

**Cross section measurements
of the Higgs boson exploiting
the $H \rightarrow \tau\tau$ decay mode
and their combination
with other decay modes
using the ATLAS detector**

Fabian Becherer



**Cross section measurements
of the Higgs boson exploiting
the $H \rightarrow \tau\tau$ decay mode
and their combination
with other decay modes
using the ATLAS detector**

DISSERTATION

zur Erkrangung des Doktorgrades der
Fakultät für Mathematik und Physik der
ALBERT-LUDWIGS-UNIVERSITÄT
Freiburg im Breisgau

vorgelegt von
Fabian Becherer
January 2022

Dekan: Prof. Dr. Michael Thoss
Betreuer der Arbeit: Prof. Dr. Markus Schumacher
Koreferent: Apl. Prof. Dr. Horst Fischer
Prüfer: Prof. Dr. Karl Jakobs
JProf. Dr. Stefan Vogel
Prof. Dr. Markus Schumacher

Datum der mündlichen Prüfung: 06. April 2022

CONTENTS

1	Introduction	1
2	The Standard Model	5
2.1	Elementary particles	5
2.2	Interactions in the Standard Model and gauge invariance	6
2.2.1	Quantum Electrodynamics	7
2.2.2	Quantum Chromodynamics	7
2.2.3	Electroweak Unification	9
2.3	Mass generation and Higgs mechanism	11
2.4	Limitations of the Standard Model	14
2.5	Extension of the Standard Model	15
3	Measurements of the Higgs boson properties	17
3.1	Observation	17
3.2	Mass	17
3.3	Cross sections	18
3.3.1	Inclusive cross sections	18
3.3.2	Simplified template cross sections	19
3.3.3	Differential cross sections	21
4	Experiment setup	23
4.1	The Large Hadron Collider	23
4.2	The ATLAS experiment	24
4.2.1	Inner detector	25
4.2.2	Calorimeter system	26
4.2.3	Muon system	27
4.2.4	Trigger	28
4.3	Simulation process of pp collisions	30
4.3.1	Partonic cross section and parton shower	31
4.3.2	Hadronization and hadron decays	32
4.3.3	Pile-up	33
4.3.4	Detector simulation	33
4.4	Data set	33
4.5	Simulated events	35
5	Reconstruction and identification of physics objects	39
5.1	Tracks and Vertices	39
5.2	Electrons	40
5.3	Muons	42
5.4	Jets	45
5.4.1	B-jet tagging	46

5.5	Hadronically decaying τ -leptons	48
5.6	Missing transverse energy	51
5.7	Invariant mass of di- τ system	52
6	Physics processes	57
6.1	Higgs boson production and decay	58
6.1.1	Higgs boson production modes	58
6.1.2	Higgs boson decay modes	59
6.1.3	The $\tau\tau$ -decay mode	59
6.2	W/Z boson production in association with jets	60
6.3	Di-boson production	61
6.4	QCD multi-jet production	61
6.5	Single top quark and top quark pair production	62
7	Experimental methods	63
7.1	Simplified Template Cross Section	63
7.1.1	Object definition	64
7.1.2	STXS Stage 0	65
7.1.3	STXS Stage 1	65
7.1.4	STXS Stage 1.2	68
7.2	Likelihood function	72
7.2.1	Best-fit values and uncertainties	75
7.2.2	Impact of individual uncertainties	75
8	τ-lepton trigger measurement	77
8.1	Signal and background processes	77
8.2	Event selection	77
8.3	Analysis strategy	78
8.3.1	Scale factor calculation	79
8.3.2	Background estimation	81
8.4	Results	88
8.5	Combination	92
9	$H \rightarrow \tau\tau$ analysis with the 2015 and 2016 data set	95
9.1	Signal and background processes	95
9.2	Event selection	95
9.2.1	Trigger requirement	95
9.2.2	Preselection	97
9.2.3	Signal categories	99
9.3	Background estimation	105
9.4	Systematic uncertainties	113
9.5	Statistical model	117
9.6	Signal strength measurement	122
9.7	Inclusive cross section measurements	122

9.7.1	Higgs boson production cross section	122
9.7.2	ggF and VBF Higgs boson production cross section	124
9.8	STXS measurement	126
10	$H \rightarrow \tau\tau$ analysis with the 2015-2018 data set	129
10.1	Signal and background processes	129
10.2	Event selection	129
10.2.1	Trigger selection	129
10.2.2	Preselection	130
10.2.3	Inclusive signal categories	130
10.2.4	Final signal categories	134
10.3	Background estimation	146
10.4	Systematic uncertainties	153
10.4.1	Theoretical uncertainties on signal	153
10.4.2	Theoretical uncertainties on background	155
10.4.3	Experimental uncertainties	156
10.5	Statistical model	158
10.5.1	Parameters of interest and normalization factors	158
10.5.2	Input observable and likelihood function for the fit	160
10.5.3	Modifications of the likelihood function	169
10.6	Results	174
10.6.1	Observed normalization factors	174
10.6.2	Behavior of nuisance parameters in the fit	178
10.6.3	Inclusive cross section measurements	186
10.6.4	STXS measurement	193
11	Combined Higgs boson measurements	203
11.1	Strategy of the combination	203
11.2	Input for the combination	204
11.3	Signal strength	205
11.4	Production cross sections	206
11.5	Product of production cross sections and branching ratios	209
11.6	STXS measurement	214
11.7	SMEFT interpretation	221
12	Extrapolation of the $H \rightarrow \tau\tau$ analysis to the data set expected at HL-LHC	229
12.1	Extrapolation strategy	229
12.1.1	Extrapolation of nominal prediction	229
12.1.2	Extrapolation of systematic uncertainties	230
12.2	Results	232
12.3	Limitations of the extrapolation	236
13	Conclusion	237

A Appendix	242
References	256

1 INTRODUCTION

Particle physics studies the fundamental constituents of nature and describes their properties and interactions. The Standard Model (SM) developed in the 1960s is the precisest description of the microcosm. It is able to describe nearly all observations in previous and current experiments. Matter is build up in the SM by half-integer spin particles, so called fermions interacting with each other by exchanging integer spin particles, so called bosons. Three out of the four fundamental interactions are included in the SM. The electromagnetic, weak and strong interaction are described with relativistic quantum field theories, while gravity is excluded. Gravity is at energy scales considered in particle physics today negligible compared to the remaining interactions.

The SM was able to describe precisely all known particles at the time of its formulation and predicted several new experimentally undiscovered particles. All of these particles have been discovered in measurement until today, such as the electroweak bosons, the top quark, the ν_τ -neutrino and a Higgs boson discovered between 1983 and 2012 [1–9]. Out of them the Higgs boson was the last discovered particle observed at CERN near Geneva in 2012. CERN was founded in 1949 as the European council for nuclear research (Conseil Européen pour la Recherche Nucléaire) and has evolved to the largest laboratory for particle physics on earth. The Large Hadron Collider (LHC) [10] based at CERN is designed to collide protons with an center-of-mass energy up to 14 TeV. Two independent analyses were performed on the two data sets of the multipurpose detectors ATLAS [11] and CMS [12] recorded in the years 2011 and 2012 at a center-of-mass energy of 7 and 8 TeV. Both analyses observed an Higgs boson with an significance of 5.9σ and 5.0σ for the ATLAS and CMS experiments [8, 9]. With these measurements the Higgs boson introduced in the SM by the Englert-Brout-Higgs-Guralnik-Hagen-Kibble mechanism, which describes the mass terms for fermions and electroweak bosons [13–18], was discovered nearly fifty years after its prediction. A combined analysis measured the Higgs boson mass m_H , which is a unknown parameter in the SM, to $m_H = (125.09 \pm 0.21(\text{stat.}) \pm 0.11(\text{sys.}))$ GeV [19]. Several subsequent analyses as measuring the coupling strength in different production and decay modes [20] or the spin of the Higgs boson [21, 22] have confirmed that the Higgs boson behaves SM-like within current uncertainties. Peter Higgs and Francois Englert were honored with the Nobel Prize for the formulation of the underlying mechanism of Higgs boson, after it was observed at CERN.

Observations such as dark matter [23–25] have shown the limitations of the SM, while at the same time no deviations from the SM predictions are observed in the measurements at the LHC. To improve the search for new particles and to close in on beyond the SM models (BSM), it is important to measure the SM processes and parameters as precisely as possible. In the years 2015 to 2018 new data sets were recorded by the experiments at the LHC with a higher center-of-mass energy of 13 TeV and a combined integrated luminosity of 139 fb^{-1} . The increased data statistics and the higher center-of-mass energy leads to a higher amount of observed Higgs bosons in the data set. This increases in particular the precision of measurements limited by statistical uncertainties, as the analysis of the $H \rightarrow \tau\tau$ decay channel, which was able to measure an evidence of the process with 4.5σ at the ATLAS experiment with the previous data set [26], but missed the observations requiring $>5\sigma$.

New analysis techniques as the Simplified Template Cross Sections (STXS) [27], which measures cross sections of exclusive regions of phase space of the Higgs boson production, can be used to explore the Higgs boson sector with finer granularity. This reduces the dependence on theoretical predictions of the signal compared to more inclusive measurements. The STXS framework allows to combine analyses of different decay channels and between the two experiments ATLAS and CMS. In this way, the full information of the recorded data can be used for combined measurements of Higgs boson production cross sections with fine granularity. The measured cross sections can be used to probe BSM models or to find hints for new physics using methods as the effective field theory of the SM (SMEFT), which extends the SM by including operator of higher mass dimensions.

The highest sensitivity to measure the Yukawa coupling of the Higgs boson at the LHC is given in the $H \rightarrow \tau\tau$ decay mode. The properties of the decay mode provide a high sensitivity to the Higgs boson production via vector boson fusion (VBF) and to Higgs boson production with high transverse momentum. These properties make the $H \rightarrow \tau\tau$ decay channel an important decay mode for analyses based on the data set recorded between 2015 and 2018. In Chapter 2 an introduction to the SM focusing on the Englert-Brout-Higgs-Guralnik-Hagen-Kibble mechanism, the Higgs boson and SMEFT is given. An overview of current and previous measurements of the Higgs boson is given in Chapter 3, while Chapter 4 describes the ATLAS experiment, the used data sets and simulated events. In Chapter 5 the reconstruction and identification methods of physics objects used in the measurements are discussed, followed by the description of all relevant physics processes for this thesis in Chapter 6. Chapter 7 describes the experimental methods used in the analyses as the STXS framework and likelihood models.

The thesis focuses on cross section measurements in the $H \rightarrow \tau\tau$ decay channel performed with data recorded between 2015 and 2018 by the ATLAS experiment, especially based on the STXS framework. Two different analyses are performed on this data set. The first analysis uses the data from 2015 to 2016 and is described in Chapter 9. It focuses on the discovery of the $H \rightarrow \tau\tau$ decay mode with the ATLAS experiment, while the second analysis described in Chapter 10 uses the full data set and is optimized for the measurement within the STXS framework. The $H \rightarrow \tau\tau$ analyses rely on the τ_h -trigger of the ATLAS detector to collect events of interest. The application of trigger requirements in data and simulated events can lead to difference between data and simulated events. These deviations are corrected in simulation by applying so called scale factors, which are measured in an auxiliary measurement based on a tag-and-probe analysis with $t\bar{t} \rightarrow \tau_\mu \tau_h$ events summarized in Chapter 8. The $H \rightarrow \tau\tau$ measurements are combined with analyses of other decay modes. These combinations are discussed in Chapter 11 focusing on STXS measurements and their interpretations within the SMEFT framework. Finally, the $H \rightarrow \tau\tau$ analysis based on the 2015 to 2018 data set is extrapolated in Chapter 12 to the expected data set of the high luminosity LHC (HL-LHC) corresponding to a integrated luminosity of 3000 fb^{-1} . This extrapolation provides an estimation of the possible sensitivity with the full data set recorded by the ATLAS detector, but also allows to study future limitations of the analysis. These studies help to develop new strategies to improve the analysis based on the upcoming data sets.

Personal contributions All analyses presented in this theses are performed within the ATLAS collaboration and based on a joint effort of about 3000 members. The two $H \rightarrow \tau\tau$ analyses described in Chapters 9 and 10 are published in Reference [26] and [28], while the τ -trigger analysis presented in Chapter 8 has not been published until now. The combined Higgs boson cross section measurements described in Chapter 11 are published in References [29–31]. The extrapolation of the $H \rightarrow \tau\tau$ measurement described in Chapter 12 will be published by ATLAS in 2022. The following list summarizes the personal contributions to the individual analyses by the author of the thesis:

- τ -trigger tag-and-probe analysis based on the $t\bar{t} \rightarrow \tau_\mu \tau_h$ events (Chapter 8)
 - Development and validation of complete analysis strategy and infrastructure
 - Preparation of simulated events and data
 - Development of event selection, signal region optimization, background estimation, derivation and treatment of systematic uncertainties
 - Estimation of efficiencies and scale factors
 - Preparation of the analysis and support for the combination with the tag-and-probe analysis based on $Z \rightarrow \tau_\mu \tau_h$ events
- $H \rightarrow \tau\tau$ analysis based on the 2015-2016 data set (Chapter 9)
 - Development of a data-driven method to estimate the misidentified τ background in the $\tau_1 \tau_h$ final state and the related systematic uncertainties
 - Implementation of the STXS framework, adjustment of theoretical uncertainties on the signal prediction for the STXS measurement and development of the fitting strategy and model
 - Evaluation of fit results and fit stability studies
- $H \rightarrow \tau\tau$ analysis based on the 2015-2018 data set (Chapter 10)
 - Optimization studies for the event selection compared to the 2015-2016 analysis toward better STXS measurements
 - Fit studies to estimate the BDT score cuts in the event selection
 - Implementation, development and validation of fitting strategies, including PoIs, NFs and NPs
 - Development and validation of correction methods for systematic uncertainties as pruning, symmetrization and smoothing
 - Determination and validation of all cross section measurements and post-fit values for NFs and NPs
- Combined Higgs boson measurements (Section 11)
 - Implementation of modification in the likelihood function of the $H \rightarrow \tau\tau$ decay channel to prepare and harmonize it for the combination

-
- Integration of the $H \rightarrow \tau\tau$ decay channel in the combined model
 - Validation studies of modified likelihood models and consistency checks
 - $H \rightarrow \tau\tau$ extrapolation analysis (Section 12)
 - Development of complete analysis strategy and technical infrastructure
 - Implementation of extrapolation for nominal prediction
 - Implementation of extrapolation for systematic uncertainties and their optimization for the $H \rightarrow \tau\tau$ decay channel
 - Estimation and validation of results including PoIs, NFs and NPs

2 THE STANDARD MODEL

This chapter describes the underlying theoretical background of this thesis. First the general concepts of the Standard Model (SM) [32–34], the elementary particles and their interactions are introduced. In addition, the description of masses of the elementary particles by the Englert-Brout-Higgs-Guralnik-Hagen-Kibble mechanism (BEGGHK mechanism) [13–18] is explained.

The SM of particle physics was developed in the 1960s and 1970s and consists of renormalizable quantum field theory (QFT). Internal symmetries of a unitary group ensure that the QFT of the SM is invariant under local gauge transformations. The model describes all known particles and their fundamental interactions at its formulation, but also has successfully predicted several new particles discovered in following experiments. For example the weak and strong gauge bosons as well as the top quark were discovered after their prediction in different experimental measurements between 1983 and 1995 [1–6]. The SM gives a theoretical description of three out of the four known fundamental interactions. Gravity is the only interaction not included in the SM, though it is most probably negligible at the energy of current collider experiments. As a consequence of the preservation of the local gauge invariance, interactions take place via the exchange of gauge bosons in the SM. These gauge fields result in the existence of spin-1 gauge bosons. The matter surrounding us is formed by quantization of half-integer spin fermion fields. Spontaneous symmetry breaking is introduced to allow massive fermions and electroweak gauge bosons in the model. This principle predicts the additional existence of a spin-0 boson, the so-called Higgs boson, which was discovered experimentally at the LHC in 2012 [8, 9]. The following sections give a more detailed introduction into the elementary particles, their interactions and the mass generation of the SM based on [32–34] and other explicitly specified references.

2.1 Elementary particles

The elementary particles and their properties are visualized in Figure 2.1. They are grouped into fermions and bosons. Fermions have half-integer spin, while bosons carry integer spin. Fermions are split into color charged quarks (u,d), (c,s) and (t,b) and into colorless leptons (ν_e, e), (ν_μ, μ) and (ν_τ, τ), which are organized in three families or generations. The matter surrounding us consists of the up quark, down quark and the electron, which together with the electron-neutrino build the first generation. The second and third generations are copies of the first with an increased mass and a different flavor quantum number. All other quantum numbers of the higher generations are equal to the ones of the corresponding particle of the first generation. Particles of higher families only appear in collider experiments, in cosmic rays or in the early universes with temperatures far above the particle masses. Every family consists of one lepton with electric charge $q=-1$, one uncharged neutrino with the same flavor, one up-type quark with an electric charge of $q = 2/3$ and one down-type quark with an electric charge of $q = -1/3$. Every elementary particle has a corresponding anti-particle, which has an opposite charge while the spin, mass, lifetime and strength of interaction are equal.

The strong, the weak and the electromagnetic interaction are described by renormalizable

Fermion generations				
	I	II	III	
mass→ charge→ spin→	2.3 MeV 2/3 1/2 u up	1.3 GeV 2/3 1/2 c charm	173 GeV 2/3 1/2 t top	0 0 1 γ photon
				125 GeV 0 0 H Higgs boson
Quarks	4.8 MeV -1/3 1/2 d down	95 MeV -1/3 1/2 s strange	4.2 GeV -1/3 1/2 b bottom	0 0 1 g gluon
	<2 eV 0 1/2 ν_e electron-neutrino	<0.2 MeV 0 1/2 ν_μ muon-neutrino	<18 MeV 0 1/2 ν_τ tau-neutrino	91.2 GeV 0 1 Z^0 Z boson
	0.5 MeV -1 1/2 e electron	106 MeV -1 1/2 μ muon	1.8 GeV -1 1/2 τ tau	80.4 GeV ± 1 1 $W^\pm$ W boson
Leptons				Gauge bosons

Figure 2.1: Elementary particles and their properties in the Standard Model [35].

gauge field theories in the SM resulting in gauge bosons with spin $s = 1$. The massless photon is the exchange particle of the electromagnetic interaction, which was the first well understood interaction. It couples to the electromagnetic charge of a particle, so that all fermions except for the neutrinos participate in it. Quantum chromodynamics (QCD) [36–38] describes the strong interaction and is transmitted by gluons. The weak interaction has three exchange particles the charged $W^\pm$ bosons and the neutral Z^0 boson and acts on all fermions. The exchange particles of the weak interaction are massive in contrast to the other exchange particle resulting in a short distance of the weak interaction. The $W^\pm$ bosons couples only to left-handed particles and right-handed anti-particles, while the Z^0 boson couples to all fermions.

The masses of fermions and weak gauge bosons are explained by spontaneous symmetry breaking the only spin-0 field, the so called Higgs field, with the associated Higgs boson.

2.2 Interactions in the Standard Model and gauge invariance

Interaction are described by local gauge symmetries in the SM. The Lagrangian density is required to be invariant under a group of global and local transformations. These symmetries constrain the underlying Lagrangian and contribute to the construction of the theories.

2.2.1 Quantum Electrodynamics

The theory of quantum electrodynamics (QED) [39–41] based on an underlying $U(1)_Q$ symmetry with the electric charge Q , describes the electromagnetic interaction within the SM. The photon is the exchange particle of this relativistic quantum field theory and couples to Q . The Lagrangian density $\mathcal{L}$ of a free fermion with the mass m_f is given by:

$$\mathcal{L} = i\bar{\Psi}\gamma_\mu\partial^\mu\Psi - \bar{\Psi}m_f\Psi, \quad (2.1)$$

where Ψ is the Dirac spinor, γ^μ the Dirac matrices and $\partial_\mu = \frac{\partial}{\partial x^\mu}$ the partial derivative [34]. Applying the Euler-Lagrange equation leads to an equation of motion, which is called Dirac equation:

$$(i\gamma^\mu\partial_\mu - m_f)\Psi = 0. \quad (2.2)$$

The fermion fields transform under the $U(1)_Q$ symmetry according to

$$\Psi \mapsto e^{-iQ\alpha(x)}\Psi, \quad (2.3)$$

with the local phase $\alpha(x)$. Considering the space-time dependence of the local phase in equation 2.1 a new term enters the Lagrangian and breaks the gauge invariance:

$$\partial_\mu\Psi \mapsto \partial_\mu\left(e^{-iQ\alpha(x)}\Psi\right) = e^{-iQ\alpha(x)}\partial_\mu\Psi - iQe^{-iQ\alpha(x)}\partial_\mu\alpha(x)\Psi. \quad (2.4)$$

An additional interaction between the fermion field and a new vector field A_μ , with $A_\mu \mapsto A_\mu + 1/e \partial_\mu\alpha(x)$, has to be introduced to restore the local phase invariance and corresponds to the replacement of the derivatives by a covariant derivative:

$$\partial_\mu \mapsto D_\mu = \partial_\mu + iqA_\mu, \quad (2.5)$$

with the coupling strength q . q is expressed in units of the elementary electric charge and can be identified with the physical photon γ . This leads to the QED Lagrangian density:

$$\mathcal{L}_{QED} = \bar{\Psi}\left(i\gamma^\mu(\partial_\mu - iqA_\mu) - m_f\right)\Psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}, \quad (2.6)$$

with the field strength tensor $F_{\mu\nu}$. Since the introduction of any mass term for the photon of the form $(-\frac{1}{2}m^2A^\mu A_\mu)$ to the equation would violate the gauge invariance, the QED requires a massless gauge boson. This is in agreement with the measured upper limit of the photon mass of $\mu_\gamma < 3 \cdot 10^{-27}$ eV [42].

2.2.2 Quantum Chromodynamics

The strong interaction of quarks and gluons, which couples to the color charge, is described by QCD and based on a $SU(3)_C$ symmetry group. Measurements have shown that quarks appear in three different color states (red, blue, green), or the corresponding anti-color for

anti-quarks [43]. The four component Dirac spinor of the quark field $\Psi(x)$ is replaced by a vector of three spinors:

$$\Psi = (\Psi_{\text{red}}(x), \Psi_{\text{blue}}(x), \Psi_{\text{green}}(x)). \quad (2.7)$$

The local $\text{SU}(3)_C$ transformation of a free quark field $\Psi(x)$ is given by:

$$\Psi(x) \mapsto \exp\left(i\frac{g_s}{2} \sum_{a=1}^8 \lambda_a \beta_a(x)\right) \Psi(x), \quad (2.8)$$

with the eight-component gauge function $\vec{\beta}(x)$, the eight Gell-Mann matrices λ_a and the coupling strength g_s [44]. Gluons interact with themselves in contrast to photons and the $\text{SU}(3)_C$ structure is non-abelian. To preserve local gauge invariance eight gauge fields G_a^μ are defined. These fields transform under the $\text{SU}(3)_C$ as:

$$G_a^\mu \mapsto G_a^\mu + \frac{1}{g_s} \partial^\mu \beta_a(x) + f_{abc} \beta_b(x) G_c^\mu, \quad (2.9)$$

where f_{abc} refers to the structure constants of the $\text{SU}(3)_C$. Similar to the QED a covariant derivative

$$D^\mu = \partial^\mu + ig_s \sum_{a=1}^8 \frac{\lambda_a}{2} G_a^\mu \quad (2.10)$$

is introduced with the gluon field-strength tensor

$$G_a^{\mu\nu} = \partial^\mu G_a^\nu - \partial^\nu G_a^\mu - g_s f_{abc} G_b^\nu G_c^\mu. \quad (2.11)$$

The Lagrangian density of the QCD for a specific quark flavor with the mass m_q is then given by:

$$\mathcal{L} = \bar{\Psi} (i\gamma_\mu D^\mu - m_q) \Psi - \frac{1}{4} \sum_{a=1}^8 G_{a,\mu\nu} G_a^{\mu\nu}. \quad (2.12)$$

A finite interaction range for the QCD is introduced by the self-interaction of the gluons, described in the last term the equation 2.11. Equation 2.12 shows that quarks with the same flavor but different color have the same masses. In agreement with experimental observation the gluon has to be massless to ensure gauge invariance. The running of the strong coupling strength results in an increase of the strong force at lower energy scales. Therefore, the force becomes weak at very short distances and the gluons act as free particles. This behavior is called asymptotic freedom [37, 38]. On the other hand, a higher spatial separation between the color-charged particles increases the interaction potential and is called confinement. Due to the confinement, quarks do not occur as free particles, instead they are bound into hadrons. Hadrons consisting of a quark and anti-quark pair are called mesons, while hadrons consisting of three (anti-)quarks are called baryons. The LHCb experiment has recently announced the observation of additional bound states with four and five quarks, which correspond to tetraquarks and pentaquarks [45, 46].

2.2.3 Electroweak Unification

Weak interaction with and without a flavor transition for quarks are observed, caused by the weak neutral exchange particle Z^0 and the charged exchange particles $W^\pm$ bosons. All three carriers have a mass, with $m_{W^\pm} = 80.4$ GeV and $m_{Z^0} = 91.2$ GeV [47]. Measurements have observed that the weak carriers $W^\pm$ boson couple only to left-handed particles and right-handed anti-particles and thus violate parity transformation (P) and charge conjugation (C) maximally [48, 49]. This behavior inspired Glashow, Salam and Weinberg to formulate a theory of the electroweak interaction [50–52], which is able to describe the weak and the electromagnetic force by using a $SU(2)_{I_W} \times U(1)_Y$ group, with the hypercharge Y and the isospin I_W . The relation between the electric charge Q , the hypercharge and the third component of the weak isospin I_W^3 is given by the Gell-Mann-Nishihima relation [53, 54]:

$$Q = I_W^3 + Y/2. \quad (2.13)$$

Fermions in right-handed states correspond to $SU(2)_{I_W}$ singlets with $I_W = 0$ and left-handed fermions are assigned to $SU(2)_{I_W}$ doublets with $I_W = 1/2$ and $I_W^3 = \pm 1/2$. Figure 2.2 shows an overview of fermions which are included in the SM together with their quantum numbers. Even though current measurements of neutrino oscillations have indicated that at least two neutrinos have a non-zero mass [55–57], this thesis will assume massless neutrinos. Therefore, only left-handed neutrinos are included in Figure 2.2 since a right-handed massless neutrino will not interact within the SM. The Cabbibo-Kobayashi-Maskawa (CKM) matrix describes

Fermion generations				Quantum numbers			
	I	II	III	I_W	I_W^3	Y	$Q[e]$
Quarks	u_L	c_L	t_L	$1/2$	$1/2$	$1/3$	$2/3$
	d'_L	s'_L	b'_L	$1/2$	$-1/2$	$1/3$	$-1/3$
	u_R	c_R	t_R	0	0	$4/3$	$2/3$
	d_R	s_R	b_R	0	0	$-2/3$	$-1/3$
Leptons	$\nu_{e,L}$	$\nu_{\mu,L}$	$\nu_{\tau,L}$	$1/2$	$1/2$	-1	0
	e_L	μ_L	τ_L	$1/2$	$-1/2$	-1	-1
	e_R	μ_R	τ_R	0	0	-2	-1

Figure 2.2: Fermions that interact within the SM and their quantum numbers [34].

the flavor eigenstates (d', s', b') with $I_3 = -1/2$ [58, 59]:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (2.14)$$

These flavor eigenstates are a mixture of their mass eigenstates (d, s, b). $|V_{ij}|^2$ gives the probability for a quark flavor transition from i to j by the exchange of a $W^\pm$ boson. Next to C and P is the invariance under the combination of charge conjugation and parity transformation (CP) broken by the complex phase of the CKM matrix [60]. A left-handed isospin doublet transform under the $SU(2)_{I_W}$ with:

$$\Psi_L(x) \mapsto \exp\left(i\frac{g}{2}\sum_{a=1}^3\tau_a\alpha_a(x)\right)\Psi_L(x), \quad (2.15)$$

where $\alpha_a(x)$ are the local phases, τ_a the (2x2) Pauli matrices as the generators of the $SU(2)_{I_W}$ group and g the coupling strength. Under the $U(1)_Y$ symmetry the transformation of a left-handed isospin doublet (L) and a right-handed singlet (R) is given by:

$$\Psi_L(x) \mapsto \exp\left(i\frac{g'}{2}Y\beta(x)\right)\Psi_L(x) \quad \text{and} \quad \Psi_R(x) \mapsto \exp\left(i\frac{g'}{2}Y\beta(x)\right)\Psi_R(x), \quad (2.16)$$

with the local phase $\beta(x)$, the hypercharge Y and the coupling strength g' . To restore the local gauge invariance three vector fields W_a^μ with $a=(1,2,3)$ for $SU(2)_{I_W}$ and one gauge field B for $U(1)_Y$ are introduced. The covariant derivatives are defined as:

$$D_L^\mu = \partial^\mu + i\frac{g}{2}\sum_{a=1}^3\tau_a W_a^\mu + i\frac{g'}{2}YB_\mu \quad \text{and} \quad D_R^\mu = \partial^\mu + i\frac{g'}{2}YB_\mu \quad (2.17)$$

and the electroweak Lagrangian density is given by:

$$\mathcal{L} = \bar{\Psi}_L(i\gamma_\mu D_L^\mu)\Psi_L + \bar{\Psi}_R(i\gamma_\mu D_R^\mu)\Psi_R - \frac{1}{4}\sum_{a=1}^3 W_{a,\mu\nu}W_a^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B_a^{\mu\nu}. \quad (2.18)$$

The field strength tensors are defined as:

$$W_a^{\mu\nu} = \partial^\mu W_a^\nu - \partial^\nu W_a^\mu - g\epsilon_{abc}W_b^\mu W_c^\nu \quad \text{and} \quad B^{\mu\nu} = \partial^\mu B^\nu - \partial^\nu B^\mu, \quad (2.19)$$

with the structure constant ϵ_{abc} of the non-abelian $SU(2)_{I_W}$ group. By introducing the weak mixing angle:

$$\cos(\Theta_W) = \frac{g}{\sqrt{g^2 + g'^2}} \quad (2.20)$$

the physical photon field A^μ and the physical boson field Z^μ can be constructed as linear combinations of the neutral $SU(2)_{I_W} \times U(1)_Y$ gauge fields:

$$\begin{pmatrix} Z^\mu \\ A^\mu \end{pmatrix} = \begin{pmatrix} \cos(\Theta_W) & -\sin(\Theta_W) \\ \sin(\Theta_W) & \cos(\Theta_W) \end{pmatrix} \begin{pmatrix} W_3^\mu \\ B^\mu \end{pmatrix} \quad (2.21)$$

and the charge eigenstates for the $W^\pm$ are given by:

$$W_\mu^\pm = \sqrt{1/2} (W_\mu^1 \mp iW_\mu^2). \quad (2.22)$$

Since a mass terms for electroweak gauge bosons in equation 2.18 would violate the gauge invariance, the electroweak theory requires massless fermions and massless electroweak gauge bosons. This is in contrast to experimental observation, which confirms massive electroweak gauge bosons and fermions [47]. To include this observation into the theory, the BEGGHK mechanism is used, which introduces an additional scalar field in the context of spontaneous symmetry breaking described in the next Section 2.3.

2.3 Mass generation and Higgs mechanism

The BEGGHK mechanism developed in the 1960s in References [13–18] explains how massive fermions and massive weak gauge bosons can be described in the SM. The method introduces mass terms by spontaneous symmetry breaking. This ensures the invariance of field theories under local gauge transformations. Hence, the SM does not break the electroweak symmetry explicitly, instead it breaks the symmetry spontaneously via the vacuum state of a doublet of complex scalar fields with four degrees of freedom in the Lagrangian. This so-called Higgs field is defined with the hypercharge $Y = 1$ and the isospin $I_W = 1/2$ by [44]:

$$\Phi = \begin{pmatrix} \Phi^+ \\ \Phi^- \end{pmatrix} = \begin{pmatrix} \Phi_1 + i\Phi_2 \\ \Phi_3 + i\Phi_4 \end{pmatrix}. \quad (2.23)$$

Using the covariant derivative in equation 2.17 a Lagrangian density for the Higgs field can be constructed, which is invariant under the $SU(2)_{I_W} \times U(1)_Y$ group:

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi) \text{ with} \quad (2.24)$$

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2, \quad \lambda > 0 \text{ and} \quad (2.25)$$

$$D_\mu \Phi = \left(\partial_\mu - \frac{1}{2} ig \vec{\sigma} \vec{W} - \frac{1}{2} ig' Y B_\mu \right) \Phi, \quad (2.26)$$

where the Higgs potential $V(\Phi)$ is the most general potential, that is invariant under the local $SU(2)_{I_W} \times U(1)_Y$ transformation and provides renormalizability. The minimum of $V(\Phi)$ corresponds to a ground state of $\Phi_0 = 0$ for $\mu^2 > 0$, while $\mu^2 < 0$ results in a degenerated set of non-zero ground states:

$$\Phi_0^\dagger \Phi_0 = \frac{-\mu^2}{2\lambda} = \frac{v^2}{2}, \quad (2.27)$$

with $v = \sqrt{-\mu^2/\lambda}$ defined as the vacuum expectation value of the Higgs field. A simplified model with two instead of four degrees of freedom of the Higgs potential is shown in Figure 2.3. The figure illustrates that at any fixed point in the potential minimum, no rotation

symmetry is given and the choice of one of these infinite points refers to a spontaneous symmetry breaking. Choosing the non-vanishing ground state

$$\Phi_0(x) = \sqrt{1/2} \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad Y = 1, I_3 = -1/2 \quad (2.28)$$

breaks spontaneously the symmetry of $SU(2)_{I_W} \times U(1)_Y$ to a $U(1)_Y$ symmetry, while the choice of $\Phi_1 = \Phi_2 = \Phi_4 = 0$ preserves the $U(1)_Q$ symmetry resulting in massless photons. The

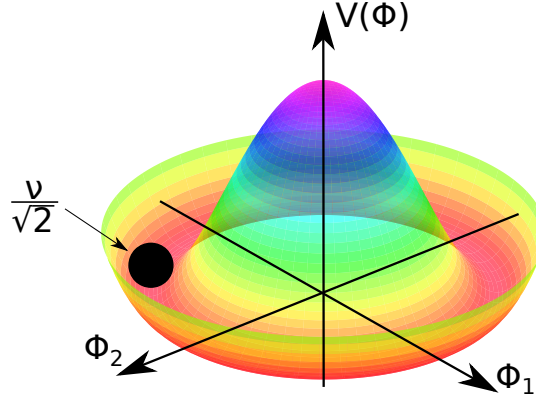


Figure 2.3: Simplified scalar potential in the SM with two degrees of freedom.

measurement of the electroweak fermi constant G_F [47] determines the vacuum expectation value to

$$v = \left(\sqrt{2} G_F \right)^{-1/2} \approx 246 \text{ GeV}. \quad (2.29)$$

The Higgs field Φ can be parameterized around v by:

$$\Phi_0(x) = \sqrt{1/2} \exp \left(i \sum_{a=1}^3 \frac{\tau_a G_a(x)}{v} \right) \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}. \quad (2.30)$$

This introduces three fields $G_a(x)$ ($a=1,2,3$) which refer to Nambu-Goldstone bosons [61–63] and a fourth field $H(x)$. The Nambu-Goldstone bosons are massless, scalar and can be eliminated by the unitary gauge transformation of $\Phi_0(x) \mapsto \exp \left(-i \sum_{a=1}^3 \tau_a G_a(x)/v \right) \Phi_0(x)$. $H(x)$ is a massive excitation from the ground state and corresponds to the Higgs boson. Therefore, the Lagrangian density after the electroweak symmetry breaking is given by [33]:

$$\begin{aligned} \mathcal{L}_{\text{Higgs}} = & \frac{1}{2} (\partial_\mu H) (\partial^\mu H) - \lambda v^2 H^2 - \lambda v H^3 - \frac{1}{4} \lambda H^4 \\ & + \frac{1}{2} \left(\frac{1}{2} v g \right)^2 W^{\mu,+} W_{\mu}^- + \frac{1}{2} \left(\frac{v g}{2 \cos(\Theta_W)} \right)^2 Z^\mu Z_\mu \\ & + g \left(\frac{v g}{2} \right) H W^{\mu,+} W_{\mu}^- + g \frac{v g}{4 \cos^2(\Theta_W)} H Z^\mu Z_\mu \\ & + \frac{g^2}{4} H^2 W^{\mu,+} W_{\mu}^- + \frac{g^2}{4 \cos^2(\Theta_W)} H^2 Z^\mu Z_\mu + \text{const}. \end{aligned} \quad (2.31)$$

This equation leads to a relation between the different weak gauge boson masses:

$$M_{W^\pm} = v \frac{g}{2} \quad \text{and} \quad M_{Z^0} = v \frac{\sqrt{g'^2 + g^2}}{2} = \frac{M_W^\pm}{\cos(\Theta_W)}. \quad (2.32)$$

The weak mixing angle Θ_W describes the relation of the masses between the charged weak gauge bosons and the neutral weak gauge boson. In addition, a new scalar particle with mass $M_H = v\sqrt{2\lambda}$ is introduced. Equation 2.31 contains self-interaction terms of the Higgs boson $H(x)$ as well as interaction terms between $H(x)$ and the electroweak bosons. The self-coupling of the Higgs field is described by the cubic (H^3) and quartic (H^4) terms. The interaction strength of the Higgs field to the vector field HVV is proportional to M_V^2/v and the one to H^2VV proportional to M_V^2/v^2 , with $V \in \{W^\pm, Z^0\}$. The defined relation between the coupling strength and the gauge boson masses can be experimentally verified.

To describe the fermion masses an additional term is introduced in $\mathcal{L}$, which is invariant under the $SU(2)_{L,W} \times U(1)_Y$ transformation. This so-called Yukawa term corresponds to the interaction between the Higgs-doublet, left-handed fermion $SU(2)_{L,W}$ -doublets and right-handed fermion singlets resulting in a mass for electrons, while neutrinos remain massless. The Yukawa term is given by:

$$\mathcal{L}_{\text{Yuk}}^{\text{lep}} = -g_e (\bar{\nu}_e, \bar{e})_L \Phi e_R + h.c., \quad (2.33)$$

with h.c. referring to the hermitian conjugate term. An additional charge conjugated Higgs doublet is defined to also enable interaction with $I_W^3 = 1/2$ quarks:

$$\Phi_C(x) = \sqrt{1/2} \begin{pmatrix} v + H(x) \\ 0 \end{pmatrix}, \quad (2.34)$$

and the Lagrangian density for the first generation is given by:

$$\mathcal{L}_{\text{Yuk}}^{\text{quark}} = -g_d (\bar{u}, \bar{d})_L \Phi d_R + g_u (\bar{u}, \bar{d})_L \Phi u_R + h.c.. \quad (2.35)$$

The interaction of higher generations can be described in the same way and the Lagrangian density of the fermion Higgs interaction is given by:

$$\mathcal{L}_{\text{Yuk}} = -M_f \bar{f} f \left(1 + \frac{H}{v} \right), \quad (2.36)$$

with the mass of the participating fermion $M_f = v g_f / \sqrt{2}$, which is proportional to the coupling constant g_f . This leads to the final Lagrangian of the SM:

$$\mathcal{L} = \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{EW}} + \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{Yuk}}. \quad (2.37)$$

Therefore, the mass of the Higgs boson remains unknown, while the coupling strength for fermions and bosons to the Higgs boson is predicted. In Chapter 3 recent measurements of the Higgs boson mass and other properties are discussed.

2.4 Limitations of the Standard Model

The SM has predicted several new particles at its formulation. All of these particles has been experimentally discovered after their prediction as the weak gauge bosons [1, 2, 4, 5, 51], the gluon [36, 64–68], the top and bottom quark [3, 6, 58, 69] and the Higgs boson [8, 9, 13–18]. In addition, the cross sections of the several interactions have been experimentally measured in agreement with the SM predictions. Figure 2.4 summarizes the cross section measurements of the ATLAS collaboration until July, 2021. The experiment was able to measure cross sections over several magnitudes without observing significant deviations from the SM predictions.

However, several observed phenomena cannot be described within the SM, such as the

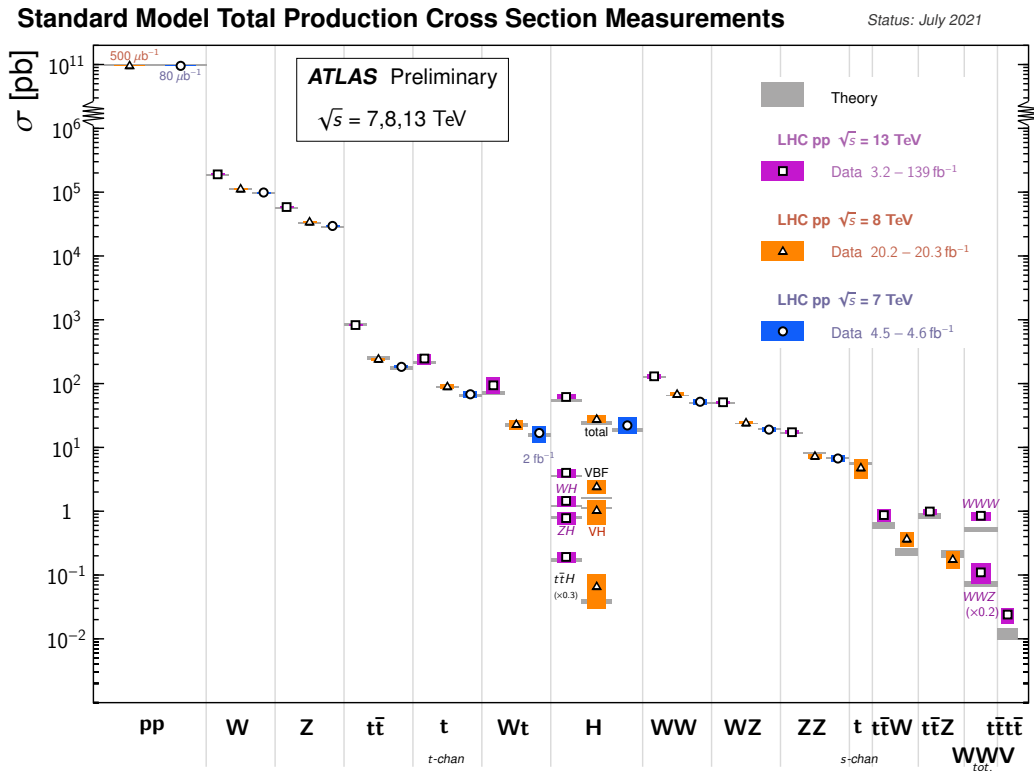


Figure 2.4: Cross sections measured with the ATLAS detector at the LHC compared to predictions of the SM [70].

fundamental force gravity [71], the baryon asymmetry of the universe [25], the presence of dark matter [23, 24] and dark energy [72, 73] or massive neutrinos [74–76]. To include these observations in the theoretical description of particle physics, the SM has to be extended. Different models have been developed to include some of these observations in an extended model beyond the SM (BSM), but until now none of these models have been experimentally

confirmed.

2.5 Extension of the Standard Model

Many different BSM models predict modifications of the Higgs boson properties and at the same time introduce new particles. As of today, none of the measurements performed at the LHC have found such new particles. These new particles can be integrated out if their mass (the scale of the new physics) is large compared to the momentum transfer in the observed physics processes. In this way, the effects on properties of the SM, as e.g. on the Higgs boson sector, can be described model independently in effective field theories. The SM effective field theory (SMEFT) uses this to encode these modifications. Kinematic effects of BSM models that act at energies Λ well above the electroweak scale ($\Lambda \gg v = 246$ GeV) can be parametrized at low energies $E \ll \Lambda$ within SMEFT by adding operators of higher mass dimensions to the SM Lagrangian:

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \sum_i \frac{f_i^{(5)}}{\Lambda} \mathcal{O}_i^{(5)} + \sum_j \frac{f_j^{(6)}}{\Lambda^2} \mathcal{O}_j^{(6)} + \sum_u \frac{f_u^{(7)}}{\Lambda^3} \mathcal{O}_u^{(7)} + \dots, \quad (2.38)$$

where $\mathcal{L}_{\text{SM}}$ is the SM Lagrangian, $\mathcal{O}^{(d)}$ are gauge and Lorentz invariant operators of dimension d built from SM fields and $f_i^{(d)}$ are real-valued dimensionless parameters, called Wilson coefficients. Each added term is suppressed by the energy scale $1/\Lambda^{(d-4)}$. Complete sets of non-redundant operators exist up to dimensions eight, e.g. for $d = 6$ the Warsaw basis [77] or the SILH basis [78, 79]. The construction of the full set of operators with higher dimensions leads to a large number of independent operators. While only one operator, the so called Weinberg operator [80], is possible for $d = 5$, 2772, 1542 and 44807 operators can be built for $d = 6$, $d = 7$ and $d = 8$ assuming three lepton generations [81]. Several assumptions can be made to reduce the number of operators, such as lepton and/or baryon number conservation, which excludes operators of dimension 5 and 7 [82, 83]. Operators with $d \geq 8$ can be neglected due to their large suppression of $1/\Lambda^4$. The assumptions of baryon number conservation and one generation of fermions reduce the number of operators to 59 for $d = 6$ [81]. At the one-loop level, the effective Lagrangian of $d = 6$ is renormalizable [84–86], so that the scale dependence of the Willson coefficients is calculable.

3 MEASUREMENTS OF THE HIGGS BOSON PROPERTIES

The Higgs boson is a consequence of the BEGGHK mechanism, which is used in the SM to introduce massive fermions and massive weak gauge bosons. It is a neutral scalar boson with spin 0 and has a short lifetime of approximately 10^{-22} s. Therefore, the Higgs boson is not directly detected at the LHC, instead it is reconstructed using its decay products. The following sections give an overview of selected properties of Higgs boson and the experimental methods used to measure them. None of these measurements or measurements not discussed in this thesis have shown a significant deviation from SM predictions. The production and decay modes of the Higgs boson are described in Section 6.1 and Chapter 7 gives a detailed description of the experimental methods used in this thesis.

3.1 Observation

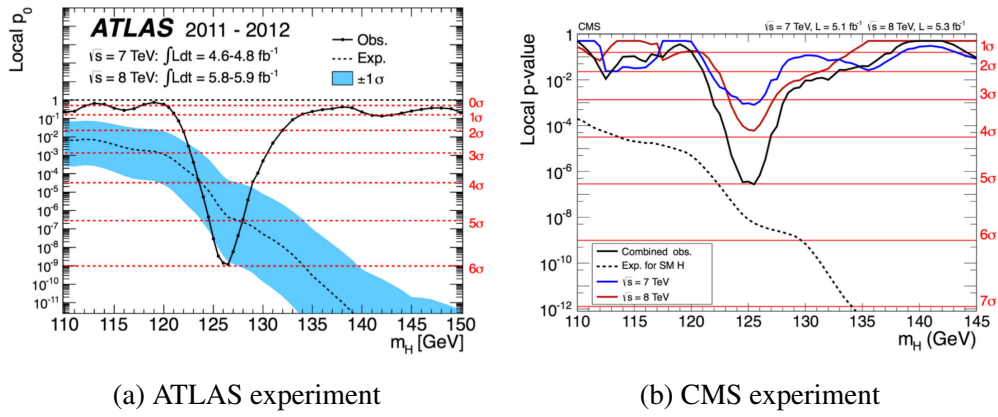


Figure 3.1: Observed (full black lines) and expected (dashed lines) p_0 -values (left y-axis) and their corresponding σ intervals (right y-axis) as a function of the hypothetical Higgs boson mass m_H for the ATLAS (a) and CMS (b) measurements with the data set recorded in 2011 and 2012 [8, 9].

The ATLAS and CMS experiments at CERN announced the observation of a Higgs boson with a mass of approximately 125 GeV in July 2012. They reached a significance of 5.9σ and 5.0σ [8, 9]. The significance is based on p_0 -values, which is a criterion for the probability that the resonance is introduced by fluctuating backgrounds. A value of 5σ corresponding to a p_0 -value of $2.85 \cdot 10^{-7}$ is necessary to claim an observation. Figure 3.1 shows the observed p_0 -values of the two experiments as a function of the particle mass m_H .

3.2 Mass

As described in Section 2.3 the mass of the Higgs boson remains as an unknown parameter in the SM. Therefore, the precise measurement of the Higgs boson mass is one of the

3.3 Cross sections

important goals of the ATLAS and CMS experiments at the LHC. The mass of Higgs boson is reconstructed using its decay products. For a precise reconstruction the resolution and the misidentification rate of the final state particles has to be small, so that only the $H \rightarrow \gamma\gamma$ and the $H \rightarrow ZZ^* \rightarrow 4l$ decay mode are used for this measurement due to their clear signal signature and the missing of neutrinos in the final state. The measurements of the Higgs boson mass are summarized in Figure 3.2. Both experiments study the Run 1 data set as well as the first part of the Run 2 data set recorded in 2015 and 2016 and a combination of both data sets [87, 88]. In addition, a combination of the two Run 1 measurements of both experiment was performed [19]. The results of the individual data sets, experiments and decay modes show no significant deviation to each other. The most precise result is given by the combined CMS measurement with $m_H = (125.38 \pm 0.14)$ GeV corresponding to a relative uncertainty of 0.1%. Therefore, the mass of the Higgs boson is already more precisely measured as the world average of the mass of the top quark with (172.76 ± 0.30) GeV and a relative uncertainty of 0.2% provided by the particle data group [89].

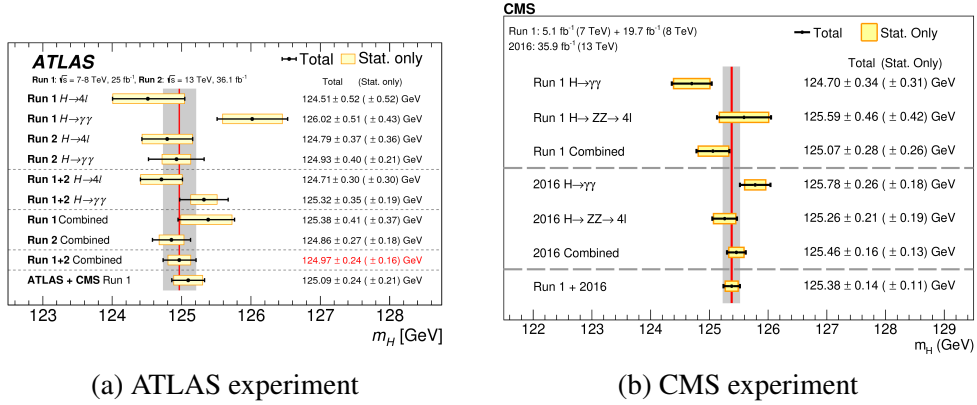


Figure 3.2: Observed mass of the Higgs boson for the ATLAS (a) and CMS (b) measurements with the Run 1 data set form 2011 and 2012 and the with the first part of the Run 2 data set from 2015 and 2016. Both figures show the combined results of the two data sets and in (a) the Run 1 combination of the two experiments is shown. The best-fit value are illustrated by black point together with their full uncertainty (black line) and their statistical uncertainty (yellow box) [87, 88].

3.3 Cross sections

3.3.1 Inclusive cross sections

The ATLAS experiment has used the $H \rightarrow \gamma\gamma$ and the $H \rightarrow ZZ^* \rightarrow 4l$ decay modes to determine the inclusive Higgs production cross section $\sigma_{pp \rightarrow H}$ as a function of the center-of-mass energy $\sqrt{s}$ [90]. The measurement is based on the data sets from 2011, 2012 and 2015-2018 and the results are shown in Figure 3.3 (a). No significant deviation from the SM

prediction is observed.

In addition, the cross sections of the dominant Higgs boson production modes at the LHC are measured by the ATLAS experiment. To reduce the extrapolation from the phase space covered by the detector to the measured production modes cross section, a fiducial volume is defined by requiring a rapidity of the Higgs boson y_H of less than 2.5. The analyses of several different Higgs boson decay modes are combined for this measurement and the data sets recorded between 2015 and 2017 are used [29]. Chapter 11 discusses this analysis in detail as a part of this thesis. All observed cross sections agree within one standard deviation with the SM prediction. The most precise measured production mode is gluon-gluon-fusion with a relative uncertainty of 8.6%.

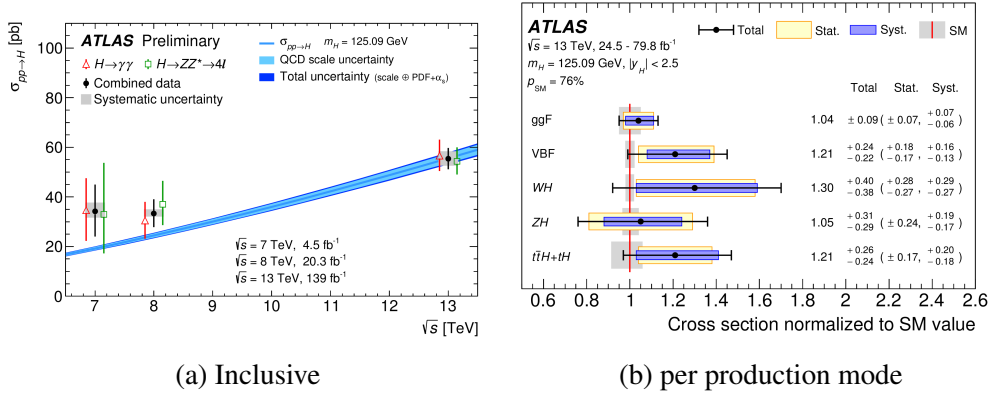


Figure 3.3: Observed values of the Higgs boson for the inclusive cross section as a function of the center of mass energy (a) and split per production mode (b). In (a) the single decay channel results are shown in red ($H \rightarrow \gamma\gamma$) and in green ($H \rightarrow ZZ^* \rightarrow 4l$), while the combined results are shown in black. The blue band shows the SM prediction with its uncertainty. In (b) the black points with corresponding black bars show the best-fit values and total uncertainties for the cross sections of the individual production modes normalized to SM prediction. The yellow boxes illustrate the statistical uncertainty and the blue boxes the systematic uncertainties on the measurement, while the gray boxes correspond to the theory uncertainty on the SM prediction [29, 90].

3.3.2 Simplified template cross sections

The Simplified Template Cross Section (STXS) [27] is an analysis strategy to measure the Higgs boson production cross section in different regions of phase space. The full phase space of the Higgs boson production is split into exclusive regions of phase space using kinematic properties. These regions of phase space are defined at particle level, while the event selection is defined at reconstruction level. Therefore, the measured cross sections are extrapolated from reconstruction level to particle level within the likelihood model. The inclusive phase space is split into the different production modes, which are further split into smaller regions

3.3 Cross sections

of phase using kinematic observables as the jet multiplicity or the transverse momentum of the Higgs boson. The STXS method uses different kinematic observables in different regions of phase space, which results in a rather complex splitting of the global phase space. A detailed description of the method is given in Section 7.1. The common definition of the measured production cross sections allows the combination across different Higgs boson decay modes. Such a combined measurement benefits from the sensitivity of different decay modes to different regions of phase space and can constrain in this way the cross sections of a large number of individual regions of phase space.

The ATLAS experiment has published several combinations using this method. The results using the data set recorded from 2015 to 2017 are visualized in Figure 3.4. No significant deviation from the SM predictions is observed. This analysis is part of this thesis and further discussed in Chapter 11.

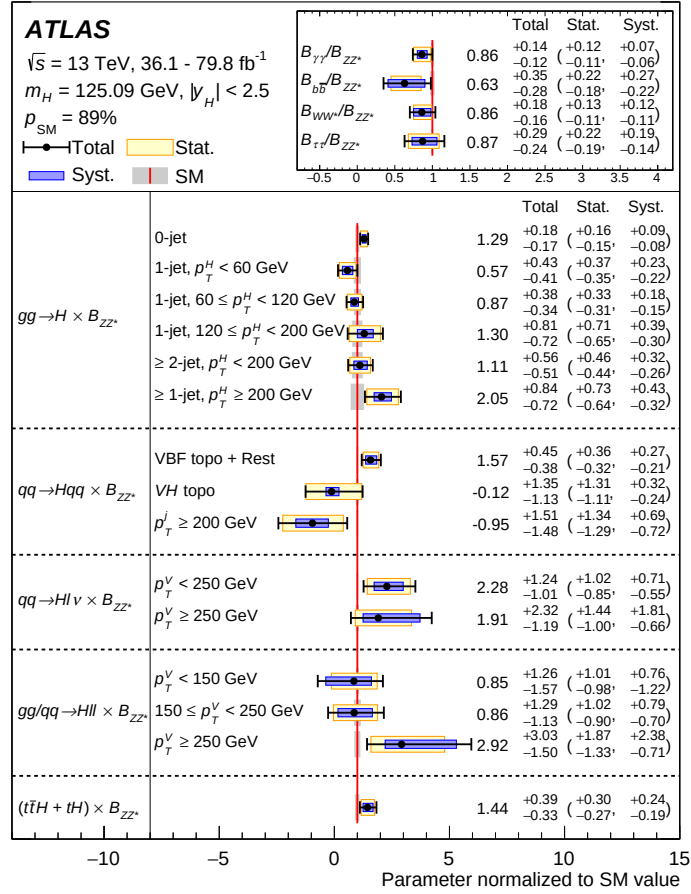


Figure 3.4: Measured cross sections for different regions of phase space for the Higgs boson production using the STXS framework [29].

3.3.3 Differential cross sections

The detector is limited by finite energy and space resolution but also by misidentification and inefficiency of the measured objects. For differential measurements the cross section of region of phase space defined on reconstruction level are measured first and then transferred to the particle level using the unfolding method [91–93]. This method corrects all effects of the detector limitations in contrast to the method used in STXS measurements. Disadvantages of differential measurements compared to the STXS method are the individual definitions used in each analysis focusing on a different Higgs boson decay mode. The definition of the fiducial phase space, the kinematic observables used for the differential measurements and the binning of them varies across the individual analyses. This inconsistent definition leads to complex combination procedure, so that differential measurements are limited to a sub-set of decay modes. This limits the statistics of differential measurements compared to the STXS measurements. Therefore, differential cross sections are usual only measured as a function of one or two kinematic observables, while STXS framework includes several different observables. Figure 3.5 shows an ATLAS measurement using the $H \rightarrow \gamma\gamma$ and the

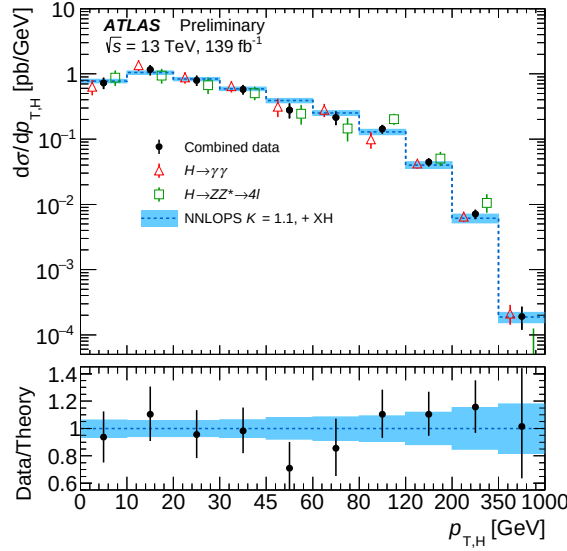


Figure 3.5: Observed cross section of the Higgs boson production as a function of the transverse momentum of the Higgs boson. The measurements of the single decay channel are shown in red ($H \rightarrow \gamma\gamma$) and in green ($H \rightarrow ZZ^* \rightarrow 4l$), while the combined result is shown in black. The SM prediction is illustrated in blue [90].

$H \rightarrow ZZ^* \rightarrow 4l$ decay modes to perform a combined differential measurement of the Higgs boson production cross section σ as a function the transverse momentum of the Higgs boson $p_{T,H}$ [90]. No significant deviation from SM predictions are observed in this measurement.

4 EXPERIMENT SETUP

4.1 The Large Hadron Collider

The LHC (Large Hadron Collider) [10] running since 2008 is a particle accelerator and collider at CERN near Geneva. It consists of two ring accelerators with a circumference of 27 km and can be filled with protons or heavy ions. It is placed between 100 and 150 meters below ground-level and has four collision points. The accelerator is filled with up to 2808 bunches with 10^{11} particles per bunch and a spacing of 25 ns for protons. 1232 superconducting dipole magnets bend the proton beams with a magnetic field up to 8.3 Tesla and additional 392 quadrupole magnets focus them.

Two main properties can describe the performance of the LHC. The instantaneous luminosity $\mathcal{L}$ for two colliding particle beams with an equal Gaussian shape distribution and is defined by [94]:

$$\mathcal{L} = \frac{N^2 \gamma_r f_{\text{rev}}}{4\pi \beta^* \varepsilon_n} n_b \cdot F, \quad (4.1)$$

with the number of particle per bunch N , the relativistic gamma-factor γ_r , the revolution frequency f_{rev} , the beta-function at the collision point β^* , the normalized transverse beam emittance ε_n , the number of bunches per beam n_b and the geometric luminosity reduction factor F considering the beam crossing angle and other effects at the interaction point. The

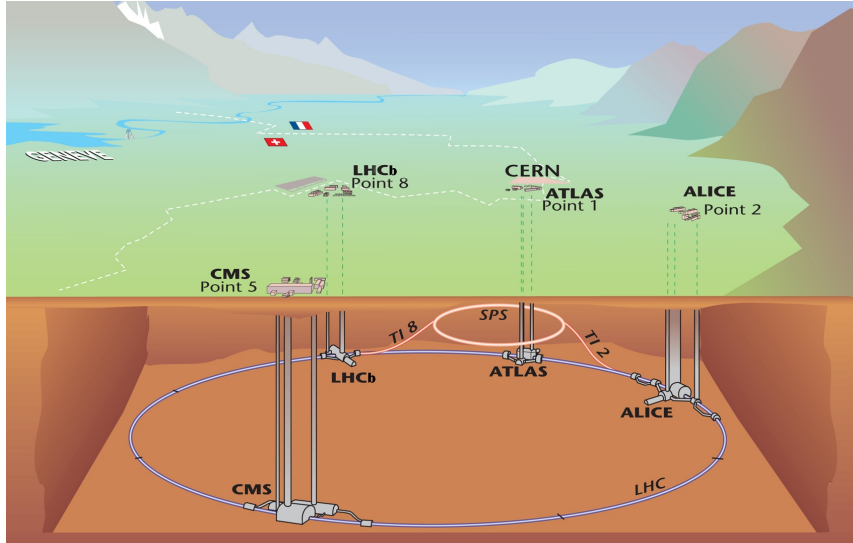


Figure 4.1: Schematic overview of the LHC with the four major experiments ALICE, ATLAS, CMS and LHCb [95].

LHC is designed for an instantaneous luminosity up to $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. The second main property is the center-of-mass energy $\sqrt{s}$, which can reach up to 14 TeV for the LHC (7 TeV

per beam).

For each of the four collision points of the LHC a unique detector is designed to answer different physical questions. These detectors are ALICE [96] and LHCb [97] as well as two multi-purpose detectors ATLAS [11] and CMS [12]. The LHC and the four main experiments are shown in Figure 4.1.

4.2 The ATLAS experiment

The ATLAS experiment [11, 98, 99], one of four major experiments at the LHC, has a detector with a diameter of 25 meters and a length of 44 meters and is the largest particle detector at the LHC. It consists of different parts built in a cylindrical way and has a weight of 7000 metric tons. From the inside to the outside the main components of the detector are: inner detector surrounded by a superconducting magnet solenoid, electromagnetic and hadronic calorimeters and muon 3x8 toroids instrumented with muon chambers. A schematic overview of the detector is given in Figure 4.2. The x-y-plane for the coordinate system of the ATLAS

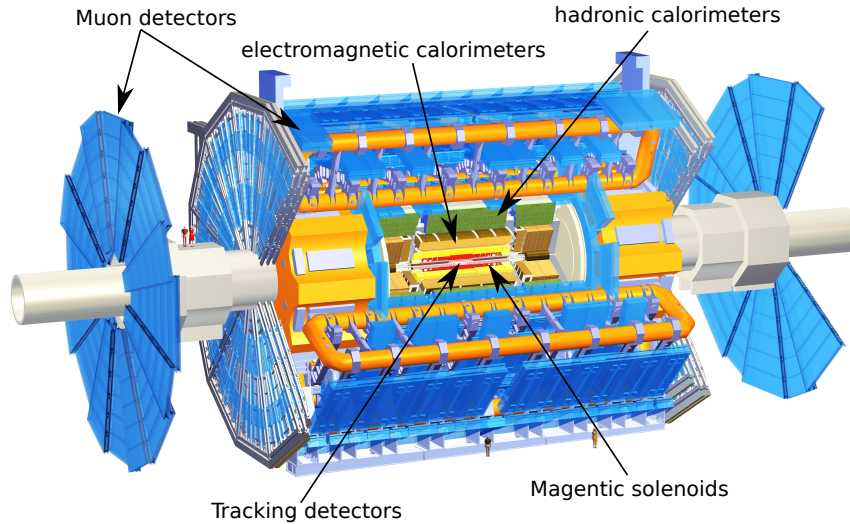


Figure 4.2: Visualization of the ATLAS detector and its main components [100].

experiment is defined transversal to the beam axis and the x-axis points towards the center of the LHC. The y-axis is orthogonal to the x-axis and points upward towards the surface. The z-axis is defined parallel to the beam axis, in such a way that a right-handed coordinate system results. The nominal interaction point is used as the origin of the system. The azimuthal angle ϕ is defined in the (x-y)-plane, while the polar angle θ is defined with respect to the z-axis. In measurements the pseudorapidity η is used instead of θ defined as

$$\eta = -\ln(\tan(\theta/2)). \quad (4.2)$$

A nearly flat distribution of the particle density in η is introduced by the variety of soft QCD processes in a proton-proton collision. The pseudorapidity is an approximation for the rapidity

of a particle at high energy or negligible mass: $y = 1/2 \ln(E+p_z/E-p_z)$, where E is the energy and p_z the projection of the particle momenta in z-direction. The Lorentz invariance of Δy then results in an invariance of $\Delta \eta$ under a boost along the z-axis. This is an important property of the system, since the boost along the z-axis is unknown for two colliding partons. The separation of two objects in the η - Φ -plane is defined by:

$$\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \Phi)^2}, \quad (4.3)$$

which is as well invariant under a boost in the z-direction. This property is used in the identification and the reconstruction of the physics objects.

4.2.1 Inner detector

The most central part of the ATLAS detector is the inner detector [101], which has a length of 6.2 m and a radius of 1.1 m. It consists of different tracking detectors and is visualized in Figure 4.3. The inner detector is designed to provide precise momentum measurements for charged particles, to reconstruct the particle charge, tracks and momenta as well as the primary and secondary vertex.

It consists of a pixel detector, a silicon micro-strip tracker (SCT), a transition radiation tracker

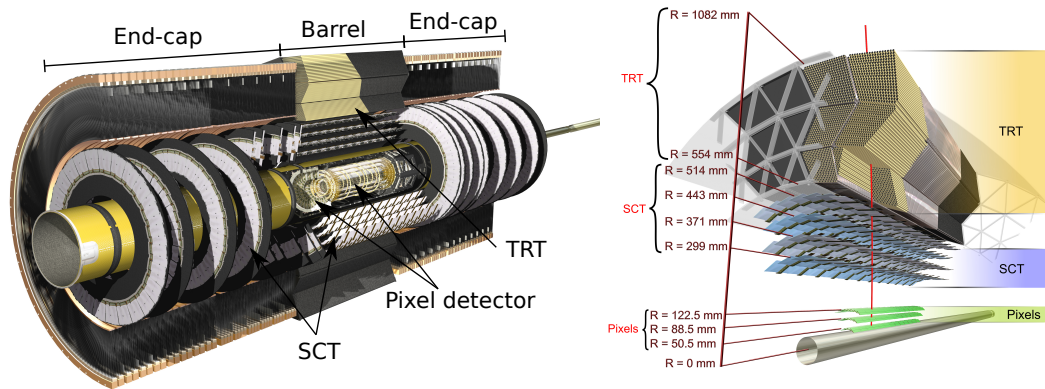


Figure 4.3: Schematic overview of the inner detector of the ATLAS experiment and its main components [100].

(TRT) and a central magnetic solenoid generating a two Tesla magnetic field, that penetrates the inner detector. The expected resolution of the inner detector regarding momentum measurements are given in Table 4.1. While the TRT covers a range up to $|\eta|=2.0$, the Pixel detector and SCT are especially important for small radii and cover a range up to $|\eta|=2.5$. Both trackers consist of a barrel region and two end-caps, which are mounted radially. The pixel detector and the SCT are placed concentrically around the beam axis in the barrel region, while the straw tubes of the TRT run parallel to the beam axis. As the most inner part of the

experiment, the inner detector copes with high radiation doses and is cooled down to reduce potential damages. It consists of the insertable b-Layer (IBL) installed during the first long shutdown of the LHC in 2013 [102]. The IBL is build cylindrical around the beam axis in a distance of 33mm and provides a precise measurement of the secondary vertex, which is in particular used for the identification of b-hadrons.

Both the barrel and the end-cap part of the pixel detector consist of three layers corresponding to 250 μm thick silicon semi-conductors with 46080 readout channels for each of the 1744 modules. Every pixel has a size between $(R - \Phi) \times z = 50 \times 400 \mu\text{m}^2$ and $(R - \Phi) \times z = 50 \times 600 \mu\text{m}^2$, such that the detector reaches an intrinsic hit resolution of $(R - \Phi) \times z = 10 \times 115 \mu\text{m}^2$.

Four layers of silicon strip detectors build the barrel part of the SCT, while the end-caps consist of nine layers, such that for each charged particle at least four measured points are ensured. The SCT has 758 active strips per module of 12 cm length and 285 μm thickness per strip and contains over 6 million readout channels. The nominal hit resolution of the detector is $(R - \Phi) \times z = 17 \times 580 \mu\text{m}^2$.

The TRT consists of straw tubes of 4 mm diameter and a length of 144 cm in the barrel region, while in the end-cap region the tubes are installed radially in wheels of 37 mm length. They are filled with gas and are stabilized by carbon fibers. Therefore, only a position measurement of $(R - \Phi) = 130 \mu\text{m}$ and none in the z-axis is performed in the barrel region. The TRT is able to provide important information for the particle identification. The transition radiation emitted by particles penetrating the interface of materials with different di-electric constants is inversely proportional to the mass of the particles. Gas-filled tubes stabilized by carbon fibers are used as the transition material, so that especially charged pions and electrons can be distinguished.

4.2.2 Calorimeter system

The calorimeter [103–105] is build up by an electromagnetic calorimeter (ECal), which is optimized to measure the energy of photons, electrons as well as positrons and a hadronic calorimeter (HCal), which is used together with the ECal to measure the energy of hadrons, jets and the missing transverse energy E_T^{miss} . The η -coverage and the energy resolution of the sub-calorimeters are given in Table 4.1.

Electromagnetic calorimeter The ECal consists of lead with a liquid argon sampling technology and has a fine granularity. It is build with accordion-shaped electrodes and lead absorber plates with a complete symmetry in the azimuthal angle Φ and is split into a barrel and two end-cap parts (see Figure 4.4). The barrel is split into two identical half-barrels, which cover a pseudorapidity range up to $|\eta| = 1.475$. The barrels consist of 3 layers with granularity of $(\Delta\eta \times \Delta\Phi) = 0.025/8 \times 0.025$, $(\Delta\eta \times \Delta\Phi) = 0.025 \times 0.025$ and $(\Delta\eta \times \Delta\Phi) = 0.05 \times 0.025$. The range between $1.375 < |\eta| < 3.2$ is covered by an inner and outer end-cap wheel (EMEC). A pre-sampler detector is added in the central region ($|\eta| < 1.8$) to compensate energy loss of photons or electrons upstream of the calorimeter. This pre-sampler is made of an active liquid argon layer.

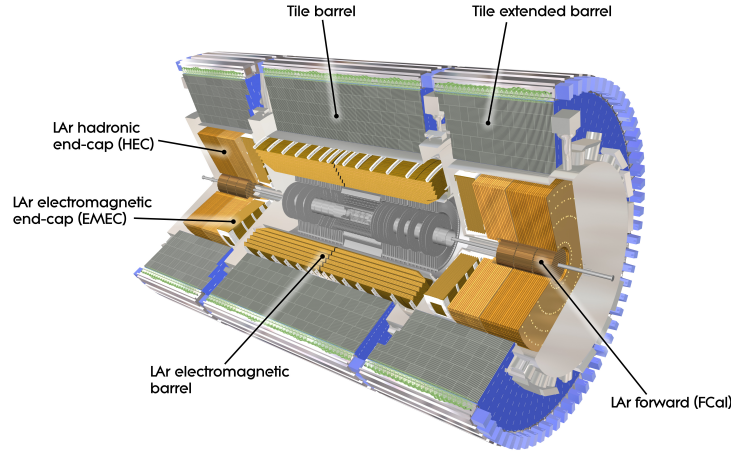


Figure 4.4: Schematic overview of the calorimeter system of the ATLAS experiment [100].

Hadronic calorimeter The HCal (see Figure 4.4) has a coarser granularity compared to the ECal and measures the energy of jets and E_T^{miss} together with other sub-detectors. The tile calorimeter consists of a barrel part with $|\eta| < 1.0$ around the ECal and two extended parts covering higher values of $0.8 < |\eta| < 1.7$. Steel plates as an absorber and scintillating tiles as active material covering a $|\eta|$ range up to 3.2 are the components of the hadronic barrel calorimeter. For the hadronic end-cap (HEC) and the forward calorimeter (FCal) liquid argon is used as a sensitive material. For the HEC only copper is used as absorber material, while for the FCal tungsten is used as well.

4.2.3 Muon system

To measure charged particles, which penetrate the inner detector and the calorimeters, a muon system [106] is used, which is visualized in Figure 4.5. It consists of a barrel region covering a $|\eta|$ range up to 1.4, the transition region with $1.4 < |\eta| < 1.6$ and the end-cap region with $1.6 < |\eta| < 2.7$. To measure the momentum of the charged particle toroidal magnets are installed, generating a field strength of 1.5-5.5 Tm in the barrel and 1-7.5 Tm in the end-cap region.

The muon system is build up by two types of precision tracking chambers, monitored drift tubes (MDT) and cathode strip chambers (CSC), and two types of trigger chambers, resistive-plate chambers (RPC) and thin-gap chambers (TGC). The muon are bent parallel to the beam axis by barrel toroid magnets in the region $|\eta| < 1.4$, by end-cap toroids in the region $1.6 < |\eta| < 2.7$ and by a combined field of barrel and end-cap toroids in the intermediate range with $1.4 < |\eta| < 1.6$.

To reconstruct the muon trajectory 1088 MDT chambers are used in the central region with $|\eta| < 2.7$, while in the innermost wheel in the end-cap region the MDTs are replaced by CSCs

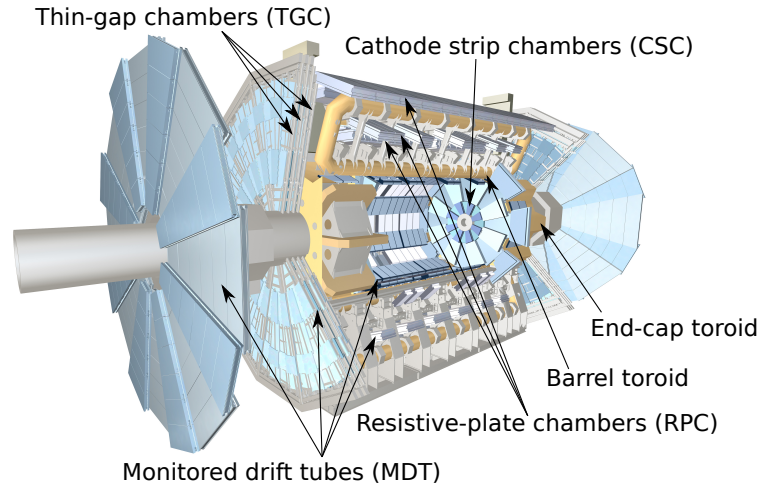


Figure 4.5: Schematic overview of the muon system of the ATLAS experiment [100].

in the region $2.0 < |\eta| < 2.7$. The MDT consists of three cylindrical layers of tubes around the beam axis in the barrel region and of three wheels perpendicular to the beam axis in the end-cap region. Each layer is build up with three layers of MDTs, except of the innermost which consists of four layers. The MDTs are aluminum tubes filled with a gas mixtures consisting 93% Ar and 7% CO₂ and a diameter of 30 mm. CSCs are multiwire proportional cambers consisting of radially oriented wirs and cathodes segmented into strips. The CSCs are filled with a gas mixture of 80% Ar and 20% CO₂. The CSC is able to provide a better spatial resolution to cope with the higher signal rates in the large η -range.

The RPC is used to trigger particles in the barrel region, while the TGC is used in the end-cap region, so that the muon system can trigger particles up to $|\eta| = 2.4$. Both systems are able to deliver a signal within 15-25 ns. An overview of the required resolution for the muon system is given in Table 4.1.

4.2.4 Trigger

Due to the high bunching crossing rate of the LHC (40 MHz) a two level trigger is used at the ATLAS experiment to reduce the output-recording rate [107]. The first level of the trigger system is a hardware-based trigger (L1) and the second level a software-based high level trigger (HLT), which are visualized together with the data acquisition in Figure 4.6. Region-of-Interests (RoIs) are determined by custom electronics at the L1 using a coarse granularity calorimeter and muon system information. L1 reduce the rate from 40 MHz to 100 kHz, with a decision time of 2.5 μ s. The RoIs are then used as an input for the HLT, which uses sophisticated selection algorithms on the full granularity detector information of the RoI. The HLT further reduces the output-rate of the L1 from 100 kHz to approximately 1 kHz on average, with a processing time of about 300 ms.

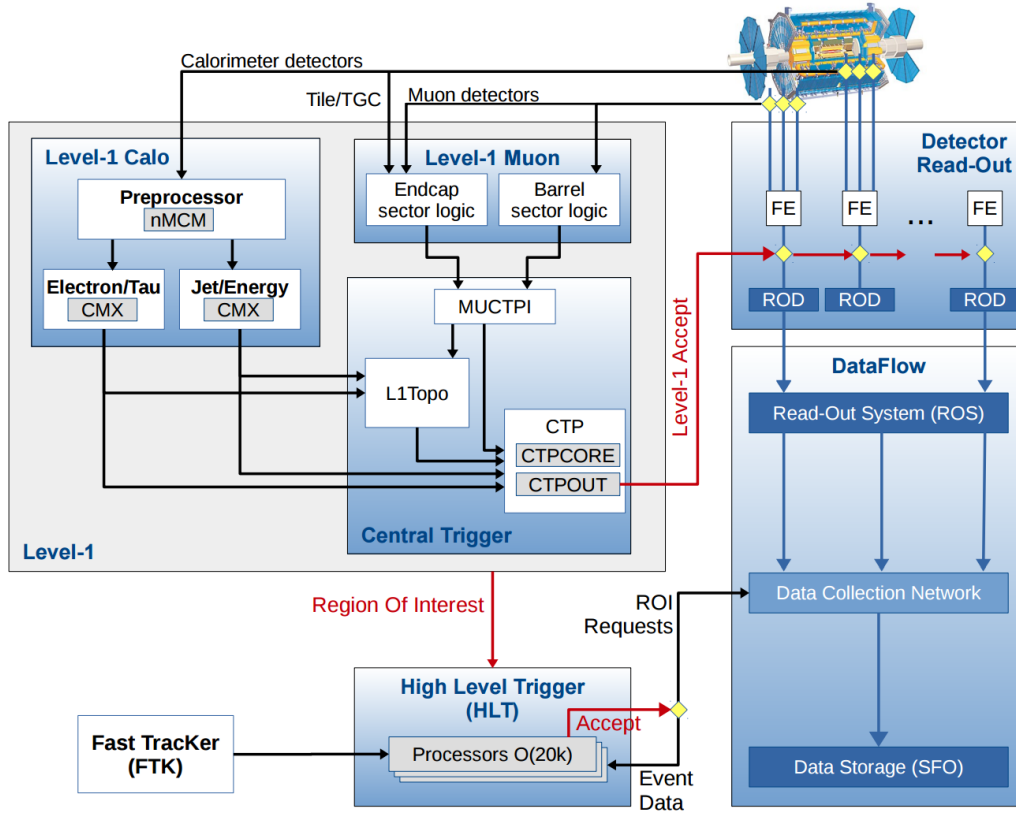


Figure 4.6: Schematic overview of the trigger system and the data acquisition of the ATLAS experiment in Run 2 [107].

Sub-detector	Required resolution	η -coverage	
		Measurement	Trigger
Inner detector	$\sigma_{p_T}/p_T = 0.05\% p_T \oplus 1\%$	± 2.5	
Electromagnetic calorimeter	$\sigma_{E/e} = 10\% / \sqrt{E} \oplus 0.7\%$	± 3.2	± 2.5
Hadronic calorimeter			
barrel and end-cap	$\sigma_{E/e} = 50\% / \sqrt{E} \oplus 3\%$	± 3.2	± 3.2
forward	$\sigma_{E/e} = 100\% / \sqrt{E} \oplus 10\%$	$3.1 < \eta < 4.9$	$3.1 < \eta < 4.9$
Muon system	$\sigma_{p_T}/p_T = 10\%$ at $p_T = 1$ TeV	± 2.7	± 2.4

Table 4.1: The resolution and the pseudorapidity (η) coverage of the different sub-detectors of the ATLAS experiment. Energy and momentum values are given in units of GeV. $a \oplus b$ refers to adding a and b in quadrature: $a \oplus b = \sqrt{a^2 + b^2}$ [11].

4.3 Simulation process of pp collisions

To compare the recorded data of the ATLAS experiment to the theory predictions simulations of the physics processes are needed. This is a quite challenging process for proton-proton collisions, since several physics phenomena with different scales of momentum transfer have to be taken into account.

The factorization theorem [108] is used to treat the different phenomena separately. The collisions can be split into a hard and soft process. The hard process describes the interaction between the partons by using perturbation theory and has a large scale of momentum transfer. For the soft process with a small scale of momentum transfer non-perturbative models with parameters derived from data are used. Soft processes are described by parton distribution functions of the protons and the hadronization (fragmentation) of the colored partons produced during the collision into colorless hadrons.

Underlying event, pile-up, initial- and final-state radiations have to be considered as well. The underlying event is caused by interactions of beam remnant, while pile-up describes particles produced in additional pp collisions during the same or neighboring bunch crossing and is described in Section 4.3.3.

To simulate these effects, which are summarized in Figure 4.7, Monte Carlo event generators with random sampling techniques are used.

This simulation process generates particle level events, which include the theoretical prediction but no detector effects. To be able to compare the simulations with the recorded data several resolution and efficiency effects of the detector have to be considered. Simulated events including these effect are called reconstruction level events. Therefore, the particle level events are passed through a simulation of the ATLAS detector based on GEANT4 [109].

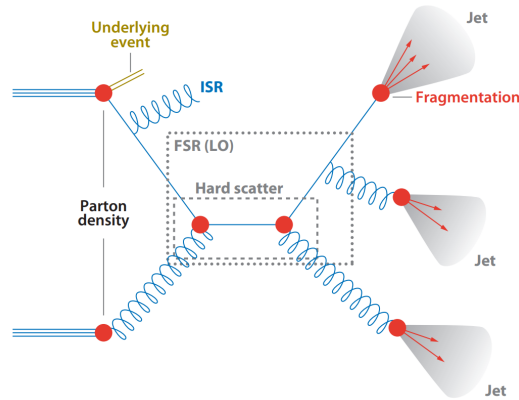


Figure 4.7: Schematic overview for the simulation of proton proton collision at particle level. ISR and FSR are examples for initial- and final-state radiation [110].

4.3.1 Partonic cross section and parton shower

The cross section $\sigma_{p_1 p_2 \rightarrow X}$ for two colliding protons p_1 and p_2 producing a particular final state X quantifies the probability that this process takes place. It can be related to the expected number of events N using the total integrated luminosity (see Equation 4.1):

$$N = \sigma_{p_1 p_2 \rightarrow X} \cdot \int \mathcal{L} dt. \quad (4.4)$$

The parton distribution functions describe the probability to find a parton with the momentum fraction x of the initial proton momentum in a collision at a momentum transfer Q^2 .

The factorization theorem, the partonic cross section $\hat{\sigma}_{ab \rightarrow X}$ and the parton distribution functions f_a and f_b of the interacting partons a and b can be used to calculate the cross section $\sigma_{p_1 p_2 \rightarrow X}$ by integrating over the momenta for all possible parton flavor combinations:

$$\sigma_{p_1 p_2 \rightarrow X} = \sum_{a,b} \int dx_a dx_b f_a(x_a, \mu_F^2) f_b(x_b, \mu_F^2) \hat{\sigma}_{ab \rightarrow X}(x_a, x_b, \mu_F, \mu_R), \quad (4.5)$$

with factorization scale μ_F and renormalization scale μ_R .

This partonic cross section can be parametrized in a power series in the strong coupling strength α_s :

$$\sigma = \sum_n \alpha_s^n \sigma_n. \quad (4.6)$$

Higher orders in α_s take into account virtual corrections, so called loops, or real emissions of additional particles in the final state. The real emissions can be grouped into so called initial state radiation (ISR), in which the additional parton is emitted by an initial state parton, while for so called final state radiations (FSR) a final state parton emits the additional parton.

Including particle momenta in loops results in ultraviolet divergences. These can be regularized via dimensional regularization by introducing the unphysical renormalization scale μ_R , such that the divergences are absorbed inside the coupling strength α_s . With this, α_s gains a dependency on μ_R and is called a running coupling strength.

Collinear real emissions results in infrared divergences. Again dimensional regularization can be used to regularize the divergence introducing the unphysical factorization scale parameter μ_F . The probability for a parton i to radiate a parton j can be described using splitting functions. Here $q \rightarrow qg$, $\bar{q} \rightarrow \bar{q}g$, $g \rightarrow q\bar{q}$ and $g \rightarrow gg$ have to be considered, which results in a set of coupled differential equations named Dokshitzer-Gribov-Lipatov-Altarelli-Parisi [111–113]. This set of equations describes the evolution of the parton distribution function between different scales of momentum transfers using μ_F and the momentum fraction of the partons as input parameters. In the calculation the separation of the perturbative and non-perturbative energy scale regime is introduced by μ_F .

The dependence of the cross section on both unphysical parameters is a result of a finite order calculation of Equation 4.6 and would vanish for $n \rightarrow \infty$. The choice of these parameters is arbitrary so that both scales are varied to estimate an approximation of the missing higher order contributions. Most of the simulation are performed with the hard scattering energy

scale for both parameters.

The quasi continuous final state radiation of partons results in parton showers, which can be described by the splitting functions. For each emission the energy scale and momentum fraction decreases. The showering is simulated from the hard scattering energy scale until a lower cutoff scale. This evolution to a lower energy scale is described by the Sudakov form factors [114]. The radiation of an additional parton from an initial state parton is called initial state radiation. It is described by PDFs and is simulated by going from the hard scattering scale to a cutoff scale, but increasing the momentum fraction.

The parton shower approach is limited to collinear and soft partons, while processes with higher energy or widely separated partons are described based on matrix element calculations. To combine these two approaches a matching or merging scheme is applied. For both schemes an energy scale threshold have to be defined, which resolves any overlap in the phase space of the resulting jets. Jets below this threshold are described using parton shower simulation, while the remaining jets are simulated with the matrix element. The two schemes use different approaches for the matrix-element calculations. While a full next-to-leading-order calculation, including loop correction, is applied for the matching scheme, a leading-order calculation with an increasing number of emitted partons is used in the merging scheme.

The beam remnants, the parts of the interacting proton, which are not involved in the hard process, can produce additional (soft) interactions. These additional interactions lead to the underlying event, which includes also multiple parton interactions between the remaining partons. While jets from underlying event are usually soft and therefore not reconstructed by the ATLAS detector, the total energy is affected by these jets. The modelling of underlying events is based on phenomenological models as for other soft interactions.

4.3.2 Hadronization and hadron decays

Quantum Chromodynamics predicts that quarks and gluons cannot exist as free particle (see Section 2.2.2). Therefore, the partons produced during the colored parton shower simulation have to be combined to colorless hadrons. This step is called hadronization or fragmentation and is affected by color connection between the partons of the hard process and the beam residuals as well. The two approaches cluster fragmentation and string fragmentation can be used to describe the hadronization.

The cluster fragmentation model described in [115, 116] use a split of gluons into quark pairs until they reach a lower cutoff scale. From these colorless clusters are formed, which subsequently decay into hadrons. This approach is used in the SHERPA Monte Carlo event generator [117].

An example for a string fragmentation model is the Lund string model [118]. The model use the increase of the potential energy between color charged partons with an increase of spatial separation. A upper threshold is set on the potential energy, so that a new quark pair is created when the threshold is crossed. This is done iteratively until the energy of all parton is low enough and only colorless hadrons remain. These hadron can be unstable, so that they decay into lighter hadrons until all hadrons are stable. This approach is used in the PYTHIA event generator [119].

4.3.3 Pile-up

Pile-up describes the interaction of additional protons. Pile-up can be grouped into in-time pile-up and out-of-time pile-up. In-time pile-up describes the interaction of two protons, which are in the same proton bunches as the hard scattered protons, while out-of-time pile-up is produced by proton of neighboring bunches. Pile-up is simulated as an inclusive sample of collision events and it is combined with the simulation of the hard process. The explicit beam condition affects the number of interactions per bunch crossing μ , which is one of the key parameter for the pile-up simulation. To achieve a good modeling of the pile-up activity, the simulated distribution are re-weighted to the distribution of μ observed in data.

4.3.4 Detector simulation

The produced simulate events are at particle level, which follow the SM prediction. The recorded data includes additionally detector effects like the infinite resolutions, particle misidentification and inefficiencies. To be able to compare the simulated events with data these detector effects have to be included in the simulation. The simulated events are corrected for this using GEANT4 [109], which simulates the interaction between the simulated final state particles with the detector materials of the ATLAS experiment. In GEANT4 a detailed description of the detector is implemented including for example passive materials, magnetic field configurations and known detector defects of particular detector parts. The software produces an output, which is identical to the format of the recorded data, so that the same trigger, identification and reconstruction software can be used on them. These events, which include detector effects and the approximations of reconstruction algorithms, are called reconstruction level events. These events are directly comparable to data, while a direct comparison to the theory prediction or other experiments is not possible anymore.

4.4 Data set

The LHC has well defined periods of time in which the accelerator is running with stable beams, so that the experiments can collect data. These time periods are called Runs. This time between the Runs is used for maintenance and upgrades of the accelerator and the detectors. Additional short shutdown as for example at the end of each data taking year are used for small upgrades or repairs. The first data taking period was Run 1 from 2011-2012 and the second one Run 2 from 2015-2018. At the beginning of the year 2022 the Run 3 data taking period is planed. In addition to the long shutdown between the two Runs end-of-year technical stops are performed after each year of data taking. Figure 4.8 show an overview of the data taking periods. While the center of mass energy was changed in Run 1 from 7 TeV (2011) to 8 TeV (2012) a constant energy of 13 TeV was used during Run 2.

This thesis will focus on the data from Run 2. This Run is additional split by an end-of-year extend technical stop (EYETS) at the end of 2016. This stop was used to start a first round of analyses on the partial Run 2 data set taken from 2015-2016. This thesis includes two

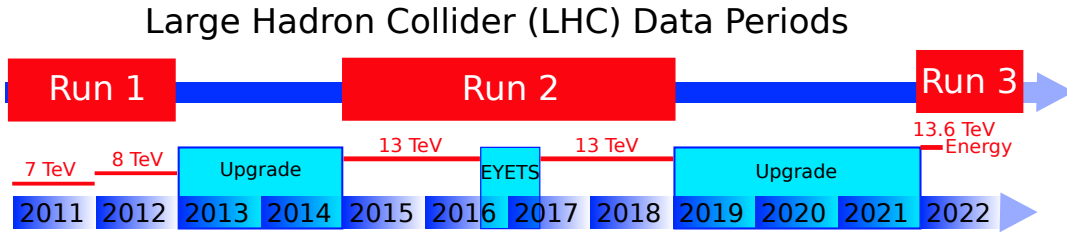


Figure 4.8: Schematic overview of the data taking periods of the LHC. EYETS corresponds to an end-of-year extend technical stop.

analyses, which are based on this data set. The remaining analysis in this thesis are based on the full Run 2 data set including the data from 2015 to 2018. Figure 4.9 shows the integrated luminosity of Run 2 as a function of time. The figure shows the luminosity for the data delivered by the LHC and for the data taken by the ATLAS experiment. Furthermore, during a small fraction of the data taking not all detector subsystems and magnetic system were fully operational. Data taken during this time periods is not used for analyses. The data which fulfill all quality criteria are stored in Good Run Lists and the luminosity of this data is also shown in Figure 4.9. The partial Run 2 data set recorded until the EYETS has an integrated luminosity of 36.1 fb^{-1} and the full Run 2 data set a integrated luminosity of 139 fb^{-1} . For the pile-up simulation the number of interaction per bunch crossing has to be measured.

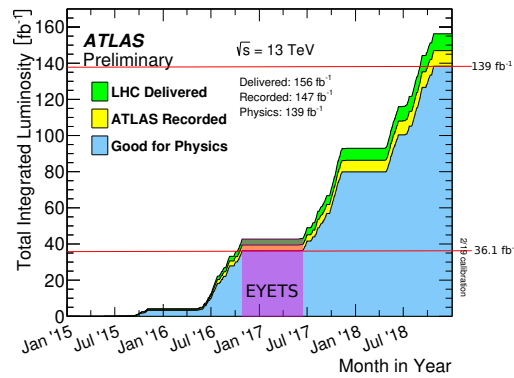


Figure 4.9: Total integrated luminosity for the ATLAS detector during Run 2 as a function of time. Green shows the luminosity that was provided by the LHC and yellow the part of the luminosity, which was collected by the ATLAS detector. Blue shows the luminosity of the part of data set, which can be used for analyses. EYETS corresponds to an end-of-year extend technical stop [120].

Figure 4.10 shows the integrated luminosity as a function of the mean number of interactions per bunch crossing (μ). The change of the beam conditions, as the bunch spacing or the bunch density, between the different data taking years of Run 2 results in an increase of the average

$\langle \mu \rangle$ from 2015 to 2017. In 2017 the beam conditions got additional change during the data taking period. This results in a two peak structure of the distribution. For the year 2018 a intermediate setting of the two condition used in 2017 was used, so that $\langle \mu \rangle$ is slightly decreased from 2017 to 2018 and the double peak structure is removed.

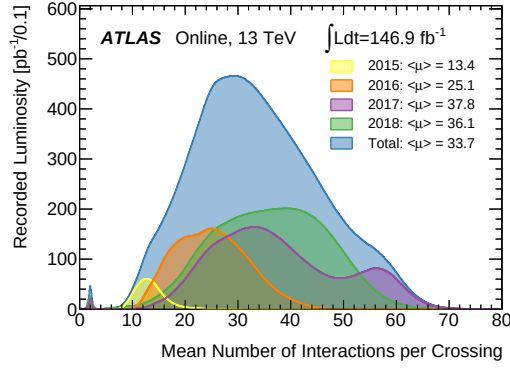


Figure 4.10: Recorded luminosity as a function of the mean number of interactions per crossing (μ) for Run 2. Blue corresponds to distribution for the full Run 2 data set, while the remaining colors corresponds to the different individual years of data taking [120].

4.5 Simulated events

This section summarizes the generators and their configurations used to simulate events for the analyses described in Chapters 8, 9 and 10. Section 11 combines analyses of different Higgs boson decay modes, which can use different generators and configuration. The generators and configuration of these analyses are described in the references summarized in Table 11.1. The includes two $H \rightarrow \tau\tau$ analyses described in Chapter 9 based on the data recorded from 2015 to 2016 and in Chapter 10 based on the data recorded from 2015 to 2018. The second analysis is based on the previous but is optimized. Next to the optimization of the analysis strategy, several generator configuration are modified. The increased sensitivity increases the contribution in the 2015 to 2018 analysis of small physics processes compared to the 2015 to 2016 analysis, in which they have a negligible impact. Therefore, additional simulation events samples of physics processes with small cross section are considered for the 2015 to 2018 analysis. Table 4.2 summarize the used generators and their configurations for both data sets. The Higgs boson production processes are split in different production modes. For the partial Run 2 analysis with the 2015 and 2016 data set the four main production modes for the $H \rightarrow \tau\tau$ decay channel are simulated. The dominant processes are the gluon gluon fusion ggF with an expected cross section of 48.52 pb, vector boson fusion VBF with 3.78 pb, associated vector boson production VH with 2.25 pb and associated top quark pair production ($t\bar{t}H$) with 0.51 pb.

The ggF production is simulated at next-to-next-to-leading order (NNLO) accuracy in QCD using the POWHEG NNLOPS framework [121–125]. The NNLO accuracy is achieved for

4.5 Simulated events

Process	Generator		PDF set		$\sigma \times \mathcal{BR}$ [pb]	
	ME	PS+UE	ME	PS+UE		
Higgs boson ($m_H = 125.09$ GeV)						
$ggF H \rightarrow \tau\tau$	Powheg	PYTHIA8	PDF4LHC15 NNLO	CTEQ6L1	3.04	N ³ LO QCD + NLO EW
$VBF H \rightarrow \tau\tau$	Powheg	PYTHIA8	PDF4LHC15 NLO	CTEQ6L1	0.24	NNLO QCD + NLO EW
$VH H \rightarrow \tau\tau$	Powheg	PYTHIA8	PDF4LHC15 NLO	CTEQ6L1	0.14	NNLO QCD + NLO EW
$t\bar{t}H H \rightarrow \tau\tau$	MG5_aMc@NLO2.2.2	PYTHIA8	NNPDF30LO	NNPDF23LO	0.05	NLO QCD + NLO EW
$t\bar{t}H H \rightarrow \tau\tau$ (full Run 2)	Powheg	PYTHIA8	NNPDF2.3		0.05	NLO QCD + NLO EW
$tH H \rightarrow \tau\tau$ (full Run 2)	MG5_aMc@NLO2.2.2	PYTHIA8	CT10		0.005	NLO QCD + NLO EW
$b\bar{b}H H \rightarrow \tau\tau$ (full Run 2)	Powheg	PYTHIA8	NNPDF2.3		0.05	NLO QCD + NLO EW
$ggF H \rightarrow WW \rightarrow \ell\nu\ell\nu$	Powheg	PYTHIA8	PDF4LHC15 NNLO	CTEQ6L1	1.12	N ³ LO QCD + NLO EW
$ggF H \rightarrow WW$ (full Run 2)	Powheg	PYTHIA8	PDF4LHC15 NNLO	CTEQ6L1	10.44	N ³ LO QCD + NLO EW
$VBF H \rightarrow WW \rightarrow \ell\nu\ell\nu$	Powheg	PYTHIA8	PDF4LHC15 NLO	CTEQ6L1	0.09	NNLO QCD + NLO EW
$VBF H \rightarrow WW$ (full Run 2)	Powheg	PYTHIA8	PDF4LHC15 NLO	CTEQ6L1	0.81	NNLO QCD + NLO EW
Other						
$Z \rightarrow \ell\ell$		SHERPA2.2.1	NNPDF30NLO		40400	NNLO
$W \rightarrow \ell\nu$		SHERPA2.2.1	NNPDF30NLO		184500	NNLO
$EW Z \rightarrow \ell\ell$		SHERPA2.2.1	NNPDF30NLO		1.91	LO
$EW W \rightarrow \ell\nu$		SHERPA2.2.1	NNPDF30NLO		20.25	LO
Di-boson		SHERPA2.2.1	NNPDF30NLO		130	NLO
$t\bar{t}$	Powheg	PYTHIA6	CT10	CTEQ6L1	2100	NNLO+NNLL
$t\bar{t}$ (full Run 2)	Powheg	PYTHIA8	NNPDF2.3		2100	NNLO+NNLL
single top	Powheg	PYTHIA6	CT10	CTEQ6L1	140	NLO
single top (full Run 2)	Powheg	PYTHIA8	NNPDF2.3		140	NLO

Table 4.2: Overview of MC generators used for the different signal and background processes and their predicted cross sections at $\sqrt{s} = 13$ TeV. Samples marked with (full Run 2) are used only for analysis with the full Run 2 data set and replace the samples listed one row above which are only used for analysis with the partial Run 2 data set from 2015 and 2016.

arbitrary inclusive ggF observable by reweighting the Higgs boson rapidity spectrum in Hj-MiNLO [126–128] to that of HNNLO [129]. The VBF and associated VH production processes are generated using POWHEG-Box v2 [123–125, 130] with the MiNLO [122] approach at next-to-leading order (NLO), interfaced to the PYTHIA 8.186 [131] parton shower model. The PDF4LHC15 [132] parametrization of the parton distribution function (PDF) and the AZNLO [133] tune is used, with the CTEQ6L1 PDF set [134], to model non-perturbative effects for the ggF , VBF and VH production modes. The QED emissions from electroweak (EW) vertices and charged leptons is modeled with PHOTOS++ version 3.52 [135]. The $t\bar{t}H$ process is generated with MG5_aMc@NLO2.2.2 [136] at NLO accuracy in QCD, interfaced to PYTHIA8 for parton showering. The NNPDF23LO [137] PDF set is used in the parton shower with the A14 [138] tune for non-perturbative effects, while the NNPDF30LO [139] PDF set is used in the matrix element.

