelprize.org/nobel_prizes/physics/laureates/2013/ (cit. on p. 1).

[5] ATLAS Collaboration, *Evidence for the Associated Production of the Higgs Boson and a Top Quark Pair with the ATLAS Detector*, Phys. Rev. D **97** (2018) 072003, arXiv: 1712.08891 [hep-ex] (cit. on pp. 1, 2, 43, 44, 48, 50, 52, 53, 55–57, 78, 79, 93–95, 97–101, 121).

[6] Particle Data Group, M. Tanabashi et al., *Review of Particle Physics*, Phys. Rev. D **98** (2018) 030001, URL: pdg.lbl.gov (cit. on pp. 3, 4, 12, 15, 23–26).

[7] S. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579 (cit. on p. 4).

[8] A. Salam, "Weak and Electromagnetic Interactions", *Elementary particle theory. Relativistic groups and analyticity. Proceedings of the Eighth Nobel Symposium*, ed. by N. Svartholm, Almquist & Wiksell, 1968 367 (cit. on p. 4).

[9] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264 (cit. on pp. 4, 9).

[10] M. Merk, I. van Vulpen and W. Hulsbergen, *Particle Physics 1: Lecture Notes Electroweak Standard Model*, Nikhef, 2017, URL: https://www.nikhef.nl/~wouterh/teaching/PP1/LectureNotes.pdf (visited on 25/05/2018) (cit. on p. 4).

[11] A. Djouadi, *The anatomy of electroweak symmetry breaking: Tome I: The Higgs boson in the Standard Model*, Physics Reports **457** (2008) 1, arXiv: hep-ph/0503172 [hep-ph] (cit. on p. 4).

[12] D. Griffiths, *Introduction to Elementary Particles*, 2nd edition, Wiley-VCH, 2008 (cit. on pp. 4, 11).

[13] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531 (cit. on p. 10).

[14] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Progress of Theoretical Physics **49** (1973) 652 (cit. on p. 10).

[15] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945 (cit. on p. 10).

[16] J. C. Collins, D. E. Soper and G. Sterman, *Factorization of Hard Processes in QCD*, Advanced Series on Directions in High Energy Physics **5** (1989), arXiv: hep-ph/0409313 (cit. on p. 12).

[17] The NNPDF Collaboration et al., *Parton distributions from high-precision collider data*, to be published in EPJC (2017), arXiv: 1706.00428 [hep-ph] (cit. on p. 13).

[18] ATLAS Collaboration, *Plot of summary of Standard Model production cross section measurements*, URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SM/ATLAS_a_SMSummary_TotalXsect/ATLAS_a_SMSummary_TotalXsect.eps (visited on 06/06/2018) (cit. on p. 14).

[19] ATLAS Collaboration, *Charged-particle distributions at low transverse momentum in $\sqrt{s}$ = 13 TeV pp interactions measured with the ATLAS detector at the LHC*, Eur. Phys. J. C **76** (2016) 502, arXiv: 1606.01133 [hep-ex] (cit. on p. 15).

[20] ATLAS Collaboration, *Measurement of the charged-particle multiplicity inside jets from $\sqrt{s}$ = 8 TeV pp collisions with the ATLAS detector*, Eur. Phys. J. C **76** (2016) 322, arXiv: 1602.00988 [hep-ex] (cit. on p. 15).

[21] P. Skands, "Introduction to QCD", *Proceedings of the 2012 Theoretical Advanced Study Institute (TASI) in Elementary Particle Physics*, ed. by K. Matchev and et al., 2013, arXiv: 1207.2389 [hep-ph] (cit. on p. 15).

[22] R. K. Ellis, W. J. Stirling and B. R. Webber, *QCD and Collider Physics*, Cambridge University Press, 2003 (cit. on p. 15).

[23] A. Buckley et al., *General-purpose event generators for LHC physics*, Physics Reports **504** (2011) 5, arXiv: 1101.2599 [hep-ph] (cit. on p. 15).

[24] T. Sjöstrand et al., *High-energy-physics event generation with Pythia 6.1*, Comput. Phys. Commun. **135** (2001) 238, arXiv: hep-ph/0010017 (cit. on pp. 15, 33, 34).

[25] T. Sjöstrand, S. Mrenna and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv: 0710.3820 [hep-ph] (cit. on pp. 15, 33, 34).

[26] M. Bahr et al., *Herwig++ physics and manual*, Eur. Phys. J. C **58** (2008) 639, arXiv: 0803.0883 [hep-ph] (cit. on pp. 15, 34).

[27] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: 0811.4622 [hep-ph] (cit. on pp. 15, 34).

[28] LHC Higgs Cross Section Working Group, *Plot of Standard Model Higgs boson production cross sections*, URL: https://twiki.cern.ch/twiki/pub/LHCPhysics/LHCHXSWGCrossSectionsFigures/Plot_Escan_H125_new_sqrt.pdf (visited on 04/06/2018) (cit. on p. 17).

[29] 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. 18–21, 45, 46).

[30] W. Beenakker et al., *NLO QCD corrections to $t\bar{t}H$ production in hadron collisions*, Nucl. Phys. B **653** (2003) 151, arXiv: hep-ph/0211352 (cit. on p. 19).

[31] S. Dawson et al.,
Associated Higgs production with top quarks at the large hadron collider: NLO QCD corrections,
Phys. Rev. D **68** (2003) 034022, arXiv: hep-ph/0305087 (cit. on p. 19).

[32] Y. Zhang et al.,
QCD NLO and EW NLO corrections to $t\bar{t}H$ production with top quark decays at hadron collider,
Phys. Lett. B **738** (2014) 1, arXiv: 1407.1110 [hep-ph] (cit. on p. 19).

[33] S. Frixione et al.,
Weak corrections to Higgs hadroproduction in association with a top-quark pair,
JHEP **09** (2014) 065, arXiv: 1407.0823 [hep-ph] (cit. on p. 19).

[34] 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. 20).

[35] 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. 20).

[36] ATLAS Collaboration, *Evidence for the $H \to b\bar{b}$ decay with the ATLAS detector*,
JHEP **12** (2017) 024, arXiv: 1708.03299 [hep-ex] (cit. on p. 21).

[37] ATLAS Collaboration, *Cross-section measurements of the Higgs boson decaying to a pair of tau leptons in proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*,
ATLAS-CONF-2018-021, 2018, URL: https://cds.cern.ch/record/2621794
(cit. on p. 21).

[38] CMS Collaboration, *Evidence for the Higgs boson decay to a bottom quark–antiquark pair*,
Phys. Lett. B **780** (2018) 501, arXiv: 1709.07497 [hep-ex] (cit. on p. 21).

[39] CMS Collaboration, *Observation of the Higgs boson decay to a pair of τ leptons*,
Phys. Lett. B **779** (2018) 283, arXiv: 1708.00373 [hep-ex] (cit. on p. 21).

[40] ATLAS Collaboration,
Search for the Decay of the Higgs Boson to Charm Quarks with the ATLAS Experiment,
Phys. Rev. Lett. **120** (2018) 211802, arXiv: 1802.04329 [hep-ex] (cit. on p. 21).