The ggF process is normalized to a next-to-next-to-next-to-leading order (N³LO) QCD calculation with NLO EW corrections applied [140–143]. The overall normalization of the VBF process is taken from an approximate-NNLO QCD cross section with NLO EW corrections applied [144–146]. The VH process is normalized to cross sections calculated at NNLO in QCD, with NLO EW radiative corrections applied [147–149]. The $t\bar{t}H$ production mode cross section is calculated at NLO in QCD with NLO EW corrections applied [150–155].

For the ggF and VBF production the $H \rightarrow WW \rightarrow \ell\nu\ell\nu$ decay channel is generated in addition to the $H \rightarrow \tau\tau$ final state. The samples are generated separately but with the same configuration as the $H \rightarrow \tau\tau$ samples.

The production of $V + jets$ is simulated with Sherpa 2.2.1 [156] generator using NLO matrix

elements for up to two partons and leading order (LO) matrix element for up to four partons using Comix [157] and Open-Loops [158–161] libraries. They are matched with Sherpa parton shower [162] using MEPS@NLO prescription [163–166]. The NNPDF30NNLO set of PDFs [164] is used and the samples are normalized to NNLO prediction [167]. The electroweak production of W/Z bosons from VBF , W/Z bosons associated with jets and di-boson processes are simulated with the Sherpa 2.2.1 generator [117]. The Comix [157] and the OpenLoops [168] matrix element generators at leading-order are used and merged with the SHERPA parton shower simulation [162] using the ME+PS@NLO prescription [169]. As a PDF set NNPDF30NNLO [137] is used in conjunction with a dedicated parton shower tuning developed by the SHERPA authors.

The $t\bar{t}$ process is generated with POWHEG-BOX v2 [123–125, 170], with the CT10 PDF set in the matrix element calculation. The POWHEG-BOX v1 [171, 172] generator is used for the electroweak top quark production, which uses the 4-flavour scheme for the NLO matrix elements calculation together with the fixed four-flavour PDF set CT10F4 [173]. The top quark spin correlation is preserved for all top processes and for top quarks in the t-channel MadSpin [174] is used for the decay. PYTHIA 6.428 [175] with the CTEQ6L1 PDF set and the corresponding Perugia 2012 tune [176] is used for the parton shower, fragmentation and the underlying event.

A full simulation of the ATLAS detector response [177] using the GEANT4 program [109] is performed for all samples. Pile-up is included by overlaying minimum-bias events simulated with PYTHIA8 using MSTW2008LOPDF [178] and the A2 [179] set of tuned parameters on each generated event.

For the analysis using the full Run 2 data set the following updates and additional samples are used.

The decays and spin correlations for τ -leptons are simulated by PYTHIA8 [180] for samples generated with POWHEG while for samples generated with Sherpa the implementation of the generator itself is used. Both implementations have been validated within the ATLAS collaboration. For PYTHIA8 a comparison to TAUOLA [181] is performed, while for Sherpa is validated in Reference [182].

For the $H \rightarrow \tau\tau$ samples the setup of the $t\bar{t}H$ production mode is updated and an additional sample for the tH production mode is included. The $t\bar{t}H$ production mode is modeled using the POWHEG-BOX v2 generator [123–125, 183, 184], with the NNPDF3.0NLO [139] PDF set at NLO. The sample is interfaced to PYTHIA8.230 [119] using the A14 [138] tune and the NNPDF2.3LO [139] PDF set. The decays of bottom and charm hadrons are performed by EVTGEN v1.6.0 [185]. The production of the tH process is generated with MG5_aMc@NLO2.2.2 [136] interfaced with PYTHIA8 [119] using the A14 [138] tune and the CT10 [173] PDF set for the matrix element.

For the $H \rightarrow WW$ process the ggF and VBF production modes are included now for all final states for the full Run 2 analysis, produced with the same settings as $H \rightarrow WW \rightarrow l\nu l\nu$ decay channel.

The $t\bar{t}$ events are generated using the POWHEG-BOX v2 [123–125, 170] generator at NLO in QCD with the NNPDF3.0NLO [139] PDF set. The samples are interfaced to Pythia8.230 [119] to model the parton shower, fragmentation and underlying event using the A14 tune

[138] and the NNPDF2.3LO PDF set [137]. The decays of bottom and charm hadrons are performed by EvtGen v1.6.0 [185].

Single top s-channel events are produced with the POWHEG-BOX v2 [123–125, 171] generator at NLO in QCD in the five-flavour scheme with the NNPDF3.0NLO [139] PDF set. The events are interfaced to Pythia8.230 [119] using the A14 [138] tune and the NNPDF2.3LO [139] PDF set.

5 RECONSTRUCTION AND IDENTIFICATION OF PHYSICS OBJECTS

In this section the reconstruction and identification techniques used to combine signal of hits and energy depositions measured in the ATLAS detector to physics objects. For the $H \rightarrow \tau\tau$ analyses electrons, muons, jets and the missing transverse energy is reconstructed and identified. Identification and isolation criteria are used differentiate electron and muons from jets described in Sections 5.2 and 5.3. Jets are grouped into three different types by defining τ_h -candidates, b-jet and all remaining jets. τ_h -candidates correspond to jets from τ_h -lepton decaying into pions and are identified with specific identification algorithms described in Section 5.5. B-jet are jet initiated by b-quarks and are identified by so called b-tagging techniques described in Section 5.4.1. The missing transverse energy corresponds to the energy of object which are not detected, as for example neutrinos which penetrates detector without an energy deposit. The reconstruction of the missing transverse energy is described in Section 5.6.

5.1 Tracks and Vertices

The inner detector of ATLAS is used to reconstruct tracks. With these tracks the primary and secondary vertices are reconstructed. The information of the pixel detector, the SCT and the TRT are used in different iterative algorithms to reconstruct the tracks [186, 187].

First clusters are formed out of pixels and strips hits that are on the same sensor of the silicon detectors. These three-dimensional clusters are seeded into the inside-out algorithm. A combinatorial Kalman filter is used to extrapolate the clusters to the outer layers of the inner detector and to build track candidates [186, 188]. Drift-circles around the wires of the TRT are matched to track candidates. One set of seeds can result in multiple track candidates. An ambiguity solver [189] is used to reject candidates that fail certain quality requirements, as for the Run 2 data tracking are at least seven hits in the pixel plus SCT layer, a transverse momentum of $p_T > 400$ MeV and a pseudorapidity of $|\eta| < 2.5$. The remaining tracks are the final output of this inside-out chain. This method is used to find the track of particles originating directly from the hard-scattering process or from short-lived particles. Tracks of particles coming from secondary vertices, which have no hits in the silicon layers are reconstructed using an outside-in algorithm, which starts from the TRT track segments and extrapolates them back to the silicon layer. For this compatible unused clusters in the pixel and SCT are used.

The primary and secondary vertices are determined by looking for intersections of sub-set of these reconstructed tracks, which fulfill tighter selection criteria to suppress misidentified tracks. Misidentified tracks are wrongly identified tracks, which do not correspond to a real particle track in the event. The best position of the vertex is estimated by a vertex-finding algorithm [191], for which the global maximum in the distribution of z-coordinates for reconstructed tracks is used as a seed. Only tracks which are compatible with the expected bunch-crossing region are considered for this. Starting from the vertex candidate an iterative χ^2 -fit is used to estimate the best position of the vertex. Therefore, the input tracks are

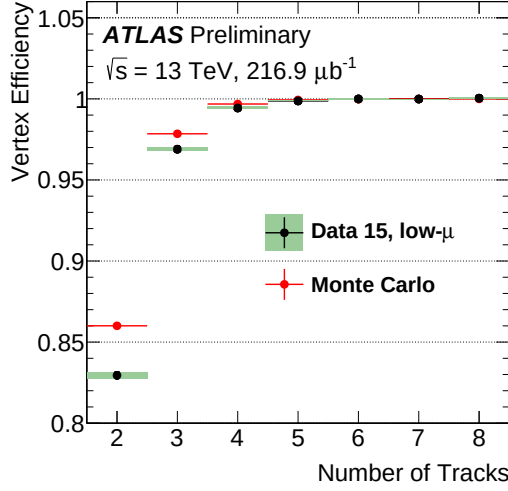


Figure 5.1: Vertex reconstruction efficiency for primary vertex candidates as a function of the number of tracks, for data in green and simulation in red [190].

weighted after each minimization step by their compatibility with the current vertex position. Outliers, which have a $P(\chi^2)$ -value of less than 8% between the vertex candidate and the trajectory are removed. All tracks that are not used for this vertex reconstruction are used to reconstruct an additional vertex. The procedure is repeated until all tracks are used for the reconstruction of one of the vertices or no further vertices can be reconstructed. A minimum of at least two tracks is required for a vertex reconstruction and as Figure 5.1 shows, an efficiency of 80% to 95% depending on the final number of associated tracks is reached [190]. The primary vertex, which is defined as the vertex of the hard scattering process, is assigned to the vertex with the highest momenta, while the remaining vertices are considered as pile-up vertices. The momenta of the vertices are reconstructed as the sum of squared transverse momenta of the assigned tracks is taken to estimate the vertices.

5.2 Electrons

For the reconstruction of electrons energy deposits in the electromagnetic calorimeter are clustered to electron candidates. The sliding-window algorithm [192, 193] is used to build electromagnetic towers of the size of $\Delta\eta \times \Delta\Phi = 0.025 \times 0.025$. A window with the size of 3×5 elements is moved across the electromagnetic tower grid to find local maxima of energy deposits. A minimum transverse energy of 2.5 GeV is required for the clusters. For overlapping clusters only the one with the highest transverse energy is kept. These clusters are matched to tracks in the inner detector using the global χ^2 -track filter [194] and requiring a minimum transverse momentum of 400 MeV on the track. The energy loss of the electrons due to bremsstrahlung in the different detector materials is considered by using a Gaussian-sum filter [195] on the matched tracks. The final cluster is build by extending the window around

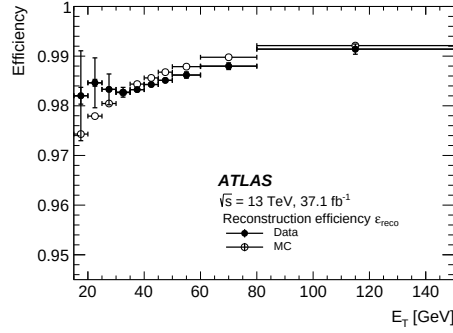


Figure 5.2: The total reconstruction efficiency for electron in data (closed circles) and in simulation (open circles) estimated in $Z \rightarrow ee$ events [193].

the seed cluster to 3×7 in the barrel region and to 5×5 in the end-cap region. The energy of the final cluster is calibrated to the original electron using multivariate analysis techniques [196, 197]. Above a transverse energy of 15 GeV an efficiency above 97% can be achieved for the final electron reconstruction as illustrated in Figure 5.2.

A likelihood method [193] is used to identify the reconstructed candidate as an electron. This method reject jet misidentified as electron and non-prompt electron from photon conversion or from the subsequent decay of heavy-flavor hadrons. The likelihood score gives the probability that a candidate is an electron and uses several information of the inner detector and the calorimeters as input. Different working points are defined using individual thresholds for the likelihood score. The working points are defined as: loose, medium and tight, where loose has the highest signal efficiency and lowest background rejection and tight has the lowest signal efficiency but the highest background rejection. The efficiencies are estimated using a tag-and-probe measurement in $Z \rightarrow ee$ and $J/\psi \rightarrow ee$ events. Figure 5.3 (a) shows the efficiencies as a function of the transverse momentum for different working points with a lower threshold on the transverse momentum at 20 GeV. The signal efficiency varies between 85% and 95%, 75% and 94% and 55% and 90% for the loose, medium and tight working point.

Additionally, an requirement on the isolation of electrons is used to further suppress background from misidentified jets and non-prompt electrons. The isolation is a quantification of the additional amount of energy in a cone of radius ΔR around the electron candidate. The isolation requirements can be grouped into track-based and calorimeter-based criteria. For the track-based method a cone around the electron trajectory is defined. A limitation of the sum of transverse momenta in this cone, excluding the contribution of the electron track and further contribution of bremsstrahlung gives an isolation requirement. The cone size is varied depending on the transverse momenta of the electron, such that the size decrease with larger transverse momentum. This cone size variation takes into account that for high transverse momentum additional particles can be close to the electron. A tag-and-probe measurement in $Z \rightarrow ee$ events defines several working points. Figure 5.3 (b) shows the measured efficiencies of these working points as a function of the transverse energy of the

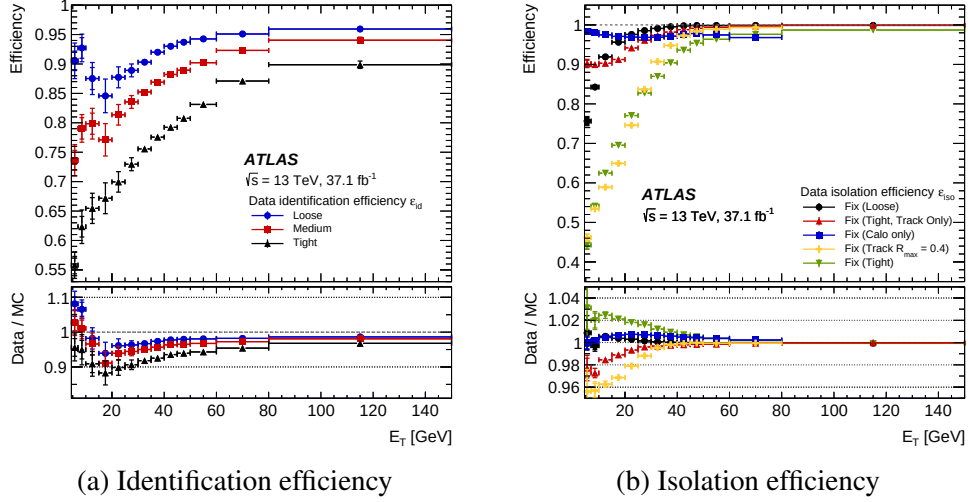


Figure 5.3: (a) The electron identification efficiency for the loose (blue), medium (red) and tight (black) working points as a function of the transverse energy (E_T). (b) The electron isolation efficiency for different work points as a function of E_T [193].

candidates. Topological cluster [198] are used for the calorimeter-based requirements. For this clusters cells are selected, which have four times the energy deposit of the expected noise level, adding iteratively neighbor cells with a energy of at least two times the expected noise level. To remove the included energy of the electron candidate a rectangle with $\Delta\eta \times \Delta\Phi = 0.125 \times 0.175$ around the electron is defined and subtracted. Simulation are used to estimate and subtract the energy contribution of the electron outside this area as well as contribution of pile-up and underlying events [193, 199].

5.3 Muons

The low energy deposit of muons with the detector material results in a clear experimental signature. They leave only a small amount of energy in the calorimeters and penetrate the outermost muon system, in contrast to other physic objects as photons, electron and hadrons. The muon are reconstructed and identified using the combined information of the inner detector, the calorimeters and the muon system.

The tracks of muon candidates are reconstructed following the description in Section 5.1. For the muon system segments of the middle detector layers are used to seed the tracking algorithm and are extrapolated to the inner and outer layers. For the CSC hit combinations in the η - and Φ -plane are used as segments, while a straight line fit on hits is used for the MDT. A global χ^2 -fit is performed to combine the segments from the different layers to muon track candidates. Selection criteria based on the quality of the fit and the number of hits are used to identify a good muon track. The reconstruction of the muon candidates uses one out of four different algorithms depending on the available information from the different detector

components [200]. The different algorithms result in four different muon types: Calorimeter-tagged (CT), segment-tagged (ST), combined (CB) and extrapolated (ME) muons. CT muons are used in the low η -region with $|\eta| < 0.1$. In this region services are located and the muon system is not fully instrumented. For these muons a track has to be matched with an energy deposit in the calorimeter compatible with the energy loss of a minimum-ionizing particle. ST muons candidates have only hits in one layer of the muon system chambers, caused by a low transverse momentum of the muon or a low acceptance of the region. One inner detector track candidate has to be matched with at least one segment in the MDT or CSC for this type. The reconstruction of CB muons uses hit information of the inner detector and the muon system. To combine those the muon system track is extrapolated and matched to the track of the inner detector. The final muon is estimated by a global refit of the hits, which takes into account the energy loss in the calorimeter. Similarly, a second reconstruction approach can be used, which extrapolates the inner detector tracks to the muon system. ME muons are reconstructed for candidates without information from the inner detector. For this type hits in at least two to three layers of the muon system chambers depending on the region are required. Additionally, the reconstructed track has to be compatible with an interaction point, to reduce cosmic muons. To avoid the double use of inner detector tracks, the muon types are prioritized in the following order CB, ST and CT. In case of an overlap with an additional ME muon, the fit quality of the track reconstruction and the number of hits are used to select the final muon. The calibration of the transverse momentum of the muon takes into account several momentum scale and resolution effects. It is performed separately for tracks of the inner detector and of the muon system. A limited knowledge of the magnetic field and the energy loss in the detector affects the muon scale. Inhomogeneities of the magnetic field, misalignment in the muon system, the spatial resolution of the hit measurement and multiple scattering spread the momentum resolution. Special runs without the magnetic field are used to determine the misalignment, while all remaining effects are estimated data-driven by performing likelihood fits in the invariant mass of $J/\Psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ events. To suppress the fraction of backgrounds from other objects misidentified as muon additional identification criteria are applied. Three different working points are defined: loose, medium and tight.

For the loose criteria all muon types are considered. CB muons are required to have at least three hits in at least two MDT or CSC layers except for the $|\eta| < 0.1$ region in which the requirement is relaxed to one hit. CT and ST muons are only considered for the $|\eta| < 0.1$ region. Since the momentum measurement is expected to differ between the inner detector and the muon system for non-prompt muons, the q/p significance is required to be less than seven for CB muons. The q/p significance corresponds to the absolute difference between the ratio of the charge and momentum measured in the inner detector and the muon system divided by the sum in quadrature of their uncertainties. The requirement to have hits in at least three MDT or CSC layers and the restriction of $2.5 < |\eta| < 2.7$ is applied on ME muons. Only CB and ME muons that fulfill the loose requirement are considered for the medium working point, while only CB muons pass the tight working point. In addition, muons of the tight working point need to have hits in at least two layers of the muon system and the χ^2 -value of the combined track fit has to be smaller than eight. To further increase the

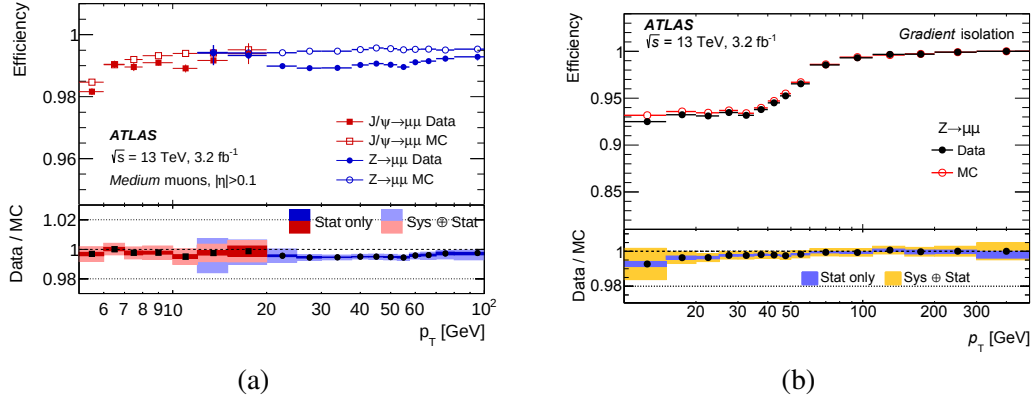


Figure 5.4: The combined reconstruction and identification efficiency for medium muons (a) and the muon gradient isolation efficiency (b) as a function of the transverse momentum (p_T) [200].

backgrounds rejection, especially for lower transverse momentum, transverse momentum cuts depending on the q/p significance and ρ' are applied, where ρ' is the absolute value of the difference between the transverse momentum measured by inner detector and the muon system divided by the transverse momentum of the combined track. The combined reconstruction and identification efficiency for the medium working point as a function of the transverse momentum of the muon is given in Figure 5.4 (a). The obtained efficiencies for $J/\Psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ events are above 98% for the full transverse momentum range. Similar to the definition for electrons calorimeter- and track-based isolation variables are defined for muons. For the calorimeter-based isolation the sum of the energies of all topological clusters in a cone of size ΔR around the muon position is used. The energy in the cone is corrected for the energy of the muon itself and for contributions from pile-up. The track-based isolation uses the sum of transverse momenta of the tracks inside a cone of size ΔR around the muon position, excluding the muon tracks. To take into account that for large transverse momenta other object can be close to the muon a variable cone size is used, which decrease with increasing momenta: $\Delta R = \min(10 \text{ GeV}/p_T, R_{\max})$, with maximum cone size $R_{\max}$. Figure 5.4 (b) shows the gradient isolation working point used for the analysis described in Chapter 9. For the gradient isolation working point the relative isolation variable, defined as the ratio between the calorimeter (track-based) isolation variables and the transverse momentum of the muon are used to reach a certain efficiency. $\Delta R = 0.2$ is used for the calorimeter-based isolation and $R_{\max} = 0.3$ for the track-based isolation. The efficiency of the working point is around 93% below a transverse momentum of the muon of 40 GeV and increases for higher values up to 99% above 80 GeV (see Figure 5.4 (b)).

5.4 Jets

As described in Section 4.3, quarks and gluons produced in proton-proton collision hadronize to a collimated spray of color-neutral hadrons called jets. Pile-up, underlying events and the large cross section of QCD interactions result in a large amount of jets. Therefore, a solid and sophisticated reconstruction of jets is crucial for many analyses.

The reconstruction of jets is based on topological clusters [201] built in the electromagnetic and hadronic calorimeters. Energy cells with a significant energy contribution of at least four times the total energy noise are used to seed the cluster finding algorithm [190, 202]. For the total energy, noise contributions from the readout electronic as well as pile-up noise is estimated via simulation [203]. The topological clusters are built iteratively around these cells by adding neighboring cells which fulfill the requirement of at least two times the noise energy. Overlapping topological clusters are merged together, to avoid the double use of a cell in two clusters. In a final step, all neighboring cells of the cluster with a positive energy deposit are added, so that the final cluster has a highly significant cells surrounded by cells with lower significance. The energy of the constructed cluster is defined as the sum of all included cells and is measured at the electromagnetic scale. A weighted average of the η - and Φ -coordinates of the individual cells is used to reconstruct the direction of the cluster. The resulting four momentum vectors are used as an input for the jet finding algorithms. For the analysis using the full Run 2 data-set recorded from 2015-2018 the optimized par-

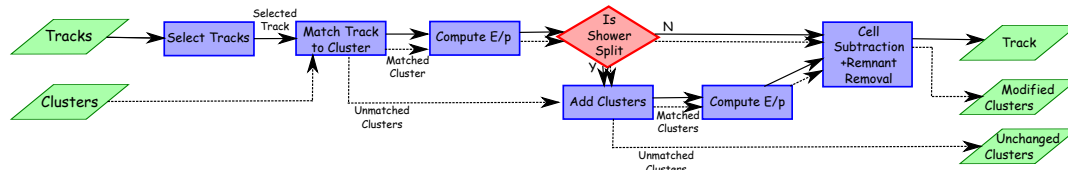


Figure 5.5: Schematic overview of the particle flow algorithm [204].

ticle flow algorithms is used [204]. A cell-based subtraction to remove overlaps between momentum and energy measurements of the inner detector and the calorimeters is performed. Figure 5.5 shows a schematic overview of the particle flow algorithm. First, the tracks of the inner detector are matched to a single topological cluster in the calorimeter. Based on the cluster position and the track momentum the expected energy deposit of the particle, which creates the track, is calculated. Most of the particles deposit energy in multiple clusters. The algorithm evaluates the probability that the particle deposit energy in multiple clusters for each of the track-cluster systems. These probabilities are used to decide if additional clusters has to be added to the system to recover the full shower energy. The expected energy is then subtracted cell by cell from the clusters. The cluster remains are removed if the remaining energy of the system is consistent with the expected shower fluctuations of a single particle. The algorithm is applied iteratively to all tracks, sorted in descending transverse momentum order. In this way, the tracking and calorimeter information is combined to reconstruct jets and soft activity in the event. The soft activity is an important input to calculate the missing transverse momentum described in the Section 5.6. The particle flow algorithms provides

a list of tracks as well as a list of topological-clusters containing unmodified clusters and modified clusters after the energy subtraction.

The list of topological jet candidates as well as the particle flow jet candidates is used separately as an input for the anti- k_t algorithms [205] implemented in FASTJET [206] to reconstruct the final jets for each method. The distances measures between two jet candidates d_{ij} and one candidate and the beamline d_{iB} are defined as:

$$d_{ij} = \min \left(\frac{1}{p_{T,i}^2}, \frac{1}{p_{T,j}^2} \right) \cdot \frac{\Delta R_{ij}^2}{R^2} \quad \text{and} \quad d_{iB} = \frac{1}{p_{T,i}^2}, \quad (5.1)$$

with the transverse momentum of the particle $p_{T,i}$ and $p_{T,j}$ and the geometrical distance ΔR_{ij} between them. The radius R defines the size of the jets and is set to 0.4. The reconstruction is performed iteratively. For all particles d_{ij} and d_{iB} are calculated. If d_{iB} is the smallest value, particle i is defined as a jet and removed from the input list, while if one of the d_{ij} is the smallest value, the corresponding two particles are replaced by a new merged particle with $p_{T,ij} = p_{T,i} + p_{T,j}$. The reconstruction stops when no input particle is left.

The reconstructed jets are corrected to restore the jet energy scale to the energy of a particle-level jet built from truth particles after the hadronization [207]. The jet four-momenta is recalculated first to point to the vertex of the hard-scatter process while keeping the energy fixed. Then the energy is corrected by subtracting contributions from pile-up. The energy is corrected from the reconstructed energy to the energy of the particle-level jet and for biases in η due to transitions between different calorimeters or between layers with different granularity. The energy correction between the reconstructed and particle-level jet include corrections for the flavor of the parton to minimize differences introduced by the differences in the fragmentation process of quarks and gluons. To account for differences between data and simulation, correction factors derived in in-situ measurements are used. Figure 5.6 shows the Φ and η -resolution for reconstructed jets using the topological cluster and for jets using the particle flow algorithms as a function of the transverse momentum of the jet p_T and the transverse momentum resolution as a function of p_T and as a function of $|\eta|$. The particle flow algorithms improves the resolution in the low transverse momentum and low $|\eta|$ range, while at high transverse momentum and $|\eta|$ both algorithms reach a similar resolution.

5.4.1 B-jet tagging

Jets initiated by b-quarks are identified by so called b-tagging techniques. The long lifetime of hadrons containing b-quarks results in a secondary vertex, which is displaced from the primary vertex by a few mm. This property is used by the b-tagging algorithm. For the analyses using the partial Run 2 data-set the MV2c20 algorithm [208, 209] is applied, while analyses based on the full Run 2 data-set use the DL1r algorithm [210, 211]. The MV2c20 algorithm identifies b-jet with a multivariate discriminant, which combines complimentary information from different algorithms: IP2D [209], IP3D [209], vertex-based SV1 [212] and JetFitter [213]. IP2D and IP3D are based on the transversal and longitudinal impact-parameters d_0 and z_0 . d_0 is defined as the closest distance between a given track and the primary vertex in the $r - \Phi$ plane, while z_0 is the difference between the z -coordinate of this point and the primary

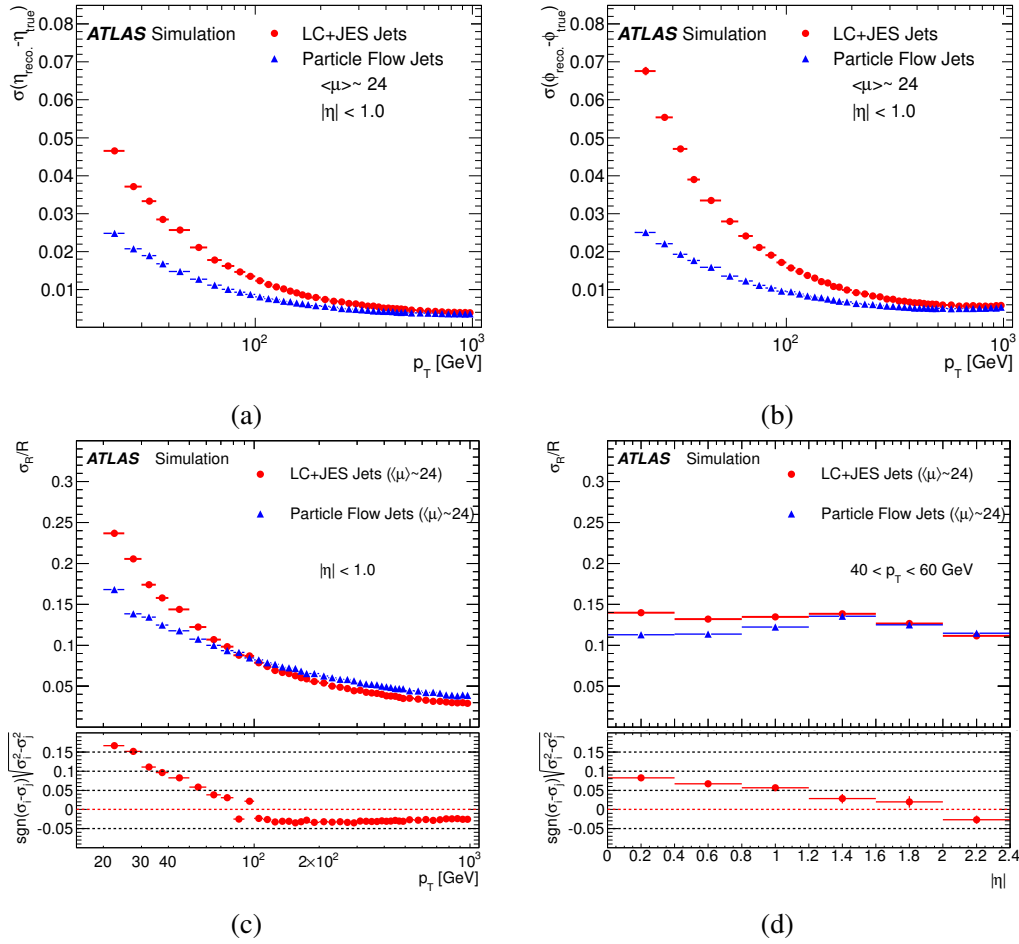


Figure 5.6: The angular resolution of reconstructed jets in η (a) and in Φ (b), as a function of the transverse momentum of the jet p_T and the jet transverse momentum resolution as function of p_T for $|\eta| < 1.0$ (c) and as a function of $|\eta|$ (d), estimated in di-jet simulation [204].

vertex. Those parameters are generally larger for tracks originating from a secondary vertex compared to tracks from the primary vertex. The DL1 is a deep feed-forward neural network trained on the same three input algorithms. For both methods, different working point are defined, which use different cuts on the final discriminant. Tighter working points have a better background suppression but a lower efficiency. Therefore, the different analysis of this thesis use individual working points depending on the studied phase space. Figure 5.7 shows for the 70% single-cut working point a comparison between the efficiency to tag b-jets and the rejection for other jet types for the different methods. The DL1r method has a slightly lower efficiency as the MV2c20 method, but the rejection of other jets is higher especially for jets with a transverse momentum below 500 GeV. Therefore, the DL1r results in a better b-jet

to background ratio for analyses with a significant contribution of jets below 500 GeV as the $H \rightarrow \tau\tau$ analysis.

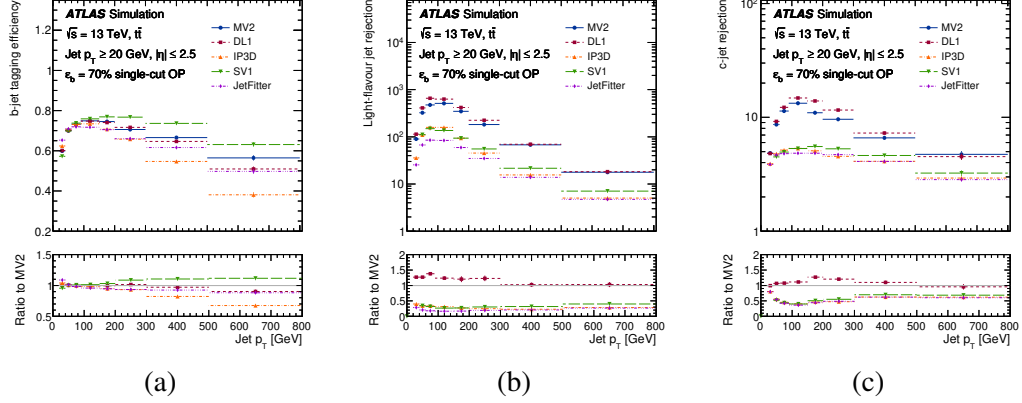


Figure 5.7: The b-jet tagging efficiency (a), light-flavour jet rejection (b) and c-jet rejection (c) as a function of the transverse momentum of the jet based on simulated $t\bar{t}$ events for the IP3D, SV1, JetFitter, MV2 and DL1 b-tagging algorithms at the 70% single-cut working point [210].

5.5 Hadronically decaying τ -leptons

Due to their relative short lifetime τ -leptons are not detected as final state objects (see Section 6.1.3). Instead, they are reconstructed based on their visible decay products, excluding for example neutrinos produced in the decay. 65% of the τ -leptons decay hadronically. Similar to other processes including hardons this decay will be reconstructed as a jet. In all analyses discussed in this thesis, the signal includes hadronically decaying τ -leptons and therefore it is necessary to identify these τ -jets. The decay can be classified into processes with one charged track (1-prong) with a branching ratio of 72% and three charged tracks (3-prong) with a branching ratio of 22%. Most of the remaining decays have a kaon in the final state. Up to three neutral particle are produced in 68% of the decays. The decay products of the τ -lepton are boosted, so that the τ -jets tend to be highly collimated and consists of a smaller amount of particles compared to quark/gluon initiated jets.

To reconstruct the τ -lepton an algorithm is seeded with all reconstructed and with the LCW scheme [215] calibrated jets with a transverse momentum of at least 10 GeV and $|\eta| < 2.5$ [202]. Candidates in the transition region with $1.37 < |\eta| < 1.52$ are excluded. The number of charged tracks found within a cone of $\Delta R < 0.2$ around the jet direction is counted to classify the decay into 1-prong and 3-prong decays. The reconstructed τ -candidate must not originate from the selected primary vertex for events with multiple interaction points. The production vertex of the τ -candidate (TV) is identified among the event vertices to increase the reconstruction efficiency. The transverse momenta of all tracks within a cone of $\Delta R < 0.2$ around the seed jet direction with $p_T > 1$ GeV and satisfying requirements on the number of

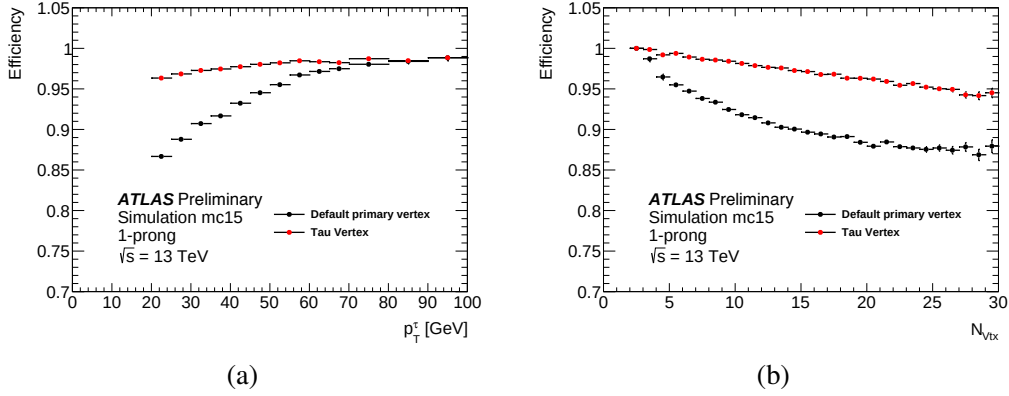


Figure 5.8: Efficiency to assign the correct vertex in 1-prong τ -lepton decays for the τ -reconstruction algorithm (red) and the default selected primary vertex (black) as a function of p_T^τ (a) and the number of reconstructed vertices in the event N_{vtx} (b) [214].

hits in the inner detector are summed and the TV is defined as the vertex to which the largest fraction of this sum is matched. Using the TV instead of the primary vertex improves the efficiency to assign the correct vertex. In particular, for events with a low p_T of the τ -lepton and for events with a higher number of reconstructed vertices the efficiency is improved, as shown in Figure 5.8.

The same cone of $\Delta R < 0.2$ is used to calculate the direction to the τ -candidate. Here the

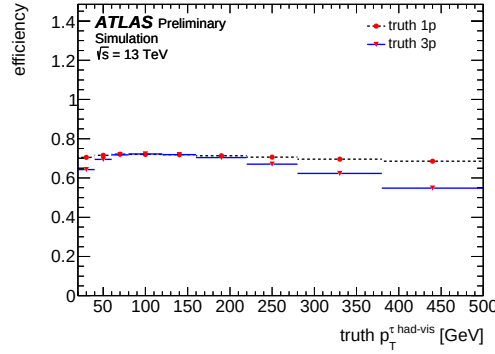


Figure 5.9: Efficiency for reconstruction of a τ -candidate with 1-prong (red point, black dashed line) and 3-prong (red triangle, blue solid line) as a function of the transverse momentum of the τ -candidate [214].

TV is used as an origin and the vectorial sum of the topological cluster in the cone is used assuming that the τ -candidate is massless. The efficiencies for reconstructing a τ -candidate with 1-prong and 3-prong are shown in Figure 5.9. The efficiencies of 1-prong candidates are higher as for 3-prong candidates, especially for small values (< 70 GeV) and for large values

(>220 GeV) of the transverse momentum of the candidate. For 1-prong events, the efficiency is relatively flat with respect to the transverse momentum with approximately 70%, while for 3-prong the efficiency varies between approximately 55 and 70%.

The reconstructed energy of the τ -candidate is calibrated back to the energy of the visible τ -lepton at particle level. The energy contribution from pile-up is subtracted and the detector response calibration (DRC) function is applied. DRC corrects for decay products that do not reach the calorimeter, so that the energy of them is not high enough to create a topological cluster or are outside the $\Delta R < 0.2$ cone. This correction is calculated based on simulated events using the particle level visible energy.

A boosted decision tree (BDT) is trained to identify τ -leptons and to suppress the contributions

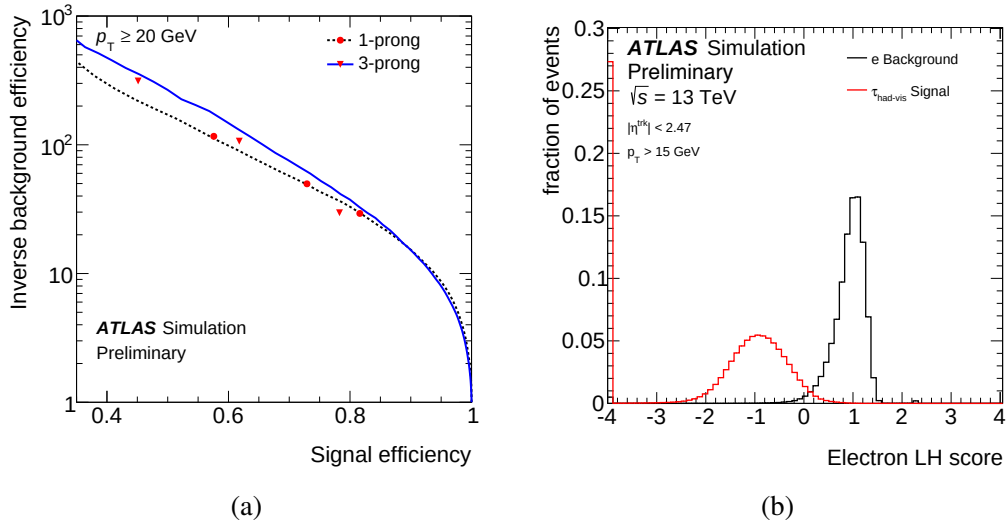


Figure 5.10: Inverse of misidentifying a QCD multi-jet as a function of the efficiency for the τ -lepton identification (a) and the electron likelihood score for τ -candidates (red) and prompt electrons (black) estimated in $Z \rightarrow ee$ events (b). The efficiency for events with 1-prong candidates are shown as a blue solid line, while for events with 3-prong candidates the efficiency is displayed with a dashed black line. The three working point loose, medium and tight are visualized with red points for 1-prong and with red rectangles for 3-prong events. [214].

of quark and gluons initiated jets, which mimic a τ -lepton signature [216]. Separated BDTs are trained for 1-prong and 3-prong decays against di-jet background events using the shower shape and tracking information. As an input, variables of the tracks and topological clusters inside the $\Delta R < 0.2$ cone are used. The output scores of the BDTs are then used as an identification criteria by defining three different working points: loose, medium and tight. Each working point corresponds to a different identification efficiency and is defined to not depend on the transverse momentum of the τ -candidate. The continuous identification efficiency and the three working points are shown in Figure 5.10 (a) for 1-prong and 3-prong

τ -candidates. The working points of 1-prong τ -leptons are defined closer to each other as for 3-prong events. The loose working point for 1-prong events has a similar efficiency as the medium working point for 3-prong events, while the two tight working points are defined with a similar efficiency. Especially for 1-prong τ -candidates, electrons with similar detector signatures as charged pions are an important background. The contribution of these events is suppressed using a likelihood-based rejection algorithm based on a discriminator developed for electron identification [217, 218], which is shown in Figure 5.10 (b).

5.6 Missing transverse energy

The missing transverse energy E_T^{miss} can be grouped into real and fake contributions. Real E_T^{miss} is created by particles, which do not interact with the detector, as the neutrinos of the SM but also potential beyond the SM particles interacting only via the weak interaction. The fake contribution can be associated with energy mismeasurements, particles which penetrate the detector in regions with low acceptance, pile-up or particles which are not detected because of the limited reconstruction efficiency of the detector.

The missing transverse energy can be reconstructed by using the measurement of all detected

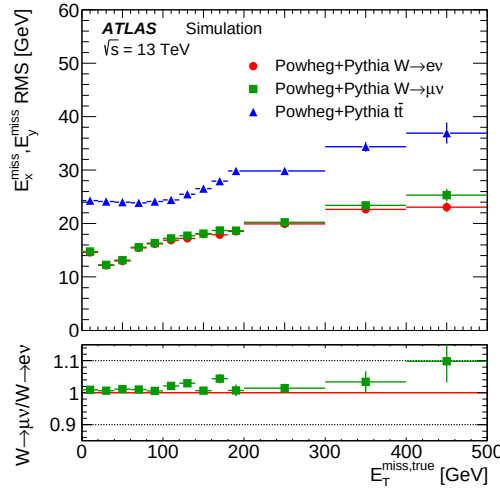


Figure 5.11: Resolution of E_T^{miss} measured by $\text{RMS}_{x(y)}^{\text{miss}}$ as a function of the particle level transverse momentum $E_T^{\text{miss, true}}$ for $W \rightarrow e\nu$ (red), $W \rightarrow \mu\nu$ (green) and $t\bar{t}$ (blue) events taken from simulation [219].

particles and momentum conservation transverse to the beam line [219, 220]. All tracks from reconstructed particles but also non-matched tracks originating from the primary vertex are considered. The vector of the missing transverse momentum is defined as the negative vectorial sum of the transverse momenta $\vec{p}_T$ of all these objects, where the contribution of reconstructed objects is called hard term, while the remaining objects are summarized as the soft term. The soft terms covers underlying event activity, soft radiations from the hard

5.7 Invariant mass of di- τ system

interaction and contribution from pile-up events. Finally, the vectorial missing transverse energy is the sum of the hard and soft term:

$$\vec{E}_T^{\text{miss}} = - \underbrace{\sum_{i \in \text{obj}} \vec{p}_{T,i}}_{\text{hard term}} - \underbrace{\sum_{j \in \text{track}} \vec{p}_{T,j}}_{\text{soft term}}, \quad (5.2)$$

with i running over all reconstructed objects and j over all tracks which are not associated with an object. $E_{T^{\text{miss}}}$ is then defined as the magnitude of this vector:

$$E_T^{\text{miss}} = |\vec{E}_T^{\text{miss}}| = \sqrt{(E_x^{\text{miss}})^2 + (E_y^{\text{miss}})^2}. \quad (5.3)$$

A priority order for overlapping objects is used, to avoid double counting of energy deposits in the hard term. The order is electrons, photons, hadronically decaying τ -leptons and jets. Therefore, all electrons are used to reconstruct the missing transverse energy, while other objects are only used if they do not overlay with any already included object. Some muons, which are typically reconstructed with the MS- and ID-tracks only (see Section 5.3), can have large deposits in the calorimeters and therefore be reconstructed as jets. For these cases, but also for jets which overlap only marginal with other objects, a corrected energy estimation is used, which is described in Reference [219]. Figure 5.11 shows the resolution of the missing transverse energy estimated in simulated $W \rightarrow e\nu$, $W \rightarrow \mu\nu$ and $t\bar{t}$ events. The best resolution is achieved in $W \rightarrow e\nu$ events, followed by a slightly higher resolution of $W \rightarrow \mu\nu$ events, while the resolution for $t\bar{t}$ events is approximately 150% of the resolution estimated for the other processes.

5.7 Invariant mass of di- τ system

The invariant mass of the di- τ system is an important discriminant in $H \rightarrow \tau\tau$ analyses. The mass is able to distinguish between the irreducible background from $Z \rightarrow \tau\tau$ events and the signal. The mass of the instable di- τ system (see Section 5.5) is reconstructed from the decay products. The decay of τ -lepton includes at least one undetected neutrino. These invisible decay product has to be considered in the mass reconstruction of the Higgs boson candidate. The two approaches used in this thesis to reconstruct the invariant mass of the di- τ system are described in this section.

Visible mass The visible mass approach uses only the visible decay products of the τ -leptons to calculate the invariant mass of the di- τ -system, so that the contributions of the missing energy are neglected. This results in a shift of the invariant mass towards lower values. The visible mass of the di- τ -system $m_{\tau\tau}^{\text{vis}}$ is defined as:

$$m_{\tau\tau}^{\text{vis}} = \sqrt{(E_{\tau_{1,\text{vis}}} + E_{\tau_{2,\text{vis}}})^2 - (\vec{p}_{\tau_{1,\text{vis}}} + \vec{p}_{\tau_{2,\text{vis}}})^2}, \quad (5.4)$$

with the reconstructed energy of the two visible τ -leptons $E_{\tau_{1(2),\text{vis}}}$ and their reconstructed visible momentum $\vec{p}_{\tau_{1(2),\text{vis}}}$.

Collinear approximation The collinear approximation described in Reference [221] uses the assumption that the invisible decay products are emitted in the same direction as the visible system and that these decay products are the only source of missing transverse momentum in the event. Using the measurement of the missing transverse energy $\vec{E}_T^{\text{miss}} = (E_x^{\text{miss}}, E_y^{\text{miss}})$ and the transverse momentum of the visible decay products of the τ -leptons $\vec{p}_{T,1(2)}^{\text{vis}} = (p_{x,1(2)}^{\text{vis}}, p_{y,1(2)}^{\text{vis}})$ the fraction of τ -momenta carried by the visible decay products can be expressed as:

$$x_{1,(2)} = \frac{p_{x,2}^{\text{vis}} p_{y,1}^{\text{vis}} - p_{y,2}^{\text{vis}} p_{x,1}^{\text{vis}}}{p_{x,2}^{\text{vis}} p_{y,1}^{\text{vis}} + (-) E_x^{\text{miss}} p_{x,1(2)}^{\text{vis}} - p_{y,2}^{\text{vis}} p_{x,1}^{\text{vis}} - (+) E_y^{\text{miss}} p_{x,1(2)}^{\text{vis}}}. \quad (5.5)$$

The collinear mass of the di- τ system can be constructed using the assumption of massless leptons with the invariant mass of the visible system m_{vis} and the momentum fractions $x_{1,(2)}$:

$$m_{\tau\tau}^{\text{coll}} = \frac{m_{\text{vis}}}{\sqrt{x_1 x_2}}. \quad (5.6)$$

The method provides a good resolution for boosted topologies, while for topologies in which the collinear approximation is not valid the invariant mass is overestimated. For back-to-back topologies of the decay products ($\vec{p}_{T,1}^{\text{vis}} \approx -\vec{p}_{T,2}^{\text{vis}}$) Equation 5.5 becomes degenerate, resulting in unphysical values for $x_{1,(2)}$ and too high values of the invariant mass. Figure 5.12 shows the normalized distribution of the reconstructed di- τ -mass in the collinear approximation for $H \rightarrow \tau\tau$ and $Z \rightarrow \tau\tau$ events in the boosted and VBF inclusive signal categories of the $H \rightarrow \tau\tau$ analysis based on the 2015 and 2016 data set described in Chapter 9. The collinear approximation is able to reconstruct the Z boson mass to 100.2 GeV in the boosted and to 99.7 GeV in the VBF category, while the reconstructed masses for the Higgs boson are 126.3 GeV and 124.7 GeV, respectively. The values in the boosted category are slightly higher than in the VBF category for both processes showing a systematic impact of the event selection on the collinear assumption. The resolution for the Higgs boson mass is slightly smaller in the VBF category with 18.9 GeV than in the boosted category with 21.1 GeV, while the resolution for the Z boson mass are very similar with 20.5 GeV in the VBF category and 19.6 GeV in the boosted category.

Missing mass calculator The missing mass calculator (MMC) [223] can be used in any di- τ -decay topology, since it does not use the collinear assumption. The method estimates the most probable kinematic configuration of the invisible system by using probability density functions for the τ -decay properties. By assuming a perfect detector resolution and no additional neutrinos in the event, the full reconstruction of the event topology requires the determination between six and eight unknown variables depending on the decay mode of the di- τ system. These observables are the components of the invisible momenta $p_{1(2)}^{\text{miss}}$ of the two τ -neutrinos and the invariant mass of the neutrino system for each leptonically decaying τ -lepton ($m_{1(2)}^{\text{miss}}$). To estimate the variables, mass-shell requirements can be used, which result

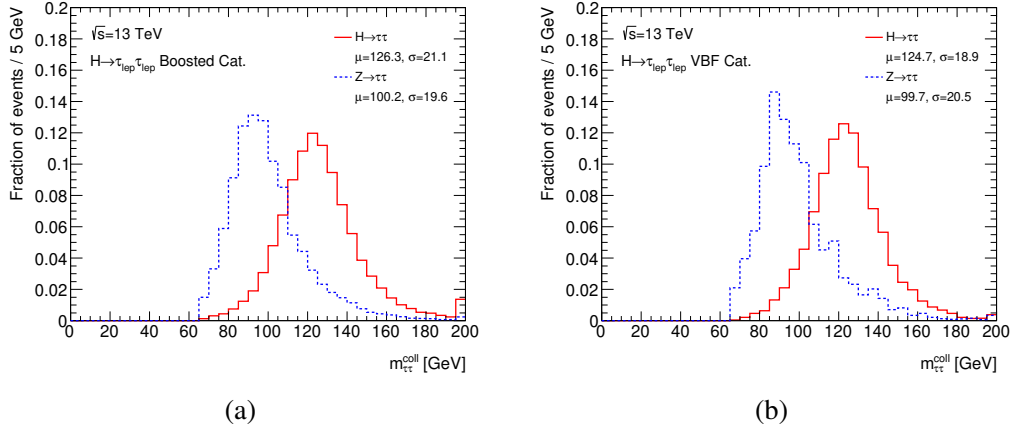


Figure 5.12: Normalized distribution of the reconstructed di- τ -mass in the collinear approximation for $H \rightarrow \tau\tau$ and $Z \rightarrow \tau\tau$ events in the boosted inclusive (a) and VBF inclusive (b) signal categories of the $H \rightarrow \tau\tau$ analysis based on the 2015 and 2016 data set (see Chapter 9) [222].

into four equations:

$$\begin{aligned}
 E_x^{\text{miss}} &= p_1^{\text{miss}} \sin(\theta_1^{\text{miss}}) \cos(\phi_1^{\text{miss}}) + p_2^{\text{miss}} \sin(\theta_2^{\text{miss}}) \cos(\phi_2^{\text{miss}}) \\
 E_y^{\text{miss}} &= p_1^{\text{miss}} \sin(\theta_1^{\text{miss}}) \sin(\phi_1^{\text{miss}}) + p_2^{\text{miss}} \sin(\theta_2^{\text{miss}}) \sin(\phi_2^{\text{miss}}) \\
 m_{\tau_1}^2 &= (m_1^{\text{miss}})^2 + (m_1^{\text{vis}})^2 + 2\sqrt{(p_1^{\text{miss}})^2 + (m_1^{\text{miss}})^2} \sqrt{(p_1^{\text{vis}})^2 + (m_1^{\text{vis}})^2} \\
 &\quad - 2p_1^{\text{vis}} p_1^{\text{miss}} \cos(\theta_1^{\text{vis}} - \theta_1^{\text{miss}}) \\
 m_{\tau_2}^2 &= (m_2^{\text{miss}})^2 + (m_2^{\text{vis}})^2 + 2\sqrt{(p_2^{\text{miss}})^2 + (m_2^{\text{miss}})^2} \sqrt{(p_2^{\text{vis}})^2 + (m_2^{\text{vis}})^2} \\
 &\quad - 2p_2^{\text{vis}} p_2^{\text{miss}} \cos(\theta_2^{\text{vis}} - \theta_2^{\text{miss}}).
 \end{aligned} \tag{5.7}$$

For hadronically decaying τ -leptons only one massless neutrino is produced in the final state, so that in Equation 5.7 $m_{1(2)}^{\text{miss}}$ is set to zero. Since these requirements are not sufficient to estimate all unknown variables, probability density functions (PDFs) depending on the τ -decay kinematics are used by the MMC algorithm to find the best solution. Simulations are used to estimate these PDFs and a scan over the unknown variables of the neutrino system is performed. The resolution of the missing transverse momentum has an impact on the performance of the MMC. To take this dependency into account, additional scans over $(E_x^{\text{miss}}, E_y^{\text{miss}})$ are performed and each scan point is weighted with a Gaussian probability depending on the scalar sum of all transverse energy deposits in the calorimeter. For each scan point of the MMC algorithm Equation 5.7 is solved and the solution is weighted according to the event likelihood calculated for these τ -decay kinematics. The solution corresponding to

the maximum of the weighted solutions for each scan is used as the final estimate for the di- τ mass $m_{\tau\tau}^{\text{MMC}}$. In rare cases between 1 and 3% of events with large fluctuations in the missing energy measurement the MMC algorithms can fail to find a solution.

Table 5.1 compares the reconstructed di- τ -mass of the MMC method with $m_{\tau\tau}^{\text{vis}}$ and $m_{\tau\tau}^{\text{coll}}$ for

	$m_{H \rightarrow \tau\tau}$ [GeV]	$\sigma_{H \rightarrow \tau\tau}$ [GeV]	$m_{Z \rightarrow \tau\tau}$ [GeV]	$\sigma_{Z \rightarrow \tau\tau}$ [GeV]	$\frac{m_{H \rightarrow \tau\tau} - m_{Z \rightarrow \tau\tau}}{\sqrt{\sigma_{H \rightarrow \tau\tau}^2 + \sigma_{Z \rightarrow \tau\tau}^2}}$
$m_{\tau\tau}^{\text{vis}}$	74.9	15.8	59.3	14.6	0.73
$m_{\tau\tau}^{\text{coll}}$	125.6	20.2	98.3	21.7	0.92
$m_{\tau\tau}^{\text{MMC}}$	123.5	16.2	96.7	17.5	1.04

Table 5.1: Mean and RMS values for the mass of the di- τ -system for the $H \rightarrow \tau\tau$ and $Z \rightarrow \tau\tau$ process using the visible mass, collinear approximation and the MMC method in the $H \rightarrow \tau_l \tau_h$ final state preselection of the analysis based on the 2015 and 2016 data set (see Chapter 9). The right column shows the separation power of the two processes [224].

the $H \rightarrow \tau\tau$ and $Z \rightarrow \tau\tau$ processes in the $H \rightarrow \tau_l \tau_h$ final state preselection of the analysis based on the 2015 and 2016 data set (see Chapter 9). The table shows the clear bias of the visible mass towards lower values due to the neglected contribution of the missing transverse energy, while the values of $m_{\tau\tau}^{\text{coll}}$ and $m_{\tau\tau}^{\text{MMC}}$ are relatively close to each other. For both processes, the best resolution is achieved for $m_{\tau\tau}^{\text{vis}}$, since the resolution does not consider the systematic shift of the mean value. The resolution of $m_{\tau\tau}^{\text{MMC}}$ is smaller for both processes compared to $m_{\tau\tau}^{\text{coll}}$. The highest separation power of the two processes with 1.04 is achieved for $m_{\tau\tau}^{\text{MMC}}$, followed by $m_{\tau\tau}^{\text{coll}}$ with 0.92 and $m_{\tau\tau}^{\text{vis}}$ with 0.73. Figure 5.13 shows the corresponding normalized

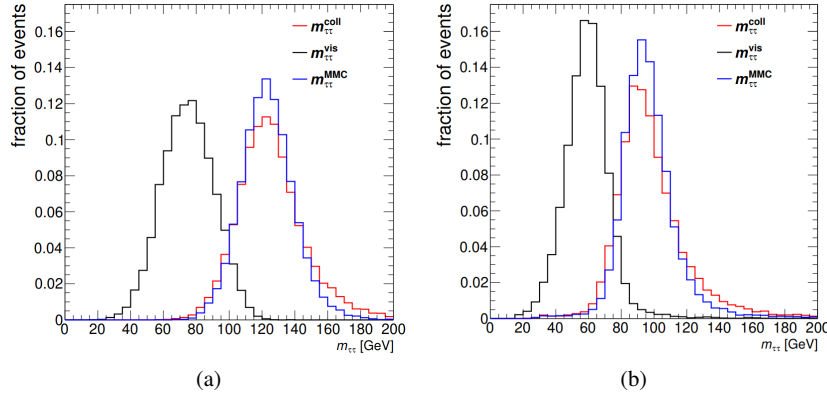


Figure 5.13: Normalized distribution of the reconstructed di- τ -mass using the visible mass (black), the collinear approximation (red) and the MMC method (blue) for $H \rightarrow \tau\tau$ (a) and $Z \rightarrow \tau\tau$ (b) events in the $H \rightarrow \tau_l \tau_h$ preselection of the analysis based on the 2015 and 2016 data set (see Chapter 9) [224].

contributions of different mass approximations summarized in Table 5.1.

6 PHYSICS PROCESSES

This section summarizes the physics processes relevant for this thesis. All processes are either included as signal or background processes in the different analyses. The analyses discussed in this thesis compare the SM predictions simulated with Monte Carlo event generators with data recorded by the ATLAS experiment. Therefore, the targeted signal processes but also background processes with similar signature as the signal process have to be considered. The used Monte Carlo event generators and their precision are described in Section 4.5. Background processes can be grouped into reducible and irreducible processes. Reducible backgrounds have a different final state than the signal process, but misidentification of physics objects and limitations in the resolution of the detector e.g. the estimation of the missing transverse energy, can lead to a misidentification of the event as a signal process. This type of background can be suppressed by tightening the event selection to a topology, in which the probability for the misidentification is smaller. Irreducible backgrounds have the same final state signature as the signal process and are therefore difficult to suppress. Here, kinematic properties can be used to differentiate signal and background processes. In the $H \rightarrow \tau\tau$ analysis the $Z \rightarrow \tau\tau$ process is an irreducible background, which can be distinguished from the signal using the reconstructed mass of the di- τ -system.

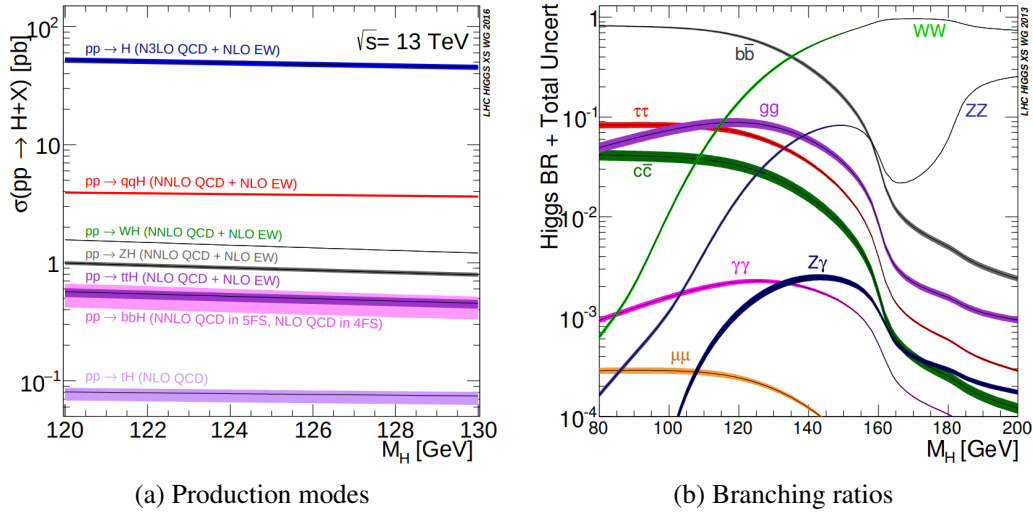


Figure 6.1: Cross sections for the main production modes of the Higgs boson at a center-of-mass energy $\sqrt{s}$ of 13 TeV (a) and branching fraction of the main decay channels of the Higgs boson (b), as a function of the Higgs boson mass M_H [27].

6.1 Higgs boson production and decay

6.1.1 Higgs boson production modes

Several physics processes contribute to the Higgs boson production, which can be classified by the phenomenology and their cross section. The dominant production modes at the LHC are visualized in Figure 6.2 as Feynman diagrams and their cross section as a function of the Higgs boson mass M_H is shown in Figure 6.1 (a). The gluon fusion ggF production has the largest cross section of 48.5 pb at a center-of-mass energy $\sqrt{s}$ of 13 TeV and $M_H = 125.09$ GeV [27]. The Higgs boson is produced via a quark q loop, in which the heavy top t and bottom b quarks are the preferred ones, since the coupling strength is proportional to the mass of the fermions.

The second leading production mode is vector boson fusion VBF with a cross section of

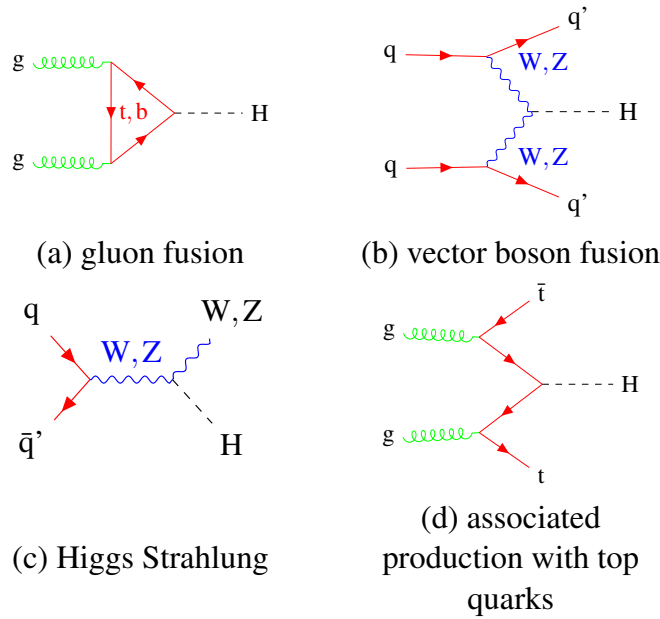


Figure 6.2: Dominant Higgs boson production processes at the LHC visualized by leading-order Feynman diagrams. The diagrams are ordered by decreasing cross sections: Gluon fusion ggF (a), vector boson fusion VBF (b), Higgs Strahlung VH (c) and associated production with top quarks $t\bar{t}H$ (d).

3.8 pb at $\sqrt{s} = 13$ TeV and $M_H = 125.09$ GeV [27]. Two quarks interact via the exchange of weak gauge bosons, which fuse into a Higgs boson. Therefore, two jets with a high invariant mass m_{jj} and a large difference in pseudorapidity $\Delta\eta_{jj}$ are present in the final state.

The Higgs Strahlung is the production of a Higgs boson in association with a weak gauge boson VH and has a cross section of 2.2 pb at $\sqrt{s} = 13$ TeV and $M_H = 125.09$ GeV [27]. A quark annihilates with an anti-quark into a weak gauge boson, which radiates a Higgs boson. The associated production of a Higgs boson with a top quark pair $t\bar{t}H$ with a cross section

of 0.5 pb at $\sqrt{s} = 13$ TeV and $M_H = 125.09$ GeV [27] is an order of magnitude smaller compared to the other production modes, due to the high invariant mass which is required to produce the heavy top quarks.

6.1.2 Higgs boson decay modes

The Higgs boson H decays into the massive weak bosons V and massive fermions f . The coupling strength to V is proportional to the squared mass of the bosons, while it is proportional to the fermion mass for f (see Section 2.3).

The dominant fermion decay modes are $H \rightarrow b\bar{b}$, $H \rightarrow \tau^+\tau^-$, $H \rightarrow c\bar{c}$ and $H \rightarrow \mu^+\mu^-$ (see Figure 6.3 (a)) and the possible bosonic decays are $H \rightarrow WW^*$, $H \rightarrow ZZ^*$, $H \rightarrow gg$, $H \rightarrow Z\gamma^*$ and $H \rightarrow \gamma\gamma$. For bosonic decay modes the Higgs boson decays directly into massive vector bosons, while the decay into massless particles is realized via a fermion or massive gauge boson loop (see Figure 6.3 (b-d)).

The branching ratio as a function of the Higgs boson mass M_H is shown in Figure 6.1 (b). The dominant decay modes for a Higgs boson with a mass of 125.09 GeV are $H \rightarrow b\bar{b}$ with a branching ratio of 57.7% and $H \rightarrow WW^*$ with a branching ratio of 21.5%. The main focus of this thesis is the $H \rightarrow \tau^+\tau^-$ decay mode with a branching ratio of 6.3%.

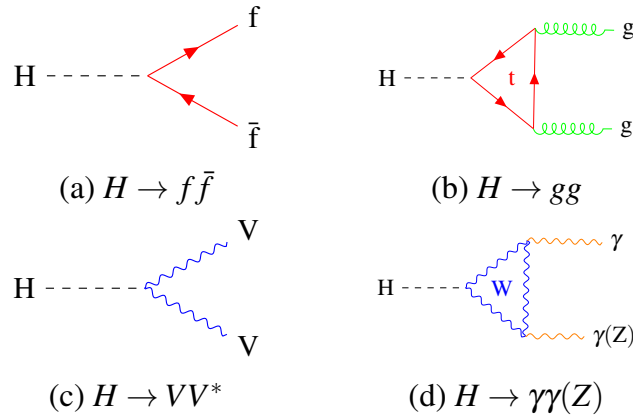


Figure 6.3: Dominant Higgs boson decay modes visualized by leading-order Feynman diagrams: $H \rightarrow f\bar{f}$ (a), $H \rightarrow gg$ (b), $H \rightarrow VV^*$ (c) and $H \rightarrow \gamma\gamma(Z)$ (d) via a W boson loop.

6.1.3 The $\tau\tau$ -decay mode

The τ -lepton has a relative short life-time ($2.9 \cdot 10^{-13}$ s [47]) and decays before it interacts with the detector. Therefore, only the decay products of the τ -leptons are measured in the ATLAS experiment. The decay of the τ -lepton is shown in Figure 6.4, while Figure 6.5 (a) summarizes the branching fractions of the τ -decay modes. The decay modes can be grouped into a leptonic decay mode with a light lepton flavor (e, μ) and two neutrinos ($\nu_\tau, \nu_{e,\mu}$) in

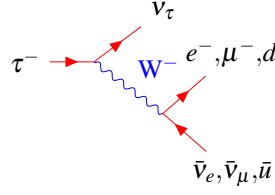
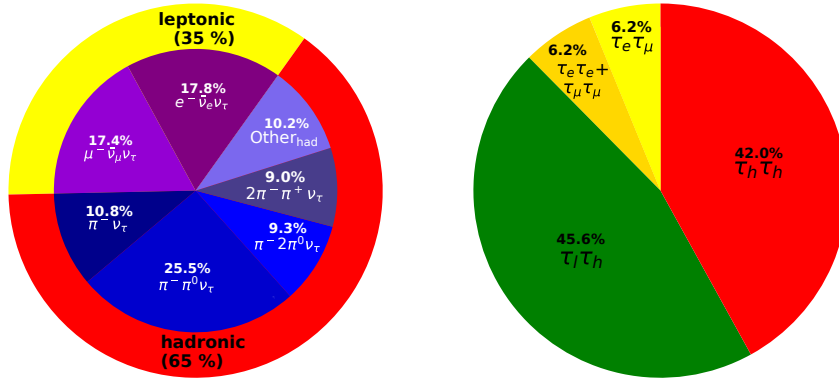


Figure 6.4: Feynman diagram of a τ^- -lepton decaying via a W^- boson into light leptons ($[e^-, \nu_e]$, $[\mu^-, \nu_\mu]$) or into quarks ($[d, \bar{u}]$), which hadronize to pions (π) or kaons (κ).



(a) Branching ratios of the τ -lepton (b) Branching ratios of the di- τ -system

Figure 6.5: Branching ratios of the decay modes of the τ -lepton (a) and the di- τ -system (b) [47].

the final state and into a hadronic decay mode with pions (π) or kaons (κ) and a τ -neutrino (ν_τ) in the final state. The hadronic decay mode has a branching fraction of 65% [47]. A di- τ -system has three different decay modes: $\tau_h \tau_h$ where both τ -leptons decay hadronically, $\tau_l \tau_l$ where both τ -leptons decay leptonically and $\tau_l \tau_h$ where one τ -lepton decays leptonically and one hadronically. The $\tau_l \tau_l$ decay mode can be additionally grouped into the same-flavor channel ($\tau_e \tau_e + \tau_\mu \tau_\mu$), in which both τ -leptons decay into the same light lepton, and the different-flavor channel ($\tau_e \tau_\mu$), in which the τ -leptons decay into light leptons with different flavors. The branching ratio of the di- τ -lepton decay modes are summarized in Figure 6.5 (b).

6.2 W/Z boson production in association with jets

Exemplary Feynman diagrams for the W/Z boson production with up to two associated jets are shown in Figure 6.6. The vector boson can either decay hadronically resulting in jets in the final state or leptonically into electrons, muons or τ -leptons.

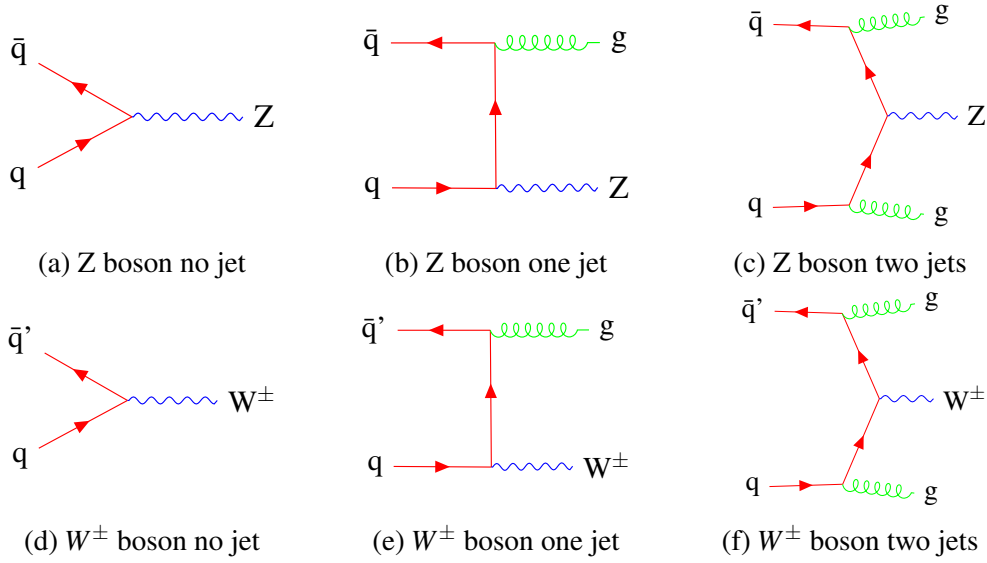


Figure 6.6: Vector boson production of a Z boson (top) and of a $W^\pm$ boson (bottom) in association with jets visualized by Feynman diagrams. From left to right the number of jets is increasing from zero jets to two jets.

6.3 Di-boson production

Di-boson events can be produced by WW -, ZZ - or WZ -production. The vector boson decays either leptonically or hadronically resulting in leptons and jets in the final state. Figure 6.7 shows exemplary Feynman diagrams for different di-boson production modes.

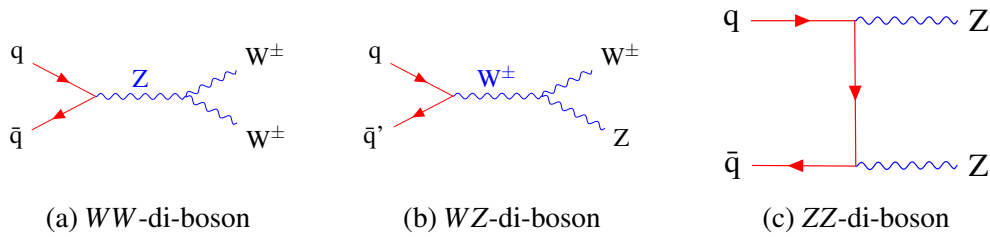


Figure 6.7: Vector boson pair production visualized by Feynman diagrams.

6.4 QCD multi-jet production

The cross sections for various QCD processes with outgoing quarks and gluons are large at the LHC. Color-charged particles hadronize and produce events with several jets in the final state. These jets can be misidentified as τ -leptons or lighter leptons, so that these events

6.5 Single top quark and top quark pair production

are an important background for analyses with a τ -lepton in the signal process. Figure 6.8 summarizes exemplary Feynman diagrams for QCD multi-jet production.

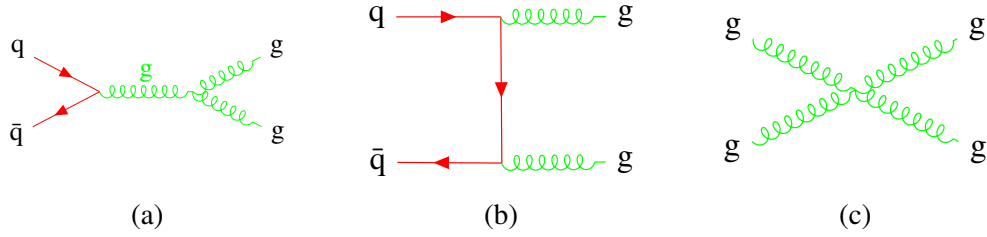


Figure 6.8: QCD multi-jet production in QCD processes with two jets visualized by Feynman diagrams.

6.5 Single top quark and top quark pair production

The single top quark production is shown in Figure 6.9 (a-c) and can be produced in the s-channel, t-channel, or in association with a W boson. The top quark preferably decays into a bottom quark and a W boson.

In the top quark pair production both top quarks decay into a bottom quark or a W boson. Exemplary Feynman diagrams are shown in Figure 6.9 (d-e).

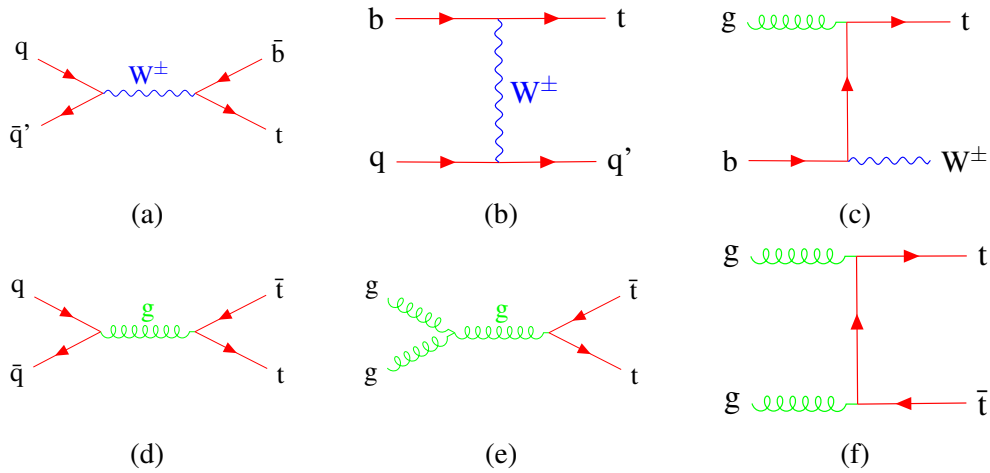


Figure 6.9: Single top quark production (top) and top quark pair production (bottom) visualized by Feynman diagrams.

7 EXPERIMENTAL METHODS

7.1 Simplified Template Cross Section

The measurement of Simplified Template Cross Sections (STXS) is a common strategy of the ATLAS and CMS experiments to study the Higgs boson properties and is described in detail in Reference [27]. The measurements of the signal strength (μ) and coupling strength (κ) make use of kinematics predicted by the SM and extrapolate from the measured phase space to the global phase space. Both increase the impact of theoretical uncertainties on the determination of cross sections and dependencies on the underlying physics model. Furthermore, the coupling strength and the signal strength are normalized to the SM prediction ($\kappa = g_{obs}/g_{SM}$, $\mu = \sigma_{obs}/\sigma_{SM}$).

The STXS technique allows the reduction of theoretical dependencies, which are directly

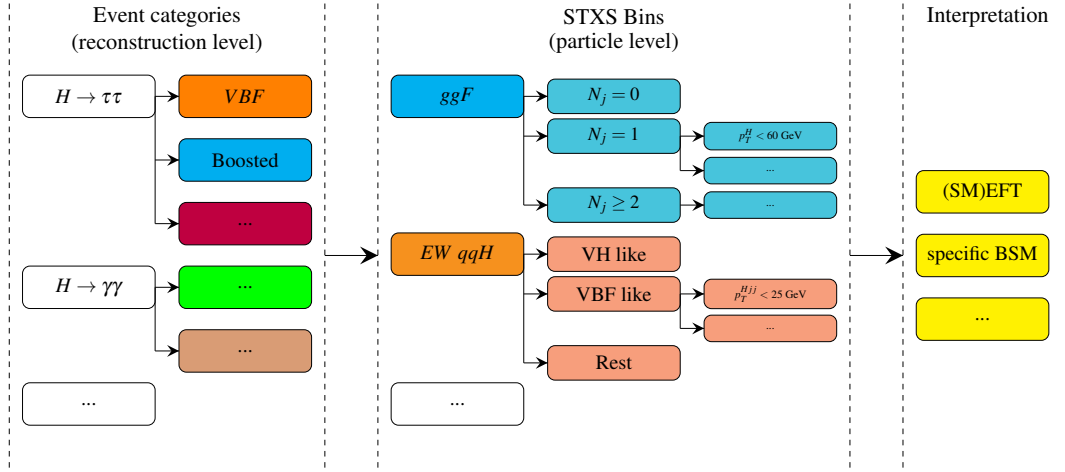


Figure 7.1: Schematic overview of the STXS framework, including the input from event selection (left), the measured cross-sections as the main result of the STXS (mid) and a list of possible interpretations based on the STXS results (right).

folded into the measurements. The parameters of interest are cross sections of exclusive regions of phase space for specific production modes defined on particle-level and will be referred to as STXS bins throughout the rest of the thesis. The definitions of these STXS bins are motivated by minimizing the dependence on theoretical uncertainties, maximizing the experimental sensitivity and isolating possible effects from BSM-physics.

The STXS bins are defined in different stages. Stage 0 is targeting inclusive production mode cross sections, while Stage 1 defines regions of phase space with a finer granularity to decrease the model dependence and increase the sensitivity to deviations from the SM predictions in specific phase space regions. In addition to Stage 1, a Stage 1.2 is defined, which is an optimized version of Stage 1 to further increase the granularity and reduce the impact of theoretical uncertainties. It was defined after the first round of publications of STXS

Stage 1 measurements with the partial Run 2 data set recorded from 2015 to 2016 and is used for the measurements with the full Run 2 data set recorded from 2015 to 2018.

A simplified schematic overview of the STXS definition is given in Figure 7.1. On the left-hand side the analysis event categories are shown, which are enriched with specific Higgs boson production processes and defined on reconstruction level. The event selection should be optimized to increase the sensitivity to STXS bins. The center of Figure 7.1 illustrates the part of the STXS method, which is determined by a global fit using the analysis categories and is the main result of the STXS measurement: Cross sections per production mode, split into exclusive kinematic STXS bins are defined on particle-level.

These STXS measurements are used as an input for subsequent interpretations, which are shown in Figure 7.1 on the right-hand side, such as the test of specific models for BSM physics and the estimation of Wilson coefficients in SMEFT (see Section 2.5), of signal strength modifiers or of coupling scale factors κ . The last two of these examples provide compatibility with earlier measurements. The experimental results should quote the full covariance matrix among the different STXS bins. In this way, the theoretical uncertainties of individual models can be included in the interpretation step to ensure a long-term usability of the measurements. The STXS measurement has some similarity to differential cross section measurements, but it combines the advantages of signal strength, fiducial and differential measurements, by allowing the combination of different analyses, increasing the granularity of the measurement and the reducing the dependency of the measurement on theoretical uncertainties.

7.1.1 Object definition

The global fit unfolds the measured number of events in the event selections on reconstruction level to the cross section in STXS bins defined on particle level. Thus, the final state objects on particle level need to be defined unambiguously for the measured bins and theoretical predictions from either analytic calculations or simulations. Within the STXS framework simple definitions of the final state objects are used. The Higgs boson is treated as a final state object to allow the combination of analyses exploiting different decay modes.

Higgs boson The STXS approach requires an on-shell Higgs boson. A requirement on the rapidity of Higgs boson y_H with $|y_H| < 2.5$ is applied, since current measurements are not sensitive beyond this rapidity range.

Leptons Electrons and muons originating from decays of vector bosons are required to be “dressed”, meaning that the momenta of final state radiated photons are added back to the momenta of them. τ -leptons are defined by the sum of their decay products. There are no restrictions on the transverse momentum or on the pseudorapidity of the leptons, such that for a leptonically decaying weak vector boson the full phase space is included.

Jets The STXS particle level jets are defined as anti- k_t jets with a jet radius of $R = 0.4$ and are built from all stable particles, including neutrinos, photons and leptons from hadronic decays or produced in the shower. Stable particles are defined by requiring a lifetime larger

than 10 ps, corresponding to the required value for a particle to pass GEANT 4 [109] in the experimental simulation. Decay products of Higgs bosons are not considered, due to its definition as a final state particle. Leptons and neutrinos originating from weak vector boson are removed and treated separately, while decay products from hadronically decaying weak vector bosons are included in the jet finding algorithm. The particle level jets are defined without a requirement on their pseudorapidity, but requiring a minimum transverse momentum p_T of 30 GeV, to reduce the extrapolation in the measurements.

7.1.2 STXS Stage 0

The Stage 0 definition of the STXS framework is targeting the inclusive production mode cross sections of the Higgs boson (H) and is summarized in Figure 7.2. The definitions of production modes differ slightly from the definitions used in previous measurements at the LHC. The VH production mode is explicitly defined as Higgs boson production in association with a leptonically decaying vector boson (V), while electroweak $EW\ qqH$ production comprises the VBF Higgs boson production and the $V(\rightarrow qq)H$ production with hadronically decaying V . Similarly, the $ggZ(\rightarrow qq)H$ production mode, in which the Z boson decays hadronically, is included in the ggF production mode. This definition takes into account that the VH and VBF processes and similarly the $ggZH$ and ggF production become ambiguous when including higher order QCD corrections if the vector boson V decays hadronically. For the $EW\ qqH$ and VH production modes the information of sub-production modes is stored to allow a comparison to the previous measurements. Therefore, the $EW\ qqH$ is split into $V(\rightarrow qq)H$ and VBF , while the VH production mode is split into quark and gluon initiated events and into production associated with a Z and W boson.

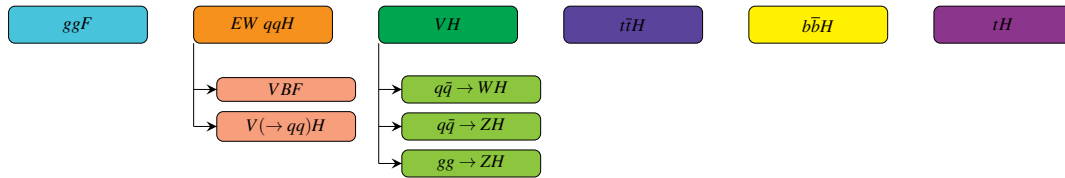


Figure 7.2: Schematic overview of the STXS bins defined at Stage 0.

7.1.3 STXS Stage 1

At Stage 1 the inclusive production mode STXS bins of the Stage 0 are split further, based on different kinematic properties of the event such as the jet multiplicity N_j or the transverse momentum of the Higgs boson p_T^H . This Section will give an overview of STXS bins at Stage 1.

$EW\ qqH$ STXS bins at Stage 1 The $EW\ qqH$ STXS bins at STXS Stage 1 are shown in Figure 7.3. The phase space is split into two different transverse momentum regions of the

leading jet at $p_T^{j_1} = 200$ GeV. The lower phase space is expected to be dominated by SM-like events, while the high STXS bin should be sensitive to potential BSM contributions. The lower phase space is split into a VBF-like, VH-like and rest phase space. The VBF like STXS bin requires $N_j \geq 2$, an invariant mass of the di-jet system of the two

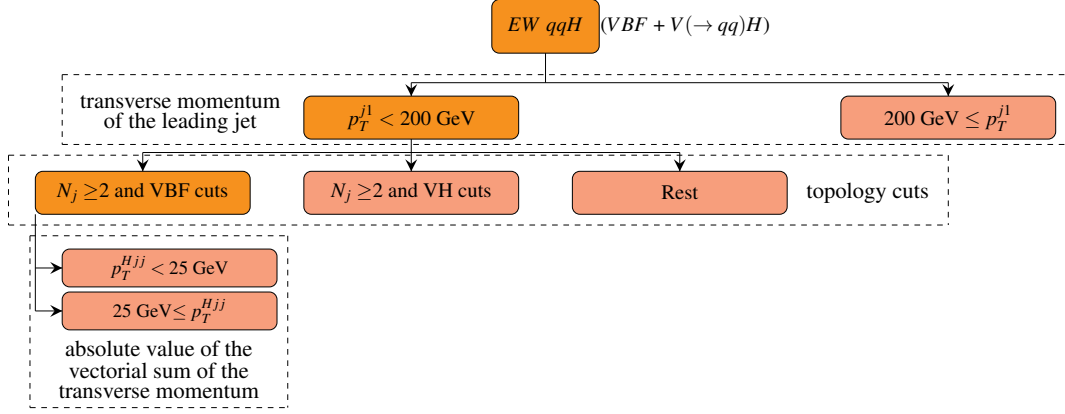


Figure 7.3: Stage 1 STXS bins for the $EW\ qqH$ production mode. The melon filled boxes refer to STXS bins, while the orange ones are only given to group the STXS bins.

jets with the largest transverse momentum of $m_{jj} \geq 400$ GeV and an absolute difference between the η -values of the jets of $\Delta\eta_{jj} = |\eta_{j_1} - \eta_{j_2}| \geq 2.8$, where j_2 corresponds to the sub-leading jet. It is defined as a compromise between the various VBF event categories used by different ATLAS and CMS analyses. This phase space is split into two STXS bins at $p_T^{Hjj} = 25$ GeV, where p_T^{Hjj} is the absolute value of the vectorial sum of the transverse momentum: $p_T^{Hjj} = |\vec{p}_T^H + \vec{p}_T^{j_1} + \vec{p}_T^{j_2}|$. This split mimics a separation into an exclusive 2-jet-like and an inclusive 3-jet-like phase space. Several analyses of ATLAS and CMS use similar kinematic variables to enrich the VBF production mode, so that this split reduces the extrapolation from the event categories to the measure phase space.

The VH like STXS bin is defined as a typical topology of associated vector boson production with a hadronically decaying vector boson by requiring $N_j \geq 2$ and $60 \leq m_{jj} < 120$ GeV.

The rest bin contains all events of the remaining phase space, including the $N_j = 0$ phase space and could be sensitive to certain BSM contributions, which do not follow the typical SM VBF signature of two forward jets.

ggF STXS bins at Stage 1 An overview of the current STXS Stage 1 definition for the ggF production mode is given in Figure 7.4. The phase space is split into $N_j = 0, 1$ and ≥ 2 . These splits are motivated by their use in the experimental analyses and by the reduction of the uncertainty for extrapolating such measurements to a more inclusive phase space.

Additionally, the $N_j \geq 2$ phase space is split into a VBF like and a non-VBF like phase space. The VBF-like phase space is defined in the same way as for the $EW\ qqH$ production mode at Stage 1, using the requirements $m_{jj} > 400$ GeV and $\Delta\eta_{jj} \geq 2.8$. The requirement of

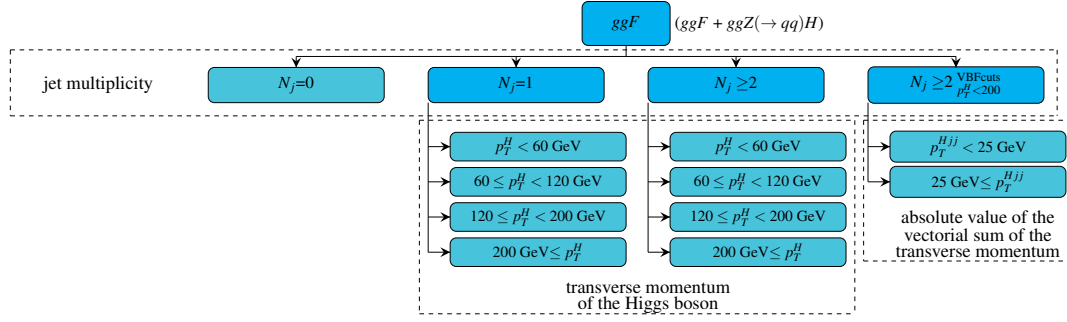


Figure 7.4: Stage 1 STXS bins for the ggF production mode. The turquoise filled boxes refer to STXS bins, while the blue ones are only given to group the STXS bins.

$p_T^H < 200$ GeV is applied on the VBF-like phase space to give priority to the STXS bin with $p_T^H \geq 200$ GeV and without a VBF-like topology. The VBF-like phase space is split into two STXS bins, where one is an exclusive 2-jet-like STXS bin ($p_T^{Hjj} < 25$ GeV) and the other an inclusive 3-jet-like bin ($p_T^{Hjj} \geq 25$ GeV). These two STXS bins are defined to measure the ggF contamination in VBF event categories. This contamination can introduce correlations between EW qqH and ggF STXS bins. By the separated definition of these STXS bins the correlation is absorbed into these STXS bins and reduced for the remaining ggF STXS bins. The $N_j = 1$ and ≥ 2 phase space regions are split according to p_T^H . The STXS bin with $0 \leq p_T^H < 60$ GeV has the highest expected number of events of all ggF STXS bins. The upper threshold is chosen to give a more even split of events between the $0 \leq p_T^H < 60$ GeV and $60 \leq p_T^H < 120$ GeV STXS bins, while it is high enough such that no resummation effects are expected. Furthermore, the threshold is sufficient to introduce a negligible bias from the p_T^H requirement. The STXS bin with $60 \leq p_T^H < 120$ GeV remains from the split into a low and a high region of p_T^H . Here, the lower threshold is high enough for the STXS bin to be treated as a hard Higgs boson production with an associated jet system in the theoretical description. The STXS bin with $120 \leq p_T^H < 200$ GeV avoids large extrapolations and it provides in the case of $N_j = 2$ a substantial part of the ggF production contribution in the event category designed for $V(\rightarrow qq)H$ production. The STXS bins beyond the top quark mass $m_t = 171$ GeV ($p_T^H \geq 200$ GeV) allow the top quark loop to be on-shell such that top quark mass effects become relevant. These high p_T^H STXS bins ensure the usability of the heavy-top expansion for the lower p_T^H STXS bins and offer the possibility to distinguish a point-like gluon fusion vertex induced by heavier BSM particles in the loop from the resolved top quark loop.

VH STXS bins at Stage 1 The VH STXS Stage 1 definition for events with a leptonically decaying vector boson is shown in Figure 7.5. The three inclusive production mode STXS bins of Stage 0 are split according to N_j and the transverse momentum of the vector boson p_T^V . Both $q\bar{q} \rightarrow VH$ STXS bins are split at $p_T^V = 150$ and 250 GeV. The resulting intermediate bin with $150 \leq p_T^V < 250$ GeV is additionally split into events with $N_j = 0$ and events with

$N_j \geq 1$. The values of p_T^V used in the split are oriented at quantities used in the $H \rightarrow b\bar{b}$ analysis, which is the most sensitive analyses for the VH STXS bins. The STXS bins with $p_T^V < 150$ GeV include events of the bosonic decay modes and events from $H \rightarrow b\bar{b}$ with $W \rightarrow l\nu$ or $Z \rightarrow ll + \nu\bar{\nu}$, which do not rely on the missing transverse energy triggers. The intermediate p_T^V STXS bins are dominated by events from $H \rightarrow b\bar{b}$ with $Z \rightarrow \nu\bar{\nu}$ due to the high threshold of the missing transverse energy trigger. The N_j split is introduced to reflect the different experimental sensitivities and to avoid the corresponding theory dependencies. The high p_T^V bins have the highest sensitivity to BSM effects. The $gg \rightarrow ZH$ production phase space is split in the same way except for the split into a high and intermediate p_T^V bin.

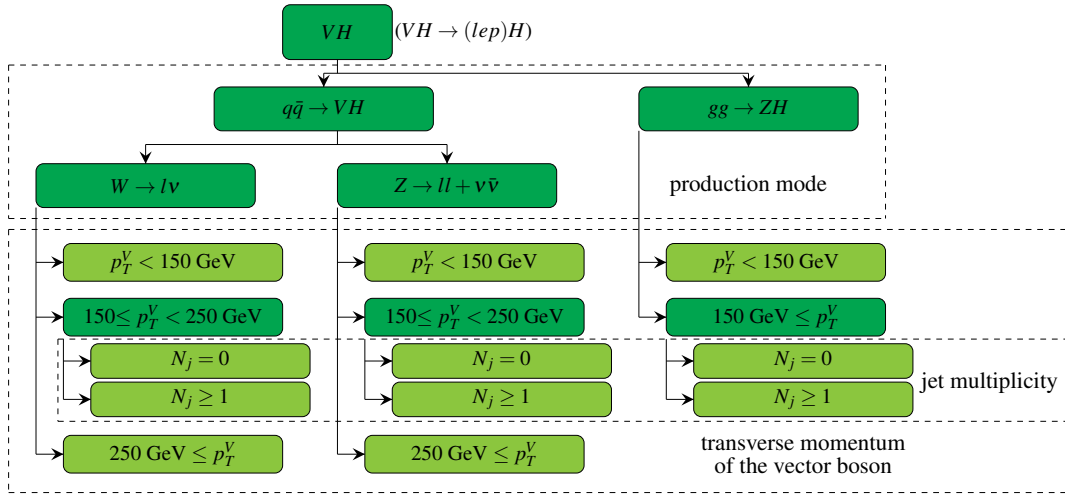


Figure 7.5: Stage 1 STXS bins for the VH production mode. The light green filled boxes refer to STXS bins, while the dark green ones are only given to group the STXS bins.

Further production modes at Stage 1 The STXS Stage 0 bins for the remaining production modes $t\bar{t}H$, $b\bar{b}H$ and tH are not further split at Stage 1.

7.1.4 STXS Stage 1.2

The STXS Stage 1 definition was optimized after the first round of publication using the data sets recorded in 2015 and 2016. The definition was modified to harmonize the STXS bins across the $EW\ qqH$ and ggF production modes and to match the event categories of the analyses better. The increased statistics of the data recorded from 2015 to 2018 allows to measure STXS bins with a finer granularity compared to the previous measurements, so that the several STXS bins at Stage 1 are further split into finer regions of phase space at Stage 1.2.

***EW qqH* STXS bins at Stage 1.2** An overview of the current STXS Stage 1.2 definition for the *EW qqH* production mode is given in Figure 7.6. The definitions of the STXS bins are changed to harmonize them with the kinematic variables used for the *ggF* production mode and to optimize the definition further to reduce extrapolation uncertainties in different regions of phase space.

Compared to Stage 1 the split in a low and high region of p_T^{j1} is replaced by a split in p_T^H , which is only applied to the VBF like phase space. The definition of the VBF like phase space is simplified compared to the previous definition by removing the $\Delta\eta_{jj}$ requirement and lowering the m_{jj} threshold to 350 GeV. In addition to the split at $p_T^{Hjj} = 25$ GeV a split at $m_{jj} = 700$ GeV is introduced to define high m_{jj} bins for analyses using the $H \rightarrow b\bar{b}$ decay, which have no sensitivity to the regions of phase space with lower m_{jj} . This results into eight different VBF like STXS bins instead of the two defined at Stage 1.

In addition to the VBF-like phase space, the rest bin of Stage 1 was redefined. The rest bin is split into N_j and m_{jj} bins for events with $N_j \geq 2$. This allows to account for the different experimental sensitivities in these regions and to reduce model dependencies and extrapolation uncertainties. The requirements on VH-like STXS bin is a simplified compared to Stage 1 by simply requiring with $60 \leq m_{jj} < 120$ GeV.

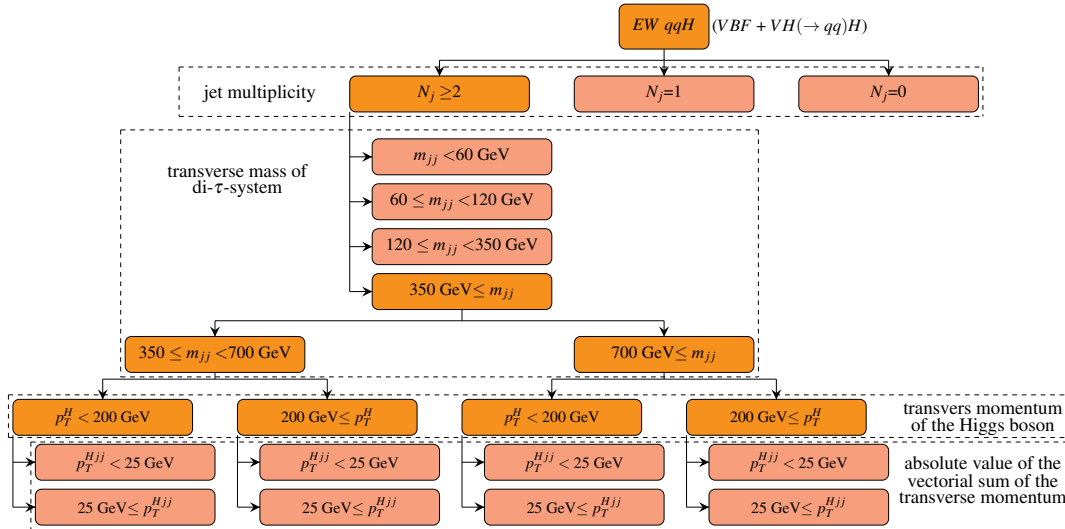


Figure 7.6: Stage 1.2 STXS bins for the *EW qqH* production mode. The melon filled boxes refer to STXS bins, while the orange ones are only given to group the STXS bins.

***ggF* STXS bins at Stage 1.2** An overview of the current STXS Stage 1.2 definition for the *ggF* production mode is given in Figure 7.7. The STXS bin definitions in the $N_j = 0$, high p_T^H and *ggF* VBF-like phase space regions are changed compared to Stage 1.

The $N_j = 0$ STXS bin is split into two p_T^H slices at 10 GeV. This threshold is defined to collect $N_j = 1$ events with a low transverse momentum of the jet p_T^j . In these events, the probability

to not reconstruct a jet is high compared to phase spaces with higher p_T^j . The split ensures that the lower p_T^H $N_j = 0$ STXS bin is rather pure, while the contamination of $N_j = 1$ events are absorbed in the higher p_T^H $N_j = 0$ STXS bin.

The VBF like STXS bin requirements are redefined for the STXS bins with $p_T^H < 200$ GeV to account for the new EW qqH STXS bins at Stage 1.2. The VBF like phase space is defined with $m_{jj} \geq 350$ GeV and is split into four STXS bins at $m_{jj} = 700$ GeV and $p_T^{Hjj} = 25$ GeV to mirror the EW qqH STXS bins.

In the phase space region with $p_T^H \geq 200$ GeV the jet multiplicity split is removed compared to Stage 1. The jet multiplicity is not well-defined in such a high boosted phase space region. Jets in this boosted topology will overlap in the detector and cannot be experimentally resolved from each other. Instead, a finer granularity in p_T^H is introduced by splitting at $p_T^H = 300, 450$ and 650 GeV.

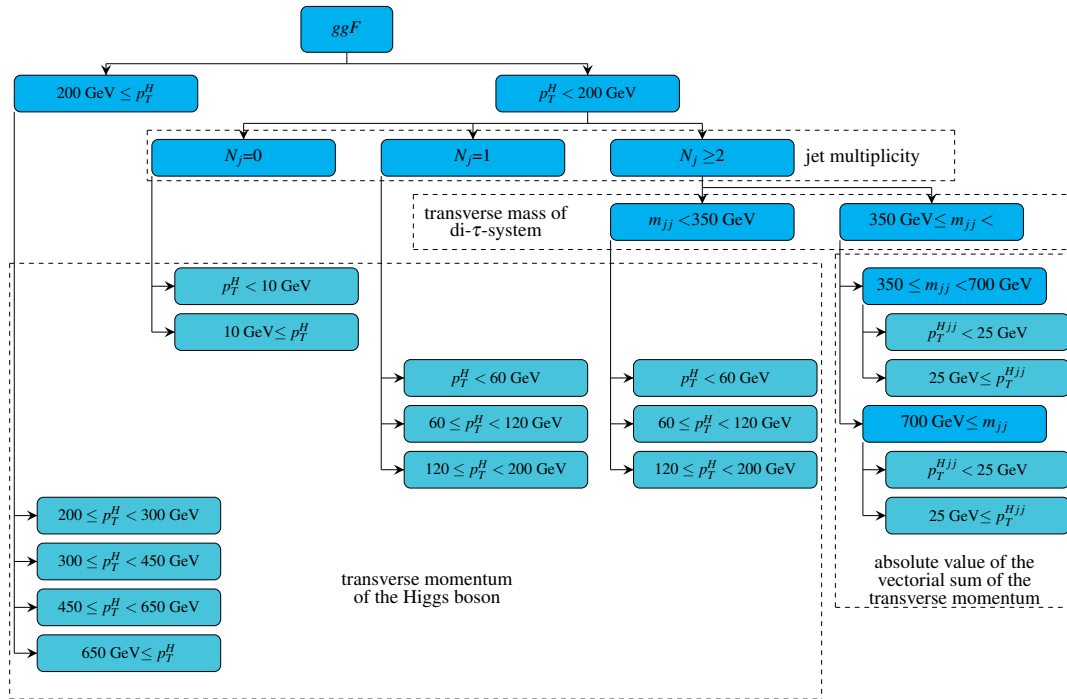


Figure 7.7: Stage 1.2 STXS bins for the ggF production mode. The turquoise filled boxes refer to STXS bins, while the blue ones are only given to group the STXS bins.

VH STXS bins at Stage 1.2 The VH STXS bins definition at Stage 1.2 is shown in Figure 7.8. In comparison to Stage 1 the granularity is increased to account for the higher sensitivity that is reached by the analyses with the full Run 2 data set. Furthermore, the phase space splitting is harmonized across the three different inclusive VH production modes.

Each production mode is split according to N_j and p_T^V . Compared to Stage 1 the granularity

in p_T^V was increased by adding splits at $p_T^V = 75$ and 400 GeV. The N_j split is applied to all p_T^V STXS bins and not only in the $150 \leq p_T^V < 250$ GeV phase space region as for Stage 1.

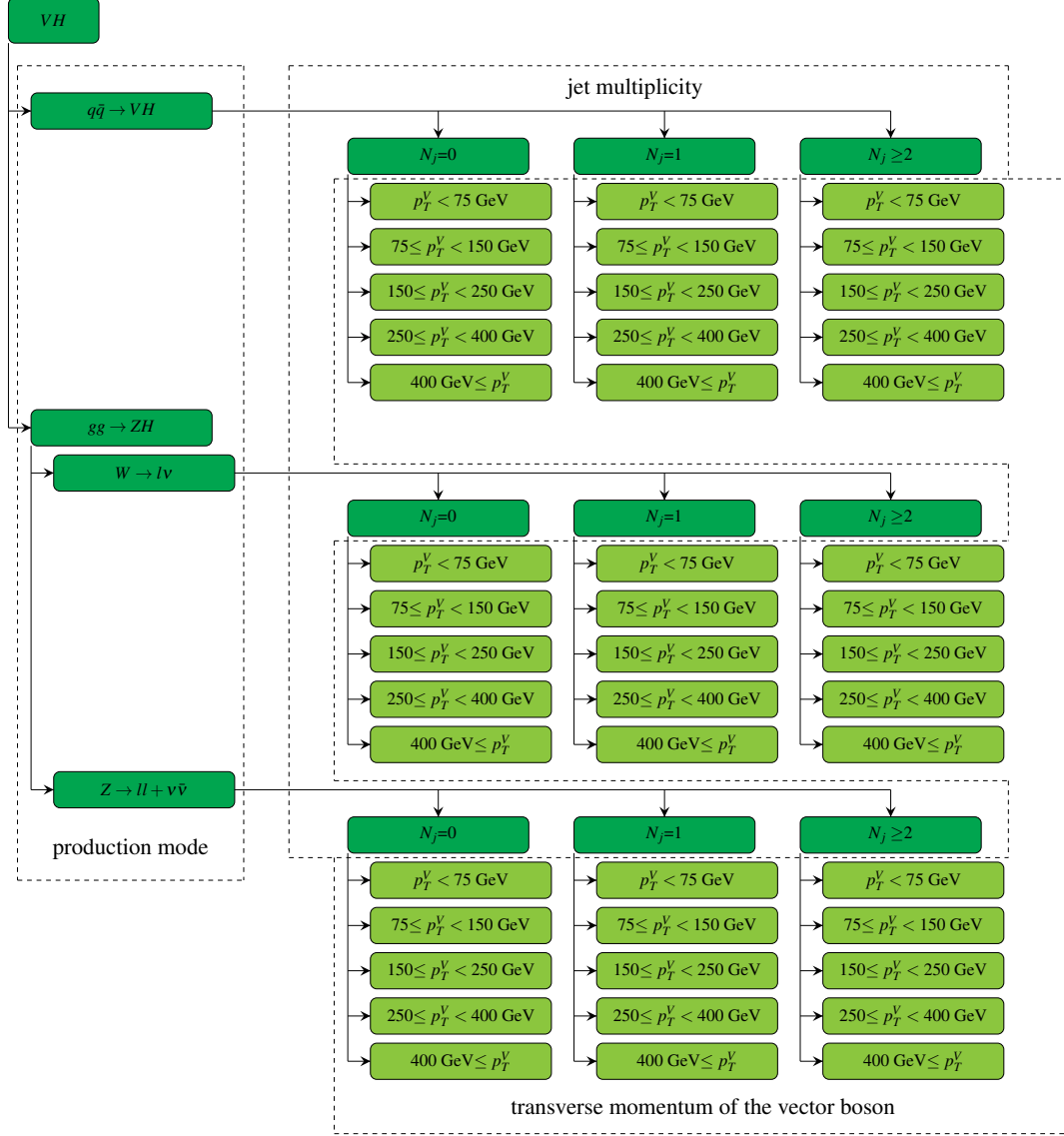


Figure 7.8: Stage 1.2 STXS bins for the VH production mode. The light green filled boxes refer to STXS bins, while the dark green ones are only given to group the STXS bins.

$t\bar{t}H$ STXS bins at Stage 1.2 In contrast to the STXS Stage 1, dedicated STXS bins for the $t\bar{t}H$ production are defined at Stage 1.2 (see Figure 7.9). The phase space is split into regions according to p_T^H . The splits are defined using the same values as for the ggF production mode

with $p_T^H = 60, 120, 200, 300$ and 450 GeV, except for the highest split at 650 GeV, which is removed compared to the ggF production mode to account for the lower experimental sensitivity to $t\bar{t}H$ production.

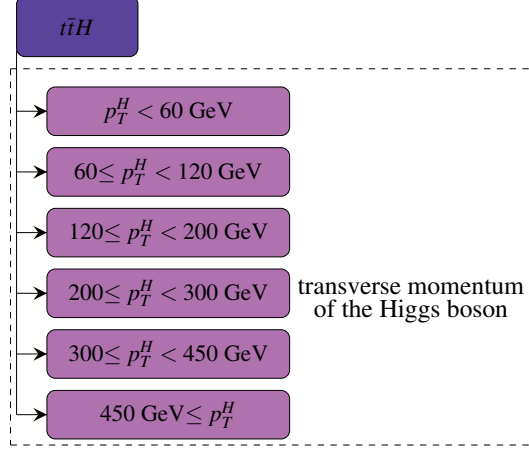


Figure 7.9: Stage 1.2 STXS bins for the $t\bar{t}H$ production mode. The light violet filled boxes refer to STXS bins, while the dark violet ones are only given to group the STXS bins.

Further production modes at Stage 1.2 The remaining production modes $b\bar{b}H$ and tH are not modified compared to Stage 1.

7.2 Likelihood function

For most of the analyses in thesis, maximum-likelihood fits are performed to data in order to extract the free model parameters. This section will give a brief overview of the utilized likelihood function. The section is written based on the descriptions in References [225] and [226].

The measurement of the signal strength can be used to introduce the likelihood function. In such an analysis the parameter of interest (PoI) is a scaling parameter r , which is free floating and acts on the expected signal s_i . The expected number of events n^{exp} in one bin i in a given event category can be expressed as the sum of s_i and background events b_i in this bin:

$$n_i^{\text{exp}} = r \cdot s_i(\vec{\theta}) + b_i(\vec{\theta}), \quad (7.1)$$

where r is the PoI acting on the signal sample and $\vec{\theta}$ a set of nuisance parameters on which s_i and b_i depend. The probability to observe n_i^{obs} events given the expected number of events n_i^{exp} can be expressed by the Poisson probability:

$$P_{\text{Pois}}(n_i^{\text{obs}} | r, \vec{\theta}) = \frac{(r \cdot s_i(\vec{\theta}) + b_i(\vec{\theta}))^{n_i^{\text{obs}}}}{n_i^{\text{obs}}!} \cdot e^{-(r \cdot s_i(\vec{\theta}) + b_i(\vec{\theta}))}. \quad (7.2)$$

If n_i^{obs} corresponds to a fixed number of observed data events this probability function is called likelihood function for the parameters r and $\vec{\theta}$: $\mathcal{L}(r, \vec{\theta} | n_i^{\text{obs}})$. If the data-set is grouped into smaller sets, for example by using more than one event category or by using the distribution of an observable in the fit model, the likelihood function is build for each selection and bin separately. The final likelihood function is then defined by the product of all individual likelihood functions:

$$\mathcal{L}(r, \vec{\theta} | n^{\text{obs}}) = \prod_i \frac{(r \cdot s_i(\vec{\theta}) + b_i(\vec{\theta}))^{n_i^{\text{obs}}}}{n_i^{\text{obs}}!} \cdot e^{-(r \cdot s_i(\vec{\theta}) + b_i(\vec{\theta}))}, \quad (7.3)$$

where i runs over all categories and bins, while n^{obs} is the total number of observed events. The best-fit value for the PoI is estimated by varying all free parameters of the likelihood and is defined as the one which maximizes the likelihood function.

This approach can be generalized for more complex measurements, in which cross sections of several independent signals are measured simultaneously. In this case the likelihood model has more than one PoI. The expected signal is split into different sub-samples for such a measurement, e.g. for the STXS measurement one sub-sample per STXS bin is defined (see Section 7.1) or for inclusive production mode cross section measurements one sample per production mode. One PoI per sub-sample is defined for these measurements:

$$\mathcal{L}(\vec{r}, \vec{\theta} | n^{\text{obs}}) = \prod_i \frac{(\sum_j (r_j \cdot s_{i,j}(\vec{\theta})) + b_i(\vec{\theta}))^{n_i^{\text{obs}}}}{n_i^{\text{obs}}!} \cdot e^{-(\sum_j (r_j \cdot s_{i,j}(\vec{\theta})) + b_i(\vec{\theta}))}, \quad (7.4)$$

where the j loops over all signal sub-samples.

The nuisance parameters NPs $\vec{\theta}$ are used to include systematic uncertainties into the likelihood function. The NPs can be varied during the maximization of the likelihood. However, most of the systematic uncertainties are estimated a priori in auxiliary measurements. To include this information into the likelihood model a constraint term in the form of a Gaussian probability density function is defined for each systematic uncertainty:

$$P_{\text{Gauss}}(\theta | \theta_0, \sigma_0) = \frac{1}{\sqrt{2\pi}\sigma_0} e^{-\frac{(\theta - \theta_0)^2}{2\sigma_0^2}}, \quad (7.5)$$

with the value for the nuisance parameter θ_0 and the variance σ_0 , both estimated in the auxiliary measurement. This term is maximized for $\theta = \theta_0$, while the term gets smaller for larger values of $(\theta - \theta_0)^2$. By adding this term to 7.4, the likelihood function gets punished for moving θ away from θ_0 . Therefore, the fit tends to keep the NPs as close as possible to auxiliary measurements and only changes them if the observed data shows a significant deviation, which can only be corrected by shifting θ with respect to θ_0 . The fit is able to change σ_0 to value σ if the included data allow so. A deviation between θ and θ_0 is called a pull. This simple implementation has two disadvantages. First the result for every NP has to be crosschecked with the value from the auxiliary measurement and second including NPs with different scales of input uncertainties can lead to an instability when maximizing the likelihood function. This can be avoided by a parametrization of the NPs. The impact of systematic uncertainties can be split into an overall normalization part η and into shape

part σ_i , where η affects only the total number of expected events, while σ_i affects the relative distribution of expected events in the different bins i . The nominal values $\eta^{\text{nom}} = 1$ and σ_i^{nom} as well as the variations $\eta^\pm$ and $\sigma_i^\pm$ for a NP are estimated in the auxiliary measurements and provided to the fit model. An interpolation and extrapolation scheme has to be used to obtain a continuous parametrization of η and σ_i based on a parameter α . An exponential extrapolation is used for the normalization, while a linear extrapolation is used for the shape uncertainty:

$$\eta(\alpha) = \begin{cases} (\eta^+)^{\alpha}, & \alpha \geq 0 \\ (\eta^-)^{-\alpha}, & \alpha < 0 \end{cases} \quad \text{and} \quad (7.6)$$

$$\sigma_i(\alpha) = \begin{cases} \sigma_i^{\text{nom}} + \alpha(\sigma_i^+ - \sigma_i^{\text{nom}}), & \alpha \geq 0 \\ \sigma_i^{\text{nom}} + \alpha(\sigma_i^{\text{nom}} - \sigma_i^-), & \alpha < 0 \end{cases}. \quad (7.7)$$

The extrapolations return the nominal values for $\alpha=0$ ($\eta(0) = 1 = \eta^{\text{nom}}$ and $\sigma_i(0) = \sigma_i^{\text{nom}}$) and for $\alpha^\pm = \pm 1$ the variations $\eta^\pm$ and $\sigma_i^\pm$. Using this parametrization scheme Equation 7.5 is simplified to:

$$P_{\text{Gauss}}(\alpha|0, 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\alpha^2}{2}}. \quad (7.8)$$

The NPs of all systematic uncertainties are depending on a individual parameter α_u , using a standard Gaussian distribution as the underlying probability density function. This solves the instability of the fit due to parameters of different scales and makes it easy to identify NPs, which get constraint or pulled in the fit.

This method is used for all systematic uncertainties except for incorporating systematic uncertainties from limited statistics in simulated events n_i^{bkg} per bin i . The sum of background events is given by the sum of weights w_j for all generated events in each bin: $n_i^{\text{bkg}} = \sum w_j$, where j runs over all generated events. The statistical uncertainty is given as: $\sigma_i = \sqrt{\sum w_j^2}$. The relative uncertainty of the background estimation is larger than the relative uncertainty on the unweighted events given by $1/\sqrt{N}$ (see [227]). The number of events that would be measured in an auxiliary measurement with the same relative uncertainty can be described by Poisson statistics using the number of unweighted event $n_i^{\text{bkg, unw}}$ defined as:

$$n_i^{\text{bkg, unw}} = \left(\frac{n_i^{\text{bkg}}}{\sigma_i} \right)^2, \quad (7.9)$$

with the uncertainty $\sqrt{n_i^{\text{bkg, unw}}}$ and the Poisson distribution with the mean $\gamma_i n_i^{\text{bkg, unw}}$ is given by:

$$P_{\text{Pois}}(n_i^{\text{bkg, unw}}|\gamma_i) = \frac{(\gamma_i n_i^{\text{bkg, unw}})^{n_i^{\text{bkg, unw}}}}{n_i^{\text{bkg, unw}}!} e^{-\gamma_i n_i^{\text{bkg, unw}}}. \quad (7.10)$$

γ_i is called the *gamma factor* and is the parameter, which has to be estimated to obtain the observed mean in the corresponding bin i . The number of unweighted events $n_i^{\text{bkg, unw}}$ can be non-integer, so that the Poisson distribution has been generalized to the gamma probability distribution P_Γ :

$$P_\Gamma(n_i^{\text{bkg, unw}}|\gamma_i) = \frac{(\gamma_i n_i^{\text{bkg, unw}})^{n_i^{\text{bkg, unw}}}}{\Gamma(n_i^{\text{bkg, unw}})} e^{-\gamma_i n_i^{\text{bkg, unw}}}, \quad (7.11)$$

with the gamma function $\Gamma(x)$.

The final likelihood function is then built by combining the likelihood function in Equation 7.3 with the additional terms:

$$\mathcal{L}(\vec{r}, \vec{\alpha}, \vec{\gamma}|n^{\text{obs}}) = \prod_i P_{\text{Pois}}(n_i^{\text{obs}}|\vec{r}, \vec{\theta}) \prod_u P_{\text{Gauss}}(\alpha_u|0, 1) \prod_j P_\Gamma(n_j^{\text{bkg, unw}}|\gamma_j). \quad (7.12)$$

7.2.1 Best-fit values and uncertainties

The best-fit values can be computed by maximizing the likelihood function with respect to $\vec{r}, \vec{\alpha}, \vec{\gamma}$. This maximization can be rather complex and computationally expensive. This can be improved by minimizing instead the double negative logarithm of the likelihood function $-2 \ln \mathcal{L}$. The products get replaced by sums, which are easier to calculate, while the same result will be achieved. The factor 2 is added, since in the asymptotic regime, in which $\mathcal{L}$ has approximately Gaussian shape, the values of $-2 \ln \mathcal{L}$ follow a χ_n^2 distribution, where the number of degrees of freedom n corresponds to the number of PoI r_i in the likelihood model. The uncertainty on the PoIs can then be extracted using confidence intervals for a confidence level (CL) $1-p$, where the p -value is the probability of obtaining a result that is as or less compatible with the null hypothesis as the actual observation. The confidence intervals correspond to regions with values of $-2 \ln \mathcal{L}$ below a threshold $F_{\chi_n^2}^{-1}(1-p)$. Here, $F_{\chi_n^2}^{-1}$ is the quintile function of the χ_n^2 distribution. The CL prescription described in Reference [228] is used for all measurements in this thesis.

7.2.2 Impact of individual uncertainties

To study the impact of the individual NPs on the central value and on the uncertainty of the PoI two different strategies can be performed. The first approach is called breakdown calculation and is used to estimate the combined impact of NPs associated to the same category of systematic uncertainties. The second approach is called NP ranking and used to order the single NPs by their impact on the central value of the PoI. Therefore, the impact of each NP on the PoI is evaluated and ordered by the size of the impact.

For the breakdown calculation the uncertainty associated to a specific sub-set of NPs is estimated by comparing the uncertainties on the PoIs estimated by two different fits. First the nominal fit with all NPs included is performed. In a second step all NPs corresponding to the studied sub-set are fixed to their post-fit values estimated in the first fit. With this setup a second fit is performed, in which effectively the sub-group of NPs is removed. The uncertainty

of these NPs is then estimated by taking the quadratic subtraction of the uncertainties on the PoI in these fits:

$$s_{\text{NPs}}^{\text{PoI}} = \sqrt{(s_{\text{nom}}^{\text{PoI}})^2 - (s_{\text{NPs rem}}^{\text{PoI}})^2}, \quad (7.13)$$

where $s_{\text{NPs}}^{\text{PoI}}$ is the uncertainty on the PoI associated to the sub-set of NPs, $s_{\text{nom}}^{\text{PoI}}$ the uncertainty on the PoI estimated in the nominal fit and $s_{\text{NPs rem}}^{\text{PoI}}$ the uncertainty on the PoI estimated in the fit with the sub-group of NPs removed.

In the NP ranking approach the impact of the $\alpha^\pm = \pm 1$ variation is estimated. Here, the up (+1) and down (-1) uncertainties are studied separately, so that in addition to the nominal fit two additional minimization per NP are necessary. The best-fit value of the PoI in the nominal fit is compared to the ones of the two alternative fits, in which the studied NP is set constant to $\alpha^\pm = \pm 1$. The impact is then calculated by taking the difference of the observed best-fit values for the PoI:

$$i_{\pm}^{\text{PoI}} = r_{\pm}^{\text{PoI}} - r_{\text{nom}}^{\text{PoI}}, \quad (7.14)$$

where $i_{\pm}^{\text{PoI}}$ is the impact of the up(down) variation of the NP on the PoI, $r_{\pm}^{\text{PoI}}$ the best-fit value of the PoI setting the NP to $\alpha \pm 1$ and $r_{\text{nom}}^{\text{PoI}}$ the best-fit value of the PoI for the nominal fit. This approach evaluates by how much the PoI best-fit value is changed if the NP value α is shifted to ± 1 . The NP introducing the largest shift, has the largest impact on the PoI and is therefore the most important NP in the measurement. The evaluation can be done with and without including the constraint on the NP, so that the impact of the NP before (pre-fit) and after (post-fit) the fit minimization can be estimated. Therefore, the method can provide an estimate of the importance of the single NPs on the PoI, but also yields a comparison of the prior value of the NP and its modification in the likelihood optimization.

8 τ -LEPTON TRIGGER MEASUREMENT

Analyses exploiting the fully hadronic decay channel, as described in Chapters 9 and 10, rely on triggering on hadronically decaying τ -leptons, called τ -trigger in the following. For every trigger it is necessary to correct for efficiency differences between simulation and data by applying scale factors (SFs) to the simulated events (see Section 8.3.1). The SFs for the τ -trigger are estimated using two tag-and-probe analyses, where one is based on $Z \rightarrow \tau_\mu \tau_h$ and one on $t\bar{t} \rightarrow \tau_\mu \tau_h$ events. In both analyses the trigger efficiencies in data ϵ_{data} and simulation ϵ_{MC} are estimated and used to calculate the final SFs (see Equation 8.1 and 8.2). The analyses use orthogonal event selections to calculate the SF. Finally, they are combined to estimate the final recommendations. This section discusses the $t\bar{t} \rightarrow \tau_\mu \tau_h$ analysis and summarizes the combination with the $Z \rightarrow \tau_\mu \tau_h$ analysis for the full Run 2 data set. The $Z \rightarrow \tau_\mu \tau_h$ analysis has a dominant contribution to the final SF for events with lower values of $p_T^{\tau_h}$, while for larger values of $p_T^{\tau_h}$ the $t\bar{t} \rightarrow \tau_\mu \tau_h$ analysis gets more important. The analyses are based on previous measurements described in Reference [229], which use the data set recorded in 2015 and 2016.

8.1 Signal and background processes

The signal process of the analysis are $t\bar{t} \rightarrow \tau_\mu \tau_h$ events, where the τ_μ is used as a tag and the τ_h as the probe of the tag-and-probe analysis. The main background contribution are top quark production events where a jet is misidentified as a τ_h -lepton. QCD multi-jet events have a sizable contribution in the low $p_T^{\tau_h}$ phase space of the analysis event selection. Here, one jet is misidentified as τ_h -lepton and one as a τ_μ -lepton. The $W \rightarrow l\nu$ and $Z \rightarrow ll$ processes can mimic the signal signature if additional objects from pile-up, underlying events or additionally radiated jets are misidentified. In addition, the $Z \rightarrow ll$ process can mimic the signal signature if one of the leptons is misidentified as τ_h -lepton. Contributions from Higgs boson production modes and other processes are neglected.

8.2 Event selection

The event selection of the analysis has two different stages. First a trigger selection is applied, followed by an event selection based on object criteria and event topology. The muon is used as the tag of the analysis and therefore two single-muon triggers are used to select the events. The two triggers are the HLT_mu26_ivarmedium and the HLT_mu50. An offline selection requirement with $p_T^\mu > 28$ GeV is applied. This requirement is slightly above the lower trigger threshold to avoid the use of trigger efficiency correction factors, which are biased by migration effects at the trigger threshold.

The event selection of the analysis consists of three different topologies: One signal region (SR), which is defined to enhance the signal contribution and to extract the final SFs and two control regions (CRs), which are used to estimate the contribution of specific background events in a data-driven way. Table 8.1 lists the event selection for the analysis. The selection can be grouped into three types. Two sets of requirements are based on the object properties of the leading muon and leading τ_h -lepton and one set on the event topology. Here, the leading

object is defined as the one with the highest p_T of all objects of the same type in the event. Further on, the leading muon is required to be the muon that fired the single-muon triggers. The leading muon has to fulfill the medium identification and the FCTightTrackOnly_FixedRad isolation criteria. Finally, an upper threshold on the absolute pseudorapidity of the muon with $|\eta| < 2.5$ is set.

The analysis is performed with τ_h -leptons that pass different identification criteria. For each working point individual efficiencies are extracted and SFs are calculated. This is necessary to provide the correct SFs for each possible τ_h -lepton analysis. Furthermore, the τ_h -lepton has to fulfill $p_T^\mu > 25$ GeV and $|\eta| < 1.37$ or $1.52 < |\eta| < 2.47$ requirements. Only events in which the τ_h -lepton candidate has one or three charged tracks ($N_{\text{tracks}} = 1$ or 3) are selected. This requirement is applied to significantly reduce the background contribution of events in which a jet is misidentified as a τ_h -lepton.

To suppress background contributions and to ensure the orthogonality to the $Z \rightarrow \tau_\mu \tau_h$ analysis, several requirements on the event topology are applied. The event must contain at least one b-jet ($N_{\text{b-jet}} > 0$), while for the $Z \rightarrow \tau_\mu \tau_h$ analysis a b-jet veto is applied. Exactly one muon ($N_\mu = 1$) and at least one additional jet ($N_{\text{jet}} > 1$) are required. To especially suppress background with a misidentified τ_h -lepton, the electric charge of the muon and τ_h -lepton are required to be opposite.

Based on this SR two CRs are defined by inverting one requirement of the SR. The first CR is called same-sign CR and is defined by requiring the electric charge of the muon and τ_h -lepton to have the same sign. The second CR is called multi-jet CR and is defined by inverting the isolation criteria of the muon.

	Muons	τ_h -lepton	Event topology
Signal region (SR)	Medium ID	Varying ID	$N_{\text{b-jet}} > 0$
	Trigger matched	$ \eta < 2.47$	$N_{jet} > 1$
	$ \eta < 2.5$	$ \eta \notin [1.37, 1.52]$	$N_\mu = 1$
	FCTightTrackOnly_FixedRad	$p_{\text{T}} > 25 \text{ GeV}$	Opposite electric charge of muon and τ_h
	$p_{\text{T}} > 28 \text{ GeV}$	$N_{\text{tracks}} = 1 \text{ or } 3$	
Same-Sign CR (relative to SR)			Same electric charge of muon and τ_h
Multi-jet CR (relative to SR)	Invert isolation		

Table 8.1: Event selection of the $t\bar{t} \rightarrow \tau_\mu \tau_h$ tag-and-probe analysis.

8.3 Analysis strategy

In contrast to the other analyses described in this thesis, this analysis is not based on a profiled likelihood fit. Instead the SFs are estimated using ratios of efficiencies calculated in data and simulated events. The following sections describe the SF calculation and the background

estimation.

8.3.1 Scale factor calculation

Two efficiencies are calculated to estimate the SF, which is used to correct the τ -trigger efficiency in simulations for analyses based on the τ -trigger. These are the τ -trigger efficiency for Monte-Carlo simulations ϵ_{MC} and the τ -trigger efficiency for data events ϵ_{data} . To calculate the efficiency for Monte-Carlo simulations only simulated $t\bar{t} \rightarrow \tau_\mu \tau_h$ signal events are used. For the data efficiency the $t\bar{t} \rightarrow \tau_\mu \tau_h$ contribution is estimated by subtracting contributions of other processes, which are predicted with Monte-Carlo simulations, from the recorded data. Both efficiencies are defined as the fraction of the τ_h -leptons that pass the τ -trigger requirement. Therefore, the collected data in the $t\bar{t} \rightarrow \tau_\mu \tau_h$ signal region before $N_{\text{data}}^{\text{SR,tot}}$ and after applying the τ -trigger requirement $N_{\text{data}}^{\text{SR,pass}}$ and the predicted simulated signal events before $N_{\text{sig}}^{\text{SR,tot}}$ and after $N_{\text{sig}}^{\text{SR,pass}}$ applying the trigger requirement are used:

$$\epsilon_{\text{data}} = \frac{N_{\text{data}}^{\text{SR,pass}} - N_{\text{bkg}}^{\text{SR,pass}}}{N_{\text{data}}^{\text{SR,tot}} - N_{\text{bkg}}^{\text{SR,tot}}}, \quad \epsilon_{\text{MC}} = \frac{N_{\text{sig}}^{\text{SR,pass}}}{N_{\text{sig}}^{\text{SR,tot}}}, \quad (8.1)$$

where $N_{\text{bkg}}^{\text{SR,tot}}$ and $N_{\text{bkg}}^{\text{SR,pass}}$ are selected background events in the signal region before and after applying the τ -trigger requirement. The final SF is then calculated by using:

$$SF = \frac{\epsilon_{\text{data}}}{\epsilon_{\text{MC}}}. \quad (8.2)$$

The SFs are calculated individually for a set of τ -trigger configurations. All configurations use a medium (medium1) online identification working point for the τ -lepton candidates and they can be grouped into three types: tracktwo, tracktwoEF and tracktwoMVA. Tracktwo indicates that τ -candidates with two tracks pass the high-level-trigger (HLT) as well. The tracktwoEF uses a tighter track requirement compared to the tracktwo trigger, to recover the efficiency for events, which include τ -lepton candidates with three tracks and have a large number of pile-up events. The contributions of such events is increased for the 2017 and 2018 data sets compared to the previous data sets. The tracktwoMVA additionally uses a boosted decision tree (BDT) using the TMVA implementation to reconstruct the energy scale of the τ -candidates (see Section 5.5). For all configuration a set of different thresholds on the online transverse momentum of the τ -lepton candidate p_T^τ at the HLT are used, e.g. tau25 indicates a requirement of $p_T^\tau > 25$ GeV. For a sub-set of trigger configurations an additional requirement on p_T^τ at the Level-1 (L1) trigger is used, e.g. L1TAU12 indicates a $p_T^\tau > 12$ GeV requirement at L1.

- tracktwo
 - HLT tau25 medium1 tracktwo
 - HLT tau35 medium1 tracktwo
 - HLT tau50L1TAU12 medium1 tracktwo

- HLT tau80L1TAU60 medium1 tracktwo
- HLT tau125 medium1 tracktwo
- HLT tau160 medium1 tracktwo
- HLT tau60 medium1 tracktwo
- HLT tau160L1TAU100 medium1 tracktwo
- tracktwoEF
 - HLT tau25 medium1 tracktwoEF (EF)
- tracktwoEF OR tracktwoMVA (EFORMVA)
 - HLT tau25 medium1 tracktwoEF OR mediumRNN tracktwoMVA
 - HLT tau35 medium1 tracktwoEF OR mediumRNN tracktwoMVA
 - HLT tau35L1TAU12IM medium1 tracktwoEF OR mediumRNN tracktwoMVA
 - HLT tau60 medium1 tracktwoEF OR mediumRNN tracktwoMVA
 - HLT tau80L1TAU60 medium1 tracktwoEF OR mediumRNN tracktwoMVA
 - HLT tau160L1TAU100 medium1 tracktwoEF OR mediumRNN tracktwoMVA

For each τ -trigger a separate set of SFs is estimated for different τ -lepton candidates with different offline requirements. All SFs are binned into the transverse momentum of the τ_h -lepton and are calculated separately for:

- τ_h -lepton identification criteria
 - BDT: loose, medium, tight
 - RNN: loose, medium, tight
- Number of charged tracks of the τ_h -lepton
 - 1-prong
 - 3-prong

The trigger menu has changed during the different data taking periods. Therefore, not all triggers are available for every data set. Table 8.2 summarizes the triggers applied in each data set. During the first technical stop (TS1) of the LHC in 2018 a new MVA based trigger was added to the trigger menu. Therefore, the 2018 data set is split into a pre-TS1 with $\mathcal{L} = 22 \text{ fb}^{-1}$ and post-TS1 with $\mathcal{L} = 38 \text{ fb}^{-1}$ data set. Studies show that no SFs for any trigger show a significant difference between the individual data taking periods, such that the SFs are calculated on a combined data set for all available periods. Therefore, this analysis provides SFs for the data sets recorded in 2017 and 2018, which are not covered by the previous measurement [229], but also recalculates the SFs for the data sets from 2015 and 2016 covered by the previous measurement. The new calculated SFs replace the old recommendation and are used for the 2015 and 2016 data set in all analyses, which are based on the full Run 2 data set.

Trigger	2015	2016	2017	2018	
				pre-TS1	post-TS1
tracktwo tau60			✓	✓	✓
tracktwo tau160L1TAU60			✓	✓	✓
tracktwo other	✓	✓	✓	✓	✓
EF				✓	✓
EFORMVA					✓

Table 8.2: Available τ_h -trigger for each data set. A check-mark illustrates the availability of the trigger, while an empty cell is shown if the trigger is not applied. The 2018 data set is split into two periods. The pre-TS1 data set was recorded before the first technical stop 2018 (TS1), while post-TS1 includes the data set after the stop.

8.3.2 Background estimation

To estimate the τ -trigger efficiency in data, the background contribution in the signal region has to be estimated and subtracted. The main background of the analysis are top quark production events with a jet misidentified as a τ_h -lepton. A sizable contribution of QCD multi-jet events appear in the low $p_T^{\tau_h}$ phase space of the analysis. This part of the phase space is dominated in the combination by the $Z \rightarrow \tau_\mu \tau_h$ analysis. A small fraction of the background are events in which an electron or muon is misidentified as a τ_h -lepton. The full background $N_{\text{bkg}}^{\text{full}}$ of the analysis can be described by:

$$N_{\text{bkg}}^{\text{full}} = N_{\text{bkg}}^{\text{QCD}} + N_{\text{bkg}}^{\text{others}, j \rightarrow \tau} + N_{\text{bkg}}^{\text{others}, l \rightarrow \tau}, \quad (8.3)$$

here the $N_{\text{bkg}}^{\text{QCD}}$ term includes all QCD multi-jet background events, $N_{\text{bkg}}^{\text{others}, j \rightarrow \tau}$ background events from other processes in which a jet is misidentified as a τ_h -lepton in particular from top quark processes and $N_{\text{bkg}}^{\text{others}, l \rightarrow \tau}$ background events from other processes with an electron or muon misidentified as a τ_h -lepton.

The QCD multi-jet background is estimated in a data-driven way, while the remaining background contributions are estimated via Monte-Carlo simulations. Figure 8.1 shows a schematic overview of the QCD multi-jet estimation. The same-sign CR is used to estimate the QCD multi-jet contribution and a transfer $R_{\text{OS/SS}}$ is used to transfer the contribution from the CR to the SR:

$$N_{\text{bkg}}^{\text{QCD}} = \left(N_{\text{data}}^{\text{SS,SR}} - N_{\text{others}}^{\text{SS,SR}} \right) \cdot R_{\text{OS/SS}}, \quad (8.4)$$

with the collected data in the same-sign CR $N_{\text{data}}^{\text{SS,SR}}$ and the background in the same-sign CR of background events not assigned to QCD multi-jet events $N_{\text{others}}^{\text{SS,SR}}$. The transfer factor is calculated in QCD multi-jet enriched CRs:

$$R_{\text{OS/SS}} = \frac{N_{\text{data}}^{\text{OS,CR}} - N_{\text{others}}^{\text{OS,CR}}}{N_{\text{data}}^{\text{SS,CR}} - N_{\text{others}}^{\text{SS,CR}}}. \quad (8.5)$$

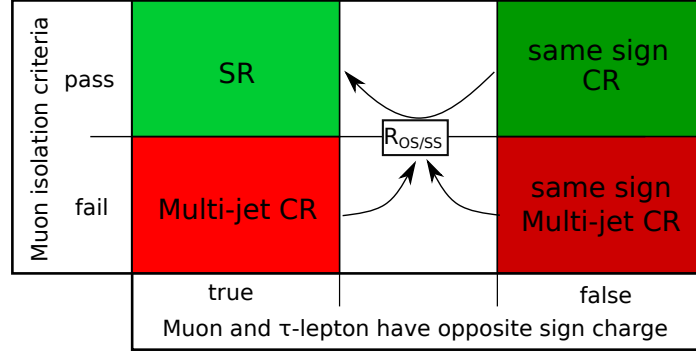


Figure 8.1: Schematic overview of the QCD multi-jet background estimation. Green boxes correspond to regions in which the leading muon pass the isolation criteria, while red boxes are assigned to QCD multi-jet control regions in which the muon fails the criteria. The $R_{OS/SS}$ factor transfer the background events in the same sign control region (green, right) to the signal region (green, left) and is estimated in the two QCD multi-jet regions (red).

A dependency of the $R_{OS/SS}$ on the transverse momentum of the τ_h -lepton is observed and considered. Figure 8.2 shows the $R_{OS/SS}$ values for the tau25 EFORMVA trigger with the 2018 data set for all BDT τ_h -ID working points and for $N_{tracks} = 1$ and 3, while Figure 8.3 shows the values for the same configurations but for the RNN τ_h -ID method. The $R_{OS/SS}$ values are similar for the different working points and for the BDT and RNN methods. For $N_{tracks} = 1$ a decreasing trend for higher values of $p_T^{\tau_h}$ is visible, while the opposite trend is shown for $N_{tracks} = 3$. For the RNN working points, especially Tight, this trend is covered by statistical fluctuations. The average value for $R_{OS/SS}$ across the $p_T^{\tau_h}$ bins is similar between $N_{tracks} = 1$ and $N_{tracks} = 3$. The values tend to be smaller after applying the τ_h -trigger compared to the values before applying the trigger requirement. The shown uncertainties correspond to the statistical uncertainties on the simulated events in the used QCD multi-jet CRs and they are propagated as systematic uncertainties to the final SFs.

A systematic uncertainty is defined using the same-sign CR after subtracting the background contributions as a function of $p_T^{\tau_h}$. The uncertainty is based on a variation of the ratio of data events in the opposite-sign and same-sign QCD multi-jet CRs. Figure 8.4 shows the shape variations on $R_{OS/SS}$ BDT for the same configuration as shown in Figure 8.2. The systematic uncertainties tend to be larger before applying τ_h -trigger. For $N_{tracks} = 1$ events the variations flip from positive to negative variations for larger values of $p_T^{\tau_h}$, while for $N_{tracks} = 3$ the opposite trend is visible. Overall the relative variations are up to 50%, except for one bin with $N_{tracks} = 3$ and τ_h -ID=Tight, where the variation is higher due to statistical fluctuations. These one-sided uncertainties are mirrored around zero to finally get up and down variations for $R_{OS/SS}$.

Table 8.3 shows the data, signal and background yields for all working points of the two τ_h identification methods (τ_h -ID) and for $N_{tracks} = 1, 3$ in the SR. In general the observed data yields are smaller for regions with $N_{tracks} = 3$ than with $N_{tracks} = 1$ and decrease for tighter working points of τ_h -ID. More data events are observed in regions with the BDT τ_h -ID method

τ_{h} -ID	N_{tracks}	Data	$N_{\text{bkg}}^{\text{QCD}}$	$N_{\text{bkg}}^{\text{others}, j \rightarrow \tau}$	$N_{\text{bkg}}^{\text{others}, l \rightarrow \tau}$	N_{signal}	
BDT	Loose	1	88933	8991 $\pm$ 201	34426 $\pm$ 178	3792 $\pm$ 38	39158 $\pm$ 93
	Loose	3	27795	3678 $\pm$ 120	8863 $\pm$ 90	1026 $\pm$ 12	13433 $\pm$ 55
	Medium	1	64252	5243 $\pm$ 156	19843 $\pm$ 138	2643 $\pm$ 36	34392 $\pm$ 87
	Medium	3	16940	1720 $\pm$ 82	3604 $\pm$ 57	520 $\pm$ 9	10724 $\pm$ 49
	Tight	1	43118	2998 $\pm$ 122	9694 $\pm$ 93	1597 $\pm$ 21	27608 $\pm$ 79
	Tight	3	10707	689 $\pm$ 50	1486 $\pm$ 38	281 $\pm$ 6	8218 $\pm$ 43
RNN	Loose	1	62466	4295 $\pm$ 154	17642 $\pm$ 130	2223 $\pm$ 26	38529 $\pm$ 92
	Loose	3	21356	1898 $\pm$ 93	5414 $\pm$ 71	634 $\pm$ 9	13088 $\pm$ 53
	Medium	1	48169	2864 $\pm$ 125	10294 $\pm$ 97	1626 $\pm$ 22	33764 $\pm$ 86
	Medium	3	13711	998 $\pm$ 69	2019 $\pm$ 44	317 $\pm$ 7	10495 $\pm$ 48
	Tight	1	34675	1621 $\pm$ 112	5044 $\pm$ 65	1074 $\pm$ 15	26551 $\pm$ 76
	Tight	3	9025	391 $\pm$ 38	747 $\pm$ 23	144 $\pm$ 5	7948 $\pm$ 41

Table 8.3: Event yields in the signal region of the $t\bar{t} \rightarrow \tau_\mu \tau_h$ tag-and-probe analysis for all identification criteria (τ_h -ID) of the BDT and RNN methods and for τ_h -leptons with one charged track and three charged tracks.

as for the corresponding working point of the RNN method. The signal contribution increases for regions with tighter working points from Loose BDT (RNN) with $<48\%$ ($<61\%$) to Tight BDT (RNN) with $>64\%$ ($>76\%$). Therefore, the expected data yields for the RNN method are smaller but the background rejection is higher.

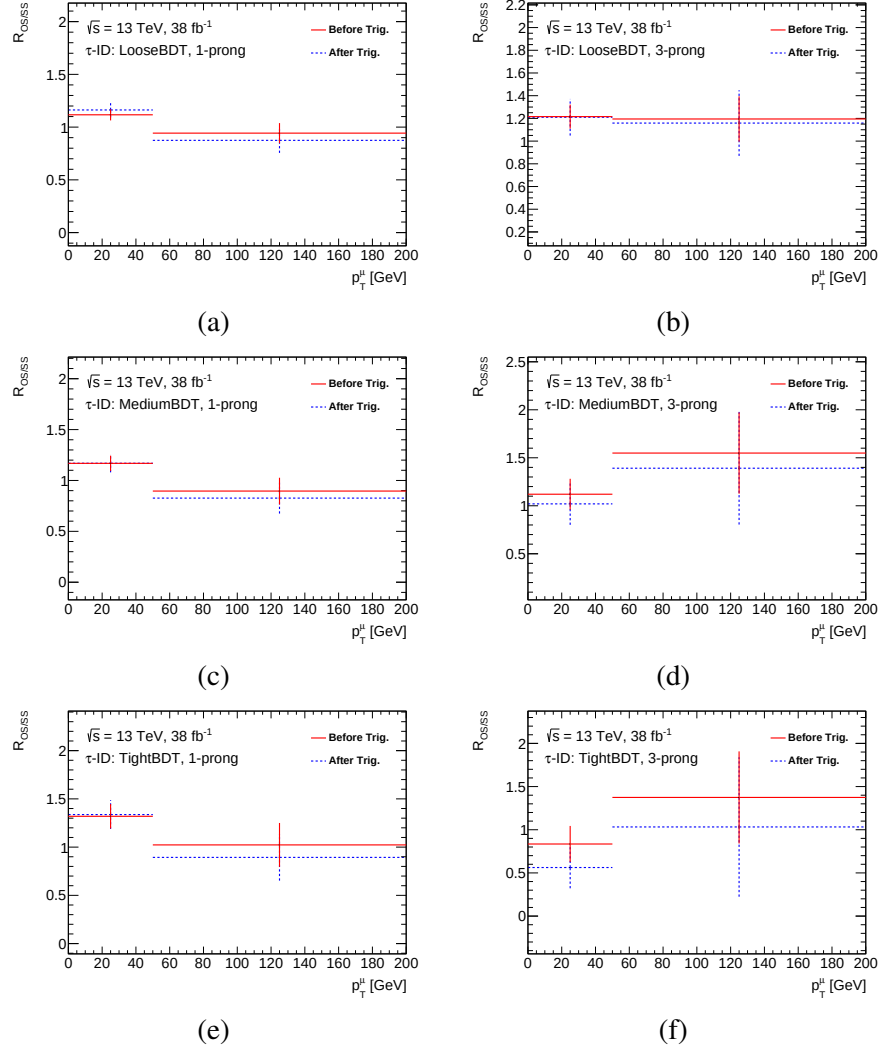


Figure 8.2: $R_{OS/SS}$ factors to calculate the SFs for the tau25 EFORMVA trigger with the 2018 data set and the BDT τ_h -identification method. From top to bottom for the Loose, Medium and Tight working point. On the left-side for τ_h -leptons with one charged track and on the right-side with three charged tracks. The $R_{OS/SS}$ values for the signal region before (red) and after applying the τ_h -trigger requirement (dashed blue) are shown.

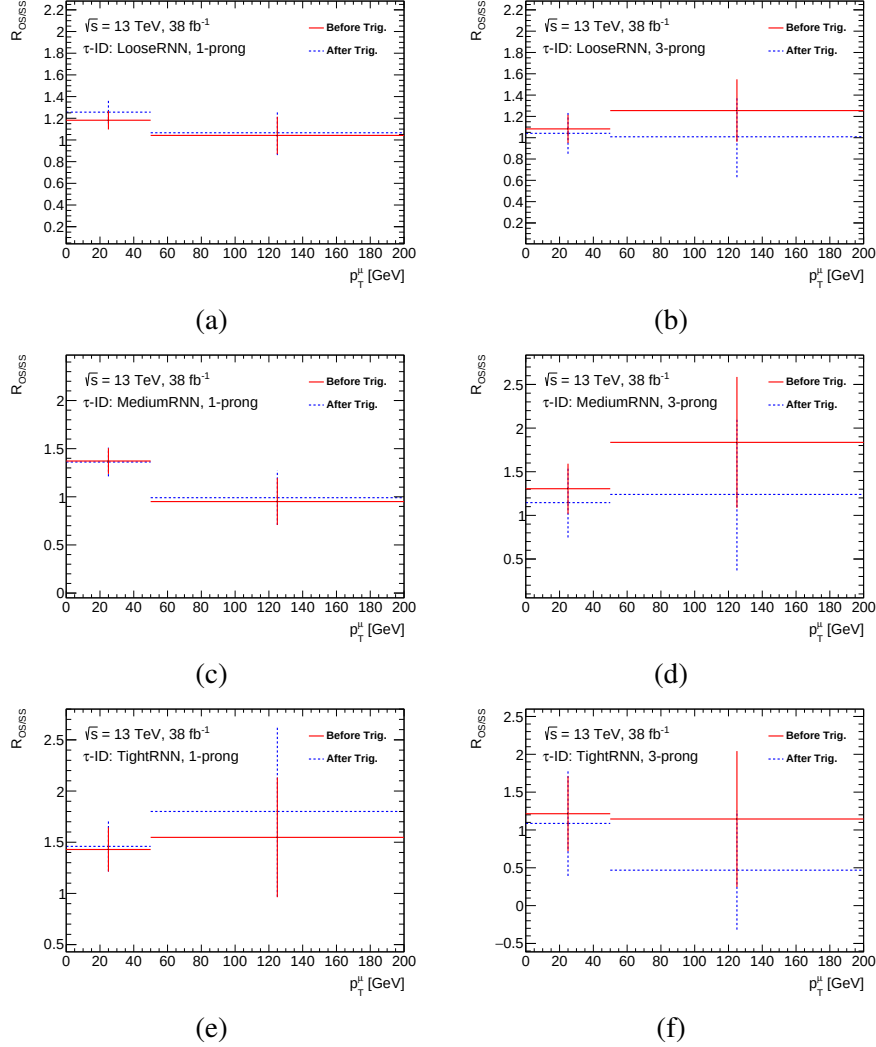


Figure 8.3: $R_{OS/SS}$ factors to calculate the SFs for the tau25 EFORMVA trigger with the 2018 data set and the RNN τ_h -identification method. From top to bottom for the Loose, Medium and Tight working point. On the left-side for τ_h -leptons with one charged track and on the right-side with three charged tracks. The $R_{OS/SS}$ values for the signal region before (red) and after applying the τ_h -trigger requirement (dashed blue) are shown.

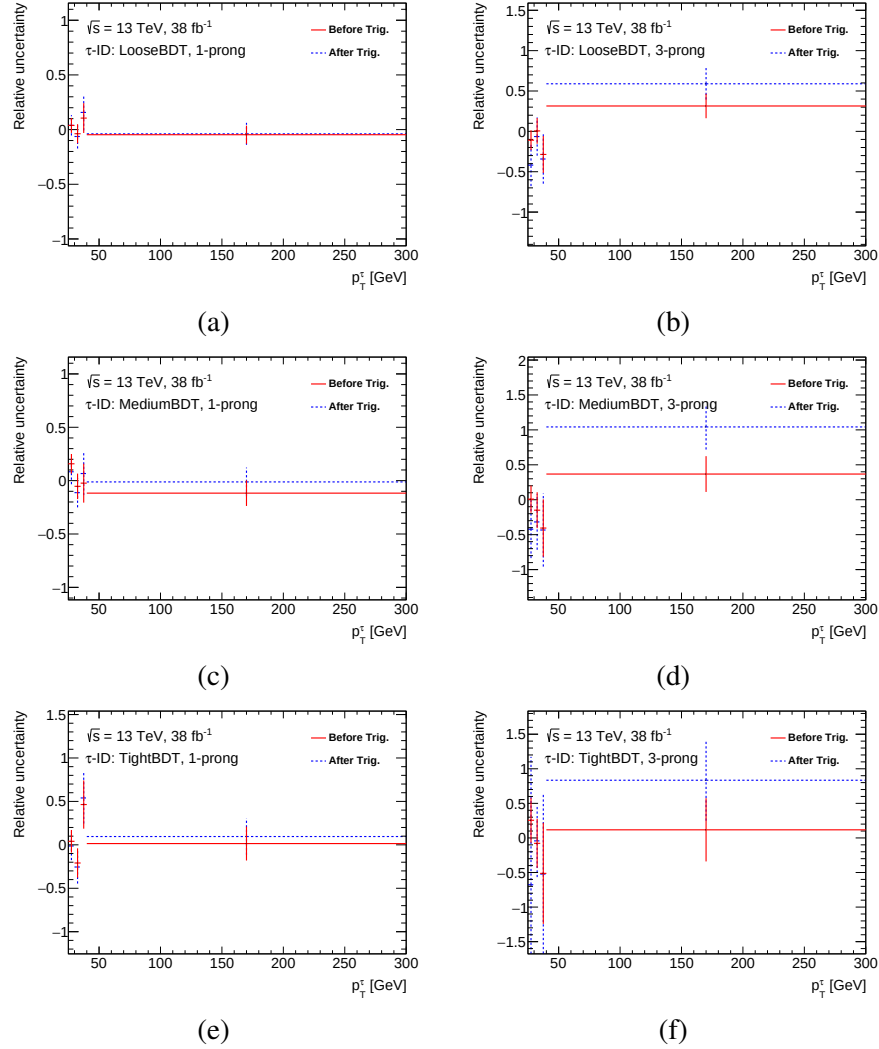


Figure 8.4: Shape variation on $R_{OS/SS}$ factors for the tau25 EFORMVA trigger with the 2018 data set and the BDT τ_h -identification method. From top to bottom for the Loose, Medium and Tight working point. On the left-side for τ_h -leptons with one charged track and on the right-side with three charged tracks. The $R_{OS/SS}$ values for the signal region before (red) and after applying the τ_h -trigger requirement (dashed blue) are shown.

Treatment of systematic uncertainties No systematic uncertainties assigned to theoretical uncertainties are considered in the analysis, while all uncertainties assigned to experimental sources are applied. The uncertainties related to the luminosity, pile-up re-weighting, muons, τ_h -leptons, jets, b-jets and the QCD multi-jet background estimation are included in the analysis. Table 8.4 summarizes the number of individual uncertainties considered in the analysis. The table splits these numbers into different sources of uncertainties for each object. More detailed discussions of the individual uncertainties are given in Section 9.4 and 10.4. All uncertainties are applied on all simulated events used to estimate $\varepsilon_{\text{data}}$ and

Leptons and jets			
Uncertainty source	Muon	τ -lepton	Jet
All	13	13	33
Resolution	3		8
Energy scale	1	2	23
Reconstruction/Identification	3	11	
Isolation	2		
Trigger	4		
Jet vertex tagger			2
QCD multi-jet estimation (RQCD)			
All	2		
Others			
Uncertainty source	b-jet	Luminosity	Pile-Up (PRW)
All	13	1	1

Table 8.4: Number of NPs related to experimental uncertainty in the fit model grouped into different sources.

ε_{MC} (see Equation 8.1). The used simulated events in the nominator and denominator of these efficiencies have a large overlap, since the calculated efficiencies are close to one. Therefore, all systematic uncertainties are treated as fully correlated before and after applying the τ -trigger requirement, by varying the nominator and denominator for each uncertainties simultaneously.

Modeling of the observables The most important variables of the analysis are $p_T^{\tau_h}$ and p_T^μ . $p_T^{\tau_h}$ is used as the final observable in which the SFs are calculated, while p_T^μ is used to calculate the shape variation of the QCD multi-jet background estimation. Figure 8.5 (a) shows the modeling for the $p_T^{\tau_h}$ observable in the SR of the 2018 data set for the Loose working point of the BDT τ_h identification requirement and 1-prong τ_h -leptons and Figure 8.5 (b) shows the p_T^μ distribution for the same configuration. For both variables the simulated events agree with the recorded data within their uncertainties. For $p_T^{\tau_h}$ a slight trend is visible in the ratio panel within the uncertainty band. The QCD multi-jet background only has a sizable contribution of both variables for lower values. The measurement of the efficiencies and the SFs are rather robust against mismodeling of the variables, since mismodeling would affect the nominators and denominators similarly and can therefore be canceled.

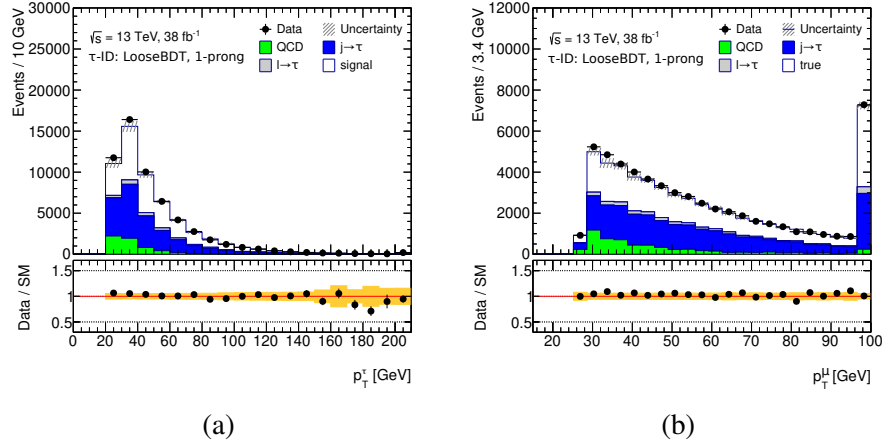


Figure 8.5: Observed events in the signal region of the 2018 data set as a function of $p_T^{\tau_h}$ (a) and p_T^μ (b) for the Loose working point of the BDT τ_h identification requirement and 1-prong τ_h -leptons. The data is shown with black dots, the signal in white, QCD multi-jet background in green, other backgrounds with jets misidentified as τ_h -leptons in blue and other backgrounds with electrons or muons misidentified as τ_h -leptons in gray.

8.4 Results

The estimated efficiencies ϵ_{data} and ϵ_{MC} , as well as the final SFs for the tau25 EFORMVA trigger setup are shown for the BDT τ identification working points in Figure 8.6 and for RNN in Figure 8.7. The ϵ_{data} and ϵ_{MC} agree for most of the displayed configurations and $p_T^{\tau_h}$ bins within their uncertainties. Therefore, most of the SFs agree within their uncertainties. The efficiencies are increasing with higher values of $p_T^{\tau_h}$ and are above 95% for most of the bins. The efficiencies raise faster with $p_T^{\tau_h}$ for 1-prong τ_h -leptons compared to 3-prong and for tighter identification working points. No systematic trend between the efficiencies of the BDT and RNN method is visible, but the uncertainties are smaller for the RNN method. This is an effect of the higher efficiencies of the RNN working point compared to the corresponding BDT working point. Due to higher efficiency more data events are selected at the working point and the statistical uncertainty is reduced. The impact of systematic uncertainties on ϵ_{data} and SF is only sizable for low $p_T^{\tau_h}$ values and especially for 3-prong τ_h -leptons.

Figure 8.8 shows the relative impact of the systematic uncertainties assigned to the different sources discussed in Section 8.3.2 for the SFs displayed in Figure 8.6. For 1-prong τ_h -leptons none of the SFs have a relative uncertainty larger than 1% sourced by one of the uncertainty groups, while for the 3-prong τ_h -leptons the uncertainties are up to 3.1%. In most of the cases a clear trend of smaller uncertainties for higher values of $p_T^{\tau_h}$ is visible. The largest values are in general assigned to the systematic uncertainties sourced by the QCD multi-jet estimation (RQCD). The small impact of systematic uncertainties is expected, due to a large overlap of events in the nominator and denominator of the efficiencies, which result in a cancellation of the systematic uncertainties.

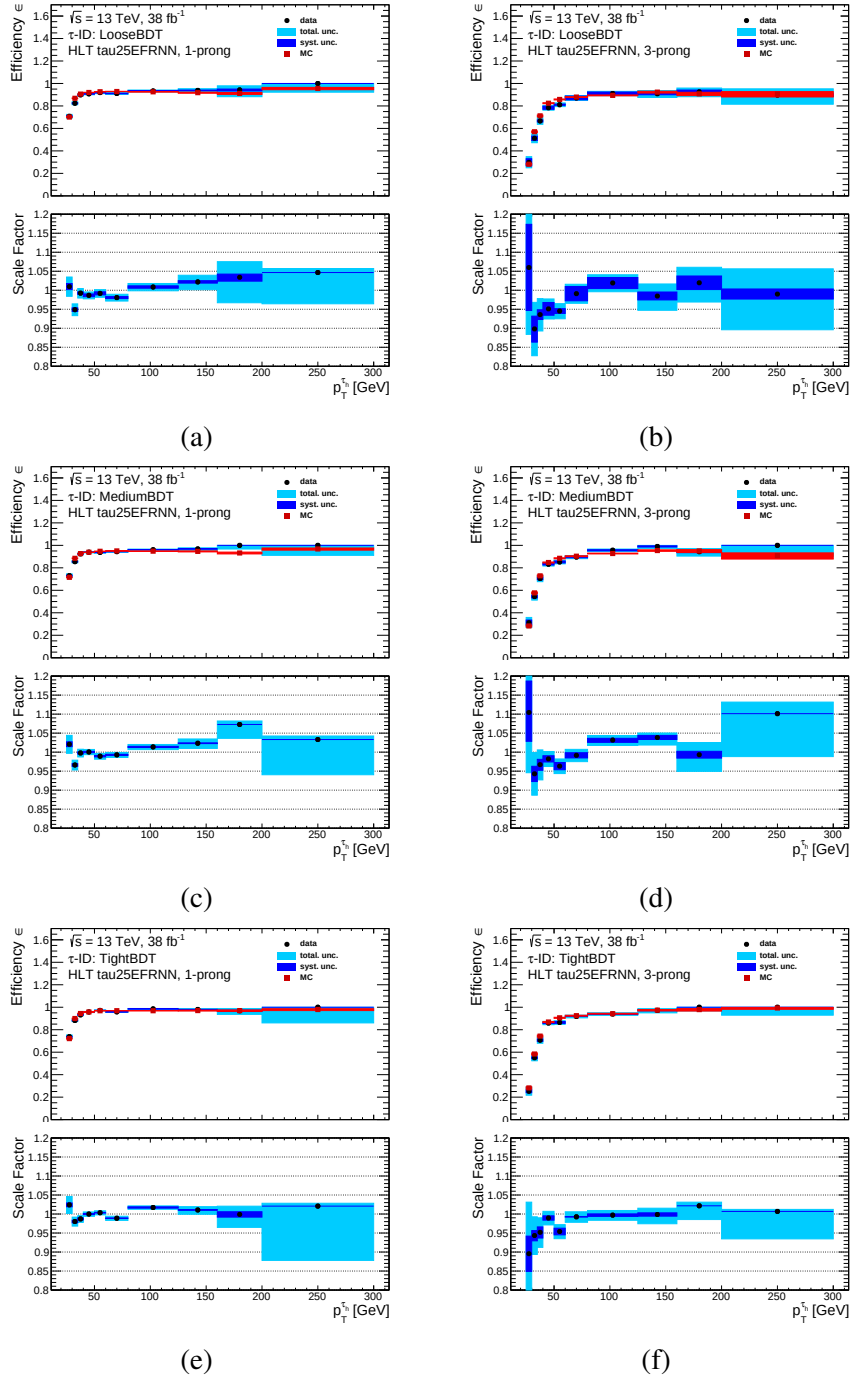


Figure 8.6: Trigger efficiencies (upper panel) in data (blue) and simulation (red) and SF (lower panel) for the tau25 EFORMVA trigger, estimated with the 2018 data set. From top to bottom for Loose, Medium and Tight BDT τ identification working point and left for 1-prong and right for 3-prong τ -leptons. For the data efficiencies and the SFs the systematic uncertainty is shown in dark blue and the total uncertainties in azure.

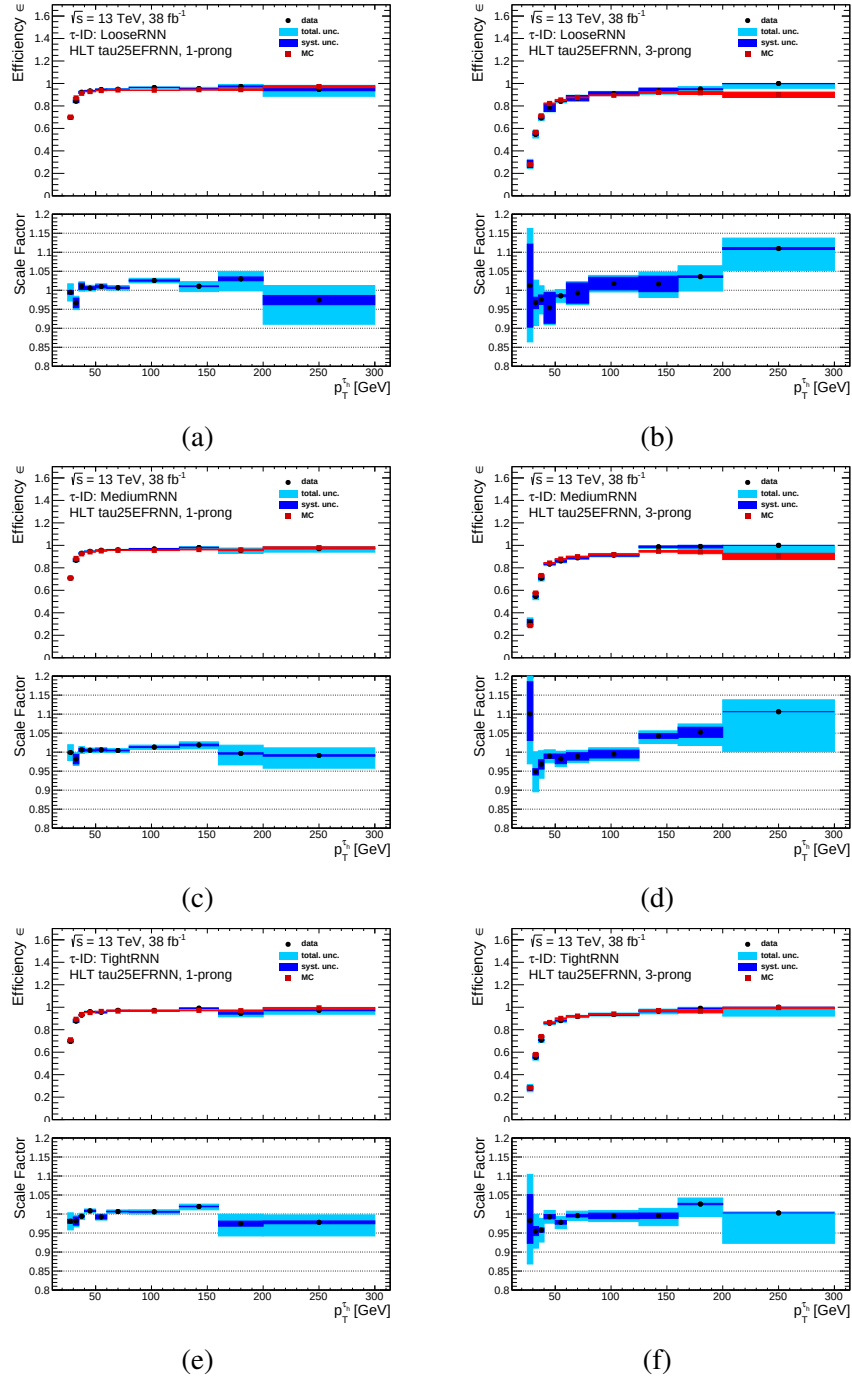


Figure 8.7: Trigger efficiencies (upper panel) in data (blue) and simulation (red) and SF (lower panel) for the tau25 EFORMVA trigger, estimated with the 2018 data set. From top to bottom for Loose, Medium and Tight RNN τ identification working point and left for 1-prong and right for 3-prong τ -leptons. For the data efficiencies and the SFs the systematic uncertainty is shown in dark blue and the total uncertainties in azure.

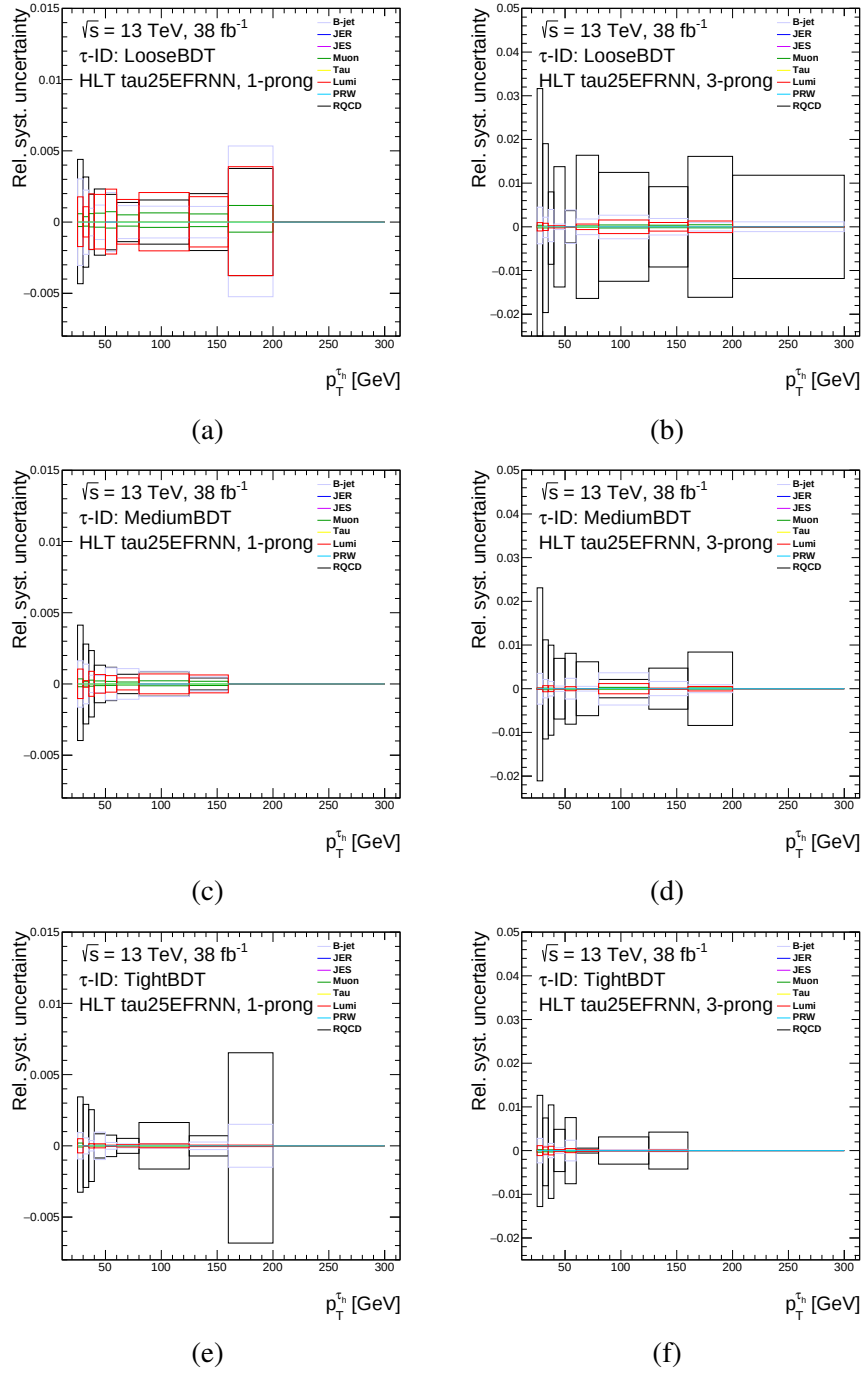


Figure 8.8: Relative uncertainty on the SF assigned to a sub-set of systematic uncertainties for the tau25 EFORMVA trigger and BDT identification criteria. The groups follow Table 8.4, except for the jet related systematic uncertainties which are split additionally into jet energy resolution (JER) and other jet related (JES) uncertainties.

8.5 Combination

The results of the $t\bar{t} \rightarrow \tau_\mu \tau_h$ tag-and-probe analysis are combined with a second tag-and-probe analysis based on $Z \rightarrow \tau_\mu \tau_h$ events using the combination framework BLUE [230, 231]. The uncertainties are treated as uncorrelated between the two analyses.

Figure 8.9 shows the combination for the tau25 EFORMVA trigger using the 2018 data set for the Medium RNN τ_h -ID and 1-prong τ -leptons. The comparison of the SFs of the single analyses and the combined values agree within their uncertainties for all bins (see Figure 8.9 (a)). The weight distribution shown in Figure 8.9 (b), corresponds to importance of the analyses for the combined value in each bin. A value of one means that combined value is derived using only this analysis, while a value of zero means that this analysis has no impact on the combined value. This distribution clearly shows the importance of both analyses. For lower values of $p_T^{\tau_h}$ (<40 GeV) the combination is clearly dominated by the $Z \rightarrow \tau_\mu \tau_h$ analysis, while for higher values the $t\bar{t} \rightarrow \tau_\mu \tau_h$ analysis is dominating. The pulls shown in Figure 8.9 (c), indicate if the combined value is pushed to larger or smaller values by the analysis. A value above zero shows that the single analysis has a SF above the combined value and therefore pulls the combined value to larger values, while a value below zero indicates the opposite. In general the $t\bar{t} \rightarrow \tau_\mu \tau_h$ and $Z \rightarrow \tau_\mu \tau_h$ analyses are alternating, such that no systematic trend of an over or under estimation of the SFs in one analysis is visible. The distribution of the combined SFs with the statistical and systematic uncertainties (Figure 8.9 (d)) illustrates that the combined measurement is dominated by statistical uncertainties. The systematic uncertainties are sizable for lower values of $p_T^{\tau_h}$, while for large values the statistical uncertainty is more and more dominating. Overall the SFs are between 0.95 and 1.05 for this configuration and therefore correspond to a correction of the simulations of up to 5%.

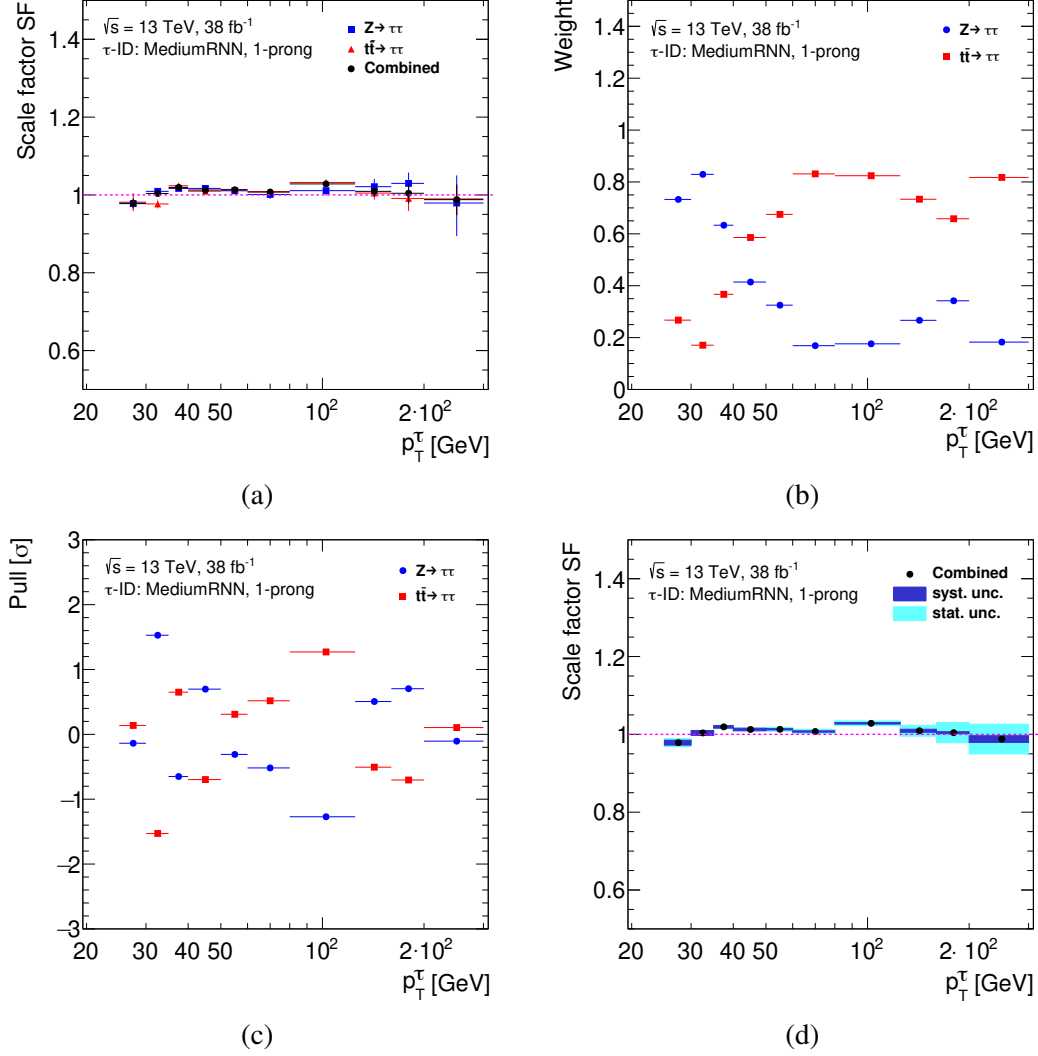


Figure 8.9: Results for the combination of the SF calculated by the $t\bar{t} \rightarrow \tau_\mu \tau_h$ (red) and $Z \rightarrow \tau_\mu \tau_h$ (blue) tag-and-probe analysis for the tau25 EFORMVA trigger using the 2018 data set for the Medium RNN τ_h -ID and 1-prong τ -leptons. The SFs of the two analyses and combination (a), the weights of the two analyses for each combined SF (b), the pull of the two analyses on each combined SF (c) and the combined SFs with their statistical and systematic uncertainties (d) are shown [232].

9 $H \rightarrow \tau\tau$ ANALYSIS WITH THE 2015 AND 2016 DATA SET

The analysis based on the Run 1 data set record in 2011 and 2012 by the ATLAS experiment achieved evidence for the $H \rightarrow \tau\tau$ coupling [233]. The main goal of this analysis is the observation of the decay channel. The analysis focuses on the two main production modes of the Higgs boson at the LHC: ggF and VBF and uses the data set recorded in 2015 and 2016 with an integrated luminosity of 36 fb^{-1} at a center-of-mass energy of 13 TeV. In addition, the measurement of the inclusive cross sections of these production modes and the first STXS measurement in this decay mode are performed. The results are published in Reference [26].

9.1 Signal and background processes

All main production modes of the Higgs boson decaying into the di- τ final state are considered as signal for the $H \rightarrow \tau\tau$ measurement (see Table 4.2). All three decay modes of the di- τ -system introduced in Section 6.1.3 are covered by the analysis.

The main irreducible background for the $H \rightarrow \tau\tau$ signal are $Z/\gamma^* \rightarrow \tau\tau$ events. Single top quark and top quark pair production are a significant background contribution in several regions of phase space included in the analysis. Additionally, $Z \rightarrow ll$ and di-boson production contribute as a significant background to the $H \rightarrow \tau_1\tau_1$ final state. Here, $Z \rightarrow ll$ includes only the events decaying into electrons and muons, while the $Z/\gamma^* \rightarrow \tau\tau$ events are treated separately. In all final states W/Z boson, top quark and QCD multi-jet production contribute as the so-called misidentified τ background, in which light leptons (electrons and muons) or jets are misidentified as τ -leptons and therefore mimic the signal.

9.2 Event selection

This section summarizes the signal event selection of the analysis. The selection is split into three different stages: Trigger selection, preselection and the selection of signal categories. The first two stages are defined to increase the $H \rightarrow \tau\tau$ signal over background ratio, while the signal categories are designed to further increase this ratio, but also to increase the signal contribution of a specific Higgs boson production mode compared to remaining production modes.

9.2.1 Trigger requirement

Only events which pass a specific set of triggers are used in the analysis. The trigger menu has changed between 2015 and 2016, so that different triggers are used for each year. The lowest unscaled triggers are summarized in Table 9.1. For each trigger an offline analysis selection requirement on the transverse momentum of the trigger object is applied, which are summarized in Table 9.2. These thresholds are slightly higher as the corresponding trigger thresholds, to avoid the use of trigger efficiency correction factors in the analysis, which are affected by migration effects at the trigger threshold.

In the $\tau_1\tau_1$ final state, single-lepton and di-lepton triggers are used [234, 235]. Overlaps between the triggers are excluded based on the transverse momentum of the lepton $p_T^{\tau_1}$ to

9.2 Event selection

Trigger menu	Data period	Chain name
Single electron	2015	HLT_e24_lhmedium_L1EM20VH, HLT_e60_lhmedium or HLT_e120_lhloose
	2016	HLT_e26_lhtight_nod0_ivarloose, HLT_e60_lhmedium_nod0 or HLT_e140_lhloose_nod0
Single muon	2015	HLT_mu20_iloose_L1MU15 or HLT_mu50
	2016	HLT_mu20_iloose_L1MU15 or HLT_mu50
Di-electron	2015	HLT_2e12_lhloose_L12EM10VH
	2016	HLT_2e17_lhvloose_nod0
Di-muon	2015	HLT_mu18_mu8noL1
	2016	HLT_mu22_mu8noL1
Electron-muon	2015	HLT_e17_loose_mu14
	2016	HLT_e17_lhloose_nod0_mu14
Di-hadronic τ	2015	HLT_tau35_medium1_tracktwo_tau25_medium1_tracktwo_L1TAU20IM_2TAU12IM
	2016	HLT_tau35_medium1_tracktwo_tau25_medium1_tracktwo

Table 9.1: Overview of the low unprescaled trigger of this analysis, grouped by the trigger menu and the data period for which they are used.

avoid a complex trigger efficiency calculation. Individual trigger sets are applied depending on the flavor of the light lepton and the used data period. In general the di-electron, di-muon and electron-muon triggers are used for lower values of $p_T^{\tau_l}$, while for higher values the single-electron and single-muon triggers are used.

Only the single-lepton triggers are used in the $\tau_l \tau_h$ final state, using the same settings as in the $\tau_l \tau_l$ final state.

Di-hadronic τ -triggers are used in the $\tau_h \tau_h$ final state. The two τ_h -lepton candidates are matched to the respective legs of the di- τ trigger using the individual single- τ trigger objects. The two legs of the di- τ trigger are assumed to be uncorrelated, so that the simulated events are scaled by the product of the correction factors of both single triggers. The correction factor estimation is published in Reference [236] and uses a similar analysis strategy as the analysis described in Chapter 8.

Trigger menu	Data period	Trigger p_T threshold	Analysis p_T threshold
Single electron	2015	$p_T(e) > 24 \text{ GeV}$	$p_T(e) > 25 \text{ GeV}$
	2016	$p_T(e) > 26 \text{ GeV}$	$p_T(e) > 27 \text{ GeV}$
Single muon	2015	$p_T(\mu) > 20 \text{ GeV}$	$p_T(\mu) > 21 \text{ GeV}$
	2016	$p_T(\mu) > 26 \text{ GeV}$	$p_T(\mu) > 27.3 \text{ GeV}$
Di-electron	2015	$p_T^{(1)}(e) > 12 \text{ GeV}, p_T^{(2)}(e) > 12 \text{ GeV}$	$p_T^{(1)}(e) > 15 \text{ GeV}, p_T^{(2)}(e) > 15 \text{ GeV}$
	2016	$p_T^{(1)}(e) > 17 \text{ GeV}, p_T^{(2)}(e) > 17 \text{ GeV}$	$p_T^{(1)}(e) > 18 \text{ GeV}, p_T^{(2)}(e) > 18 \text{ GeV}$
Di-muon	2015	$p_T^{(1)}(\mu) > 18 \text{ GeV}, p_T^{(2)}(\mu) > 8 \text{ GeV}$	$p_T^{(1)}(\mu) > 19 \text{ GeV}, p_T^{(2)}(\mu) > 10 \text{ GeV}$
	2016	$p_T^{(1)}(\mu) > 22 \text{ GeV}, p_T^{(2)}(\mu) > 8 \text{ GeV}$	$p_T^{(1)}(\mu) > 24 \text{ GeV}, p_T^{(2)}(\mu) > 10 \text{ GeV}$
Electron-muon	2016	$p_T(e) > 17 \text{ GeV}, p_T(\mu) > 14 \text{ GeV}$	$p_T(\mu) > 18 \text{ GeV}, p_T(\mu) > 15 \text{ GeV}$
	2016	$p_T(e) > 17 \text{ GeV}, p_T(\mu) > 14 \text{ GeV}$	$p_T(\mu) > 18 \text{ GeV}, p_T(\mu) > 15 \text{ GeV}$
Di-hadronic τ	2015	$p_T^{(1)}(\tau_h) > 35 \text{ GeV}, p_T^{(2)}(\tau_h) > 25 \text{ GeV}$	$p_T^{(1)}(\tau_h) > 40 \text{ GeV}, p_T^{(2)}(\tau_h) > 30 \text{ GeV}$
	2016	$p_T^{(1)}(\tau_h) > 35 \text{ GeV}, p_T^{(2)}(\tau_h) > 25 \text{ GeV}$	$p_T^{(1)}(\tau_h) > 40 \text{ GeV}, p_T^{(2)}(\tau_h) > 30 \text{ GeV}$

Table 9.2: Summary of the thresholds applied in this analysis on the transverse momentum p_T of the trigger objects and the corresponding p_T requirements of the triggers.

9.2.2 Preselection

For each of the three final states an individual preselection is applied to suppress background events and to ensure orthogonality to analysis of other Higgs boson decay modes. The orthogonality is necessary to avoid that data events are used twice in combined analyses as described in Chapter 11. The preselection requirements are summarized in Table 9.3.

	$\tau_l \tau_l$	$\tau_l \tau_h$	$\tau_h \tau_h$
	$\tau_e \tau_e / \tau_\mu \tau_\mu$	$\tau_e \tau_\mu$	
(1)	$N_{\tau_e/\tau_\mu}^{\text{loose}}=2, N_{\tau_h}^{\text{loose}}=0$	$N_{\tau_e/\tau_\mu}^{\text{loose}}=1, N_{\tau_h}^{\text{loose}}=1$	$N_{\tau_e/\tau_\mu}^{\text{loose}}=0, N_{\tau_h}^{\text{loose}}=2$
(2)	τ_e/τ_μ : Medium ID, gradient iso.	τ_e/τ_μ : Medium ID, gradient iso.	
(3)		τ_h : Medium ID	τ_h : Tight ID
(4)	Opposite charge	Opposite charge	Opposite charge
(5)	$m_{\tau\tau}^{\text{coll}} > m_Z - 25 \text{ GeV}$	$m_T < 70 \text{ GeV}$	
(6)	$30 < m_{ll} < 75 \text{ GeV}$	$30 < m_{ll} < 100 \text{ GeV}$	
(7)	$E_T^{\text{miss}} > 55 \text{ GeV}$	$E_T^{\text{miss}} > 20 \text{ GeV}$	$E_T^{\text{miss}} > 20 \text{ GeV}$
(8)	$\Delta R_{\tau\tau} < 2.0$	$\Delta R_{\tau\tau} < 2.5$	$\Delta R_{\tau\tau} < 2.5$
(9)	$ \Delta\eta_{\tau\tau} < 1.5$	$ \Delta\eta_{\tau\tau} < 1.5$	$ \Delta\eta_{\tau\tau} < 1.5$
(10)	$0.1 < x_1 < 1.0$	$0.1 < x_1 < 1.4$	$0.1 < x_1 < 1.4$
(11)	$0.1 < x_2 < 1.0$	$0.1 < x_2 < 1.2$	$0.1 < x_2 < 1.4$
(12)	$p_T^{jl} > 40 \text{ GeV}$	$p_T^{jl} > 40 \text{ GeV}$	$p_T^{jl} > 70 \text{ GeV}, \eta_{jl} > 3.2$
(13)	$N_{b\text{-jet}} = 0$	$N_{b\text{-jet}} = 0$	

Table 9.3: Preselection requirements for the three different decay channels $\tau_l \tau_l$, $\tau_l \tau_h$ and $\tau_h \tau_h$. In the $\tau_l \tau_l$ decay channel the same flavour $\tau_e \tau_e / \tau_\mu \tau_\mu$ and different flavour $\tau_e \tau_\mu$ final states are separated.

(1) **Two τ -leptons:**

In all three final states no additional τ -leptons are allowed. Therefore, the number of leptonically and hadronically decaying τ -leptons is required to be exactly the number of expected τ -leptons for the corresponding final state. For this requirement all τ -leptons, which pass the loose identification criteria, are taken into account (see Chapter 5).

(2)+(3) **τ -lepton identification:**

The τ_l has to fulfill the medium identification and the gradient isolation requirement described in Section 5.2 and Section 5.3. For $\tau_l \tau_h$ the τ_h must pass the medium identification criteria and for $\tau_h \tau_h$ the tight identification criteria (see Section 5.5).

(4) **Opposite charge:**

In all decay channels the two τ -leptons must have a opposite charge.

The requirements (2)-(4) on the τ -leptons are used to suppress background events in which electrons, muons or jets are misidentified as a τ -lepton.

(5)+(6) **Mass requirements:**

In $\tau_l \tau_l$, a requirement on the collinear mass (see Section 5.7) with respect to the mass of the Z boson is applied, which ensures orthogonality to the analysis focusing on the

$H \rightarrow WW$ decay mode. Additionally, the invariant mass of the di-lepton system (m_{ll}) is used in the selection. The requirement depends on the flavor composition of the di-leptonic final state and is tighter for the same-flavor state to suppress background events from $Z \rightarrow ll$. Figure 9.1 shows the m_{ll} distribution for the same-flavor and different-flavor final state before applying this requirement. In the $\tau_1 \tau_h$ channel the transverse mass:

$$m_T = \sqrt{2p_T^{\tau_1} E_T^{\text{miss}} (1 - \cos \Delta\phi_{\text{lep}, E_T^{\text{miss}}})}, \quad (9.1)$$

where $\Delta\phi_{\text{lep}, E_T^{\text{miss}}}$ is the azimuthal angle between the τ_1 -lepton and the missing energy, has to be greater than 70 GeV.

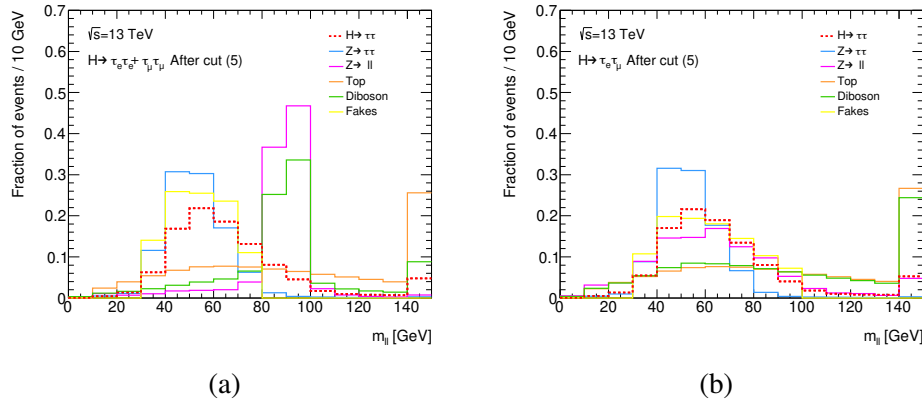


Figure 9.1: Invariant mass of the di-lepton system (m_{ll}) for different physics processes after requirement (5) for the same-flavor (a) and the different-flavor final state of the $\tau_1 \tau_1$ final state. The yield of each process is normalized to one [222].

(7) **Missing transverse energy:**

In all final states a lower threshold on the E_T^{miss} is applied to suppress background events without neutrinos, again a tighter requirement for the same flavor $\tau_1 \tau_1$ final state is applied to reduce background events from $Z \rightarrow ll$.

(8) **Angular difference:**

To reduce non-resonant background contributions, requirements on the angular distance between the visible decay products of the τ -leptons ($\Delta R_{\tau\tau}$) and on their pseudorapidity difference ($|\Delta\eta_{\tau\tau}|$) are applied in all final states.

(9) **Momentum fractions:**

To suppress background events, in which the direction of E_T^{miss} is not in agreement with the expected kinematics of the di- τ system, requirements on the momentum fractions $x_{1(2)}$, described in Equation 5.5, are applied.

(10) **Jet requirement**

As least one jet is required in each final state to increase the fraction of Higgs boson events produced by ggF and VBF productions by applying a lower threshold on the transverse momentum of the leading jet (p_T^j). While the VBF production mode produce at least two additional jets in the final state (see Section 6.1), most of the ggF events are produced without an additional jet, but requiring this boosted topology enhances the $gg \rightarrow H \rightarrow \tau\tau$ to background ratio significantly.

(11) **B-jet veto**

Background events from top quark production are suppressed by a b-jet veto in the $\tau_l \tau_l$ and $\tau_l \tau_h$ final states using the MV2c20 algorithms described in Section 5.4.1. A working point corresponding to a 85% average identification efficiency of b-jets in $t\bar{t}$ production events is used for this requirement.

9.2.3 Signal categories

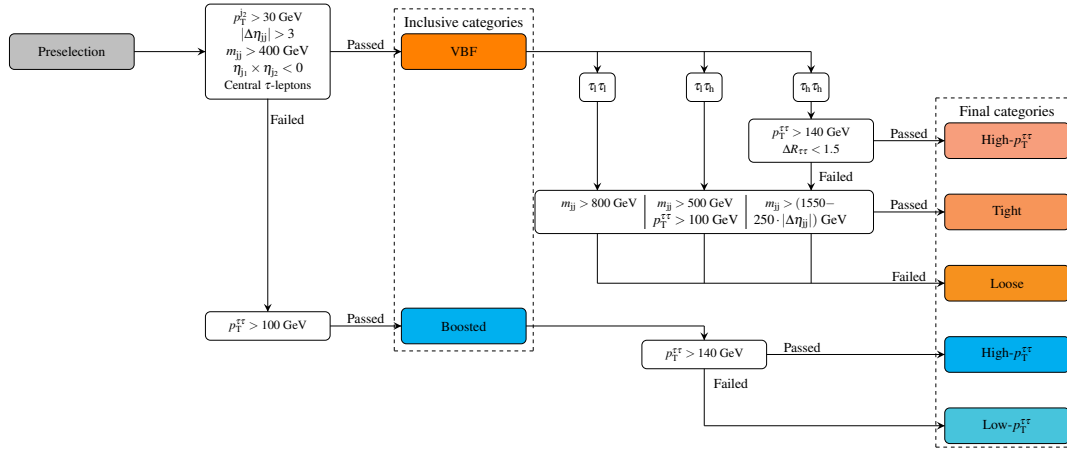


Figure 9.2: Overview of the signal category definitions. Orange boxes correspond to categories designed for the VBF production mode and blue boxes to categories designed for the ggF production mode. White boxes illustrate requirements, which have to be passed or failed by an event to end up in a specific category. For the VBF production the three different final states $\tau_l \tau_l$, $\tau_l \tau_h$ and $\tau_h \tau_h$ have individual definitions for the final signal categories.

The analysis focuses on the two main production modes ggF and VBF . Based on the preselection, an inclusive signal category per final state for each of the production modes is defined. The requirements of both inclusive signal categories are harmonized across the final states. To increase the significance, both inclusive categories are split into two sub-categories. The definitions of the sub-categories are optimized for each final state separately. Additionally, a third signal category for the VBF production mode is defined in the $\tau_h \tau_h$ final state, so that the analysis is based on 13 signal categories. The requirements for the individual signal categories are illustrated in Figure 9.2.

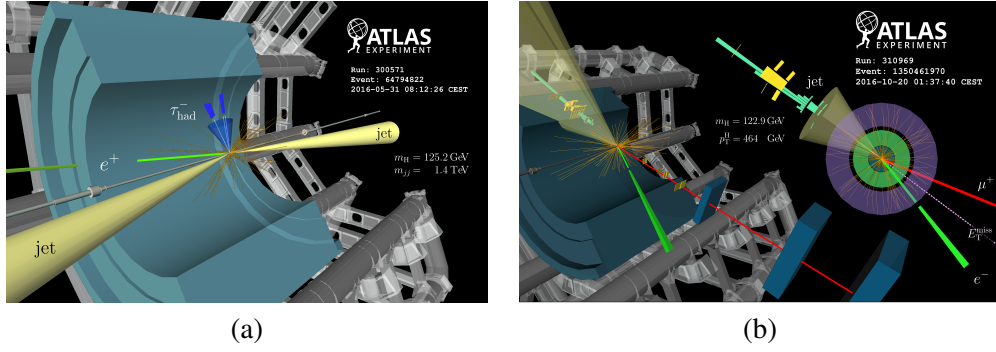


Figure 9.3: Display of an event in the inclusive VBF category of the $\tau_l \tau_h$ final state (a) and of the inclusive boosted category of the different flavor $\tau_l \tau_l$ final state (b). In (a) the transverse momenta of the τ -leptons are $p_T^{\tau_l} = 45.9$ GeV and $p_T^{\tau_h} = 32.5$ GeV. The τ_l -lepton is highlighted in green, the τ_h -lepton in blue and the two additional jets in yellow. In (b) the τ_e -lepton is highlighted in green, the τ_μ -lepton in red and the additional jets with a yellow cone [26].

VBF categories The *VBF* production mode has a specific final state signature (see Section 6.1). In addition to the particles produced in the Higgs boson decay, two jets are expected in the final state. Therefore, a second jet is required with a transverse momentum p_T^j of at least 30 GeV and the invariant mass of the di-jet system m_{jj} has to be larger than 400 GeV. Requirements on the direction of the two jets are applied by requiring a large difference in η between them ($|\Delta\eta_{jj}| > 3$) and that they are in opposite hemispheres ($\eta_{j1} \times \eta_{j2} < 0$). The η -values of the two selected τ -leptons have to be between the values of the two jets (central τ -lepton). A comparison between data and simulated events in this inclusive category is shown in Figure 9.4 as a function of m_{jj} for all three final states.

This inclusive category is split into two sub-categories for $\tau_l \tau_l$ and $\tau_l \tau_h$, while three sub-categories are defined for $\tau_h \tau_h$.

For $\tau_l \tau_l$ and $\tau_l \tau_h$ the inclusive VBF categories are split into a Tight and Loose category. In $\tau_l \tau_l$ the Tight category requires $m_{jj} > 800$ GeV, while in $\tau_l \tau_h$ the threshold is reduced to 500 GeV but supplemented with an additional requirement on the transverse momentum of the di- τ system with $p_T^{\tau\tau} > 100$ GeV. The Loose categories include all events which pass the inclusive but fail the Tight selection.

For $\tau_h \tau_h$ a High- $p_T^{\tau\tau}$ sub-category is defined with $p_T^{\tau\tau} > 400$ GeV and $\Delta R_{\tau\tau} < 1.5$. The remaining events of the inclusive VBF category are assigned to the Tight category if m_{jj} exceeds a threshold depending on $|\Delta\eta_{jj}|$: $m_{jj} > (1550 \text{ GeV} - 250 \text{ GeV} \cdot |\Delta\eta_{jj}|)$, while all other events are associated with the Loose category.

Figure 9.3 (a) shows an exemplary event display for a $\tau_l \tau_h$ candidate in the VBF Tight selection. The display visualizes the two central τ -leptons and the two additional jets in the forward and backward region of the detector.

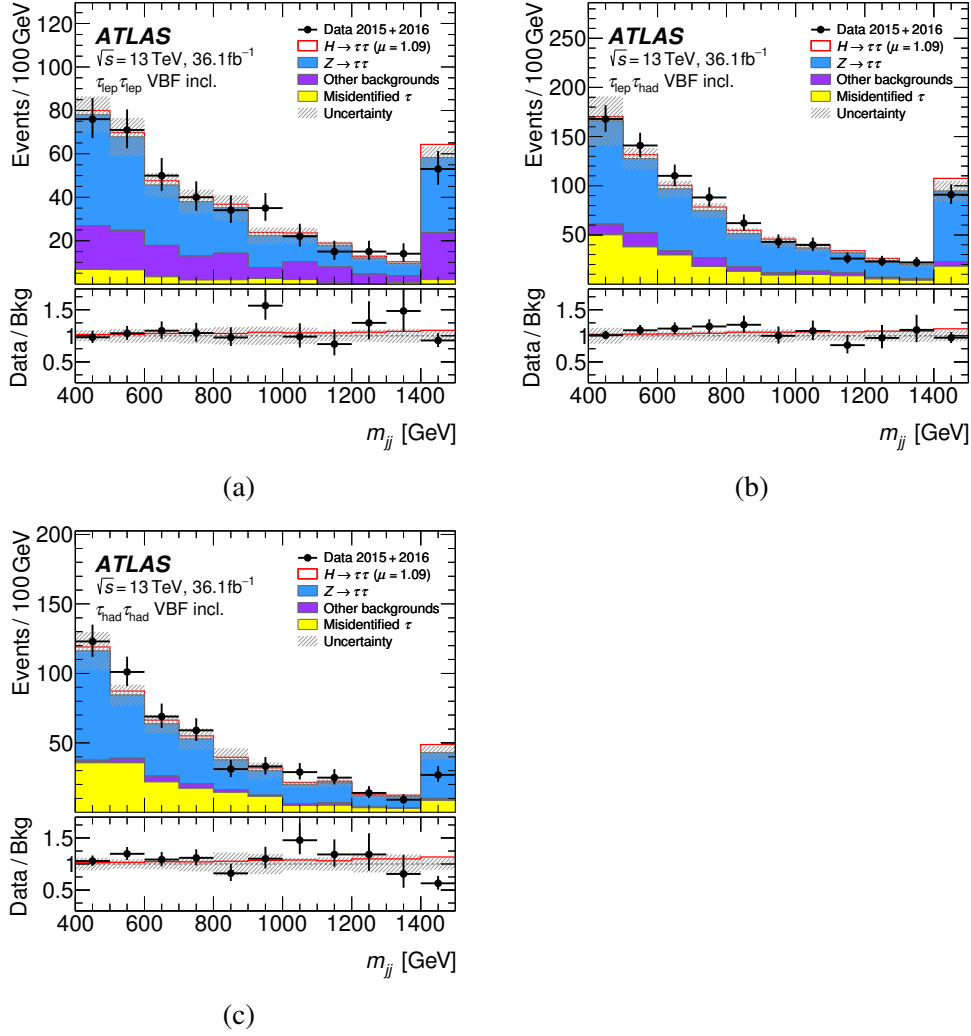


Figure 9.4: Comparisons between data and SM predictions as computed by the fit of the invariant mass of the di-jet system m_{jj} in the boosted inclusive categories for the $\tau_l \tau_l$ (a) and $\tau_l \tau_h$ (b) and $\tau_h \tau_h$ (c) final state [26].

Boosted categories To suppress background from QCD multi-jet events especially the analysis focuses on ggF production events with at least one additional jet in the final state. The Higgs boson recoils with this jet resulting in a large transverse momentum of the di- τ -lepton system. Figure 9.3 (a) shows an exemplary event display for a boosted event. The same requirements for all three final state are used to define the boosted signal categories. The events have to fail the inclusive VBF category to avoid any overlap between the selection, but pass the $p_T^{\tau\tau} > 100$ GeV requirement. A comparison between data and simulated events in this inclusive category is shown in Figure 9.5 as a function of $p_T^{\tau\tau}$ for all three final states.