[41] ATLAS Collaboration, *Search for the Dimuon Decay of the Higgs Boson in pp Collisions at $\sqrt{s}$ = 13 TeV with the ATLAS Detector*, Phys. Rev. Lett. **119** (2017) 051802,
arXiv: 1705.04582 [hep-ex] (cit. on p. 21).

[42] ATLAS and CMS Collaborations,
Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s}$ = 7 and 8 TeV,
JHEP **08** (2016) 045, arXiv: 1606.02266 [hep-ex] (cit. on pp. 22, 23, 56).

[43] ATLAS and CMS Collaborations, *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s}$ = 7 and 8 TeV with the ATLAS and CMS Experiments*,
Phys. Rev. Lett. **114** (2015) 191803, arXiv: 1503.07589 [hep-ex] (cit. on p. 22).

[44] ATLAS Collaboration,
Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector,
Eur. Phys. J. C **75** (2015) 476, arXiv: 1506.05669 [hep-ex] (cit. on p. 22),
Erratum: Eur. Phys. J. C **76** (2016) 152.

[45] CMS Collaboration, *Combined search for anomalous pseudoscalar HVV couplings in $VH(H \to b\bar{b})$ production and $H \to VV$ decay*, Phys. Lett. B **759** (2016) 672, arXiv: 1602.04305 [hep-ex] (cit. on p. 22).

[46] ATLAS Collaboration, *Test of CP Invariance in vector-boson fusion production of the Higgs boson using the Optimal Observable method in the ditau decay channel with the ATLAS detector*, Eur. Phys. J. C **76** (2016) 658, arXiv: 1602.04516 [hep-ex] (cit. on p. 22).

[47] CDF Collaboration, *Observation of Top Quark Production in $\overline{p}p$ Collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74** (1995) 2626, arXiv: hep-ex/9503002 (cit. on p. 23).

[48] D0 Collaboration, *Search for High Mass Top Quark Production in $p\overline{p}$ Collisions at $\sqrt{s} = 1.8$ TeV*, Phys. Rev. Lett. **74** (1995) 2422, arXiv: hep-ex/9411001 (cit. on p. 23).

[49] S. W. Herb et al., *Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions*, Phys. Rev. Lett. **39** (1977) 252 (cit. on p. 23).

[50] M. L. Perl et al., *Evidence for Anomalous Lepton Production in $e^+ - e^-$ Annihilation*, Phys. Rev. Lett. **35** (1975) 1489 (cit. on p. 23).

[51] M. Czakon, P. Fiedler and A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$*, Phys. Rev. Lett. **110** (2013) 252004, arXiv: 1303.6254 [hep-ph] (cit. on p. 24).

[52] M. L. Perl, *The Tau Lepton*, Annual Review of Nuclear and Particle Science **30** (1980) 299 (cit. on p. 25).

[53] CKMfitter Group, J. Charles et al., Eur. Phys. J. C **41** (2005) 1, updated results and plots available at: URL: http://ckmfitter.in2p3.fr (cit. on p. 25).

[54] M. Davier, A. Höcker and Z. Zhang, *The physics of hadronic tau decays*, Reviews of Modern Physics **78** (2006) 1043, arXiv: hep-ph/0507078 (cit. on p. 25).

[55] CMS Collaboration, *Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons*, JHEP **05** (2014) 104, arXiv: 1401.5041 [hep-ex] (cit. on p. 25).

[56] ATLAS Collaboration, *Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector*, JHEP **04** (2015) 117, arXiv: 1501.04943 [hep-ex] (cit. on p. 25).

[57] N. Wermes, *Drawing of hadronic tau lepton decay*, ATLAS internal communication, 2014 (cit. on p. 26).

[58] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001 (cit. on p. 27).

[59] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003 (cit. on pp. 27–29, 32).

[60] CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004 (cit. on p. 27).

[61] LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005 (cit. on p. 27).

[62] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002 (cit. on p. 27).

[63] TE-EPC-LPC in LHC, *Drawing of CERN Accelerator Complex*, URL: http://te-epc-lpc.web.cern.ch/te-epc-lpc/machines/pagesources/Cern-Accelerator-Complex.jpg (visited on 12/04/2018) (cit. on p. 28).

[64] ATLAS Collaboration, *Number of Interactions per Crossing in 2015–2016*, 2016, URL: https://atlas.web.cern.ch/Atlas/GROUPS/DATAPREPARATION/PublicPlots/2017/DataSummary/figs/intlumivstimeRun2DQ.eps (visited on 12/04/2018) (cit. on p. 29).

[65] ATLAS Collaboration, *Computer Generated Image of the Whole ATLAS Detector*, 2008, URL: https://cds.cern.ch/record/1095924 (cit. on p. 30).

[66] 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. 32).

[67] ATLAS Collaboration, *Trigger Menu in 2016*, ATL-DAQ-PUB-2017-001, 2017, URL: https://cds.cern.ch/record/2242069 (cit. on p. 32).

[68] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2015*, Eur. Phys. J. C **77** (2017) 317, arXiv: 1611.09661 [hep-ex] (cit. on p. 33).

[69] ATLAS Collaboration, *Total Integrated Luminosity and Data Quality in 2015–2017*, 2017, URL: https://atlas.web.cern.ch/Atlas/GROUPS/DATAPREPARATION/PublicPlots/2017/DataSummary/figs/intlumivstimeRun2DQ.eps (visited on 12/04/2018) (cit. on p. 33).

[70] ATLAS Collaboration, *Delivered Luminosity versus time for 2011–2017 (p–p data only)*, 2017, URL: https://atlas.web.cern.ch/Atlas/GROUPS/DATAPREPARATION/PublicPlots/2017/DataSummary/figs/intlumivsyear.eps (visited on 12/04/2018) (cit. on p. 33).

[71] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823, arXiv: 1005.4568 [physics.ins-det] (cit. on p. 33).

[72] S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250 (cit. on p. 33).

[73] 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. 33).

[74] ATLAS Collaboration, *Summary of ATLAS Pythia 8 tunes*, ATL-PHYS-PUB-2012-003, 2012, URL: https://cds.cern.ch/record/1474107 (cit. on p. 33).

[75] A. Martin et al., *Parton distributions for the LHC*, Eur. Phys. J. C **63** (2009) 189, arXiv: 0901.0002 [hep-ph] (cit. on p. 33).

[76] ATLAS Collaboration, *Modelling of the $t\bar{t}H$ and $t\bar{t}V(V = W, Z)$ processes for $\sqrt{s} = 13$ TeV ATLAS analyses*, ATL-PHYS-PUB-2016-005, 2016, URL: https://cds.cern.ch/record/2120826 (cit. on p. 34).

[77] ATLAS Collaboration, *Multi-boson simulation for 13 TeV ATLAS analyses*, ATL-PHYS-PUB-2016-002, 2016, URL: https://cds.cern.ch/record/2119986 (cit. on p. 34).