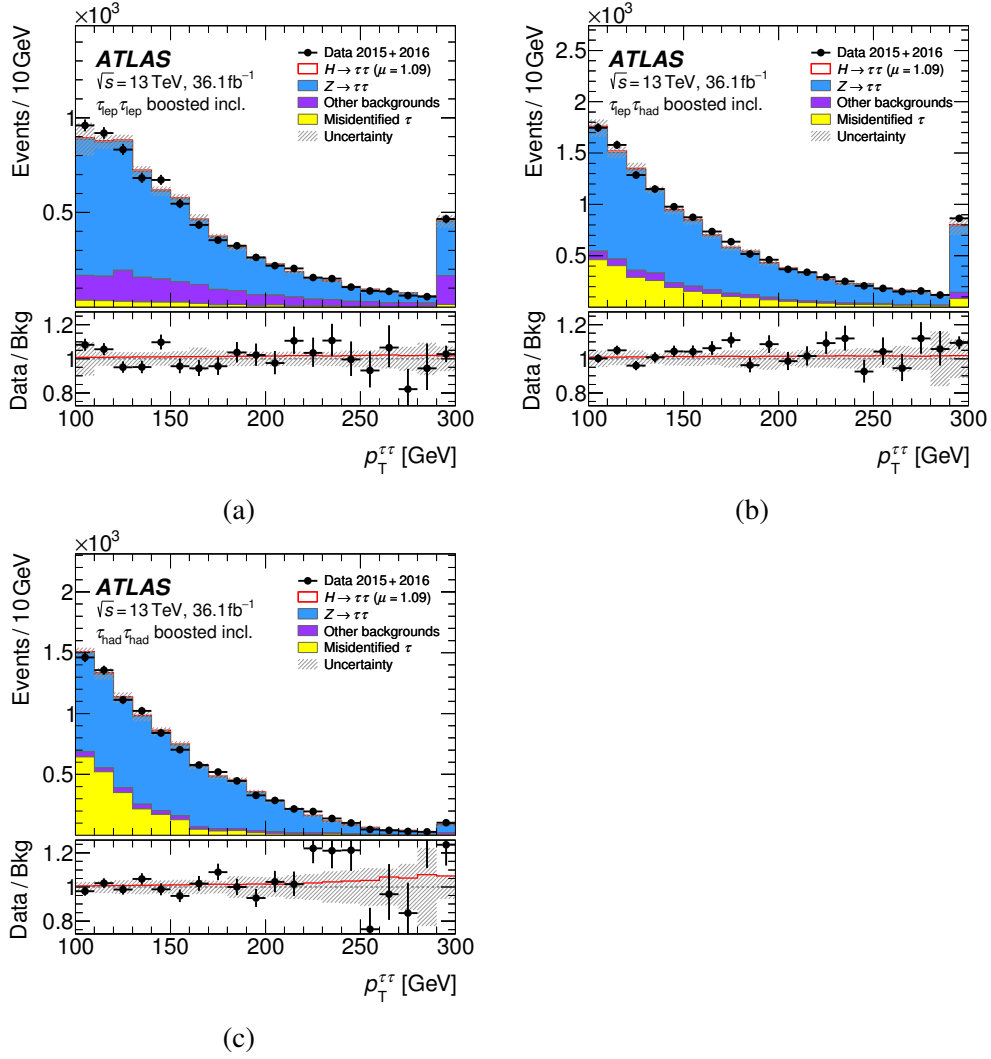


Figure 9.5: Comparisons between data and SM predictions as computed by the fit of the transverse momentum of the Higgs boson candidate $p_T^{\tau\tau}$ in the boosted inclusive categories for the $\tau_l\tau_l$ (a) and $\tau_l\tau_h$ (b) and $\tau_h\tau_h$ (c) final state [26].

Similarly to the VBF category, the inclusive boosted category is split into two sub-categories by increasing the $p_T^{\tau\tau}$ threshold to 140 GeV for the High- $p_T^{\tau\tau}$ category, while the remaining events of the inclusive boosted category are assigned to the Low- $p_T^{\tau\tau}$ category.

Process composition in signal categories The relative contribution of different physics processes in the signal categories is shown in Figure 9.6. The largest relative signal contributions are reached in the VBF categories, where the Tight (High- $p_T^{\tau\tau}$ in $\tau_h\tau_h$) have higher signal contributions compared to the Loose categories. $Z \rightarrow \tau\tau$ events are the dominant background

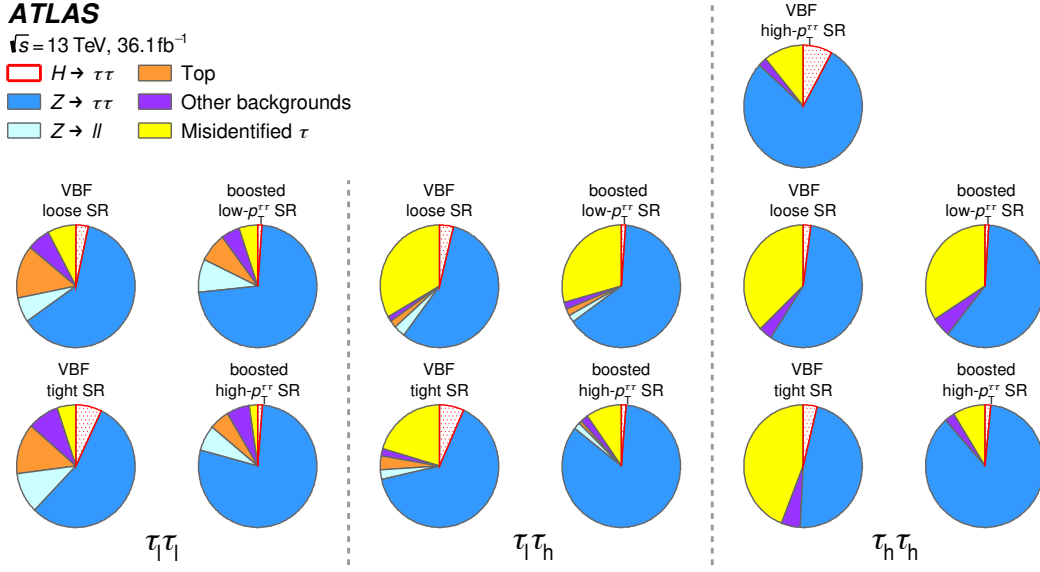


Figure 9.6: Relative contribution of processes in the signal categories. The Higgs boson signal is illustrated in white with a red frame. The background processes are grouped into $Z \rightarrow \tau\tau$ (dark blue), $Z \rightarrow ll$ (light blue), Top quark production (orange), misidentified τ (yellow) and other backgrounds (purple). Misidentified τ corresponds to event in which at least one τ -lepton is a misidentified electron, muon or jet. Other backgrounds include all remaining non-listed processes described in Chapter 6. The contributions are shown for each of the final states separately, where the left regions correspond to the $\tau_1 \tau_1$, the middle to the $\tau_1 \tau_h$ and the right to the $\tau_h \tau_h$ final state [26].

contribution in all categories. The misidentified τ background contribution increases with an increasing number of τ_h -leptons. The top quark production background has a non-negligible contribution for $\tau_1 \tau_1$ and $\tau_1 \tau_h$. $Z \rightarrow ll$ events are a sizable background in $\tau_1 \tau_1$. Data-driven corrections are used for each important background in the individual final states. Table 9.4 shows the absolute event yields for the ggF , VBF and other $H \rightarrow \tau\tau$ production modes compared to the background yields for every signal category and the preselection. The numbers illustrate the improvement of the signal to background ratio from the preselection to the final signal categories and show that the VBF and boosted categories enhance the corresponding VBF and ggF contributions.

9.2 Event selection

$\tau_l \tau_l$					
Category	Background	ggF	VBF	Other $H \rightarrow \tau\tau$	Signal/Background
Preselection	10700±100	72.9±0.5 64%	30.3±0.02 27%	10.6±0.3 9%	0.011
Boosted incl.	7700±100	60.2±0.5 72%	14.0±0.2 16%	9.6±0.3 12%	0.011
Boosted Low- $p_T^{\tau\tau}$	4230±90	30.5±0.3 73%	6.8±0.1 16%	4.5±0.2 11%	0.010
Boosted High- $p_T^{\tau\tau}$	3480±40	29.7±0.3 71%	7.2±0.1 17%	5.0±0.2 12%	0.012
VBF incl.	410±10	4.4±0.1 23%	14.8±0.2 76%	0.22±0.05 1%	0.047
VBF loose	240±10	2.5±0.1 35%	4.5±0.1 63%	0.10±0.02 2%	0.030
VBF tight	177±8	1.9±0.1 15%	10.3±0.1 84%	0.12±0.02 1%	0.070
$\tau_l \tau_h$					
Preselection	22480±170	213.4±1.9 71%	69.7±0.4 23%	18.9±0.4 6%	0.013
Boosted incl.	12711±140	125.6±1.4 76%	25.9±0.2 16%	13.9±0.3 8%	0.013
Boosted Low- $p_T^{\tau\tau}$	6461±70	61.2±1.0 75%	12.1±0.2 15%	7.9±0.2 10%	0.013
Boosted High- $p_T^{\tau\tau}$	6250±60	64.3±1.0 75%	13.8±0.2 16%	7.9±0.2 9%	0.014
VBF incl.	809±25	10.9±0.4 25%	32.5±0.3 74%	0.3±0.1 1%	0.054
VBF loose	484±17	3.9±0.3 35%	7.1±0.1 64%	0.13±0.04 1%	0.023
VBF tight	325±18	7.0±0.3 21%	25.4±0.2 78%	0.2±0.1 1%	0.100
$\tau_h \tau_h$					
Preselection	14020±110	137±2 67%	48.4±0.4 24%	17.4±0.4 9%	0.014
Boosted incl.	9640±80	105.4±1.4 74%	22.0±0.3 15%	14.9±0.4 11%	0.015
Boosted Low- $p_T^{\tau\tau}$	5590±70	49.5±1.0 74%	10.3±0.2 16%	6.4±0.2 10%	0.012
Boosted High- $p_T^{\tau\tau}$	4050±50	55.9±1.0 73%	11.8±0.2 16%	8.4±0.3 11%	0.019
VBF incl.	520±18	8.2±0.2 27%	21.5±0.3 72%	0.30±0.01 1%	0.058
VBF loose	119±9	1.3±0.1 43%	1.7±0.1 56%	0.04±0.01 1%	0.026
VBF tight	218±13	2.8±0.3 27%	7.5±0.2 72%	0.10±0.01 1%	0.048
VBF High- $p_T^{\tau\tau}$	183±9	4.2±0.2 25%	12.3±0.2 74%	0.20±0.01 1%	0.091

Table 9.4: Predicted event yields for background events, ggF production, VBF production and other $H \rightarrow \tau\tau$ production events and the relative contributions of different Higgs boson production modes to the signal for the preselection and all signal categories. The Signal over Background ratio is shown in the most right column, where all $H \rightarrow \tau\tau$ production modes are counted as signal.

9.3 Background estimation

All main background process are corrected in a data-driven way to minimize the impact of the limited background sample size. As shown in the previous section, the background composition differs between the final states, so that separate background estimations are used for each of them.

$Z \rightarrow \tau\tau$ background The Drell-Yan process $pp \rightarrow Z/\gamma^* \rightarrow \tau\tau$ is the dominant irreducible background for all signal categories. The contribution varies between 50% and 90% of the total background. To validate this important background a $Z \rightarrow \tau\tau$ validation region is defined,

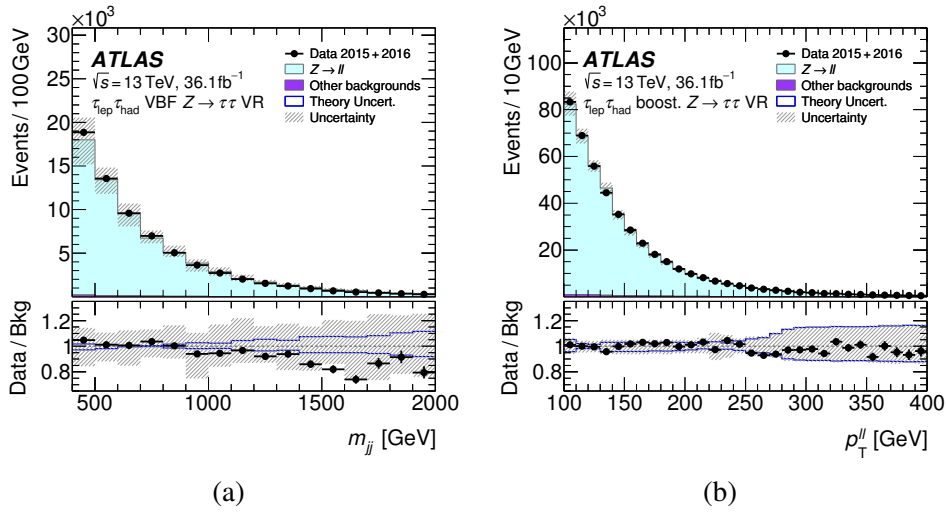


Figure 9.7: Comparison of observed $Z \rightarrow ll$ events with data in the $Z \rightarrow \tau\tau$ validation region for the $\tau_{\text{lep}}\tau_{\text{had}}$ final state as a function of the invariant mass of the di-jet system (a) and the transverse momentum of the di-lepton system (b). In (a) for the validation region of the inclusive VBF category and in (b) for the inclusive boosted category [26].

which consists of $Z \rightarrow ll$ events. This validation region is designed to test the kinematics of additional jets produced in the Z +jets production, while the modeling of the τ -decay itself has been extensively studied by a $Z \rightarrow \tau\tau$ tag-and-probe analysis [214]. Figure 9.7 shows the observed distributions of $Z \rightarrow ll$ events compared to the data as a function of two variables. The simulated events describe the data for both distributions within their uncertainties. A slight trend within the uncertainties is visible for m_{jj} . To evaluate the impact of this trend, a study is performed in which the measurement is repeated with reweighted simulated events to correct this trend. This test shows that the reweighting and therefore the visible trend has no impact on the final result. The normalization of the $Z \rightarrow \tau\tau$ background is estimated in the fit using normalization factors (NFs). These factors are correlated across the final state but separated for the VBF and boosted signal categories.

Misidentified τ background In each final state, events with misidentified τ -leptons are a significant background contribution. The contributions from light leptons misidentified as τ_h is very small and can be neglected in the misidentified τ background estimation. A specific data-driven method is used in each final state, to estimate the misidentified τ -lepton background.

In the $\tau_l \tau_l$ final state, the main sources of misidentified τ background are QCD multi-jet events,

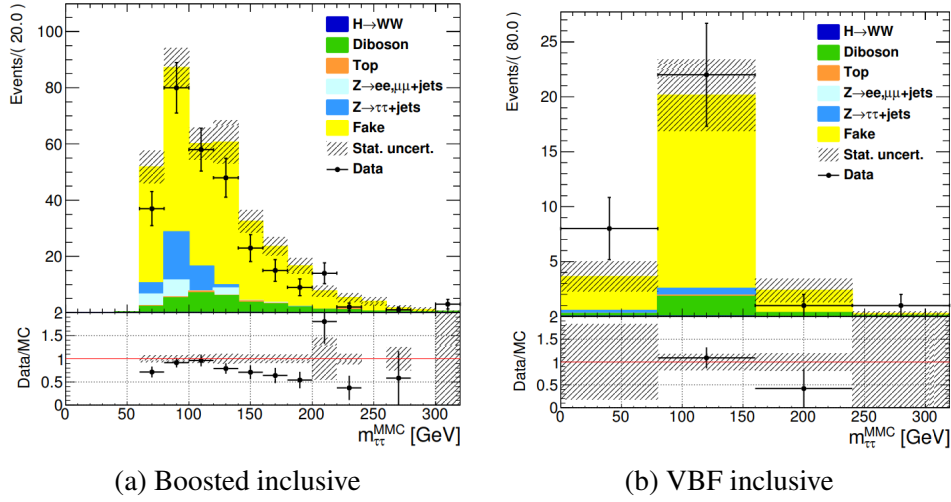


Figure 9.8: Closure test for misidentified τ background in the $\tau_l \tau_l$ final state for the boosted inclusive (a) and VBF inclusive (b) signal categories as a function of the mass of the di- τ -system.

W boson production in association with jets and semi-leptonically decaying $t\bar{t}$ production events. The contributions of these sources are treated as one background. To estimate this background a control region is defined in a phase space orthogonal to the signal category. For this region the τ_l with the second-highest p_T has to fail the gradient isolation criteria. The τ_l identification criteria is additionally relaxed to the loose working point to increase the statistics in this control region. The misidentified τ_l background in this region is finally estimated by taking the data and subtracting contributions from top quark and $Z \rightarrow ll$ background, which produce real electrons or muons. The resulting template is then transferred from the control region to the signal category by applying scale factors. These scale factors correct the difference in the acceptance of the misidentified τ background in the two regions due to inversion of the isolation criteria. They are estimated for each combination of lepton flavor from events that pass the $\tau_l \tau_l$ signal category selection but have same-sign charged leptons. Systematic uncertainties for the estimated shape and normalization of this method are calculated as a function of the p_T of the sub-leading τ_l -lepton. Events with leptons passing a same-sign charge requirement are used for a closure test (see Figure 9.8). This test results in an uncertainty ranging from 20% to 65%. For the heavy-flavor content an additional systematic uncertainty with 50% is estimated by using the isolation correction factors that are computed

from samples selected with an inverted b-jet requirement. Finally, minor contributions of uncertainties coming from the fractional composition of the misidentified τ background in top quark decays, QCD multi-jet events and W boson production are taken into account.

In the $\tau_1 \tau_h$ final state events with misidentified τ_1 -leptons can be neglected. Events with a jet misidentified as a τ_h are mainly sourced by QCD multi-jet events and W boson production in association with jets. The background is estimated in a control region defined by inverting the τ_h -lepton identification criteria. Contribution of background events in this region without a jet misidentified as a τ_h are subtracted using simulations. To transfer the background contribution from the control region to the signal region, transfer factors FF are used to correct for differences in the fake contribution due to the inversion of the identification criteria. The FF is calculated separately for QCD multi-jet events and W boson production in association with jets in control regions via:

$$FF_i = \frac{N_{CR_i}^{\text{pass } \tau\text{-ID}}}{N_{CR_i}^{\text{fail } \tau\text{-ID}}}, \quad (9.2)$$

with the number of events with a misidentified τ -lepton of the process i in the control region with applying the τ identification criteria $N_{CR_i}^{\text{pass } \tau\text{-ID}}$ and with inverting the τ identification criteria $N_{CR_i}^{\text{fail } \tau\text{-ID}}$. For W boson production in association with jets, a control region is defined by inverting the requirement on the transverse mass m_T , while for the QCD multi-jet events the τ_1 isolation criteria is inverted. The number of events with a misidentified τ -lepton in the control region of the specific process is estimated by using the data in the control region and subtracting the contribution of events without a jet that is misidentified as a τ -lepton and of other processes:

$$FF_i = \frac{N_{CR_i, \text{data}}^{\text{pass } \tau\text{-ID}} - N_{CR_i, \text{not } j \rightarrow \tau}^{\text{pass } \tau\text{-ID}} - N_{CR_i, \text{other}}^{\text{pass } \tau\text{-ID}}}{N_{CR_i, \text{data}}^{\text{fail } \tau\text{-ID}} - N_{CR_i, \text{not } j \rightarrow \tau}^{\text{fail } \tau\text{-ID}} - N_{CR_i, \text{other}}^{\text{fail } \tau\text{-ID}}}, \quad (9.3)$$

with the number of events in the control region with applying (inverting) the τ identification criteria for data $N_{CR_i, \text{data}}^{\text{pass (fail) } \tau\text{-ID}}$, for events without jets misidentified as a τ -lepton $N_{CR_i, \text{not } j \rightarrow \tau}^{\text{pass (fail) } \tau\text{-ID}}$ and for other processes $N_{CR_i, \text{other}}^{\text{pass (fail) } \tau\text{-ID}}$. FF_i shows a dependence on the number of charged tracks in the τ_h -lepton decay n_{trk}^τ and on the transverse momentum of the τ -lepton p_T^τ , so that they are estimated as a function of these observables.

To estimate the correct relative contributions of the two sources of the misidentified τ background in the signal categories a data-driven method is used. For this, a transfer factor IF from the QCD multi-jet control region to signal category is estimated using control regions defined by requiring same-sign charged τ -leptons and correcting for events with a real lepton or muon:

$$IF_i = \frac{N_{SR_{SS}, \text{data}}^{\text{fail } \tau\text{-ID}} - N_{SR_{SS}, \text{real lepton}}^{\text{fail } \tau\text{-ID}}}{N_{CR_{\text{multi-jet}, SS}, \text{data}}^{\text{fail } \tau\text{-ID}} - N_{CR_{\text{multi-jet}, SS}, \text{real lepton}}^{\text{fail } \tau\text{-ID}}}, \quad (9.4)$$

with the number of events in the control region with inverting the τ identification criteria and requiring the same-sign criteria for data (events with real electron or muon) $N_{SR_{SS}, \text{data}}^{\text{fail } \tau\text{-ID}}$ (real lepton)

and the number of these events in the control region with additionally inverting the τ isolation criteria $N_{\text{CR}_{\text{multi-jet,SS,data}}(\text{real lepton})}^{\text{fail } \tau\text{-ID}}$. This factor is estimated separately for events with an electron and a muon due to the different isolation criteria. IF is estimated as a function of the transverse momentum and of the absolute η -angle of the electron or muon p_T^l and $|\eta_l|$, to account for observed dependencies. With these factors, the relative contribution of QCD multi-jet events can be estimated, while the remaining misidentified τ background events are assumed to be sourced by W boson production in association with jets. Figure 9.9 summarizes the different control regions of this method with their requirements relative to the signal region. The transfer factors FF_i and IF can be used to estimate the contribution of events

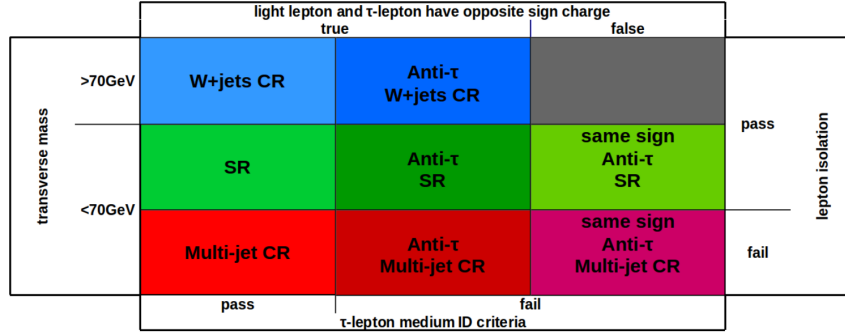


Figure 9.9: Overview of control regions used to estimate the misidentified τ background contribution in $\tau_l \tau_h$. The axis shows the different selection criteria that are changed relative to the signal region to define each control region.

with misidentified τ -lepton $N_{\text{mis. } \tau}$ as:

$$N_{\text{mis. } \tau} = \left(N_{\text{SR,data}}^{\text{fail } \tau\text{-ID}} - N_{\text{SR,not } j \rightarrow \tau}^{\text{fail } \tau\text{-ID}} \right) \cdot FF, \quad (9.5)$$

here $N_{\text{SR,data(not } j \rightarrow \tau)}^{\text{fail } \tau\text{-ID}}$ is the data (background events without a jet misidentified as a τ -lepton) in the signal region with inverted τ identification criteria and FF is the final combined transfer factor defined as:

$$FF = FF_{\text{multi-jet}} \cdot R_{\text{multi-jet}} + FW (1 - R_{\text{multi-jet}}), \quad (9.6)$$

with the individual transfer factors for the QCD multi-jet (W) background $R_{\text{multi-jet}} (W)$ and the relative contribution of QCD multi-jet events $R_{\text{multi-jet}}$. $R_{\text{multi-jet}}$ is defined with the transfer factor IF as:

$$R_{\text{multi-jet}} = \frac{IF \cdot \left(N_{\text{CR}_{\text{multi-jet,data}}}^{\text{fail } \tau\text{-ID}} - N_{\text{CR}_{\text{multi-jet,real lepton}}}^{\text{fail } \tau\text{-ID}} \right)}{N_{\text{SR,data}}^{\text{fail } \tau\text{-ID}} - N_{\text{SR,not } j \rightarrow \tau}^{\text{fail } \tau\text{-ID}}}, \quad (9.7)$$

with the number of events in the QCD multi-jet control regions with inverted τ identification criteria for data $N_{\text{CR}_{\text{multi-jet,data}}}^{\text{fail } \tau\text{-ID}}$ and simulated events with real electrons or muons

$N_{\text{CR multi-jet, real lepton}}^{\text{fail } \tau\text{-ID}}$. Observed dependencies of $R_{\text{multi-jet}}$ on n_{trk}^{τ} , p_{T}^{τ} , the lepton flavor (electron or muon) and on the difference between the τ_{h} -lepton and the missing transverse energy $E_{\text{T}}^{\text{miss}}$ in the angle ϕ ($\Delta\phi_{\tau_{\text{h}}, E_{\text{T}}^{\text{miss}}}$) are considered by estimating $R_{\text{multi-jet}}$ as a function of the observables. All factors are estimated separately for the preselection, the inclusive boosted

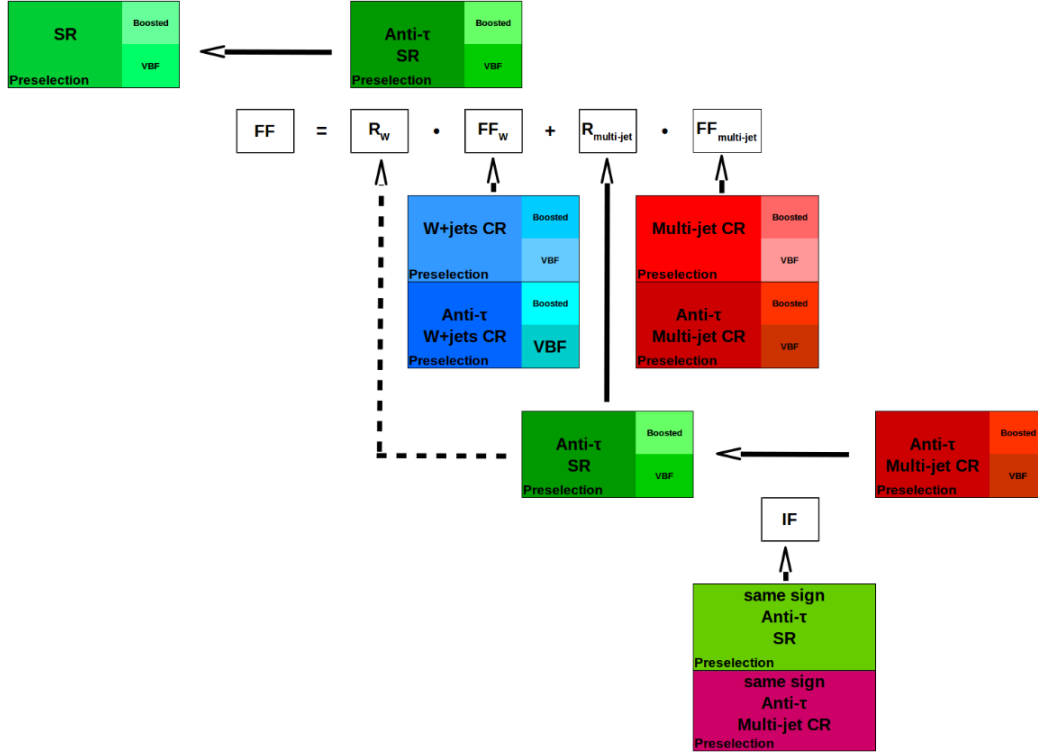


Figure 9.10: Overview of the method used to calculate the misidentified τ background contribution in the $\tau_1\tau_h$ final state. First the IF is calculated to transfer the QCD multi-jet background from the corresponding control region to the signal region with inverted τ -ID requirement. This contribution is used to estimate the relative composition of QCD multi-jet and W+jets events in this region. This is used together with the individual FF_i estimated in control regions defined for the two background processes to calculate the final transfer factor FF . FF transfers then the τ misidentified background contribution from the signal region with inverted τ -ID requirement to the signal region.

category and VBF category, except for IF , which is calculated at preselection level for all categories, to increase the data statistics in the estimation. Figure 9.10 summarizes the full fake factor method of the $\tau_1\tau_h$ final state by showing all used control regions and the transfer factors that are used between the different regions. Figure 9.11 shows the estimated IF and a sub-set of $R_{\text{multi-jet}}$ as well as FF_i values. The full set of $R_{\text{multi-jet}}$ and FF_i values are shown in Figures A.1-A.5 in the appendix. The values of IF increase for electrons with increasing $|\eta_l|$ and p_{T}^{τ} , while the opposite trend is visible for muons. The $R_{\text{multi-jet}}$ values for events

9.3 Background estimation

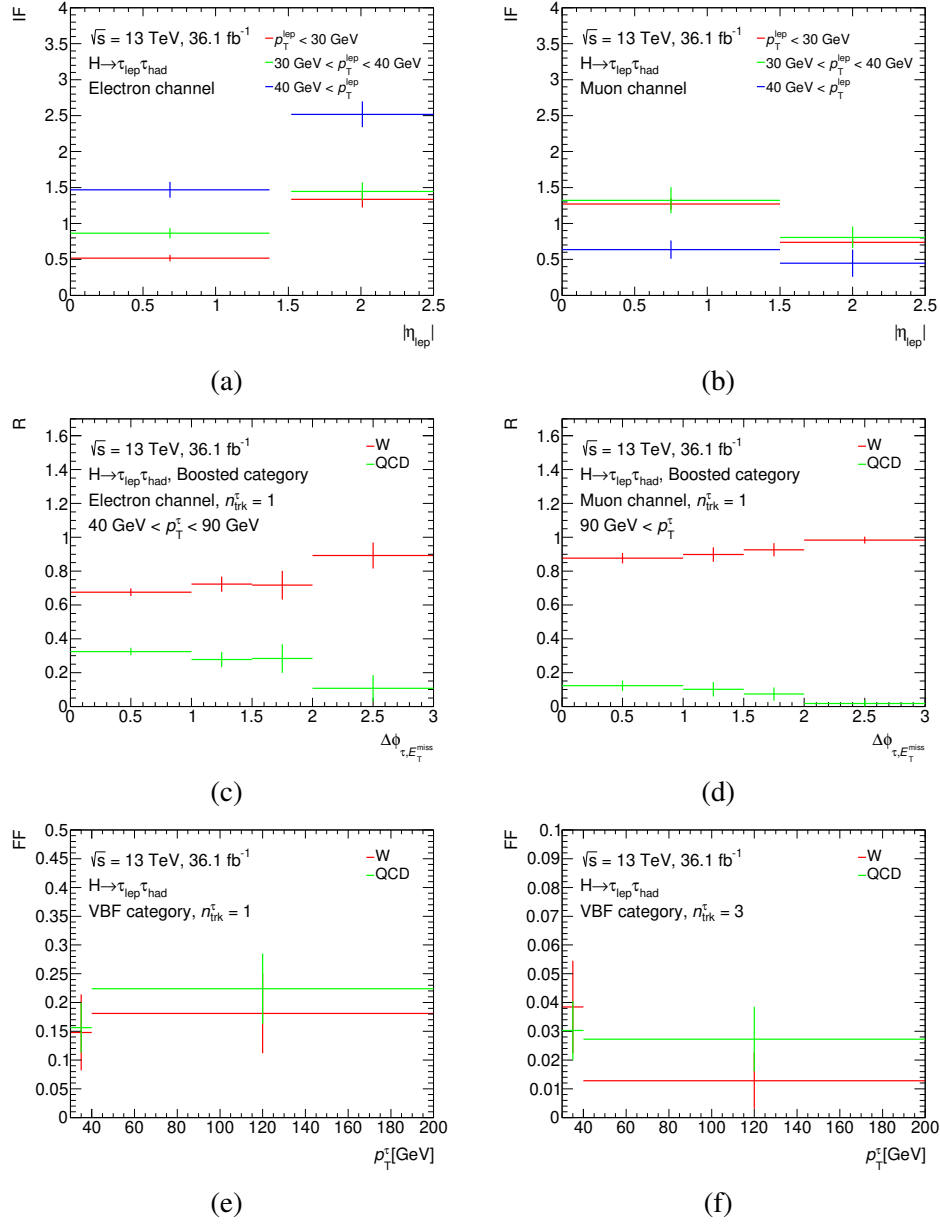


Figure 9.11: IF values for electrons (a) and muons (b) as a function of p_T^τ and $|\eta_l|$ are shown in the top row. The middle row shows the $R_{\text{multi-jet}}$ for the boosted inclusive category with $n_{\text{trk}}^\tau = 1$, for electrons and $40 < p_T^\tau < 90 \text{ GeV}$ (c) and for muons and $p_T^\tau \geq 90 \text{ GeV}$ (d) as function of $\Delta\phi_{\tau, E_T^{\text{miss}}}$. In the bottom row the FF_i values for QCD multi-jet and W events in the VBF inclusive category with $n_{\text{trk}}^\tau = 1$ (e) and $n_{\text{trk}}^\tau = 3$ (f) are displayed as a function of p_T^τ [224].

with a muon and for events with an electron of the boosted category with $n_{\text{trk}}^\tau = 1$ decrease with increasing $\Delta\phi_{\tau_h, E_T^{\text{miss}}}$. In general, the QCD multi-jet contribution is smaller for events including a muon and decreases with increasing p_T^τ . The FF values are larger for $n_{\text{trk}}^\tau = 1$ events as for $n_{\text{trk}}^\tau = 3$ events and for QCD multi-jet events as for W boson background events. For events with $n_{\text{trk}}^\tau = 1$ the FF values increase with increasing p_T^τ , while a decrease is visible for events with $n_{\text{trk}}^\tau = 3$.

Four different groups of systematic uncertainties are considered for this method. The first group is the statistical uncertainties of the events in all control regions with values between 3% and 33%. In addition, a systematic uncertainty on $R_{\text{multi-jet}}$ is estimated by varying the subtracted background in the calculation of the relative contribution by 35%, where 35% corresponds to the largest difference in the background contributions observed between the same-sign and opposite-sign control regions. This results in a relative systematic uncertainty between 1% and 3% on the final misidentified τ background. Finally, a closure test is performed in a control region with two same-sign charged τ -leptons. Any non-closure that exceeds the statistical uncertainties is taken into account as an additional systematic uncertainty (see Figure 9.12).

In $\tau_h \tau_h$ QCD multi-jet contributions are modeled using a template extracted from data in a

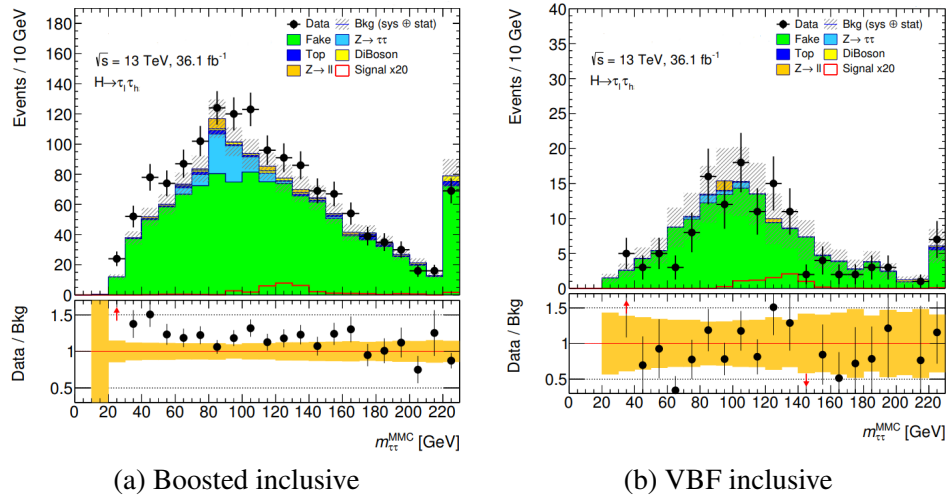


Figure 9.12: Closure test for misidentified τ_h background in the $\tau_l \tau_h$ final state for the boosted inclusive (a) and VBF inclusive (b) signal categories as a function of the mass of the di- τ -system.

dedicated control region, for which same-sign charged τ -leptons are required. To increase the statistics of this control region, events with τ_h -candidates with two tracks are added. Contributions of other background processes are subtracted. Scale factors depending on the difference in ϕ of the τ -leptons ($\Delta\phi_{\tau\tau}$) are used to transfer this background to the signal categories. The scale factors are derived by comparing the template in the control region with a second control region in which the τ -leptons have opposite charge and the τ -lepton with the

second-highest p_T fails the tight identification criteria but passes the medium requirement. To gain statistical power, the requirements on $\Delta R_{\tau\tau}$ and $|\Delta\eta_{\tau\tau}|$ are loosened compared to the preselection. The normalization of the misidentified τ background in the $\tau_h\tau_h$ final state is constrained in the fit by data within the signal categories. For this an NF is defined, which is correlated across all $\tau_h\tau_h$ signal categories.

Top quark production background As Figure 9.6 shows background events from top quark production are a negligible contribution in the $\tau_h\tau_h$ final state, so that the top quark production background is taken from simulation. For the remaining final states, a significant contribution of this background is expected. In these final states the normalization of the top quark production background is estimated in the fit via NFs, while the templates are taken from simulation. For the estimation of these NFs, additional control regions are included into the fit model. One control region per final state and per inclusive signal category is defined by inverting the b-jet veto requirement of the signal categories. Figure 9.13 shows the relative contribution of the physics processes in these control regions. The purity of the top quark production background varies between 55 and 75%. The NFs of these control regions are correlated between the sub-categories of one inclusive signal category in each final state.

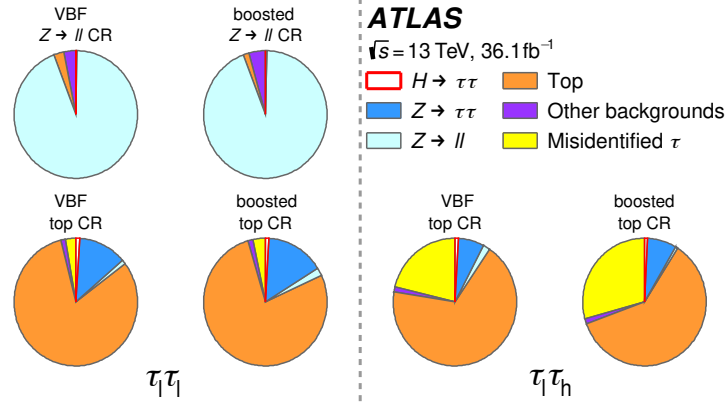


Figure 9.13: Relative contributions of processes in the control regions. The Higgs boson signal is illustrated in white with a red frame. The background processes are grouped to $Z \rightarrow \tau\tau$ (dark blue), $Z \rightarrow ll$ (light blue), Top quark production (orange) misidentified τ (yellow) and other backgrounds (purple). The contributions are shown for each final states separately, where the left regions correspond to the $\tau_1\tau_1$ and the right to the $\tau_1\tau_h$ final state [26].

$Z \rightarrow ll$ background $Z \rightarrow ll$ events are an important background in $\tau_1\tau_1$. The normalization of this background is estimated in the fit via NFs. Control regions are used to constrain these NFs. The control regions are defined by requiring the same-flavor $\tau_1\tau_1$ inclusive signal selection but changing the m_{ll} requirement from $30 < m_{ll} < 75$ GeV to $80 < m_{ll} < 100$ GeV. The relative contribution of the physics processes in these regions are shown in Figure 9.13.

Both regions show a purity of $Z \rightarrow ll$ background events of over 90%. The normalization of the $Z \rightarrow ll$ background is correlated across the control region and the corresponding signal sub-categories. In the remaining final states, this background is taken from simulation and normalized to the theoretical prediction of the SM.

9.4 Systematic uncertainties

In addition to the systematic uncertainties on the estimation of the misidentified τ background described in Section 9.3, systematic uncertainties from underlying physics models and detector effects are considered in the analysis. These uncertainties can be grouped into three different categories: theoretical uncertainties on the signal prediction, theoretical uncertainties on the background prediction and experimental uncertainties. The impact of these systematic uncertainties on the measurements is discussed in Sections 9.6-9.8.

Theoretical uncertainties on the signal prediction The estimation and implementation of theoretical uncertainties on the signal prediction follows the recommendation of the LHC Higgs Cross Section Working Group described in [27]. The impact of these uncertainties on the total cross section, the acceptance in different signal categories and on the shape of the $m_{\tau\tau}^{\text{MMC}}$ distribution used as the observable in the statistical model are evaluated individually. Three main sources of uncertainties are considered (see Section 4.3):

- Uncertainties due to missing higher orders in perturbation theory estimation, evaluated by variations of the QCD scales (μ_R and μ_F)
- Non-perturbative parts of the calculation, such as underlying event and hadronization modeling
- Uncertainties on experimental input parameters, such as parton density functions (PDFs) or the value of the strong coupling parameter (α_s)

The uncertainties on the total cross section are directly provided by [27] and are summarized in Table 9.5. The QCD scale uncertainties have an important impact on the ggF and ZH production, while PDF uncertainties have a similar impact on all production modes. The strong coupling strength has a relative uncertainty over 1% for the ggF production mode only. The QCD scale uncertainties are evaluated by varying the renormalization and factorization scales (μ_R and μ_F) described in Section 4.3 by the factor 0.5 and 2 with a constraint of $0.5 \geq \mu_R/\mu_F \geq 2$. The impact of these variations on the shape of the $m_{\tau\tau}^{\text{MMC}}$ distribution is generally small. The QCD scale uncertainties are correlated for the Higgs strahlung productions VH , namely WH and ZH .

For the ggF production these simple variations of the scale parameters are not sufficient. Experimental requirements of the analysis selection which impose a binning in the jets multiplicity introduce large contributions on the order of $\log p_{\text{cut}}/m_{\hat{H}}$ resulting in large uncertainties. The simple variation of the scale factors underestimates these uncertainties due to accidental cancellations. Therefore, a specific treatment of nine independent systematic uncertainties is used for the ggF production mode, which follows the description in [27].

The nine uncertainties are covering the following sources:

Process	QCD scale [%]	PDF [%]	α_s [%]
ggF	+3.9 (-3.9)	± 1.8	± 2.5
VBF	+0.4 (-0.3)	± 2.1	± 0.5
WH	+0.5 (-0.7)	± 1.7	± 0.9
ZH	+3.8 (-3.1)	± 1.3	± 0.9

Table 9.5: Relative theoretical uncertainties on total cross section for the four main production modes of the Higgs boson for QCD scale uncertainties due to missing higher orders in perturbation theory (QCD scale), parton density function input parameters (PDF) and the strong coupling constant (α_s) . [27]

- Four scale variations
 - $\Delta\mu$: Factorization and renormalization scale variations
 - $\Delta\phi$: Resummation scale variation
 - $\Delta_{\text{cut}}^{0/1} (\Delta_{\text{cut}}^{1/2})$: $0 \leftrightarrow 1$ ($1 \leftrightarrow 2$) jet bin migration
- Two VBF topology uncertainties
 - Variation of the VBF phase space
 - Veto of a third jet
- Two p_T^H -shape of the Higgs boson uncertainties
 - p_T^H 0-60 GeV $\leftrightarrow$ 60- ∞ GeV
 - p_T^H 0-120 GeV $\leftrightarrow$ 120- ∞ GeV
- One top mass dependency uncertainty (m_{top})

These uncertainties on the ggF production are expected to have a large impact on the ggF production mode cross section in the $H \rightarrow \tau\tau$ analysis.

Uncertainties of the used parton shower (PS) method (see Section 4.3) and the underlying event are estimated comparing two different models, namely Pythia8 [131] and Herwig7 [237]. The parton shower uncertainty are taken into account for the dominant production modes ggF , VBF and VH and are decorrelated between the boosted and VBF categories for ggF and VBF . In addition a parametrization uncertainty for the VBF parton shower model is considered

PDF uncertainties are evaluated following the recommendation described in [132], which includes 30 eigenvector variations and two variations of α_s . These uncertainties are treated as independent NPs in the likelihood model. The acceptance uncertainties of these theoretical uncertainties on the signal in the different event selections are summarized in Table A.1 in the appendix. Table 9.6 summarizes the final number of NPs included in the fit model to cover the theoretical uncertainties on the signal prediction, while Table A.2 in the appendix lists all NPs with their descriptions.

Uncertainty source	Number of NPs
Signal theory	49
PDF	30
α_s	1
QCD scale	12
Parton-shower	6

Table 9.6: Number of NPs related to signal theory uncertainties in the fit model grouped into different sources.

Theoretical uncertainties on the background prediction For $Z \rightarrow \tau\tau$ background events each considered source of systematic uncertainties is parametrized with the following strategy:

- A set of independent NPs per each inclusive signal category and per final state to account for the impact of the variation on the $m_{\tau\tau}^{\text{MMC}}$ distribution as well as for migration effects between different sub-categories of the same inclusive category. These are evaluated in each channel after the variation has been normalized to the nominal prediction in the inclusive signal categories.
- One NP per inclusive signal category to account for the different selection requirements of the inclusive categories across the final states. These are evaluated as the impact of the variation on the relative normalization of each final state with respect to the total event prediction.

Five different sources of theoretical uncertainties are considered for the $Z \rightarrow \tau\tau$ background:

- PDF: Based on internal weight variations of SHERPA [117]
- $\Delta\mu$: Factorization and renormalization scale variations
- CKKW: Jet-to-parton matching uncertainties, evaluated using particle-level parametrization as a function of the jet multiplicity and the transverse momentum of the Z -boson
- QSF: Resummation scale used for the soft gluon emission.
- PS and underlying-event: Comparison of MG5_aMc [136] and SHERPA [117] for $Z \rightarrow ll$ events

The PS uncertainties are evaluated on $Z \rightarrow ll$ events instead of $Z \rightarrow \tau\tau$ to increase the statistics. A comparison of these uncertainties on the $Z \rightarrow \tau\tau$ and $Z \rightarrow ll$ events shows a good agreement. For the slight trend observed in m_{jj} for $Z \rightarrow \tau\tau$ and $Z \rightarrow ll$ (see Section 9.3), one systematic uncertainty per process is considered by taking the difference when comparing the nominal prediction with a prediction reweighted to correct this trend.

For the $Z \rightarrow ll$ and top quark production backgrounds theoretical systematic uncertainties are studied and found to be significantly smaller than the statistical uncertainty of simulated events, so that they are neglected in the analysis.

Table 9.7 summarizes the final 42 NPs included in the fit model to cover the theoretical uncertainties on the background prediction, while Table A.2 in the appendix lists all NP with their descriptions.

Uncertainty source	Number of NPs
Background theory	42
Z PDF	8
Z $\Delta\mu_R, \mu_F$	8
Z CKKW	8
Z QSF	8
Z PS	8
m_{jj} modeling	2

Table 9.7: Number of NPs related to background theory uncertainties in the fit model grouped into different sources.

Experimental uncertainties Uncertainties on the reconstructed objects, the pile-up models, the luminosity and the misidentified τ background are considered as experimental uncertainties. For each object, the recommendation of the reconstruction method described in Chapter 5 is used.

For muons, electrons and jets resolutions, scale, trigger and efficiency uncertainties are considered. In addition, scale and efficiency uncertainties are evaluated for τ_h -leptons. Therefore, 13 independent NPs for electrons, 15 for muons and τ_h -leptons each are considered. The impact of the electron and muon uncertainties on the analysis is expected to be small, while the impact of the τ_h -leptons is sizable.

The systematic uncertainties on jets have an significant impact on this analysis due the requirement of jets in the signal categories and the correlation between the jet objects and the $m_{\tau\tau}^{\text{MMC}}$ distribution. Therefore, eleven NPs covering the jet energy resolution, 21 parameters covering the jet energy scale and one covering the efficiency of the jet vertex tagger (JVT) are evaluated. Similar to the jet uncertainties, a large impact of uncertainties related to the estimation of the missing transverse energy is expected. Resolution and scale uncertainties are considered, which impact the analysis via the missing transverse energy requirement in the event selections and through shape effects on the $m_{\tau\tau}^{\text{MMC}}$ distribution.

To cover uncertainties on the b-jet veto requirement in $\tau_1\tau_1$ and $\tau_1\tau_h$, 14 NPs on the flavor tagging efficiency are considered, which are expected to have a minor impact on the measurement. Two NPs are considered for the pile-up reweighting and one for the luminosity measurement (see Section 4.3). Finally, NPs related to the misidentified τ background estimation described in Section 9.4 are implemented. Table 9.8 summarizes the number of NPs included in the fit model that are related to experimental uncertainties, while Table A.3 in the appendix lists all NPs with their descriptions.

Leptons and jets				
Uncertainty source	Electrons	Muons	τ -leptons	Jets
All	13	15	15	33
Resolution	1	2		11
Energy scale	6	1	3	21
Reconstruction/Identification	3	6	4	
Isolation	1	2		
Trigger	2	4	6	
Charge identification	2			
Electron veto			2	
Jet vertex tagger				1
Misidentified- τ background				
Uncertainty source	$\tau_l \tau_l$	$\tau_l \tau_h$	$\tau_h \tau_h$	
All	2	7	3	
Others				
Uncertainty source	b-jet	E_T^{miss}	Luminosity	Pile-Up
All	13	3	1	1

Table 9.8: Number of NPs related to experimental uncertainty in the fit model grouped into different sources .

9.5 Statistical model

To extract the signal parameters a profiled likelihood fit is used as described in Section 7.2. All systematic uncertainties described in the previous Section 9.4 are included as NPs. For the background estimation, as described in Section 9.3, several NFs are included. Figure 9.14 summarizes the fit model. The figure shows the VBF categories and their control regions on the left side, while the right side illustrates the boosted categories and their control regions. The likelihood model consists of thirteen signal categories, four top quark production control regions and two $Z \rightarrow ll$ control regions. For every control region one NF is defined, which acts on the control region and on every corresponding signal category, so that four top quark production and two $Z \rightarrow ll$ NFs are included in the likelihood model. Two $Z \rightarrow \tau\tau$ NFs are included, where one is applied on the $Z \rightarrow \tau\tau$ background in all VBF categories and one in all boosted signal categories. Finally, one free NF is applied on the misidentified τ background in all $\tau_h \tau_h$ signal categories, so that nine free NFs on background contributions are included in the likelihood model.

On each signal sample per STXS Stage 1 bin, an individual normalization factor is defined as a PoI (see Section 7.1.3). However, due to the definition of signal categories and the limited data statistics, the full set of PoIs cannot be measured. Instead, three different measurements are performed using individual assumptions on the PoIs to reduce the number of free parameters in the fit model. For each measurement, a different set of PoIs on the signal sample are correlated. In addition PoIs of minor signal contribution are set to the SM prediction, so that

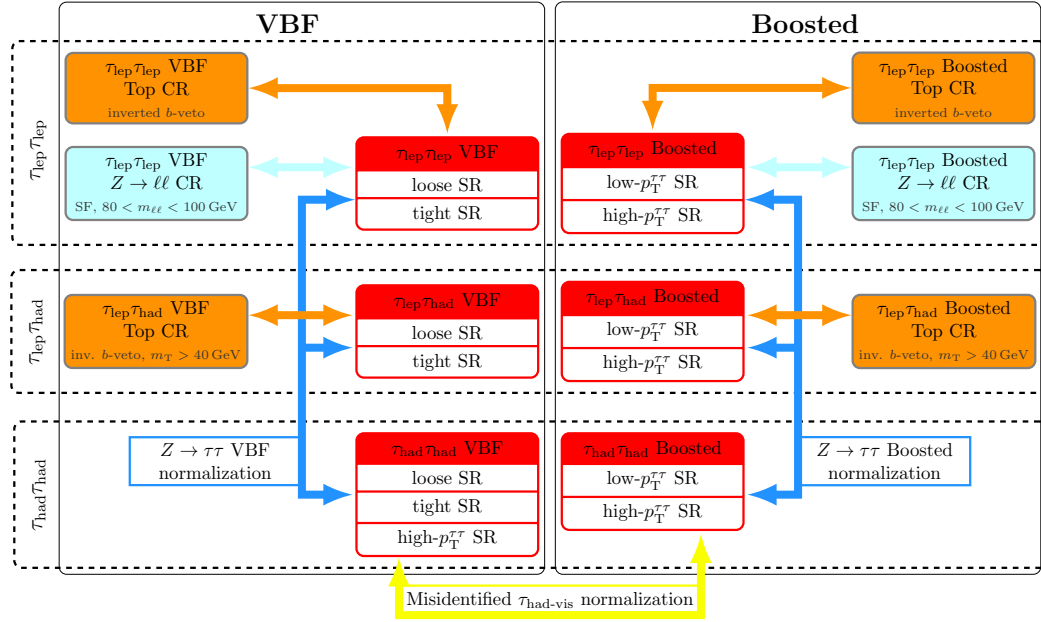


Figure 9.14: Schematic overview of the fit model. The red boxes correspond to signal categories, the orange boxes to top quark production control regions and the cyan boxes to $Z \rightarrow \ell\ell$ control regions. Each arrow corresponds to one NF, where blue acts only on $Z \rightarrow \tau\tau$, cyan on $Z \rightarrow \ell\ell$, orange on top production and yellow on misidentified τ background events. Arrows pointing to more than one signal category or control region illustrate that the NFs are correlated across these regions [26].

for each measurement a different set of final PoIs is defined.

To further increase the significance of the fit model an additional observable is used in the likelihood model. In each signal category the mass distribution of the di- τ -system calculated with the missing mass calculator $m_{\tau\tau}^{\text{MMC}}$ (see Section 5.7) is used as an input for the fit model, while for all control regions only the number of expected events are used. The binning in $m_{\tau\tau}^{\text{MMC}}$ varies across the signal categories depending on the given data statistics. Figure 9.15 and Figure 9.16 show the observed $m_{\tau\tau}^{\text{MMC}}$ distribution with the binning used in the fit model for all VBF and Boosted signal categories. The figures show that the $Z \rightarrow \tau\tau$ background peaks around the Z boson mass (≈ 90 GeV), while the $H \rightarrow \tau\tau$ signal peaks around Higgs boson mass (≈ 125 GeV). For all other background processes the distribution is rather flat. Therefore, the use of $m_{\tau\tau}^{\text{MMC}}$ as an observable in the likelihood model increases the signal to background separation. The observed event yields compared to the recorded data are shown in Table 9.9. In all signal categories the observed events agree within their uncertainties with the recorded data.

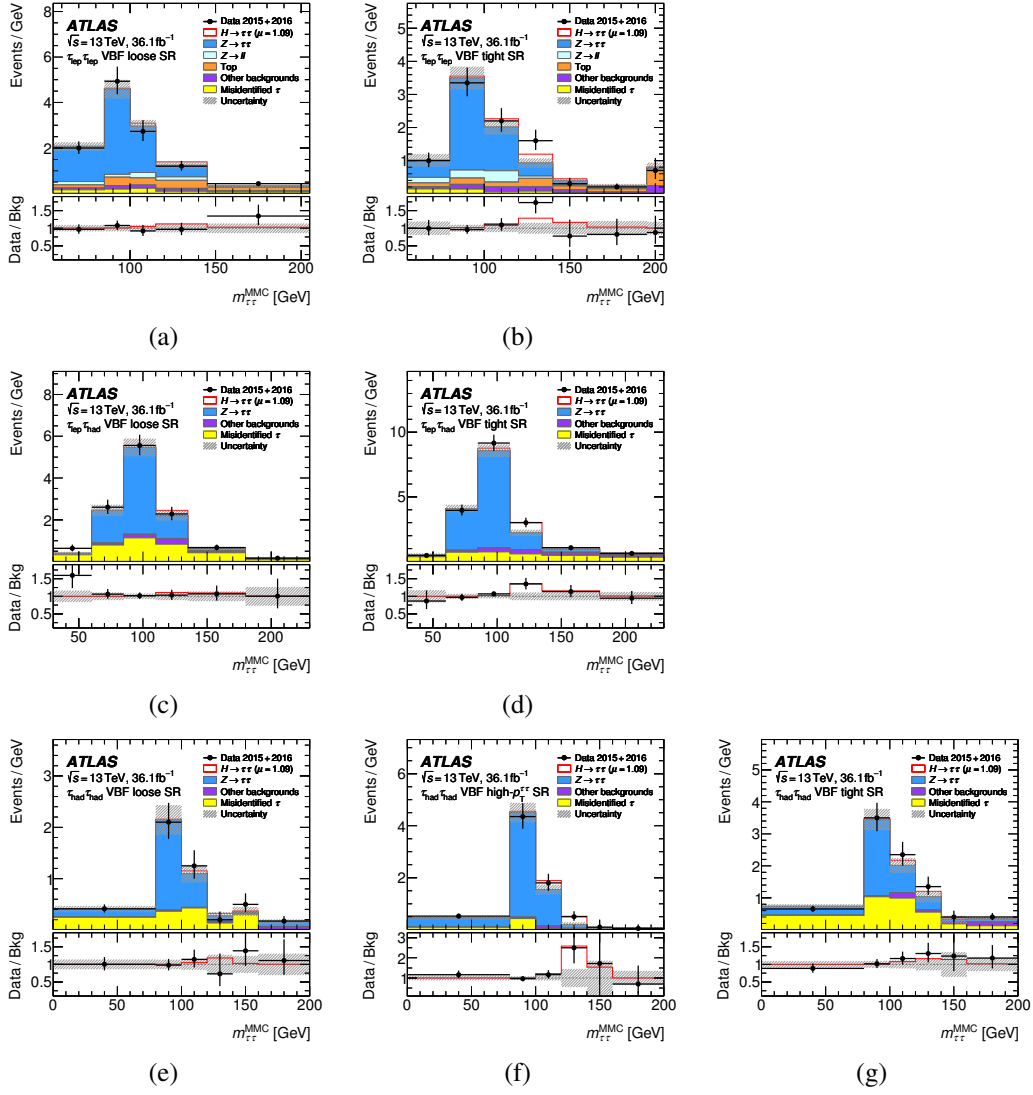


Figure 9.15: Observed $m_{\tau\tau}^{\text{MMC}}$ distribution of the $H \rightarrow \tau\tau$ analysis based on the 2015-2016 data set binned as in the fit model for all VBF signal categories, from top to bottom for the $\tau_1\tau_1$, $\tau_1\tau_h$ and $\tau_h\tau_h$ final states [26].

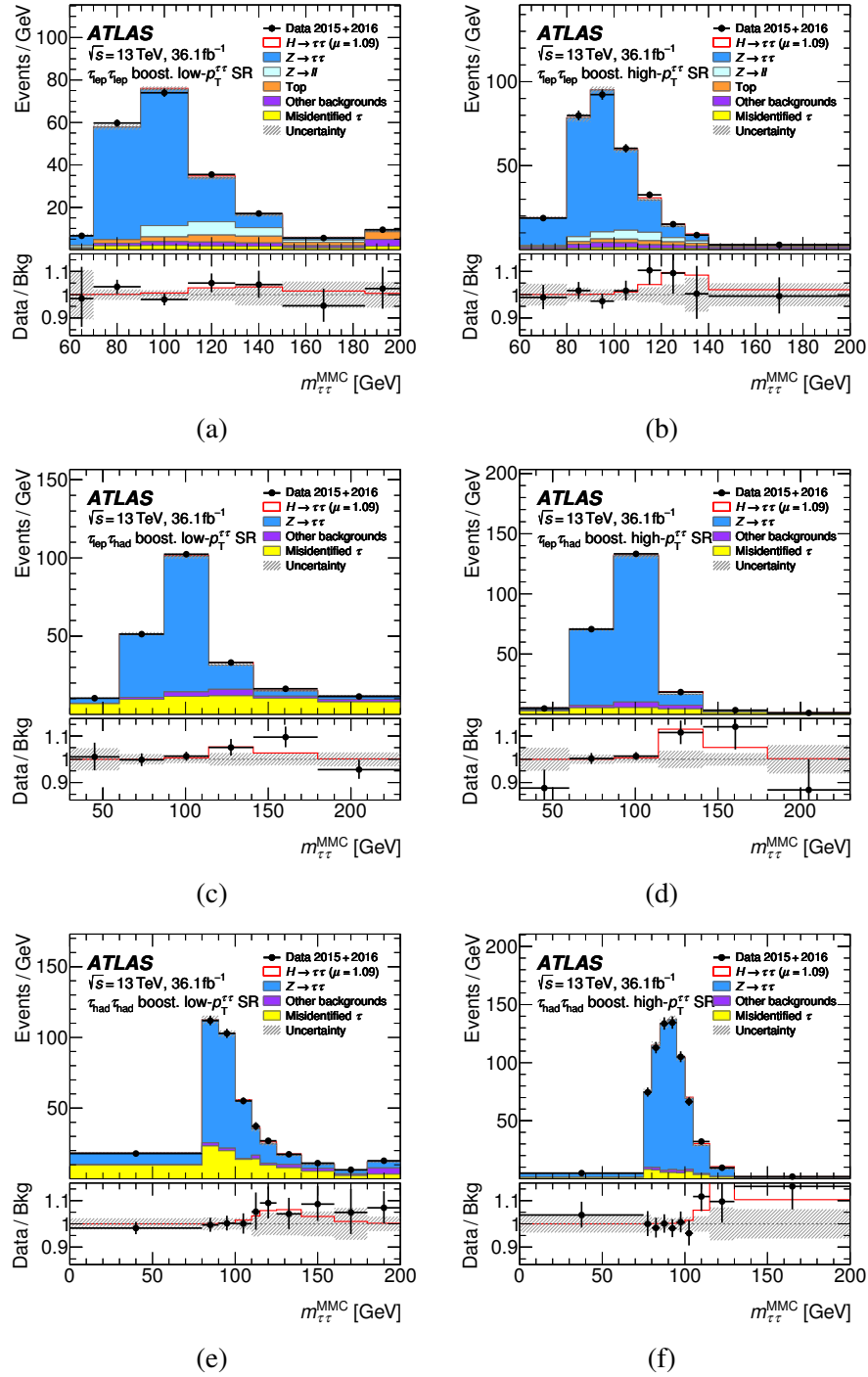


Figure 9.16: Observed $m_{\tau\tau}^{\text{MMC}}$ distribution of the $H \rightarrow \tau\tau$ analysis based on the 2015-2016 data set binned as in the fit model for all boosted signal categories, from top to bottom for the $\tau_l \tau_l$, $\tau_l \tau_h$ and $\tau_h \tau_h$ final states [26].

	$\tau_{\text{lep}}\tau_{\text{lep}}$ VBF		$\tau_{\text{lep}}\tau_{\text{lep}}$ boosted	
	Loose	Tight	Low- $p_{\text{T}}^{\tau\tau}$	High- $p_{\text{T}}^{\tau\tau}$
$Z \rightarrow \tau\tau$	151 $\pm$ 13	107 $\pm$ 12	2977 $\pm$ 90	2687 $\pm$ 64
$Z \rightarrow \ell\ell$	15.1 $\pm$ 4.9	20.3 $\pm$ 6.6	360 $\pm$ 54	236 $\pm$ 31
Top	33.0 $\pm$ 6.4	25.1 $\pm$ 4.5	321 $\pm$ 50	189 $\pm$ 29
VV	11.8 $\pm$ 2.2	10.7 $\pm$ 1.5	194.1 $\pm$ 8.5	195.3 $\pm$ 8.8
Misidentified τ	18.3 $\pm$ 9.6	9.6 $\pm$ 4.8	209 $\pm$ 92	80 $\pm$ 35
$gg\text{F}, H \rightarrow W W^*$	1.2 $\pm$ 0.2	1.4 $\pm$ 0.3	11.8 $\pm$ 2.6	16.4 $\pm$ 1.7
VBF, $H \rightarrow W W^*$	1.7 $\pm$ 0.2	4.1 $\pm$ 0.5	2.9 $\pm$ 0.3	2.9 $\pm$ 0.3
$gg\text{F}, H \rightarrow \tau\tau$	2.6 $\pm$ 0.9	1.8 $\pm$ 0.9	34.4 $\pm$ 9.2	33.8 $\pm$ 9.5
VBF, $H \rightarrow \tau\tau$	5.3 $\pm$ 1.5	11.3 $\pm$ 3.0	7.7 $\pm$ 2.1	8.2 $\pm$ 2.3
$WH, H \rightarrow \tau\tau$	< 0.1	< 0.1	2.5 $\pm$ 0.7	3.1 $\pm$ 0.9
$ZH, H \rightarrow \tau\tau$	< 0.1	< 0.1	1.3 $\pm$ 0.4	1.6 $\pm$ 0.4
$t\bar{t}H, H \rightarrow \tau\tau$	< 0.1	0.1 $\pm$ 0.1	1.5 $\pm$ 0.5	1.2 $\pm$ 0.4
Total background	232 $\pm$ 13	178 $\pm$ 12	4075 $\pm$ 61	3408 $\pm$ 54
Total signal	8.0 $\pm$ 2.2	13.2 $\pm$ 3.5	47 $\pm$ 12	48 $\pm$ 12
Data	237	188	4124	3444

	$\tau_{\text{lep}}\tau_{\text{had}}$ VBF		$\tau_{\text{lep}}\tau_{\text{had}}$ boosted	
	Loose	Tight	Low- $p_{\text{T}}^{\tau\tau}$	High- $p_{\text{T}}^{\tau\tau}$
$Z \rightarrow \tau\tau$	178 $\pm$ 18	323 $\pm$ 21	4187 $\pm$ 92	5347 $\pm$ 82
$Z \rightarrow \ell\ell$	10.0 $\pm$ 3.0	12.7 $\pm$ 3.1	130 $\pm$ 37	115 $\pm$ 16
Top	5.8 $\pm$ 1.6	17.9 $\pm$ 4.6	121 $\pm$ 20	57 $\pm$ 10
Misidentified τ	103 $\pm$ 16	101 $\pm$ 15	1895 $\pm$ 80	605 $\pm$ 29
Other backgrounds	4.0 $\pm$ 1.6	9.3 $\pm$ 1.9	115.0 $\pm$ 7.8	129.0 $\pm$ 8.8
$gg\text{F}, H \rightarrow \tau\tau$	3.8 $\pm$ 1.1	7.1 $\pm$ 1.9	62 $\pm$ 16	66 $\pm$ 22
VBF, $H \rightarrow \tau\tau$	7.6 $\pm$ 2.2	24.7 $\pm$ 6.8	11.9 $\pm$ 3.4	14.0 $\pm$ 4.0
$WH, H \rightarrow \tau\tau$	< 0.1	0.1 $\pm$ 0.0	3.9 $\pm$ 1.1	5.4 $\pm$ 1.4
$ZH, H \rightarrow \tau\tau$	< 0.1	< 0.1	1.8 $\pm$ 0.5	2.8 $\pm$ 0.7
$t\bar{t}H, H \rightarrow \tau\tau$	< 0.1	< 0.1	0.1 $\pm$ 0.0	0.2 $\pm$ 0.1
Total background	301 $\pm$ 17	463 $\pm$ 21	6448 $\pm$ 81	6253 $\pm$ 80
Total signal	11.5 $\pm$ 3.2	32.0 $\pm$ 8.2	80 $\pm$ 20	89 $\pm$ 26
Data	318	496	6556	6347

	$\tau_{\text{had}}\tau_{\text{had}}$ VBF			$\tau_{\text{had}}\tau_{\text{had}}$ boosted	
	Loose	Tight	High- $p_{\text{T}}^{\tau\tau}$	Low- $p_{\text{T}}^{\tau\tau}$	High- $p_{\text{T}}^{\tau\tau}$
$Z \rightarrow \tau\tau$	67.3 $\pm$ 9.2	100 $\pm$ 12	141 $\pm$ 12	3250 $\pm$ 130	3582 $\pm$ 82
Misidentified τ	45.0 $\pm$ 5.4	96.4 $\pm$ 9.2	20.0 $\pm$ 2.9	1870 $\pm$ 140	364 $\pm$ 53
Other backgrounds	4.4 $\pm$ 1.4	11.6 $\pm$ 1.7	4.4 $\pm$ 0.7	281 $\pm$ 21	109.9 $\pm$ 9.2
$gg\text{F}, H \rightarrow \tau\tau$	1.1 $\pm$ 0.4	2.0 $\pm$ 0.7	3.5 $\pm$ 1.0	41 $\pm$ 11	48 $\pm$ 14
VBF, $H \rightarrow \tau\tau$	1.4 $\pm$ 0.5	6.4 $\pm$ 1.8	11.2 $\pm$ 3.0	9.0 $\pm$ 3.4	10.7 $\pm$ 2.9
$WH, H \rightarrow \tau\tau$	< 0.1	< 0.1	< 0.1	3.3 $\pm$ 0.9	4.4 $\pm$ 1.2
$ZH, H \rightarrow \tau\tau$	< 0.1	< 0.1	< 0.1	2.4 $\pm$ 0.7	2.9 $\pm$ 0.8
$t\bar{t}H, H \rightarrow \tau\tau$	< 0.1	< 0.1	< 0.1	1.6 $\pm$ 0.5	1.9 $\pm$ 0.5
Total background	116.7 $\pm$ 9.4	208 $\pm$ 12	165 $\pm$ 12	5401 $\pm$ 78	4057 $\pm$ 64
Total signal	2.6 $\pm$ 0.8	8.6 $\pm$ 2.4	14.9 $\pm$ 3.8	57 $\pm$ 15	68 $\pm$ 18
Data	121	220	179	5455	4103

Table 9.9: Observed event yields of the $H \rightarrow \tau\tau$ analysis based on the 2015-2016 data set binned as in the fit model for all signal categories, from top to bottom for the $\tau_l\tau_l$, $\tau_l\tau_h$ and $\tau_h\tau_h$ final states [26].

9.6 Signal strength measurement

The signal strength measurement is designed to discover the $H \rightarrow \tau\tau$ coupling with the ATLAS detector. The signal strength μ is defined as the observed cross section of the $H \rightarrow \tau\tau$ decay mode $\sigma_{H \rightarrow \tau\tau}^{\text{obs}}$ divided by the corresponding SM prediction $\sigma_{H \rightarrow \tau\tau}^{\text{SM}}$: $\mu = \sigma_{H \rightarrow \tau\tau}^{\text{obs}} / \sigma_{H \rightarrow \tau\tau}^{\text{SM}}$ and is used to estimate the significance of the signal excess relative to the background-only hypothesis as described in Section 3.1. For this measurement, one common PoI for all $H \rightarrow \tau\tau$ signal samples is considered in the likelihood model. This setup uses the assumption that the relative contribution of the different signal samples to the combined signal is SM-like. The observed significance corresponds to 4.4σ , while the expected significance is 4.1σ . This corresponds to an evidence ($>3\sigma$ but $<5\sigma$) of the $H \rightarrow \tau\tau$ decay mode, while for an observation $>5\sigma$ is required. To reach this threshold with the data recorded by the ATLAS detector a combined fit using this analysis and the previous analysis based on the Run 1 data set [233] is performed. All NPs and NFs on background processes are treated as uncorrelated between the two analyses, while the NFs on the signal processes are fully correlated to one PoI. The observed (expected) significance for the combined measurement amounts to 6.4σ (5.4σ) and corresponds to the first observation of the $H \rightarrow \tau\tau$ decay mode with the ATLAS experiment.

9.7 Inclusive cross section measurements

The focus of the measurements discussed in this section is the determination of cross sections of different Higgs boson production modes in the $H \rightarrow \tau\tau$ decay channel. Two different measurements with individual sets of PoIs are performed. The first setup targets the measurement of the inclusive cross section of the Higgs boson production and assumes that the relative contributions of all signal processes to the inclusive Higgs boson production are SM-like. Therefore, one common PoI for all $H \rightarrow \tau\tau$ production modes is defined by correlating the individual PoIs of all signal samples in the likelihood model. The final PoI is defined as the inclusive Higgs boson production cross section σ_H multiplied by the $H \rightarrow \tau\tau$ branching ratio $\mathcal{B}_{H \rightarrow \tau\tau}$: $\sigma_{H \rightarrow \tau\tau} = \sigma_H \times \mathcal{B}_{H \rightarrow \tau\tau}$, where the ggF , VBF , VH and $t\bar{t}H$ production modes are considered for σ_H (see Section 6.1). In the second setup the measurement of individual cross sections for the main production modes of the Higgs boson are targeted. Therefore, one PoI per production mode is defined, which are uncorrelated across different production modes. As described in Section 9.2.3 the signal categories are focusing on the ggF and VBF production modes. For a stable likelihood model all PoIs of the remaining Higgs boson production modes are set constant to the SM prediction in this setup, so that two final PoIs are defined in this setup: $\sigma_{H \rightarrow \tau\tau}^{ggF}$ and $\sigma_{H \rightarrow \tau\tau}^{VBF}$.

9.7.1 Higgs boson production cross section

The observed value of the inclusive Higgs boson production cross section of the $H \rightarrow \tau\tau$ decay mode is:

$$\sigma_{H \rightarrow \tau\tau}^{\text{obs}} = 3.77^{+0.60}_{-0.59}(\text{stat.})^{+0.87}_{-0.74}(\text{syst.}) \text{ pb}, \quad (9.8)$$

which is consistent with SM prediction of

$$\sigma_{H \rightarrow \tau\tau}^{\text{SM}} = 3.46 \pm 0.13 \text{ pb.} \quad (9.9)$$

The measurement is dominated by the systematic uncertainties. Table 9.10 shows the relative

Uncertainty source	Impact $\Delta\sigma/\sigma_{H \rightarrow \tau\tau}$ [%]			
	Observed		Expected	
Theoretical unc. on signal	+13.4	-8.7	+12.0	-7.8
Background sample size	+10.8	-9.9	+10.1	-9.7
Jets and E_T^{miss}	+11.2	-9.1	+10.4	-8.4
Background normalization	+6.3	-4.4	+6.3	-4.4
Misidentified τ leptons	+4.5	-4.2	+3.4	-3.2
Theoretical uncert. on background	+4.6	-3.6	+5.0	-4.0
τ_h -leptons	+4.4	-2.9	+5.5	-4.0
B-jets	+3.4	-3.4	+3.0	-2.3
Luminosity	+3.3	-2.4	+3.1	-2.2
Electrons and muons	+1.2	-0.9	+1.1	-0.8

Table 9.10: Relative impact of systematic uncertainties on the inclusive cross section grouped into different categories. The groups are listed in decreasing order of their impact [26].

uncertainties of systematic uncertainties grouped into different sources estimated with the breakdown approach (see Section 7.2). The dominant systematic uncertainties are theoretical uncertainties on the signal prediction with +13.4 (-8.7)%, followed by the statistical uncertainties of the background simulation and systematic uncertainties related to jets and missing transverse energy. The main contribution to the theoretical uncertainties on the signal are uncertainties on the QCD calculation of ggF and on the parton shower modeling for VBF . The high impact of jets and missing transverse energy uncertainties is a result of the limited detector energy resolution, which are used in this analysis for the signal to background separation and impact the shape of $m_{\tau\tau}^{\text{MMC}}$. Figure 9.17 shows the ten NPs with the largest impact on the measurement estimated with the NP ranking approach (see Section 7.2). Four uncertainties correspond to QCD calculation uncertainties on ggF , where the QCD calc. of ggF , $p_T^H \geq 120$ GeV has the largest impact. Two uncertainties related to the jet energy resolution of the detector are ranked on the second and seventh position. Furthermore, uncertainties on the background normalization of $Z \rightarrow \tau\tau$ and $Z \rightarrow ll$ events, on the b-jet identification and on the luminosity calculation are ranked under the top ten systematic uncertainties. The pulls and constraints of these uncertainties are shown as well in Figure 9.17. Most of the leading uncertainties have minor pulls and small constraints except for the jet and b-jet related uncertainties. The pulls on constraints on these systematic uncertainties show that this analysis gives the likelihood model the possibility to reduce and correct the input of these uncertainties, which are estimated in auxiliary measurements. The sensitivity of the $H \rightarrow \tau\tau$ analysis on these NPs can be explained by relatively high data statistics of the analysis in events with

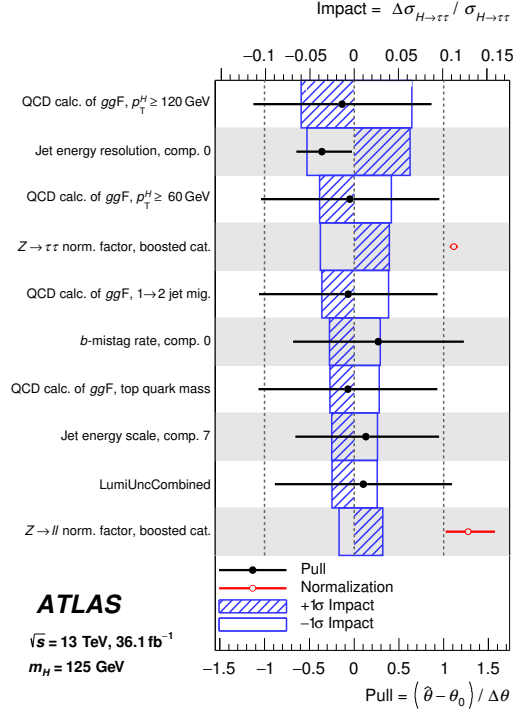


Figure 9.17: Relative impact of the ten systematic uncertainties with the highest impact on the inclusive measurement. The uncertainties are listed in decreasing order of their impact. The upper x-axis corresponds to the impact of the systematic uncertainty on the the inclusive cross section if the systematic is varied by one standard deviation and is visualized with blue boxes. The lower x-axis corresponds to the pull and constraint of the systematic uncertainty, which is illustrated with black points and black lines [26].

at least one jet and the dependency of the $m_{\tau\tau}^{\text{MMC}}$ distribution on these. Due to the $m_{\tau\tau}^{\text{MMC}}$ dependency, the systematic uncertainties have a large shape contribution in the fit model, so that the fit can pull and constrain these in event categories with large data statistics compared to the auxiliary measurement, such as the boosted categories. This is given especially for the jet related uncertainties, since the calculation of these uncertainties is still based the Run 1 data set, with an integrated luminosity of 26 fb $^{-1}$. An evaluation of these uncertainties based on the Run 2 data set was not available at this time.

9.7.2 ggF and VBF Higgs boson production cross section

In a second configuration of the analysis the inclusive production mode cross section for the two leading production modes ggF and VBF are determined. Therefore, one PoI for the ggF production mode and one for the VBF production mode are defined in the likelihood model. The normalization of the remaining production modes are set constant to the SM prediction.

The observed values for the ggF production mode $\sigma_{H \rightarrow \tau\tau}^{ggF} = \sigma_H^{ggF} \times \mathcal{B}_{H \rightarrow \tau\tau}$ and the VBF production mode $\sigma_{H \rightarrow \tau\tau}^{VBF} = \sigma_H^{VBF} \times \mathcal{B}_{H \rightarrow \tau\tau}$ are:

$$\begin{aligned}\sigma_{H \rightarrow \tau\tau}^{ggF, \text{obs}} &= 3.10 \pm 1.0 \quad (\text{stat.}) \quad {}^{+1.6}_{-1.3} \quad (\text{syst.}) \quad \text{pb}, \\ \sigma_{H \rightarrow \tau\tau}^{VBF, \text{obs}} &= 0.28 \pm 0.09 \quad (\text{stat.}) \quad {}^{+0.11}_{-0.09} \quad (\text{syst.}) \quad \text{pb}.\end{aligned}\quad (9.10)$$

The SM prediction are:

$$\begin{aligned}\sigma_{H \rightarrow \tau\tau}^{ggF, \text{SM}} &= 3.05 \pm 0.13 \quad \text{pb}, \\ \sigma_{H \rightarrow \tau\tau}^{VBF, \text{SM}} &= 0.237 \pm 0.006 \quad \text{pb}.\end{aligned}\quad (9.11)$$

The observed cross section values agree within their uncertainties with their SM predictions. The measurement of $\sigma_{H \rightarrow \tau\tau}^{ggF}$ is dominated by systematic uncertainties, while the impact of the systematic uncertainties and data statistics are of the same size for $\sigma_{H \rightarrow \tau\tau}^{VBF}$. The impact of individual sources of systematic uncertainties is similar to the one for the measurement of the inclusive production modes cross section described in the previous paragraph. The leading experimental systematic uncertainties are related to jets and missing transverse momentum. The leading theoretical uncertainties are QCD calculation uncertainties for the ggF production mode and parton shower uncertainties for the VBF production mode.

Figure 9.18 shows the best-fit value as well as the 68% and the 95% confidence limit (CL)

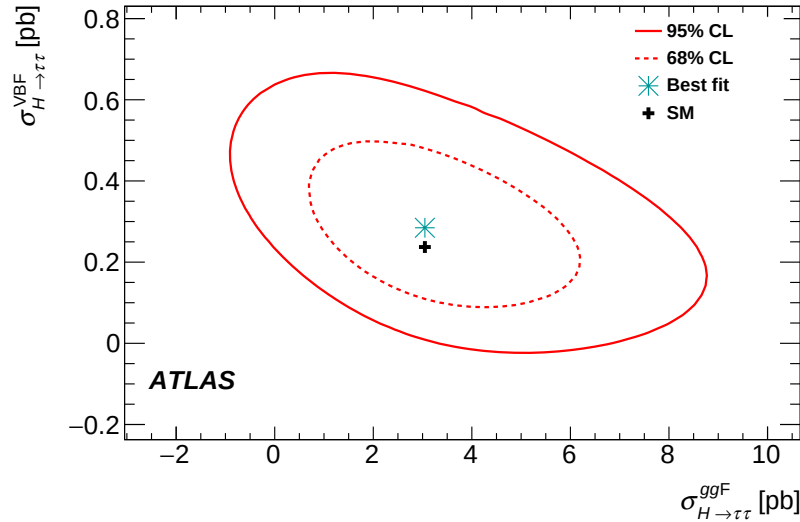


Figure 9.18: Best-fit values and the 68% as well as the 95% confidence limits (CL) for the simultaneous measurement of the inclusive ggF and VBF Higgs production mode in the $H \rightarrow \tau\tau$ decay channel. The x-axis corresponds to the ggF and the y-axis to the VBF cross section. The black cross shows the best-fit value, while the blue cross marks the SM prediction. The dashed red line shows the 68% CL and the solid red line the 95% CL area of the measurement [26].

contours of this measurement in the $(\sigma_{H \rightarrow \tau\tau}^{ggF}, \sigma_{H \rightarrow \tau\tau}^{VBF})$ -plane. The SM prediction is covered by the 68% CL around the best-fit value. The 68% and 95% CL ellipse are both rotated anti-clock wise, which illustrates the high anti-correlation of the two measured cross sections with a value of -52%.

9.8 STXS measurement

The analysis strategy described in the previous sections is used to preform the first STXS measurement in the $H \rightarrow \tau\tau$ decay mode. For this, the Stage 1 definition of the STXS framework (see Section 7.1.3) is considered. The analysis event selection is optimized to measure the inclusive ggF and VBF production modes and not to measure cross sections within the STXS framework. This measurement is designed to define and measure the best configuration of STXS bins for the given event selections. Figure 9.19 shows the

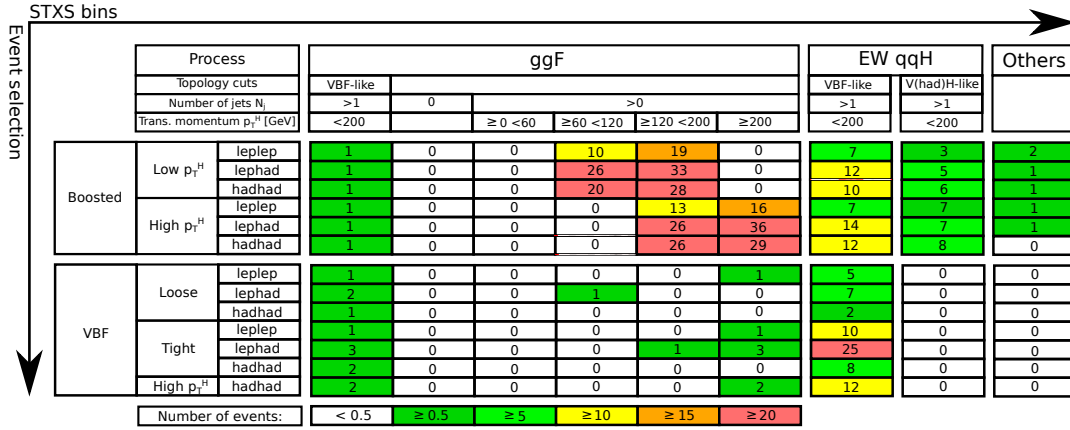


Figure 9.19: Signal event yields of STXS bins (columns) in the analysis signal categories (rows). The colored background indicates the size of event yield, where the color changes from green (0.5-10 events) over yellow (10-15 events) and orange (15-20 events) to red (over 20 events) for higher yields.

expected event yields of the different Stage 1 STXS bins in the signal categories. The boosted signal categories are dominated by the ggF production mode. The event category split at $p_T^H = 140$ GeV is not in agreement with the STXS p_T^H bin borders. The boosted low categories are mainly populated with events of the ggF $60 \leq p_T^H < 120$ GeV and ggF $120 \leq p_T^H < 200$ GeV STXS bins and the boosted high categories with events of the ggF $120 \leq p_T^H < 200$ GeV and ggF $200 \text{ GeV} \leq p_T^H$ STXS bins. All boosted categories have additional signal contributions from the VBF and $V(\rightarrow qq)H$ production modes. For a stable fit result, each considered STXS bin should have a signal category in which the STXS bin is dominating. To achieve this, the ggF $60 \leq p_T^H < 120$ GeV and ggF $120 \leq p_T^H < 200$ GeV bins are merged together to one STXS bin, while the ggF $200 \text{ GeV} \leq p_T^H$ STXS bin is kept separated. The VBF categories are all dominated by the VBF like STXS bin. Figure 9.20 shows the relative contributions of the STXS bins in the different categories. The categories are

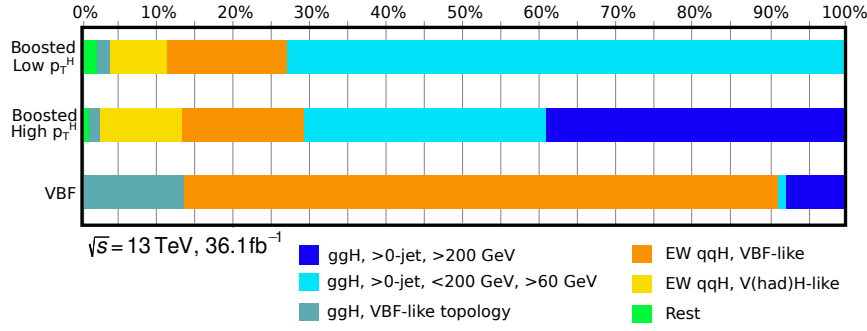


Figure 9.20: Relative contribution of STXS bins (colors) to the full expected $H \rightarrow \tau\tau$ signal in each signal category (rows). The categories are summed up across the final state and all VBF categories are summed up.

summed up across the final states and for all VBF categories, since they show no significant differences across each other. The boosted low categories and the VBF categories are dominated by their corresponding STXS bins $ggF\ 60 \leq p_T^H < 200\text{ GeV}$ and $EW\ qqH\ VBF\text{-like}$, while the boosted high category has a similar contribution of the $ggF\ 60 \leq p_T^H < 200\text{ GeV}$ and $ggF\ 200\text{ GeV} \leq p_T^H$ STXS bins. The normalization of these three STXS bins: $ggF\ 60 \leq p_T^H < 200\text{ GeV}$ and $ggF\ 200\text{ GeV} \leq p_T^H$ and the $EW\ qqH\ VBF\text{-like}$ are kept as PoIs in the likelihood model. The contribution of other STXS bins is relatively small, so that the NFs of them are set constant to the SM prediction in the likelihood model.

Figure 9.21 compares the observed cross section for the three STXS bins with their SM

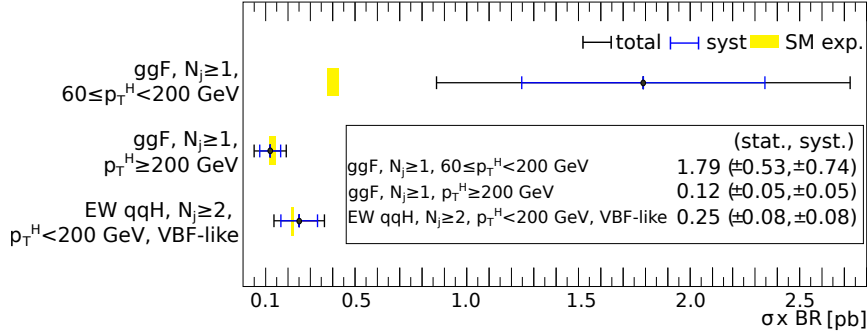


Figure 9.21: Observed cross sections for STXS bins in the $H \rightarrow \tau\tau$ decay channel, compared to the SM prediction illustrated with yellow boxes. The full uncertainties on the measured PoIs are shown with black bars, while the systematic uncertainties on them is shown with blue bars.

predictions and Table 9.11 summarizes the observed cross section values as well as the SM

9.8 STXS measurement

	Process	STXS bin			$\sigma_{H \rightarrow \tau\tau}^{obs}$ [pb]			$\sigma_{H \rightarrow \tau\tau}^{SM}$ [pb]
(1)	$EW\ qqH$	$N_j \geq 2$	$p_T^H < 200\text{ GeV}$	$VBF\text{-like}$	0.25	± 0.08 (stat.)	± 0.08 (syst.)	0.22 ± 0.01
(2)	ggF	$N_j \geq 1$	$200\text{ GeV} \leq p_T^H$		0.12	± 0.05 (stat.)	± 0.05 (syst.)	0.14 ± 0.03
(3)	ggF	$N_j \geq 1$	$60 \leq p_T^H < 200\text{ GeV}$		1.79	± 0.53 (stat.)	± 0.74 (syst.)	0.40 ± 0.05

Table 9.11: Observed ($\sigma_{H \rightarrow \tau\tau}^{obs}$) and SM expected ($\sigma_{H \rightarrow \tau\tau}^{SM}$) cross sections for measured STXS bins in the $H \rightarrow \tau\tau$ decay channel [26].

predictions. For the $EW\ qqH$ VBF -like and ggF $200\text{ GeV} \leq p_T^H$ STXS bins, the observed values agree within one standard deviation with the SM prediction, while the ggF $60 \leq p_T^H < 200\text{ GeV}$ STXS bin agrees within 1.5 standard deviations. The measurement of the ggF $60 \leq p_T^H < 200\text{ GeV}$ STXS bin is dominated by systematic uncertainties due to the large extrapolation from the event selection to the measured phase space defined by the STXS bin. For the remaining measured STXS bins the statistical uncertainty is of the same size as the systematic uncertainty. However, also for these STXS bins the event selection does not agree well with the STXS bin definitions, so that the sensitivity is reduced by large extrapolation uncertainties. To improve the measurement, the event selection has to be defined closer to the STXS bin definitions to reduce the impact of systematic uncertainties. For this, a new analysis is performed using the full Run 2 data set and an event selection optimized for the STXS measurement, which is described in Chapter 10.

10 $H \rightarrow \tau\tau$ ANALYSIS WITH THE 2015-2018 DATA SET

After the observation of the $H \rightarrow \tau\tau$ coupling with the ATLAS experiment analysis described in Chapter 9, a new analysis focusing on the measurement of cross sections, especially within the STXS framework (see Section 7.1), is performed. This analysis is based on the full Run 2 data set taken from 2015 to 2018. The previous analysis [233] is used as a baseline but the event selection, background estimation and the fit model are optimized. This section will describe the changes compared to the previous analysis and discuss the results of the optimized analysis. The results are published in Reference [28].

10.1 Signal and background processes

Additional production processes of the Higgs boson are included compared to the previous analysis (see Section 4.5). The tH and $b\bar{b}H$ production modes are added, which have a small contribution compared to the other production modes, but become more relevant compared to the previous analysis due to the increased precision of the measurement. The same background processes as for the 2015 to 2016 analysis are considered.

10.2 Event selection

This section summarizes the signal event selection of the analysis. The selection is split into three different stages: Trigger selection, preselection and event categories. The first two selections are defined to increase the $H \rightarrow \tau\tau$ signal over background ratio, while the event categories are designed to further increase this ratio but also to increase the signal contribution of a specific STXS bin (see Section 7.1) compared to other STXS bins.

10.2.1 Trigger selection

The trigger selection for the 2015 and 2016 data set, described in Table 9.1, is used again for this analysis. For the 2017 and 2018 data sets the same single-lepton and di-lepton triggers are used as for the 2016 data set. Only for the $\tau_h\tau_h$ final state, two new triggers are used for the 2017 and 2018 data set (see Table 10.1). The same trigger and offline analysis requirements on p_T , as summarized in Table 9.2, are used for this analysis. For the two additional di- τ_h triggers, the same thresholds as for the 2016 di- τ_h trigger are applied.

Trigger menu	Data period	Chain name
Di-hadronic τ	2017	HLT_tau35_medium1_tracktwo_tau25_medium1_tracktwo_03dR30_L1DR_TAU20ITAU12U_J25
	2018	HLT_tau35_medium1_tracktwoEF_tau25_medium1_tracktwoEF_03dR30_L1DR_TAU20ITAU12U_J25

Table 10.1: Overview of additional unprescaled di- τ_h triggers of this analysis compared to the 2015 to 2016 analysis, grouped by the corresponding data period.

10.2.2 Preselection

Table 10.2 summarizes the preselection of the analysis. Compared to the previous analysis the $\tau_l \tau_l$ selection has changed as well as the criteria used for the τ_l and τ_h identification and isolation. In $\tau_l \tau_l$ the same flavor final state is excluded from the analysis. In this final state the $Z \rightarrow ll$ background has a significant contribution and a large impact on the signal extraction, while it has only a small contribution in other final states. In the previous measurement, a dedicated control regions combined with NFs were used to model this background. The $\tau_l \tau_l$ final state has a small branching ratio of 6.2% and a minor contribution to the cross section measurements in the $H \rightarrow \tau\tau$ decay mode. By removing this final state, the control regions for the $Z \rightarrow ll$ background can be removed. This gives the possibility to use these $Z \rightarrow ll$ control regions to constrain the $Z \rightarrow \tau\tau$ background (see Section 10.3). The new working points for τ_l and τ_h , based on new methods described in Section 5.2, 5.3 and 5.5, are considered. The remaining requirements of the preselection follow the selection of the previous analysis described in Section 9.2.2.

	$\tau_e \tau_\mu$	$\tau_l \tau_h$	$\tau_h \tau_h$
(1)	$N_{\tau_e/\tau_\mu}^{\text{loose}}=2, N_{\tau_h}^{\text{loose}}=0$	$N_{\tau_e/\tau_\mu}^{\text{loose}}=1, N_{\tau_h}^{\text{loose}}=1$	$N_{\tau_e/\tau_\mu}^{\text{loose}}=0, N_{\tau_h}^{\text{loose}}=2$
(2)	τ_e/τ_μ : Medium ID τ_e/τ_μ : FCLoose/FCTightTrackOnly	τ_e/τ_μ : Medium ID τ_e/τ_μ : FCLoose/FCTightTrackOnly	
(3)		τ_h : RNN Medium ID	τ_h : RNN Medium ID
(4)	Opposite charge	Opposite charge	Opposite charge
(5)	$m_{\tau\tau}^{\text{coll}} > m_Z - 25 \text{ GeV}$	$m_T < 70 \text{ GeV}$	
(6)	$30 < m_{ll} < 100 \text{ GeV}$		
(7)	$E_T^{\text{miss}} > 20 \text{ GeV}$	$E_T^{\text{miss}} > 20 \text{ GeV}$	$E_T^{\text{miss}} > 20 \text{ GeV}$
(8)	$\Delta R_{\tau\tau} < 2.0$	$\Delta R_{\tau\tau} < 2.5$	$0.6 < \Delta R_{\tau\tau} < 2.5$
(9)	$ \Delta\eta_{\tau\tau} < 1.5$	$ \Delta\eta_{\tau\tau} < 1.5$	$ \Delta\eta_{\tau\tau} < 1.5$
(10)	$0.1 < x_1 < 1.0$	$0.1 < x_1 < 1.4$	$0.1 < x_1 < 1.4$
(11)	$0.1 < x_2 < 1.0$	$0.1 < x_2 < 1.2$	$0.1 < x_2 < 1.4$
(12)	$p_T^{\text{ll}} > 40 \text{ GeV}$	$p_T^{\text{ll}} > 40 \text{ GeV}$	$p_T^{\text{ll}} > 70 \text{ GeV}, \eta_{\text{ll}} > 3.2$
(13)	$N_{\text{b-jet}} = 0$	$N_{\text{b-jet}} = 0$	$N_{\text{b-jet}} = 0$ except for ttH category

Table 10.2: Preselection requirements for the three different final states $\tau_e \tau_\mu$, $\tau_l \tau_h$ and $\tau_h \tau_h$.

10.2.3 Inclusive signal categories

The signal categories are optimized to measure cross sections within the STXS framework. Therefore, new event selections covering additional or finer regions of phase space are defined compared to the previous analysis described in Section 9.2.3. The analysis aims to measure the four dominant production modes: ggF , VBF , $V(\rightarrow qq)H$ and $t\bar{t}H$. One inclusive signal category is defined for each of these production modes.

A schematic overview of the inclusive signal categories is shown in Figure 10.1. For the ggF , VBF and $V(\rightarrow qq)H$ production modes, the inclusive categories are defined in every final state of the di- τ -system, while the $t\bar{t}H$ production category is only defined in the $\tau_h \tau_h$ decay channel to avoid any overlap with the multi-lepton $t\bar{t}H$ analysis [238], which is used in

Chapter 11.

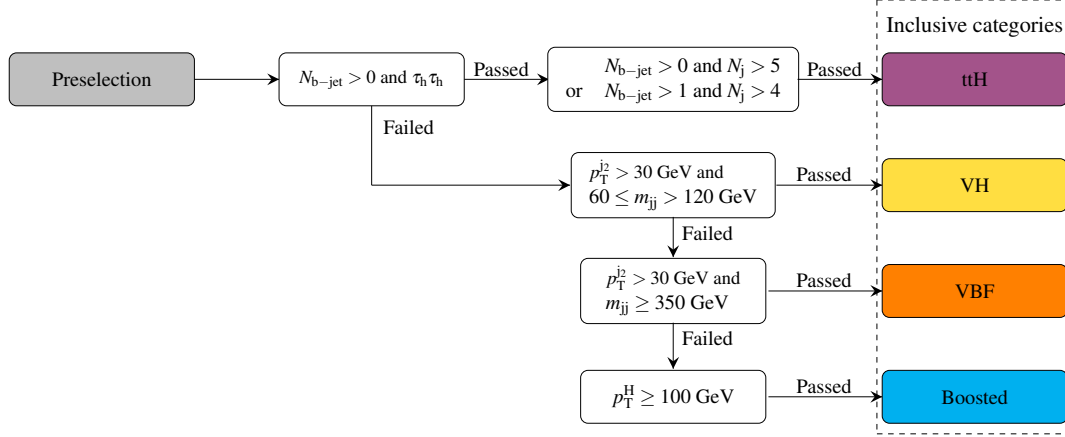


Figure 10.1: Schematic overview of the inclusive signal category requirements. The color illustrates for which production mode the category is designed: Purple for the $t\bar{t}H$, yellow for the $V(\rightarrow qq)H$, orange for the VBF and blue for the ggF production mode. White boxes illustrate requirements, which have to be passed or failed by an event to end up in a specific category.

$t\bar{t}H$ signal category A dedicated analysis of the ATLAS experiment measures the $t\bar{t}H$ production cross section in the di- τ decay channel with at least one leptonically decaying τ -lepton. However, the full hadronic decay mode of the di- τ system is not covered by this analysis. Therefore, the new event selection is introduced for the full hadronic final state only. As described in Section 6.1, the $t\bar{t}H$ production modes have two associated top quarks in addition to the Higgs boson, which decay most likely into jets. Especially, the decay into b-jets has a large branching ratio. Therefore, an inclusive $t\bar{t}H$ signal category can be designed by requiring a large number of jets N_j and b-jets $N_{b\text{-jet}}$. Two different selections are combined to the inclusive $t\bar{t}H$ category by requiring either $N_j > 5$, where at least one jet is b-jet tagged ($N_{b\text{-jet}} > 0$) or $N_j > 4$ with $N_{b\text{-jet}} > 1$.

VH signal category The VH category is designed to measure the EW qqH $60 \leq m_{jj} < 120$ GeV STXS bin and should enhance the contribution of $V(\rightarrow qq)H$ events, in which the V boson decays hadronically into two jets. To select this production mode two additional jets assigned to the V boson decay are expected in the final state. For the signal category a second jet with a transverse momentum of at least 30 GeV ($p_T^{jj} > 30$ GeV) and an invariant mass of the di-jet system m_{jj} around the W and Z boson mass, with $60 \leq m_{jj} < 120$ GeV, is required.

VBF signal category The inclusive VBF category is optimized for this analysis with respect to the previous analysis by harmonizing the selection towards the VBF-like STXS bin

definition of the $EW\ qqH$ production mode. A second jet is required with $p_T^{j_2} \geq 30$ GeV and $m_{jj} \geq 350$ GeV has to be fulfilled. The absolute difference in η of the two jets $|\eta_{jj}|$ must be larger than 3 and the two jets are required to be in opposite hemispheres ($\eta_{j_1}\eta_{j_2} > 0$). These requirements reduce further background contributions and in particular the contribution from ggF production in the VBF category.

Boosted signal category The boosted inclusive signal category is defined to measure ggF production. It focuses on the boosted ggF production with at least one additional jet and high values of p_T^H . To avoid any overlap with other signal categories, an explicit veto against them is applied, so that any event that passing one of the other selection requirements will be excluded for the boosted category. A threshold of $p_T^H \geq 100$ GeV is set to reduce background contributions, especially from QCD multi-jet events.

Process contributions The expected event yields in each inclusive signal category for all three final states are shown in Table 10.3. $\tau_h\tau_h$ and $\tau_l\tau_h$ have a similar size of total expected

Process	Background	ggF	VBF	VH	$t\bar{t}H$	Other $H \rightarrow \tau\tau$	S/B
$\tau_h\tau_h$							
ttH incl.	749 $\pm$ 33	2.38 $\pm$ 0.41	0.23 $\pm$ 0.03	1.10 $\pm$ 0.16	12.30 $\pm$ 2.02	1.69 $\pm$ 0.14	0.024
VH incl.	7297 $\pm$ 138	78.69 $\pm$ 13.71	8.69 $\pm$ 1.32	42.90 $\pm$ 3.91	0.74 $\pm$ 0.08	0.48 $\pm$ 0.05	0.018
VBF incl.	3449 $\pm$ 85	41.75 $\pm$ 9.67	112.80 $\pm$ 11.26	1.78 $\pm$ 0.18	0.19 $\pm$ 0.03	0.51 $\pm$ 0.06	0.046
Boosted incl.	49568 $\pm$ 401	425.60 $\pm$ 36.12	100.40 $\pm$ 7.72	37.21 $\pm$ 2.49	2.41 $\pm$ 0.19	1.94 $\pm$ 0.12	0.011
$\tau_l\tau_h$							
VH incl.	9516 $\pm$ 145	92.38 $\pm$ 20.45	9.73 $\pm$ 1.61	43.78 $\pm$ 3.92	0.48 $\pm$ 0.05	0.36 $\pm$ 0.03	0.015
VBF incl.	3636 $\pm$ 77	74.44 $\pm$ 27.95	136.43 $\pm$ 12.97	1.70 $\pm$ 0.17	0.10 $\pm$ 0.01	0.37 $\pm$ 0.04	0.059
Boosted incl.	47297 $\pm$ 262	438.04 $\pm$ 47.91	100.08 $\pm$ 6.93	35.98 $\pm$ 2.90	1.25 $\pm$ 0.10	1.41 $\pm$ 0.08	0.012
$\tau_l\tau_l$							
VH incl.	3567 $\pm$ 65	27.88 $\pm$ 5.05	2.54 $\pm$ 0.43	11.78 $\pm$ 1.20	0.10 $\pm$ 0.01	0.11 $\pm$ 0.02	0.012
VBF incl.	1295 $\pm$ 36	13.72 $\pm$ 3.25	36.53 $\pm$ 3.50	0.35 $\pm$ 0.04	0.01 $\pm$ 0.01	0.18 $\pm$ 0.02	0.039
Boosted incl.	17176 $\pm$ 139	133.56 $\pm$ 11.54	30.48 $\pm$ 2.27	10.93 $\pm$ 0.95	0.38 $\pm$ 0.03	0.70 $\pm$ 0.04	0.010

Table 10.3: Predicted event yields for background events and the ggF , VBF , VH , $t\bar{t}H$ and other production modes of the Higgs boson in the inclusive signal categories. In addition, the ratio between the inclusive $H \rightarrow \tau\tau$ signal and the total background (S/B) is given for each category.

events, while the expected events in $\tau_l\tau_l$ are only one-third. The ttH inclusive category, which is only defined in $\tau_h\tau_h$, has the smallest event yields with 18 $H \rightarrow \tau\tau$ events and 749 background events. The remaining categories are defined in every final state. In all of them, the largest expected yields are given in the boosted categories with 49568, 47297 and 17176 background events and 568, 577 and 176 $H \rightarrow \tau\tau$ events in the $\tau_h\tau_h$, $\tau_l\tau_h$ and $\tau_l\tau_l$ final states. The VH categories have around four times less expected events with between 3567 and 9516 background events. The event yields in the VBF categories are 3449, 3636 and 1295 background events in the $\tau_h\tau_h$, $\tau_l\tau_h$ and $\tau_l\tau_l$ final states.

The inclusive $H \rightarrow \tau\tau$ signal to background ratio differs between 0.010 and 0.059. The highest ratios are given for the inclusive VBF categories with at least 0.039, followed by the ttH category with 0.024. The VH categories have a signal to background ratio between 0.012

and 0.020. The boosted categories have the lowest ratio around 0.01.

The relative contributions of different Higgs boson production modes to the inclusive $H \rightarrow \tau\tau$ signal are summarized in Table 10.4. In the ttH category the highest contribution is assigned

Process	ggF	VBF	VH	$t\bar{t}H$	Other $H \rightarrow \tau\tau$
$\tau_h \tau_h$ [%]					
ttH incl.	13.45	1.30	6.21	69.49	9.55
VH incl.	59.84	6.61	32.62	0.56	0.37
VBF incl.	26.59	71.83	1.13	0.12	0.32
Boosted incl.	74.99	17.69	6.56	0.42	0.34
$\tau_l \tau_h$ [%]					
VH incl.	62.96	6.63	29.84	0.33	0.25
VBF incl.	34.94	64.04	0.80	0.05	0.17
Boosted incl.	75.95	17.35	6.24	0.22	0.24
$\tau_l \tau_l$ [%]					
VH incl.	65.74	5.99	27.78	0.24	0.26
VBF incl.	27.01	71.92	0.69	0.02	0.35
Boosted incl.	75.86	17.31	6.21	0.22	0.40

Table 10.4: Contribution of individual Higgs boson production modes to the inclusive $H \rightarrow \tau\tau$ signal in each inclusive signal category. All contributions are shown in percent.

to $t\bar{t}H$ events with 69%, followed by ggF with 13%. The remaining signal categories have similar relative contribution of Higgs boson production modes across the different final states. The highest purity is given in the boosted categories with 75% to 76% ggF events, followed by VBF events with 17% to 18%. The VBF categories are dominated by the VBF production mode with up to 72% followed by the ggF production mode with between 27% and 35%. Therefore, all of these production modes are dominated by their corresponding production mode. The VH categories have a VH production mode contribution between 28% and 33%. Therefore, these categories are not dominated by the VH production mode, instead the ggF production mode has the highest contribution with between 60% and 66%. This illustrates the high ggF contamination in this VH category.

Figure 10.2 shows the relative contribution of background processes and the inclusive $H \rightarrow \tau\tau$ signal in each inclusive signal category. All categories are dominated by $Z \rightarrow \tau\tau$ background events with a contribution between 62% and 83%, except for the ttH category with 38% $Z \rightarrow \tau\tau$ and 33% top quark production events. The misidentified τ background contribution is reduced compared to the previous analysis, due to the new jet object reconstruction method based on particle flow objects (see Section 5.4). The relative contribution of this background increases for an increasing number of τ_h -leptons in the final state and differs between 8% in the $\tau_e \tau_\mu$ boosted category and 27% in the $\tau_h \tau_h$ VBF category. Similarly, the relative contribution of top quark production background events is reduced for the $\tau_e \tau_\mu$ (<13%) and $\tau_l \tau_h$ (<3%) categories compared to the previous analysis. This reduction shows the improved top quark production background rejection of the new b-jet identification algorithm, which is used in this analysis for the b-jet veto (see Section 5.4). The $\tau_e \tau_\mu$ categories have a relative contribution of $Z \rightarrow ll$ events below 2%, while this contribution in the previous analysis for the $\tau_e \tau_\mu + \tau_{e(\mu)} \tau_{e(\mu)}$ final state was up to 15%. The relative contribution of other background

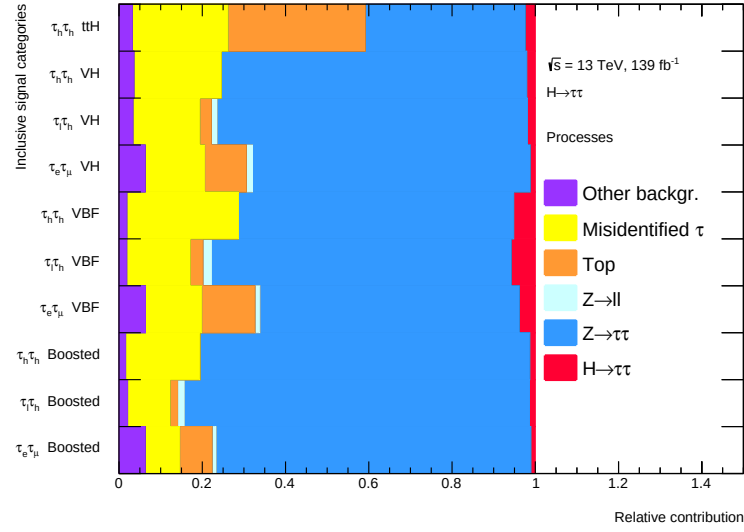


Figure 10.2: Relative contribution of processes in the inclusive signal categories. The Higgs boson signal is illustrated in red. The background processes are grouped to $Z \rightarrow \tau\tau$ (dark blue), $Z \rightarrow ll$ (light blue), top quark production (orange), misidentified τ (yellow) and other backgrounds (purple). Other backgrounds include all remaining non-listed processes described in Chapter 6. The contributions are shown separately for each final state.

processes varies between 2% and 6%.

10.2.4 Final signal categories

Each inclusive signal category is split into sub-categories, which will be used in the final likelihood model. The ttH , VH and VBF categories are split into two sub-categories, while the boosted category is split into six sub-categories. The goal for the ttH , VH and VBF categories is to increase the significance for the cross section measurement of the corresponding inclusive production mode by introducing a better signal to background separation. For this purpose, the inclusive categories are split into one sub-category with an increased signal to background ratio, while the second sub-category includes the remaining events of the inclusive categories and has a lower signal to background ratio. Boosted decision trees BDTs are trained using the TMVA implementation [239] and the output scores are used to divide the inclusive categories. In contrast to this, the boosted categories are not split to increase the significance for the inclusive cross section measurement. Instead, they are split into sub-categories depending on transverse momentum of the Higgs boson p_T^H and the jet multiplicity N_j . These regions of phase space follow the STXS bin definitions at Stage 1.2 (see Section 7.1.4). The final signal event categories are summarized in Figure 10.3 and detailed descriptions of the individual categories are given in the following paragraphs.

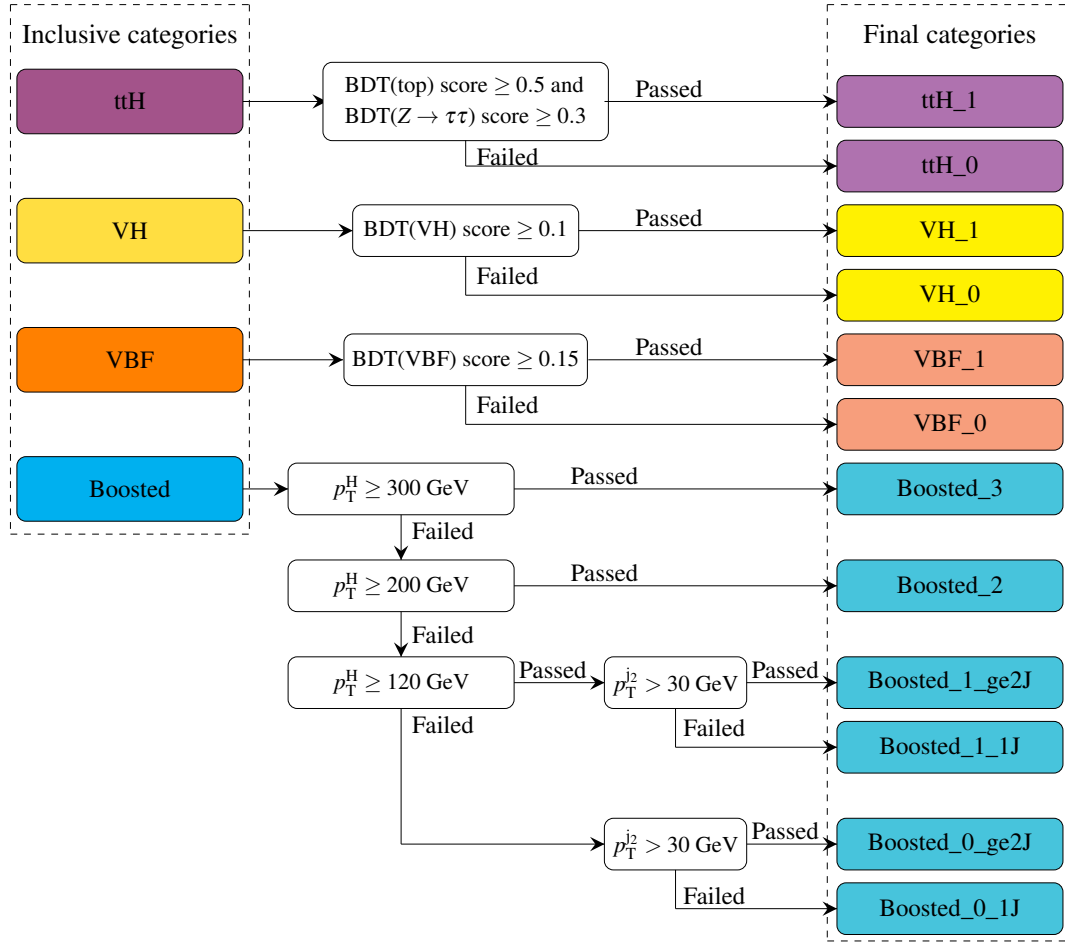


Figure 10.3: Schematic overview of the inclusive signal category split into the final signal categories used in the fit model. The colored boxes correspond to inclusive (darker color) and to final signal categories (lighter colors). The color type shows for which production mode the category is designed: Purple boxes for the $t\bar{t}H$, yellow for the $V(\rightarrow qq)H$, orange for the VBF and blue for the ggF production mode. White boxes illustrate requirements, which have to be passed or failed by an event to end up in a specific final category.

$t\bar{t}H$ signal categories The inclusive signal category has a similar relative background contribution from $Z \rightarrow \tau\tau$ and top quark production events (see Figure 10.2). Therefore, both background processes have to be reduced to increase the signal to background ratio in this category. To achieve a reduction of both backgrounds, two separated BDTs are trained. In both, the inclusive $t\bar{t}H \rightarrow \tau\tau$ production is used as signal and trained against $Z \rightarrow \tau\tau$ events in one BDT, while it is trained against top quark production events in the second BDT. Table 10.5 summarizes the variables used in the training of the two BDTs. A simple rectangular requirement in the plane of the two BDT scores is used to split the inclusive $t\bar{t}H$

10.2 Event selection

	Variable	VBF	$V(\rightarrow qq)H$	$t\bar{t}H$ vs $t\bar{t}$	$t\bar{t}H$ vs $Z \rightarrow \tau\tau$
Jet properties	Invariant mass of 2 leading jets	•	•		
	p_T^{jj}	•	•		
	Product of η of 2 leading jets	•			
	Sub-leading jet p_T	•			
	Leading jet η				•
	Sub-leading jet η				•
	Scalar sum of all jets p_T			•	•
	Scalar sum of all b -tagged jets p_T				•
	Best W -candidate di-jet invariant mass			•	•
	Best t -quark-candidate three-jet invariant mass			•	•
Angular distances	$\Delta\phi_{jj}$	•			
	$ \Delta\eta_{jj} $	•	•		
	ΔR_{jj}		•		
	$\Delta R_{\tau\tau,jj}$		•		
	$\Delta R_{\tau\tau}$		•	•	
	Smallest $\Delta R_{\text{any 2 jets}}$			•	
	$ \Delta\eta_{\tau\tau} $			•	•
τ prop.	$p_T^{\tau\tau}$			•	
	Sub-leading τ p_T				•
	Sub-leading τ η				•
p_T cand.	p_T^{Hjj}	•	•		
	p_T^{jj}/p_T^{jj}		•		
$\tilde{E}_T^{\text{miss}}$	Missing transverse momentum E_T^{miss}		•	•	•
	Smallest $\Delta\phi_{\tau, \tilde{E}_T^{\text{miss}}}$				•

Table 10.5: Variables used in the four BDTs implemented in the analysis. For each BDT, the presence of a • indicates that the variable is used for it. The symbol τ stands for any reconstructed τ lepton candidate (τ_e , τ_μ or τ_h) as appropriate in each final state. The symbols $\tau\tau$ and jj indicate the vectorial sums of the momenta of two τ -lepton candidates and of the two leading jets, respectively [28].

category into two sub-categories. A variety of more complex two-dimensional combinations of requirements on the two BDT scores are studied, but none of these show any significant improvement of the final significance. Figure 10.4 shows the post-fit distributions for both BDT scores in the inclusive $t\bar{t}H$ category. The $t\bar{t}H_1$ category is defined by applying the two requirements: $\text{BDT}(\text{top}) \text{ score} \geq 0.5$ and $\text{BDT}(Z \rightarrow \tau\tau) \text{ score} \geq 0.3$, while the $t\bar{t}H_0$ category includes the remaining events of the inclusive $t\bar{t}H$ category that fail these requirements (see Figure 10.3). The expected event yields for the inclusive $H \rightarrow \tau\tau$ signal process and background processes in $t\bar{t}H_0$ and $t\bar{t}H_1$ are shown in Figure 10.5 (a), while Figure 10.5 (b) shows the relative contribution of them to the total yields. The expected yield in $t\bar{t}H_0$ are 744 events and in $t\bar{t}H_1$ 60 events. The largest contribution corresponds to the $Z \rightarrow \tau\tau$ events and top quark production with 38% and 34% in $t\bar{t}H_0$ and 44% and 23% in $t\bar{t}H_1$. The signal

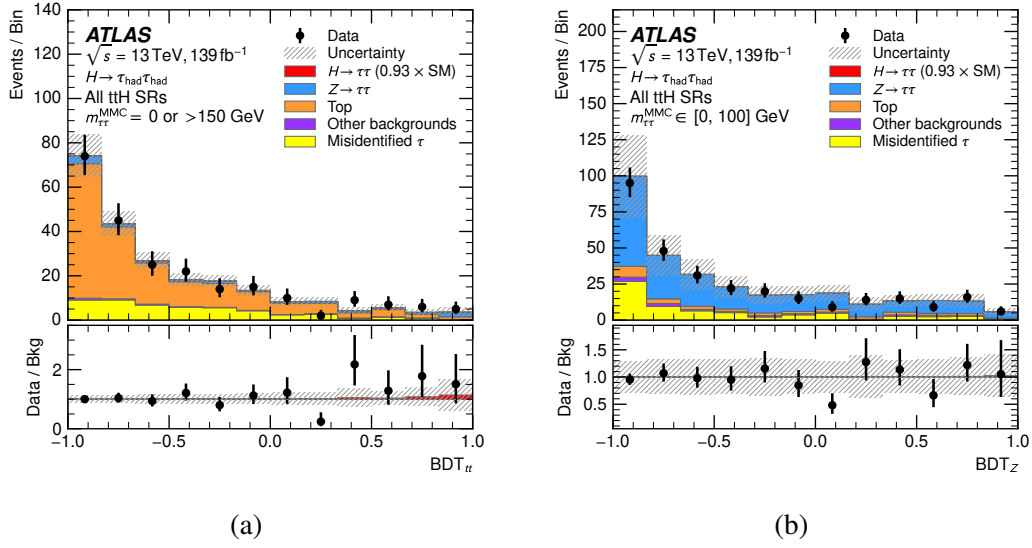


Figure 10.4: Post-fit event yield for the $t\bar{t}H$ inclusive category estimated with the inclusive $H \rightarrow \tau\tau$ PoI setup as a function of the $\text{BDT}(\text{top})$ score (a) and $\text{BDT}(Z \rightarrow \tau\tau)$ score (b) [28].

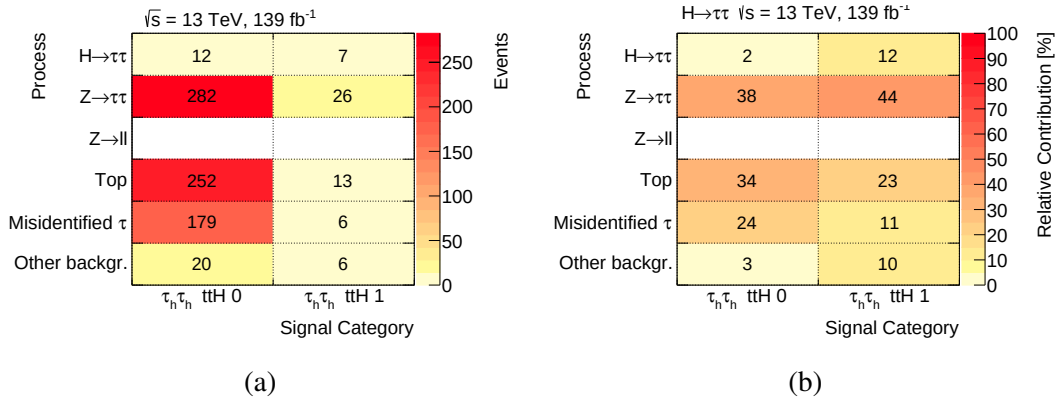


Figure 10.5: Absolute (a) and relative (b) event yields for the $t\bar{t}H_0$ and $t\bar{t}H_1$ categories for the inclusive $H \rightarrow \tau\tau$ signal and background processes.

contribution increases for $t\bar{t}H_1$ compared to $t\bar{t}H_0$ from 2% to 12%, while the misidentified τ background contribution is reduced from 24% to 11%. The two BDTs used to separate the two signal categories therefore increase the signal to background ratio for $t\bar{t}H_1$ and reduce especially the top quark production and misidentified τ background contributions.

VH signal categories In contrast to the inclusive $t\bar{t}H$ category, in which the signal is dominated by the corresponding $t\bar{t}H$ production mode, the signal in the inclusive VH category

is not dominated by the targeted signal process. Table 10.3 shows that the expected number of ggF production events is twice as large as the expected VH production events. In the

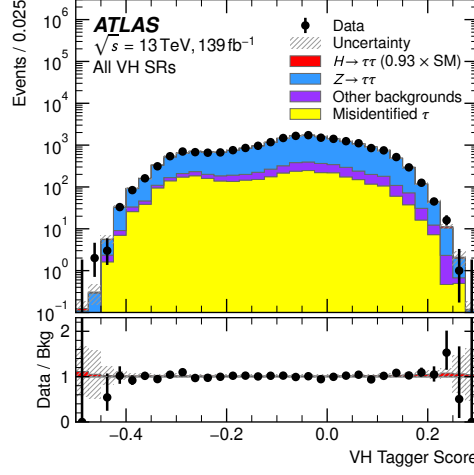


Figure 10.6: Post-fit event yield for the VH inclusive categories summed up over all final states estimated with the inclusive $H \rightarrow \tau\tau$ PoI setup as a function of the BDT(VH) score [28].

likelihood model the reconstructed mass of the di- τ -system is used as an observable (see Section 10.5), which increases the signal to background separation, but not the separation of different Higgs boson production modes. Therefore, the inclusive VH category is split into two sub-categories to increase the VH contribution compared to other Higgs boson production modes. A BDT is trained using the VH production as signal and other Higgs boson production modes as background. Table 10.5 summarizes the input variables used for this BDT. Most of the input variables correspond to properties of the two leading jets in the events. These jets most likely originate from the V boson for the $V(\rightarrow qq)H$ production, while they correspond to either pile-up jets or jets initiated by quarks or gluons radiated in the Higgs boson production for other production modes. A requirement on the BDT score, which is shown in Figure 10.6, is used to define a signal category VH_1, which has an increased VH ratio, while all events of the VH inclusive category that fail this requirement are assigned to the VH_0 category. The significance of measuring the VH signal using these two categories was estimated for several different BDT score requirements. A split at BDT(VH) score = 0.1 has shown the best significance and is used for the final definition of the VH_0 and VH_1 categories in all three final states (see Figure 10.3). Figure 10.7 shows the expected event yields of the inclusive Higgs boson signal and background processes in the VH_0 and VH_1 categories. The total yields for the VH_0 categories are 7035, 9134 and 3441 events and for the VH_1 categories 743, 730 and 313 events in the $\tau_h\tau_h$, $\tau_l\tau_h$ and $\tau_l\tau_l$ final states. Therefore, the total yields are around ten times larger in the VH_0 categories compared to the VH_1 categories. The used BDT optimized to separated VH production from other Higgs

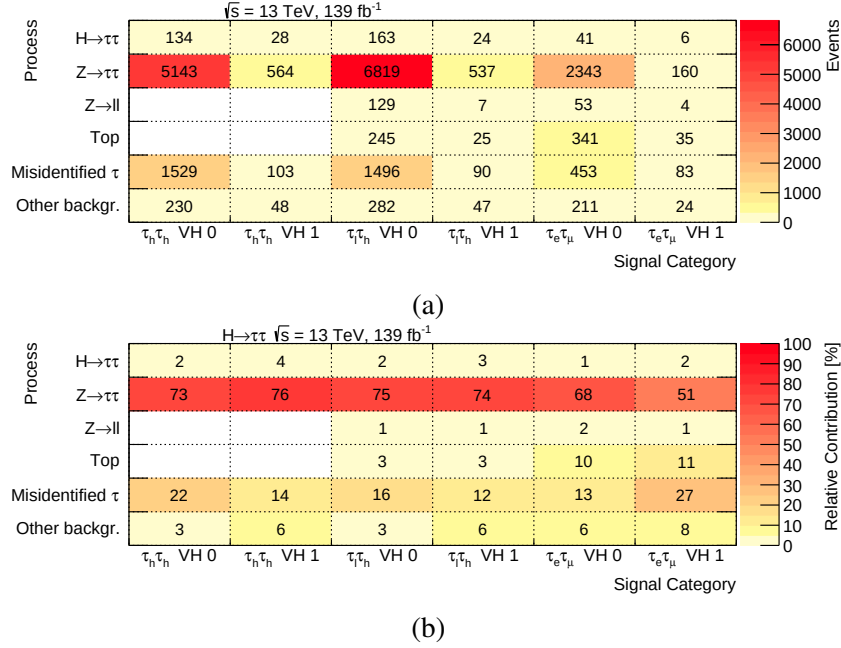


Figure 10.7: Absolute (a) and relative (b) event yields for the VH_0 and VH_1 categories for the inclusive $H \rightarrow \tau\tau$ signal and background processes.

boson production modes also improves the relative inclusive signal contribution for the VH_1 categories compared to the VH_0 from 2% to 4%, 2% to 3% and 1% to 2% in the $\tau_h\tau_h$, $\tau_l\tau_h$ and $\tau_l\tau_l$ final states. The main background contribution in all signal categories are $Z \rightarrow \tau\tau$ events with 73% to 76% in the final states with at least one hadronically decaying τ -lepton and with 51% to 68% in the $\tau_l\tau_l$ final state. In the $\tau_l\tau_l$ final state the misidentified τ background contribution is smaller for the VH_1 category compared to the VH_0 category from 13% to 27%, while in the remaining final states the contribution is reduced.

VBF signal categories The inclusive VBF category has a signal purity of the targeted signal between the purities in the ttH and VH inclusive categories. The signal to $Z \rightarrow \tau\tau$ ratio is higher than in the VH category, while other background events are not as important as for the ttH category. The contamination of other Higgs boson production modes in the VBF inclusive category is higher than for the ttH category but not at the same size as for the VH category. Therefore, a BDT using the VBF production mode as signal and $Z \rightarrow \tau\tau$ and ggF production events as background is trained in the VBF inclusive category. In this way both the VBF to $Z \rightarrow \tau\tau$, but also the VBF to ggF separation can be increased. Observables based on the kinematics of the two leading jets are used as input variables for the BDT (see Table 10.5). Figure 10.8 shows the post-fit event yields as a function of the BDT(VBF) score in the VBF inclusive category summed up over all final states. The inclusive VBF category is split into a category VBF_1, with a higher relative contribution of VBF production events, by requiring

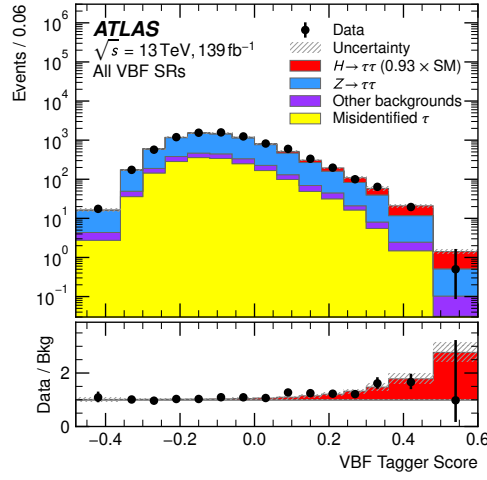


Figure 10.8: Post-fit event yield for the VBF inclusive categories summed up over all final states estimated with the inclusive $H \rightarrow \tau\tau$ PoI setup as a function of the BDT(VBF) score [28].

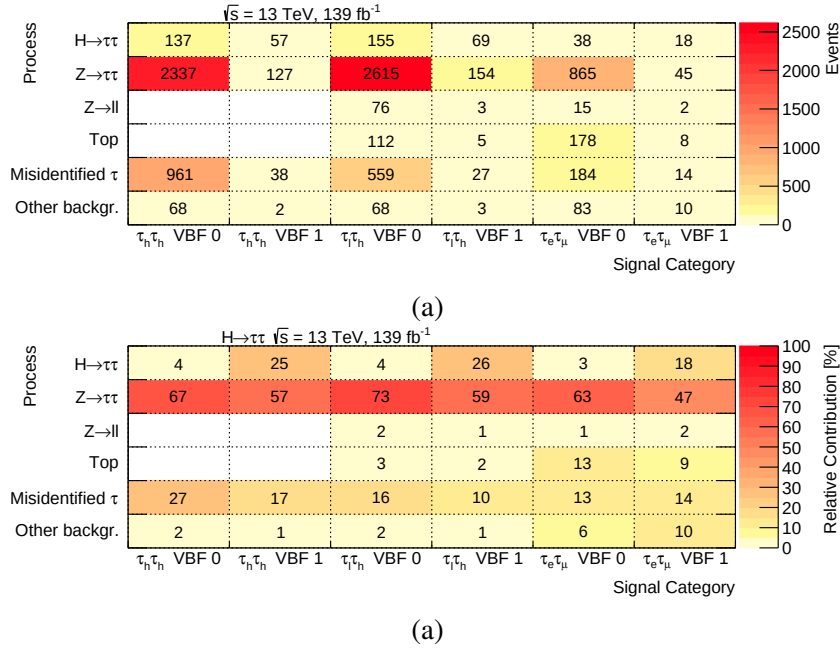


Figure 10.9: Absolute (a) and relative (b) event yields for the VBF_0 and VBF_1 categories for the inclusive $H \rightarrow \tau\tau$ signal and background processes.

BDT(VBF) score ≥ 0.15 , while all events of the inclusive VBF category, which fail this requirement are assigned to the VBF_0 category (see Figure 10.3). Figure 10.9 shows the expected yields for the inclusive Higgs boson signal and background processes in the VBF_0 and VBF_1 categories for all final states. The total event yields in the VBF_0 categories are 3504, 3585 and 1363 events and in the VBF_1 categories 224, 261 and 97 events in the $\tau_h \tau_h$, $\tau_l \tau_h$ and $\tau_l \tau_l$ final states. While the $H \rightarrow \tau\tau$ signal contribution for the VBF_1 categories differs between 18% and 26%, the signal contribution in the VBF_0 categories is around 4% for all final state. Therefore, the additional requirement on the BDT score increases the signal contribution of the inclusive VBF signal categories (4%-6%), while the contribution in the VBF_0 categories are below these values. For all final states the relative contribution of $Z \rightarrow \tau\tau$ events is smaller for the VBF_1 than for the VBF_0 categories. The misidentified τ background is slightly higher in the VBF_1 than in the VBF_0 category for the $\tau_l \tau_l$ final state, while in the remaining final state the contribution is smaller.

Boosted signal categories Due to the large expected statistics in the inclusive boosted categories, the categories can be split in more than two sub-categories. In contrast to the other categories no BDT is used in the boosted category to increase the signal to background ratio. Instead, the inclusive categories are split into smaller regions of phase space by using requirements on the jet multiplicity and transverse momentum of the Higgs boson p_T^H . Figure 10.10 shows the post-fit event yield for the boosted inclusive categories summed up over all final states as a function of p_T^H . Four different regions as a function of p_T^H are defined: $100 \leq p_T^H < 120$ GeV, $120 \leq p_T^H < 200$ GeV, $200 \leq p_T^H < 300$ GeV and $300 \text{ GeV} \leq p_T^H$. In addition, the regions with $p_T^H < 200$ GeV are split into 1-jet and ≥ 2 -jet categories, by

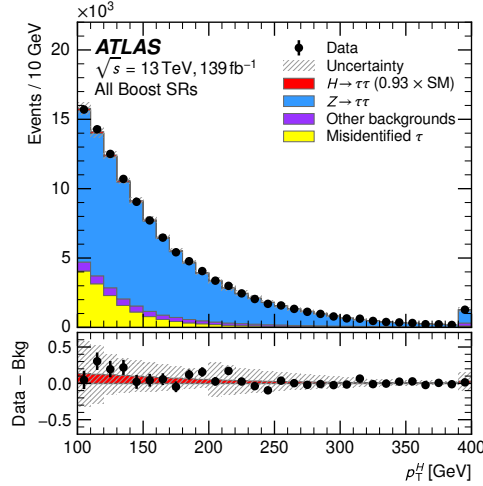


Figure 10.10: Post-fit event yield for the boosted inclusive categories summed up over all final states estimated with the inclusive $H \rightarrow \tau\tau$ PoI setup as a function of the transverse momentum of the Higgs boson p_T^H [28].

requiring $p_T^{j_2} > 30$ GeV for ≥ 2 -jet categories, while events which fail this requirement are assigned to the 1-jet categories. Figure 10.3 summarizes the six boosted signal categories. The categories are named boosted_0-boosted_3, where the number increases for categories with a higher requirement on p_T^H . The categories boosted_0 and boosted_1 have either _1J or _ge1J as a suffix corresponding to 1-jet and ≥ 2 -jet categories. These signal categories are targeting individual STXS bins of Stage 1.2 (see Figure 10.12 (b)). Figure 10.11 shows the expected

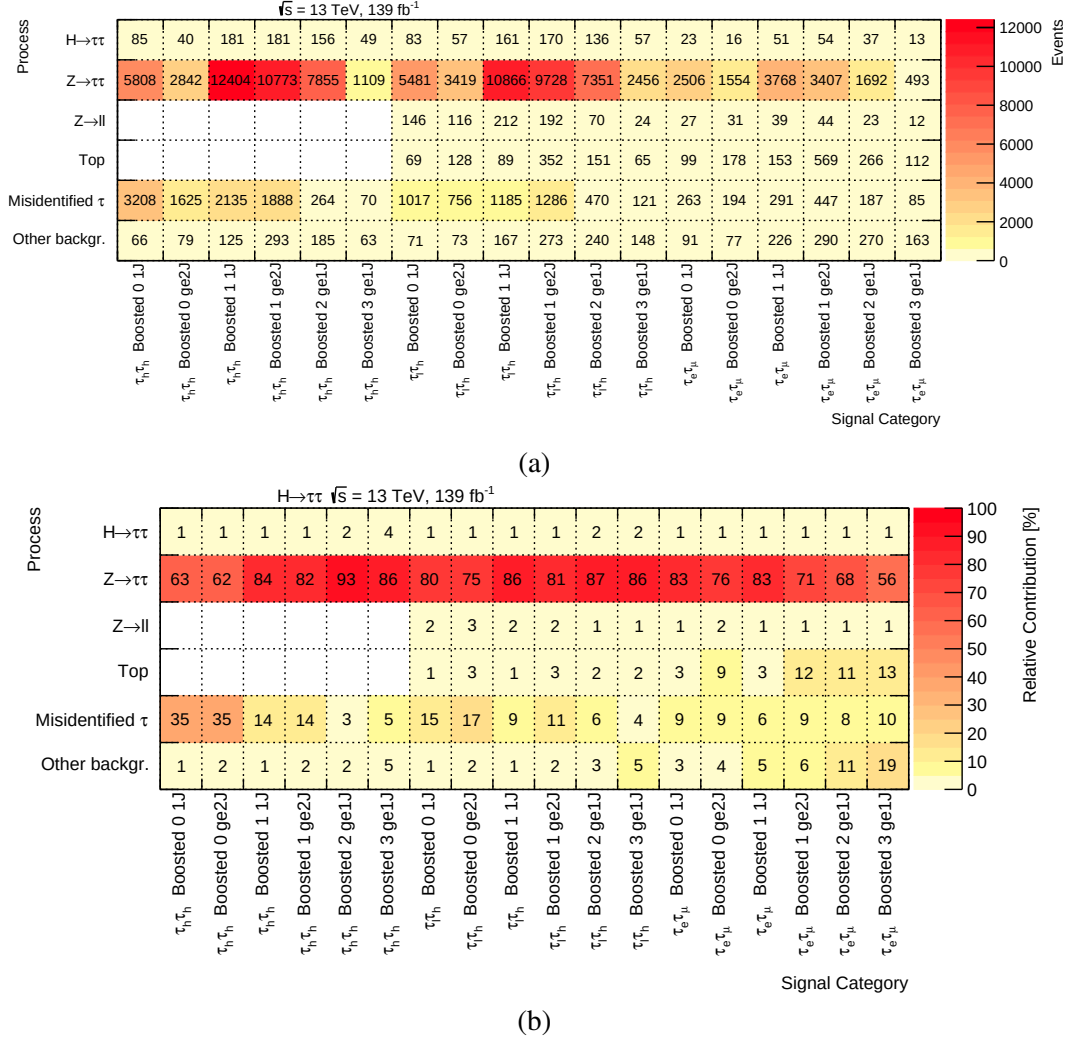


Figure 10.11: Absolute (a) and relative (b) event yields for the boosted categories for the inclusive $H \rightarrow \tau\tau$ signal and background processes.

yields of physics processes in the boosted signal categories for all final states. The boosted_1 categories have the largest expected total yields with 12001 to 148146 events in final states with at least one hadronically decaying τ lepton and between 4529 and 4811 events in the

$\tau\tau$ final state. The total yields decrease for boosted categories with larger p_T^H requirements and are overall similar or smaller for ≥ 2 -jet categories compared to the corresponding 1-jet category. The relative signal contribution increases for higher p_T^H categories from 1% up to 4%, while the contribution of misidentified τ background decreases from 35% up to 3%. The top quark production background contribution is relatively flat across the p_T^H categories with 3% to 13% in the $\tau\tau$ final state and with 1% to 3% in the $\tau\tau_h$ final state. The top quark production modes contribution is smaller in inclusive 1-jet categories compared to their corresponding ≥ 2 -jet categories. In the $\tau_e\tau_\mu$ final state an increase of other backgrounds is visible for categories with larger p_T^H from 3% to 19%, corresponding to an increase of the relative contribution of di-boson background events.

Comparison of signal categories and STXS bins at Stage 1.2 The VH and VBF categories are designed to measure specific STXS bins of the $EW\ qqH$ production process at Stage 1.2. Figure 10.12 (a) shows a comparison of the VH and VBF signal categories and the $EW\ qqH$ STXS bins at Stage 1.2. The figure illustrates that VH inclusive category matches the $EW\ qqH$ STXS bin with $N_j \geq 2$, $60 \leq m_{jj} < 120$ GeV, so that both final VH categories are designed to target this STXS bin. The inclusive VBF category targets the $EW\ qqH$ STXS bins with $N_j \geq 2$, $350 \text{ GeV} \leq m_{jj}$. The split in VBF_0 and VBF_1 is only designed to reduce background contribution but not to target a sub-set of these STXS bins. Therefore, it is not expected that the likelihood model based on these categories can disentangle the different p_T^H and p_T^{Hjj} STXS bins in this region of phase space. Therefore, these bins will be merged together for the rest of this analysis and quoted as one $EW\ qqH\ N_j \geq 2$, $350 \text{ GeV} \leq m_{jj}$ STXS bin. Several STXS bins have no or only a small overlap with the signal categories so that the contribution of these STXS bins is expected to be negligible for the analysis.

Figure 10.12 (b) compares the boosted signal categories with the STXS 1.2 bins for the ggF production mode. The event selection mimics the STXS bin definition for the STXS bins with $60 \text{ GeV} \leq p_T^H$, with the exception that the boosted signal categories require at least $100 \text{ GeV} \leq p_T^H$ and that one common signal category is defined for all STXS bins with $300 \text{ GeV} \leq p_T^H$. The difference in the lower threshold is needed to reduce large background contributions in the event selection, especially from QCD multi-jet events, while the common signal category for $300 \text{ GeV} \leq p_T^H$ is needed to achieve a reasonable data statistics in this signal category. The four STXS bins on the right side of Figure 10.12 (b) surrounded by a red box, correspond to ggF production STXS bins with a VBF-like topology. These bins are defined by requiring $N_j \geq 2$, $350 \text{ GeV} \leq m_{jj}$, which corresponds to the definition of the $EW\ qqH$ STXS bin, which is targeted with the VBF categories, except for an additional requirement of $p_T^H < 200 \text{ GeV}$. These STXS bins have no specific signal category, however they are expected to have non-negligible contributions in the VBF categories and boosted categories with at least two jets. Similar to the $EW\ qqH$ STXS bins none of these categories are sensitive to disentangle these four ggF VBF-like bins, so that they are combined to one common STXS bin in this analysis with: $N_j \geq 2$, $350 \text{ GeV} \leq m_{jj}$, $p_T^H < 200 \text{ GeV}$. The remaining ggF STXS bins at Stage 1.2 are not explicitly covered by the signal categories and their contributions are negligible.

The $t\bar{t}H$ categories, similar to the VH and VBF categories, are not designed to measure specific

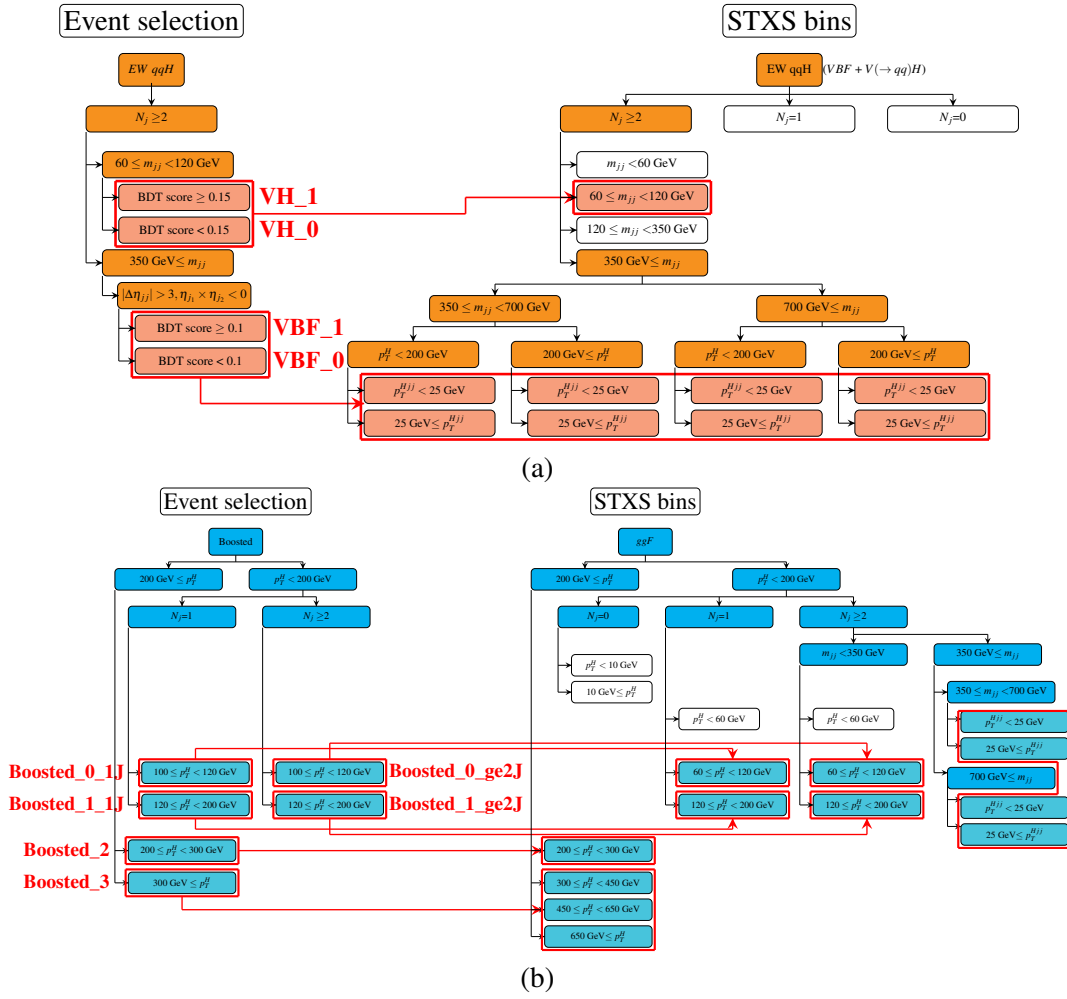
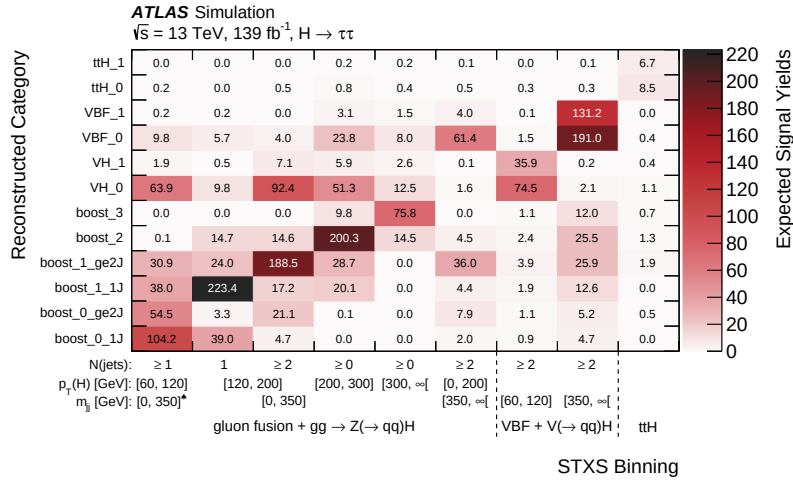


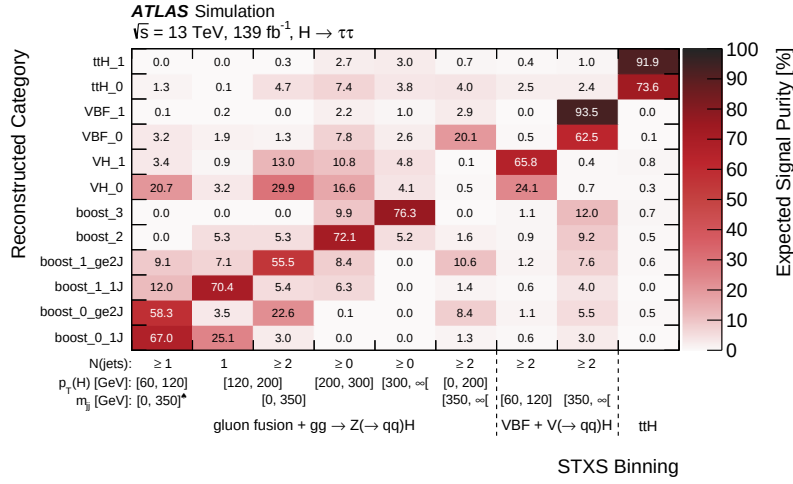
Figure 10.12: Comparison of the signal categories defined on reconstruction level on the left side and the Stage 1.2 STXS bins defined on truth level on the right side for the EW qqH (a) and ggF (b) production mode. For the left side, light blue and melon filled boxes refer to final signal categories, while dark blue and orange boxes refer to requirements in the inclusive signal categories. On the right side light blue and melon filled boxes refer to STXS bins, while the dark blue and orange ones are only given to group the STXS bins. White STXS bins have none or only small overlaps with the signal categories. The signal categories are connected to their corresponding STXS bins by red arrows. STXS bins surrounded by a red box have common signal categories.

STXS bins of the $t\bar{t}H$ production at Stage 1.2. Therefore, all $t\bar{t}H$ STXS bins are combined to one common $t\bar{t}H$ STXS bin for this analysis.

Figure 10.13 shows the expected yields of each of the STXS bins in this reduced Stage 1.2 definition. The yields are summed up for each signal category across the final states, since the



(a)



(b)

Figure 10.13: Absolute (a) and relative (b) contribution of STXS bins at the reduced Stage 1.2 in the final signal categories. The contributions are estimated summed up across all final states for each signal category and are normalized to the total yield of all listed STXS bins [28].

relative contribution of the STXS bins are very similar between them. The ttH_0 and ttH_1 categories have a high purity of $t\bar{t}H$ signal events with 74% and 92%, as well as the VBF_0 and VBF_1 categories have a high purity of the EW qqH $N_j \geq 2, 350 \text{ GeV} \leq m_{jj}$ STXS bin with 63% and 94%. The split of the inclusive VH categories, which are dominated by the

ggF production (see Table 10.3) results in VH_1 categories, which are dominated by the $EW\ qqH$ $N_j \geq 2$, $60 \leq m_{jj} < 120$ GeV with 66%, while the VH_0 categories are still dominated by STXS bins corresponding to the ggF production mode combined relative contribution of 75%. In all boosted categories the targeted ggF STXS bin is dominating with a relative contribution between 56% and 76%. The purity increases for boosted categories with higher values of p_T^H . The ggF VBF-like STXS bin with $N_j \geq 2$, $350 \text{ GeV} \leq m_{jj}$, $p_T^H < 200$ GeV, has a relative contribution between 8% and 11% for the boosted categories with $p_T^H < 200$ GeV and at least two jets, and 20% in the VBF_0 categories. The contribution of this STXS bin in the remaining categories is below 5%.

10.3 Background estimation

The background processes are estimated via simulations and data-driven techniques. The dominant background for this analysis is the $Z \rightarrow \tau\tau$ process with a contribution of 79% of the total background and up to 90% in single signal categories. This background is estimated using Monte Carlo simulations, which are validated with $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ events that can be selected with high purity. In the remaining analysis these two processes will be abbreviated as $Z \rightarrow ll$. To mimic as well as possible the kinematics of the $Z \rightarrow \tau\tau$ process in the signal categories, the selected $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ events are modified using a rescaling algorithm called object-based embedding.

The second largest background are misidentified τ background events, in which a jet is misidentified as an electron, muon or τ_h -lepton. This background is estimated with data-driven techniques developed separately for each final state.

Processes including a decaying top quark are a sizable contribution in the $\tau_e\tau_\mu$ and $\tau_l\tau_h$ final states as well as in the ttH signal categories. This background is estimated using Monte Carlo simulation to predict the background shape, while the absolute normalization is estimated via NFs using additional control regions.

The remaining background events are fully derived by Monte Carlo simulation and have a relative contribution between 1% and 20% in the signal categories.

$Z \rightarrow \tau\tau$ background A new technique was developed to estimate the $Z \rightarrow \tau\tau$ background in this analysis. The idea of the method is based on the embedding procedure [240] used in the $H \rightarrow \tau\tau$ analysis based on the Run 1 data set [233]. In this procedure $Z \rightarrow \mu\mu$ data events are used. The muons in these events are removed on detector cell level and simulated τ -lepton are inserted instead. This way, only the τ -lepton decay itself is estimated via simulation, while the rest of the event is taken from data. To use this technique the information about the energy deposit in each detector cell has to be used, which makes this technique rather complicated and time consuming. Therefore, a new technique was developed using the same concept but implemented on object level.

Fully reconstructed $Z \rightarrow ll$ simulated events are used and parametrized to mimic the $Z \rightarrow \tau\tau$ kinematics. The method can be staged in three steps. First, the effects of the reconstruction and trigger efficiency of the electrons and muons are unfolded from the events. In this way the reconstructed $Z \rightarrow ll$ are unfolded to $Z \rightarrow ll$ events with truth-level leptons (neglecting

resolution effects). On these events, a re-weighting based on kinematic distributions is performed, which corrects for the kinematic difference between electrons/muons and τ -leptons. Finally, the reconstruction and trigger efficiencies for τ -lepton are applied to the events. This parametrization is derived separately for τ_e , τ_μ and τ_h -leptons. Therefore, three different sets of re-weighted events are calculated, one for each of the three di- τ final states. In this way the same set of events could be used in each final state, so that each of these events would be used three times in the analysis. To avoid this the $Z \rightarrow ll$ events are randomly split into three equal sized sub-samples before applying the object-based embedding. For each of the three final states only one of these sub-samples is used.

This method can be validated by comparing the object-based embedded $Z \rightarrow ll$ Monte Carlo simulation with the $Z \rightarrow \tau\tau$ simulation. It can be used to validate the prediction of simulations by comparing the events with data in various observables of interest, but also to construct control regions that allow the estimation of the $Z \rightarrow \tau\tau$ normalization for each signal category. Figure 10.14 shows the validation of the technique for two example distributions. For both distributions the deviation are covered by the statistical uncertainties and no trend is visible. Dedicated uncertainties are assigned to the final re-weighted $Z \rightarrow ll$ events to account for small difference in modeling. These uncertainties are found to be below 1% for all regions.

In the likelihood model of this analysis the technique is used to estimated the $Z \rightarrow \tau\tau$

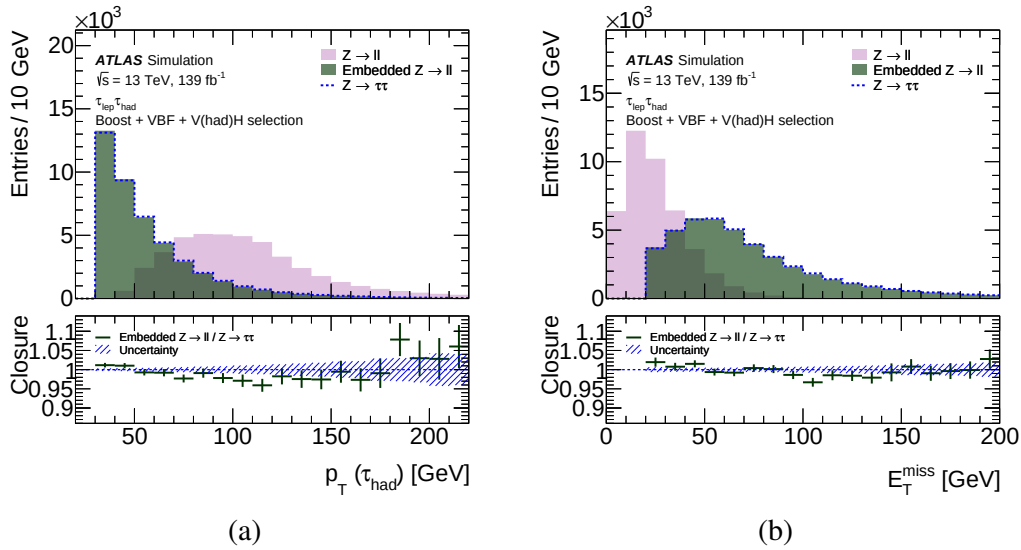


Figure 10.14: Comparison between $Z \rightarrow ll$ simulated events before and after the object-based parametrization and $Z \rightarrow \tau\tau$ simulated events as a function of $p_T^{\tau_h}$ (a) and E_T^{miss} (b). The distributions include all events of the inclusive boosted, VBF and VH signal categories. The lower panels show the ratio between the reweighted $Z \rightarrow ll$ events and $Z \rightarrow \tau\tau$ events, while the uncertainty band includes only statistical uncertainties [28].

background normalization for each signal category. Therefore, $Z \rightarrow ll$ control regions with a high purity of $Z \rightarrow ll$ events and kinematic properties as close as possible to each signal

category need to be constructed and re-weighted with the object-based embedding procedure. Finally, a common free normalization factor on the object-based embedded $Z \rightarrow ll$ events in the control region and on the $Z \rightarrow \tau\tau$ events in the corresponding signal region is defined to allow the likelihood fit to correct the $Z \rightarrow \tau\tau$ normalization based on the object-based embedded $Z \rightarrow ll$ control region.

The construction of these control regions is based on a modified version of the preselection for the $\tau_e\tau_\mu$ final state. Table 10.6 compares the preselection for the $\tau_e\tau_\mu$ final state with the

	Preselection SR $\tau_e\tau_\mu$	Preselection $Z \rightarrow ll$ CR $ee(\mu\mu)$
(1)	$N_e^{\text{loose}}=1, N_\mu^{\text{loose}}=1, N_{\tau_h}^{\text{loose}}=0$	$N_e^{\text{loose}}=2(0), N_\mu^{\text{loose}}=0(2), N_{\tau_h}^{\text{loose}}=0$
(2)	e/μ : Medium ID e/μ : FCLoose/FCTightTrackOnly	e/μ : Medium ID e/μ : FCLoose/FCTightTrackOnly
(4)	Opposite charge	Opposite charge
(5)	$m_{ll}^{\text{coll}} > m_Z - 25 = \text{GeV}$	
(6)	$30 < m_{ll} < 100 \text{ GeV}$	$m_{ll} > 80 \text{ GeV}$
(7)	$E_T^{\text{miss}} > 20 \text{ GeV}$	
(8)	$\Delta R_{e\mu} < 2.0$	$\Delta R_{ee} < 2.5$
(9)	$ \Delta\eta_{\tau\tau} < 1.5$	$ \Delta\eta_{\tau\tau} < 1.5$
(10)	$0.1 < x_1 < 1.0$	
(11)	$0.1 < x_2 < 1.0$	
(12)	$p_T^{\text{jl}} > 40 \text{ GeV}$	$p_T^{\text{jl}} > 40 \text{ GeV}$
(13)	$N_{\text{b-jet}} = 0$	

Table 10.6: Comparison of preselection requirements for the signal categories of $\tau_e\tau_\mu$ and the $Z \rightarrow ll$ control regions.

preselection for the $Z \rightarrow ll$ control regions. By requiring exactly two electrons or exactly two muons the orthogonality to the preselection of the signal categories is ensured. The remaining requirements follow the signal region preselection with some exceptions. Several requirements have been optimized to increase the data statistics or to account for kinematic differences of the $Z \rightarrow ll$ events. No neutrinos are expected in $Z \rightarrow ll$ events, so that no requirements on m_{ll}^{coll} , $x_{1(0)}$ and E_T^{miss} are applied. For m_{ll} , the thresholds are changed to increase the relative contribution of $Z \rightarrow ll$ events and the remove contribution from $Z \rightarrow \tau\tau$ in the control regions. These modifications introduce a bias between the signal categories and their corresponding control regions. This bias is approximately 2% and covered by the uncertainties, so that this effect can be ignored for this procedure. The ΔR requirement between the leptons is relaxed from 2 to 2.5, where 2.5 corresponds to the threshold used for the $\tau_l\tau_h$ and $\tau_h\tau_h$ final states. The b-jet veto requirements are removed for the $Z \rightarrow ll$ selection to increase the statistics.

On these events, the object-based embedding procedure is applied to re-weight the $Z \rightarrow ll$ events to the kinematics of $Z \rightarrow \tau\tau$ events. This process is applied on the predicted simulated $Z \rightarrow ll$ events, but the same re-weighting is also applied on the recorded data in this $Z \rightarrow ll$ control region. To construct control regions out of these preselected events for every signal category, all requirements of the corresponding signal category are applied, so that for each

of the signal categories one object-based $Z \rightarrow ll$ control region with the same kinematic properties is defined.

The only exceptions are the ttH categories, which require a large number of jets and b-jets in the signal categories. For these categories the selected $Z \rightarrow ll$ events at preselection are not sufficient to provide a reasonable data statistics to constrain the $Z \rightarrow \tau\tau$ normalization in the likelihood model. Instead, the normalization of the $Z \rightarrow \tau\tau$ background in ttH categories is estimated only in the ttH signal categories themselves. Therefore, a common normalization factor is applied on the $Z \rightarrow \tau\tau$ background in the ttH_0 and ttH_1 categories.

Figure 10.15 shows the relative contribution of the physics processes in $Z \rightarrow \tau\tau$ control regions for the remaining signal categories. The relative contribution of the object-based embedded $Z \rightarrow ll$ process is over 90% for all of these control regions. The remaining 10% of the control region are mainly sourced by top quark and W +jets productions.

τ misidentified background Compared to the previous analysis the methods used to estimate the misidentified τ background have changed for the $\tau_e\tau_\mu$ and $\tau_h\tau_h$ final state, while the same method is used for the $\tau_l\tau_h$ final state. For the $\tau_l\tau_h$ final state the method is only optimized to account for the increased data statistics and the new regions of phase space, which are included in this analysis. In all three final states the misidentified τ background is estimated in a similar procedure by selecting data events in a misidentified τ control region with the same requirements as for the signal categories, but loosening or inverting the criteria on the identification or isolation of the leptons. To correct for kinematic and normalization differences between the signal categories and their misidentified τ control region transfer factors are estimated.

In the $\tau_e\tau_\mu$ final state this background is estimated using the matrix-method technique [241]. The misidentified τ control region is defined by using the signal category requirements but inverting the isolation criteria on the electron and muon. For the electron additionally the identification criteria is loosened to increase the statistics. The expected number of misidentified τ background in the signal category is derived by a set of equations, in which the efficiencies for real ϵ_{real} and fake ϵ_{fake} leptons is related to the observed data yields. The efficiencies are estimated separately for muons and electrons as a function of p_T and η of the leptons. ϵ_{fake} is estimated using data events with inverted opposite charge requirement on the leptons, while ϵ_{real} is computed with simulated events. Several sources of uncertainties are taken into account for this estimation: statistical uncertainties on the derived ϵ_{real} and ϵ_{fake} , dependencies of ϵ_{real} on the used physics processes, dependencies of ϵ_{fake} on the number of jets and b-jets in the event and uncertainties on the normalization of the real lepton contribution in the measurement.

In the $\tau_l\tau_h$ final state the same fake factor method as described in Section 9.3 is used. However, the increased data statistics compared to the previous analysis allows to optimize the estimation further. Therefore, the parametrization of the fake factors as a function of p_T and η of the τ_h -lepton is estimated with finer granularity.

The misidentified τ background of the $\tau_h\tau_h$ final state is estimated with a fake-factor approach, similar to the $\tau_l\tau_h$ final state. However, in this final state the method is changed to simultaneously account for events with one and two jets misidentified as τ_h -leptons. The fake

factors are estimated in the $W + jets$ enriched control regions defined for the $\tau_l \tau_h$ final state, but with the additional requirement of the trigger-matching of the τ_h -lepton. To account for systematic uncertainties on the misidentified τ background two alternative sets of fake factors are estimated. These are estimated using control regions with two τ_h -leptons, where for the first set the $|\Delta\eta_{\tau\tau}|$ requirement is inverted and for the second the opposite charge requirement of the two τ_h -leptons is inverted. The differences between fake factors based on these event selections and the nominal estimation are used to derive uncertainties on the composition of the misidentified τ background. Furthermore, the statistical uncertainties on the fake factors and uncertainties related to the choice of the parametrization are taken into account.

For each final state, the modeling of the misidentified τ background is studied in dedicated validation regions. For the $\tau_e \tau_\mu$ and the $\tau_l \tau_h$ final state the opposite charge requirement on the two leptons is inverted, while for the $\tau_h \tau_h$ final state the $|\Delta\eta_{\tau\tau}|$ requirement is inverted. Figure 10.16 shows the validation of the modeling for the inclusive VBF categories in all three final states. The predicted events agree with the observed data within the given uncertainties and no significant mismodeling is observed in all three final states. Similarly, no significant deviations are observed in the remaining signal categories.

Top quark production background The top quark production background contribution for the $\tau_e \tau_\mu$ and the $\tau_l \tau_h$ final states are estimated in the same way as in the previous analysis described in Section 9.3. Therefore, one dedicated control region for each inclusive signal category is defined. The top quark production background has a small contribution in the $\tau_h \tau_h$ final state, except for the $t\bar{t}H$ categories. In these categories the shape for top quark production background is taken from simulations, but the normalization is estimated from data. Therefore, a common NF is applied on both $t\bar{t}H$ signal categories. Figure 10.15 shows the relative contribution of the different physics processes in the six top quark control regions defined for the $\tau_e \tau_\mu$ and the $\tau_l \tau_h$ final states. The purity of the top quark production background in these control regions differs between 70% and 80%.

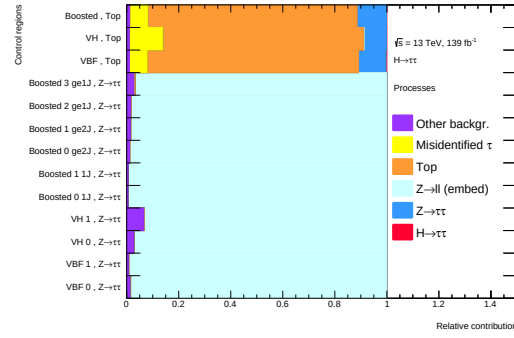
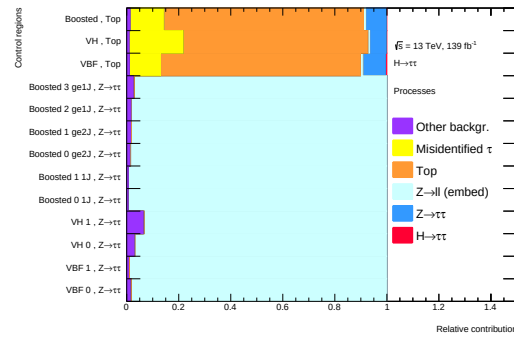
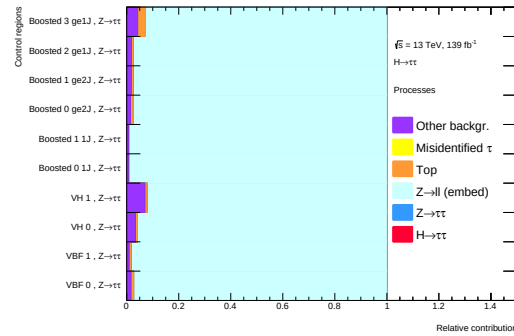
(a) $\tau_e \tau_\mu$ (b) $\tau_l \tau_h$ (c) $\tau_h \tau_h$

Figure 10.15: Relative composition from different processes in all control regions (listed on the y-axis) used in the likelihood model. The Higgs boson signal is illustrated in red. The background processes are grouped to $Z \rightarrow \tau\tau$ (dark blue), top quark production (orange), misidentified τ (yellow) and other backgrounds (purple). The $Z \rightarrow ll$ (light blue) corresponds to simulated $Z \rightarrow ll$ events in all top control regions, while in the $Z \rightarrow \tau\tau$ control regions $Z \rightarrow ll$ corresponds to object-based embedded $Z \rightarrow ll$ events. Other backgrounds include all remaining non-listed processes described in Chapter 6. The contributions are shown for each of the final states separately.

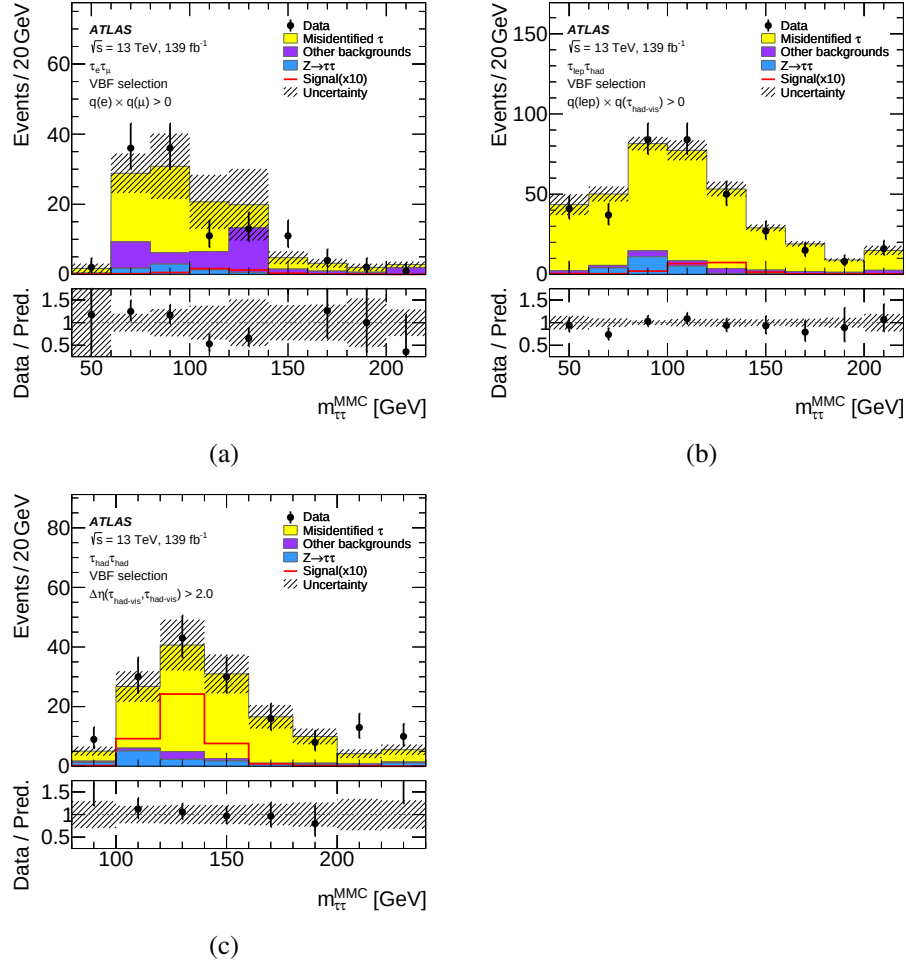


Figure 10.16: Validation of the τ misidentified background modeling in the inclusive VBF category for the $\tau_e \tau_\mu$ (a), $\tau_l \tau_h$ (b) and $\tau_h \tau_h$ (c) final states. The uncertainty band includes both, the statistical uncertainty due to the limited size of the simulated events and the systematic uncertainties of the data-driven methods [28].

10.4 Systematic uncertainties

The sets of systematic uncertainties, which are included in the measurement are based on the implementation of the previous analysis. The uncertainties on the misidentified τ background has been updated to the new methods described in the previous Section 10.3. Additionally, systematic uncertainties due to the new object-based embedding method have been included. The following section introduces all uncertainties grouped into three categories: theoretical uncertainties on signal, theoretical uncertainties on background and experimental uncertainties.

10.4.1 Theoretical uncertainties on signal

The LHC Higgs Cross Section Working Group provides common recommendations for the estimation and implementation of theoretical uncertainties on the Higgs boson production. This analysis follows these recommendations and uses an optimized scheme compared to the previous analysis. In general, the uncertainties cover the same three sets of sources as described in Section 9.4.

The PDF and α_s uncertainties are implemented and estimated as in the previous analysis, so that 30 eigenvectors variations are included for the PDF uncertainties and two variations of α_s . The corresponding NPs are correlated across the different Higgs production modes. As for the previous analysis, the QCD scale uncertainties are evaluated by varying the renormalization and factorization scales (μ_R and μ_F). Therefore, the values of the uncertainties have not changed with respect to Table 9.5, except for the $t\bar{t}H$ production. These uncertainties are neglected in the previous measurement, but in the new measurement the impact of the uncertainties is sizable, especially for the $t\bar{t}H$ categories. The size of these uncertainties are: QCD scale +5.8% (-9.2%), PDF $\pm 3\%$ and α_s 2%. As in the previous analysis the simple variation of the QCD scale is not sufficient for this measurement. The set of individual uncertainties, which are used to avoid an underestimation of the QCD scale uncertainties have been updated for ggF production compared to Section 9.4 and a similar procedure for the remaining production modes has been introduced [27, 242, 243].

For the ggF production mode the same QCD scale uncertainties as for the previous analysis are implemented but extended by five additional uncertainties, where one systematic covers the migration at the new p_T^H splitting for 0-jet events and four the new migration between the new m_{jj} STXS bins:

- One p_T^H -dependent uncertainty for the 0-jet bin
- Four m_{jj} migration uncertainties based on the STXS 1.2 bin definitions
 - m_{jj} 0-350 GeV $\leftrightarrow$ m_{jj} 350- ∞ GeV
 - m_{jj} 0-700 GeV $\leftrightarrow$ m_{jj} 700- ∞ GeV
 - m_{jj} 0-1000 GeV $\leftrightarrow$ m_{jj} 1000- ∞ GeV
 - m_{jj} 0-1500 GeV $\leftrightarrow$ m_{jj} 1500- ∞ GeV

In addition to these uncertainties a shape uncertainty has been assigned to the ggF production with more than one jet in the VBF and VH categories. These uncertainties take into account differences of the matrix-element calculations in these regions of phase space. For this, the nominal prediction of the ggF production is compared to an alternative Monte Carlo prediction generated by using the FxFx prescription to merge the jet multiplicities [244]. Furthermore, the use of jet related variables, which can mimic a veto on a third jet, in the VH and VBF BDT can lead to an underestimation of the QCD scale uncertainties. Two additional systematic uncertainties following the Stewart-Tackmann approach are introduced to cover these [245].

For the EW qqH production ten independent uncertainties for the QCD scale calculation are included in the measurement:

- One uncertainty related to the inclusive cross section
- One uncertainty related to the two jet requirement
- One uncertainty related to the p_T^H cut at 200 GeV
- One uncertainty related to the p_T^{Hjj} cut at 25 GeV
- Six m_{jj} migration uncertainties based on the STXS 1.2 bin definitions
 - m_{jj} 0-60 GeV $\leftrightarrow$ m_{jj} 60- ∞ GeV
 - m_{jj} 0-120 GeV $\leftrightarrow$ m_{jj} 120- ∞ GeV
 - m_{jj} 0-350 GeV $\leftrightarrow$ m_{jj} 350- ∞ GeV
 - m_{jj} 0-700 GeV $\leftrightarrow$ m_{jj} 700- ∞ GeV
 - m_{jj} 0-1000 GeV $\leftrightarrow$ m_{jj} 1000- ∞ GeV
 - m_{jj} 0-1500 GeV $\leftrightarrow$ m_{jj} 1500- ∞ GeV

For the $t\bar{t}H$ production mode six independent QCD scale uncertainties are defined:

- One uncertainty related to the inclusive cross section
- Five p_T^H migration uncertainties based on the STXS 1.2
 - p_T^H 0-60 GeV $\leftrightarrow$ p_T^H 60- ∞ GeV
 - p_T^H 0-120 GeV $\leftrightarrow$ p_T^H 120- ∞ GeV
 - p_T^H 0-200 GeV $\leftrightarrow$ p_T^H 200- ∞ GeV
 - p_T^H 0-300 GeV $\leftrightarrow$ p_T^H 300- ∞ GeV
 - p_T^H 0-450 GeV $\leftrightarrow$ p_T^H 450- ∞ GeV

Parton shower and matrix element uncertainties are estimated and included for the ggF , VBF , WH , ZH and $t\bar{t}H$ production modes independently. Alternative Monte Carlo samples are produced for the parton shower uncertainties using the matrix-element generator as for the nominal prediction, but varying the parton-shower model from PYTHIA8 to HERWIG7 [119,

237]. For the matrix-element uncertainties alternative predictions with the same parton-shower model, but different matrix-element generator are compared. Table 10.7 summarizes the final number of NPs that are included in the likelihood model to account for the theoretical uncertainties on the signal prediction, while Table A.4 in the appendix lists all NPs with their descriptions.

Uncertainty source	Number of NPs
Signal theory	77
PDF	30
α_s	1
QCD scale	36
Parton-shower	5
Matrix-element	5

Table 10.7: Number of NPs related to theoretical signal prediction in the fit model grouped into different sources.

10.4.2 Theoretical uncertainties on background

The strategy of the parametrization of the theoretical uncertainties on the $Z \rightarrow \tau\tau$ background has changed compared to the previous analysis due to the implementation of the new object-based embedding method. The same six sources of uncertainties are taken into account: PDF, $\Delta\mu_R, \Delta\mu_F$, CKKW, QSF, underlying-event and parton-shower (see Section 9.4) [136, 246, 247].

The PDF uncertainty is based on 100 weight variations of SHERPA [117], which are combined to one final uncertainty. To account for kinematic differences across the categories in the analysis, a fully decorrelated NP scheme is used for the PDF uncertainty. Therefore, the normalization and shape component of the uncertainty are decorrelated and furthermore both components are decorrelated across the final states and each signal category. The correlation of the NPs between signal categories and their $Z \rightarrow \tau\tau$ control region are kept.

The factorization and renormalization scale variations ($\Delta\mu_R, \Delta\mu_F$) are estimated in the same way as for the previous analysis, but they are implemented with the same decorrelation scheme that is applied on the PDF uncertainties.

The remaining theoretical uncertainties on the $Z \rightarrow \tau\tau$ background are estimated and implemented in the same way as described in Section 9.4, except for a decorrelation of the NPs between the ttH categories and the remaining categories. This decorrelation takes into account, that no $Z \rightarrow \tau\tau$ control regions are defined for the ttH categories.

In contrast to the previous analysis the same set of NPs as for the $Z \rightarrow \tau\tau$ background is also applied to $Z \rightarrow ll$ events. To account for theoretical uncertainties on the top quark production and di-boson background in the object-based embedded control regions, two uncertainties on the total cross section of these processes are included. Finally, the uncertainties on the modeling of the minor $Z \rightarrow \tau\tau$ electroweak (EWK) contribution is known to be underestimated for some variables used in the VBF BDT. Therefore, an additional uncertainty on the QCD to

EWK ratio between the VBF_0 and VBF_1 categories is applied on the $Z \rightarrow \tau\tau$ background. The new signal categories designed for the $t\bar{t}H$ production mode have a significant impact from top quark production background, so that theoretical uncertainties on this background are included additionally in the fit model. The following sources of uncertainties are covered in the analysis:

- PDF: based on 30 eigenvector variations [248, 249]
- Matrix-element: Comparison of alternative predictions using the same parton-shower model but a different matrix-element generator [136]
- Parton-shower: Comparison of alternative predictions using the different parton-shower models but the same matrix-element generator [237, 249, 250]
- Final state radiation: Based on event-weights calculated from the reweighting of the nominal prediction [251]
- Initial state radiation: Based on event-weights calculated from the reweighting of the nominal prediction [251]

Table 10.8 summarize the number of final independent NPs assigned to theoretical uncertainties on the background prediction in the likelihood model, while Table A.4 in the appendix lists all NPs with their descriptions.

Uncertainty source	Number of NPs
Background theory	172
Z PDF	64
Z $\Delta\mu_R, \Delta\mu_F$	64
Z Others	9
Top PDF	30
Top Others	5
Others	1

Table 10.8: Number of NPs related to theoretical background prediction in the fit model grouped into different sources.

10.4.3 Experimental uncertainties

The same sources of experimental uncertainties as for the previous analysis are taken into account (see Section 9.4). However, due to the use of optimized reconstruction, identification and trigger requirements the individual uncertainty schemes have changed. The increased data statistics and the new region of phase space included in the event selection increased the sensitivity to specific sources of experimental uncertainties, so that for these a more complex uncertainty scheme is used.

For reconstructed muons the uncertainties recommended in [200] are used. These cover

Leptons and jets				
Uncertainty source	Electrons	Muons	τ -leptons	Jets
All	32	12	27	45
Resolution	1	3		13
Energy scale	1	1	5	30
Reconstruction/Identification	26	4	11	
Isolation	1	2		
Trigger	1	2	8	
Charge identification	2			
Electron veto			3	
Jet vertex tagger				2
Misidentified- τ background				
Uncertainty source	$\tau_l \tau_l$	$\tau_l \tau_h$	$\tau_h \tau_h$	
All	15	42	7	
Others				
Uncertainty source	b-jet	E_T^{miss}	Luminosity	Object-based embedding
All	25	3	1	30

Table 10.9: Number of NPs related to experimental uncertainty in the fit model grouped into different sources.

energy momentum selection and scale, reconstruction, identification, isolation and trigger efficiency uncertainties. Uncertainties assigned for the same sources are also included for electrons [252]. In addition, two systematic uncertainties covering the charge identification of electrons are included. First studies of the likelihood model has shown small pulls of the electron identification NP so that a more complex scheme compared to the muons was included to account for kinematic dependencies of this uncertainty source. The uncertainties related to τ -lepton reconstruction and identification are estimated using the approach in [202]. They cover uncertainties related to energy scale, reconstruction, identification, trigger and electron veto efficiency. For uncertainties related to jets, the R4 CategoryReduction FullJER described in [204] is used. This scheme includes 30 NPs covering energy scale uncertainties and 13 NPs assigned to energy resolution uncertainties [253]. Uncertainties related to the jet vertex tagger and the b-tagging are covered by dedicated NPs. Since the used working point for the b-jet identification is different the $\tau_h \tau_h$ compared to the other final states, the NPs are decorrelated across those. Uncertainties on the missing transverse energy [220] and the luminosity [254] calculated with LUCID-2 [255] are included as well. Finally, NPs related to the object-based embedding method and the data-driven estimation of the misidentified τ background are included. For the object-based embedding one independent parameter per signal categories is implemented, while for the misidentified τ background specific parameters per final state are defined. For the $\tau_e \tau_\mu$ final state 15 uncertainties are included and seven for the $\tau_h \tau_h$. In the $\tau_l \tau_h$ final state 14 uncertainties are included, which are decorrelated

across the inclusive signal categories. Table 10.9 summarize the number of NPs assigned to experimental uncertainties in the likelihood model, while Table A.5 in the appendix lists all NPs with their descriptions.

10.5 Statistical model

10.5.1 Parameters of interest and normalization factors

As for the previous analysis a profiled likelihood fit is used to estimated the PoIs, NFs and the impact of systematic uncertainties implemented as NPs (see Section 7.2).

Several NFs are implemented to estimate the final background contribution in the likelihood

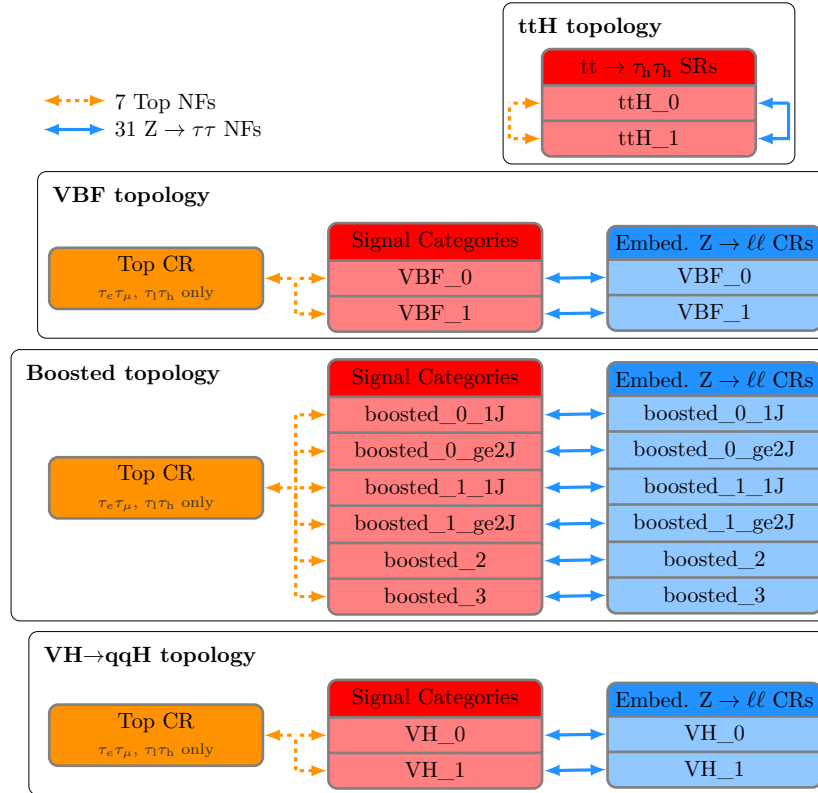


Figure 10.17: Schematic overview of NFs in the likelihood model. The red boxes correspond to signal categories, the orange boxes to top quark production control regions and the blue boxes to object-based embedded $Z \rightarrow \ell\ell$ control regions. Each arrow corresponds to one NF, where blue ones act only on $Z \rightarrow \tau\tau$ events in signal categories and on the object-based embedded $Z \rightarrow \ell\ell$ events in the control region, while orange ones act on events produced by top quark production. Arrows pointing to more than one signal category or control region illustrated that the NF is correlated across these regions [28].

model. Figure 10.17 illustrates the implementation of the NFs. Seven NFs for top quark

production background and 31 NFs for $Z \rightarrow \tau\tau$ events are implemented. Except for the $t\bar{t}H$ topology each NF has a dedicated control region. For NFs scaling the top quark production background one control region per inclusive signal category and per final state is defined, while for the $Z \rightarrow \tau\tau$ NFs, one object-based embedded control region per final signal category is defined. For the $t\bar{t}H$ categories no control regions are defined, instead the NFs are correlated across the two signal categories $t\bar{t}H_0$ and $t\bar{t}H_1$.

The analysis measures three different configurations of cross section measurements. Com-

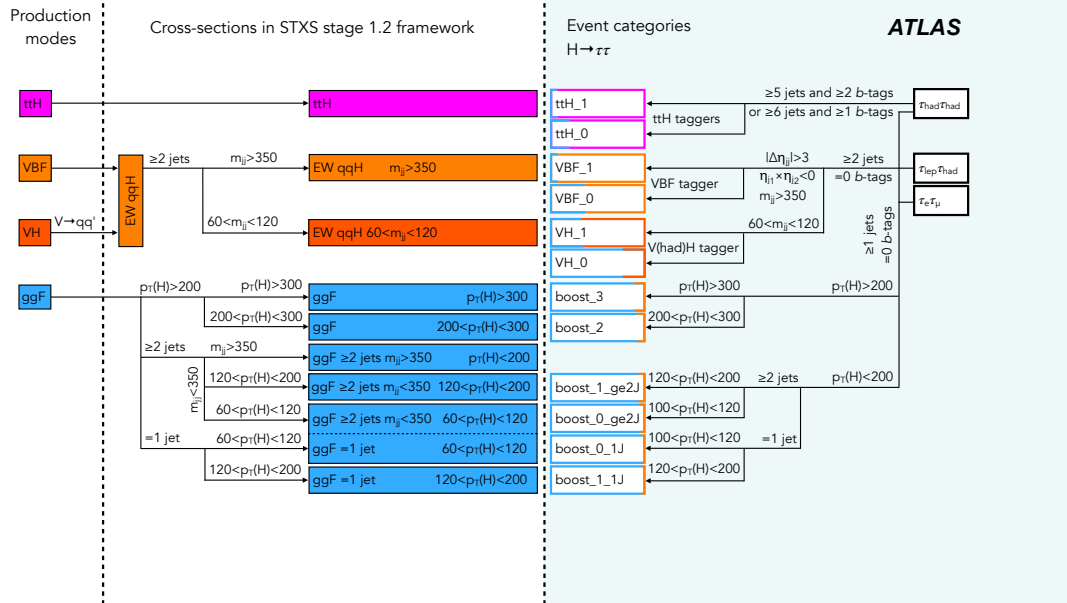


Figure 10.18: Schematic overview of PoIs in the four and nine PoI fit model. The left side of the sketch shows the PoI definition, while on the right side the dedicated signal categories for each PoI are shown. The colors surrounding the signal categories illustrate the relative contribution of the PoIs defined in the 4-PoI setup [28].

pared to the previous analysis the region of phase space with a rapidity of the Higgs boson with $|y| > 2.5$, called forward region, is excluded from the signal for all measurements. This region of phase space is not covered by the detector and therefore the measurements in $|y| \leq 2.5$ would be extrapolated to this regions resulting in extrapolation uncertainties. This results in a lower SM prediction for the inclusive measurement compared to the 2015-2016 analysis. The first setup measures the inclusive $H \rightarrow \tau\tau$ ($\sigma_H \times BR_{\tau\tau}$) cross section (1-PoI) assuming that the relative contributions of Higgs boson production modes behaves SM-like. For this setup, one PoI scales the summed contribution of all $H \rightarrow \tau\tau$ signal samples included in the likelihood model. The second setup measures inclusive cross sections for the four dominant production modes ggF , VBF , VH and $t\bar{t}H$ (4-PoI). One individual PoI for each of these production modes is defined, while event yields of the remaining production modes are set constant to the SM prediction, so that four PoIs are measured simultaneously. The third setup targets a

measurement of STXS bins with finer granularity (9-PoI). In this configuration one PoI for each of the STXS bins of the reduced Stage 1.2 configuration, as illustrated in Figure 10.12 by the red boxes, is defined with one exception. For the two lowest p_T^H STXS bins of ggF one common PoI is defined across the two jet multiplicity STXS bins. Additionally, one PoI scaling all $t\bar{t}H$ STXS bins is included. The event yields for the remaining STXS bins of the $EW\ qqH$ and ggF production as well as for the other production modes are set constant to the SM prediction. Nine PoI are defined in this setup: one $t\bar{t}H$ PoI, two $EW\ qqH$ PoIs and six ggF PoIs. Figure 10.18 summarize the PoIs for the 4-PoI and 9-PoI setup. On the left side, the definition of the PoIs is visualized, while on the right side, the signal categories dedicated to each PoI are shown. Each PoI in this setup has at least one dedicated signal category, except for the ggF , $N_j \geq 2$, $350\text{ GeV} \leq m_{jj}$, $p_T^H < 200\text{ GeV}$ STXS bins.

10.5.2 Input observable and likelihood function for the fit

In each of the signal categories the mass of the di- τ -system calculated with the missing mass calculator $m_{\tau\tau}^{\text{mmc}}$ is used as an input observable to the fit, to further increase the signal to background separation in the likelihood model. The binning is optimized to increase the signal to background separation for each signal category independently, depending on the expected data statistics and the background compositions. Figures 10.19-10.24 show the observed event yields in the $m_{\tau\tau}^{\text{mmc}}$ distribution for all signal categories. The binning of the shown distributions corresponds to the binning used in the likelihood model. The figures are produced using the final fit result of the 1-PoI configuration, in which the inclusive $H \rightarrow \tau\tau$ cross section is measured. Overall a good agreement between data and prediction within the uncertainties is observed. The figures show that the binning optimization chooses a finer binning for the region with large data statistics such as the boosted categories shown in Figure 10.20, 10.22 and 10.24, while a coarser binning is used in the remaining signal categories. However, Figures 10.19 (b,d,f), 10.21 (b,d) and 10.23 (b,d) show the high purity of signal contribution in the vh_1 , vbf_1 and tth_1 categories, so that a good signal to background separation is already achieved with a coarser binning compared to the boosted regions.

Similar to the previous analysis described in Chapter 9 only the expected yields are used in each control region.

In Tables 10.10-10.15 the observed event yields of the 1-PoI fit are compared to the recorded data for each signal category. In all signal categories combined 2201 $H \rightarrow \tau\tau$ and 141456 background events are observed. Therefore, a signal to background ratio of 1.6% is selected. The highest signal to background ratio is achieved in the VBF_1 categories with 28.0%, 27.6% and 33.7% for the $\tau_h\tau_h$, $\tau_l\tau_h$ and $\tau_e\tau_\mu$ final state, while the lowest ratio is observed in the low p_T^H boosted categories of the $\tau_h\tau_h$ final state with 0.8% and 0.7% for boosted_0_1J and boosted_0_ge1J categories. The categories with the largest data statistics are the intermediate boosted signal categories of the $\tau_h\tau_h$ and $\tau_l\tau_h$ final state with between 11886 and 14389 recorded data events. In all signal categories a good agreement between data and the post-fit prediction within their uncertainties is observed.

	VBF_0	VBF_1	VH_0	VH_1	ttH_0	ttH_1
$Z \rightarrow \tau\tau$	2051 ± 50	114.9 ± 10.8	4636 ± 84	539 ± 24	265 ± 24	20 ± 4
Fake	1027 ± 68	39.6 ± 5.3	1627 ± 110	112 ± 10	182 ± 17	6.5 ± 1.7
Top					239 ± 26	15.2 ± 3.4
Other backgrounds	57.1 ± 5.2	1.7 ± 0.6	209 ± 12	43.13 ± 2.93	16.4 ± 1.5	5.05 ± 0.49
$ggF, H \rightarrow \tau\tau$	38.5 ± 9.5	3.2 ± 1.8	72 ± 14	6.8 ± 1.7	1.96 ± 0.41	0.42 ± 0.08
$VBF, H \rightarrow \tau\tau$	72 ± 10	40 ± 5	8.1 ± 1.3	0.5 ± 0.1	0.23 ± 0.03	< 0.03
$WH, H \rightarrow \tau\tau$	1.00 ± 0.14	< 0.03	15.2 ± 2.5	9.8 ± 1.5	0.24 ± 0.03	< 0.03
$ZH, H \rightarrow \tau\tau$	0.8 ± 0.1	< 0.03	12.4 ± 2.3	5.4 ± 1.1	0.69 ± 0.15	0.15 ± 0.02
$\bar{t}tH, H \rightarrow \tau\tau$	0.19 ± 0.03	< 0.03	0.52 ± 0.07	0.22 ± 0.03	7.0 ± 1.6	5.4 ± 1.3
$tH, H \rightarrow \tau\tau$	0.41 ± 0.06	< 0.03	0.25 ± 0.03	0.07 ± 0.01	0.92 ± 0.13	0.41 ± 0.06
$b\bar{b}H, H \rightarrow \tau\tau$	0.10 ± 0.02	< 0.03	0.14 ± 0.02	< 0.03	0.27 ± 0.04	0.09 ± 0.02
Total background	3135 ± 84	156 ± 12	6472 ± 136	694 ± 26	703 ± 33	46.7 ± 5.3
Total signal	113 ± 15	43.6 ± 5.2	109 ± 16	23.0 ± 3.2	12 ± 2	6.6 ± 1.4
Total	3248 ± 84	200 ± 12	6580 ± 135	717 ± 26	715 ± 33	53.3 ± 5.5
Data	3318	197	6532	720	727	49

Table 10.10: Observed events yields in the VBF, VH and ttH signal categories of the $\tau_h \tau_h$ final state estimated in the 1-PoI setup. In the VBF and VH categories, the top production processes are estimated together with the other backgrounds (Di-boson, $Z \rightarrow ll$, $H \rightarrow$ others) in the fit. Uncertainties include statistical and systematic components [28].

	VBF_0	VBF_1	VH_0	VH_1
$Z \rightarrow \tau\tau$	2362 ± 59	162 ± 12	6724 ± 112	535 ± 23
Fake	611 ± 49	30 ± 3	1315 ± 126	80 ± 8
Top	107 ± 12	5.3 ± 1.5	243 ± 25	27 ± 5
Other backgrounds	139 ± 17	5.8 ± 2.4	396 ± 39	50 ± 4
$ggF, H \rightarrow \tau\tau$	71 ± 28	3.5 ± 1.2	87.3 ± 20.3	5.1 ± 2.2
$VBF, H \rightarrow \tau\tau$	84.4 ± 11.2	52 ± 6	9.3 ± 1.6	0.5 ± 0.2
$WH, H \rightarrow \tau\tau$	0.83 ± 0.11	< 0.03	17.4 ± 2.6	8.1 ± 1.2
$ZH, H \rightarrow \tau\tau$	0.86 ± 0.12	< 0.03	13.3 ± 2.5	5.0 ± 0.9
$\bar{t}tH, H \rightarrow \tau\tau$	0.10 ± 0.01	< 0.03	0.35 ± 0.05	0.13 ± 0.02
$tH, H \rightarrow \tau\tau$	0.26 ± 0.04	< 0.03	0.17 ± 0.03	< 0.03
$b\bar{b}H, H \rightarrow \tau\tau$	0.09 ± 0.01	< 0.03	0.16 ± 0.02	< 0.03
Total background	3219 ± 75	203 ± 13	8677 ± 143	692 ± 24
Total signal	158 ± 30	56 ± 7	129 ± 23	19 ± 3
Total	3377 ± 76	259 ± 13	8806 ± 143	711 ± 24
Data	3402	267	8780	743

Table 10.11: Observed events yields in the VBF and VH signal categories of the $\tau_l \tau_h$ final state estimated in the 1-PoI setup. Uncertainties include statistical and systematic components [28].

	VBF_0	VBF_1	VH_0	VH_1
$Z \rightarrow \tau\tau$	820 ± 29	49.3 ± 6.5	2424 ± 58	186 ± 13
Fake	90 ± 21	3.3 ± 5.3	214 ± 42	33 ± 14
Top	165 ± 15	9 ± 2	346 ± 29	33 ± 5
Other backgrounds	96.1 ± 8.5	11.9 ± 1.6	259 ± 23	28 ± 2
$ggF, H \rightarrow \tau\tau$	12.7 ± 3.2	1.04 ± 0.31	26 ± 5	1.8 ± 0.6
$VBF, H \rightarrow \tau\tau$	23 ± 3	14.3 ± 1.8	2.41 ± 0.43	0.13 ± 0.02
$WH, H \rightarrow \tau\tau$	0.21 ± 0.03	< 0.03	4.3 ± 0.7	2.45 ± 0.42
$ZH, H \rightarrow \tau\tau$	0.14 ± 0.02	< 0.03	3.6 ± 0.8	1.5 ± 0.3
$\bar{t}tH, H \rightarrow \tau\tau$	< 0.03	< 0.03	0.08 ± 0.01	< 0.03
$tH, H \rightarrow \tau\tau$	0.13 ± 0.02	< 0.03	0.04 ± 0.01	< 0.03
$b\bar{b}H, H \rightarrow \tau\tau$	< 0.03	< 0.03	0.05 ± 0.01	< 0.03
Total background	1154 ± 35	66.4 ± 8.5	3230 ± 63	280 ± 18
Total signal	51.9 ± 5.4	22.4 ± 1.8	50 ± 7	7.1 ± 0.9
Total	1206 ± 35	88.8 ± 8.6	3280 ± 63	287 ± 18
Data	1215	98	3277	286

Table 10.12: Observed events yields in the VBF and VH signal categories of the $\tau_e \tau_\mu$ final state estimated in the 1-PoI setup. Uncertainties include statistical and systematic components [28].

10.5 Statistical model

	Boosted_0_1J	Boosted_0_ge2J	Boosted_1_1J	Boosted_1_ge2J	Boosted_2_ge1J	Boosted_3_ge1J
$Z \rightarrow \tau\tau$	5635 ± 115	2640 ± 67	11863 ± 134	10076 ± 125	7252 ± 93	973 ± 30
Fake	3388 ± 224	1729 ± 118	2312 ± 155	2072 ± 140	293 ± 32	54 ± 20
Other backgrounds	61 ± 9	74.2 ± 11.3	116 ± 19	251 ± 14	157.3 ± 10.5	53.6 ± 5.5
$ggF, H \rightarrow \tau\tau$	54.4 ± 9.7	23.1 ± 4.1	112.8 ± 20.5	109.1 ± 20.6	96.2 ± 17.2	30 ± 7
$VBF, H \rightarrow \tau\tau$	11.3 ± 2.0	5.8 ± 0.9	27.6 ± 4.7	24.6 ± 4.2	23.7 ± 3.6	7.3 ± 1.1
$WH, H \rightarrow \tau\tau$	2.1 ± 0.6	1.5 ± 0.3	3.8 ± 1.1	7.0 ± 1.1	4.6 ± 0.7	2.5 ± 0.7
$ZH, H \rightarrow \tau\tau$	1.4 ± 0.3	1.1 ± 0.3	2.7 ± 0.9	5.3 ± 1.0	3.7 ± 0.5	1.5 ± 0.2
$i\bar{i}H, H \rightarrow \tau\tau$	< 0.03	0.27 ± 0.04	< 0.03	1.01 ± 0.14	0.8 ± 0.1	0.35 ± 0.05
$t\bar{t}H, H \rightarrow \tau\tau$	< 0.03	0.06 ± 0.01	< 0.03	0.30 ± 0.04	0.39 ± 0.05	0.09 ± 0.01
$b\bar{b}H, H \rightarrow \tau\tau$	0.19 ± 0.03	0.07 ± 0.01	0.21 ± 0.03	0.30 ± 0.04	0.21 ± 0.03	0.05 ± 0.01
Total background	9084 ± 244	4444 ± 132	14291 ± 199	12398 ± 188	7702 ± 96	1080 ± 32
Total signal	69.5 ± 11.0	32 ± 5	147 ± 23	148 ± 22	130 ± 18	41.7 ± 7.2
Total	9153.5 ± 243.5	4476 ± 132	14438 ± 198	12546 ± 187	7832 ± 95	1122 ± 32
Data	9163	4503	14389	12585	7800	1124

Table 10.13: Observed events yields in the boosted signal categories of the $\tau_h \tau_h$ final state estimated in the 1-PoI setup. Uncertainties include statistical and systematic components [28].