[78] ATLAS Collaboration,
Simulation of top-quark production for the ATLAS experiment at $\sqrt{s}$ = 13 TeV,
ATL-PHYS-PUB-2016-004, 2016, URL: https://cds.cern.ch/record/2120417
(cit. on p. 34).

[79] 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. 34).

[80] ATLAS Collaboration, *ATLAS Run 1 Pythia8 tunes*, ATL-PHYS-PUB-2014-021, 2014,
URL: https://cds.cern.ch/record/1966419 (cit. on p. 34).

[81] The NNPDF collaboration: R. D. Ball et al., *Parton distributions for the LHC Run II*,
JHEP **04** (2015) 040, arXiv: 1410.8849 [hep-ph] (cit. on p. 34).

[82] M. H. Seymour and A. Siodmok,
Constraining MPI models using σ_{eff} and recent Tevatron and LHC Underlying Event data,
JHEP **10** (2013) 113, arXiv: 1307.5015 [hep-ph] (cit. on p. 34).

[83] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024,
arXiv: 1007.2241 [hep-ph] (cit. on p. 34).

[84] F. Cascioli, P. Maierhofer and S. Pozzorini, *Scattering Amplitudes with Open Loops*,
Phys. Rev. Lett. **108** (2012) 111601, arXiv: 1111.5206 [hep-ph] (cit. on p. 34).

[85] T. Gleisberg and S. Hoeche, *Comix, a new matrix element generator*, JHEP **12** (2008) 039,
arXiv: 0808.3674 [hep-ph] (cit. on p. 34).

[86] S. Schumann and F. Krauss,
A Parton shower algorithm based on Catani-Seymour dipole factorisation, JHEP **03** (2008) 038,
arXiv: 0709.1027 [hep-ph] (cit. on p. 34).

[87] S. Hoeche et al., *QCD matrix elements + parton showers: The NLO case*, JHEP **04** (2013) 027,
arXiv: 1207.5030 [hep-ph] (cit. on p. 34).

[88] P. Skands, *Tuning Monte Carlo generators: The Perugia tunes*, Phys. Rev. D **82** (2010) 074018,
arXiv: 1005.3457 [hep-ph] (cit. on p. 34).

[89] J. Pumplin et al.,
New Generation of Parton Distributions with Uncertainties from Global QCD Analysis,
JHEP **07** (2002) 012, arXiv: hep-ph/0201195 (cit. on p. 34).

[90] P. M. Nadolsky et al., *Implications of CTEQ global analysis for collider observables*,
Phys. Rev. D **78** (2008) 013004, arXiv: 0802.0007 [hep-ph] (cit. on p. 34).

[91] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244,
arXiv: 1207.1303 [hep-ph] (cit. on p. 34).

[92] S. Frixione, G. Ridolfi and P. Nason,
A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction,
JHEP **09** (2007) 126, arXiv: 0707.3088 [hep-ph] (cit. on p. 34).

[93] 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. 34).

[94] S. Alioli et al.,
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 p. 34).

[95] 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 p. 34).

[96] ATLAS Collaboration, *An computer generated image representing how ATLAS detects particles*, 2013, URL: https://cds.cern.ch/record/1505342 (cit. on p. 35).

[97] 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. 36).

[98] 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. 36).

[99] 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. 36).

[100] 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. 36).

[101] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton–proton collision data*, ATLAS-CONF-2016-024, 2016,
URL: https://cds.cern.ch/record/2157687 (cit. on pp. 36, 83).

[102] 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. 36, 83).

[103] ATLAS Collaboration, *Properties of jets and inputs to jet reconstruction and calibration with the ATLAS detector using proton–proton collisions at $\sqrt{s}$ = 13 TeV*, ATL-PHYS-PUB-2015-036,
2015, URL: https://cds.cern.ch/record/2044564 (cit. on p. 37).

[104] M. Cacciari, G. P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063,
arXiv: 0802.1189 [hep-ph] (cit. on p. 37).

[105] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*,
Phys. Rev. D **96** (2017) 072002, arXiv: 1703.09665 [hep-ex] (cit. on pp. 37, 84).

[106] 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 pp. 37, 84).

[107] ATLAS Collaboration, *Performance of b-jet identification in the ATLAS experiment*,
JINST **11** (2016) P04008, arXiv: 1512.01094 [hep-ex] (cit. on pp. 37, 84).

[108] ATLAS Collaboration, *Expected performance of the ATLAS b-tagging algorithms in Run-2*, ATL-PHYS-PUB-2015-022, 2015, URL: https://cds.cern.ch/record/2037697 (cit. on p. 37).

[109] ATLAS Collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, ATL-PHYS-PUB-2016-012, 2016, URL: https://cds.cern.ch/record/2160731 (cit. on p. 37).

[110] 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://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2015-045 (cit. on pp. 38–40, 83).

[111] C. Deutsch, *Drawing of tau jet and quark/gluon jet*, ATLAS internal communication, 2017 (cit. on p. 38).

[112] ATLAS Collaboration, *Local Hadronic Calibration*, ATL-LARG-PUB-2009-001-2, 2008, URL: https://cds.cern.ch/record/1112035 (cit. on p. 38).

[113] 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 pp. 40–42).

[114] ATLAS Collaboration, *Reconstruction of hadronic decay products of tau leptons with the ATLAS experiment*, Eur. Phys. J. C **76** (2016) 295, arXiv: 1512.05955 [hep-ex] (cit. on p. 41).

[115] ATLAS Collaboration, *Expected performance of missing transverse momentum reconstruction for the ATLAS detector at* $\sqrt{s} = 13$ *TeV*, ATL-PHYS-PUB-2015-023, 2015, URL: https://cds.cern.ch/record/2037700 (cit. on p. 41).

[116] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector in the first proton–proton collisions at* $\sqrt{s} = 13$ *TeV*, ATL-PHYS-PUB-2015-027, 2015, URL: https://cds.cern.ch/record/2037904 (cit. on p. 41).

[117] N. Bruscino, *A gateway to new physics: direct measurement of the top Yukawa coupling to the Higgs boson*, urn:nbn:de:hbz:5n-48058, PhD Thesis: Universität Bonn, 2017, URL: http://hss.ulb.uni-bonn.de/2017/4805/4805.htm (cit. on p. 45).

[118] ATLAS Collaboration, *Search for the Standard Model Higgs boson decaying into* $b\bar{b}$ *produced in association with top quarks decaying hadronically in pp collisions at* $\sqrt{s} = 8$ *TeV with the ATLAS detector*, JHEP **05** (2016) 160, arXiv: 1604.03812 [hep-ex] (cit. on p. 45).

[119] ATLAS Collaboration, *Search for the Standard Model Higgs boson produced in association with top quarks and decaying into a* $b\bar{b}$ *pair in pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Rev. D **97** (2018) 072016, arXiv: 1712.08895 [hep-ex] (cit. on pp. 45, 56).

[120] ATLAS Collaboration, *Measurements of fiducial cross-sections for $t\bar{t}$ production with one or two additional b-jets in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector*, Eur. Phys. J. C **76** (2016) 11, arXiv: 1508.06868 [hep-ex] (cit. on p. 45).