	Boosted_0_1J	Boosted_0_ge2J	Boosted_1_1J	Boosted_1_ge2J	Boosted_2_ge1J	Boosted_3_ge1J
$Z \rightarrow \tau\tau$	5583 ± 100	3228 ± 69	10927 ± 130	9546 ± 118	7195 ± 92	2413 ± 46
Fake	1119 ± 58	832 ± 44	1324 ± 76	1434 ± 81	536 ± 35	139 ± 12
Top	65 ± 7	123 ± 16	81 ± 13	317 ± 25	129 ± 12	52 ± 8
Other backgrounds	214 ± 40	177 ± 22	374 ± 40	447 ± 33	300 ± 17	164 ± 6
$ggF, H \rightarrow \tau\tau$	45.4 ± 11.6	44.8 ± 15.2	99.5 ± 19.7	$\pm 123.5 \pm 29.4$	91.2 ± 24.4	33.5 ± 8.9
$VBF, H \rightarrow \tau\tau$	12.1 ± 2.0	7.2 ± 1.1	26.9 ± 4.1	23.6 ± 3.7	21.6 ± 3.2	8.6 ± 1.4
$WH, H \rightarrow \tau\tau$	1.7 ± 0.3	2.2 ± 0.6	3.5 ± 2.0	6.5 ± 1.2	4.4 ± 0.8	3.3 ± 1.0
$ZH, H \rightarrow \tau\tau$	1.2 ± 0.3	1.3 ± 0.2	2.4 ± 0.3	4.7 ± 0.6	3.0 ± 0.4	1.8 ± 0.2
$i\bar{i}H, H \rightarrow \tau\tau$	< 0.03	0.11 ± 0.01	< 0.03	0.55 ± 0.08	0.35 ± 0.05	0.22 ± 0.03
$t\bar{t}H, H \rightarrow \tau\tau$	< 0.03	0.09 ± 0.01	< 0.03	0.28 ± 0.04	0.17 ± 0.02	< 0.03
$b\bar{b}H, H \rightarrow \tau\tau$	0.13 ± 0.02	0.13 ± 0.02	0.17 ± 0.02	0.21 ± 0.03	0.12 ± 0.02	0.04 ± 0.01
Total background	6981 ± 112	4360 ± 82	12706 ± 143	11743 ± 135	8160 ± 94	2768 ± 49
Total signal	61 ± 13	56.1 ± 15.5	133 ± 22	160 ± 31	121.6 ± 25.1	47.8 ± 9.2
Total	7042 ± 112	4416 ± 81	12839 ± 143	11903 ± 134	8281 ± 93	2816 ± 49
Data	7094	4374	12779	11886	8236	2848

Table 10.14: Observed events yields in the boosted signal categories of the $\tau_l \tau_h$ final state estimated in the 1-PoI setup. Uncertainties include statistical and systematic components [28].

	Boosted_0_1J	Boosted_0_ge2J	Boosted_1_1J	Boosted_1_ge2J	Boosted_2_ge1J	Boosted_3_ge1J
$Z \rightarrow \tau\tau$	2642 ± 64	1523 ± 42	3913 ± 69	3453 ± 68	1734 ± 37	469 ± 20
Fake	101 ± 31	85 ± 23	117 ± 36	179 ± 42	88 ± 24	37 ± 13
Top	101 ± 8	187 ± 20	157 ± 11	569 ± 44	258 ± 16	104 ± 11
Other backgrounds	118 ± 17	101 ± 16	273 ± 14	325 ± 30	294 ± 8	173 ± 5
$ggF, H \rightarrow \tau\tau$	16.6 ± 3.2	11 ± 2	35 ± 7	36.8 ± 7.1	25.5 ± 4.6	8.7 ± 2.2
$VBF, H \rightarrow \tau\tau$	3.4 ± 0.6	2.2 ± 0.3	8.7 ± 1.6	7.8 ± 1.1	6.2 ± 0.9	2.1 ± 0.3
$WH, H \rightarrow \tau\tau$	0.44 ± 0.06	0.57 ± 0.13	1.35 ± 0.56	2.22 ± 0.57	1.29 ± 0.18	0.87 ± 0.35
$ZH, H \rightarrow \tau\tau$	0.29 ± 0.04	0.33 ± 0.05	0.73 ± 0.10	1.57 ± 0.22	0.85 ± 0.12	0.41 ± 0.06
$i\bar{i}H, H \rightarrow \tau\tau$	< 0.03	< 0.03	< 0.03	0.20 ± 0.03	0.08 ± 0.01	0.07 ± 0.01
$t\bar{t}H, H \rightarrow \tau\tau$	< 0.03	0.08 ± 0.01	< 0.03	0.13 ± 0.02	0.11 ± 0.02	< 0.03
$b\bar{b}H, H \rightarrow \tau\tau$	< 0.03	< 0.03	0.06 ± 0.01	0.11 ± 0.01	0.06 ± 0.01	< 0.03
Total background	2952 ± 65	1890 ± 47	4439 ± 70	4506 ± 75	2357 ± 43	777 ± 25
Total signal	30 ± 4	20 ± 3	65 ± 9	69 ± 10	51 ± 7	18 ± 3
Total	2982 ± 65	1910 ± 47	4504 ± 70	4575 ± 75	2407 ± 42	795 ± 25
Data	2973	1877	4458	4594	2325	743

Table 10.15: Observed events yields in the boosted signal categories of the $\tau_e \tau_\mu$ final state estimated in the 1-PoI setup. Uncertainties include statistical and systematic components [28].

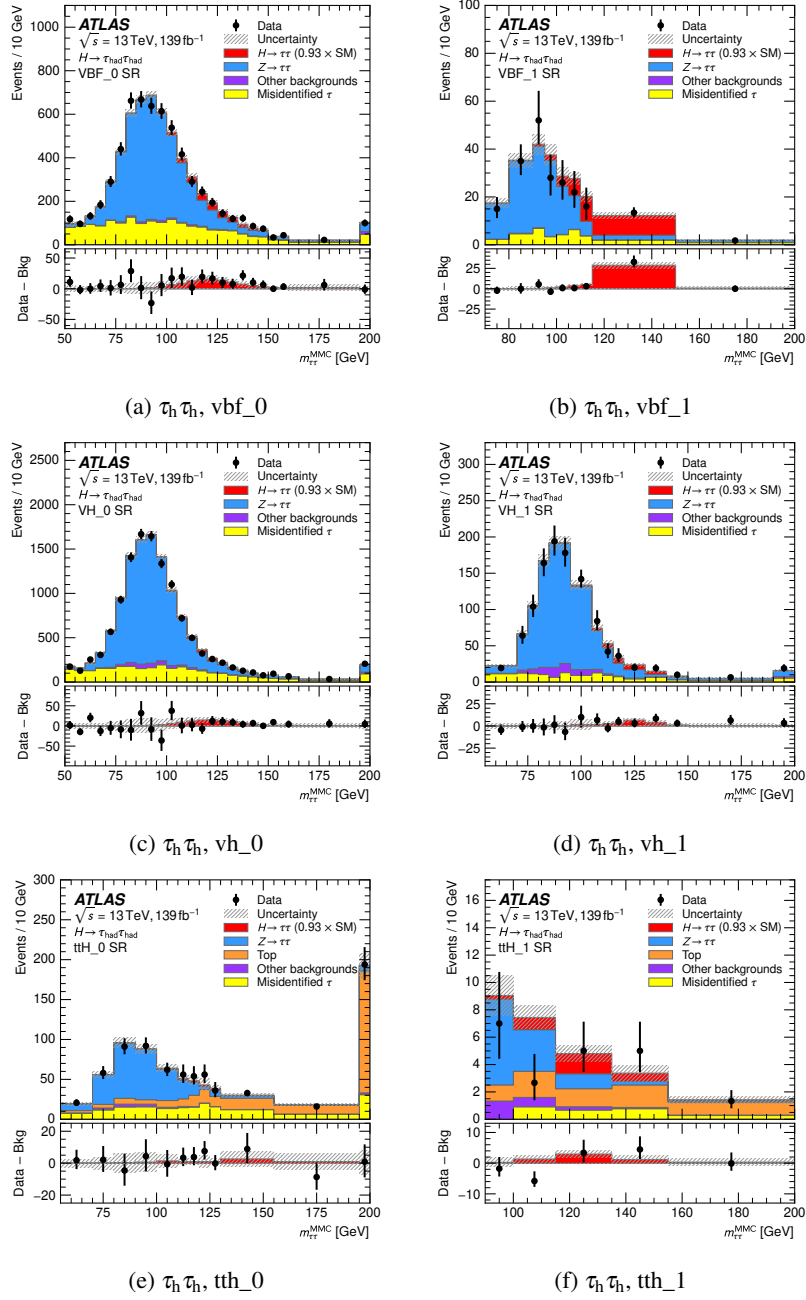


Figure 10.19: Post-fit event yield for the $\tau_h \tau_h$, vbf (top row), vh (middle row) and ttH (bottom row) signal categories estimated with the inclusive $H \rightarrow \tau\tau$ PoI setup as a function of $m_{\tau\tau}^{\text{MMC}}$ with the same binning as used in the likelihood model [28].

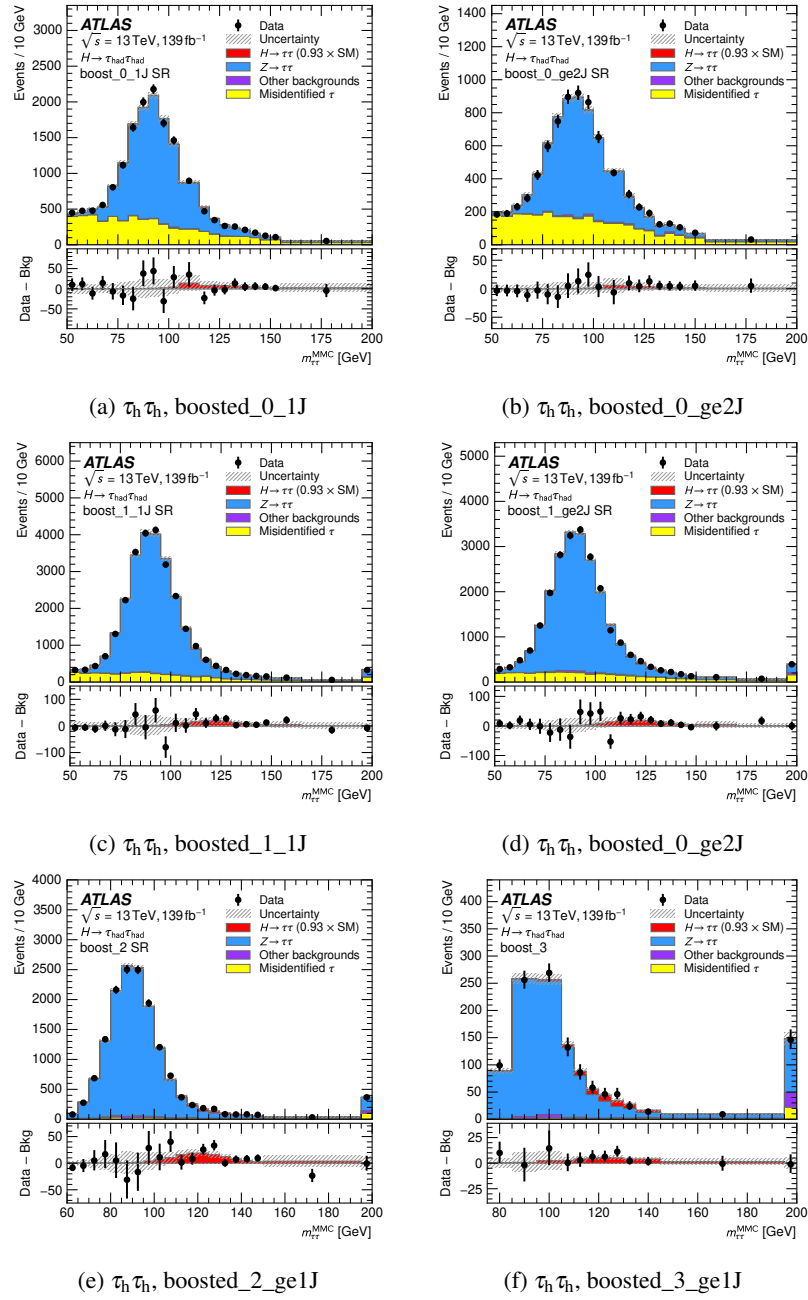


Figure 10.20: Post-fit event yield for the $\tau_h \tau_h$, boosted signal categories estimated with the inclusive $H \rightarrow \tau\tau$ PoI setup as a function of $m_{\tau\tau}^{\text{MMC}}$ with the same binning as used in the likelihood model [28].

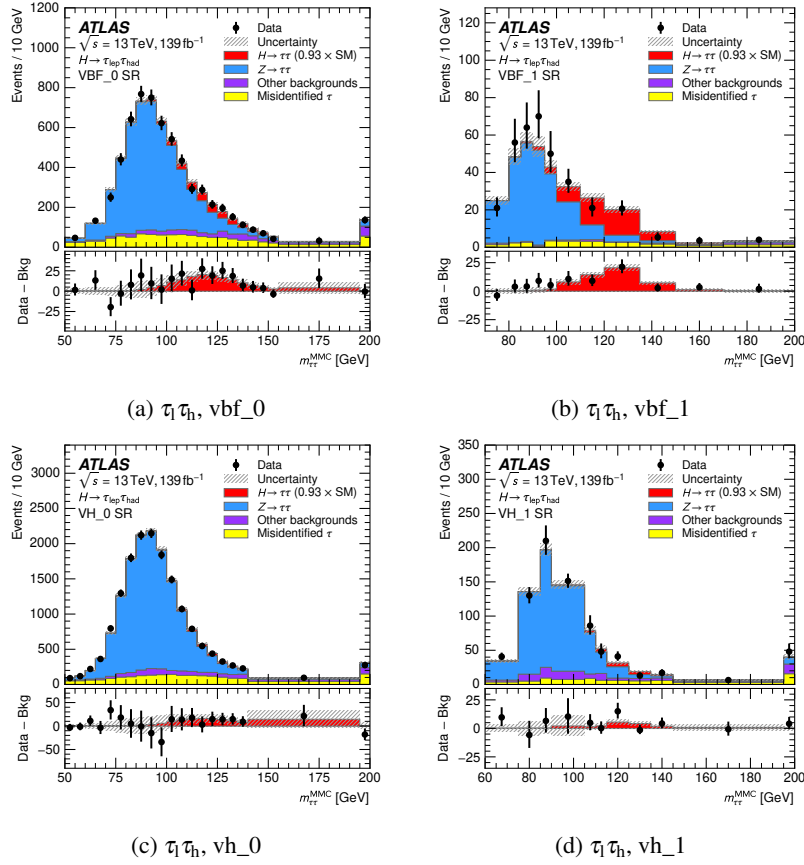


Figure 10.21: Post-fit event yield for the $\tau_1 \tau_h$, vbf (top row) and vh (bottom row) signal categories estimated with the inclusive $H \rightarrow \tau\tau$ PoI setup as a function of $m_{\tau\tau}^{\text{MMC}}$ with the same binning as used in the likelihood model [28].

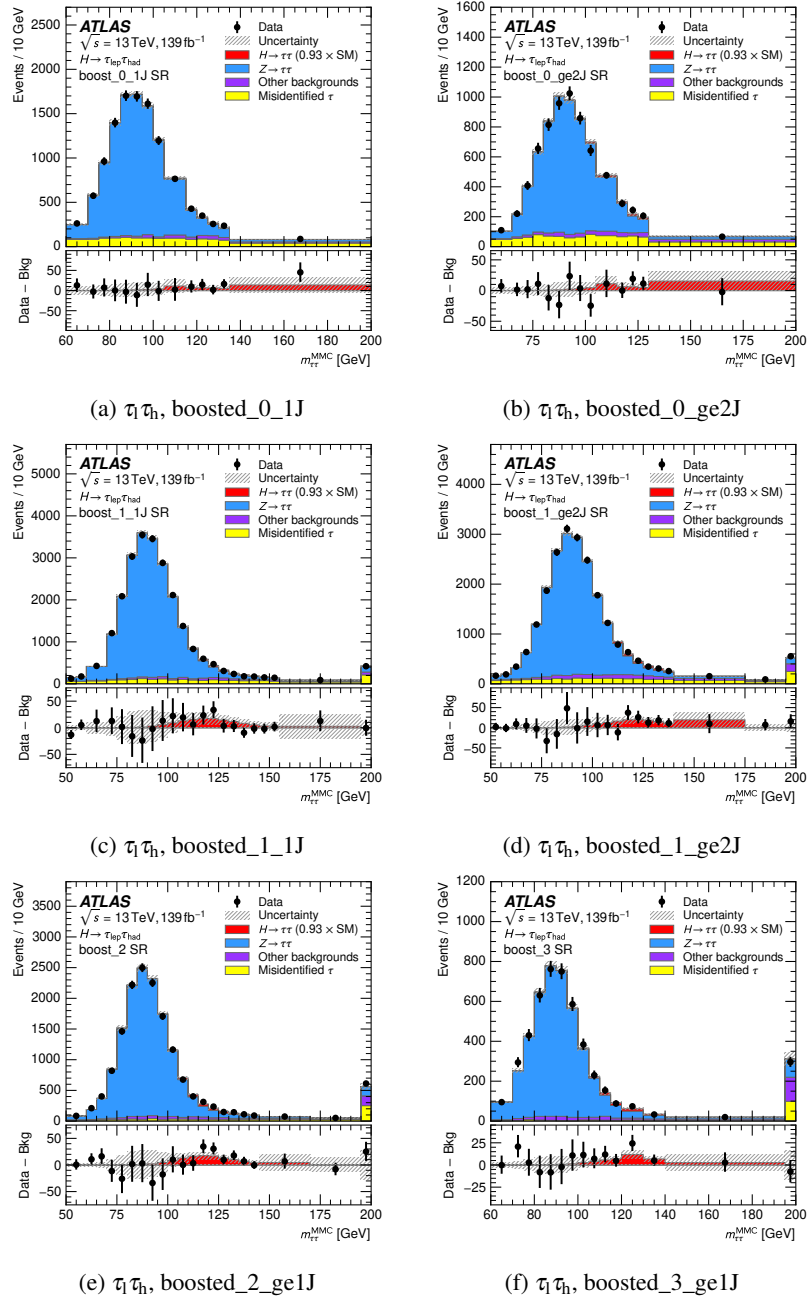


Figure 10.22: Post-fit event yield for the $\tau_1 \tau_h$, boosted signal categories estimated with the inclusive $H \rightarrow \tau\tau$ PoI setup as a function of $m_{\tau\tau}^{\text{MMC}}$ with the same binning as used in the likelihood model [28].

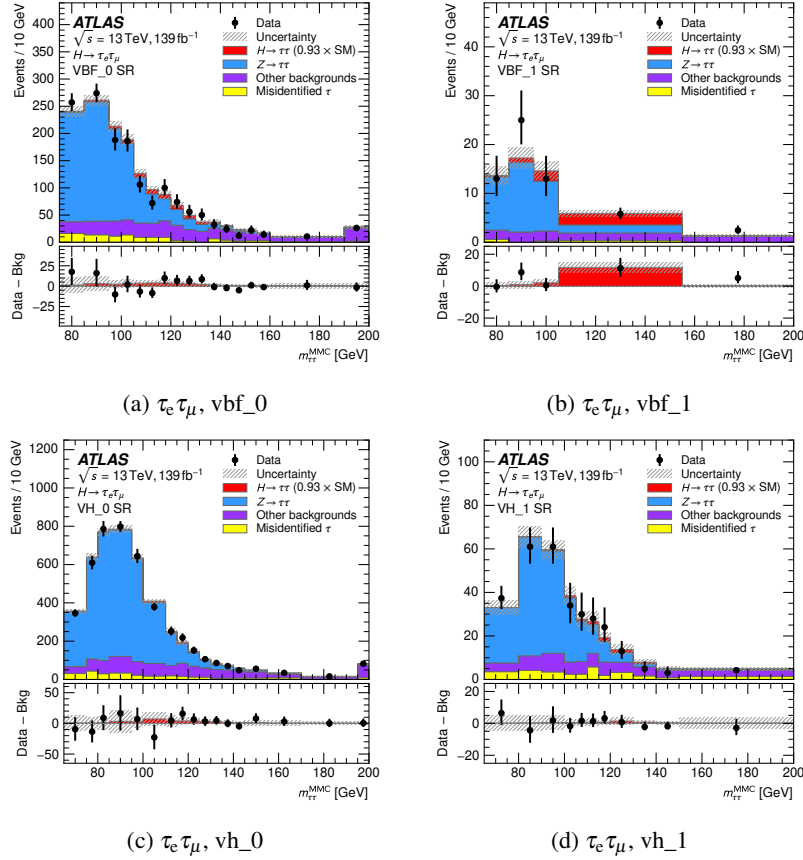


Figure 10.23: Post-fit event yield for the $\tau_c \tau_\mu$, vbf (top row) and vh (bottom row) signal categories estimated with the inclusive $H \rightarrow \tau\tau$ PoI setup as a function of $m_{\tau\tau}^{\text{MMC}}$ with the same binning as used in the likelihood model [28].

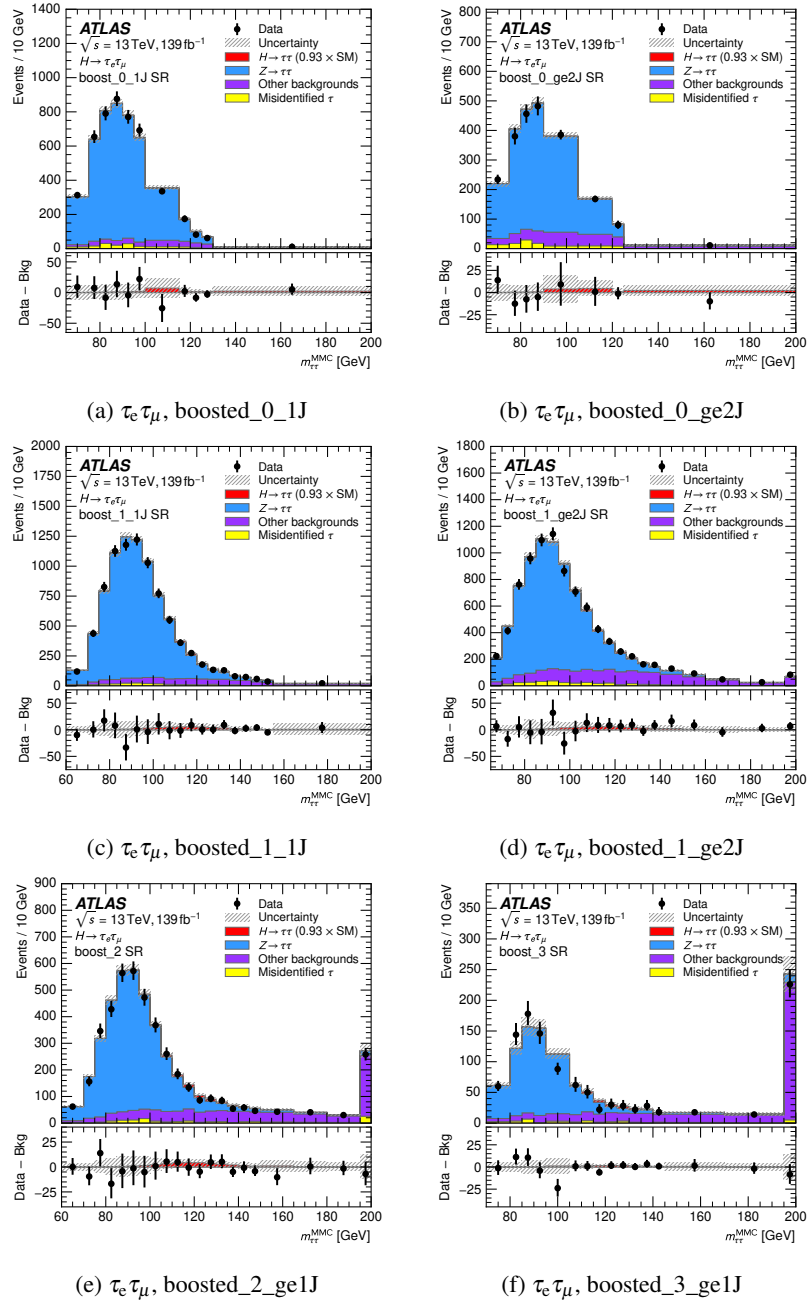


Figure 10.24: Post-fit event yield for the $\tau_e \tau_\mu$, boosted signal categories estimated with the inclusive $H \rightarrow \tau\tau$ PoI setup as a function of $m_{\tau\tau}^{\text{MMC}}$ with the same binning as used in the likelihood model [28].

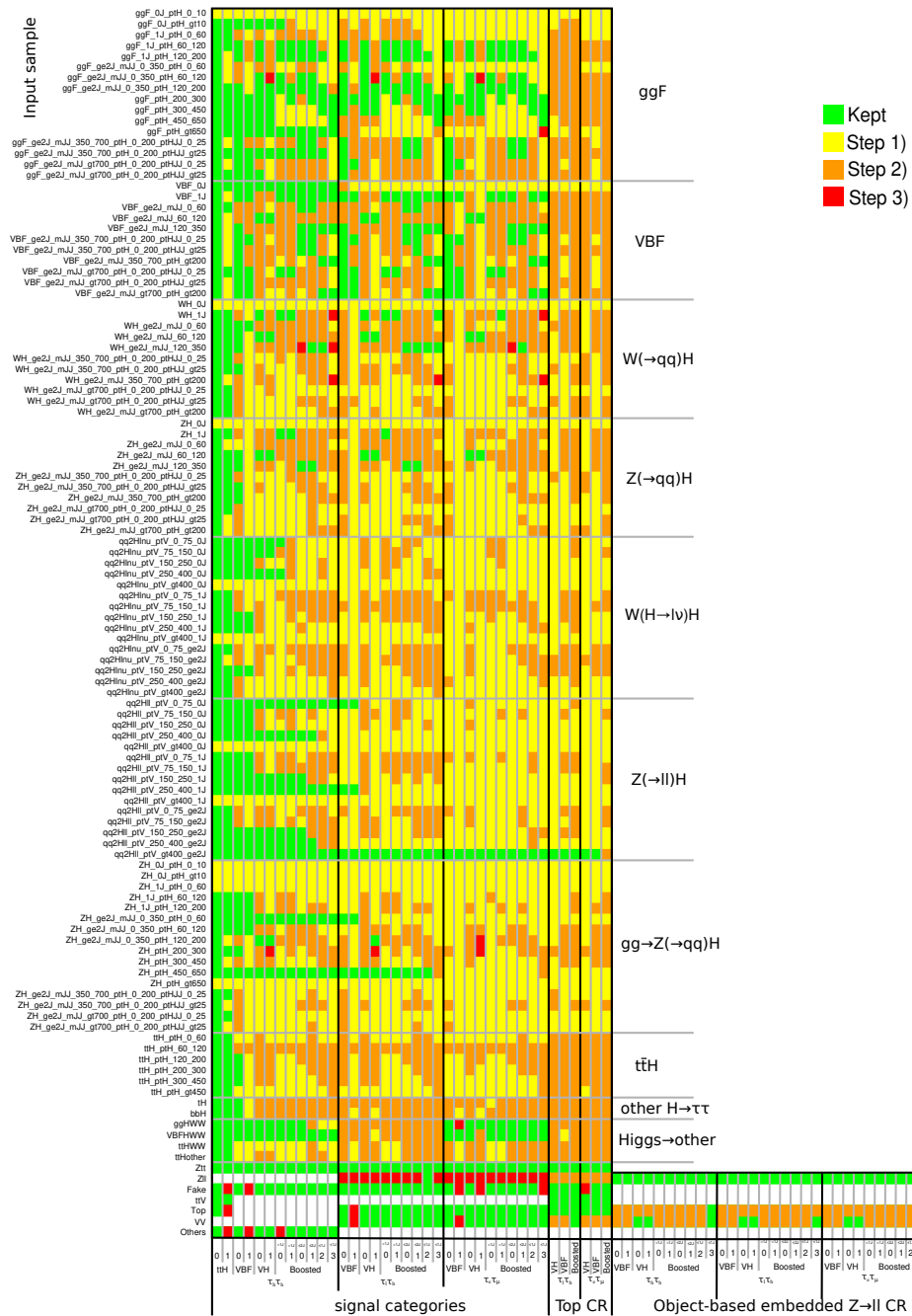
10.5.3 Modifications of the likelihood function

The complexity of the analysis caused by the large number of signal categories, control regions, NFs and NPs introduces several challenges for the fit model. Furthermore, the implementation of the STXS Stage 1.2 makes it necessary to treat the signal prediction for each STXS bin at this stage as an individual sample. The shape effects of many NPs in the likelihood model suffer from the relatively small sample sizes and the resulting large statistical uncertainty. Such that in these cases small systematic variations may in fact be dominated by statistical noise instead of real variations introduced by the systematic uncertainty itself. Such noise in the likelihood model causes instabilities and increases the impact of some uncertainties artificially. A lot of effort has been invested to develop methods to systematically remove and correct the shape systematic uncertainties before building the fit model. These methods suppress the noise introduced by statistical fluctuations, without accidentally removing genuine and significant shape variations. The algorithms can be split to three stages: 1) Pre-pruning, 2) Symmetrization and Smoothing and 3) Pruning.

Pre-pruning This step is implemented to remove nominal predictions and/or systematic uncertainties of samples with negligible contributions. On one hand, this decreases the rather long running time of the fit, but on the other hand, it also removes systematic uncertainties with high statistical fluctuations which cause instabilities in the following steps. The pre-pruning has three steps:

1. Remove systematic uncertainties and the nominal prediction for negligible samples
2. Remove systematic uncertainties for non-negligible but very small samples
3. Remove systematic uncertainties for samples dominated by the statistical uncertainty of the simulated sample size

All three steps are applied to each sample in all signal categories and control regions separately. In step 1) all samples are removed from a signal category or control region if the sample has an expected event yield of less than 0.01 events. In step 2) different requirements for signal categories and control regions are used. In signal categories the uncertainties on all samples with an expected event yield of less than 1% relative to the expected signal yield are removed. In control regions, the expected event yield has to be less than 1% relative to the sum of all events in the region. In step 3) the relative statistical uncertainty caused by the limited sample size is used. If this uncertainty exceeds 10% of the yield of a sample all other systematic uncertainties on this sample are removed in this region. Figure 10.25 summarizes the removed samples and systematic uncertainties in the pre-pruning step. For all samples, except for the $Z \rightarrow \tau\tau$ contribution, uncertainties are removed in a sub-set of categories and regions. Most of signal samples are removed in step 1) or the systematic uncertainties on them are removed in step 2). Many of STXS bins are not expected to contribute to a measured phase space or are expected to contribute only to a sub-set of signal categories so that they do not fulfill the requirements of these steps. The limited amount of $Z \rightarrow ll$ events causes their systematic uncertainties to be removed in step 3) in several categories and regions. For most of the other



samples no systematic uncertainties are removed in the last step, if they already passed step 1) and 2) without failing the requirements of these steps. The sum of yields, which are removed from the fit by failing step 1), is less than 1 event per signal category or control region and therefore negligible.

Symmetrization and Smoothing After pre-pruning, each systematic uncertainty is symmetrized and smoothed per input sample. These steps are implemented to remove fluctuations in the shape variation of the systematic uncertainties caused by small sample sizes. Especially, bin entries of systematic variations with negative events, bins with an up and down variation pointing into the same direction and large spikes in the entries of the observable used in the fit have to be corrected to stabilize the likelihood fit. Negative bin entries are removed by setting the bin value to 10^{-6} events. To reduce the fluctuation in the shape of the observable used in the fit model the smoothing algorithm "353QH twice" of ROOT is used, which is described in [256–258]. Therefore, the algorithm starts by studying the first bin entry of the observable used in the fit. It estimates if the systematic uncertainty in this bin is statistically significant. If so, it keeps the original systematic uncertainty in this bin and the algorithm starts again with the neighboring bin entry of the observable. If the systematic uncertainty is not significant, the bin is combined with the neighboring bin by summing the bin entries of the nominal prediction and separately summing the bin entries of the systematic variations. The systematic uncertainty of this combined bin is tested against the nominal prediction. The systematic uncertainties of both bins are set to the size of the combined systematic uncertainty if it is significant and the algorithm moves on. If not, the next bin is added as well until the requirement is fulfilled or all bins are combined and therefore the full shape variation of the systematic uncertainty is removed by setting the systematic uncertainty of all bin entries to the same relative variation. This algorithm is optimized for relatively flat distributions and can therefore not be applied to the $m_{\tau\tau}^{\text{mmc}}$ distribution. However, the relative variations of the systematic uncertainties is expected to be rather flat and is used instead. Fluctuations due to small sample sizes can introduce occasional large spikes in the relative variations which cannot be handled by the algorithm. Furthermore, ROOT applies the smoothing from left to right to the bin entries, which is optimized for observables with decreasing population for higher values, but the $m_{\tau\tau}^{\text{mmc}}$ has the highest expected events around the mass of the Z boson. To improve the algorithm for this analysis a local smoothing is applied first. In this, all bins with an absolute relative variation of more than 50% and bins for which the relative variation has a opposite sign compared to the variations of both neighboring bins are smoothed locally with their neighboring bins to reduce the number of spikes and sign changes. In a next step, two neighboring bins are smoothed if the relative variation between these bins differs by more than 20%. To ensure that this smoothing is not introducing new large steps, it is applied iteratively. After the local smoothing a symmetrization and correction procedure is applied on the systematic variations to correct bins with a relative up and down variation pointing in the same direction. In each of these bins, both variations are set to the absolute average of the two variations. For the more significant variation the sign is kept, while the smaller variation is set to the opposite sign. This correction by hand can introduce spikes in the relative variation so that the smoothing algorithm of ROOT is applied again after this correction has been applied.

The VH and ttH categories especially suffer from small sample sizes and are dominated by data statistics and the limited sample size of the simulation. Therefore, these regions have higher fluctuations and the systematic uncertainties are less important than in other categories. All systematic uncertainties are fully symmetrized in these regions by centering the relative up and down variation in every bin around zero. The same approach is applied to energy scale uncertainties of the τ -lepton and to jet related uncertainties with relative variation on the event yield below 1.2%. The impact of normalization uncertainties from these systematic uncertainties on the PoIs is expected to be small, however their shape uncertainty can still have a larger impact on the fit. The full symmetrization of these systematic uncertainties ensures that the impact is kept, while statistical fluctuations are reduced. All of these steps have an impact on the shape but also on the normalization of the variations. However, the statistical uncertainty on the normalization is much smaller than on the individual bin variations which are used for the shape variations. To avoid that these manipulations of the shape variations have a large impact on the normalization uncertainties they are reset to original values if the shape variation is pruned away. Therefore, the manipulated variations are only kept for uncertainties which have a non-negligible shape component.

Pruning After the symmetrization and smoothing procedure explained in the previous paragraph, the shape and/or normalization component of each NP or the statistical uncertainty on the background prediction in one bin of the $m_{\tau\tau}^{\text{mmc}}$ distribution is removed from the fit model in a signal category or control region of the fit according to the following rules:

- Bin-by-bin statistical uncertainties of simulations
 - The statistical uncertainty on the background prediction in one bin is removed, if the relative statistical uncertainty on this bin is smaller than 1%
- Normalization component of NP
 - The normalization component is removed, if relative systematic uncertainty of one NP on the event yield is smaller than 1%
- Shape component of NP
 - The shape component of one NP is removed, if the relative difference between the systematic uncertainties across all bins is less than 1%
 - The shape component of one NP is removed, if the difference between the systematic uncertainties across all bins of the $m_{\tau\tau}^{\text{mmc}}$ distribution binned into [0,90,110,140, ∞ GeV] agree within the one- σ interval of their statistical uncertainty

Examples of systematic uncertainties Figure 10.26 shows the original distribution and the manipulated distribution after applying the smoothing and symmetrization method for six examples of systematic uncertainties. The examples are chosen to show extreme cases of input variations to validate that these cases are corrected in a reasonable way. Figure 10.26 (a) shows

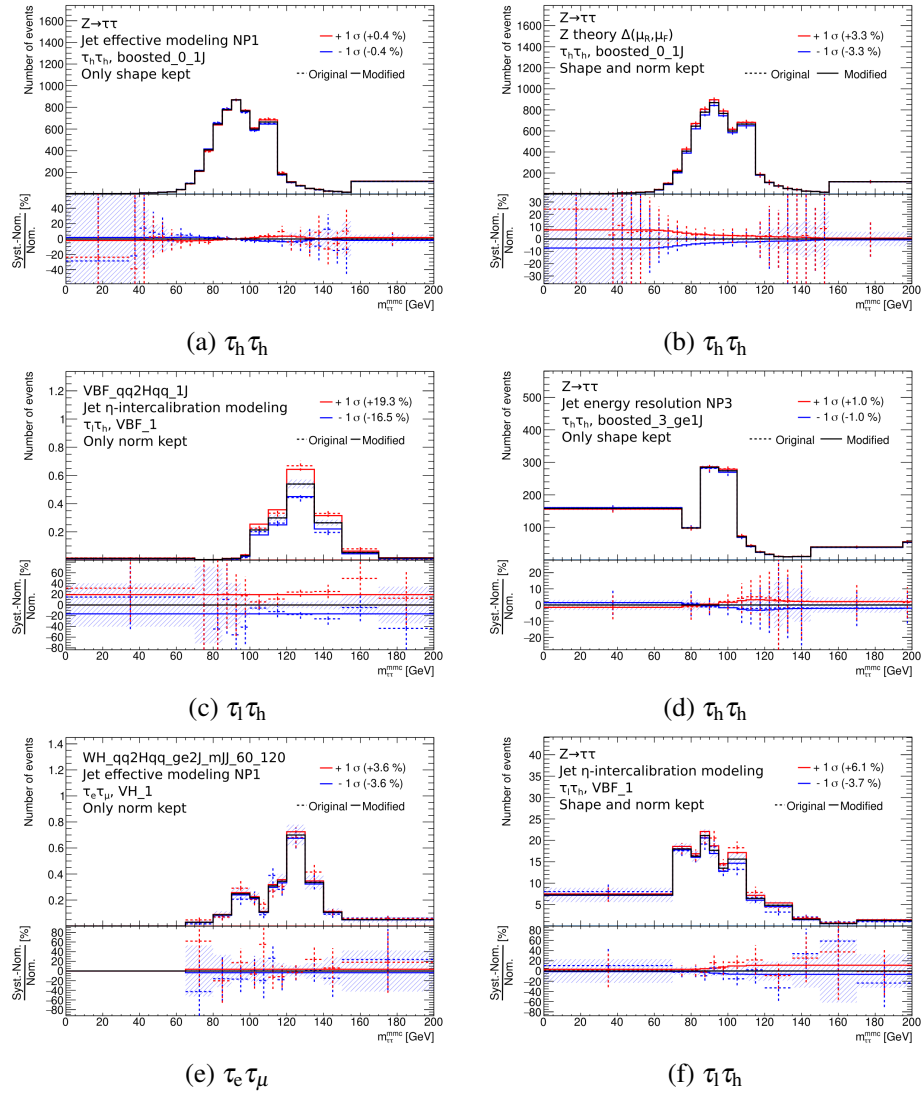


Figure 10.26: Examples of envelopes for the $Z \rightarrow \tau\tau$ (a,b,d,f), the VBF_qq2Hqq_1J (c) and the WH_qq2Hqq_ge2J_mJJ_60_120 (e) input samples in the $\tau_h \tau_h$ boosted_0_1J (a,b), the $\tau_l \tau_h$ vbf_1 (c,f), the $\tau_h \tau_h$ boosted_3_ge1J (d) and the $\tau_e \tau_\mu$ vh_1 (e) signal categories. (b) shows the envelope for an Z theory NP and the remaining ones show jet related NPs. Each legend defines if the shape and/or normalization (norm) component of the systematic uncertainties is kept after the pruning. The dashed lines corresponds to the systematic uncertainty before smoothing and symmetrization, while the solid line shows the systematics after these steps. Black corresponds to the nominal prediction, red to the up variation and blue to the down variation.

an example in which a reasonable shape systematic is calculated even with sign flips in the original histogram introduced by statistical fluctuations. Figures 10.26 (b) and (d) show cases without large statistical fluctuations in which the algorithm keeps most of the original shape. Figure 10.26 (c) gives an example for a systematic uncertainty, which has a shape introduced by a statistical fluctuation, which is removed by the algorithm, while the asymmetry of the normalization uncertainty is kept. Figure 10.26 (e) shows an example for a systematic uncertainty in the low data statistics region VH_1. Several of the bins have an up and down variation pointing in the same direction introduced by statistical fluctuations. Additionally, the up and down variations have several sign flips between neighbor bins. The algorithm is able to calculate a reasonable normalization uncertainty using the full symmetrization approach. Figure 10.26 (f) gives an example of a systematic variation with large statistical fluctuation for large values of $m_{\tau\tau}^{\text{mmc}}$, while for lower values only small fluctuation are given. The algorithms is able to produce a reasonable shape and normalization uncertainty.

10.6 Results

This section summarizes the results of the cross section measurements using the 1-PoI, 4-PoI and the 9-PoI fit setup. In Section 10.6.1 the observed values for the NFs to correct the background predictions are shown and their compatibility across the different fit setups is studied. The behavior of all NPs in the likelihood model is shown and compared between the different fit setups in Section 10.6.2. The extracted inclusive cross sections using the 1-PoI and 4-PoI fit setups and the impact of systematic and statistical uncertainties on these results are summarized in Section 10.6.3. Finally, the cross section measurements of the 9-PoI fit based on the reduced STXS Stage 1.2 configuration (see Figure 10.12) and the impact of the systematic and statistical uncertainties on them is discussed in Section 10.6.4.

10.6.1 Observed normalization factors

A large number of NFs are included in the likelihood model to estimate the normalization of several background contributions. This section summarizes the results for the NFs from all three fit setups (1-, 4- and 9-PoI) and compares them to study the stability of the background estimation across the different likelihood models.

Figure 10.27 shows the best-fit values and uncertainties for all NFs applied to the top quark production background. No significant deviation from the SM prediction is visible for them. The largest correction is applied to the top quark production background in the ttH categories, in which the contribution is increased by 23%. The uncertainties of all NFs are dominated by systematic uncertainties. A good agreement between the estimated NFs between the different fit setups is observed. This confirms the stability of the estimation of the top quark production background normalization across the different PoI configurations.

Figure 10.29 and 10.28 compares the best-fit values of the NFs to normalize the $Z \rightarrow \tau\tau$ background from the three different PoI configurations. One NF per signal category and their corresponding object-based embedded control region is defined, except for the ttH categories, in which one common NFs is defined for both categories and no control region is included. The NFs show no significant deviation from the SM prediction of one. The ttH

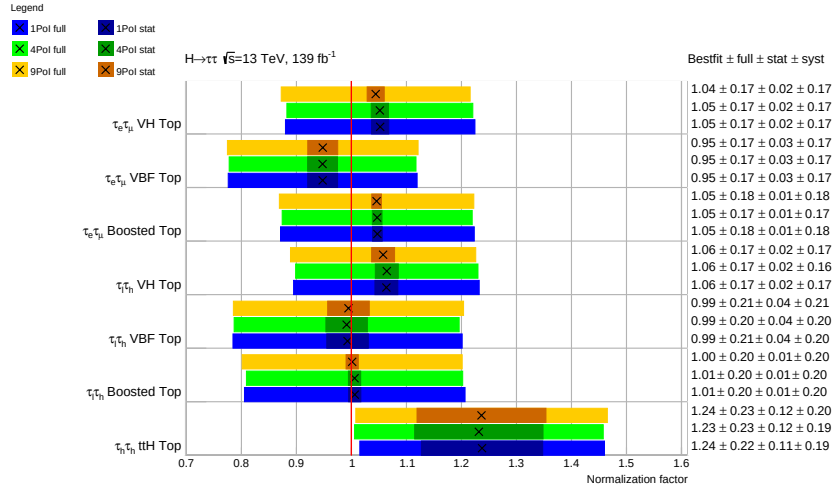


Figure 10.27: Measured NFs to normalize the top quark production background for the 1-PoI (blue), 4-PoI (green) and 9-PoI (orange) setup. The crosses mark the best-fit values, while the colored boxes show their uncertainties. The lighter color illustrates the full uncertainty, while the darker corresponds to the statistical uncertainty. On the right side the best-fit values, the full, the statistical and the systematic uncertainties for each NF are listed. The red line corresponds to the SM prediction of one.

NF agrees within a 1.2σ interval with the SM prediction, while all other NFs agree within an one- σ interval. The NFs for the VH_0, VH_1 and VBF_1 categories are larger than one and for VBF_0 smaller. The NFs for the boosted categories tend to have a trend from values larger than one to values smaller than one for categories with increasing p_T^H values, which is especially visible in the $\tau_h \tau_h$ final state. The highest correction of the background normalization is estimated in the ttH categories 19%. Most of the NFs are dominated by systematic uncertainties, except for the ones estimated in low statistics categories as VBF_1 and ttH. The estimated best-fit values and uncertainties for the 1-, 4- and 9-PoI setup show a good consistency, which confirms the robustness of the $Z \rightarrow \tau\tau$ background normalization estimation against a change of the PoI configuration.

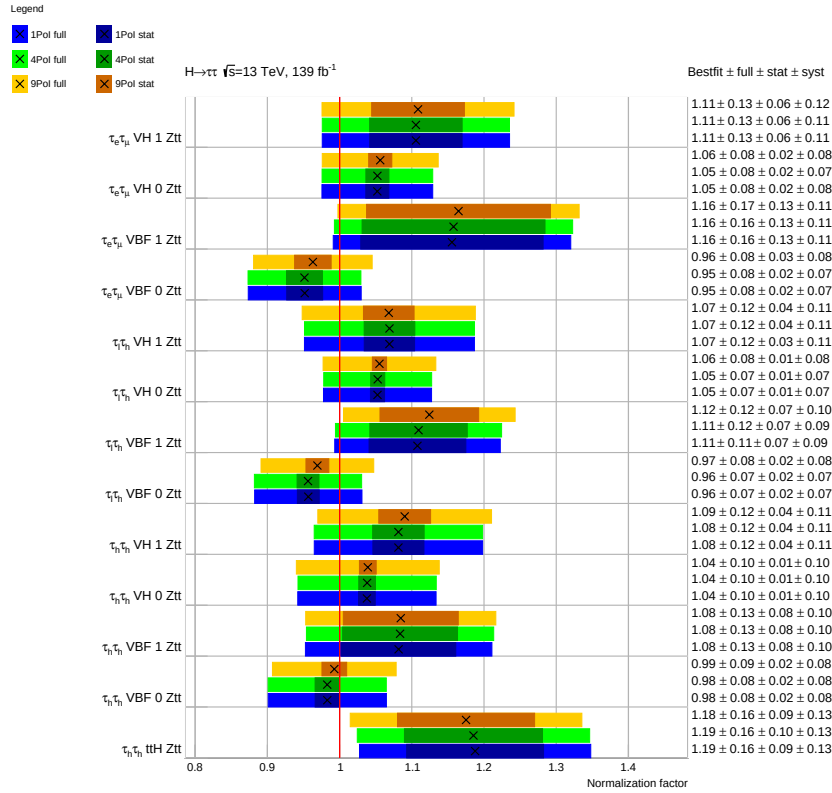


Figure 10.28: Measured NFs to normalize the $Z \rightarrow \tau\tau$ background for the 1-PoI (blue), 4-PoI (green) and 9-PoI (orange) setup in the VBF, VH and ttH categories. The crosses mark the best-fit values, while the colored boxes show their uncertainties. The lighter color illustrates the full uncertainty, while the darker corresponds to the statistical uncertainty. On the right side the best-fit values, the full, the statistical and the systematic uncertainties for each NF are listed. The red line corresponds to the SM prediction of one.

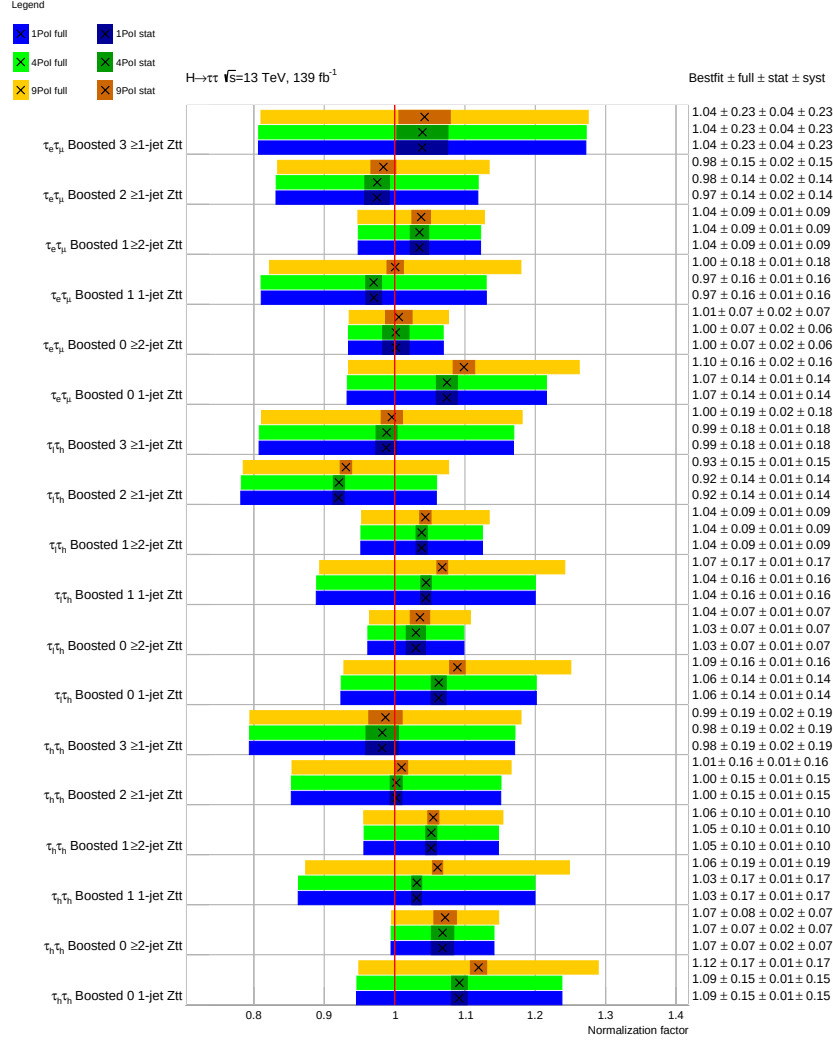


Figure 10.29: Measured NFs to normalize the $Z \rightarrow \tau\tau$ background for the 1-PoI (blue), 4-PoI (green) and 9-PoI (orange) setup in the boosted categories. The crosses mark the best-fit values, while the colored boxes show their uncertainties. The lighter color illustrates the full uncertainty, while the darker corresponds to the statistical uncertainty. On the right side the best-fit values, the full, the statistical and the systematic uncertainties for each NF are listed. The red line corresponds to the SM prediction of one.

10.6.2 Behavior of nuisance parameters in the fit

In this section the behavior of all NPs, which pass the pruning criteria, are studied for the 1-, 4- and 9-PoI setup. To do so, the pull and constraint of each NP is estimated. The pull of an NP illustrates whether the post-fit value of the NP agrees pre-fit value or not (See Section 7.2). A pull of zero corresponds to a perfect agreement, while non-zero values illustrates that the fit prefers an NP value that differs from the auxiliary measurement. The constraint shows by how much the fit can reduce the systematic uncertainty given by the auxiliary measurement. If the uncertainty of the auxiliary measurement is unchanged no constraint is expected, while for reduced uncertainties a value below one is expected. A constraint value above one would corresponds to an increasing of the size of the systematic uncertainty. This behavior is not expected and would hint to an instable likelihood model. In general, a strong constraint of a systematic uncertainty illustrates that the fit is very sensitive to this NP and therefore larger pulls are expected for this NP, while large pulls of NPs with a constraint close to one are not expected. Tables A.4 and A.5 in the appendix give an description of all NPs discussed in the this section.

Theoretical uncertainties on the signal First the NPs' parameters of uncertainties related to the theoretical prediction of the signal, described in Section 10.4, are studied. Figure 10.30 summarizes the pulls and constraints for all NPs related to parton-shower and matrix-element uncertainties. Most of the NPs show small pulls and no strong constraints, except for the parton shower uncertainty in the VBF production, which has a visible pull and constraint. This uncertainty is estimated by the comparison of two generators (see Section 10.4) and has a large impact in the VBF categories. Therefore, the fit is expected to have sensitivity to this NP.

Figure 10.31 (a) shows the pulls and constraints for the NPs related to PDF and α_s uncertain-

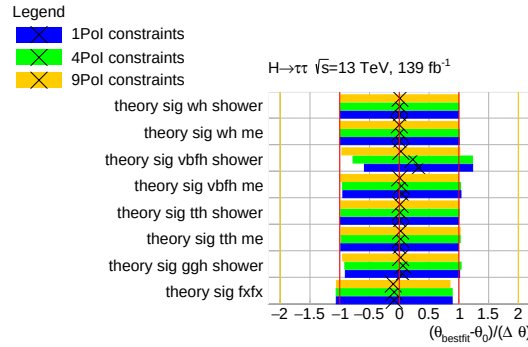


Figure 10.30: Pulls and constraints for the NPs estimated in the 1-PoI (blue), 4-PoI (green) and 9-PoI (orange) for the parton-shower, matrix-element (me, fxfx) related theoretical uncertainties on the signal.

ties, while Figure 10.31 (a) shows the results for the QCD scale uncertainties. Some minor pulls and constraints are visible for QCD scale NPs, while none are visible for PDF and α_s NPs.

Values for all NPs related to theoretical uncertainties on the signal prediction are a very

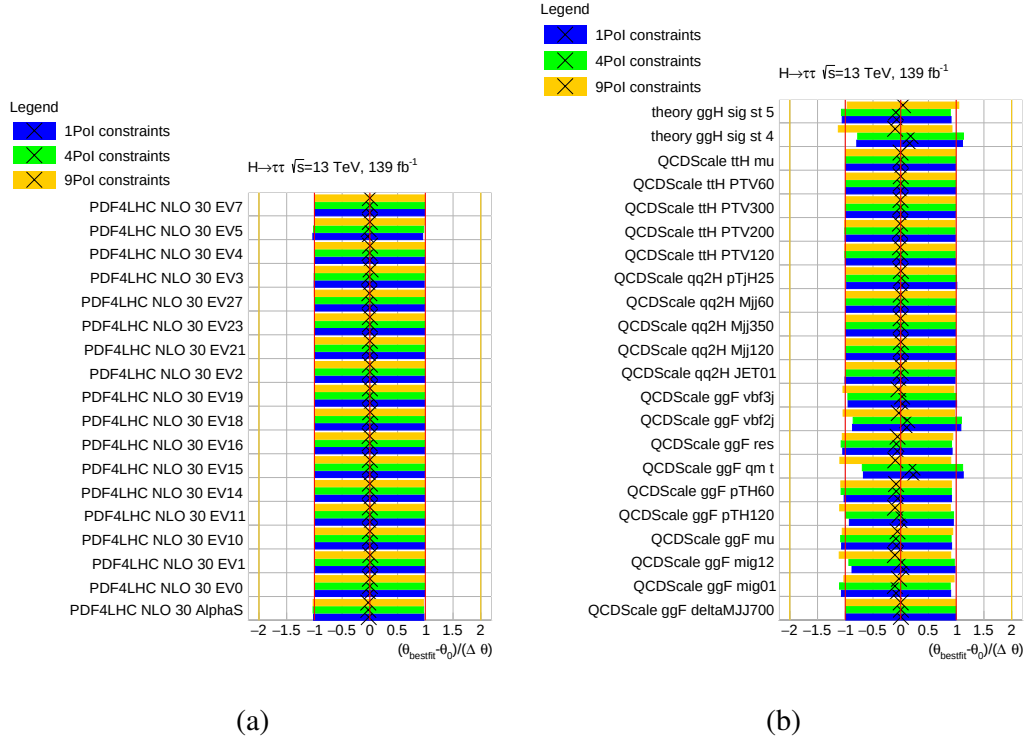


Figure 10.31: Pulls and constraints for the NPs estimated in the 1-PoI (blue), 4-PoI (green) and 9-PoI (orange) for the PDF and α_s (a) and QCD scale (b) related theoretical uncertainties on the signal.

consistent between the 1-, 4- and 9-PoI fits. Therefore, the set of NPs for these uncertainties is very robust against changing the PoI configuration scheme in the likelihood model.

Theoretical uncertainties on the background Figure 10.32 and Figure 10.33 (a) show the pulls and constraints of all NPs related to theoretical uncertainties on the $Z \rightarrow ll/\tau\tau$ background prediction. For the PDF and QCD scale uncertainties a fully decorrelated set of NPs across the different signal categories is used, which means that for each of them, 64 NPs are defined in the likelihood model. Figure 10.32 (a) shows 27 PDF related and (b) 45 QCD scale related NPs, so that over 50% of the PDF and 30% of the QCD scale NPs are pruned. Overall, several minor pulls and constraints are observed for these NPs, which are caused by kinematic difference between the signal categories and their object-based embedded $Z \rightarrow ll$ control regions. This shows that the reweighting of the $Z \rightarrow ll$ events does

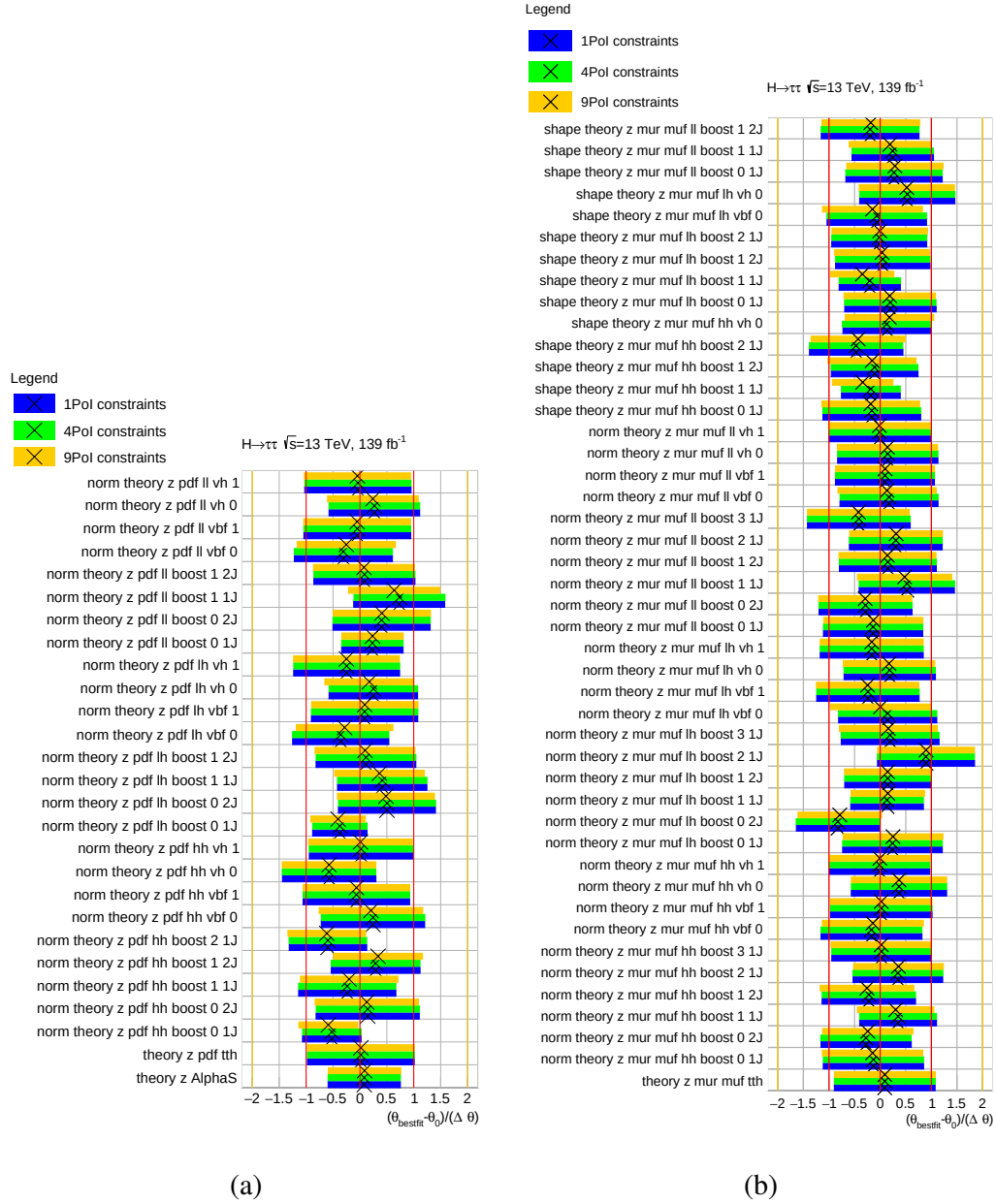


Figure 10.32: Pulls and constraints for the NPs estimated in the 1-PoI (blue), 4-PoI (green) and 9-PoI (orange) for the PDF and α_s (a) and QCD scale (b) related theoretical uncertainties on the $Z \rightarrow ll/\tau\tau$ background prediction.

not match the kinematic properties of the $Z \rightarrow \tau\tau$ events perfectly. However, the observed pulls and constraints have shown no significant impact on the best-values of the PoIs. The remaining NPs related to the theoretical prediction of the $Z \rightarrow ll/\tau\tau$ background are shown

in Figure 10.33 (a) and have minor pulls, except for the NP related to the QCD to EWK production ratio of $Z \rightarrow ll/\tau\tau$ background events in the VBF categories. The impact of this NP on the PoIs is negligible and therefore no additional investigations have been performed to study this behavior.

The NPs related to theoretical uncertainties on other backgrounds are shown in Figure 10.33 (b).

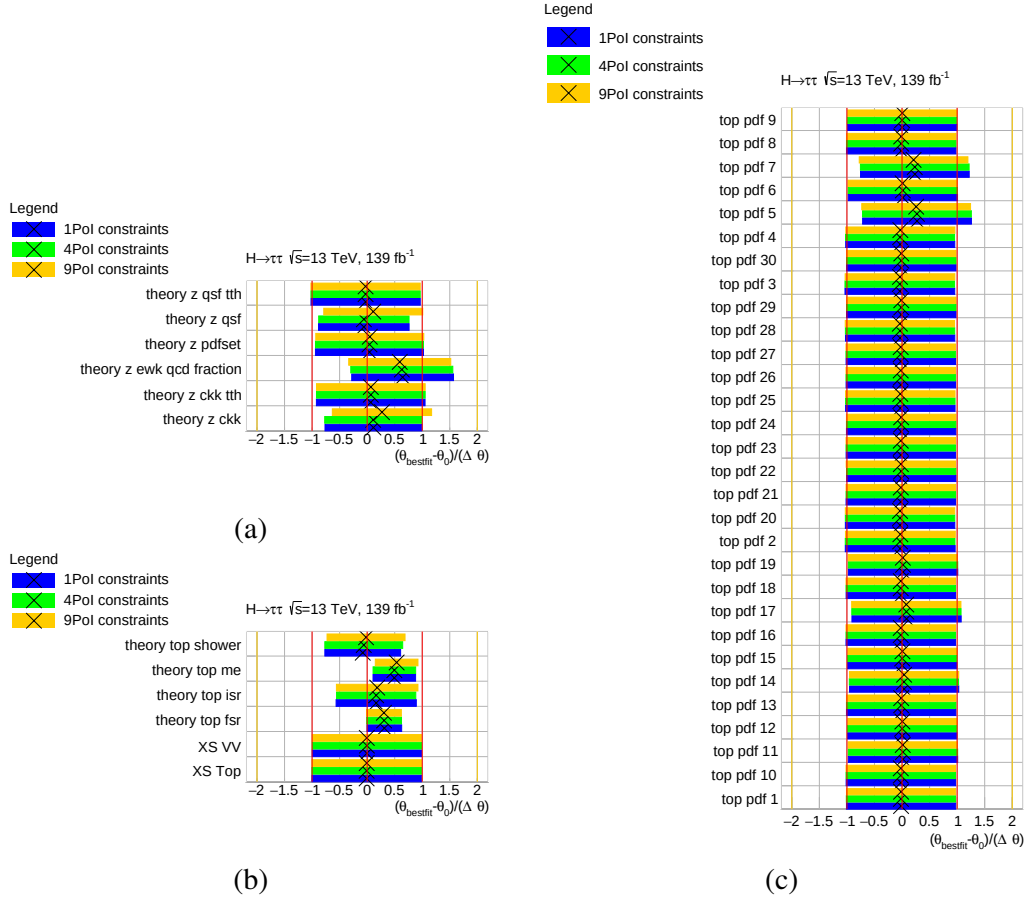


Figure 10.33: Pulls and constraints for the NPs estimated in the 1-PoI (blue), 4-PoI (green) and 9-PoI (orange) for the remaining theoretical uncertainties on the $Z \rightarrow ll/\tau\tau$ background prediction (a), for PDF related uncertainties on top (b) and for other theoretical uncertainties on the top and the di-boson (VV) production background (c).

Except for the NP assigned to the cross section prediction of di-boson events in the object-based embedded control regions, all NPs are related to the top quark production background. These uncertainties have a sizable impact in the $t\bar{t}H$ categories only. The prediction of these uncertainties suffer from the low data statistics in the $t\bar{t}H$ regions and are artificially inflated by statistical fluctuation arising from the limited statistics of the simulated events. Especially, the NPs related to the matrix-element and the final state radiation uncertainties are pulled and

constrained. These pulls are mainly introduced by shape effects of the NPs in the ttH categories but also by the boosted signal categories, which have a large data statistics to constrain them. These NPs interplay with the NF on the top quark production background in the ttH categories. The value above one of this NF observed in Figure 10.27 compensates for the pulls of these NPs.

Experimental uncertainties The pulls and constraints of NPs related to the jet systematic uncertainties, except for those on b-jet and τ_h -jets, are shown in Figure 10.35. Several of them show rather large pulls and strong constraints. They are mainly driven by a large shape variation of these uncertainties in the $m_{\tau\tau}^{\text{mmc}}$ distribution. Especially, the boosted signal region with the large data statistics is able to use this shape information to constrain and pull these NPs. A pull of these NPs effectively moves the $Z \rightarrow \tau\tau$ peak in the distribution of the $m_{\tau\tau}^{\text{mmc}}$ observable. The dependencies of the $m_{\tau\tau}^{\text{mmc}}$ shape on the jet related uncertainties is caused by the impact of the jet resolution and energy scale on E_T^{miss} , especially on the soft term. Therefore, the E_T^{miss} estimation, which is used in the MMC algorithm (see Section 5.7), is highly affected by the jet uncertainties. Furthermore, the uncertainties in modeling of pile-up and underlying event have a strong impact on the selected phase space with a large number of jets and/or τ_h -jets. On the other hand, the large data statistics in the boosted regions allows to have a high sensitivity to measure these effects. These behaviors of the jet related NPs is well known from the previous analyses [26, 233]. The same effect is visible for NPs related to uncertainties in E_T^{miss} and is illustrated in Figure 10.34 (b) and impacts also some of the NPs related to b-jet uncertainties shown in Figure 10.34 (a).

Figure 10.36 shows the NPs related to uncertainties in modeling of electrons and muons

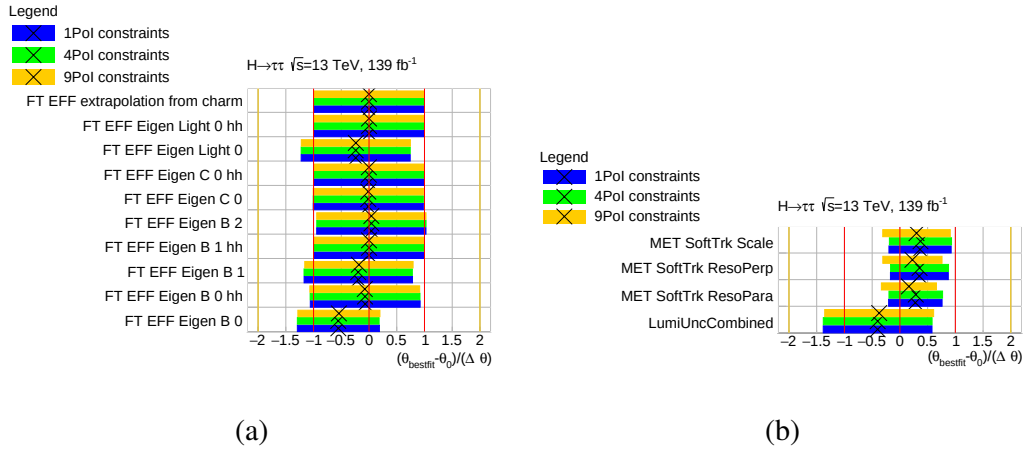


Figure 10.34: Pulls and constraints for the NPs estimated in the 1-PoI (blue), 4-PoI (green) and 9-PoI (orange) for b-jet (a) and E_T^{miss} and luminosity (b) related uncertainties.

(a) and to τ -leptons (b). The scheme for the electron uncertainties related to the object identification, which is able to corrected for kinematic dependencies, shows only small pulls

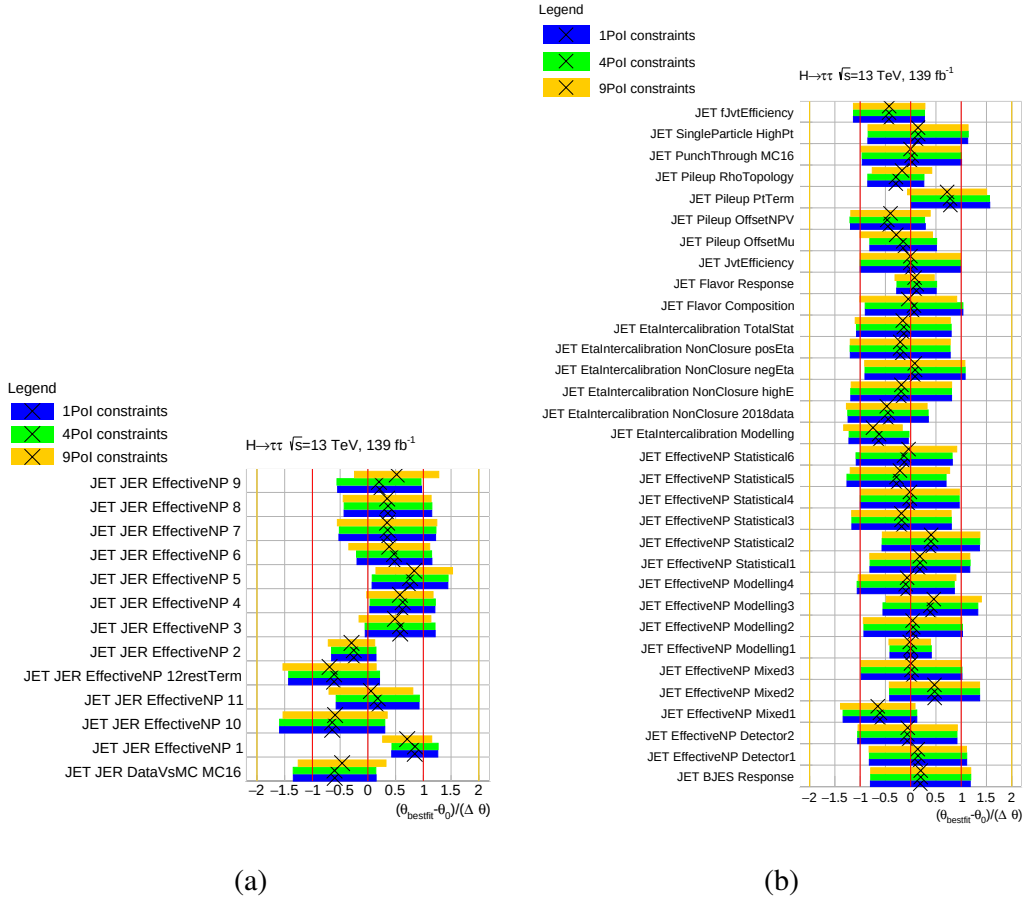


Figure 10.35: Pulls and constraints for the NPs estimated in the 1-PoI (blue), 4-PoI (green) and 9-PoI (orange) for jet energy resolution (a) and jet energy scale, flavor and vertex tagger (b) related uncertainties.

and constraints, while some NPs of the simplified scheme related to the muons are pulled. These uncertainties have only a sizable impact in the $\tau_e \tau_\mu$ final state, which has itself only a minor impact on the combined result. Therefore, these pulls have a negligible impact on the final results and are not studied in more detail. Some of the NPs related to the τ -lepton energy scale show pulls and constraints. The largest pull and constraint is visible for the NP related to the τ -lepton reconstruction efficiency. The estimation of this systematic uncertainties relies fully on simulated events as described in [202]. Therefore, the auxiliary measurement do not correct for differences between the simulation and data. The main reason for the fully simulation based estimation is the lack of an event selection with a high purity of τ_h -leptons, which could be used for a tag-and-probe measurement. This analysis provides such a selection via the object-based embedded control regions. Therefore, this analysis is expected to be more sensitive to the measurement of this uncertainty compared to the auxiliary measurement

itself and is able to correct for differences between simulation and data. Therefore, the prior of this NP is inflated in the likelihood model by a factor of four. In this way, the Gaussian constraint on the best-fit value of the NP is removed and the NP basically acts as a free floating parameter, which can be directly measured in the likelihood fit.

The NPs related to the data-driven method of the misidentified τ background are shown in

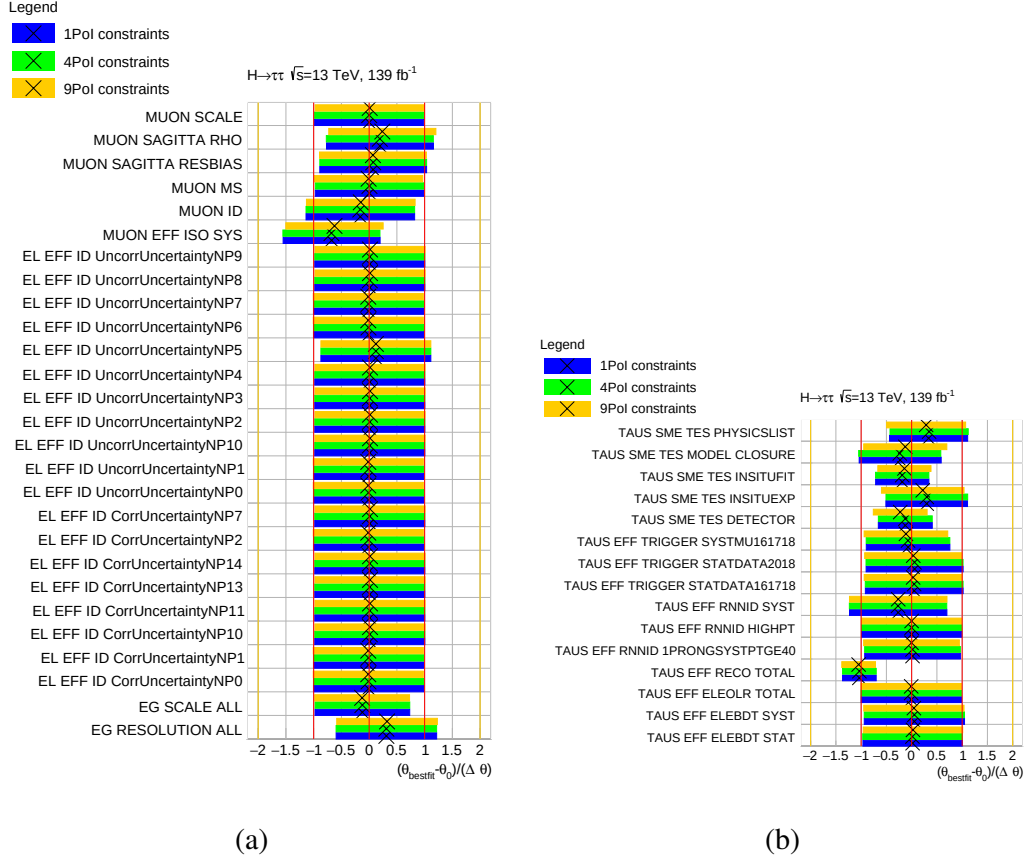


Figure 10.36: Pulls and constraints for the NPs estimated in the 1-PoI (blue), 4-PoI (green) and 9-PoI (orange) for electron and muon (a) and τ -lepton (b) related uncertainties.

Figure 10.37. For the $\tau_1 \tau_h$ (see Figure 10.37 (a)) most of the NPs show rather small constraints and pulls. The largest pull is visible for the NP related to the subtraction of other background contributions in the used control regions for τ_h -lepton decay into jets with one charged track in the boosted categories. This NP is estimated ad-hoc in a conservative way by simply applying a large variation on the yield of the subtracted background. Therefore, a pull and constraint of this parameter is reasonable and also for most of the other NPs related to the misidentified τ background estimation in the $\tau_1 \tau_h$ final state and in the other final states (see Figure 10.37 (b)). Figure 10.38 shows the NPs related to the migration uncertainties of the object-based embedding method used to estimate the normalization of the $Z \rightarrow \tau\tau$ background. The NPs show

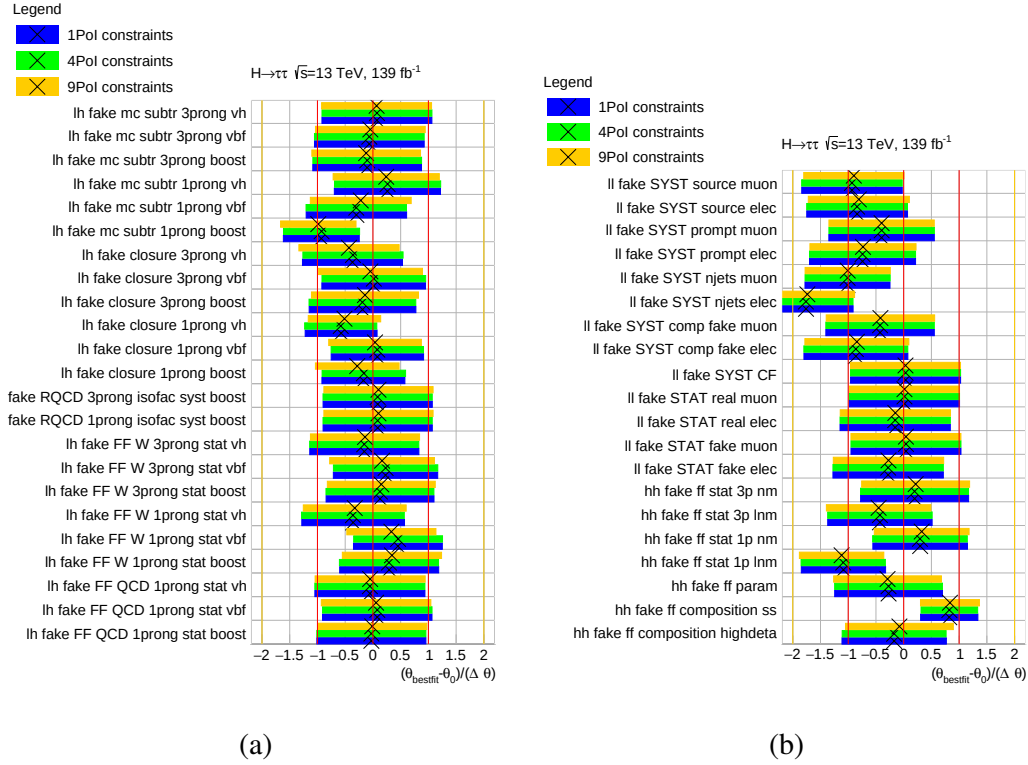


Figure 10.37: Pulls and constraints for the NPs estimated in the 1-PoI (blue), 4-PoI (green) and 9-PoI (orange) of the τ misidentified background estimation in $\tau_1\tau_h$ (a) as well as $\tau_e\tau_\mu$ and $\tau_h\tau_h$ (b).

relatively small constraints and pulls for all control regions. However, the NPs tend to correct for migration differences between the control regions with similar kinematic properties. For example the uncertainties on the $Z \rightarrow \tau\tau$ prediction in the $\tau_1\tau_h$ boosted_3_ge1J region is pulled to a negative value, while the NP for the $\tau_1\tau_h$ boosted_2_ge1J region is pulled to positive values.

None of the NPs related to experimental uncertainties show a significant difference between the 1-, 4- and 9-PoI fit setup.

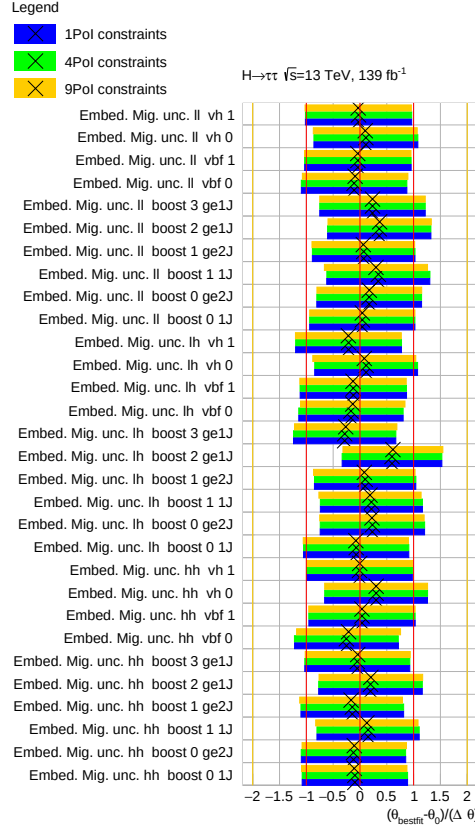


Figure 10.38: Pulls and constraints for the NPs estimated in the 1-PoI (blue), 4-PoI (green) and 9-PoI (orange) of the object-based embedding method used for the estimation of the $Z \rightarrow \tau\tau$ background.

10.6.3 Inclusive cross section measurements

In the 1-PoI fit the inclusive cross section of the $H \rightarrow \tau\tau$ decay channel is targeted and uses the assumption that the relative contribution of all Higgs boson production modes follow the SM prediction. The 4-PoI measures the cross section of the four main Higgs boson production modes ggF , VBF , VH and $t\bar{t}H$ simultaneously, while for all other production modes the cross section is assumed to be SM-like.

The observed inclusive cross section of the $H \rightarrow \tau\tau$ decay channel is:

$$\sigma_{H \rightarrow \tau\tau}^{obs} = 2.94 \pm 0.21(\text{stat})^{+0.37}_{-0.32}(\text{syst})\text{pb}, \quad (10.1)$$

which agrees with the SM prediction of 3.17 ± 0.09 pb with a p -value of 58%. The first row of lines of Table 10.16 summarizes the results of the 1-PoI fit and shows the individual impact of the four main sources of uncertainties (see Section 10.4) on the result. The measurement is dominated by systematic uncertainties with $^{+0.37}_{-0.32}$ p. The highest impact is estimated for the

theoretical uncertainties on the signal prediction with ± 0.26 pb, followed by experimental uncertainties with ± 0.19 pb. The impact of theoretical uncertainties on the background prediction is small. Compared to the previous analysis based on the data set with 36 fb^{-1} (see Section 9.7.1) the statistical uncertainty is reduced from 16% to 7% and the systematic uncertainties from 21% to 12%.

The consistency of the fit results across the different regions of phase space used in the fit

	SM prediction		Result		Stat. unc.	Syst. unc. [pb]		
Production Mode	[pb]		[pb]		[pb]	Th. sig.	Th. bkg.	Exp.
inclusive $H \rightarrow \tau\tau$	3.17	± 0.09	2.94	± 0.41	± 0.21	± 0.26	± 0.05	± 0.19
$t\bar{t}H$	0.0313	± 0.0032	0.033	± 0.037	± 0.031	± 0.010	± 0.010	± 0.010
VH	0.1176	± 0.0025	0.115	± 0.070	± 0.058	± 0.016	± 0.005	± 0.021
ggF	2.77	± 0.09	2.65	± 0.85	± 0.41	± 0.56	± 0.07	± 0.45
VBF	0.220	± 0.005	0.197	± 0.041	± 0.028	± 0.024	± 0.005	± 0.012

Table 10.16: Measured inclusive $H \rightarrow \tau\tau$ cross section (first row) and cross sections for the four dominant production modes ggF , VBF , VH and $t\bar{t}H$ (lower four rows) are shown in the second column, while their SM predictions are shown in the first column. The four right hand columns list separately the contributions to the total uncertainty in the measurements from statistical uncertainties (Stat. unc.) as well as from systematic uncertainties (Syst. unc.) on the signal prediction (Th. sig.), on the background prediction (Th. bkg.), and from the experimental performance (Exp.) [28].

model is studied with reduced likelihood models. For each fit setup includes only a sub-set of signal categories together with the corresponding control regions. For a stable fit result all NFs and NPs, which are not applied in the used signal categories or control regions are excluded. Seven sub-fits are preformed: Three fits in which only the signal categories of one final state are included ($\tau_e \tau_\mu$, $\tau_l \tau_h$, $\tau_h \tau_h$) and four fits in which only the categories of one inclusive signal category are included (tth, vbf, vh, boosted). The results for the final state split are shown in Figure 10.39 (a) and the results for the inclusive signal category split are shown in Figure 10.39 (b). In both figures, the results are divided by the SM prediction and the combined result is shown for comparison. All sub-fits agree within a one- σ interval with the combined results, except of the $\tau_e \tau_\mu$ sub-fit, which agrees within 1.3σ . A compatibility test between the sub-fits and the combined fit is performed. The measurements in the final states show a compatibility of 30%, while the measurements in the inclusive categories show a compatibility of 72%.

The results for the PoIs of the 4-PoI setup, in which the cross sections of the four main production modes are measured is summarized in Table 10.16 (lower four rows). All four observed cross sections agree within a one- σ uncertainty interval with the SM prediction. The highest precision is reached for the VBF production mode with a relative uncertainty of 21% (± 0.041 pb), followed by the ggF production mode with 32% (± 0.85 pb). The relative uncertainty for the VH production is 61% (± 0.070 pb) and for the $t\bar{t}H$ production

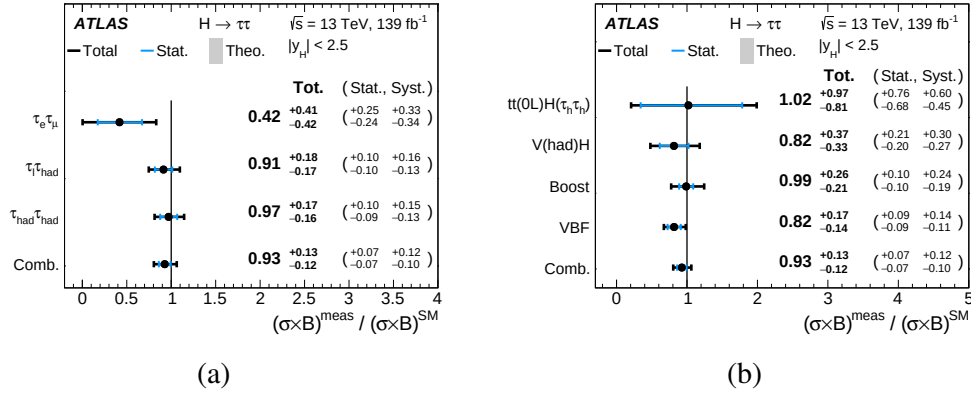


Figure 10.39: Cross section measurement using different sub-sets of the full likelihood model. (a) shows the fit results for using only one of the final states and (b) for using only categories of one inclusive signal category. In both figures, the result for the full likelihood model is shown. The cross sections are shown normalized to the SM prediction [28].

112% (± 0.037 pb). The measured ggF cross section is dominated by systematic uncertainties. The highest contribution is assigned to theoretical uncertainties on the signal prediction with ± 0.56 pb, followed by experimental uncertainties with ± 0.45 pb. The VBF measurement has a similar impact from the statistical and systematic uncertainties. The systematic uncertainties are dominated by the theoretical uncertainties on the signal prediction. The VH and $t\bar{t}H$ production modes are dominated by statistical uncertainties. Figure 10.40 (a) summarizes the measured cross sections. All results are divided by the SM prediction. The figure illustrates the good agreement of the measurement with the SM prediction, for which an overall compatibility of 98% is calculated. The correlations between the PoIs is illustrated in Figure 10.40 (b). The largest anti-correlation is observed between ggF and VH (-29%) and between ggF and VBF (-24%). These large anti-correlations are mainly introduced by the high contamination of ggF events in the VBF and VH categories (see Figure 10.13). The correlations between the remaining PoIs is below 10%. The high purity of the $t\bar{t}H$ signal in the $t\bar{t}H$ categories results in correlations with all other production modes below 2.1%. The observation of the $H \rightarrow \tau\tau$ decay was achieved in the previous analysis (see Section 9.6), but none of the individual production modes have been observed so far in the $H \rightarrow \tau\tau$ decay mode. Therefore, the 4-PoI setup is also used to estimate the significance of the two production modes with the highest sensitivity: ggF and VBF . For the VBF production mode an observed (expected) significance of 5.3σ (6.2σ) is measured, so that this production mode is observed for the first time in the $H \rightarrow \tau\tau$ channel. For the ggF production modes an observed (expected) significance of 3.9σ (4.6σ) is estimated, so that an evidence but no observation of the production mode in the $H \rightarrow \tau\tau$ decay channel is achieved.

Compared to the previous analysis, the relative statistical uncertainties could be reduced for the ggF production mode from 32% to 15% and for the VBF production from 32% to 13%. The systematic uncertainties are reduced from 48% to 27% for ggF and from 36% to 14% for

VBF.

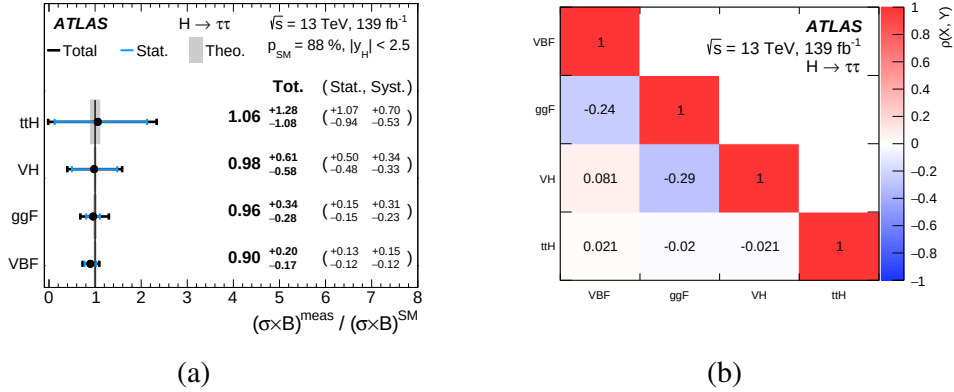


Figure 10.40: Results of the cross section measurement of the four inclusive main production modes. (a) shows the measured cross section with the uncertainties, normalized to the SM prediction, while (b) shows the correlation matrix of the PoIs [28].

Impact of systematic uncertainties Table 10.17 summarizes the relative impact of different sets of systematic uncertainties on the inclusive cross section measurements using the breakdown approach described in Section 7.2. The left column of the table shows that for the inclusive $\sigma_{H \rightarrow \tau\tau}$ cross section measurement the dominant set of experimental uncertainties is related to jets and E_T^{miss} with an impact of 4.53% and to statistical uncertainties of the background prediction due to the limited sample size with 3.99%. Some sizable contributions can be assigned to uncertainties related to τ_h -leptons (2.16%) and misidentified τ background (2.05%), while the remaining uncertainties have a minor impact on the measurement (<2%). Compared to the previous analysis (see Table 9.10) the relative impact of all systematic uncertainties could be reduced. However, the impact of the theoretical uncertainties on the signal prediction is more dominant compared to the previous measurement.

The measurement of the ggF and the VBF cross sections is dominated by systematic uncertainties (28.56% and 15.05%) and in particular by theoretical uncertainties on the signal prediction (21.23% and 12.37%). Overall, the relative impact of the different sets of systematic uncertainties are very similar for the ggF , VBF and inclusive $H \rightarrow \tau\tau$ cross section measurements. The only difference between ggF and VBF measurements is that the data statistics has a similar impact (14.16%) as the systematic uncertainties for VBF , while the impact of the data statistics is relatively small for ggF . The measurement of VH and $t\bar{t}H$ cross sections is dominated by data statistics (35.64% and 94.64%). The set of systematic uncertainties with the highest impact for both cross sections measurements is the size of the background sample (24.41% and 38.05%). For the VH cross section measurement a sizable impact of uncertainties related to the theoretical prediction of the signal (13.69%), to jets and E_T^{miss} (13.28%) and to misidentified τ -leptons (10.93%) is observed. For the measurement of the $t\bar{t}H$ cross section the theoretical uncertainties on the top quark production background

prediction (28.23%), the theoretical uncertainties on the signal prediction (23.14%) and uncertainties related to jets and E_T^{miss} (19.45%) have the largest impact after the size of the background sample.

Figure 10.41 shows the ten NPs with the highest impact on the inclusive $\sigma_{H \rightarrow \tau\tau}$ cross section

Set of uncertainty	Impact on [%]				
	$\Delta\sigma/\sigma_{H \rightarrow \tau\tau}$	$\Delta\sigma/\sigma_{H \rightarrow \tau\tau}^{\text{ggF}}$	$\Delta\sigma/\sigma_{H \rightarrow \tau\tau}^{\text{VBF}}$	$\Delta\sigma/\sigma_{H \rightarrow \tau\tau}^{\text{VH}}$	$\Delta\sigma/\sigma_{H \rightarrow \tau\tau}^{\text{tH}}$
All uncertainties	13.92	32.32	20.40	60.78	111.14
Data statistics	7.23	15.59	14.16	50.26	94.62
All systematic uncertainties	11.97	28.56	15.05	35.64	70.25
Theoretical unc. on signal	8.74	21.23	12.37	13.69	23.14
Jet and E_T^{miss}	4.53	12.87	4.68	13.28	19.45
Background sample size	3.99	9.52	5.69	24.41	38.05
τ_h -leptons	2.16	5.23	1.45	4.09	5.19
Misidentified τ -leptons	2.05	5.13	2.34	10.93	4.82
Luminosity	1.83	2.20	1.67	1.23	1.32
Theoretical unc. on Z background	1.19	2.41	2.12	1.02	3.02
Theoretical unc. on top background	1.08	1.05	1.11	4.19	28.23
B-jets	0.43	0.63	0.67	0.61	9.07
Electrons and muons	0.43	1.67	0.22	0.82	0.66

Table 10.17: Summary of the relative impact of different sets of systematic uncertainties on the measurements of the cross sections in the 1- and 4-PoI setups [28].

measurement using the NP ranking approach (see Section 7.2). Eight of the ten NPs are related to theoretical uncertainties on the signal prediction. Seven of them are QCD scale uncertainties on the ggF production, while the leading one is related to the parton shower uncertainty in the VBF production. In addition, one NP related to jet uncertainties and one NP related to the luminosity are shown.

Figure 10.42 (a) shows the ten NPs with the largest impact on the measurement of $\sigma_{H \rightarrow \tau\tau}^{\text{ggF}}$ in the 4-PoI setup. Eight of them are related to theoretical uncertainties on the signal prediction, six to QCD scale uncertainties and two related to the parton shower uncertainties in the ggF and VBF production modes. Due to the large size of the parton shower uncertainty in the VBF production, its impact on the measurement of $\sigma_{H \rightarrow \tau\tau}^{\text{ggF}}$ is sizable due to contamination from VBF production in the boosted categories. The remaining two NPs are related to uncertainties on E_T^{miss} and the misidentified τ background.

Figure 10.42 (b) displays the top ten ranked NPs on the measurement of $\sigma_{H \rightarrow \tau\tau}^{\text{VBF}}$. Five of the ten NPs are related to theoretical uncertainties on the signal prediction. The impact of the leading the parton shower uncertainties in the VBF production is three times larger than the second highest ranked NP. Therefore, the remaining NPs have a rather small impact on the measurement. Four of the remaining NPs are related to uncertainties on jets.

The top ten ranked NPs for the measurements of $\sigma_{H \rightarrow \tau\tau}^{\text{VH}}$ and $\sigma_{H \rightarrow \tau\tau}^{\text{tH}}$ are shown in Figures 10.42 (c) and (d). Due to the high impact of the data statistics and of the size of the background sample on these cross section measurements, the importance of the ten highest ranked NPs is relatively small. For both, NPs related to theoretical uncertainties on the signal prediction and to uncertainties on jets are highly ranked. For the $\sigma_{H \rightarrow \tau\tau}^{\text{tH}}$ measurement,

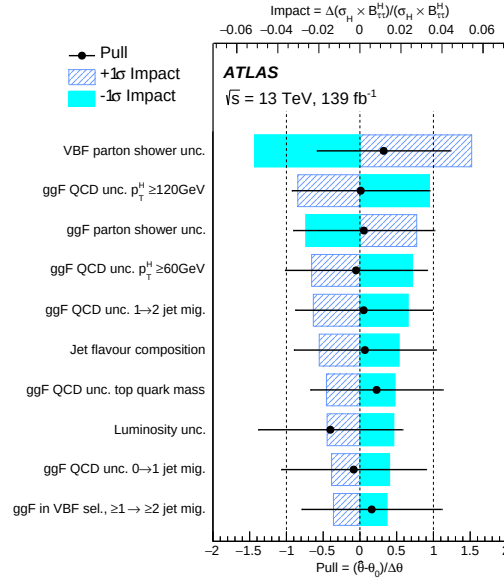


Figure 10.41: Relative impact of the ten NPs with the largest impact on the inclusive $\sigma_{H \rightarrow \tau\tau}$ cross section measurement. The uncertainties are listed in decreasing order of their impact. The upper x-axis corresponds to the impact of the variation of the NPs by one- σ on the inclusive cross section measurement and is visualized by blue boxes. The lower x-axis corresponds to the pulls and constraints of the NPs, which are illustrated with black points and lines [28].

two NFs scaling the top quark production and $Z \rightarrow \tau\tau$ predictions and two NPs related to theoretical uncertainties on the top production background prediction are ranked under the top ten NPs.

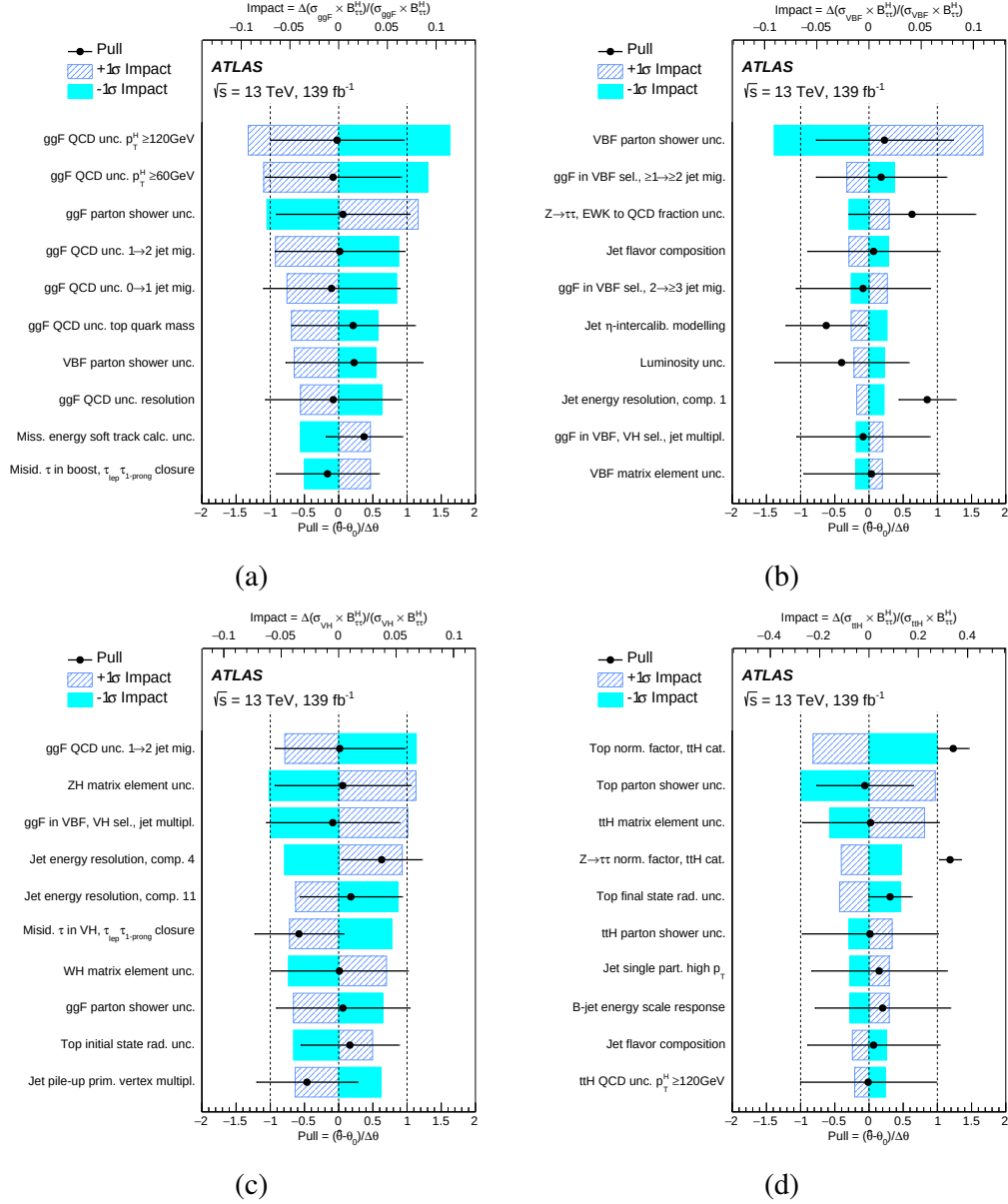


Figure 10.42: Relative impact of the ten NPs with the largest impact on the inclusive $\sigma_{H \rightarrow \tau\tau}^{\text{ggF}}$ (a), $\sigma_{H \rightarrow \tau\tau}^{\text{VBF}}$ (b), $\sigma_{H \rightarrow \tau\tau}^{\text{VH}}$ (c) and $\sigma_{H \rightarrow \tau\tau}^{\text{ttH}}$ (d) cross section measurements. The uncertainties are listed in decreasing order of their impact. The upper x-axis corresponds to the impact of the variation of the NPs by one- σ on the inclusive cross section measurements and is visualized by blue boxes. The lower x-axis corresponds to the pulls and constraints of the NPs, which are illustrated with black points and lines [28].