[121] ATLAS Collaboration, *Measurements of Higgs boson properties in the diphoton decay channel with 36 fb^{-1} of pp collision data at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Phys. Rev. D **98** (2018) 052005, arXiv: 1802.04146 [hep-ex] (cit. on pp. 45, 56).

[122] ATLAS Collaboration, *Measurement of the Higgs boson coupling properties in the $H \to ZZ^* \to 4\ell$ decay channel at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, JHEP **03** (2018) 095, arXiv: 1712.02304 [hep-ex] (cit. on pp. 46, 56).

[123] B. Stapf, *Studies of $t\bar{t}H(\tau\tau)$ event topologies and kinematics with the ATLAS experiment*, BONN-IB-2015-03, Master's Thesis: Universität Bonn, 2015, URL: https://www.hep1.physik.uni-bonn.de/results/data/internal/stapf.pdf (cit. on p. 48).

[124] CMS Collaboration, *Evidence for associated production of a Higgs boson with a top quark pair in final states with electrons, muons, and hadronically decaying τ leptons at $\sqrt{s}$ = 13 TeV*, JHEP **08** (2018) 066, arXiv: 1803.05485 [hep-ex] (cit. on p. 56).

[125] CMS Collaboration, *Observation of $t\bar{t}H$ production*, Phys. Rev. Lett. **120** (2018) 231801, arXiv: 1804.02610 [hep-ex] (cit. on p. 56).

[126] ATLAS Collaboration, *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, Phys. Lett. B **784** (2018) 173, arXiv: 1806.00425 [hep-ex] (cit. on p. 56).

[127] ATLAS Collaboration, *Search for the associated production of the Higgs boson with a top quark pair in multilepton final states with the ATLAS detector*, Phys. Lett. B **749** (2015) 519, arXiv: 1506.05988 [hep-ex] (cit. on pp. 56, 57).

[128] CMS Collaboration, *Search for $t\bar{t}H$ Production in the $H b\bar{b}$ Decay Channel with Leptonic $t\bar{t}$ Decays in Proton-Proton Collisions at $\sqrt{s}$ = 13 TeV*, Submitted to JHEP (2018), arXiv: 1804.03682 [hep-ex] (cit. on p. 56).

[129] CMS Collaboration, *Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at sqrt(s) = 13 TeV*, JHEP **11** (2017) 047, arXiv: 1706.09936 [hep-ex] (cit. on p. 56).

[130] CMS Collaboration, *Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, JHEP **11** (2018) 185, arXiv: 1804.02716 [hep-ex] (cit. on p. 56).

[131] CMS Collaboration, *Combined Measurements of the Higgs Boson's Couplings at $\sqrt{s}$ = 13 TeV*, CMS-PAS-HIG-17-031, 2018, URL: http://cds.cern.ch/record/2308127 (cit. on p. 56).

[132] E. Gross and L. Zivković, *$t\bar{t}H \to t\bar{t}\tau^+\tau^-$—toward the measurement of the top-Yukawa coupling*, Eur. Phys. J. C **59** (2009) 731 (cit. on p. 57).

[133] ATLAS Collaboration, *Search for additional heavy neutral Higgs and gauge bosons in the ditau final state produced in 36 fb^{-1} of pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, JHEP **01** (2018) 055, arXiv: 1709.07242 [hep-ex] (cit. on p. 60).

[134] CMS Collaboration, *Measurement of the $Z/\gamma^* \to \tau\tau$ cross section in pp collisions at $\sqrt{s}$ = 13 TeV and validation of τ lepton analysis techniques*, Eur. Phys. J. C **78** (2018) 708, arXiv: 1801.03535 [hep-ex] (cit. on p. 60).

[135] T. Hastie, R. Tibshirani and J. Friedman,
The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer, 2013 (cit. on p. 69).

[136] A. Hoecker et al., *TMVA - Toolkit for Multivariate Data Analysis*, 2007,
arXiv: physics/0703039 [physics.data-an] (cit. on pp. 69, 71).

[137] G. James et al., *An Introduction to Statistical Learning: With Applications in R*, Springer, 2013 (cit. on p. 71).

[138] F. James, *MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1*, 1994 (cit. on p. 82).

[139] G. Cowan, *Statistical Data Analysis*, Clarendon Press, 1998 (cit. on p. 82).

[140] S. Brandt, *Data Analysis: Statistical and Computational Methods for Scientists and Engineers*, 4th ed., Springer, 2014 (cit. on p. 82).

[141] K. Cranmer et al.,
HistFactory: A tool for creating statistical models for use with RooFit and RooStats,
CERN-OPEN-2012-016, 2012 (cit. on p. 82).

[142] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, 2003,
arXiv: physics/0306116 [physics.data-an] (cit. on p. 82).

[143] G. Cowan et al., *Asymptotic formulae for likelihood-based tests of new physics*,
Eur. Phys. J. C **71** (2011) 1554, arXiv: 1007.1727 [physics.data-an],
Erratum: Eur. Phys. J. C **73** (2013) 2501 (cit. on pp. 82, 87).

[144] ATLAS Collaboration,
Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC,
Eur. Phys. J. C **76** (2016) 653, arXiv: 1608.03953 [hep-ex] (cit. on p. 84).

[145] ATLAS Collaboration,
Luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC,
Eur. Phys. J. C **71** (2011) 1630, arXiv: 1101.2185 [hep-ex] (cit. on p. 84).

[146] J. R. Klein and A. Roodman, *BLIND ANALYSIS IN NUCLEAR AND PARTICLE PHYSICS*,
Annual Review of Nuclear and Particle Science **55** (2005) 141 (cit. on p. 87).

APPENDIX A

MC plots in control regions

Figure A.1 shows the transverse momentum of the τ_{had} in all control regions used in the fake τ_{had} estimate of $2\ell(\text{OS})1\tau_{\text{had}}$ (see Sec. 5.3.1).