10.6.4 STXS measurement

In this section the measured cross sections in the 9 STXS bins and their uncertainties are discussed and compared to the SM prediction. The fit measures the cross sections of five ggF STXS bins:

- $ggF, N_j \geq 2, 350 \text{ GeV} \leq m_{jj}, p_T^H < 200 \text{ GeV}$ (ggF VBF-like)
- $ggF, N_j \geq 2, m_{jj} < 350 \text{ GeV}^\spadesuit, 60 \leq p_T^H < 120 \text{ GeV}$ ($ggF, p_T^H : [60, 120]$)
- $ggF, N_j = 1, m_{jj} < 350 \text{ GeV}, 120 \leq p_T^H < 200 \text{ GeV}$ (ggF 1-jet $p_T^H : [120, 200]$)
- $ggF, N_j \geq 2, m_{jj} < 350 \text{ GeV}, 120 \leq p_T^H < 200 \text{ GeV}$ ($ggF \geq 2\text{-jet } p_T^H : [120, 200]$)
- $ggF, N_j \geq 1, 200 \leq p_T^H < 300 \text{ GeV}$ ($ggF, p_T^H : [200, 300]$)
- $ggF, N_j \geq 1, 300 \text{ GeV} \leq p_T^H$ ($ggF, p_T^H : [300, \infty]$)

The $\spadesuit$ indicates that the $m_{jj} < 350 \text{ GeV}$ requirement is only applied on events with at least two jets. In addition, cross sections of two $EW qqH$ STXS bins and one $t\bar{t}H$ STXS bin are measured:

- $EW qqH, N_j \geq 2, 60 \leq m_{jj} < 120 \text{ GeV}$ ($EW qqH$ VH-like)
- $EW qqH, N_j \geq 2, 350 \text{ GeV} \leq m_{jj}$ ($EW qqH$ VBF-like)
- $t\bar{t}H$

The short term defined for each STXS bin in the round brackets will be used throughout the rest of this section to improve readability.

Compared to the previous measurement based on the data set recorded in 2015 and 2016 (see Section 9.8), the granularity of this STXS measurement is increased. In the previous measurement only two STXS bins for the ggF production mode with the requirements $N_j \geq 1, 60 \leq p_T^H < 200 \text{ GeV}$ and $N_j \geq 1, p_T^H \geq 200 \text{ GeV}$ have been measured and for the remaining production modes only one $EW qqH$ VBF-like STXS bin was defined as described in Section 7.1 with slightly different requirements compared to this analysis. Therefore, the current measurement explores new regions of phase space by including the VH and $t\bar{t}H$ signal categories to measure the $EW qqH$ VH-like and $t\bar{t}H$ STXS bins and increases the granularity in the boosted ggF phase space.

Figure 10.43 and Table 10.18 summarize the measured values of the nine STXS bins and the uncertainties on the measurement. In the ratio panel of Figure 10.43 the good agreement of the measurement with the SM prediction is visible in particular. None of the STXS bins differs with more than one- σ from their SM prediction and the estimated p -value of the compatibility between the measurement and the SM prediction is 96%.

The smallest relative uncertainties are estimated for the $EW qqH$ VBF-like (25%), the $ggF, p_T^H : [200, 300]$ (36%) and the $ggF, p_T^H : [300, \infty]$ (40%) STXS bins. The remaining ggF STXS bins are measured with a relative uncertainty between 75% and 206%. The measurement of the $t\bar{t}H$ STXS bin has a relative uncertainty of 111% and the measurement of the $EW qqH$

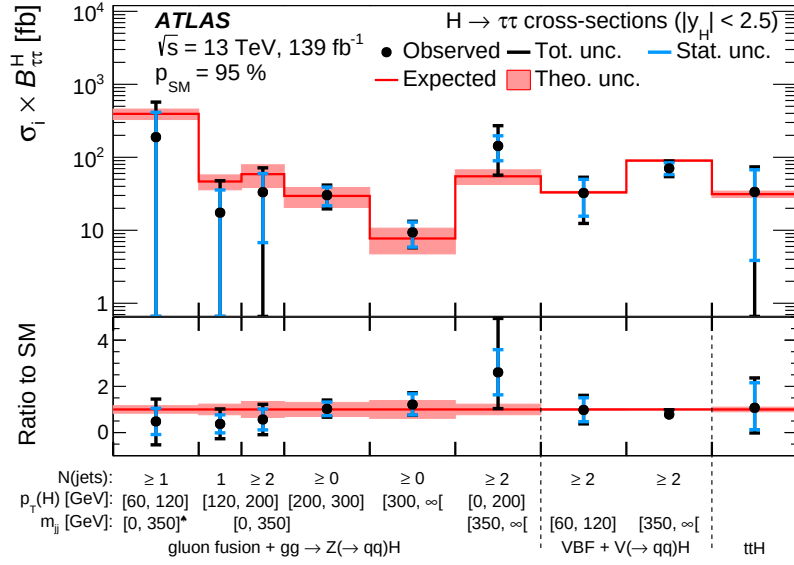


Figure 10.43: Results of the STXS measurement with nine independent PoIs. The measured cross sections are marked with black dots and their total (statistical) uncertainties are shown with black (blue) error bars. The SM prediction for each STXS bin is indicated by a red line and the theoretical uncertainty on the prediction by a red band. The spades symbol ♠ indicates, that the requirement on m_{jj} is only applied for $N_j \geq 2$ [28].

VH-like STXS bin has 63% relative uncertainty.

Table 10.18 also shows the uncertainties from data statistics, theoretical uncertainties on the signal and background prediction and from experimental sources. For most of the STXS bins, the precision of the measurement is dominated by the data statistics in contrast to the inclusive measurements. The measurement of the finer granularity of phase space increases the relative statistical uncertainty by decreasing the amount of data used to estimate the independent cross sections but also removes partially theoretical uncertainties on the signal prediction by decreasing the model dependence and the extrapolation from the detected phase space to the measured STXS bin. The STXS bins, which are not dominated by data statistics are the ggF STXS bins with $p_T^H \leq 200$ GeV. Particularly, the measurement of the ggF $p_T^H : [60, 120]$ STXS bin is dominated by experimental uncertainties. Most of the events from this STXS bin populate the boosted_0 signal categories. These large uncertainties can be assigned to the necessary extrapolation from these two signal categories to the phase space of the measured STXS bin. The signal categories have a requirement of $100 \leq p_T^H < 120$ GeV, while the phase space of the STXS bin is defined by $60 \leq p_T^H < 120$ GeV. For the ggF VBF-like STXS bin no specific signal categories are defined and therefore the STXS bin is measured in regions of phase space designed for other STXS bins. This increases the impact of experimental uncertainties on this measured STXS bins. Furthermore, the main contribution of this signal

STXS bins				SM prediction		Result		Stat. unc.	Syst. unc. [fb]		
Process	m_{ij} [GeV]	p_T^H [GeV]	N_j	[fb]		[fb]		[fb]	Th. sig.	Th. bkg.	Exp.
ggF	[0,350]♣	[60,120]	≥ 1	394	± 60	189	± 360	± 220	± 59	± 152	± 240
		[120,200]	$= 1$	47	± 11	17	± 30	± 18	± 4	± 4	± 16
	[0,350]	[120,200]	≥ 2	59	± 20	33	± 39	± 27	± 10	± 10	± 23
		[200,300]	≥ 1	30.5	± 9	30.3	± 11.0	± 8.6	± 2.9	± 0.8	± 5.6
		[300, ∞]	≥ 1	7.7	± 3.0	9.35	± 3.80	± 3.50	± 1.00	± 0.22	± 1.20
	[350, ∞]	[0,200]	≥ 2	55	± 13	143	± 110	± 54	± 58	± 6	± 71
$EW\ qqH$	[60,120]		≥ 2	33.1	± 1.1	33	± 20	± 17	± 4	± 2	± 6
	[350, ∞]		≥ 2	90.1	± 2.2	71	± 17	± 13	± 10	± 2	± 4
$t\bar{t}H$				31.3	± 3.2	34	± 37	± 32	± 7	± 10	± 8

Table 10.18: Observed cross sections and uncertainties in the 9-PoI fit setup. The SM predictions for each parameter are computed using the inclusive cross section calculations and the simulated event samples. On the right-side the contributions to the total uncertainty in the measurements from statistical (Stat. unc.) or systematic uncertainties (Syst. unc.) in the signal prediction (Th. sig.), background prediction (Th. bkg.), and in experimental performance (Exp.) are given separately. The spades symbol ♣ indicates, that the requirement on m_{jj} is only applied for $N_j \geq 2$ [28].

populates the VBF categories, which suffer from a large impact from theoretical uncertainties on the signal, particularly the parton shower uncertainty in the VBF production mode. The STXS bins for the $EW\ qqH$ and $t\bar{t}H$ production have a similar contribution of uncertainties from data statistics, theoretical uncertainties on the signal and background prediction and from experimental sources as the VBF , VH and $t\bar{t}H$ cross sections of the 4-PoI setup.

A comparison of the measured STXS bins and their an uncertainties to the STXS measurement of the previous analysis is shown in Table 10.19. The table lists the STXS bins of this measurement based on Stage 1.2 and the STXS bins at Stage 1 used in the previous measurement (see Section 7.1). This comparison shows the improvements in the STXS measurement between the two analyses. The three STXS bins of the current analysis could not be extracted in the previous measurement due to missing signal categories or low sensitivity. The contribution of these STXS bins has been fixed to their SM prediction in the likelihood fit. For the three ggF STXS bins with an upper threshold on p_T^H of 200 GeV or below and $m_{jj} < 350$ GeV, the comparison is rather difficult. Only one STXS bin of the previous measurement can be assigned to them, which shows the increased granularity of the current measurement. This STXS bin was measured with a precision of 51%, while the uncertainties on the three measured STXS bins of the current analysis are between 116% and 206%. However, in the previous measurement the measured cross section is 4.5 times large than its SM prediction, while in the current measurement all three bins are measured to be 0.4 and 0.6 times their SM prediction. The two STXS bins $ggF\ p_T^H : [200, 300]$ and $ggF\ p_T^H : [300, \infty[$ of the current measurement are combined in the previous measurement to one STXS bin. The relative uncertainty of the previous measurement on this STXS bin is with 59% larger than the relative uncertainties of both STXS bins of the current analysis (36% and 40%). This shows the large

STXS bins (Stage 1.2)				Current analysis (Stage 1.2)		Previous analysis (Stage 1)	
Process	m_{jj} [GeV]	p_T^H [GeV]	N_j	σ_{obs}/σ_{SM}	Uncertainty [%]	σ_{obs}/σ_{SM}	Uncertainty [%]
ggF	[0,350]♠	[60,120]	≥ 1	0.48	206.3	4.48	50.9
		[120,200]	$= 1$	0.37	172.5		
	[0,350]	[120,200]	≥ 2	0.57	116.1	0.86	58.9
		[200,300]	≥ 1	1.02	36.3		
		[300,∞[	≥ 1	1.21	40.4		
	[350,∞[	[0,200]	≥ 2	2.61	74.9	Set to SM prediction	
$EW\ qqH$	[60,120]		≥ 2	0.98	63.1	Set to SM prediction	
	[350,∞[		≥ 2	0.78	24.6	1.14	45.3
$t\bar{t}H$				1.07	111.2	Set to SM prediction	

Table 10.19: Measured values for STXS bins divided by their SM predictions (σ_{obs}/σ_{SM}) and the relative uncertainty on each measured STXS bin for the current and previous analyses (see Chapter 9). The requirements in the four left hand columns corresponds to the STXS Stage 1.2 definition used in the current analysis. The STXS bins of previous analysis based on Stage 1 are listed in the same row as STXS bins of Stage 1.2 with a similar but not the same definition (see Section 7.1). The spades symbol ♠ indicates, that the requirement on m_{jj} is only applied for $N_j \geq 2$.

improvement of the current analysis by optimizing the signal category definition, in particular the harmonization of the requirements on p_T^H in the selection of the signal categories with the definition in the STXS bins. The relative uncertainty on the measurement of the $EW\ qqH$ VBF-like bin is decreased from 45% to 25% for the new analysis.

Figure 10.44 shows the correlation matrix of the nine PoIs. The correlation varies between -40% and +10%. The largest anti-correlation is observed for the $EW\ qqH$ VBF-like and ggF VBF-like STXS bins. This anti-correlation is introduced by the high contribution of this ggF STXS bin in the VBF signal categories. For measured ggF STXS bins with a similar topology an anti-correlation between -30% and -10% is observed.

Impact of systematic uncertainties The impact of different sets of systematic uncertainties on the measured ggF and on the remaining measured STXS bins are shown in Tables 10.20 and 10.21. The impact is estimated using the breakdown approach described in Section 7.2. Systematic uncertainties related to jets and E_T^{miss} but also the size of the background samples are dominating contributions to the experimental uncertainties on the measurement of ggF STXS bins. In particular for the STXS bins $ggF\ p_T^H : [60, 120]$ and $ggF\ 1\text{-jet } p_T^H : [120, 200]$, the relative impact of this set of systematic uncertainties amounts to 70% to 108%. The impact of the theoretical uncertainties on the signal prediction is smaller or of the same size as uncertainties related to jets and E_T^{miss} and the size of the background samples on the measurements of the ggF STXS bins. In contrast to the measurement of the inclusive

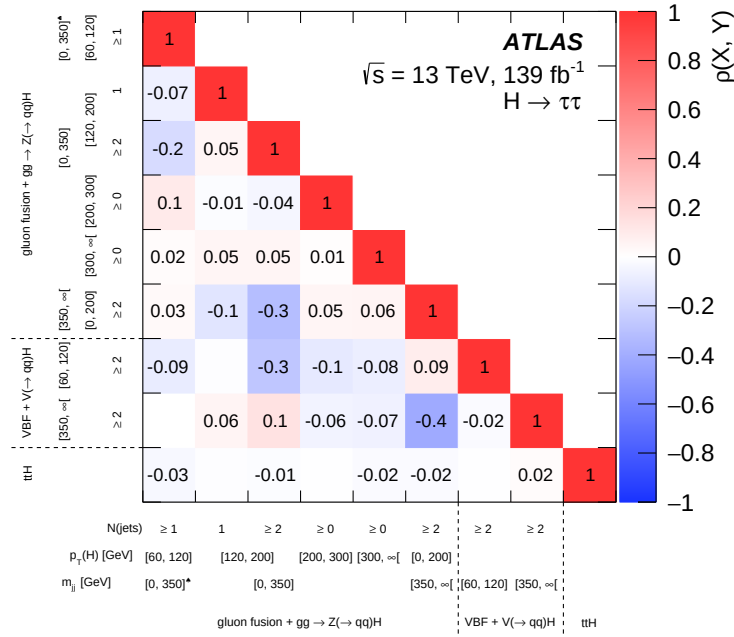


Figure 10.44: Observed correlation matrix of the nine PoIs fit setup. The color code indicates a correlation between two PoIs in red (positive value) and anti-correlation in blue (negative value). The spades symbol ♠ indicates, that the requirement on m_{jj} is only applied for $N_j \geq 2$ [28].

ggF cross section, in which theoretical uncertainties dominate the total uncertainty[†] (see Table 10.17). This shows that the increased granularity of the STXS measurements reduces the model dependency and therefore the impact of theoretical uncertainties on the signal prediction of the measurement. In addition, a sizable impact of uncertainties related to τ_h -leptons (5% to 35%) and misidentified τ -leptons (4% to 57%) is observed for the ggF STXS bins, while the impact of the remaining uncertainties on the measurement is rather small.

The experimental uncertainties on the measurement of the $EW\ qqH$ VH-like and $t\bar{t}H$ STXS bins are dominated by the size of the background sample (26% and 38%) (see Table 10.21), followed by uncertainties related to misidentified τ -leptons for the $EW\ qqH$ VH-like STXS bin (12%) and by theoretical uncertainties on the predictions of the top quark background for the $t\bar{t}H$ STXS bin (28%). The systematic uncertainties on the measurement of the $EW\ qqH$ VBF-like STXS bin is dominated by theoretical uncertainties on the signal prediction (14%), while the size of the background sample is the dominating contribution to the experimental uncertainties (7%).

The ten NPs and NFs with the largest impact on each measured STXS bin are shown in Figures 10.45-10.46. The impact is estimated with the NP ranking approach (see Section 7.2). In the measurement of the ggF VBF-like STXS bin (see Figure 10.45 (a)) NP related to theoretical uncertainties on the ggF signal prediction in the VBF topology and to uncertainties

10.6 Results

Process N_j m_{jj} [GeV] p_T^H [GeV]	Impact on [%]					
	ggF					
	≥ 2	≥ 1	$= 1$	≥ 2	≥ 1	≥ 1
	[350,∞[[0,200]	[0,350]♣ [60,120]	[120,200]	[0,350] [120,200]	[200,300]	[350,∞[
All uncertainties	74.9	206.3	172.5	116.1	36.3	40.4
Data statistics	37.5	118.4	104.0	78.0	28.6	37.7
All systematic uncertainties	65.6	171.7	141.3	89.3	24.7	21.6
Theoretical unc. on signal	40.6	31.1	25.1	29.2	9.9	10.6
Jet and E_T^{miss}	42.3	101.3	81.8	54.9	15.3	8.3
Background sample size	26.0	107.8	71.8	46.1	11.5	11.2
τ_h -leptons	13.8	35.1	35.2	17.5	5.0	5.6
Misidentified τ -leptons	16.8	57.0	24.2	19.5	4.1	4.0
Luminosity	2.7	9.3	5.8	2.0	2.2	2.1
Theoretical unc. on Z background	2.7	9.3	5.8	2.0	2.2	2.1
Theoretical unc. on top background	2.9	22.4	3.6	11.3	1.8	2.5
B-jets	1.4	3.3	4.1	4.4	0.2	0.7
Electrons and muons	1.2	9.8	9.4	9.5	0.8	1.4

Table 10.20: Summary of the relative impact of different sets of systematic uncertainty on the measurements of the ggF STXS bins in the 9-PoI setup.

Process N_j m_{jj} [GeV]	Impact on [%]		$t\bar{t}H$
	EW	qqH	
	≥ 2 [60,120]	≥ 2 [350,∞[	
All uncertainties	63.1	24.6	111.2
Data statistics	52.9	18.3	94.8
All systematic uncertainties	36.1	17.0	70.3
Theoretical unc. on signal	11.9	13.5	22.9
Jet and E_T^{miss}	11.2	5.4	19.8
Background sample size	25.8	7.0	37.9
τ_h -leptons	3.6	1.8	5.2
Misidentified τ -leptons	12.0	1.5	5.0
Luminosity	1.2	1.5	1.2
Theoretical unc. on Z background	1.1	2.9	3.3
Theoretical unc. on top background	5.4	0.9	28.2
B-jets	1.0	0.6	9.1
Electrons and muons	1.6	0.2	0.9

Table 10.21: Summary of the relative impact of different sets of systematic uncertainty on the measurements of the EW qqH and $t\bar{t}H$ STXS bins in the 9-PoI setup.

on jets are highly ranked. On the measurement of the ggF p_T^H : [60, 120], ggF 1-jet p_T^H : [120, 200] and ggF ≥ 2 -jet p_T^H : [120, 200] NPs related to the theoretical prediction of the $Z \rightarrow ll/\tau\tau$ background prediction, NFs on the $Z \rightarrow \tau\tau$ background and NPs related to uncertainties on jets and E_T^{miss} are highly ranked (see Figures 10.45 (b)-(d)).

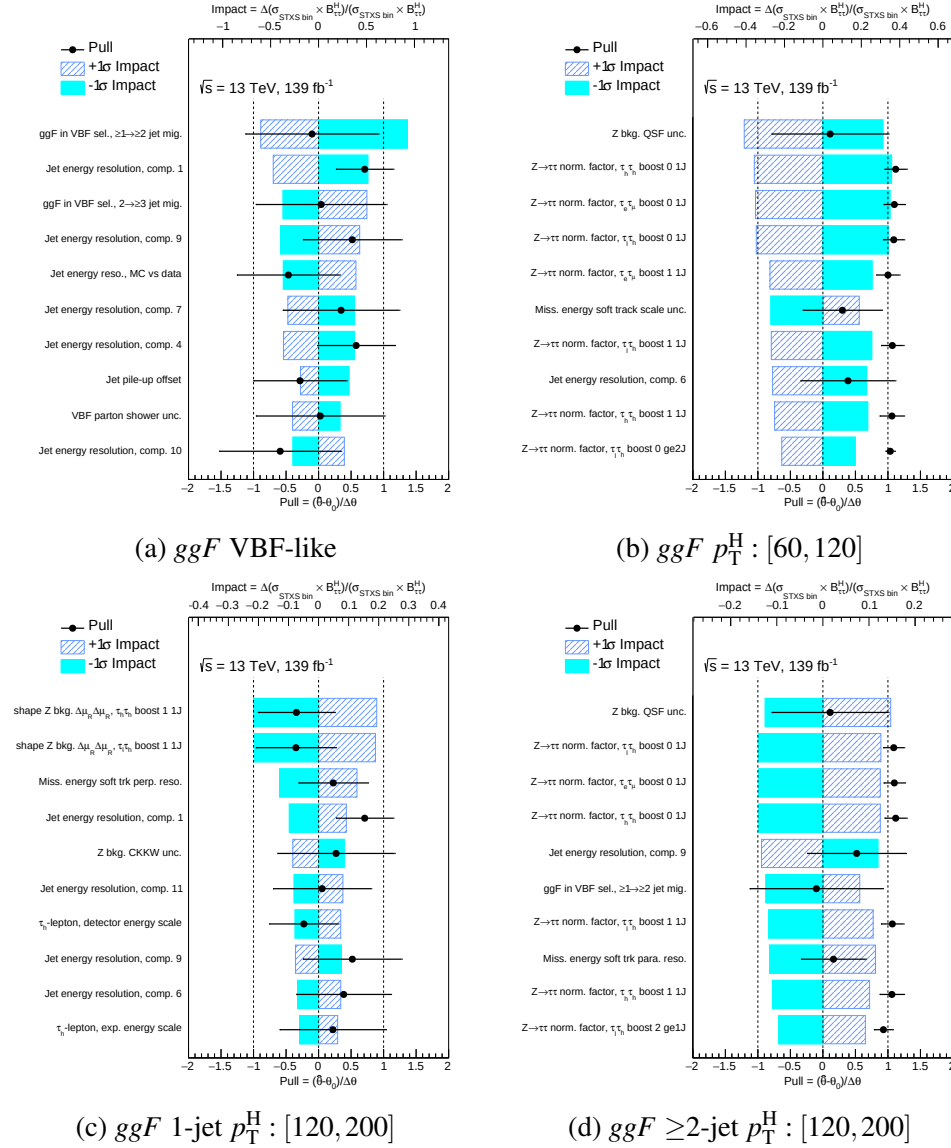


Figure 10.45: Relative impact of the ten NPs with the largest impact on the measurement of the ggF VBF-like (a), $ggF p_T^H : [60, 120]$ (b), ggF 1-jet $p_T^H : [120, 200]$ (c) and $ggF \geq 2$ -jet $p_T^H : [120, 200]$ (d) STXS bins. The uncertainties are listed in decreasing order of their impact. The upper x-axis corresponds to the impact of the variation of the NPs by one- σ on the inclusive cross section measurements and is visualized by blue boxes. The lower x-axis corresponds to the pulls and constraints of the NPs, which are illustrated with black points and lines.

Figures 10.45 (a) and (b) show the NPs with the largest impacts on the measurements of the ggF $p_T^H : [200, 300]$ and ggF $p_T^H : [300, \infty[$ STXS bins. For both STXS bins, the NP related to the parton shower uncertainty in the ggF production has the largest impact, followed mainly by NPs related to uncertainties on jets and E_T^{miss} .

The NPs with the largest impact on the measurement of the EW qqH VH-like STXS bins are related to theoretical uncertainties in the matrix element calculation in the VH production, to theoretical uncertainties on the ggF signal prediction and to uncertainties on misidentified τ -leptons (see Figure 10.47 (a)). The dominating NP on the measurement of the EW qqH VBF-like STXS bins is similarly to the measurement of the inclusive cross section of the VBF production mode (see Figure 10.42 (b)) related to the parton shower uncertainty in the VBF production, while the impact of other NPs is rather small (see Figure 10.47 (b)). Figure 10.47 (c) shows the ten NPs and NFs with the largest impact on the measurement of the $t\bar{t}H$ STXS bin. Both NFs defined to scale the top quark production mode and the $Z \rightarrow \tau\tau$ background have a large impact on the $t\bar{t}H$ STXS bin. In addition, NPs related to the theoretical uncertainties on the top background prediction and to uncertainties on jets are ranked under the top ten NPs.

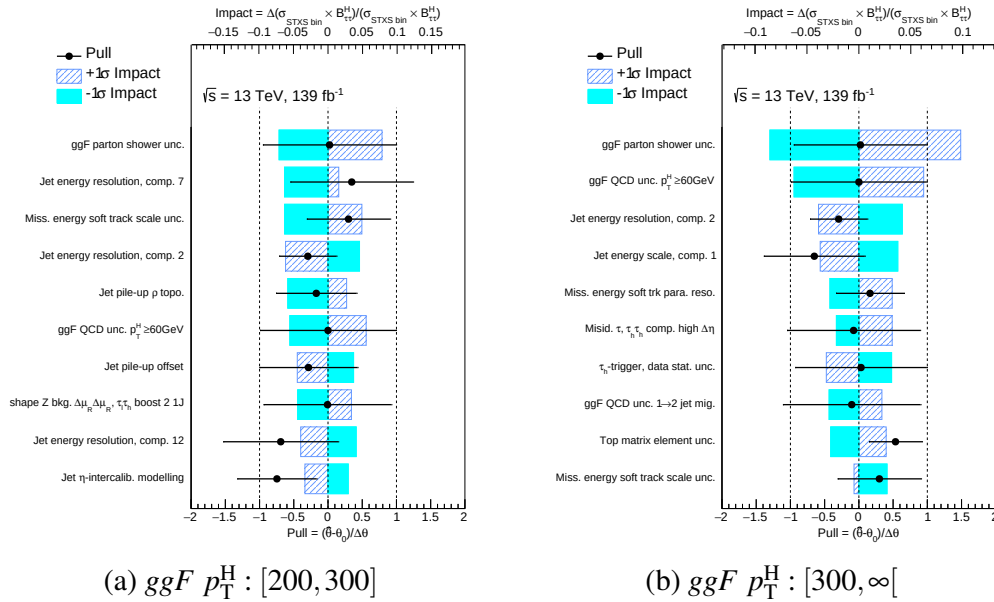


Figure 10.46: Relative impact of the ten NPs with the largest impact on the measurement of the ggF $p_T^H : [200, 300]$ (a) and ggF $p_T^H : [300, \infty[$ (b) STXS bins. The uncertainties are listed in decreasing order of their impact. The upper x-axis corresponds to the impact of the variation of the NPs by one- σ on the cross section measurements and is visualized by blue boxes. The lower x-axis corresponds to the pulls and constraints of the NPs, which are illustrated with black points and lines.

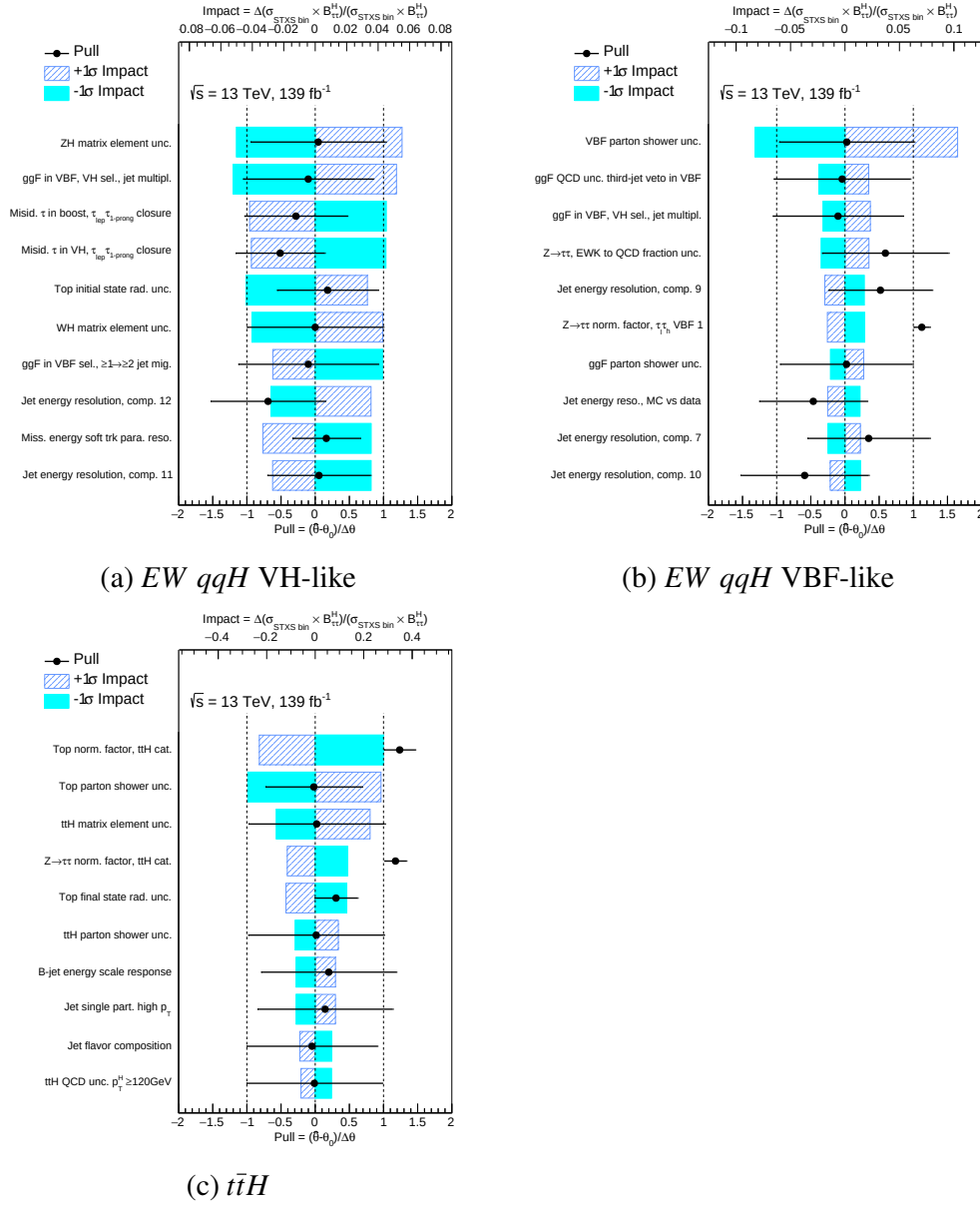


Figure 10.47: Relative impact of the ten NPs with the largest impact on the measurement of the EW qqH VH-like (a), EW qqH VBF-like (b) and $t\bar{t}H$ (c) STXS bins. The uncertainties are listed in decreasing order of their impact. The upper x-axis corresponds to the impact of the variation of the NPs by one- σ on the cross section measurements and is visualized by blue boxes. The lower x-axis corresponds to the pulls and constraints of the NPs, which are illustrated with black points and lines.

11 COMBINED HIGGS BOSON MEASUREMENTS

The two analyses described in Chapters 9 and 10 are used as input for combined Higgs boson measurements. Therefore, these analyses are combined with several other Higgs boson cross section measurements performed in other decay channels. The analyses are used for two different combined measurements (Comb1 [29, 30] and Comb2 [31]). Comb1 is based on analyses using the data sets from 2015 up to 2017 with an integrated luminosity of up to $\mathcal{L} = 79.8 \text{ fb}^{-1}$, while Comb2 uses analysis based on the data sets from 2015 up to 2018 with up to $\mathcal{L} = 139 \text{ fb}^{-1}$. The $H \rightarrow \tau\tau$ analysis described in Chapter 9 is used as input from Comb1, while the second $H \rightarrow \tau\tau$ analysis (see Chapter 10) is used from Comb2. In this chapter both analyses will be discussed and compared to each other.

11.1 Strategy of the combination

To be able to combine the individual Higgs boson measurements, several requirements have to be fulfilled. The individual event selections of the analyses are not allowed to have any overlap, to avoid the double use of data events. Therefore, analyses with similar final states topologies introduce inverted requirements to each other in the event selection (see Section 9.2.2). Dedicated overlap studies are performed to ensure the orthogonality.

The combinations of the Higgs boson cross section measurements are performed on likelihood level. Therefore, the likelihood model of each analysis is provided to the combination team. The inputs are combined by multiplying the individual likelihood functions in the same way as two likelihood functions of two different event selections in one analysis are combined (see Section 7.2). The best-fit values and their uncertainties are estimated by the minimization of the combined negative log-likelihood function. Therefore, the likelihood function of each analysis must have the same definition of PoIs, which should be measured in the combined fit. This is ensured by STXS framework and its common STXS bin definition. For Comb1, two different sets of likelihood functions are provided, where one is using the STXS Stage 0 definition for the PoIs only and one the STXS Stage 1 definition. For the Comb2 the STXS Stage 1.2 is used for all measurements. The benefit of using the STXS Stage 1.2 (or Stage 1) is the fine granularity of the PoIs, which are defined in the likelihood functions. If the combined model is not able to measure the full granularity, the PoIs can be correlated in the combined workspace to reduce the number of free parameters in the model. But a split of individual PoI in the combination to measure a finer granularity is not possible. Therefore, the used STXS Stage defines the finest granularity and all available borders of measured regions of phase in the combination. For the NPs, similar requirements have to be fulfilled. Several NPs as theoretical uncertainties on signal and background samples and experimental uncertainties arising from the ATLAS detector are common for all analyses. These uncertainties should be fully correlated across the different analyses in the combined likelihood function. For this, the sign of the impact of the up and down variations for each NP should be defined in the same way. Especially, for theoretical uncertainties as parton shower uncertainties, which are calculated by taking the difference between two different generators, the sign of the variations are chosen arbitrarily. For a correct combination of the NPs, a common strategy of defining

these variations was developed. The combination of the NPs in the combined likelihood is ensured by a common naming scheme for all NPs. The combined likelihood automatically correlates all NPs with the same name. On the other hand, NPs related to analysis specific uncertainties, as the uncertainties related to the misidentified τ background (see Section 10.4), should not have any correlation to NPs of other analyses. This is ensured by adding analysis specific suffixes to these NPs. NFs, which are defined to normalize background contributions in a specific region of phase space, are treated as uncorrelated in the same way.

11.2 Input for the combination

Table 11.1 summarizes all analyses used for the combined measurement discussed in this chapter. Comb1 corresponds to the combined measurement that uses the $H \rightarrow \tau\tau$ analysis with $\mathcal{L} = 36 \text{ fb}^{-1}$ described in Chapter 9 and Comb2 the one described in Chapter 10 with $\mathcal{L} = 139 \text{ fb}^{-1}$. Comb1 uses several analyses based on the data set recorded in 2015 and 2016 with an integrated luminosity of $\mathcal{L} = 36 \text{ fb}^{-1}$ and analyses using the data set taken from 2015 to 2017 with $\mathcal{L} = 79.8 \text{ fb}^{-1}$. For Comb2 most of the analyses are based on the full Run 2 data set with $\mathcal{L} = 139 \text{ fb}^{-1}$. Table 11.1 refers in addition to all publications of the individual analyses. Both combinations use e.g. analyses of the $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ^*$, $H \rightarrow b\bar{b}$, $H \rightarrow WW^*$ and $H \rightarrow \tau\tau$ decay modes. But also analyses to measure e.g. the $t\bar{t}H$ production mode and to study Higgs boson decays into particles which can not be measured by the ATLAS detector ($H \rightarrow \text{invisible}$). In the listed publications the used data samples and simulations, event selection and background estimation are discussed. This thesis does not discuss these important steps of the analyses, instead it will focus on the results of the combined measurement and their interpretations. In both combinations not all analyses are used for all measurements and interpretations. Some of the analyses provided a simplified likelihood model with a reduced set of PoIs and therefore could not be used for more differential measurements.

Analysis	Comb1			Comb2		
	Data set	$\mathcal{L} [\text{fb}^{-1}]$	Ref.	Data set	$\mathcal{L} [\text{fb}^{-1}]$	Ref.
(1) $H \rightarrow \gamma\gamma$	2015-17	79.8	[259]	2015-18	139	[260]
(2) $H \rightarrow ZZ^* \rightarrow 4l$ ($t\bar{t}H$, $H \rightarrow ZZ^* \rightarrow 4l$)	2015-17	79.8	[261]	2015-18	139	[262]
(3) VH , $H \rightarrow b\bar{b}$	2015-17	79.8	[263, 264]	2015-18	139	[265–267]
(4) $H \rightarrow \mu\mu$	2015-17	79.8	[268]	2015-18	139	[269]
(5) $H \rightarrow WW^* \rightarrow e\nu\mu\nu$	2015-16	36.1	[270]	2015-18	139	[271]
(6) $H \rightarrow \tau\tau$	2015-16	36.1	[26]	2015-18	139	[272]
(7) VBF , $H \rightarrow b\bar{b}$	2015-16	24.5-30.6	[273]	2015-18	126	[274]
(8) $t\bar{t}H$ multilepton	2015-16	36.1	[275]	2015-16	36.1	[275]
(9) $t\bar{t}H$, $H \rightarrow b\bar{b}$	2015-16	36.1	[276–278]	2015-18	139	[279]
(10) $H \rightarrow \text{invisible}$	2015-16	36.1	[280–283]	2015-18	139	[284]
(11) Off-shell $H \rightarrow ZZ^* \rightarrow 4l$ and $H \rightarrow ZZ^* \rightarrow 2l2\nu$	2015-16	36.1	[285]	-	-	

Table 11.1: Analyses used in the two combined measurements. For each input analysis the corresponding publications, the used data sets and the its integrated luminosities $\mathcal{L}$ are shown [29, 31].

11.3 Signal strength

The global signal strength measurement for both measurements follows the strategy described in Section 9.6. For a specific production mode i and decay mode j , the signal strength corresponds to a specific modifier μ_{ij} . The cross section σ_i and the branching ratio $\mathcal{B}_j$ cannot be measured separately without further assumptions in the combination. Therefore, the modifiers are defined as:

$$\mu_{ij} = \frac{\sigma_i^{\text{obs}}}{\sigma_i^{\text{SM}}} \times \frac{\mathcal{B}_j^{\text{obs}}}{\mathcal{B}_j^{\text{SM}}}, \quad (11.1)$$

where σ_i^{obs} and $\mathcal{B}_j^{\text{obs}}$ are the observed cross section of the production mode i and the observed branching ratio of decay channel j , while σ_i^{SM} and $\mathcal{B}_j^{\text{SM}}$ are their corresponding SM predictions. In both combinations, the analyses (1)-(9) in Table 11.1 are used to extract the signal strength. The measurement is performed under the assumption that the measured signal strength in each of the different input analyses is the same. The observed values for the signal strength measurement of Comb1 and Comb2 are:

$$\begin{aligned} \mu_{\text{Comb1}} &= 1.11 \pm 0.09 = \pm 0.05(\text{stat.}) \pm 0.05(\text{exp.}) \pm 0.05(\text{sig.}) \pm 0.03(\text{bkg.}), \\ \mu_{\text{Comb2}} &= 1.06 \pm 0.06 = \pm 0.03(\text{stat.}) \pm 0.03(\text{exp.}) \pm 0.04(\text{sig.}) \pm 0.02(\text{bkg.}), \end{aligned} \quad (11.2)$$

where (stat.) is the statistical uncertainty, (exp.) the uncertainty related to experimental uncertainties and (sig.) and (bkg.) the theoretical uncertainties on the signal and background prediction.

The individual impact of the different sets of uncertainties are estimated using the breakdown approach (see Section 7.2). Figure 11.1 shows the $-2 \ln \mathcal{L}$ functions for these measurements. By removing the NPs of the studied uncertainties, the $-2 \ln \mathcal{L}$ curve is tightened resulting in smaller values for the confidence intervals. The uncertainties from this set of NPs can be calculated with the quadratic difference between the nominal intervals including all uncertainties and the reduced intervals. The uncertainties on the Comb1 measurement is 8%, while for Comb2 it is reduced to 6%. The statistical and experimental uncertainties are 5% and 3% for Comb1 and Comb2. The reduction for the theoretical uncertainties on the signal and background is approximately 1%. The comparison of the individual sets of uncertainties shows the limitations of future coupling strength measurements. A further increase of data statistics and improved detector understanding can reduce the statistical and experimental uncertainties further, while the impact of theoretical uncertainties on the signal and background predictions can be reduced by updated predictions from theory groups only. The signal strength measurement of Comb2 is already limited by the uncertainties on theoretical predictions and therefore, it will not significantly benefit from the increased data statistics in upcoming measurements. However, a change of the analysis strategy from signal strength measurements to measurements of cross sections with fine granularity based on the STXS framework can reduce the impact of the theoretical uncertainties (see Section 7.1).

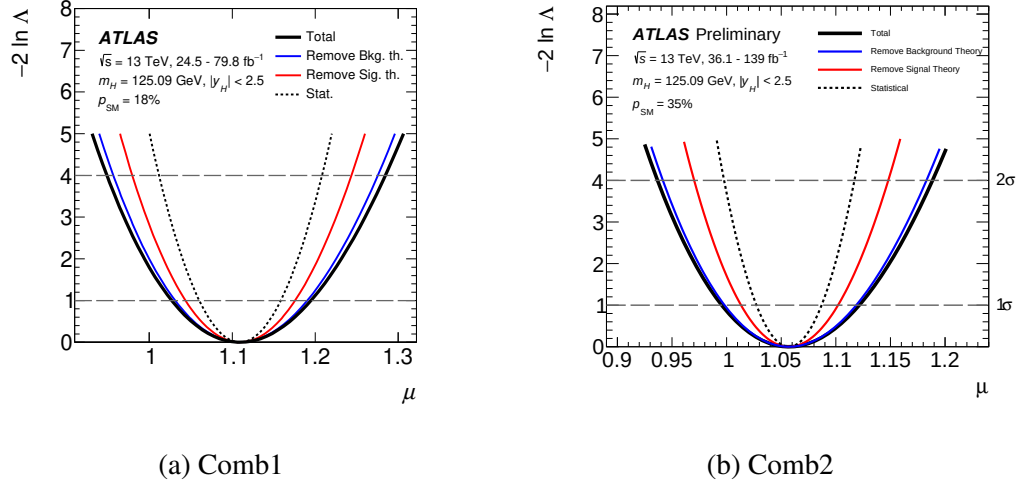
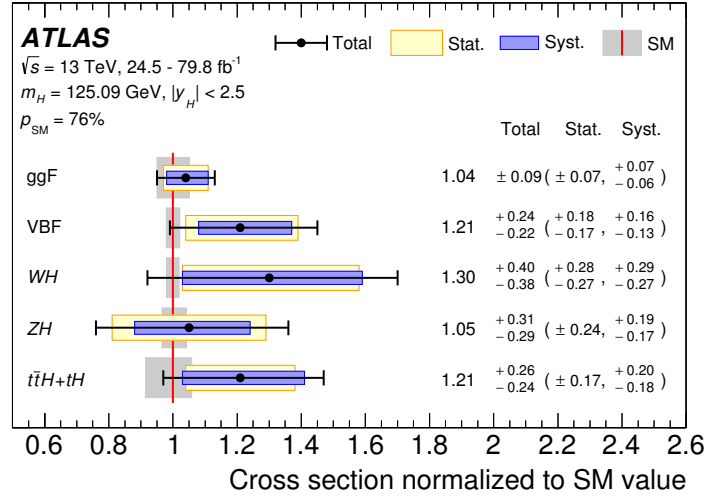


Figure 11.1: The $-2 \ln \mathcal{L}$ curves for the signal strength measurements of Comb1 (a) and Comb2 (b) for the nominal fit including all uncertainties (black) and for fits excluding individual sets of NPs to estimate the impact of them on the full uncertainty. For red (blue) the theoretical uncertainties on the signal (background) prediction are removed, while for the dashed line corresponds to a fit with statistical uncertainties only, in which all systematic uncertainties are removed. The two horizontal dashed lines mark the one- and two- σ confidence intervals [29, 31].

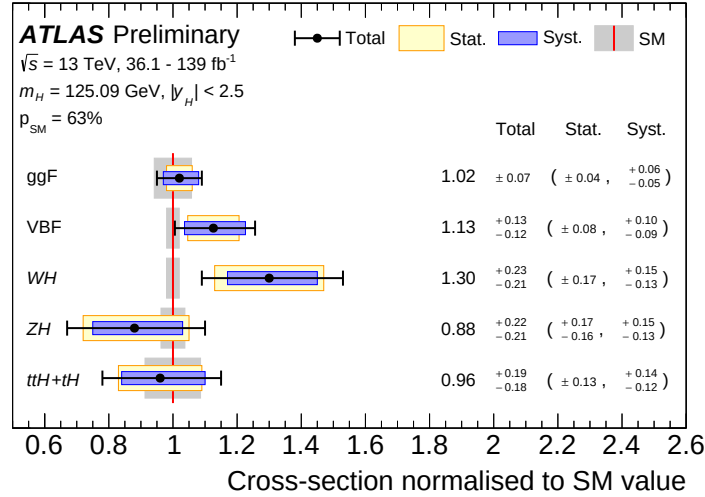
11.4 Production cross sections

For the inclusive cross section measurements of the Higgs boson production modes of Comb1 and Comb2 the analyses (1)-(9) in Table 11.1 are used. In both combinations the inclusive cross sections of the six dominant production modes ggF , VBF , ZH , WH , $t\bar{t}H$ and tH are measured, where the PoI on $t\bar{t}H$ and tH are fully correlated in the likelihood model, so that five PoIs are measured. The branching ratio for the individual decay modes are set constant to the SM prediction for this measurement.

Figure 11.2 shows the best-fit values and their uncertainties for the five measured cross sections of Comb1 and Comb2. Both analyses measure the cross section of the ggF production mode with the highest precision with 9% and 7% for Comb1 and Comb2, followed by the cross section of the VBF production mode with 19% and 12%. The measurement of the ggF cross section has the same impact of systematic and statistical uncertainties (7%) for Comb1, while for Comb2 the impact of systematic uncertainties (6%) is larger than the impact of statistical uncertainties (4%). The statistical uncertainties on the VBF measurements are 15% and 9% for the Comb1 and Comb2, while the systematic uncertainties are 12% and 9%. Therefore, the precision of the VBF measurement is limited by data statistics for Comb1 and by systematic uncertainties for Comb2. Relative uncertainties of 21% and 20% are estimated on the measurements of the $t\bar{t}H + tH$ cross section for Comb1 and Comb2. For Comb1



(a) Comb1



(b) Comb2

Figure 11.2: The observed best-fit values and their uncertainties for measurements of inclusive cross sections of Comb1 (a) and Comb2 (b). The measured cross sections are visualized normalized to their SM predictions and the statistical (systematic) uncertainties on them are shown in yellow (blue). The total uncertainties are given by the black bars and the uncertainties on the SM predictions are visualized by gray boxes [29, 31].

the statistical uncertainties (14%) on the $t\bar{t}H$ cross section are smaller than the systematic uncertainties (16%), while they have the same size for Comb2 (14%). The measurements of the WH and ZH cross sections have the largest uncertainties in Comb1 with 30% and 29%.

In Comb2 the uncertainties on the WH and ZH measurements are 17% and 25%. In contrast to Comb1, is the precision of the WH measurement in Comb2 better than for the $t\bar{t}H + tH$ measurement. The statistical and systematic uncertainties for the WH measurements are both 22% for Comb1, while for Comb2 they are 13% and 11%. So that in both measurements the impacts of systematic and statistical uncertainties are similar to each other. The measurements of the ZH cross section for Comb1 and Comb2 have statistical uncertainties of 23% and 19% and systematic uncertainties of 17% and 16%. The precision of the measurement of the ZH cross section is limited by data statistics for Comb1, while in Comb2 statistical and systematic uncertainties have similar impacts on the measurement.

The largest improvements between the two combinations are observed for the measurements of the WH and the VBF cross section, while the $t\bar{t}H + tH$ cross section is measured with a similar precision in both combinations. The precision of the measurement of the $t\bar{t}H$ cross section is driven by the $t\bar{t}H$ multi-lepton analysis, which is unchanged between the two combinations (see Table 11.1). A large improvement is observed for the measurements of the VBF and ZH cross sections between the two combinations. Both measurements are limited by data statistics in Comb1, while the increased data set used for Comb2 reduces the impact of the data statistics. The measurement of the ggF cross section, which has a large contribution to the total uncertainty from theoretical uncertainties on the signal predictions, is only slightly improved between the two combinations.

Compatibility tests between the five measured cross sections of Comb1 and Comb2 and their SM predictions estimate p -values of 76% and 63%. Therefore, no significant deviation from the SM predictions is observed in both measurements.

Figure 11.3 displays the correlation matrices of the Comb1 and Comb2 measurements. For

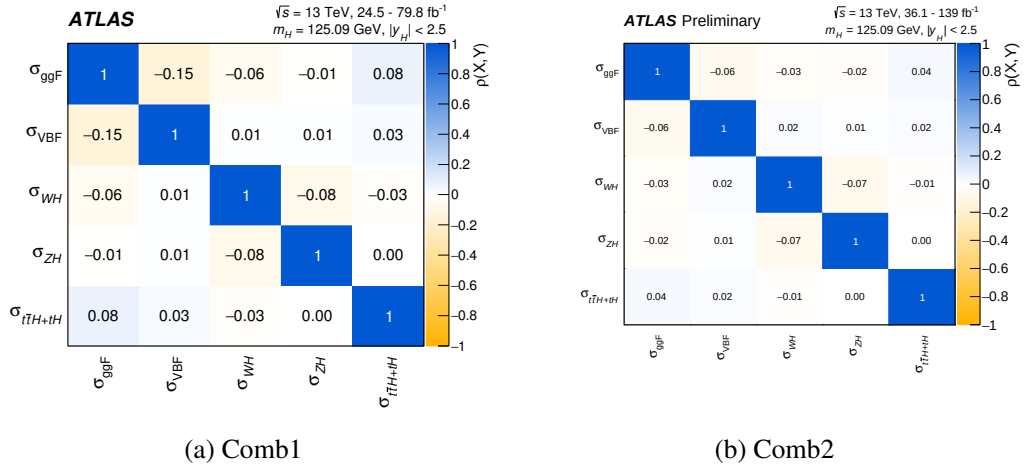


Figure 11.3: The observed correlation matrices for the 5-PoI measurements of Comb1 (a) and Comb2 (a) [29, 31].

Comb1 the largest correlation is observed between the cross sections of the VBF and ggF productions modes (-15%), followed by the correlation between the cross sections of the WH

and ZH production modes (-8%) and between the cross sections of ggF and $t\bar{t}H + tH$ (+8%). For Comb2, the largest correlation is observed between the cross sections of the WH and ZH production modes with -7% and between the cross sections of the VBF and ggF production modes with -6%. The correlations of the remaining measured cross sections are between -6% and +3% for Comb1 and between -3% and +4% for Comb2. Overall, the correlation across the different production modes is reduced for Comb2 compared to Comb1. Especially, the anti-correlation between the ggF and VBF production mode could be reduced from -15% to -6%.

11.5 Product of production cross sections and branching ratios

The measurements of the inclusive production mode cross sections in Section 11.4 assume that the branching ratios for individual decay channels behave SM-like. In the measurement discussed in this section, the product of the production mode cross section and the branching ratios of the decay modes $\sigma_i^{\text{obs}} \times \mathcal{B}_j^{\text{obs}}$ are defined as PoIs. Therefore, one PoI per production mode and per decay channel is included in the likelihood model. For each analysis, cross sections with low sensitivity in the measurement are fixed to their SM prediction. Most of the analyses used a new event selection for the Comb2 compared to Comb1 covering a larger region of phase space, so that the number of PoIs increases for Comb2. New event categories have been included e.g. in the $H \rightarrow \tau\tau$ analysis covering the VH and $t\bar{t}H$ production modes (see Section 10.2.3).

The best-fit values and their uncertainties for the measurements of $\sigma_i^{\text{obs}} \times \mathcal{B}_j^{\text{obs}}$ are shown in Figure 11.4 for Comb1 and in Figure 11.5 for Comb2. For Comb1 and Comb2 the ggF production mode is measured with the highest precision in the $\gamma\gamma$ (15% and 11%), ZZ^* (15% and 12%) and WW^* (18% and 12%) decay modes of the Higgs boson. The measurement of the VBF cross section has the smallest uncertainties in the $\gamma\gamma$ (27%) and ZZ^* (34%) decay modes for Comb1, while for Comb2 the $\tau\tau$ (19%) and WW^* (17%) decay modes have a similar precision as the $\gamma\gamma$ (18%) decay mode for Comb1. For the VH production mode the highest precision is observed in the bb decay mode of the Higgs boson. For Comb1, the VH production mode in bb is measured with an uncertainty of 22%, while for Comb2 the ZH and WH production modes are measured separately in bb with uncertainties of 26% and 23%. The $t\bar{t}H + tH$ production modes is measured in the $\gamma\gamma$ decay mode with an uncertainty of 35% for Comb1 and 28% for Comb2.

The measurement of the ggF cross section in the $\gamma\gamma$ and ZZ^* decay modes is limited by data statistics in both combinations, while the measurement in WW^* and $\tau\tau$ is limited by systematic uncertainties. For both combinations, the uncertainties on the measurements of the VBF production are dominated by statistical uncertainties in all decay modes. For both combinations, the precision of the measurements of the VH and $t\bar{t}H + tH$ production modes is limited by systematic uncertainties in the bb decay mode, while in the remaining decay channels the statistical uncertainties limit the precision, except for the $t\bar{t}H + tH$ production mode in WW^* for Comb2.

The size of the uncertainty reduction on $\sigma_i^{\text{obs}} \times \mathcal{B}_j^{\text{obs}}$ between Comb1 and Comb2 differs

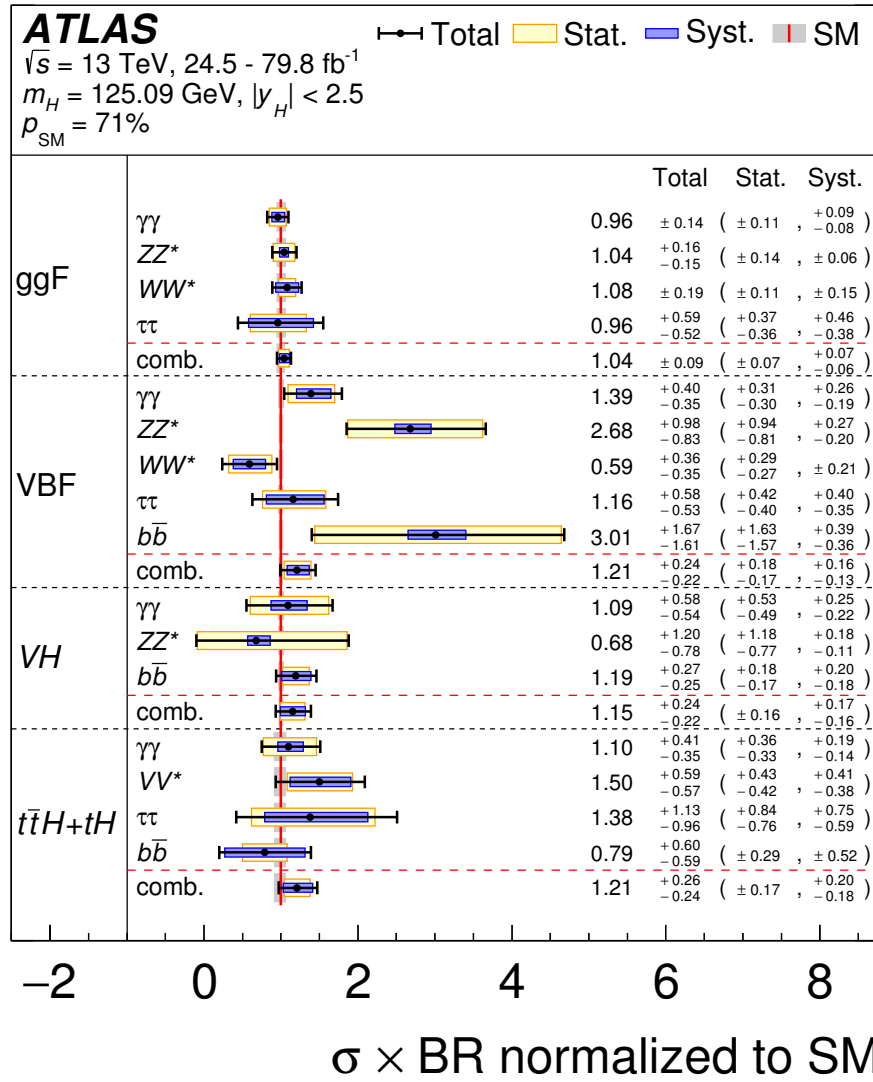


Figure 11.4: Best-fit values and uncertainties for the measurement of inclusive production modes cross sections multiplied by individual decay mode branching ratios ($\sigma_i^{\text{obs}} \times \mathcal{B}_j^{\text{obs}}$) for Comb1. The results are visualized normalized to their SM predictions. The total uncertainties are shown with a black bars, the statistical uncertainties with yellow boxes and the systematic uncertainties with blue boxes. The uncertainties on the SM predictions are marked by a gray boxes [29].

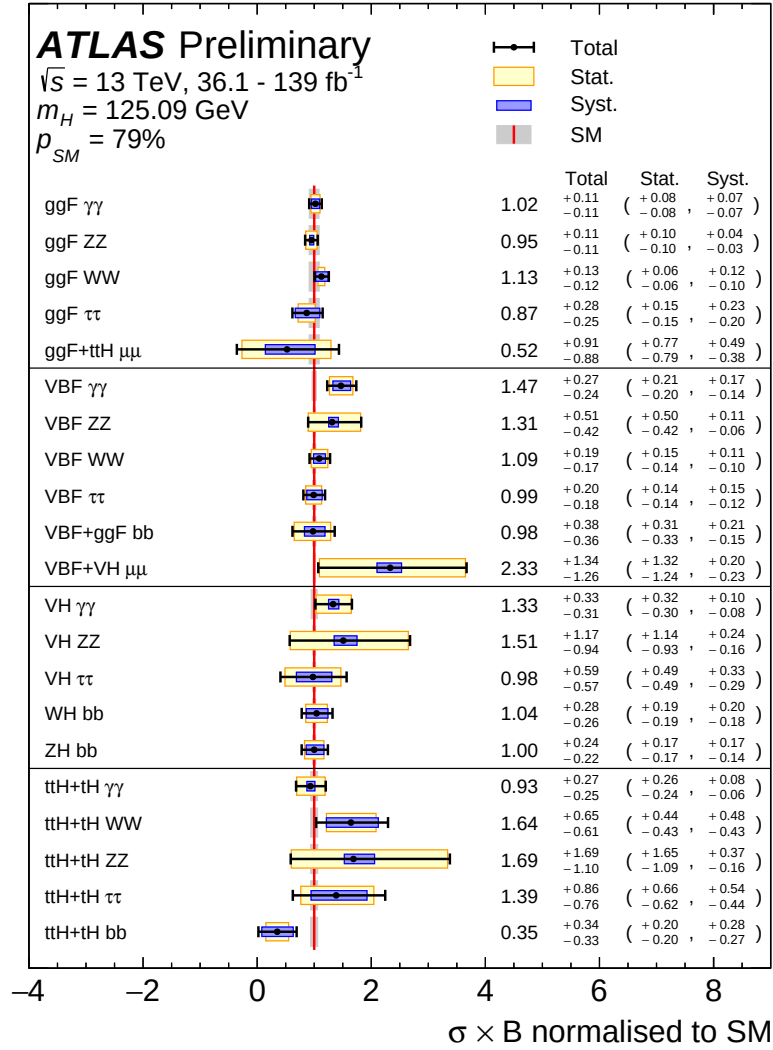


Figure 11.5: Best-fit values and uncertainties for the measurement of inclusive production modes cross sections multiplied by individual decay mode branching ratios ($\sigma_i^{\text{obs}} \times \mathcal{B}_j^{\text{obs}}$) for Comb2. The results are visualized normalized to their SM predictions. The total uncertainties are shown with a black bars, the statistical uncertainties with yellow boxes and the systematic uncertainties with blue boxes. The uncertainties on the SM predictions are marked by a gray boxes [31].

between the different production and decay modes. The largest reductions are visible for measurement of $\sigma_i^{\text{obs}} \times \mathcal{B}_j^{\text{obs}}$ with uncertainties dominated by statistical uncertainties and for $\sigma_i^{\text{obs}} \times \mathcal{B}_j^{\text{obs}}$ measurements which are estimated in the analyses with the largest increase of data statistics between Comb1 and Comb2. In particular, the measurement of $\sigma_{VH}^{\text{obs}} \times \mathcal{B}_{ZZ^*}^{\text{obs}}$ is dominated by statistical uncertainties and its uncertainty is strongly reduced from 146% to 70% from Comb1 to Comb2. Similarly, the uncertainty of $\sigma_{VBF}^{\text{obs}} \times \mathcal{B}_{\tau\tau}^{\text{obs}}$, which is mainly extracted in the $H \rightarrow \tau\tau$ analysis, is reduced from 50% to 19%. Here, the statistical uncertainty on the measurement is not so dominating as for $\sigma_{VH}^{\text{obs}} \times \mathcal{B}_{ZZ^*}^{\text{obs}}$, but the $H \rightarrow \tau\tau$ analysis increased the data statistics from 36 fb^{-1} to 139 fb^{-1} between Comb1 and Comb2.

Both measurements show no significant deviation from the SM predictions and their compatibilities with the SM predictions are 71% (Comb1) and 79% (Comb2).

The correlation matrices for both measurements are shown in Figure 11.6. The largest (anti-)correlations for the Comb1 measurement are -45% for the $t\bar{t}H + tH$ production in VV^* and $\tau\tau$, which are both mainly measured by the analysis (9) of Table 11.1 and -44% for the ggF and VBF cross sections in the $\tau\tau$ decay channel, which are both mainly measured in analysis (6). Similarly, sizable correlations are observed between the ggF and VBF cross sections in $\gamma\gamma$ (-22%) and ZZ^* (-28%) for Comb1, while the correlations between the remaining parameters are between -14 and +10%.

The largest (anti-)correlation in Comb2 is observed for the ggF and VH production mode in the $\tau\tau$ decay mode with -33%. No corresponding value is shown for Comb1, since the VH production mode in the $\tau\tau$ decay mode was set constant to the SM prediction in Comb1. For Comb2, the anti-correlation between the $t\bar{t}H + tH$ production cross sections in VV^* and in $\tau\tau$ is -32% and between the ggF and VBF cross sections in the $\tau\tau$ decay channel -27%. For Comb2, the new $H \rightarrow \tau\tau$ analysis using the signal event selections described in Section 10.2.4 is used as an input. These selections reduce the contamination of ggF events in the signal categories designed for the VBF signal compared to the previous analysis used for Comb1. The smaller contamination reduces the observed correlation between the two production modes in the $\tau\tau$ decay mode for Comb2 compared to Comb1. Similar optimizations are performed in other input analyses between Comb1 and Comb2, so that overall the (anti-)correlations, observed in Figure 11.6, are reduced for Comb2 compared Comb1.

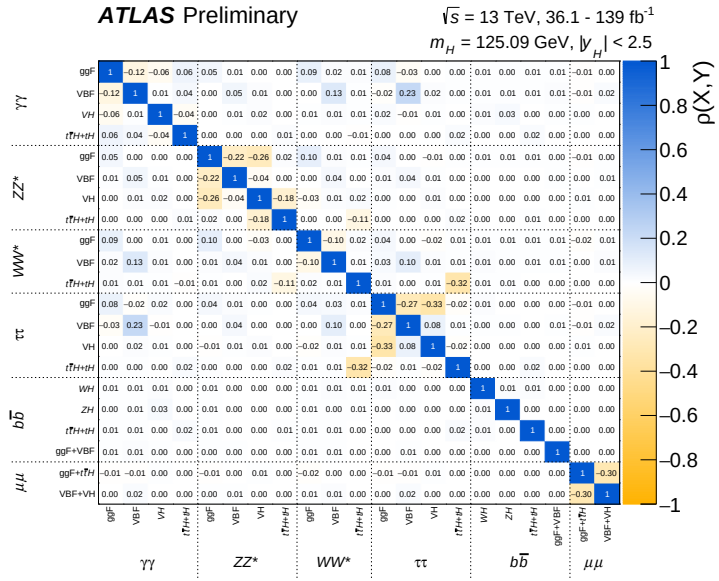
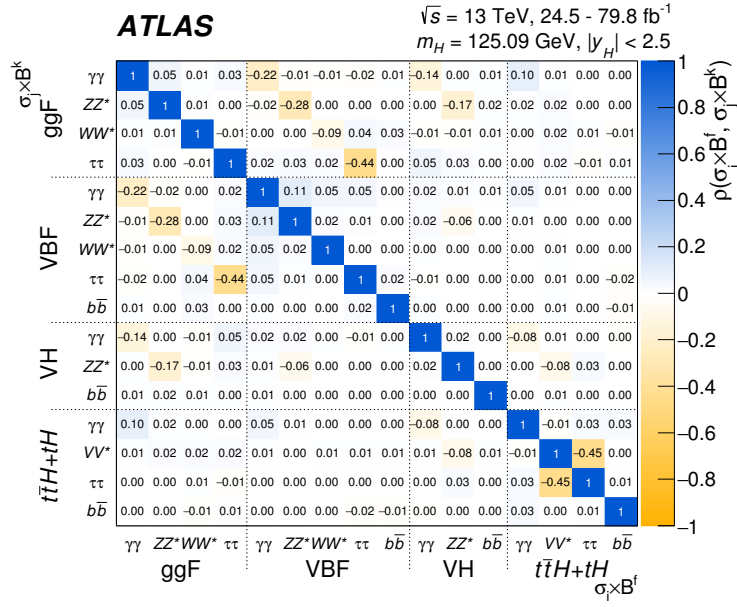


Figure 11.6: Correlation matrices for the measurement of inclusive production modes cross sections multiplied by individual decay mode branching ratios ($\sigma_i^{\text{obs}} \times \mathcal{B}_j^{\text{obs}}$) for Comb1 (a) and Comb2 (b) [29, 31].

11.6 STXS measurement

For Comb1 and Comb2, a measurement of cross sections with a finer granularity is performed using the STXS framework. For Comb1 the measurement is based on the STXS Stage 1, while Comb2 uses the STXS Stage 1.2. For both combinations several STXS bins have been combined to reduce the number of PoIs in the measurement. For Comb2 the analyses (1)-(3), (5)-(7) and (9) of Table 11.1 are used for the STXS measurement, while for Comb1 only the analyses (1)-(3) and (5)-(7) are used. The remaining input analyses were not able to include the needed STXS information in their likelihood model until these combination were performed. The branching ratios of the individual decay modes are free floating in both measurements. To make use of the well known branching ratio of the ZZ^* decay mode a re-parametrization of the PoIs and NFs is performed. Therefore, the PoIs are defined as the cross section of the STXS bins multiplied by the ZZ^* branching ratio: $\text{PoI}_i = \sigma_i \times \mathcal{B}_{ZZ^*}$, where σ_i is the cross section of the STXS bin i . By introducing these PoIs the NFs on the branching ratios for the remaining decay channel must be replaced by $\mathcal{B}_j/\mathcal{B}_{ZZ^*}$, where $\mathcal{B}_j$ is the branching ratio of the decay channel j .

Figure 11.7 summarizes the definition of the PoIs for Comb1. 15 PoIs are defined in the likelihood function, where six PoIs corresponds to the ggF , three to the $EW\ qqH$, five to the $V(\rightarrow ll)H$ and one to the $t\bar{t}H + tH$ production mode. For the ggF production mode, jet multiplicity and p_T^H STXS bins are defined. The $EW\ qqH$ production is split into a high p_T^j , a VH -like and a $VBF + \text{Rest}$ -like STXS bin. The $V(\rightarrow ll)H$ production modes has individual STXS bins defined for the W - and Z -boson production mode and for different p_T^V values. The cross sections of the remaining STXS bins (see Section 7.1) are set constant to their SM predictions in the likelihood function.

Figure 11.8 shows the definition of the PoIs for Comb2. In this fit based on the STXS Stage 1.2 bins 37 independent PoIs are defined. For ggF 13 PoIs are defined, for $EW\ qqH$ eight, for $V(\rightarrow ll)H$ nine, for $t\bar{t}H$ six and for tH one. Compared to Comb1 the PoIs for the ggF production modes have been changed in the following way:

- The ggF , 0-jet STXS bin is split into two bins at a p_T^H value of 10 GeV.
- The ggF , ≥ 2 -jet, $p_T^H < 200$ GeV STXS bin is split into five STXS bins based on m_{jj} and p_T^H .
- The ggF , ≥ 1 -jet, $p_T^H \geq 200$ GeV STXS bin is split into three STXS bins based on p_T^H .

The definition of the $EW\ qqH$ bins are changed as:

- The $p_T^j \geq 200$ GeV STXS bin is redefined to a STXS bin with $p_T^H \geq 200$ GeV and $m_{jj} \geq 350$ GeV.
- The $VBF + \text{Rest}$ -like STXS bin is split into four VBF -like STXS bins with $p_T^H < 200$ GeV and $m_{jj} \geq 350$ GeV with splits based on m_{jj} and two Rest bins with one including events with ≤ 1 -jet and one with ≥ 2 -jet and $m_{jj} < 60$ GeV or $120 \leq m_{jj} < 350$ GeV.

For the $V(\rightarrow ll)H$ production mode the granularity for the W - and Z -boson production in p_T^V has been increased from 5 to 9 STXS bins. The inclusive $t\bar{t}H + tH$ STXS bin has been

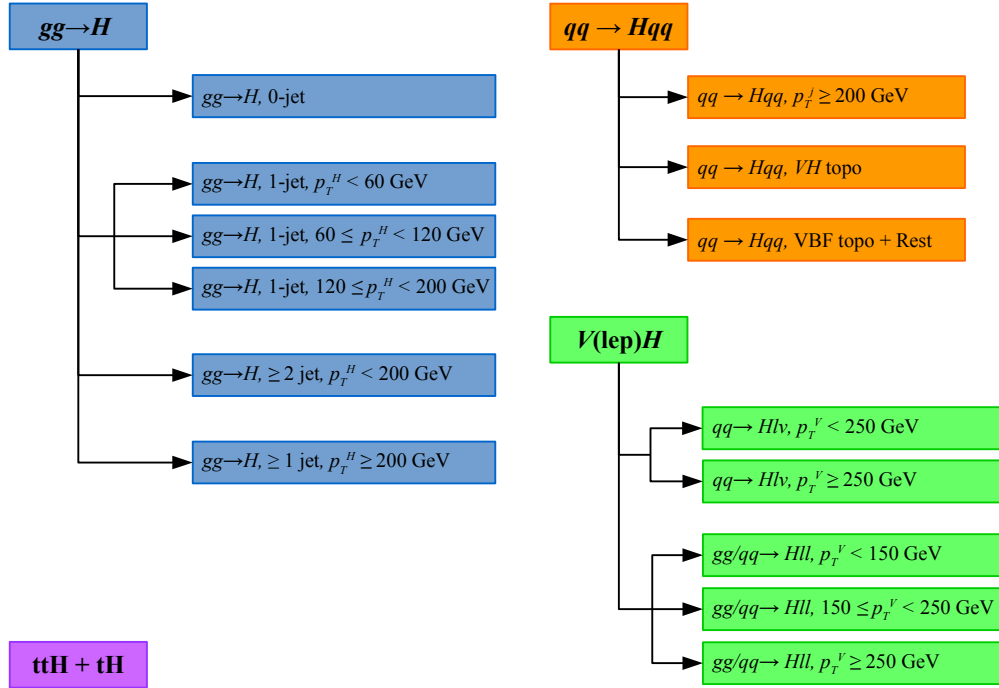


Figure 11.7: Final set of PoIs for the Comb1 STXS measurement based on the STXS Stage 1. The STXS bins are grouped by their production modes, where blue corresponds to ggF , orange to $EW\ qqH$, green to $V(\rightarrow ll)H$ and purple to $t\bar{t}H + tH$. Not listed STXS bins are set constant to their SM prediction in the likelihood model [29].

separated into one tH and six $t\bar{t}H$ STXS bins depending on p_T^H .

Figures 11.9 and 11.10 show the best-fit values and uncertainties of the PoIs for the STXS measurement of Comb1 and Comb2. No significant deviation from the SM predictions is visible and the compatibility for the Comb1 measurement with the SM prediction is 89% and for the Comb2 measurement 92%.

For Comb1 the best precision is estimated for the ggF , 0-jet STXS bin with a relative uncertainty of 13%, while the largest relative uncertainty is estimated for the $EW\ qqH$, $p_T^j > 200$ GeV STXS bin with 158%. For most of the STXS bins the statistical uncertainty is dominating in the Comb1 measurement. The highest relative impact of systematic uncertainties is visible for the $EW\ qqH, VBF + Rest$ -like and $t\bar{t}H + tH$ STXS bin. For both bins the granularity has been increased for the Comb2 measurement to reduce the impact of theoretical uncertainties on the signal in these STXS bins.

For the Comb2 measurement the smallest relative uncertainty is estimated for the ggF , 0-jet, $p_T^H \geq 10$ GeV STXS bin with 13% and therefore has the same relative uncertainty as the inclusive ggF , 0-jet STXS bin in the Comb1 measurement. The STXS bin with the largest

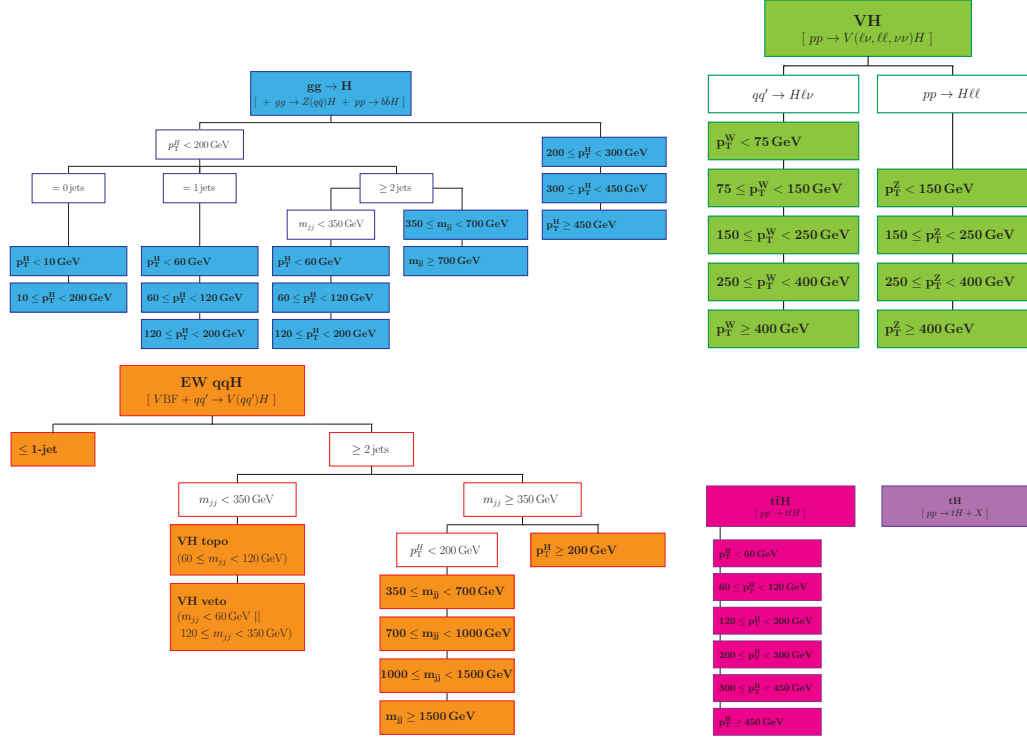


Figure 11.8: Final set of PoIs for the Comb2 STXS measurement based on the STXS Stage 1.2. The STXS bins are grouped by the production modes, where blue corresponds to ggF , orange to $EW\ qqH$, green to $V(\rightarrow ll)H$, pink to $t\bar{t}H$ and purple to the tH production. Not listed STXS bins are set constant to the SM prediction in the likelihood model [31].

relative uncertainty is the $t\bar{t}H$, $p_T^H \geq 450$ GeV STXS bins with over 1000%. This size of the relative uncertainty is mainly driven by a large downwards fluctuation of data events in this region of phase space, so that the observed best-fit is at 16% of the cross section predicted by the SM. All of the measured STXS bins in Comb2 are dominated by statistical uncertainties. The highest relative impacts of systematic uncertainties are estimated for the ggF , 0-jet, $p_T^H \geq 10$ GeV and ggF , 1-jet, $p_T^H < 60$ GeV STXS bins (see Figure 11.10).

Figures 11.11 and 11.12 show the correlation matrices for the STXS Comb1 and STXS Comb2 measurements. For Comb1 the correlations vary between -73% and +47%, while for Comb2 they vary between -67% and +57%. In both measurement, the largest correlations are visible across the NFs on the branching ratios, while large anti-correlations are visible between these NFs and most of the PoIs. For both measurement, correlations between -27% and +44% are visible across the different STXS bins of the ggF production mode. Furthermore, both measurements show an anti-correlation between the $EW\ qqH$ and the ggF , ≥ 1 -jet STXS bins. A mixture of stronger anti-correlations and correlations are visible between the STXS bins of the $V(\rightarrow ll)H$ production mode.

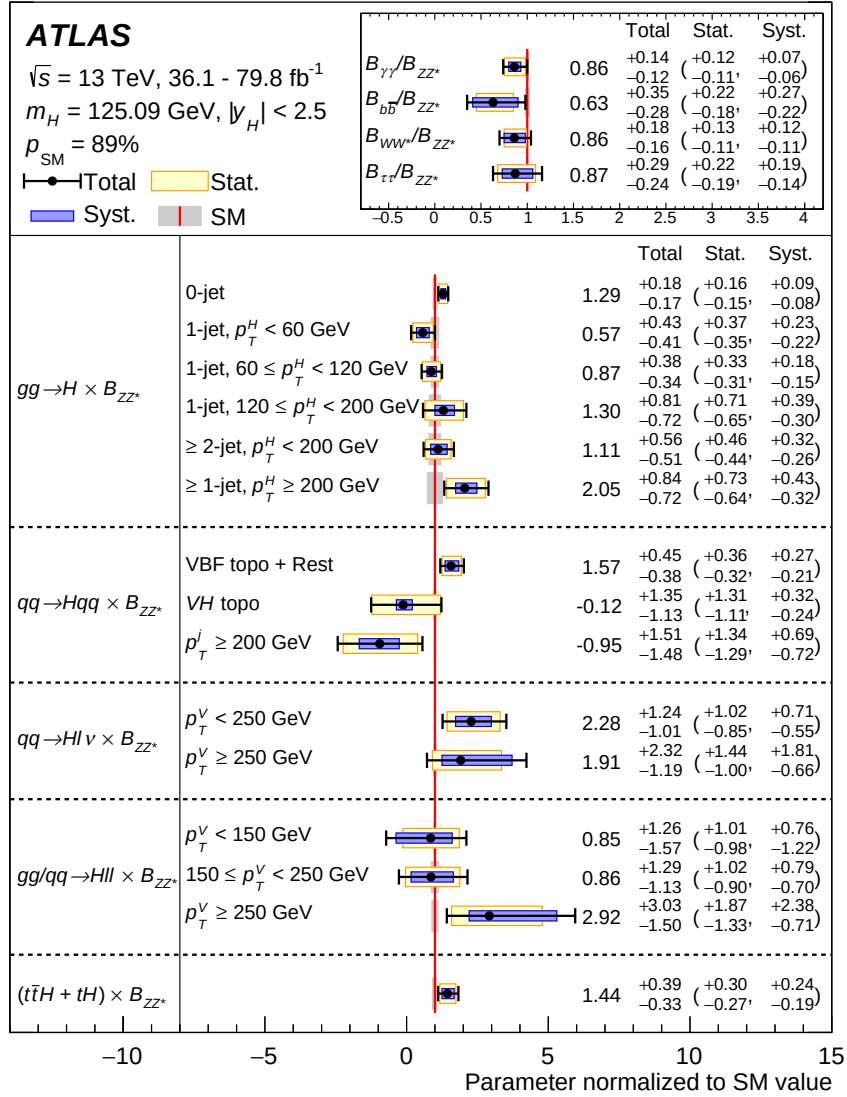


Figure 11.9: Best-fit values and uncertainties for the STXS measurement and the parametrized branching ratio (top right) of Comb1. The results are visualized normalized to the SM prediction. The total uncertainty is shown with a black bar, the statistical uncertainty with a yellow box and the systematic uncertainty with a blue box. On the right-hand side, the best-fit values, the total, statistical and systematic uncertainties are given. The uncertainty on the SM prediction is marked by a gray box [29].

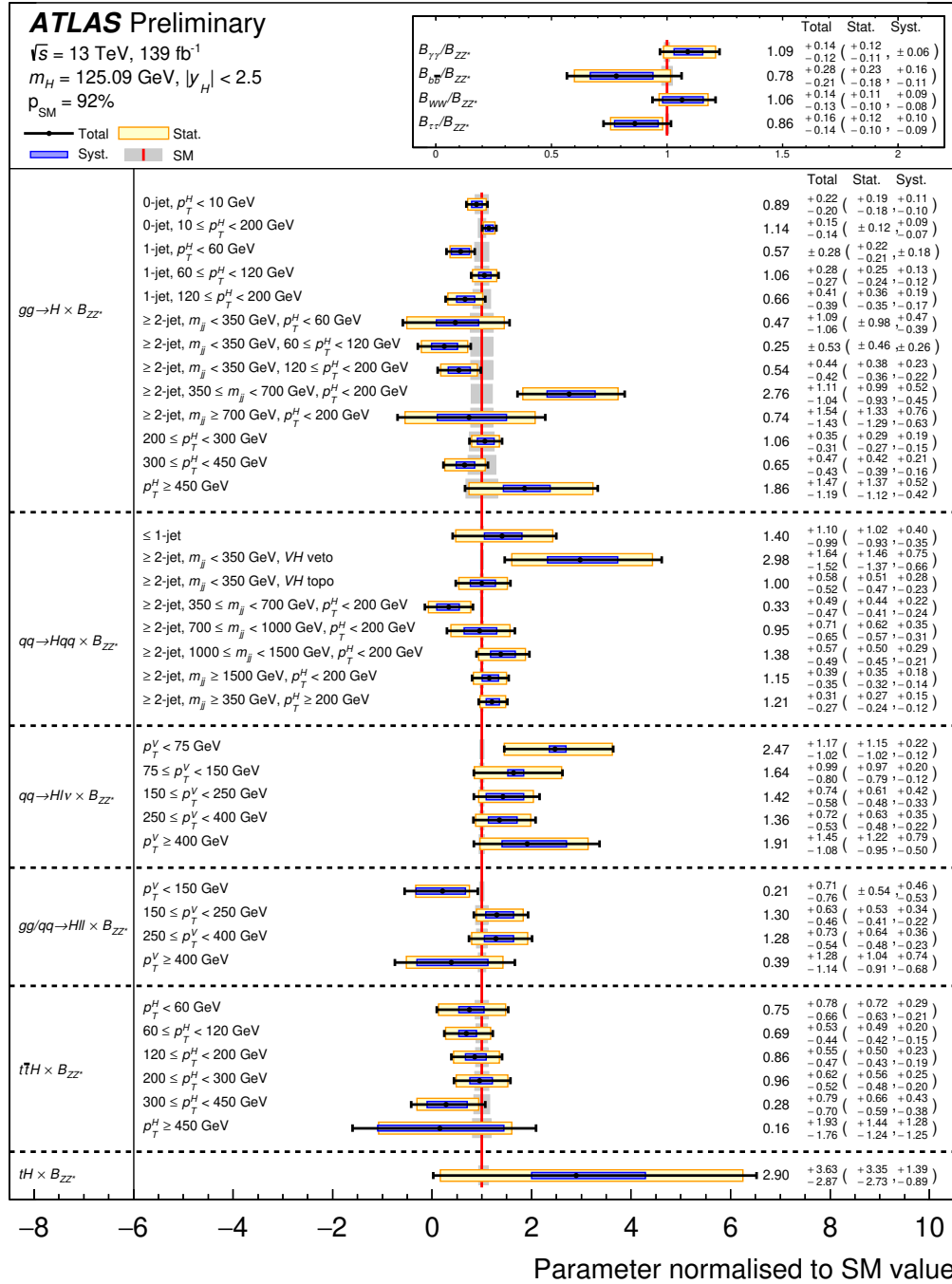


Figure 11.10: Best-fit values and uncertainties for the STXS measurement and the parametrized branching ratio (top right) of Comb2. The results are visualized normalized to the SM prediction. The total uncertainty is shown with a black bar, the statistical uncertainty with a yellow box and the systematic uncertainty with a blue box. On the right-hand side, the best-fit values, the total, statistical and systematic uncertainties are given. The uncertainty on the SM prediction is marked by a gray box [31].

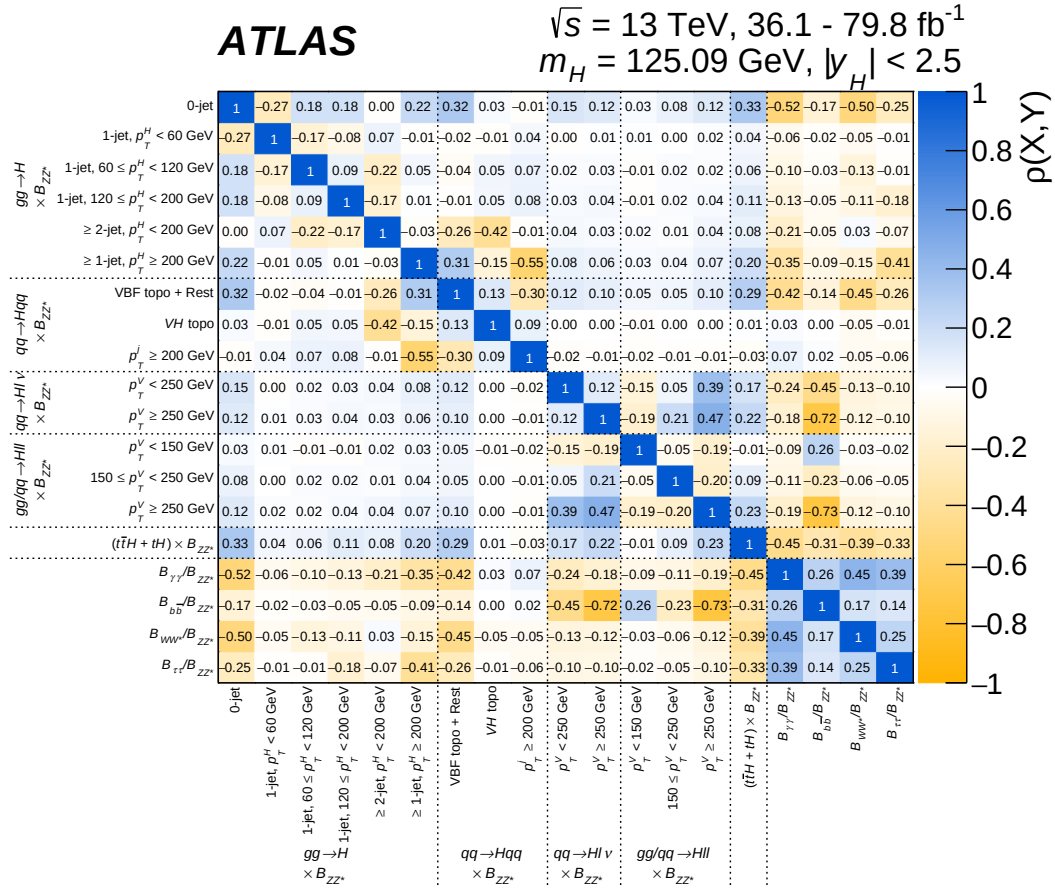


Figure 11.11: Correlation matrices of the STXS Comb1 measurement for the PoIs and the NFs on the branching ratios [29].



11.7 SMEFT interpretation

The STXS measurements described in Section 11.6 can be used for an interpretation within SMEFT (see Section 2.5). Both combinations, Comb1 and Comb2, are used for such an interpretation and are described in detail in References [30, 31]. Both interpretations are based on the Warsaw basis [77], which forms a complete set of dimension six operators, while operators with higher dimensions are not considered in the interpretations. To reduce the number of operators lepton and baryon conservation as well as SM gauge symmetries are assumed. Furthermore, only operators are considered, which are CP even operators and have a contribution of at least 0.1% on at least one measured STXS bin at their SM prediction corresponding to a Wilson coefficient $c_i=1$ and assuming $\Lambda = 1$ TeV. The goal of the measurements is to constrain the corresponding Wilson coefficients of all operators which either directly or indirectly impact the Higgs boson couplings within the SM [286, 287].

The most relevant SMEFT operators impacting the Higgs boson production and decay in the considered region of phase space are defined in Table 11.2, together with their corresponding Wilson coefficients and examples of Feynman diagrams in which they play a role. The full set of the 32 Wilson coefficients c_j considered in both interpretations is summarized in the appendix in Table A.6. The impact on different SM processes of each operator varies. Therefore, the operators can be grouped into different sets of operators based on the SM processes with highest sensitivity to the operators. One group can be defined assigned to the Higgs boson process. This group includes e.g. the corresponding operators of the Wilson coefficients c_{HW} , c_{eH} , c_{uG} and c_{dH} . These operator directly impact the vertex in which the Higgs boson interacts (see Figure 11.2). A second group can be defined for Z,W boson couplings. E.g. the operators of the Wilson coefficients $c_{Hl}^{(1)}$, c_{He} , $c_{Hq}^{(3)}$ and c_{Hd} are assigned to this group. This group acts on processes in which a Z or W boson coupling is involved. A third group can be defined to the top quark. This group includes e.g. the operators of the Wilson coefficients c_{qq} , c_{uu} and $c_{ud}^{(8)}$. These operators are mainly acting on vertices in which top quarks are involved. Furthermore, other groups can be defined as e.g. assigned to the Bhabha scattering, multi-boson couplings or to B anomalies.

In both analyses the impact of operators is estimated via the UFO model implemented in Madgraph [136] using the lowest order calculations in QCD for all production and decay modes. SMEFTsim [288, 289] is used to calculate the Higgs boson production modes with tree-level diagrams, assuming a $U(3)^5$ flavor symmetry. This symmetry corresponds to the unbroken flavor symmetry present in the SM outside the Yukawa sector. The Fermi constant as well as the Z and W boson masses are used as an input for these simulations. The ggF , $gg \rightarrow ZH$ and $H \rightarrow gg$ cross sections are calculated at next-to-leading order accuracy in QCD with SMEFTatNLO [290]. The SMEFT-SM interference terms in $H \rightarrow \gamma\gamma$ are calculated at next-to-leading order accuracy in QED [291]. No modification of background processes due to SMEFT effects are considered in the analyses.

The measured cross sections of the STXS bins σ_{STXS} are related to the EFT operators via the squared matrix element, so that these cross sections can be parametrized as:

$$\sigma_{\text{STXS}} = \sigma_{\text{SM}} + \sigma_{\text{int}} + \sigma_{\text{BSM}}, \quad (11.3)$$

Coefficient	Operator	Example process
c_{HDD}	$(H^\dagger D^\mu H)^* (H^\dagger D_\mu H)$	
c_{HG}	$H^\dagger H G_{\mu\nu}^A G^{A\mu\nu}$	
c_{HB}	$H^\dagger H B_{\mu\nu} B^{\mu\nu}$	
c_{HW}	$H^\dagger H W_{\mu\nu}^I W^{I\mu\nu}$	
c_{HWB}	$H^\dagger \tau^I H W_{\mu\nu}^I B^{\mu\nu}$	
c_{eH}	$(H^\dagger H)(\bar{l}_p e_r H)$	
$c_{Hl}^{(1)}$	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{l}_p \gamma^\mu l_r)$	
$c_{Hl}^{(3)}$	$(H^\dagger i \overleftrightarrow{D}_\mu^I H)(\bar{l}_p \tau^I \gamma^\mu l_r)$	
c_{He}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{e}_p \gamma^\mu e_r)$	
$c_{Hq}^{(1)}$	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{q}_p \gamma^\mu q_r)$	
$c_{Hq}^{(3)}$	$(H^\dagger i \overleftrightarrow{D}_\mu^I H)(\bar{q}_p \tau^I \gamma^\mu q_r)$	
c_{Hu}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{u}_p \gamma^\mu u_r)$	
c_{Hd}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{d}_p \gamma^\mu d_r)$	

Coefficient	Operator	Example process
c_{uG}	$(\bar{q}_p \sigma^{\mu\nu} T^A u_r) \tilde{H} G_{\mu\nu}^A$	
c_{uW}	$(\bar{q}_p \sigma^{\mu\nu} u_r) \tau^I \tilde{H} W_{\mu\nu}^I$	
c_{uB}	$(\bar{q}_p \sigma^{\mu\nu} u_r) \tilde{H} B_{\mu\nu}$	
$c_{qq}^{(1)}$	$(\bar{q}_p \gamma_\mu q_t)(\bar{q}_r \gamma^\mu q_s)$	
$c_{qq}^{(3)}$	$(\bar{q}_p \gamma_\mu \tau^I q_r)(\bar{q}_s \gamma^\mu \tau^I q_t)$	
c_{qq}	$(\bar{q}_p \gamma_\mu q_t)(\bar{q}_r \gamma^\mu q_s)$	
$c_{qq}^{(31)}$	$(\bar{q}_p \gamma_\mu \tau^I q_t)(\bar{q}_r \gamma^\mu \tau^I q_s)$	
c_{uu}	$(\bar{u}_p \gamma_\mu u_r)(\bar{u}_s \gamma^\mu u_t)$	
$c_{uu}^{(1)}$	$(\bar{u}_p \gamma_\mu u_r)(\bar{u}_s \gamma^\mu u_t)$	
$c_{qu}^{(1)}$	$(\bar{q}_p \gamma_\mu q_t)(\bar{u}_r \gamma^\mu u_s)$	
$c_{ud}^{(8)}$	$(\bar{u}_p \gamma_\mu T^A u_r)(\bar{d}_s \gamma^\mu T^A d_t)$	
$c_{qu}^{(8)}$	$(\bar{q}_p \gamma_\mu T^A q_r)(\bar{u}_s \gamma^\mu T^A u_t)$	
$c_{qd}^{(8)}$	$(\bar{q}_p \gamma_\mu T^A q_r)(\bar{d}_s \gamma^\mu T^A d_t)$	
c_G	$f^{ABC} G_\mu^{A\nu} G_\nu^{B\rho} G_\rho^{C\mu}$	

Table 11.2: The most relevant SMEFT operators of dimension six for the Higgs boson coupling analysis described in Section 11.7. The table shows the Wilson coefficients, their corresponding operators and examples of diagrams in which they are relevant [30].

where σ_{SM} is the cross section predicted by the SM, σ_{BSM} the additional contribution to the cross section introduced by the operators with dimension six and σ_{int} the contribution to the cross section resulting from interference terms of the SM and the dimension six operators in the squared matrix element. Here, the linear interference terms have only one contribution from dimension six. They are proportional to c_i and are suppressed by a factor $1/\Lambda^2$. The dimension six squared terms are suppressed by a factor $1/\Lambda^4$, due to the fact that two dimension six operators are involved in the term (proportional to $c_i c_j$ or c_i^2). This means

that these quadratic terms have the same suppression as linear interference terms between dimension eight and SM operators. For a consistent measurement of SMEFT effects up to a suppression of $1/\Lambda^2$, only the interference terms of the dimension six operators linear in c_i are considered in the two analyses. In Comb1 a second measurement is performed additionally in which the quadratic terms of the dimension six operators are added without considering the interference terms of dimension eight operators with the same suppression. This setup is used to study the impact of the higher dimension operators on the measurement.

The impact of the measured STXS bins and decay modes for Comb1 on the most relevant SMEFT operators relative to their SM predictions are shown in Figure 11.13 and for Comb2 in the appendix in Figure A.7. Several studies of the expected sensitivity are performed for Comb1 and Comb2. Some important observations for Comb1 will be discussed in the following, while a detailed discussion of similar observations for Comb2 are described in detail in Reference [30]. For Comb1 the parameter $c_{Hq}^{(3)}$ can be measured individually. The sensitivity to this parameter is mainly driven by the VH production mode and in particular by the $VH \rightarrow b\bar{b}$ measurement. The parameters c_{qq} , $c_{qq}^{(3)}$, $c_{qq}^{(31)}$, $c_{uu}^{(1)}$, $c_{ud}^{(8)}$, $c_{qu}^{(1)}$, $c_{qu}^{(8)}$ and c_G are only constraint by the $t\bar{t}H$ measurement, while the parameters c_{HG} , $|c_{uG}|$ and $|c_{uH}|$ are constraint by $t\bar{t}H$ and ggF measurements. The $H \rightarrow \gamma\gamma$ measurement is sensitive to the parameters c_{HW} , c_{HB} and c_{HWB} . Through a correlation of the VH and VBF measurements, weaker constraints on c_{uW} and c_{uB} arise. The parameters $c_{Hl}^{(1)}$, $c_{Hl}(3)$, c_{He} and $c_{Hq}^{(1)}$ are weakly constraint by common measurements in the $VH \rightarrow b\bar{b}$ measurement, resulting in a notable correlation between these parameters in the measurements. The parameters $c_{H\Box}$, c_{dH} and $|c_{eH}|$ impact the measurement of the cross section as a global scale factor. This effect can be absorbed by a rescaling of all other c_i resulting in a degeneration of the fit. Therefore, these parameters are set constant to zero in the analysis, to avoid instabilities of the fit model.

These and further observations show that the data sets contain insufficient information to constrain all coefficients considered in the analysis input, so that a reduced set of fit parameters is build instead. The unconstrained degrees of freedom in data not always correspond to individual coefficients, but instead to a linear combination of them. Therefore, a new modified basis of Wilson coefficients is estimated using a methodology based on the expected covariance matrix of the measurement and an eigenvector decomposition. This method is described in detail for the Comb1 in Reference [30].

In Figure 11.14 the rotation matrices from the Warsaw to the final fit basis are shown for Comb1 and Comb2. The increased data statistics and especially the new included regions of phase space in Comb2 compared to Comb1 allow the measurement of more independent coefficients. Ten free parameters are defined for Comb1, while thirteen are defined for Comb2. However, the new data set also modifies the optimal linear combinations of parameter for the fit model, so that Comb2 does not only measure more independent parameters but also parameters based on different individual or linearly combined Wilson coefficients. The 10 free parameters for Comb1 are the linear combinations:

- $c_{Hq}^{(3)}$
- $c_{HW,HB,HWB,HDD,uW,uB}^{[1]}$, $c_{HW,HB,HWB,HDD,uW,uB}^{[2]}$, $c_{HW,HB,HWB,HDD,uW,uB}^{[3]}$

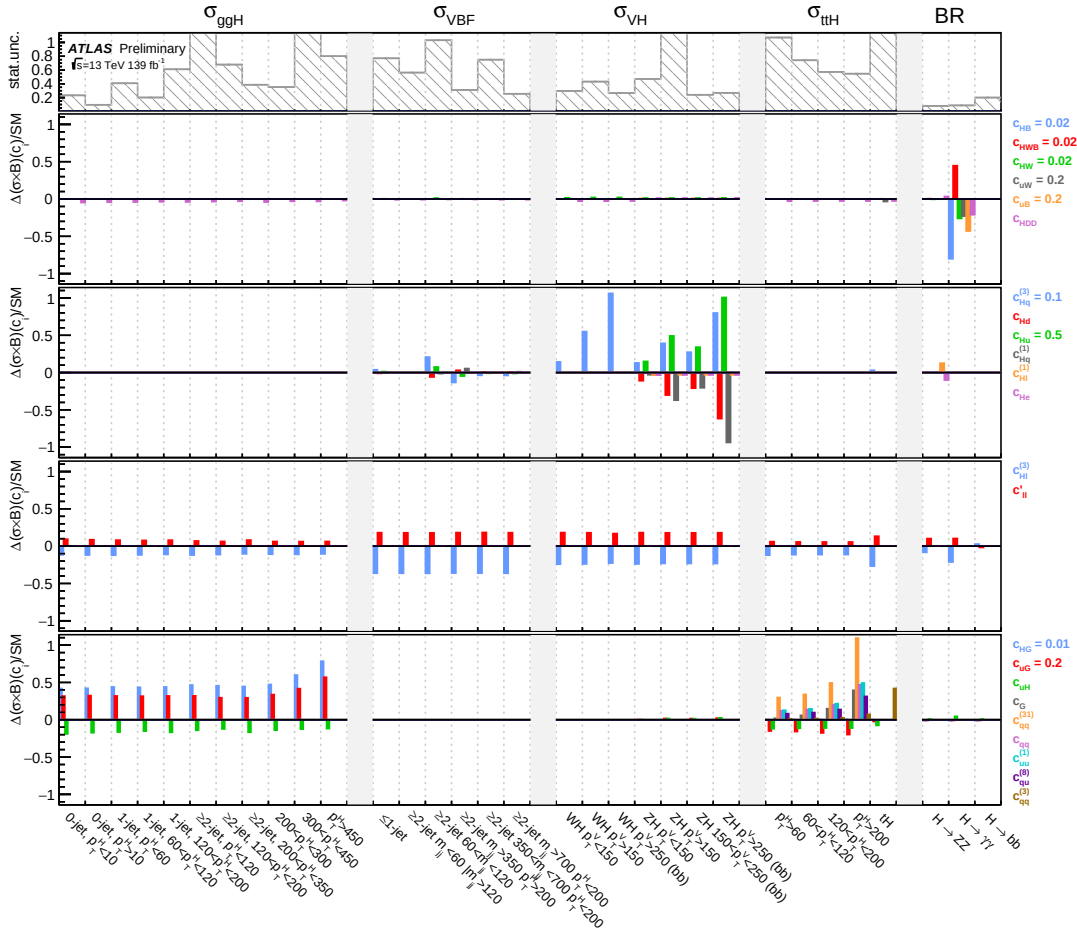


Figure 11.13: Impact of the most relevant SMEFT operators on the STXS bins and decay modes, relative to the SM predictions for Comb1. Only the interaction terms with a suppression of $1/\Lambda^2$ are considered. A unit variation is considered for all coefficients c_i , if not otherwise specified in the legend. The statistical uncertainties on the STXS bins are shown in the top panel to judge the experimental sensitivity to constrain the operators from the data. The statistical uncertainty of low precision STXS bins are clipped in the plot for a better presentation of the other uncertainties [30].