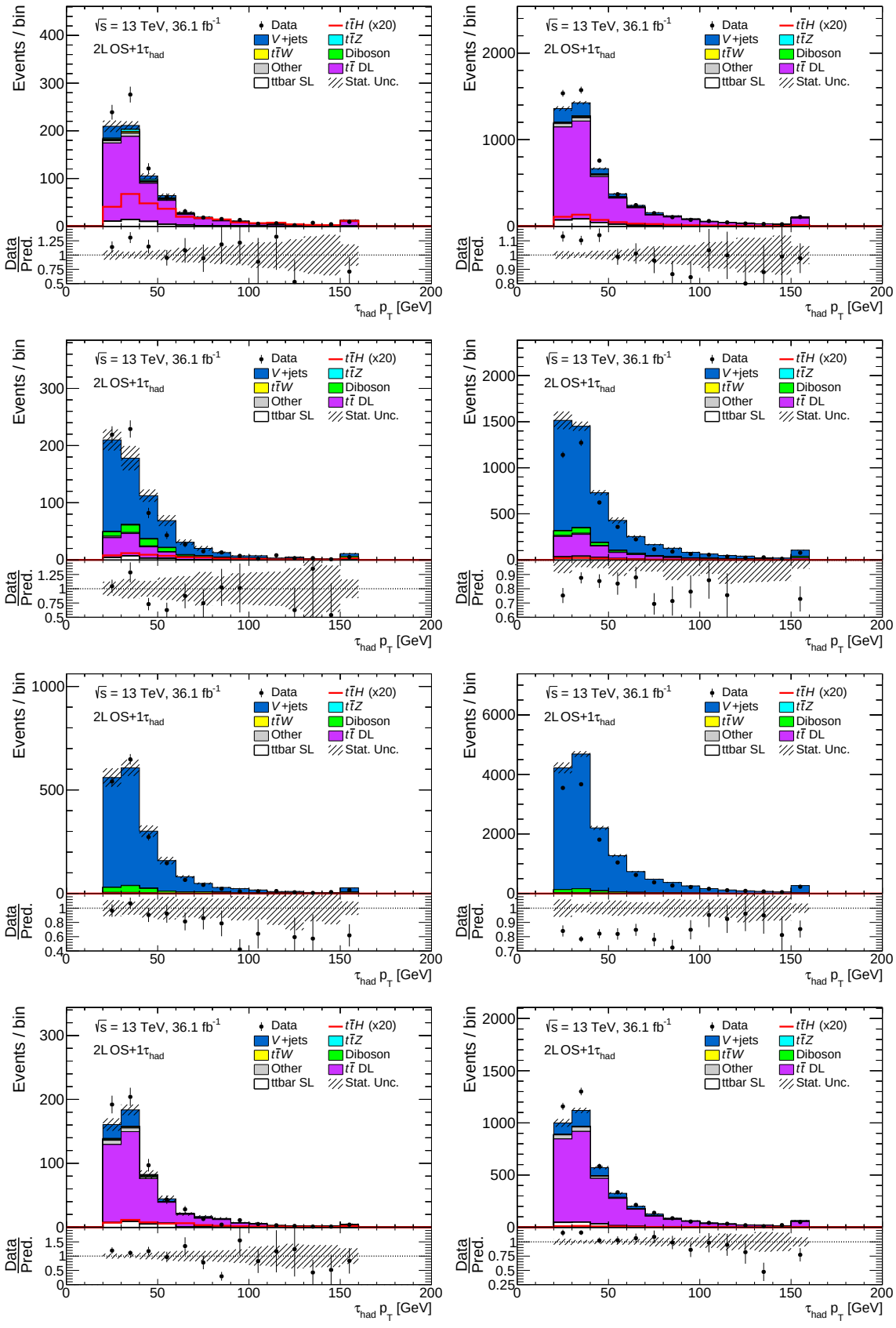


Figure A.1: Transverse momentum of τ_{had}. Left: numerator of FF. Right: denominator of FF.
1st row: Signal region.
2nd row: NJet≥ 3, NBTag= 0, Z Veto.
3rd row: NJet≥ 3, NBTag= 0, Z Selection.
4th row: NJet= 2, NBTag= 1, Z Veto.

APPENDIX B

MC plots of BDT inputs in signal region

Figures B.1 to B.3 show all input observables to the BDT used to reject $t\bar{t}$ background in $2\ell(\mathrm{OS})1\tau_{\mathrm{had}}$ (see Sec. 5.4).

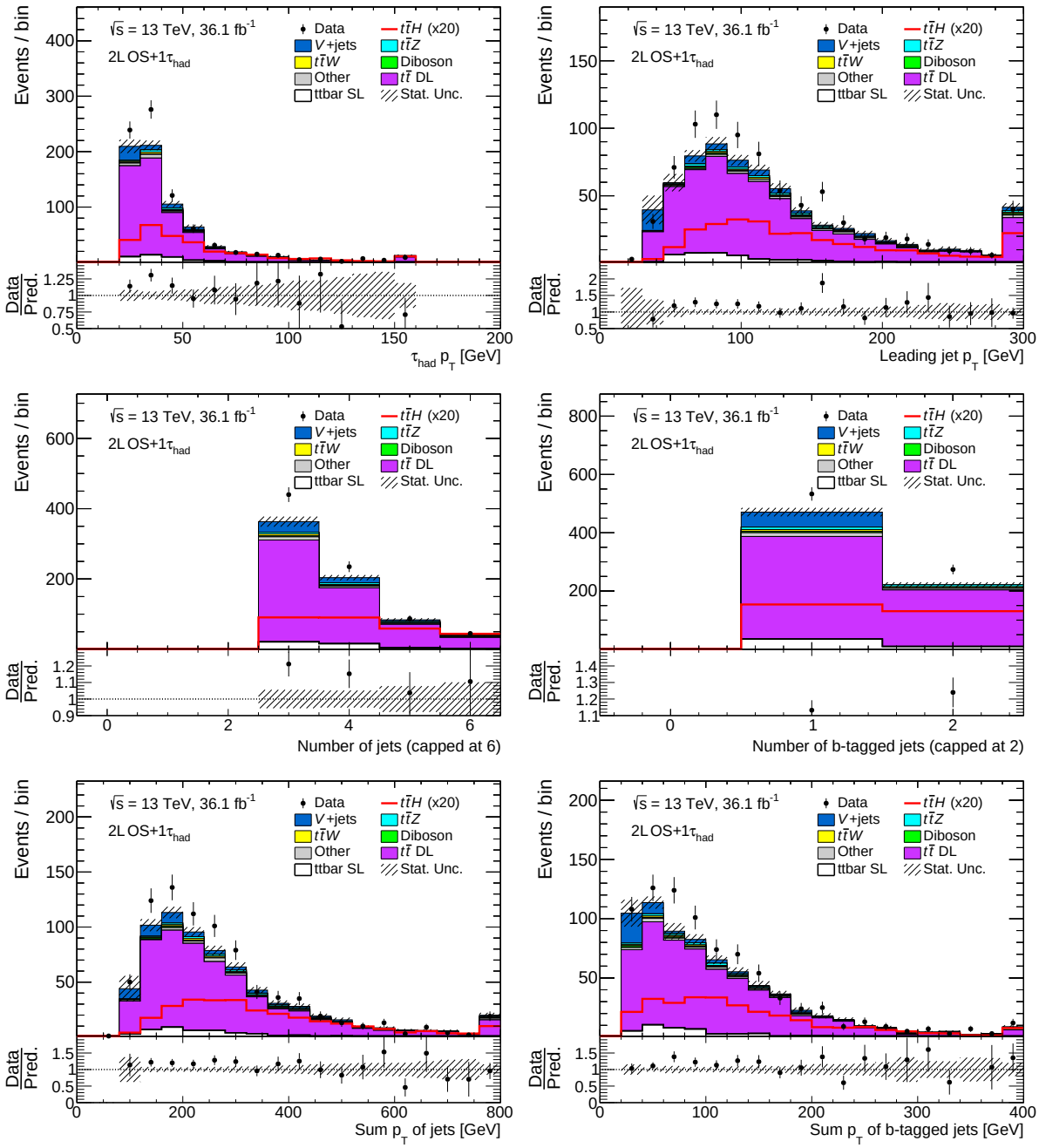


Figure B.1: BDT input observables. Top row: Leading tau p_{T}, Leading jet p_{T}. Middle row: Number of jets, Number of b–tagged jets. Bottom row: Scalar sum of jet p_{T}. Scalar sum of b–tagged jet p_{T}.

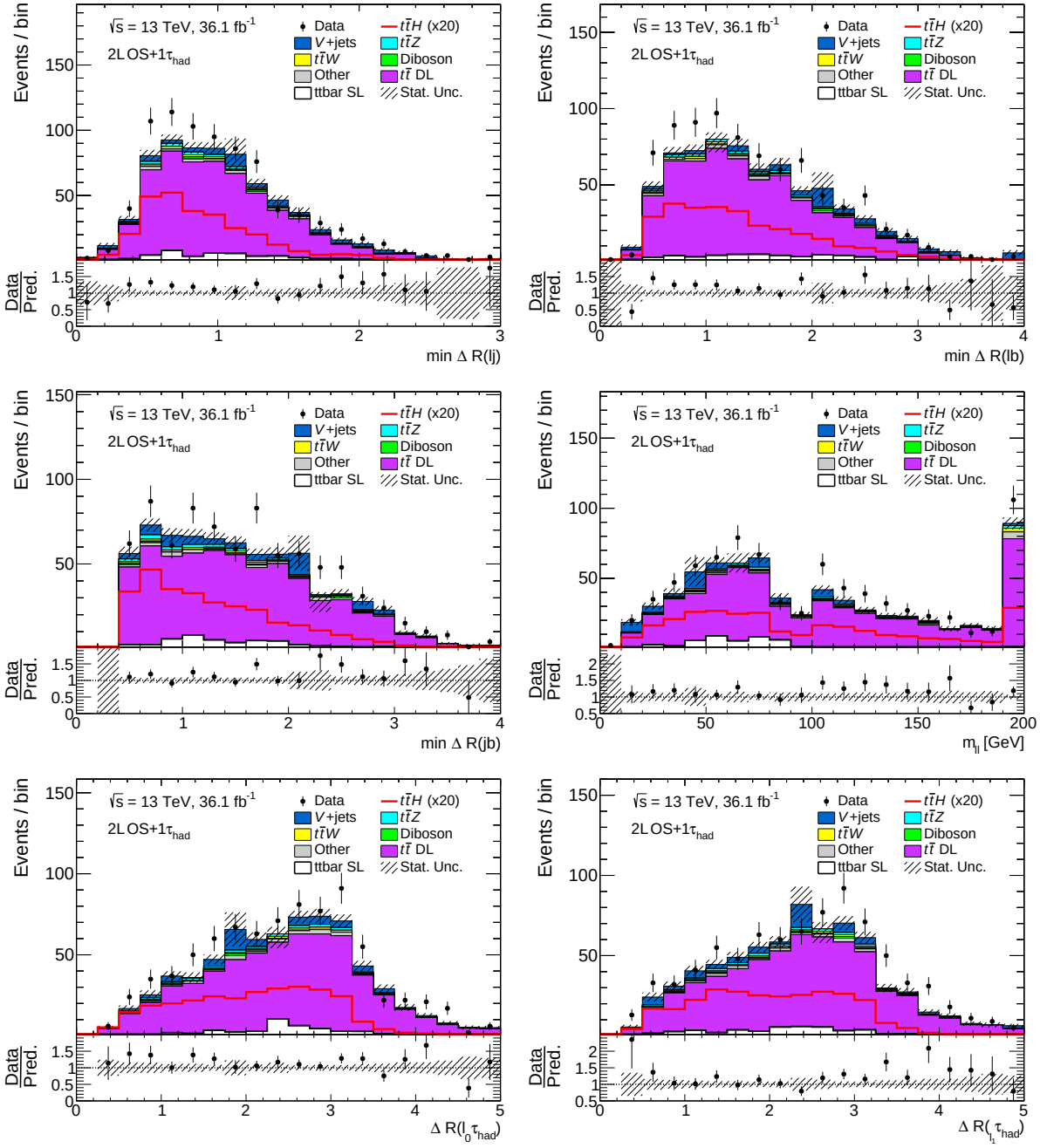


Figure B.2: BDT input observables. Top row: Minimum ΔR between a lepton and a jet. Minimum ΔR between a lepton and a b–tagged jet. Middle row: Minimum ΔR between a jet and a b–tagged jet. Invariant mass of the dilepton system. Bottom row: ΔR between leading lepton and tau. ΔR subleading lepton and tau.

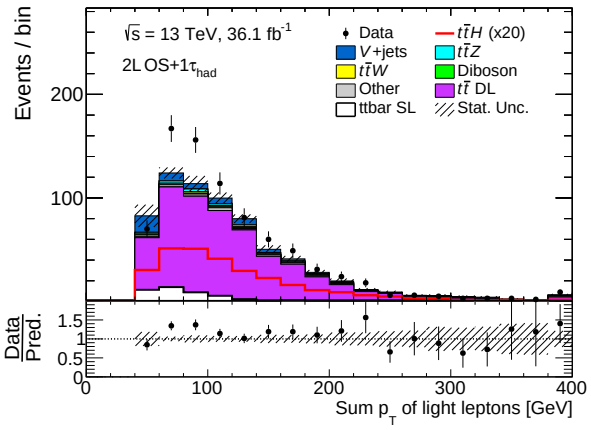


Figure B.3: BDT input observable. Scalar sum of lepton p_{T}

APPENDIX C

Cross sections of processes

Table C.1 lists the cross sections that are used to normalise the MC samples. MC samples are generated with a finite number of events N_{MC}, which corresponds to an equivalent integrated luminosity $\mathcal{L}_{\text{MC}} = \frac{N_{\text{MC}}}{\sigma}$. Each sample is then weighted with a factor $\frac{\mathcal{L}_{\text{data}}}{\mathcal{L}_{\text{MC}}}$ to normalise to the actual integrated luminosity $\mathcal{L}_{\text{data}}$ recorded by ATLAS. The equivalent luminosity of all samples which are expected to contribute in the phase space of interest should be at the very least as large as the data luminosity.

Process	Cross section in pb	QCD scale uncertainty in %	PDF+α_S uncertainty in %	Order of calculation
$t\bar{t}H$	0.51	$^{+\ 5.8}_{-\ 9.2}$	±3.6	NLO QCD+EWK
$tHqb$	0.074	$^{+\ 6.5}_{-14.7}$	±3.7	NLO QCD
tHW	0.015	$^{+\ 4.9}_{-\ 6.7}$	±6.3	NLO QCD
$t\bar{t}W$	0.60	$^{+12.9}_{-11.5}$	±3.4	NLO QCD+EWK
$t\bar{t}(Z/\gamma^* \to \ell\ell)$	0.12	$^{+\ 9.6}_{-11.3}$	±4.0	NLO QCD+EWK
$t\bar{t}t\bar{t}$	0.0092	$^{+30.8}_{-25.6}$	$^{+5.5}_{-5.9}$	NLO QCD
$t\bar{t}W^+W^-$	0.0099	$^{+10.9}_{-11.8}$	±2.1	NLO QCD
$t\bar{t}$	832	$^{+\ 2.4}_{-\ 3.5}$	±4.2	NNLO QCD + NNLL
$t\bar{t}\gamma$	5.7	±50		NLO QCD
tZ	0.61	±50		LO QCD
tWZ	0.16	±50		NLO QCD
s-channel single top	10	± 4		NLO QCD
t-channel single top	217	± 4		NLO QCD
Wt single top	72	± 5		NLO QCD + NNLL
$VV(\to \ell\ell XX)$	37	±50		NLO QCD
$Z \to \ell^+\ell^-$	2 070	± 5		NNLO QCD

Table C.1: Cross sections and their associated uncertainties that are used to normalised the MC samples. Uncertainties due to QCD and PDF+α_S scales are given. The processes $t\bar{t}\gamma$, tZ, tWZ, and $VV(\to \ell\ell XX)$ are given an inclusive uncertainty including extrapolation to the analysis phase space [5].

APPENDIX D

Systematic uncertainties of fake τ_{had} estimate in $2\ell(\text{OS})1\tau_{\text{had}}$ channel

Figure D.1 shows the effect of varying the NPs related to the statistical uncertainties of the individual τ_{had} p_{T} bins of the fake factor.

Figure D.2 shows the effect of varying the NPs related to different extraction regions of the fake factor, as well as the effect of $t\bar{t}H$ and truth-matched background subtraction in the side-bands.

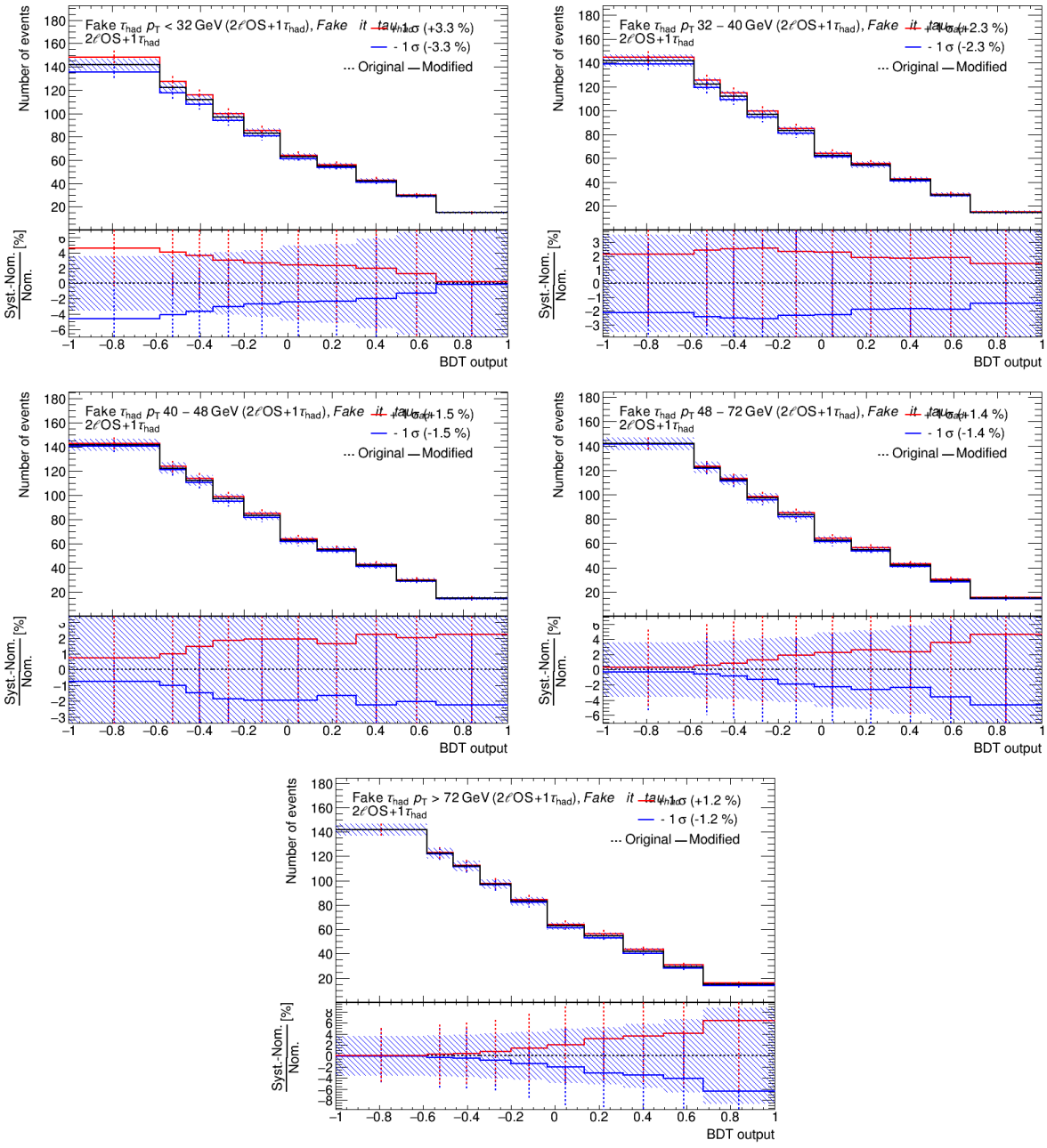


Figure D.1: Shape variations of BDT in $2\ell(OS)1\tau_{had}$ channel due to FF uncertainties.
1st-3rd row: Statistical uncertainties of the p_T bins of the fake factors.

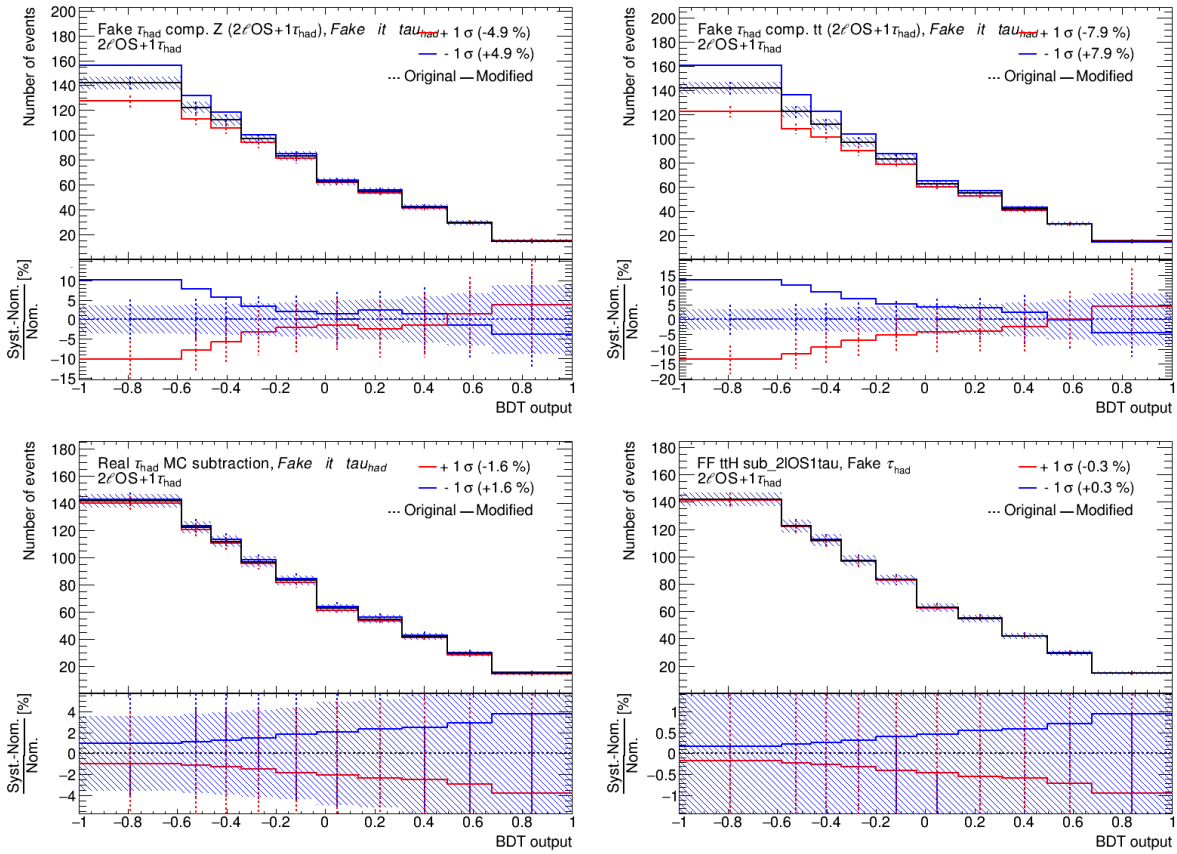
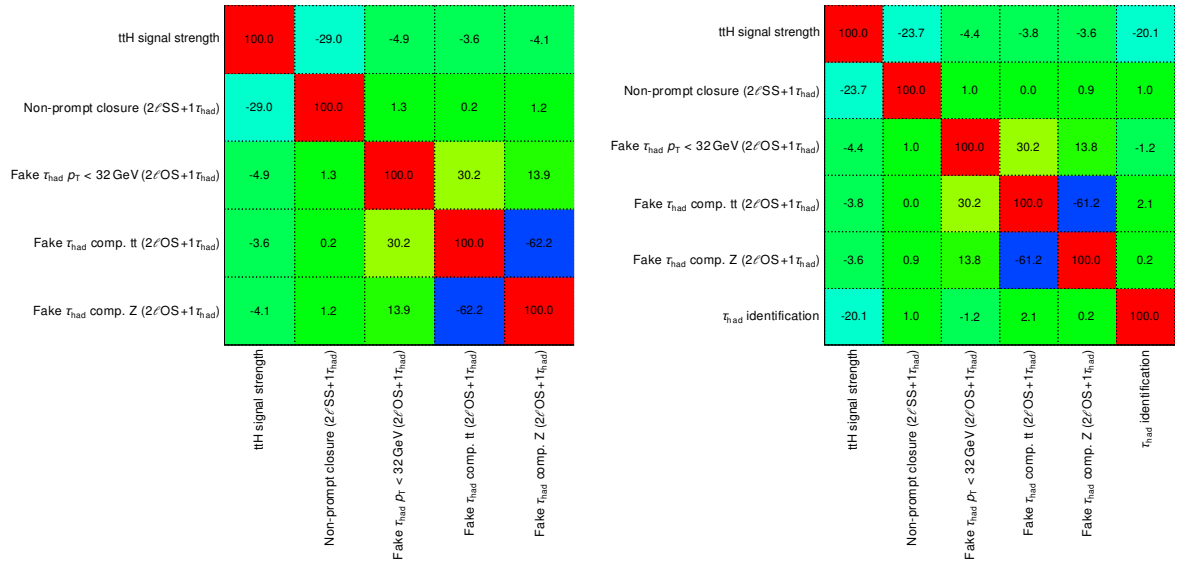


Figure D.2: Shape variations of BDT in $2\ell(\mathrm{OS})1\tau_{\mathrm{had}}$ channel due to FF uncertainties.
1st row (left): FF from Z enriched.
1st row (right): FF from $t\bar{t}$ enriched.
2nd row: Effect on FF from subtraction of real τ_{had} and $t\bar{t}H$ processes in ERs.

APPENDIX E

Correlations of estimated parameters in fit of dilepton+τ_{had} channels

Figure E.1 shows the correlation matrices of the maximum likelihood fit of the combined dilepton+τ_{had} channels with Asimov and observed data.



(a) (b)

Figure E.1: Expected (a) and observed (b) correlations between nuisance parameters and the signal strength. Only correlations that are larger than 20 % are shown.

List of Figures

List of Tables

Glossary

area under the ROC curve also known as AUC or ROC integral is a summary of the classifier performance for all possible thresholds. While in a real use case a specific threshold is chosen or a binned distribution is used, the AUC is still a good indicator for general performance. 71

boosted decision tree is a machine learning algorithm used for classification of events into signal- and background-like categories. See also Sec. 5.4. 39, 135

fake factor is the ratio of events that have a τ_{had} that passes the identification criterion to the ones in which it fails. 60, 135

ROC curve is a graphic that summarises the performance of a classifier. It shows the background rejection (1−efficiency) as a function of the signal efficiency. The better a classifier is the more towards the top right corner the ROC curve will be. 71

Acronyms

FF fake factor. 60, 63–66, 68, 135, *Glossary:* fake factor

BDT boosted decision tree. 39, 69–72, 135, *Glossary:* boosted decision tree

Acknowledgements

First and foremost, I would like to thank my thesis advisor Norbert Wermes for the opportunities and support given, as well as Jürgen Kroseberg for being the best supervisor and mentor that I could have hoped for.

Thanks also go to the members of my examination committee: Prof. Dr. Ian Brock, Prof. Dr. Drees and Prof. Dr. Litt.

I am also grateful to my parents and grandparents who brought me into this world, fostered my intellectual curiosity, pushed me to achieve, believed in my abilities and gave me confidence. I thank my siblings for always being there for me when needed.

Along the way of the journey that this PhD and thesis was, I met many people, collaborated with colleagues in Bonn and around the world, made new friends. I regard these all of these highly. I also appreciate the people that motivated me during my undergraduate studies to pursue this academic path in the first place.

I cannot imagine having done any of the past couple of years without my friend Elisabeth. Thanks also to Birgit for proof reading words and emotional support. I want to mention Irina and Leo in the same sentence, whose company I enjoy and have enjoyed, separately. Thanks to the Italians for adopting me when necessary, with special mention to Giulia and Andrea. My heart warms when I think about the people I met in Norway. I have fond memories of DnD and party nights with Christian and Vince.

Thanks, Richenel and Nicole, for witnessing the culmination of years worth of work with me.