- $C_{Hu,Hd,Hq}^{[1]}$
- $C_{Hl^{(1)},He}^{[1]}$
- $C_{Hl^{(2)},ll'}^{[1]}$
- $C_{HG,uG,uH,top}^{[1]}, C_{HG,uG,uH,top}^{[2]}, C_{HG,uG,uH,top}^{[3]}$

The indices for each parameter refer to the Wilson coefficients that are used in the linear com-

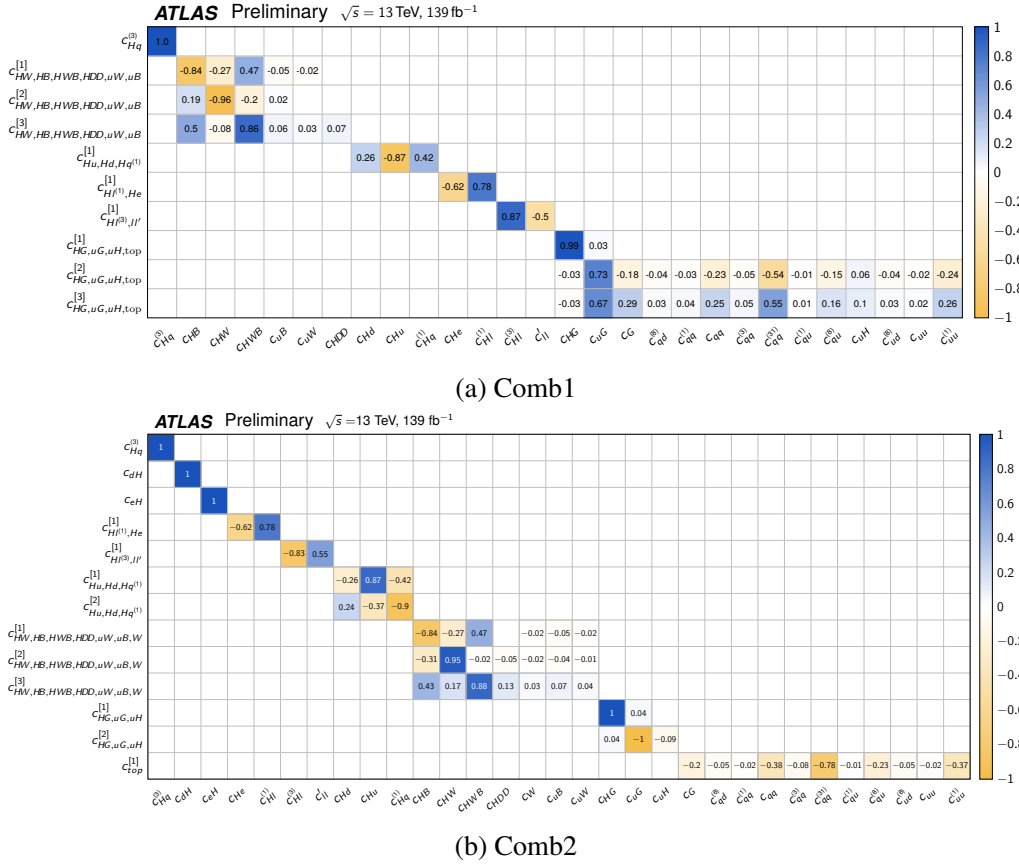


Figure 11.14: Rotation matrices for Comb1 (a) and Comb2 (b) from the Warsaw basis (x-axis) to the final fit basis (y-axis) [30, 31].

bination. The definition of the linear combination for each of these parameter is summarized in the appendix in Table A.7.

The final PoIs for Comb2 are:

- $c_{Hq}^{(3)}, c_{dH}, c_{eH}, c_{top}^{[1]}$
- $c_{HI^{(1)},He}^{[1]}$
- $c_{HI^{(3)},ll'}^{[1]}$
- $c_{Hu,Hd,Hq^{(1)}}^{[1]}, c_{Hu,Hd,Hq^{(1)}}^{[2]}$
- $c_{HW,HB,HWB,HDD,uW,uB,W}^{[1]}, c_{HW,HB,HWB,HDD,uW,uB,W}^{[2]}, c_{HW,HB,HWB,HDD,uW,uB,W}^{[3]}$
- $c_{HG,uG,uH}^{[1]}, c_{HG,uG,uH}^{[2]}$

Figures 11.15 and 11.16 summarize the observed results for Comb1 and Comb2. For a better visualization, the PoIs in the figures are grouped into three different sub-plots with different x-axis ranges. For Comb1 the results for a fit using only the interference terms (linear) and additionally using the information of the quadratic terms of the dimension six (linear+quadratic) are shown. The estimated 68% confidence-intervals (CIs) for the linear model of Comb1 vary between [1.3,27] on $c_{HI^{(1)},He}^{[1]}$ and [-0.015,0.00007] on $c_{HG,uG,uH,top}^{[1]}$, while the 95% CIs on them are [-2.9,36] and [-0.022,0.008]. In the linear+quadratic model the 68% CIs on these parameters are reduced to [-7,7] and [-0.014,-0.0016] and for the 95% CIs to [-10,10] and [-0.019,0.005]. Generally, the CIs on the estimated parameters are reduced for the linear+quadratic fit compared to the linear fit, suggesting a non-negligible impact of the dimension six operator terms suppressed by $1/\Lambda^4$ on the measurement. However, including the quadratic terms into the likelihood model results in rather non-Gaussian dependencies of the likelihood function on some of the measured parameters. These parameters can show two local minima per parameter in the minimization of the negative log-likelihood function. In particular, for the parameter $c_{HW,HB,HWB,HDD,uW,uB}^{[1]}$ two global minima with the same negative log-likelihood value are observed. In Comb2, the smallest CIs are estimated on the parameters $c_{HW,HB,HWB,HDD,uW,uB,W}^{[1]}$ and $c_{HG,uG,uH}^{[1]}$ with [-0.004,0.005] ([-0.009,0.01]) and [-0.01,-0.0018] ([-0.013,0.0021]) for 68% (95%).

In all three measurements no significant deviation of a coefficient from zero and therefore from the SM prediction is observed. A comparison of the two combinations is difficult, given the increased numbers of freedom and the use of new linear combinations of Wilson coefficients for Comb2 compared to Comb1. The $C_{Hq}^{(3)}$ Wilson coefficient can be used for a sensitivity comparison between Comb1 and Comb2, since it is included as a single fit parameter in both combination. The 95% confidence interval has been reduced from [-0.18,0.09] to [-0.08,0.1] and the 68% interval from [-0.08,0.05] to [-0.04,0.05], which shows the increased sensitivity of the new analysis.

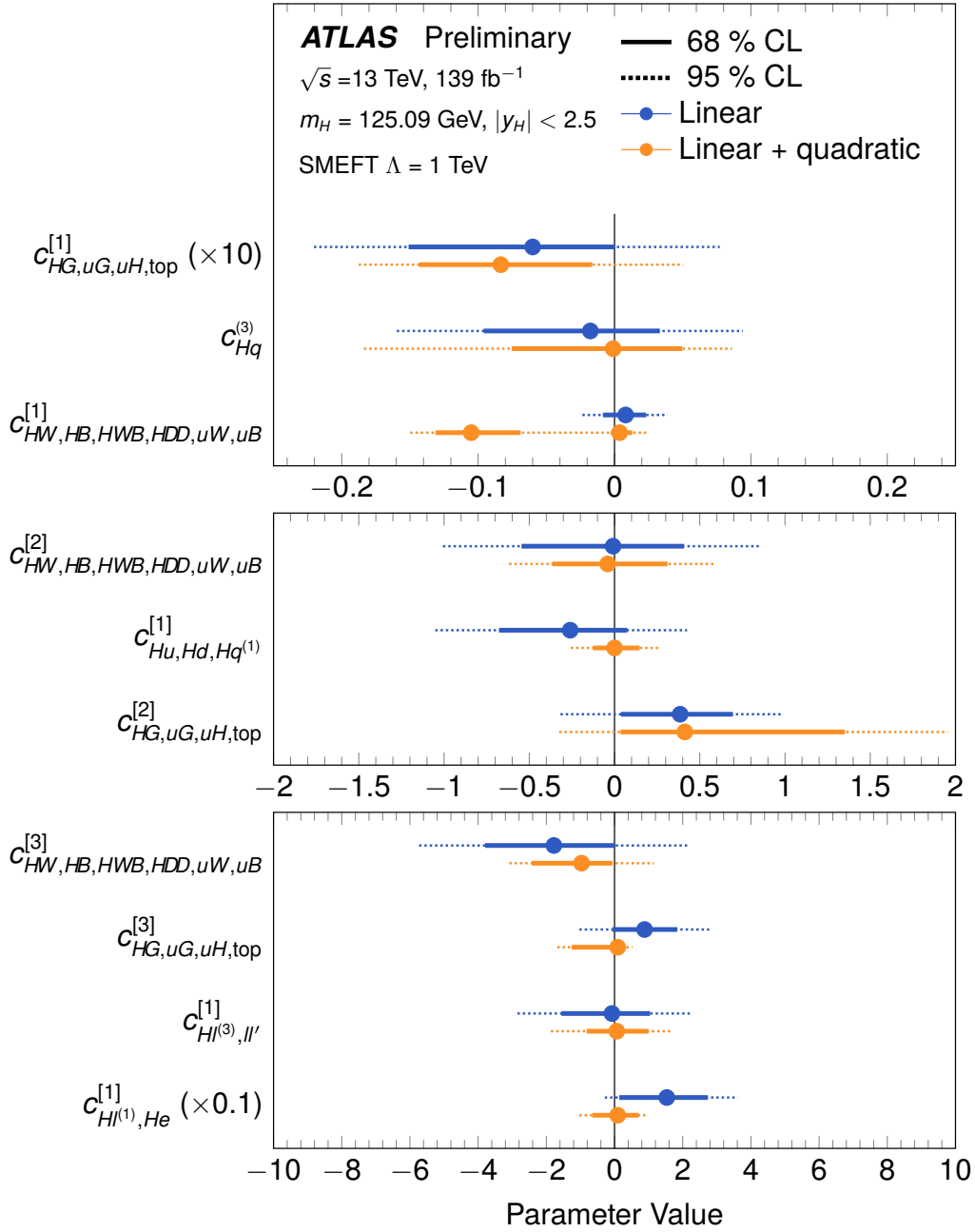


Figure 11.15: Summary of observed measurements of coefficients with the SMEFT linearized model (blue) for Comb1. The solid lines corresponds to 68% and the dashed lines to 95% confidences level intervals. The results of a second fit setup additionally including the quadratic terms of dimension six operators are shown in orange [30].

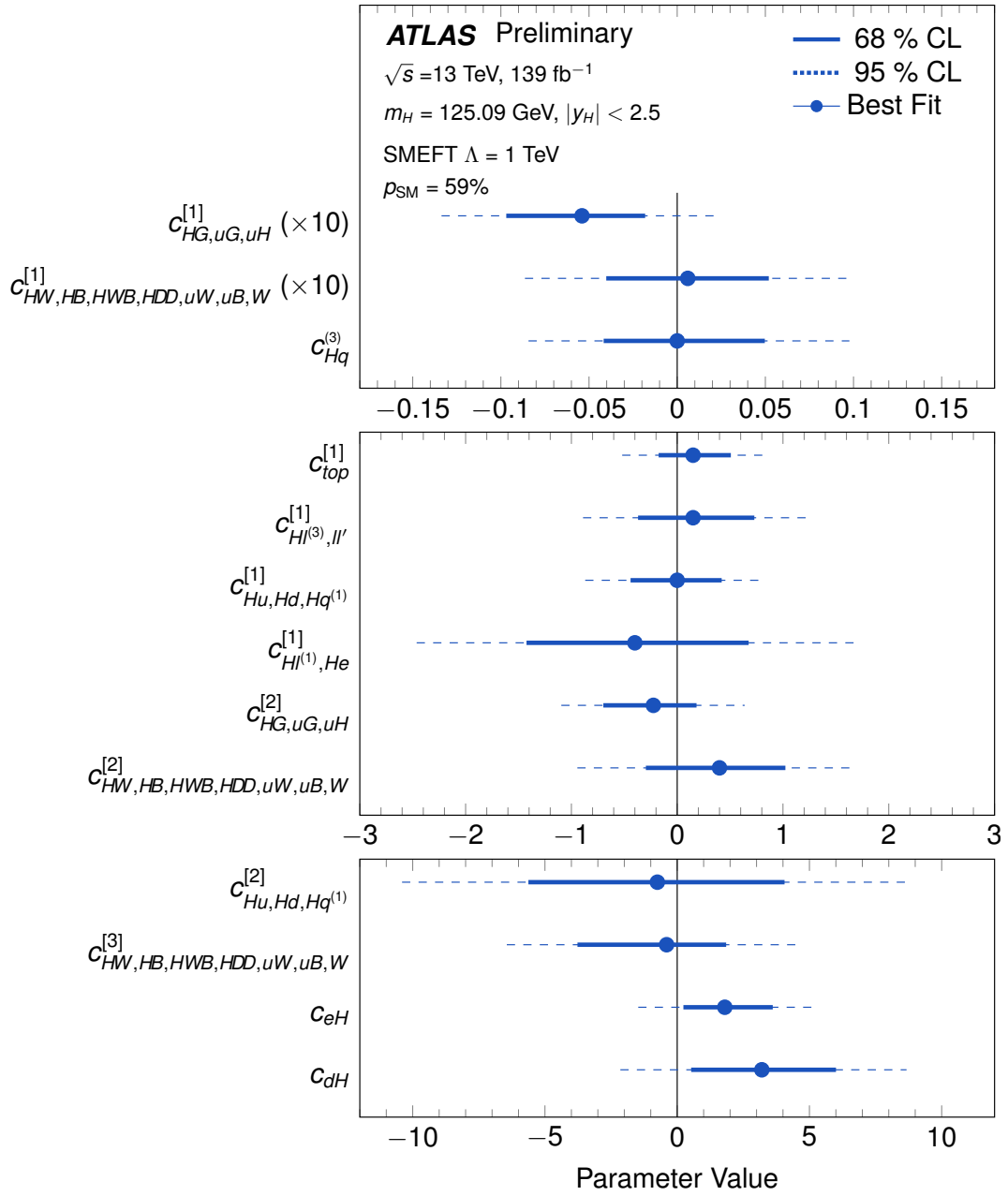


Figure 11.16: Summary of observed measurements of coefficients with the SMEFT linearized model for Comb2. The solid lines corresponds to 68% and the dashed lines to 95% confidence level intervals [31].

12 EXTRAPOLATION OF THE $H \rightarrow \tau\tau$ ANALYSIS TO THE DATA SET EXPECTED AT HL-LHC

The next data taking period Run 3 (see Section 4.4) of the LHC will start in 2022 and takes place until 2025. Afterwards a long shutdown of the LHC is used to upgrade the LHC and the detectors to the so called high luminosity LHC (HL-LHC). The upgrade will allow the LHC to provide an instantaneous luminosity of up to $7.5 \times 10^{-34} \text{ cm}^{-2} \text{ s}^{-1}$ and an integrated luminosity of 3000 fb^{-1} in ten years of data taking. Additionally, the center-of-mass energy will be increased from 13 to 14 TeV. It is also expected that several experimental and theoretical systematic uncertainties can be reduced. Prospect studies based on current analyses are performed to explore the sensitivity of them for the data set of the HL-LHC. This chapter describes the prospect analysis of the $H \rightarrow \tau\tau$ decay channel based on the analysis described in Chapter 10.

12.1 Extrapolation strategy

The likelihood model of the $H \rightarrow \tau\tau$ analysis based on the 2015-2018 data set is correct for the higher integrated luminosity and developments which affect the size of systematic uncertainties for this extrapolation. Therefore, no optimization of the analysis strategy is performed.

To extrapolate the analysis to the nominal prediction of the signal but also of the background contribution have to be adjusted for the larger integrated luminosity and center-of-mass energy. This correction is discussed in Section 12.1.1. The procedure to scale the systematic uncertainties to size expected for the HL-LHC data set is described in Section 12.1.2. The extrapolation follows the recommendations of the ATLAS Upgrade Physics Working Group, with slight modifications to account for particularities of the $H \rightarrow \tau\tau$ decay channel.

12.1.1 Extrapolation of nominal prediction

The nominal prediction of the signal and background processes is corrected to the higher expected integrated luminosity of 3000 fb^{-1} of data set at the HL-LHC, but also the prediction of the cross section is corrected for the higher center-of-mass energy of 14 TeV. To do so, correction factors are insert in the likelihood model as constant NFs.

The expected event yields of signal and background processes are scaled up from the current luminosity $\mathcal{L}_{\text{Run2}}$ to the expected luminosity $\mathcal{L}_{\text{HL-LHC}}$ by applying one correction factor c_{lumi} to all samples with: $c_{\text{lumi}} = \mathcal{L}_{\text{HL-LHC}}/\mathcal{L}_{\text{Run2}} = 3000/139 = 21.58$.

The production cross sections of signal and background processes are corrected for the increased center-of-mass energy at the HL-LHC by adding one correction factor per production mode $c_{\text{XS},i}$ in the likelihood model. Table 12.1 summarizes all applied correction factors. The main Higgs boson production modes ggF and VBF are scaled by 1.13, while the VH and $b\bar{b}H$ production are scaled by 1.1 and 1.14. The largest correction is applied to the production mode including at least one top quark with 1.21 for the $t\bar{t}H$ and the tH .

The background processes are scaled by 1.1 except of the top quark background, which is scaled by 1.16.

Process	Scaling factor $c_{XS,i}$
ggF	1.13
VBF	1.13
VH	1.10
$t\bar{t}H$	1.21
$b\bar{b}H$	1.14
tH	1.21
Top quark bkg.	1.16
τ misidentified bkg.	1.10
Other bkg.	1.10

Table 12.1: Correction factors $c_{XS,i}$ to correct the production cross section for the physics processes for the increased center-of-mass energy from 13 TeV to 14 TeV for the HL-LHC.

12.1.2 Extrapolation of systematic uncertainties

Several uncertainties are expected to be reduced for the HL-LHC data set. In particular, detector upgrades, new strategies for auxiliary measurements and updates of theoretical predictions will reduce several systematic uncertainties that are considered in the $H \rightarrow \tau\tau$ analysis. Therefore, correction factors $c_{NP,i}$ are applied on the different uncertainties to reduce the normalization component of them. Shape effects on the observable $m_{\tau\tau}^{mmc}$ are not changed. Table 12.2 summarizes the correction factors applied to different sets of systematic uncertainties. The systematic uncertainties are scaled in general by values between 1 and 0.5. The theoretical uncertainties on the signal and background prediction are scaled by 0.5, while the uncertainty on the luminosity is scaled by 0.59. Uncertainties assigned to the reconstruction and identification of the physics object, such as muons, electrons, jets, b-jets and τ_h -leptons are in general scaled by 1. Exceptions are a sub-set of systematic uncertainties assigned to the τ_h -leptons and to the b-jet tagging. All NPs corresponding to uncertainties on the b-tagging efficiencies against light quark jets are scaled by 0.5, while the remaining are scaled by 1. Systematic uncertainties on the τ_h -lepton assigned to the limited size of simulated samples in auxiliary measurements are set to zero. This follows the recommendation of the ATLAS Upgrade Physics Working Group to scale statistical uncertainties of simulations to zero, while the remaining systematic uncertainties on the τ_h -lepton are scaled by 1.

The statistical uncertainties on the data-driven backgrounds, such as the misidentified τ background are scaled according to increased data statistics of the HL-LHC. Therefore, the statistical uncertainty on the misidentified τ background in each bin is scaled down by $1/\sqrt{c_{lumi}} \approx 0.59$, while systematic uncertainties related to the data-driven method, such as closure uncertainties, are scaled by 1.

The statistical uncertainties on other background predictions arising from the limit size of simulated samples are corrected as well. The recommendation of the ATLAS Upgrade Physics

Uncertainty sources	Scaling factor $c_{\text{NP},i}$
Theory	0.5
Luminosity	0.59
Muon, Electron	1
Jet	1
B-jet Light	1
B-jet other	0.5
τ_h -lepton method	1
τ_h -lepton stat.	0
E_T^{miss}	0.5
Misidentified τ method	1
Misidentified τ stat.	0.59
Background sample size	0.59

Table 12.2: Correction factors $c_{\text{NP},i}$ to correct the impact of systematic uncertainties on the PoI for improved auxiliary measurements. The τ -lepton and τ misidentified uncertainties are split into uncertainties sourced by statistical uncertainties (stat.) of simulated events or data and uncertainties caused by the used method to calculate the object or background.

Working Group is to set these uncertainties to zero. This is based on the idea that the long time scale until the data of the HL-LHC is taken enough simulation events are produced, so that these uncertainties can be neglected. However, the specific topology of the region of phase space measured with the $H \rightarrow \tau\tau$ analysis leads to small acceptances of simulated events into the event selection. Therefore, the analysis suffers from a large impact of these uncertainties. To reflect this in the extrapolation the uncertainty of this source is not set to zero, instead it is assumed that the analysis will have the same relative amount of simulated events compared to data as for the current analysis. To achieve this the statistical uncertainties on simulated events are scaled by $1/\sqrt{c_{\text{lumi}}} \approx 0.59$. This configuration is called scenario 1. To study the impact of this alternative treatment a second scenario 2 is performed, which uses the recommendation to set the statistical uncertainties on the simulated events to zero. This second scenario will be studied and discussed for the STXS Stage 1.2 9-PoI setup, since this setup has the largest number of freedom and the largest impact of these uncertainties in the current setup. Therefore, the largest difference between the two scenarios is expected for this fit configuration.

The reduction of the normalization component of several systematic uncertainties increases the number of systematic uncertainties that fail the pruning criteria on the relative size of the normalization component (see Section 10.5.3). This can lead to an underestimation of the systematic uncertainty. Therefore, the pruning threshold is lowered from 1% to 0.4% to ensure that the pruning does not remove relevant uncertainties.

12.2 Results

To extract the expected precision for this extrapolated analysis a pseudo-data set is created. For this pseudo-data set it is assumed that the SM is realized by nature. The corrected nominal prediction for all signal and background events are summed up to create the pseudo-data set and finally the statistical uncertainties on each bin are set to the square root of the expected events. Using this pseudo-data set in the likelihood model results in no deviations of PoIs, NFs and NPs from their corresponding prediction. Therefore, for all PoIs and NFs a best-fit value of 1 and for all NPs no pulls are expected.

Table 12.3 and Figure 12.1 compare the estimated inclusive cross section results for the HL-LHC scenario 1 and the current $H \rightarrow \tau\tau$ analysis (see Chapter 10). The total uncertainty

Production Mode	Data set	$\frac{(\sigma_i \times \mathcal{B}_{\tau\tau})_{\text{obs}}}{(\sigma_i \times \mathcal{B}_{\tau\tau})_{\text{SM}}}$	Syst. unc.				
			Stat. unc.	Total	Th. sig.	Th. bkg.	Exp.
$H \rightarrow \tau\tau$	Run 2	1.0 ± 0.127	± 0.063	± 0.109	± 0.083	± 0.017	± 0.061
	HL-LHC	1.0 ± 0.046	± 0.013	± 0.044	± 0.038	± 0.007	± 0.023
$t\bar{t}H$	Run 2	1.0 ± 1.062	± 0.884	± 0.638	± 0.188	± 0.300	± 0.257
	HL-LHC	1.0 ± 0.239	± 0.169	± 0.182	± 0.095	± 0.051	± 0.105
VH	Run 2	1.0 ± 0.581	± 0.469	± 0.325	± 0.126	± 0.041	± 0.182
	HL-LHC	1.0 ± 0.137	± 0.096	± 0.100	± 0.056	± 0.011	± 0.057
ggF	Run 2	1.0 ± 0.328	± 0.138	± 0.274	± 0.212	± 0.033	± 0.173
	HL-LHC	1.0 ± 0.114	± 0.028	± 0.111	± 0.101	± 0.015	± 0.057
VBF	Run 2	1.0 ± 0.180	± 0.116	± 0.133	± 0.110	± 0.022	± 0.059
	HL-LHC	1.0 ± 0.065	± 0.023	± 0.061	± 0.056	± 0.009	± 0.026

Table 12.3: Estimated cross sections and uncertainties for the inclusive $H \rightarrow \tau\tau$ and for the four dominant production modes (ggF , VBF , VH and $t\bar{t}H$) divided by the SM prediction for the 2015-2018 analysis (Run 2) and the HL-LHC extrapolation (HL-LHC). On the right-side the contributions to the total uncertainty from statistical (Stat. unc.), systematic (Syst. unc.) and experimental uncertainties (Exp.) as well as theoretical uncertainties on the signal prediction (Th. sig.) and on background prediction (Th. bkg.) are quoted separately.

for the inclusive 1-PoI fit shown in the first row of the table is reduced from 13% to 5%. The statistical uncertainty is reduced from 6% to 1%, while the systematic uncertainties are reduced from 11% to 4%. The theoretical uncertainties on the signal prediction are halved from 8% to 4%, while the theoretical uncertainties on the background and the experimental uncertainties are reduced by more than 50%. Therefore, the extrapolated 1-PoI measurement is dominated by systematic uncertainties and the impact of theoretical uncertainties on the signal is further increased compared to the current analysis. Figure 12.1 shows that the uncertainty on the measurement of the inclusive $H \rightarrow \tau\tau$ cross section is expected to be smaller than theoretical uncertainty on the SM cross section prediction, even after applying

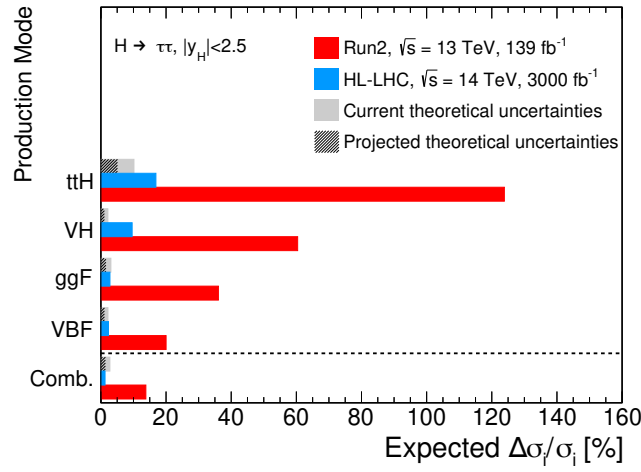


Figure 12.1: Relative uncertainties on the measured inclusive cross sections for the 2015-2018 analysis (Run 2) analysis (red) and the extrapolation to the HL-LHC (blue). The theoretical uncertainties on the prediction of each cross section is shown in gray for their current size and in dashed when applying the scaling correction on the theoretical uncertainties for the HL-LHC extrapolation.

the same reduction on them as used for the theoretical uncertainties on the signal prediction in the fit model.

For the 4-PoI setup the relative improvement of the cross section vary for the different production modes. The statistical dominated production modes in the 2015-2018 analysis $t\bar{t}H$ and VH show a larger reduction of the total uncertainty from 106% to 23% and from 58% to 14%, while the ggF and VBF production modes are reduced from 32% to 11% and from 18% to 7%. For all production modes the systematic uncertainties are larger than the statistical uncertainties, however the VH and $t\bar{t}H$ production modes have a similar impact of statistical and systematic uncertainties. The HL-LHC extrapolation for the VBF and ggF production modes are dominated by theoretical uncertainties on the signal prediction. The systematic uncertainties on the $t\bar{t}H$ and VH production modes are mainly arising from theoretical uncertainties on the signal prediction but also experimental uncertainties are important. For all four production modes the relative impact of theoretical uncertainties on the signal prediction is increased, while the relative impact of uncertainties on the background prediction are reduced compared to the current analysis. The expected significance of the $t\bar{t}H$ production mode is 4.39σ and for the VH production mode 7.74σ , so that VH production is expected to be observed with the HL-LHC data set also in the $H \rightarrow \tau\tau$ decay mode, while for the $t\bar{t}H$ production mode only achieving an evidence is expected. To observe the $t\bar{t}H$ production mode the measurement of $t\bar{t}H$ production mode decay into final state with electrons and muons have to be included. The ggF production mode has an expect significance $\gg 5\sigma$ and the VBF production mode was already observed with the Run 2 data set. Figure 12.1 shows

12.2 Results

STXS bins				Data set	$\frac{(\sigma_i \times \mathcal{B}_{\tau\tau})_{\text{obs}}}{(\sigma_i \times \mathcal{B}_{\tau\tau})_{\text{SM}}}$	Syst. unc.				
Process	m_{ij} [GeV]	p_{T}^{H} [GeV]	N_{j}			Stat. unc.	Total	Th. sig.	Th. bkg.	Exp.
ggF	[0,350]♣	[60,120]	≥ 1	Run 2	1.0 ±0.914	±0.560	±0.738	±0.106	±0.325	±0.490
				HL-LHC	1.0 ±0.197	±0.096	±0.173	±0.043	±0.056	±0.129
		[120,200]	$=1$	Run 2	1.0 ±0.646	±0.390	±0.527	±0.131	±0.086	±0.334
				HL-LHC	1.0 ±0.161	±0.073	±0.146	±0.071	±0.029	±0.095
	[0,350]	[120,200]	≥ 2	Run 2	1.0 ±0.745	±0.507	±0.581	±0.122	±0.136	±0.473
				HL-LHC	1.0 ±0.189	±0.092	±0.167	±0.056	±0.037	±0.134
		[200,300]	≥ 1	Run 2	1.0 ±0.374	±0.288	±0.260	±0.107	±0.029	±0.200
				HL-LHC	1.0 ±0.098	±0.054	±0.083	±0.050	±0.012	±0.056
		[300,∞[	≥ 1	Run 2	1.0 ±0.480	±0.450	±0.252	±0.108	±0.033	±0.159
				HL-LHC	1.0 ±0.113	±0.082	±0.081	±0.049	±0.011	±0.049
	[350,∞[	[0,200]	≥ 2	Run 2	1.0 ±1.616	±0.997	±1.308	±0.601	±0.147	±0.999
				HL-LHC	1.0 ±0.541	±0.218	±0.499	±0.295	±0.049	±0.363
EW qqH	[350,∞[	≥ 2	Run 2	1.0 ±0.208	±0.150	±0.147	±0.124	±0.025	±0.041	
			HL-LHC	1.0 ±0.075	±0.029	±0.070	±0.065	±0.011	±0.023	
	[60,120]	≥ 2	Run 2	1.0 ±0.635	±0.528	±0.367	±0.118	±0.056	±0.198	
			HL-LHC	1.0 ±0.144	±0.102	±0.105	±0.052	±0.013	±0.064	
$t\bar{t}H$				Run 2	1.0 ±1.160	±1.002	±0.716	±0.204	±0.302	±0.260
				HL-LHC	1.0 ±0.239	±0.169	±0.182	±0.095	±0.051	±0.106

Table 12.4: Estimated cross sections and uncertainties of the 9-PoI STXS measurement for the 2015-2018 analysis (Run 2) and the HL-LHC extrapolation (HL-LHC). On the right-side the contributions to the total uncertainty from statistical (Stat. unc.), systematic (Syst. unc.) and experimental uncertainties (Exp.) as well as theoretical uncertainties on the signal prediction (Th. sig.) and on background prediction (Th. bkg.) are quoted separately. The spades symbol ♣ indicates, that the requirement on m_{jj} is only applied for $N_{\text{j}} \geq 2$.

that the size of the expected uncertainties on the measured ggF and VBF production mode is similar to the theoretical uncertainties on the SM prediction of them, while the measured VH and $t\bar{t}H$ are still larger than the theoretical uncertainties on their SM predictions.

The results of the STXS Stage 1.2 measurement with the 9-PoI setup for the extrapolation scenario 1 as well as for the 2015-2018 analysis are shown in Table 12.4 and in Figure 12.2. The relative uncertainties on the cross sections measured in the extrapolation vary between 54% and 8%. The EW qqH STXS bin with at least two jets and $350 \text{ GeV} \leq m_{\text{jj}}$ (8%) followed by the two ggF STXS bins with $200 \leq p_{\text{T}}^{\text{H}} < 300 \text{ GeV}$ (10%) and $p_{\text{T}}^{\text{H}} \geq 300 \text{ GeV}$ (11%) show the smallest relative uncertainties. The largest uncertainty (54%) is observed for the ggF STXS bin with at least two jets and $350 \text{ GeV} \leq m_{\text{jj}}$. Most of the STXS bins are dominated by systematic uncertainties. For the lower p_{T}^{H} STXS bins of ggF the experimental uncertainties are dominating, while for the higher p_{T}^{H} STXS bins the theoretical uncertainty on the signal predication has a similar size. The EW qqH and $t\bar{t}H$ STXS bins behave as the corresponding inclusive cross section in the 4-PoI setup. Similar to the 1-PoI and 4-PoI setup, the relative impact of theoretical uncertainties on the signal is increased for all STXS bins

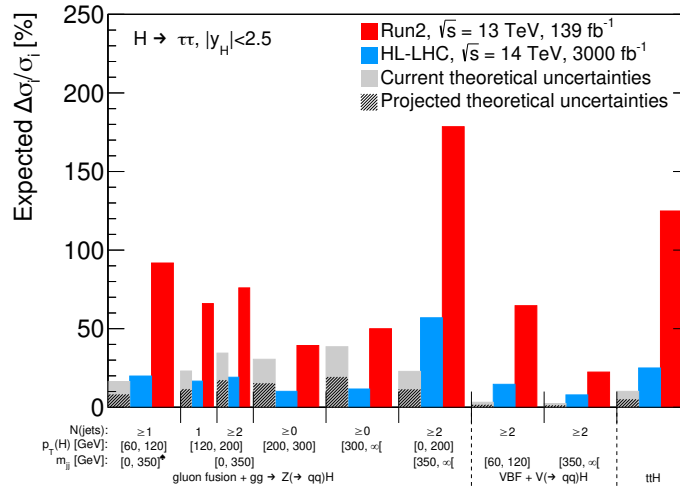


Figure 12.2: Relative uncertainties on the measured STXS bins in the 9-PoI fit for the 2015-2018 analysis (Run 2) analysis (red) and the extrapolation scenario 1 to the HL-LHC (blue). The theoretical uncertainties on the prediction of each cross section is shown in gray for their current size and in dashed when applying the scaling correction on the theoretical uncertainties for the HL-LHC extrapolation. The spades symbol ♠ indicates, that the requirement on m_{jj} is only applied for $N_j \geq 2$.

compared to the current analysis. The largest relative reduction of the total uncertainty is observed for the STXS bins which are currently dominated by statistical uncertainties as the $t\bar{t}H$ STXS bin. The largest reduction of systematic uncertainties is visible for the theoretical uncertainties on the background prediction. Figure 12.2 shows that the size of the expected uncertainties on the measured ggF STXS bins with $120 \leq p_T^H < 200$ GeV are similar to the theoretical uncertainties on the SM prediction of them, while they are smaller than the theoretical uncertainties on the SM prediction for the ggF STXS bins with $200 \text{ GeV} \leq p_T^H$. For the remaining ggF STXS the expected uncertainties are still larger than their uncertainties on the prediction and the remaining STXS show a similar picture as the corresponding inclusive production mode cross sections in the 4-PoI setup.

Table 12.5 shows the result for the 9-PoI fit setup extrapolated with scenario 1 and with scenario 2. Setting the statistical uncertainties on the simulated events to zero instead to scale them with the increased luminosity reduced the relative systematic uncertainties by 1% to 23%, resulting in a reduction of the total relative uncertainties by 1% to 21%. The lowest reduction is visible for the EWqqH STXS bin with at least 2 jets and $350 \text{ GeV} \leq m_{jj}$, while the largest reduction is observed for the ggF STXS bin with at least two jets and $350 \text{ GeV} \leq m_{jj}$. The reduction of the systematic uncertainties of the remaining bins vary between 6% and 16%. The comparison shows that for nearly all STXS bin the impact of the statistical uncertainty on the background contribution is still a sizable uncertainty if the number of simulated event cannot significantly be increased relative to the increased data statistics.

12.3 Limitations of the extrapolation

STXS bins				Scenario	$\frac{(\sigma_j \times \mathcal{B}_{\tau\tau})_{\text{obs}}}{(\sigma_j \times \mathcal{B}_{\tau\tau})_{\text{SM}}}$		Syst. unc.
Process	m_{jj} [GeV]	p_{T}^{H} [GeV]	N_j				
ggF	[0,350]♣	[60,120]	≥ 1	1	1.0	± 0.197	± 0.173
				2	1.0	± 0.183	± 0.157
		[120,200]	$= 1$	1	1.0	± 0.161	± 0.146
				2	1.0	± 0.151	± 0.134
	[0,350]	[120,200]	≥ 2	1	1.0	± 0.189	± 0.167
				2	1.0	± 0.181	± 0.157
		[200,300]	≥ 1	1	1.0	± 0.098	± 0.083
				2	1.0	± 0.095	± 0.079
		[300, ∞ [	≥ 1	1	1.0	± 0.113	± 0.081
				2	1.0	± 0.110	± 0.075
	[350, ∞ [	[0,200]	≥ 2	1	1.0	± 0.541	± 0.499
				2	1.0	± 0.520	± 0.476
EW qqH	[350, ∞ [		≥ 2	1	1.0	± 0.075	± 0.070
				2	1.0	± 0.074	± 0.069
	[60,120]		≥ 2	1	1.0	± 0.144	± 0.105
				2	1.0	± 0.139	± 0.097
$t\bar{t}H$				1	1.0	± 0.239	± 0.182
				2	1.0	± 0.232	± 0.172

Table 12.5: Estimated cross sections and uncertainties of the 9-PoI STXS measurement for both HL-LHC extrapolation (HL-LHC) scenarios. Scenario 1 includes statistical uncertainties on the simulation scaled by the luminosity, while for scenario 2 these uncertainties are set to zero. The spades symbol ♠ indicates, that the requirement on m_{ij} is only applied for $N_j \geq 2$.

12.3 Limitations of the extrapolation

The method used for this HL-LHC extrapolation can be used to estimate the expected sensitivity for the current measurement with the increased luminosity. However, the method has several limitations. The extrapolation includes the correction for the higher center-of-mass energy, the reduction of statistical uncertainties and correction of the normalization component for several systematic uncertainties, while no optimization of the analysis strategy is considered. Therefore, optimizations as the chosen binning in the final discriminant observable (here $m_{\tau\tau}^{\text{MMC}}$), the correlation scheme of the PoIs or an optimized estimation of data-driven background are not included. The simple scaling of the nominal predictions and of systematic variations scales up statistical fluctuation, while a real increase of recorded data and simulated events would reduce these fluctuation and therefore improve the modeling of the observable used as input in the fit and of observables used in the event selection. The increased data set would furthermore allow to explore new regions of phase space. All these limitations have to be considered when the results of the extrapolation are interpreted.

13 CONCLUSION

In this thesis, the measurements of cross sections exploiting the $H \rightarrow \tau\tau$ decay mode as well as their combination with other Higgs boson decay modes using data of proton-proton collisions with a center-of-mass energy of 13 TeV recorded by the ATLAS experiment are presented. The results for the $H \rightarrow \tau\tau$ decay channel (see Sections 9 and 10) and the combination of these results with the remaining decay modes are discussed (see Chapter 11). Two consecutive sets of analyses are presented, where one is based on the 2015 to 2016 data set with an integrated luminosity of $\mathcal{L} = 36 \text{ fb}^{-1}$, while the second one is based on an optimized analysis strategy and on the full Run 2 data set recorded from 2015 to 2018 with $\mathcal{L} = 139 \text{ fb}^{-1}$. Furthermore, the full analysis is used to estimate the expected sensitivity of the $H \rightarrow \tau\tau$ decay mode at the HL-LHC with $\mathcal{L} = 3000 \text{ fb}^{-1}$ (see Chapter 12).

The $H \rightarrow \tau\tau$ analyses rely on τ -triggers for the $H \rightarrow \tau_h \tau_h$ decay channel. The trigger requirements can have a different efficiency in data compared to simulated events, which have to be corrected. Two the tag-and-probe analyses estimate the scale factors to correct the efficiencies in simulated events to the efficiencies observed in data. This thesis studies the analysis based on $t\bar{t} \rightarrow \tau_\mu \tau_h$ events and its combination with the analysis based on $Z \rightarrow \tau_\mu \tau_h$ events in Chapter 8. These combined scale factors vary between 0.95 and 1.05 resulting in an efficiency correction of up to 5%.

The $H \rightarrow \tau\tau$ analysis based on the 2015 and 2016 data sets focuses on the observation of the decay mode and on the inclusive cross section measurement of the two main production modes ggF and VBF . Additionally, a first STXS measurement in the $H \rightarrow \tau\tau$ decay mode is performed. The analysis based on the full Run 2 data set recorded from 2015 to 2018 focuses on the cross section measurement of the inclusive production modes ggF , VBF , VH and $t\bar{t}H$ as well as on the measurement of STXS bins. The likelihood model of this analysis is used in the HL-LHC extrapolation analysis to estimate the expected sensitivity of the analysis for a data set with $\mathcal{L} = 3000 \text{ fb}^{-1}$ at $\sqrt{s} = 14 \text{ TeV}$.

The analysis strategy has been modified between the 2015 to 2016 and the 2015 to 2018 analyses. The 2015 to 2016 analysis is based on 13 signal categories and six control regions, while for the 2016 to 2018 analysis 32 signal categories and 36 control regions are defined. Each control region is used to normalize a specific background contribution in the fit model via one NF. For both analyses, the $m_{\tau\tau}^{\text{MMC}}$ is used as an input to the fit model for all signal categories, while only the total number of events is used in the control regions. Careful investigations and detailed studies are performed on systematic uncertainties, which are included as NPs in the fit model.

The extrapolation strategy of the $H \rightarrow \tau\tau$ decay mode measurements is changed between the 2015 to 2016 and the 2015 to 2018 analyses. In the analysis of the 2015 and 2016 data sets, the measurements are extrapolated to the full phase space, while a fiducial volume is defined for the analysis of the 2015 to 2018 data set by requiring a rapidity of the Higgs boson of less than 2.5. The region of phase space above this threshold is not measured by the ATLAS experiment and therefore the results in this phase space are based purely on extrapolations. This results in different SM predictions for the measurements in the two analyses.

The observation of the $H \rightarrow \tau\tau$ coupling is achieved by combining the 2015 to 2016 data set

analysis with the Run 1 measurement [233] resulting in an observed (expected) significance of 6.4 (5.4) σ . The VBF production mode is observed with 5.3 (6.2) σ and an evidence for the ggF production mode is achieved with 3.9 (4.6) σ with the analysis based on the 2015 to 2018 data set. The HL-LHC extrapolation shows that the VH is expected to be observed with 7.74 σ , while an evidence of the $t\bar{t}H$ production mode with 4.39 σ is estimated. The results for the inclusive cross sections for the $H \rightarrow \tau\tau$ decay channel based on the 2015

Type	Data set	$\sigma_i \times \mathcal{B}_{\tau\tau}$ [pb]	Stat. Unc. [pb]	Syst. Unc. [pb]
observed	2015-2016	3.77	± 0.60	± 0.81
observed	2015-2018	2.94	± 0.21	± 0.35
expected	HL-LHC	3.17	± 0.04	± 0.14

Table 13.1: Cross sections and uncertainties of the inclusive $H \rightarrow \tau\tau$ measurement for all three analyses described in this thesis. The observed values for the analysis based on the 2015 to 2016 and the 2015 to 2018 data sets are shown, while the expected values for the HL-LHC extrapolation based on the 2015-2018 analysis are listed.

to 2016 and the 2015 to 2018 data sets analyses as well as for the HL-LHC extrapolation analysis are summarized in Table 13.1. The statistical uncertainties are reduced from 16% to 7% between the two analyses and is expected to decrease to 1.2% for the HL-LHC analysis, while the systematic uncertainty is reduced from 21% to 12% and is expected to be 4.4% for the HL-LHC analysis. The increased data statistics reduces the statistical uncertainties and the reduction of systematic uncertainties is driven by optimized reconstruction and identification methods, optimized auxiliary measurements and an improved analysis strategy. The precision of all three measurements is limited by systematic uncertainties. Both observed cross sections agree within the uncertainties with their SM expectation of (3.46 ± 0.13) pb for the 2015 to 2016 data set and (3.17 ± 0.09) pb for the 2015 to 2018 data set.

Table 13.2 summarizes the inclusive measurements of the dominant production modes cross sections in the $H \rightarrow \tau\tau$ decay channel (4-PoI fit). The most sensitive production mode is VBF followed by ggF for all analyses. The statistical (systematic) uncertainties on the measurement of VBF production modes are 32% (36%) and 16% (15%) for the 2015 to 2016 and the 2015 to 2018 analyses and is expected to be 2% (6%) for the extrapolation. On the measurements of the ggF production cross section a statistical (systematic) uncertainty of 32% (46%) and 16% (29%) is estimated for the 2015 to 2016 and the 2015 to 2018 analyses and the expected uncertainty for the extrapolation is 3% (11%). The VH and $t\bar{t}H$ production modes are only measured in the 2015 to 2018 data set analysis and in the extrapolation based on this analysis. The statistical (systematic) uncertainties on the measurements of the VH and $t\bar{t}H$ production modes are 50% (36%) and 95% (70%) for the 2015 to 2018 analysis and are expected to be 10% (10%) and 17% (18%) for the extrapolation.

The ggF production mode is dominated by systematic uncertainties for all analyses, while for VBF the relative size for systematic and statistical uncertainties is similar for the analysis

Production Mode	Type	Data set	$\sigma_i \times \mathcal{B}_{\tau\tau}$ [pb]	Stat. Unc. [pb]	Syst. Unc. [pb]
ggF	observed	2015-2016	3.10	± 1.00	± 1.45
	observed	2015-2018	2.65	± 0.41	± 0.72
	expected	HL-LHC	2.77	± 0.08	± 0.31
VBF	observed	2015-2016	0.28	± 0.09	± 0.10
	observed	2015-2018	0.197	± 0.028	± 0.029
	expected	HL-LHC	0.220	± 0.005	± 0.013
VH	observed	2015-2016			
	observed	2015-2018	0.120	± 0.058	± 0.039
	expected	HL-LHC	0.118	± 0.011	± 0.012
$t\bar{t}H$	observed	2015-2016			
	observed	2015-2018	0.033	± 0.031	± 0.020
	expected	HL-LHC	0.031	± 0.005	± 0.005

Table 13.2: Cross sections and uncertainties of the inclusive production modes of all three analyses described in this thesis. The observed values for the analysis based on the 2015-2016 and the 2015-2018 data sets are shown, while the expected values for the HL-LHC extrapolation based on the 2015-2018 analysis are listed. The VH and $t\bar{t}H$ production modes are not estimated with the 2015-2016 analysis.

based on the Run 2 data sets. Only for the extrapolation, VBF is dominated by systematic uncertainties. The remaining production modes are dominated by statistical uncertainties for the 2015 to 2018 data set analysis, while both have a similar impact of statistical and systematic uncertainties for the HL-LHC extrapolation. No significant deviation from their SM predictions is observed for the measurements of the inclusive production mode cross sections for the 2015 to 2016 and the 2015 to 2018 analyses.

Table 13.3 summarizes the observed results for the STXS measurement of the $H \rightarrow \tau\tau$ decay channel with nine independent PoIs based on the 2015 to 2018 data sets and the corresponding expected results of the HL-LHC extrapolation. The highest precision in the $H \rightarrow \tau\tau$ STXS measurement based on the 2015 to 2018 data sets is achieved for the ggF production mode with $200 \leq p_T^H < 300$ GeV and $p_T^H \leq 300$ GeV with 36% and 40% as well as for the $EW qqH$ production mode with at least two jets and $350 \text{ GeV} \leq m_{jj}$ with 24%. The expected relative uncertainties for these STXS bins in the HL-LHC extrapolation are 9.8%, 11.3% and 7.5%. The uncertainties on the measurement of other ggF STXS bins varies between 75% and 206%, while the expected uncertainties for these STXS bins are between 16% and 54% for the extrapolation. The uncertainties on the measured $EW qqH$ production mode STXS bin with at least two jets and $60 \leq m_{jj} < 120$ GeV and on the $t\bar{t}H$ STXS bin are 63% and 111%. The expected uncertainties on the measurement of these two STXS bins are 14% and

STXS bins								
Process	m_{ij} [GeV]	p_{T}^{H} [GeV]	N_{j}	Type	Data set	$\sigma_{\text{i}} \times \mathcal{B}_{\tau\tau}$	Stat. unc.	Syst. unc.
ggF	[0,350]♣	[60,120]	≥ 1	observed	2015-2018	189	± 220	± 323
				expected	HL-LHC	394	± 38	± 68
		[120,200]	$= 1$	observed	2015-2018	17	± 18	± 24
				expected	HL-LHC	47	± 3.4	± 6.9
	[0,350]	[120,200]	≥ 2	observed	2015-2018	33	± 27	± 28
				expected	HL-LHC	59	± 5	± 10
		[200,300]	≥ 1	observed	2015-2018	30.3	± 8.6	± 6.8
				expected	HL-LHC	30.5	± 1.7	± 2.5
		[300,∞[	≥ 1	observed	2015-2018	9.4	± 3.5	± 1.3
				expected	HL-LHC	7.7	± 0.6	± 0.6
	[350,∞[	[0,200]	≥ 2	observed	2015-2018	143	± 54	± 91
				expected	HL-LHC	55	± 12	± 27
$EW\ qqH$	[350,∞[	≥ 2	observed	2015-2018	33	± 17	± 11	
			expected	HL-LHC	33.1	± 0.96	± 2.32	
	[60,120]	≥ 2	observed	2015-2018	71	± 13	± 12	
			expected	HL-LHC	90.1	± 9.2	± 9.5	
$t\bar{t}H$				observed	2015-2018	34	± 32	± 20
				expected	HL-LHC	31.3	± 5.3	± 5.7

Table 13.3: Estimated cross sections and uncertainties of the 9-PoI STXS measurement for the $H \rightarrow \tau\tau$ analysis based on the 2015-2018 data set and for the extrapolation of this analysis to the HL-LHC. The spades symbol ♠ indicates, that the requirement on m_{jj} is only applied for $N_j \geq 2$.

24% for the HL-LHC.

For most of the measured STXS bins the uncertainties are dominated by statistical uncertainties, while for the extrapolation all STXS bins are dominated by systematic uncertainties. The reduction of the total uncertainties is large compared to the reduction in the inclusive measurements shown in Table 13.1 and 13.2. The increased granularity increases the relative statistical uncertainties on the individual cross sections. The impact of systematic uncertainties is reduced as well, due to the reduced extrapolation and the reduced model dependence of the measurement. The limitation of the HL-LHC measurement by systematic uncertainties can therefore probably be reduced in future analyses by using a finer granularity of STXS bins in the fit model.

No significant deviation from the SM prediction is observed for the STXS measurement in the $H \rightarrow \tau\tau$ decay mode based on the data set recorded from 2015 to 2018.

The $H \rightarrow \tau\tau$ analyses are used for combined Higgs boson cross section measurements. This

thesis discusses the results of these combined measurements based on ten different analyses (see Chapter 11). The combined measured global signal strength μ of the Higgs boson for the most recent combination is:

$$\mu = 1.06 \pm 0.06 = \pm 0.03(\text{stat.}) \pm 0.03(\text{exp.}) \pm 0.04(\text{sig.}) \pm 0.02(\text{bkg.}) . \quad (13.1)$$

This is the most precise measurement of the Higgs boson coupling with the ATLAS detector. The uncertainties of the measurements are dominated by theoretical uncertainties on the signal prediction.

Table 13.4 summarizes the measurement of the six dominant Higgs boson production modes.

Production Mode	σ_i [pb]	Stat. Unc. [pb]	Syst. Unc. [pb]
ggF	45.7	± 1.8	± 2.4
VBF	4.0	± 0.3	± 0.4
WH	1.56	± 0.20	± 0.17
ZH	0.70	± 0.13	± 0.11
$t\bar{t}H + tH$	0.56	± 0.08	± 0.08

Table 13.4: Cross sections and uncertainties of the inclusive production modes for the combined measurement of the Higgs boson with the 2015-2108 data set.

The most precised measured production mode is the ggF production mode with an relative uncertainty of 7%, followed by the VBF production mode with 12%. The uncertainties on the remaining production modes vary between 17% and 25%.

No significant deviation from the SM predictions is observed for the measurement of the combined inclusive production mode cross sections.

A combined measurement within the STXS framework is performed estimating 41 independent PoIs. This measurement corresponds to the measurement with the finest granularity of the Higgs boson sector of the SM and can be used for several interpretations, as e.g. within SMEFT. No significant deviation from the SM predictions is observed for the cross section measurement of these 41 STXS bins.

An interpretation of these results within SMEFT is discussed in this thesis. The interpretation considers SMEFT operators up to the mass dimension six and is based on the Warsaw basis. Lepton and baryon conservation as well as SM gauge symmetries are assumed. To reduce the number of free parameters in the fit model, only CP even operators with a contribution of at least 0.1% to at least one measured STXS bin at their SM prediction corresponding to a Wilson coefficient $c_i = 1$ are considered. Finally, an eigenvector decomposition is performed to find a reduced basis which can be constrained by the measurement. The basis is built by linear combinations of the Wilson coefficients of the considered SMEFT operators. In the measurement of this reduced set of 13 parameters, no significant deviations from the SM predictions are observed.

A APPENDIX

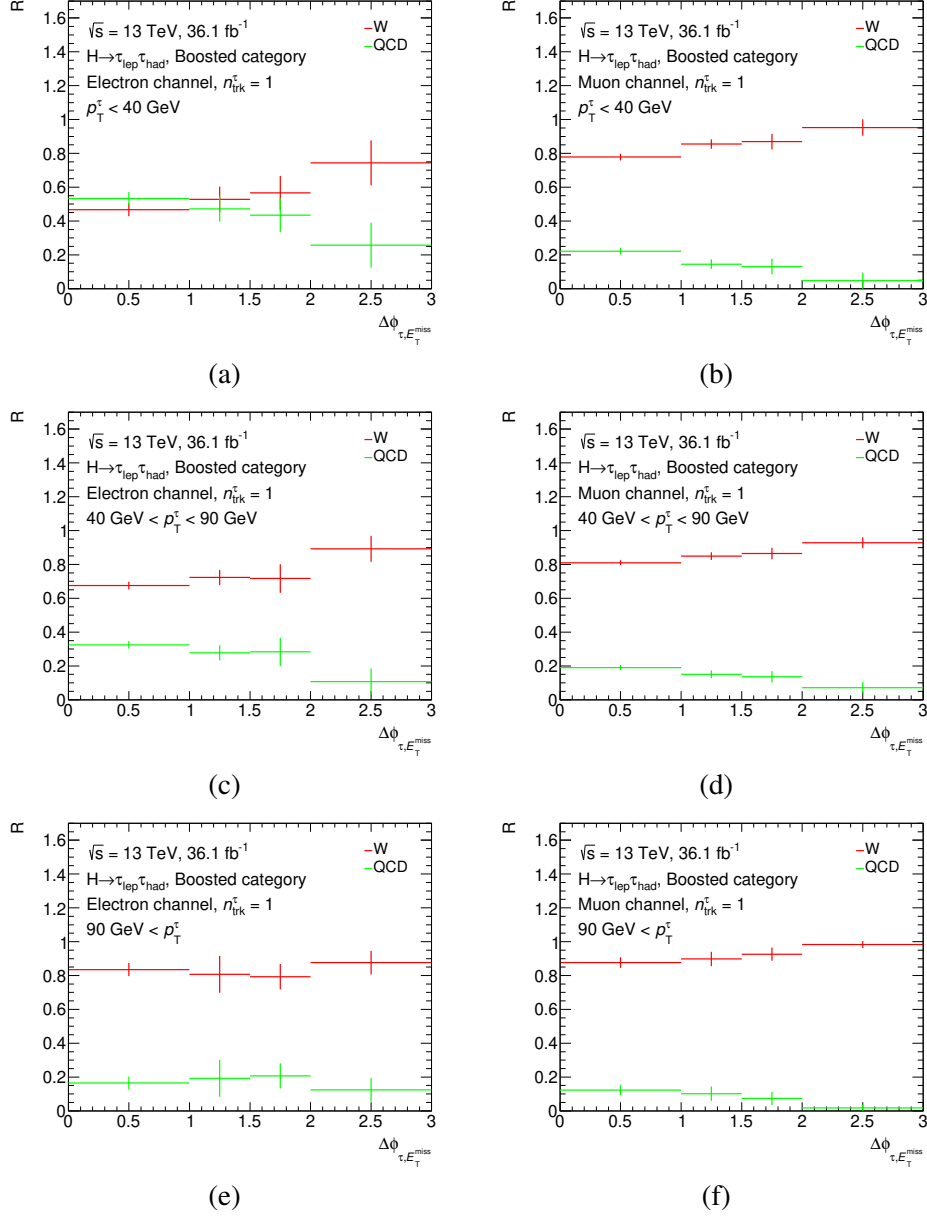


Figure A.1: $R_{\text{multi-jet}}$ values for electrons (left) and muons (right) for events with $n_{\text{trk}}^{\tau} = 1$ in the boosted inclusive category of the $H \rightarrow \tau\tau$ analysis based on the 2015-2016 data set as function of $\Delta\phi_{\tau, E_T^{\text{miss}}}$. From top to bottom estimated for increasing values of p_T^{τ} [224].

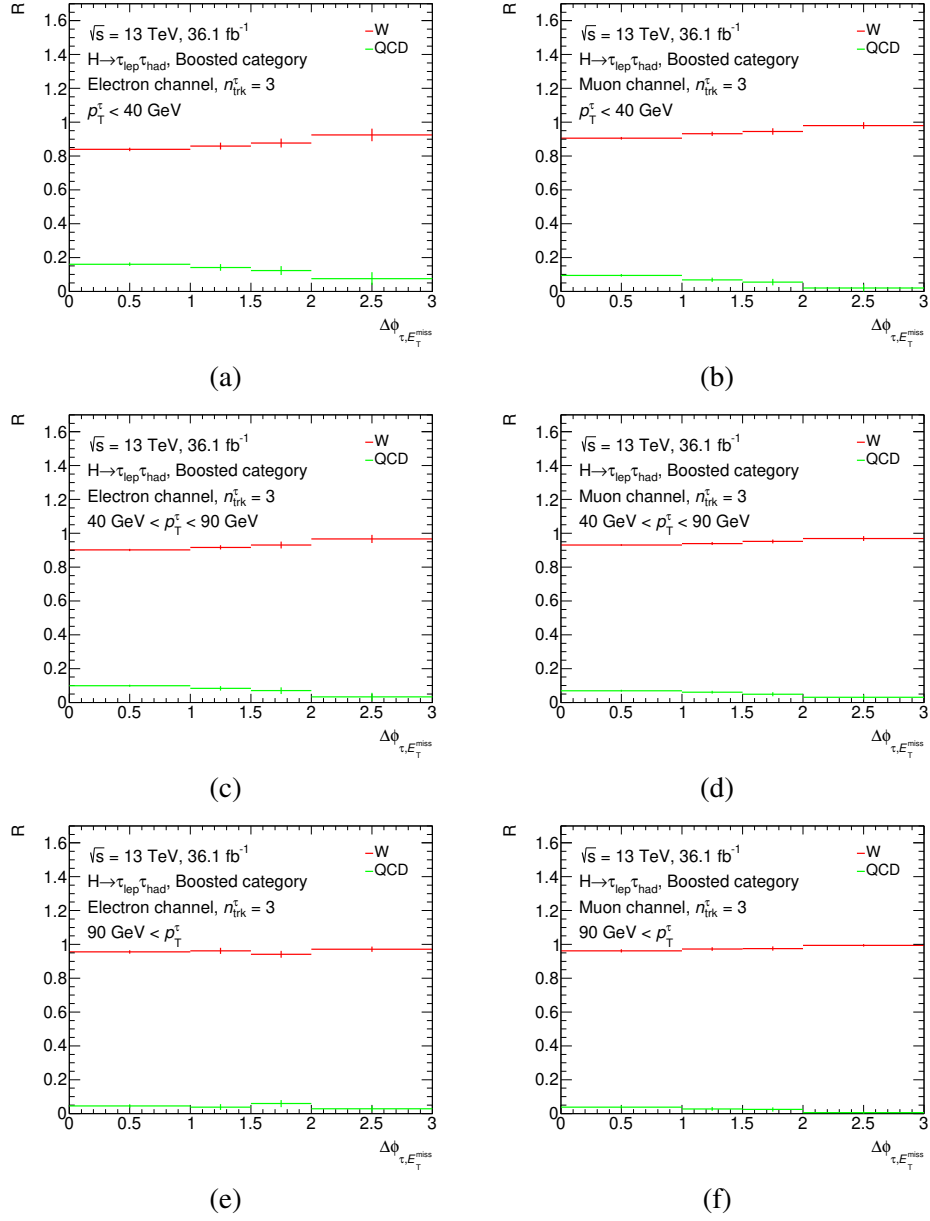


Figure A.2: $R_{\text{multi-jet}}$ values for electrons (left) and muons (right) for events with $n_{\text{trk}}^{\tau} = 3$ in the boosted inclusive category of the $H \rightarrow \tau\tau$ analysis based on the 2015-2016 data set as function of $\Delta\phi_{\tau, E_T^{\text{miss}}}$. From top to bottom estimated for increasing values of p_T^{τ} [224].

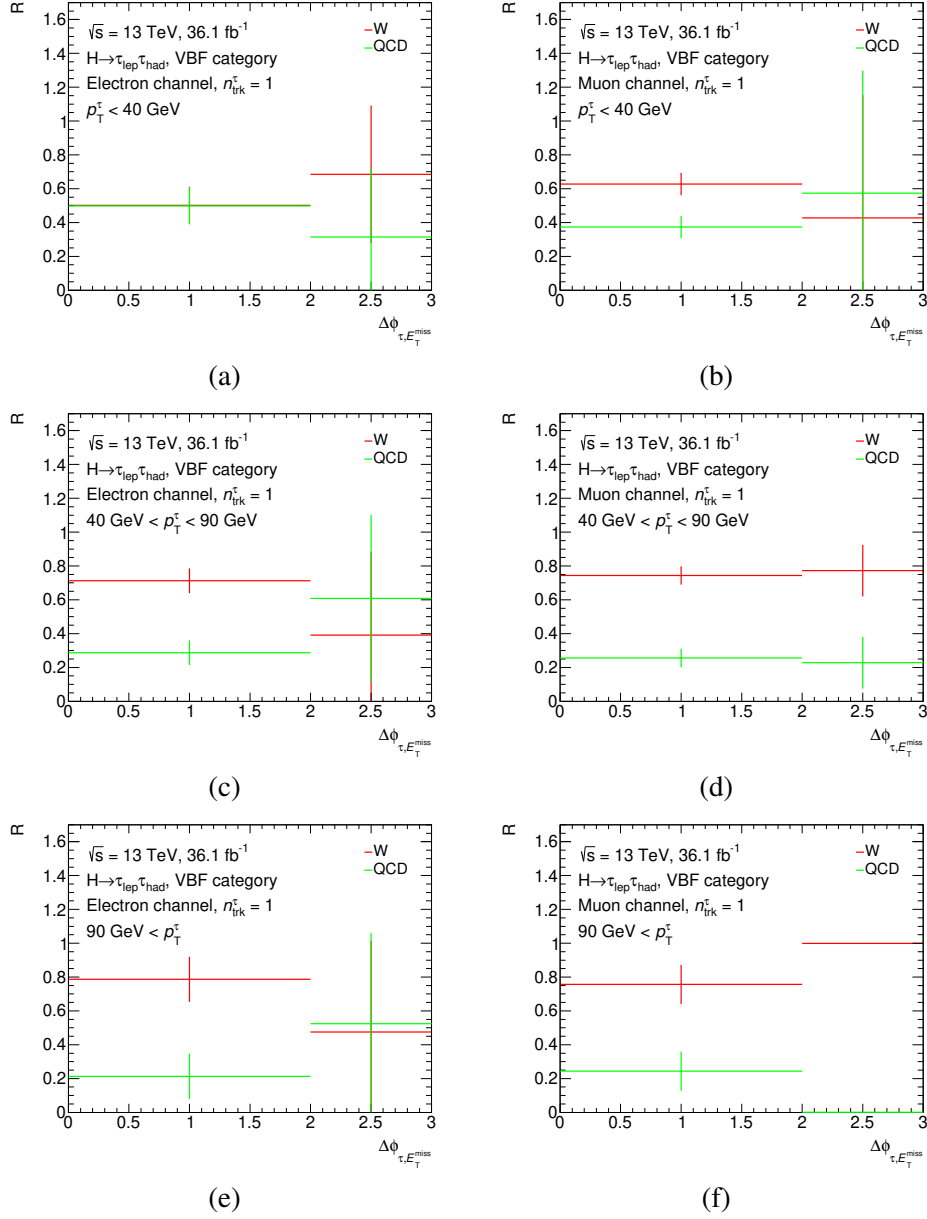


Figure A.3: $R_{\text{multi-jet}}$ values for electrons (left) and muons (right) for events with $n_{\text{trk}}^\tau = 1$ in the VBF inclusive category of the $H \rightarrow \tau\tau$ analysis based on the 2015-2016 data set as function of $\Delta\phi_{\tau, E_T^{\text{miss}}}$. From top to bottom estimated for increasing values of p_T^τ [224].

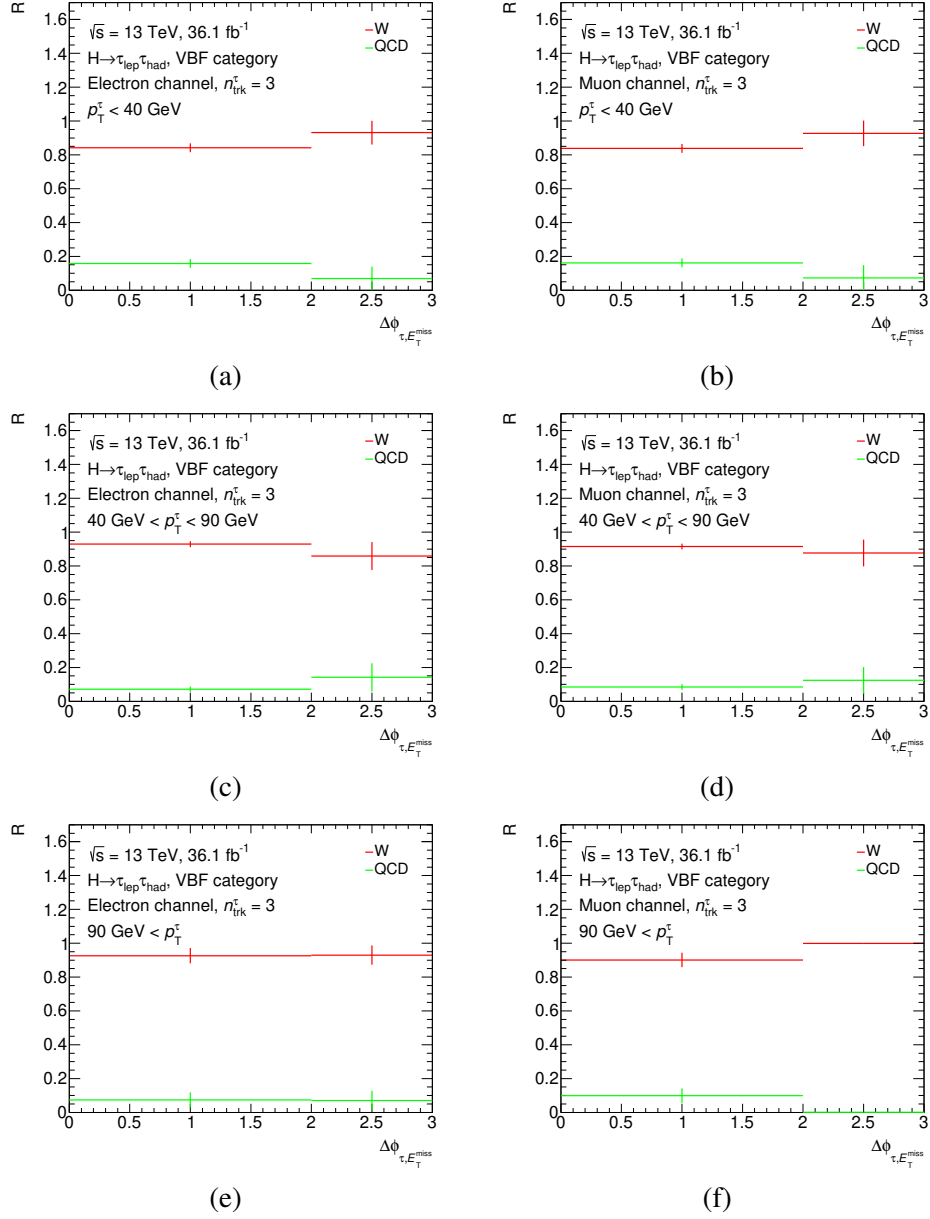


Figure A.4: $R_{\text{multi-jet}}$ values for electrons (left) and muons (right) for events with $n_{\text{trk}}^\tau = 3$ in the VBF inclusive category of the $H \rightarrow \tau\tau$ analysis based on the 2015-2016 data set as function of $\Delta\phi_{\tau, E_T^{\text{miss}}}$. From top to bottom estimated for increasing values of p_T^τ [224].

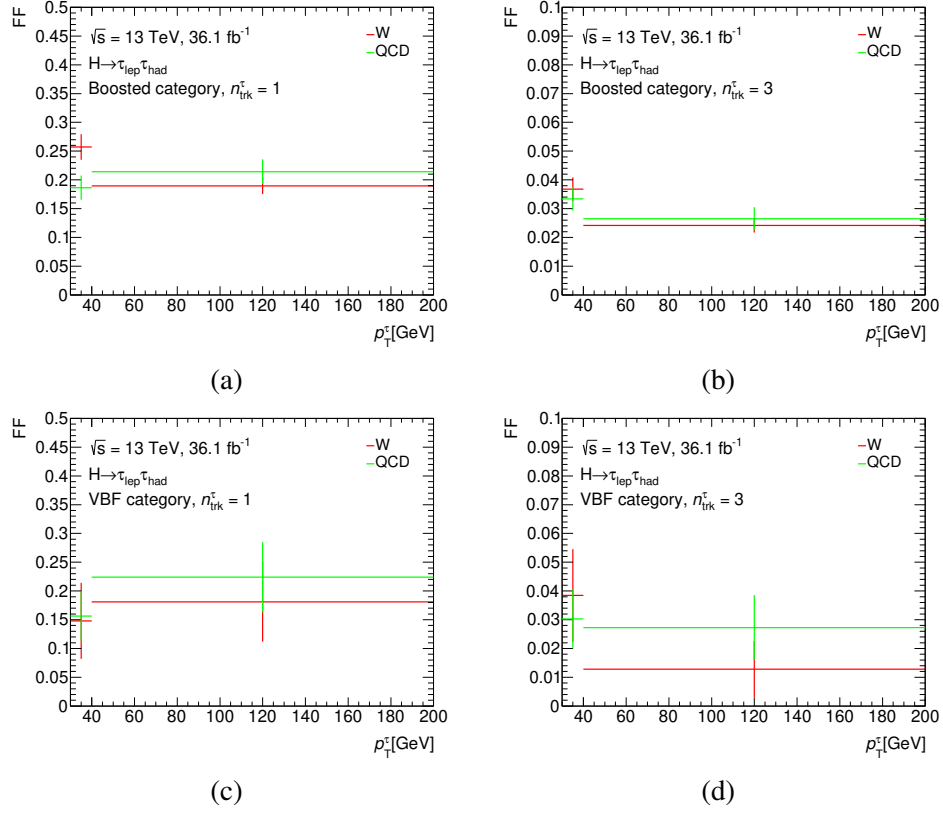


Figure A.5: FF values for for events with $n_{\text{trk}}^\tau = 1$ (left) and $n_{\text{trk}}^\tau = 3$ (right) in the VBF (top) and boosted (bottom) inclusive categories of the $H \rightarrow \tau\tau$ analysis based on the 2015-2016 data set as function of p_T^τ [224].

	Boosted						VBF					
	$\tau_1 \tau_1$			$\tau_1 \tau_h$			$\tau_1 \tau_1$			$\tau_1 \tau_h$		
	High- $p_T^{\tau\tau}$	Low- $p_T^{\tau\tau}$	High- $p_T^{\tau\tau}$	Low- $p_T^{\tau\tau}$	High- $p_T^{\tau\tau}$	Low- $p_T^{\tau\tau}$	Tight	Loose	High- $p_T^{\tau\tau}$	Tight	Loose	Loose
ggF $\Delta\mu$	7.4%	6.8%	7.5%	6.8%	7.4%	6.6%	4.3%	3.6%	4.4%	2.9%	5.2%	2.2%
ggF $\Delta\phi$	6.5%	7.2%	7.2%	6.1%	7.1%	6.2%	4.3%	3.5%	4.3%	2.7%	5.2%	2.2%
ggF $\Delta_{cut}^{0/1}$	5.6%	6.1%	5.6%	6.1%	5.6%	6.3%	2.6%	2.5%	2.9%	2.7%	2.8%	1.8%
ggF $\Delta_{cut}^{0/1}$	8.9%	5.3%	9.2%	4.9%	8.7%	3.8%	7.4%	5.3%	6.9%	1.6%	10.1%	2.2%
ggF VBF	0.4%	0.5%	0.2%	0.5%	0.3%	0.6%	9.6%	10.8%	9.2%	11.0%	8.0%	14.0%
ggF $N_j < 3$	0.1%	0.03%	0.1%	0.02%	0.1%	0.2%	0.5%	-0.8%	0.6%	-1.3%	0.5%	-4.6%
ggF $p_T^H 60 \text{ GeV}$	9.7%	7.4%	9.7%	7.2%	9.7%	7.2%	7.9%	7.3%	7.9%	3.2%	9.5%	4.2%
ggF $p_T^H 120 \text{ GeV}$	13.8%	7.4%	13.8%	6.9%	13.8%	6.2%	11.0%	10.0%	10.9%	3.4%	13.8%	6.0%
ggF m_{top}	7.5%	0.3%	8.1%	0.8%	5.8%	0.1%	6.4%	3.1%	6.2%	0.9%	6.1%	0.1%
ggF PDF, α_s	1.9%	1.5%	1.9%	1.5%	5.5%	5.5%	2.0%	1.5%	1.7%	1.4%	2.3%	2.0%
ggF PS	-4.8%	-8.8%	-24.0%	-5.1%	4.7%	9.2%	26.0%	23.0%	-13.0%	18.0%	1.8.0%	9.0%
VBF $\Delta\mu$	1.4%	0.7%	1.4%	0.6%	2.1%	1.0%	1.6%	0.8%	1.1%	0.5%	1.4%	2.2%
VBF PDF, α_s	0.3%	0.3%	0.4%	0.2%	1.0%	1.4%	0.5%	0.3%	0.4%	0.2%	0.9%	1.3%
VBF PS	9.9%	4.3%	5.4%	7.9%	2.2%	18.0%	0.7%	1.3%	10.7%	5.5%	9.2%	-10.0%
WH $\Delta\mu$	5.2%	1.9%	7.3%	7.5%	5.2%	2.5%	7.0%	13.0%	14.0%	9.0%	8.0%	7.0%
WH PDF, α_s	0.5%	0.3%	0.4%	0.2%	1.5%	1.8%	0.8%	1.7%	0.8%	1.1%	2.3%	2.0%
ZH $\Delta\mu$	5.2%	1.9%	7.3%	2.6%	5.2%	2.5%	7.0%	25.0%	14.0%	9.0%	8.0%	7.0%
ZH PDF, α_s	0.5%	0.5%	0.4%	0.2%	1.5%	1.8%	0.9%	1.7%	0.8%	1.1%	2.3%	2.0%

Table A.1: The acceptance uncertainties of the theoretical uncertainties on the signal in the different event selection for the $H \rightarrow \tau\tau$ analysis based on the data set recorded in 2015 and 2016. The left column defines the evaluated production mode and uncertainty source, which are defined in Section 9.4.

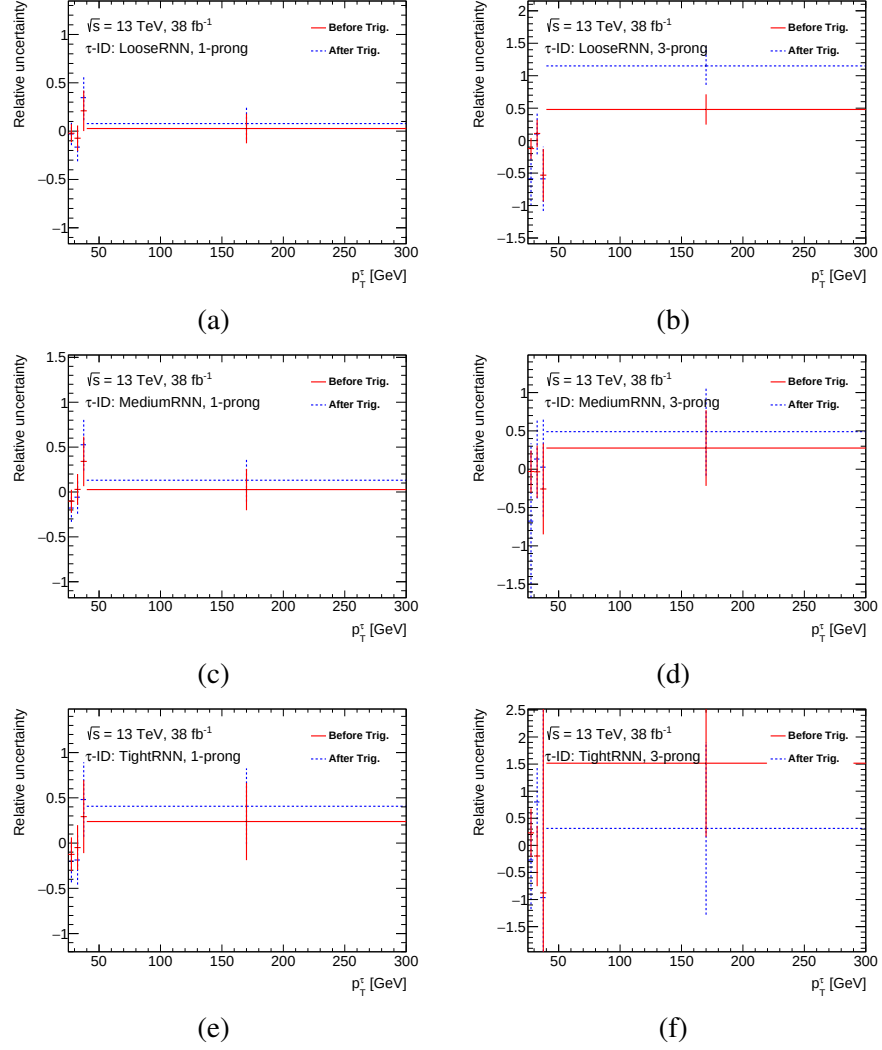


Figure A.6: Shape variation on $R_{\text{OS/SS}}$ factors for the tau25 tracktwoEF OR tracktwoMVA trigger with the 2018 data set and the RNN τ_h -identification method. From top to bottom for the Loose, Medium and Tight working point. On the left-side for τ_h -leptons with one charged track and on the right-side with three charged tracks. The $R_{\text{OS/SS}}$ value for the signal region before applying the is shown in red and after applying the τ_h -trigger requirement in dashed blue.

	Nuisance Parameter	Description
Theoretical uncertainties on signal prediction	Theo_ggH_shower	ggF, parton shower, parametrization
	Theo_ggH_sig_qcd_0(1-8)	ggF, QCD scale uncertainties component 0 (1-8)
	Theo_ggH_sig_jetveto	ggF, QCD scale uncertainties for third-jet veto
	Theo_boost_ggH_PSUE	ggF, parton shower and underlying event in boosted categories
	Theo_vbf_ggH_PSUE	ggF, parton shower and underlying event in VBF categories
	Theo_boost_VBFH_PSUE	VBF, parton shower and underlying event in boosted categories
	Theo_vbf_VBFH_PSUE	VBF, parton shower and underlying event in VBF categories
	Theo_boost_VH_PSUE	VH, parton shower and underlying event
	Theo_sig_alphaS	signal, α_s uncertainty
	Theo_sig_pdf_0(1-29)	signal, PDF uncertainty eigenvector 0 (1-29)
	Theo_VBFH_MUR_MUF	VBF, QCD scale uncertainties
	Theo_VBFH_shower	VBF, parton shower, parametrization
	Theo_VH_MUR_MUF	VH, QCD scale uncertainties
Theoretical uncertainties on background prediction	ZttTheory_PDF_boost(vbf)_leplep	$Z \rightarrow \tau\tau$, PDF, $\tau_1\tau_1$ boosted (VBF) categories
	ZttTheory_PDF_boost(vbf)_lephad	$Z \rightarrow \tau\tau$, PDF, $\tau_1\tau_h$ boosted (VBF) categories
	ZttTheory_PDF_boost(vbf)_hadhad	$Z \rightarrow \tau\tau$, PDF, $\tau_h\tau_h$ boosted (VBF) categories
	ZttTheory_PDF_Relative_boost(vbf)	$Z \rightarrow \tau\tau$, migration PDF, boosted (VBF) categories
	ZttTheory_CKK_boost(vbf)_leplep	$Z \rightarrow \tau\tau$, merging scale, $\tau_1\tau_1$ boosted (VBF) categories
	ZttTheory_CKK_boost(vbf)_lephad	$Z \rightarrow \tau\tau$, merging scale, $\tau_1\tau_h$ boosted (VBF) categories
	ZttTheory_CKK_boost(vbf)_hadhad	$Z \rightarrow \tau\tau$, merging scale, $\tau_h\tau_h$ boosted (VBF) categories
	ZttTheory_CKK_Relative_boost(vbf)	$Z \rightarrow \tau\tau$, migration merging scale, boosted (VBF) categories
	ZttTheory_MGvsSH_boost(vbf)_leplep	$Z \rightarrow \tau\tau$, parton shower, $\tau_1\tau_1$ boosted (VBF) categories
	ZttTheory_MGvsSH_boost(vbf)_lephad	$Z \rightarrow \tau\tau$, parton shower, $\tau_1\tau_h$ boosted (VBF) categories
	ZttTheory_MGvsSH_boost(vbf)_hadhad	$Z \rightarrow \tau\tau$, parton shower, $\tau_h\tau_h$ boosted (VBF) categories
	ZttTheory_MGvsSH_Relative_boost(vbf)	$Z \rightarrow \tau\tau$, migration parton shower, boosted (VBF) categories
	ZttTheory_MUR_MUF_boost(vbf)_leplep	$Z \rightarrow \tau\tau$, QCD scale uncertainties, $\tau_1\tau_1$ boosted (VBF) categories
	ZttTheory_MUR_MUF_boost(vbf)_lephad	$Z \rightarrow \tau\tau$, QCD scale uncertainties, $\tau_1\tau_h$ boosted (VBF) categories
	ZttTheory_MUR_MUF_boost(vbf)_hadhad	$Z \rightarrow \tau\tau$, QCD scale uncertainties, $\tau_h\tau_h$ boosted (VBF) categories
	ZttTheory_MUR_MUF_Relative_boost(vbf)	$Z \rightarrow \tau\tau$, migration QCD scale uncertainties, boosted (VBF) categories
	ZttTheory_QSF_boost(vbf)_leplep	$Z \rightarrow \tau\tau$, resummation scale, $\tau_1\tau_1$ boosted (VBF) categories
	ZttTheory_QSF_boost(vbf)_lephad	$Z \rightarrow \tau\tau$, resummation scale, $\tau_1\tau_h$ boosted (VBF) categories
	ZttTheory_QSF_boost(vbf)_hadhad	$Z \rightarrow \tau\tau$, resummation scale, $\tau_h\tau_h$ boosted (VBF) categories
	ZttTheory_QSF_Relative_boost(vbf)	$Z \rightarrow \tau\tau$, migration resummation scale, boosted (VBF) categories
	Ztt_mjj_rew	$Z \rightarrow \tau\tau$, m_{jj} reweighting
	Zll_mjj_rew	$Z \rightarrow ll$, m_{jj} reweighting

Table A.2: NPs of the theoretical uncertainties in the different event selection for the $H \rightarrow \tau\tau$ analysis based on the data set recorded in 2015 and 2016. NPs with similar naming and source are condensed in one row using () for the alternative NPs.

	Nuisance Parameter	Description
Experimental	EG_RESOLUTION_ALL	electron energy resolution
	EG_SCALE_ALLCORR	electron energy scale
	EG_SCALE_E4SCINTILLATOR	electron energy scale, scintillator
	EG_SCALE_LARCALIB_2015(2016)PRE	electron energy scale, intercalibration 2015 (2016)
	EG_SCALE_LARTEMP_2015(2016)PRE	electron energy scale, LAr temperature 2015 (2016)
	EL_RECO(ISO)_TOTAL	electron reconstruction (isolation) efficiency
	EL_ID_TOTAL(NONPROMPT_D0)	electron identification efficiency (non-prompt electrons)
	EL_EFF_TRIG_2015(2016)_TOTAL	electron trigger for 2015 (2016)
	MUON_EFF_STAT(SYST)	muon identification efficiency statistic (systematic)
	MUON_ISO_STAT(SYST)	muon isolation efficiency statistic (systematic)
	MUON_EFF_TTVA_STAT(SYST)	muon track-to-vertex association efficiency statistic (systematic)
	MUONS_ID	muon momentum resolution, inner detector (ID)
	MUONS_MS	muon momentum resolution, muon spectrometer (MS)
	MUONS_SAGITTA_RESBIAS	muon momentum scale, sagitta, residual bias
	MUONS_SAGITTA_RHO	muon momentum scale, sagitta, p_T imbalance ID and MS
	MUONS_SCALE	muon momentum scale
	MUON_EFF_TRIG_STAT(SYST)2015(2016)	muon trigger statistic (systematic) for 2015(2016)
	JER_CROSSCALIBFWD	jet energy resolution, Run-1 calibration, forward region
	JER_NOISEFWD	jet energy resolution, noise forward region
	JER_NP_0-8	jet energy resolution
	JES_BJES	jet energy scale, b-jets
	JES_EffectiveNP_1 (2-8)	jet energy scale, component 1 (2-8)
	JES_EtaInter_Model(NonClosure,Stat)	jet energy scale, η -intercalibration, model (non-closure,statistic)
	JES_Flavor_Comp(_zll)	jet energy scale, flavour-dependent jet composition ($Z \rightarrow ll$)
	JES_Flavor_Resp	jet energy scale, flavour-dependent jet response
	JES_HighPt	jet energy scale, jets with $p_T > 1.8$ TeV
	JES_PU_OffsetMu	jet energy scale, pile-up, interactions per crossing
	JES_PU_OffsetNPV	jet energy scale, pile-up, number primary vertices
	JES_PU_PtTerm(Rho)	jet energy scale, pile-up, p_T dependence (p_T density)
	JES_PunchThrough	jet energy scale, jets not fully contained in calorimeter
	JVT	efficiency jet vertex tagger
	FT_EFF_Eigen_b_0(1,2)	b-tagging efficiency, b-jets component 0 (1,2)
	FT_EFF_Eigen_c_0(1-3)	b-tagging efficiency, c-jets component 0 (1-3)
	FT_EFF_Eigen_light_0(1-4)	b-tagging efficiency, light quark jets component 0 (1-4)
	FT_EFF_extrapolation(_from_charm)	extrapolation of scale factors from b(c)-jets to τ -jets
	TAU_EFF_ELEORL_TRUEELE(TRUEHADTAU)	τ_h electron overlap removal efficiency, real electron (τ_h)
	TAU_EFF_ID_TOTAL(HIGHPT)	τ_h identification efficiency total (for high p_T)
	TAU_EFF_RECO_TOTAL(HIGHPT)	τ_h reconstruction efficiency total (for high p_T)
	TAU_TES_INSITU	τ_h energy scale, in-situ measurement
	TAU_TES_DETECTOR(MODEL)	τ_h energy scale, detector geometry (GEANT4 model)
	TAU_EFF_TRIG_STATDATA(MC)2015(2016)	τ_h trigger data (simulation) statistic for 2015(2016)
	TAU_EFF_TRIG_SYST2015(2016)	τ_h trigger systematic for 2015(2016)
	MET_SoftTrk_ResoPara(Perp)	E_T^{miss} soft term resolution, parallel (perpendicular)
	MET_SoftTrk_Scale	E_T^{miss} soft term scale
	PRW_DATASF	pile-up reweighting, scale factor
	ll_fakes_stat	$\tau_l \tau_l$ misidentified τ background statistic
	ll_fakes_nonclosure	$\tau_l \tau_l$ misidentified τ background closure test
	lh_fakes_stat_boost(vbf)(_top)	$\tau_l \tau_h$ misidentified τ background, boosted (vbf) category (top cr) statistic
	lh_fakes_rvar_boost(vbf)	$\tau_l \tau_h$ misidentified τ background $R_{\text{multi-jet}}$ systematic
	lh_fakes_nonclosure	$\tau_l \tau_h$ misidentified τ background closure test
	hh_fakes_contamination	$\tau_h \tau_h$ misidentified τ , CR contamination
	hh_fakes_dPhiWeightStat	$\tau_h \tau_h$ misidentified τ statistic
	hh_fakes_nonclosure	$\tau_h \tau_h$ misidentified τ background closure test

Table A.3: NPs of the experimental in the different event selection for the $H \rightarrow \tau\tau$ analysis based on the data set recorded in 2015 and 2016. NPs with similar naming and source are condensed in one row using () for the alternative NPs.

	Nuisance Parameter	Description
Theoretical uncertainties on signal prediction	Theory sig ggh shower	ggF, parton shower, parametrization
	Theory sig vbfb shower	VBF, parton shower, parametrization
	Theory sig wh shower	WH, parton shower, parametrization
	Theory sig zh shower	ZH, parton shower, parametrization
	Theory sig tth shower	t̄tH, parton shower, parametrization
	Theory sig ggh me	ggF, matrix element, parametrization
	Theory sig vbfb me	VBF, matrix element, parametrization
	Theory sig wh me	WH, matrix element, parametrization
	Theory sig zh me	ZH, matrix element, parametrization
	Theory sig tth me	t̄tH, matrix element, parametrization
	Theory sig ggh fxfx	ggF, matrix element in VBF and VH categories, parametrization
	theory sig alphaS	signal, α_s uncertainty
	PDF4LHC NLO 30 EV0(1-29)	signal, PDF uncertainty eigenvector 0 (1-29)
	theory ggH sig st 4(5)	ggF, QCD scale Stewart-Tackmann third jet-veto
	QCDScale ggF res	ggF, QCD scale resolution
	QCDScale ggF mu	ggF, QCD scale factorization and renormalization
	QCDScale ggF qm t	ggF, QCD scale top mass dependency
	QCDScale ggF mig01(12)	ggF, QCD scale migration 0(1) to 1(2) jet selection
	QCDScale ggF pTH60(120)	ggF, QCD scale migration for p_T^H boarder at 60 (120) GeV
	QCDScale ggF vbf2(3)j	ggF, QCD scale migration 1(2) to 2(3) jet VBF selection
	QCDScale ggF deltaMJJ350(700,1000,1500)	ggF, QCD scale migration for m_{jj} boarder at 350 (700,1000,1500) GeV
	QCDScale qq2H mu	EW qqH, QCD scale factorization and renormalization
	QCDScale qq2H JET01	EW qqH, QCD scale migration 0 to 1 jet selection
	QCDScale qq2H Mjj60(120,350,700,1000,1500)	EW qqH, QCD scale migration for m_{jj} boarder at 60 (120,...,1500) GeV
	QCDScale qq2H pT200	EW qqH, QCD scale migration for p_T^H boarder at 200 GeV
	QCDScale qq2H pTjH25	EW qqH, QCD scale migration for p_T^{Hjj} boarder at 25 GeV
	QCDScale tth mu	t̄tH, QCD scale factorization and renormalization
	QCDScale tth pTH60(120,200,300,450)	t̄tH, QCD scale migration for p_T^H boarder at 60 (120,...,450) GeV
Theoretical uncertainties on background prediction	theory z mur muf envelope [region]	Z production, QCD scale uncertainties for each [region] as ll vbf 0
	theory z ckk (tth)	Z production, merging scale (for tth categories)
	theory z qsf (tth)	Z production, resummation scale (for tth categories)
	theory z pdfset (tth)	Z production, PDF set comparison (for tth categories)
	theory z alphaS	Z production, α_s uncertainty
	theory z pdf [region]	Z production, PDF uncertainties for each [region] as ll vbf 0
	theory z EWK / QCD uncertainty	Z production, EWK to QCD ratio
	theory top parton shower	top quark production, parton shower, parametrization
	theory top me	top quark production, matrix-element, parametrization
	theory top ISR (FSR)	top quark production, initial (final) state radiation
	theory top PDF1(2-30)	top quark production, PDF uncertainty eigenvector 1 (2-30)

Table A.4: NPs of the theoretical uncertainties in the different event selection for the $H \rightarrow \tau\tau$ analysis based on the data set recorded in 2015 and 2018. NPs with similar naming and source are condensed in one row using () for the alternative NPs, while [region] illustrates that an individual NP per signal category is defined.

Experimental Nuisance Parameter	Description
EG RESOLUTION ALL	electron energy resolution
EG SCALE ALL	electron energy scale
EL EFF ID CorrUncertaintyNP0(1-15)	electron correlated identification efficiency component 0(1-15)
EL EFF ID SIMPLIFIED UncorrUncertaintyNP0(1-11)	electron uncorrelated identification efficiency component 0(1-11)
MUONS ID	muon momentum resolution, inner detector (ID)
MUONS MS	muon momentum resolution, muon spectrometer (MS)
MUONS SCALE	muon momentum scale
MUONS SAGITTA RESBIAS	muon momentum scale, sagitta, residual bias
MUONS SAGITTA RHO	muon momentum scale, sagitta, p_T imbalance ID and MS
MUON ISO STAT(SYS)	muon isolation efficiency statistic (systematic)
MUON EFF RECO STAT(SYST) (LOWPT)	muon reconstruction efficiency statistic (systematic) (for low p_T)
MUON EFF TrigStat(Syst)Uncertainty	muon trigger statistic (systematic)
JET EffectiveNP Detector1(2)	jet energy scale, detector component 1(2)
JET EffectiveNP Mixed1(2,3)	jet energy scale, component 1(2,3)
JET EffectiveNP Modelling1(2-4)	jet energy scale, modelling component 1(2-4)
JET EffectiveNP Statistical1(2-6)	jet energy scale, statistic component 1(2-6)
JET EtaIntercalibration Modelling(TotalStat)	jet energy scale, η -intercalibration, modelling (statistic)
JET EtaIntercalibration NonClosure 2018data (highE)	jet energy scale, η -intercalibration, non-closure for 2018 (high energy)
JET EtaIntercalibration NonClosure neg(pos)Eta	jet energy scale, η -intercalibration, non-closure for negative (positive) η -values
JET Flavor Composition(Response)	jet energy scale, flavor-dependent jet composition (response)
JET PileUp OffsetMu	jet energy scale, pile-up, interactions per crossing
JET PileUp OffsetNPV	jet energy scale, pile-up, number primary vertices
JET PileUp PtTerm(Rho)	jet energy scale, pile-up, p_T dependence (p_T density)
JET PunchThrough MC16	jet energy scale, jets not fully contained in calorimeter for 2016
JET SingleParticle HighPt	jet energy scale for large p_T
JER DataVsMC_MC16	jet energy resolution comparison for data and simulation
JER EffectiveNP 1(2-11)	jet energy resolution component 1(2-11)
JER EffectiveNP 12restTerm	jet energy resolution rest term
JET (f)JvtEfficiency	efficiency jet (forward) vertex tagger
FT EFF Eigen b 0(1,2) (hh)	b-tagging efficiency, b-jets component 0 (1,2) (for $\tau_h \tau_h$)
FT EFF Eigen c 0(1,2) (hh)	b-tagging efficiency, c-jets component 0 (1,2) (for $\tau_h \tau_h$)
FT EFF Eigen light 0(1-3) (hh)	b-tagging efficiency, light quark jets component 0 (1-3) (for $\tau_h \tau_h$)
FT EFF extrapolation(from charm) (hh)	extrapolation of scale factors from b(c)-jets to τ -jets (for $\tau_h \tau_h$)
FT EFF Eigen light 4 hh	b-tagging efficiency, light quark jets component 4 for $\tau_h \tau_h$
TAUS TRUEELECTRON EFF ELEBDT STAT(SYST)	τ_h electron overlap removal efficiency statistic (systematic), real electron
TAUS TRUEHADTAU EFF ELEOLR STAT(SYST)	τ_h electron overlap removal efficiency statistic (systematic), real τ_h
TAUS TRUEHADTAU EFF RECO TOTAL	τ_h reconstruction efficiency total
TAU EFF ID SYST(HIGHPT)	τ_h identification efficiency systematic (for high p_T)
TAUS TRUEHADTAU EFF RNNID 1(3)PRONGSTATSYSTPT2025	τ_h identification efficiency statistic for 1(3)-prong, $20 < p_T < 25$ GeV
TAUS TRUEHADTAU EFF RNNID 1(3)PRONGSTATSYSTPT2530	τ_h identification efficiency statistic for 1(3)-prong, $25 < p_T < 30$ GeV
TAUS TRUEHADTAU EFF RNNID 1(3)PRONGSTATSYSTPT3040	τ_h identification efficiency statistic for 1(3)-prong, $30 < p_T < 40$ GeV
TAUS TRUEHADTAU EFF RNNID 1(3)PRONGSTATSYSTPTGE40	τ_h identification efficiency statistic for 1(3)-prong, $40 \text{ GeV} < p_T$
TAUS SME TES INSITUEXP(FIT)	τ_h energy scale, in-situ expected (observed) measurement
TAUS SME TES PHYSICSLIST(MODELCLOSURE)	τ_h energy scale, physics model (GEANT4 model)
TAUS SME TES DETECTOR	τ_h energy scale, detector model
TAUS TRUEHADTAU EFF TRIGGER STATDATA(MC)161718	τ_h trigger data (simulation) statistic for 2016-2018
TAUS TRUEHADTAU EFF TRIGGER STATDATA(MC)2018	τ_h trigger data (simulation) statistic for 2018
TAUS TRUEHADTAU EFF TRIGGER SYST161718(2018)	τ_h trigger systematic for 2016-2018 (2018)
TAUS TRUEHADTAU EFF TRIGGER SYSTMU161718(2018)	τ_h trigger pile-up for 2016-2018 (2018)
MET SoftTrk ResoPara(Perp)	E_T^{miss} soft term resolution, parallel (perpendicular)
MET SoftTrk Scale	E_T^{miss} soft term scale
ll fake(real) STAT elec(muon)	$\tau \tau$ misidentified τ , statistic for misidentified (real) electron (muon)
ll fake SYST CF	$\tau \tau$ misidentified τ , for charge-flip events
ll fake SYST prompt elec(muon)	$\tau \tau$ misidentified τ , for prompt electron (muon)
ll fake SYST source elec(muon)	$\tau \tau$ misidentified τ , source for electron (muon)
ll fake SYST comp fake elec(muon)	$\tau \tau$ misidentified τ , composition for electron (muon)
ll fake SYST njets fake elec(muon)	$\tau \tau$ misidentified τ , jet multiplicity for electron (muon)
lh fake FF QCD(W) 1(3)prong stat [category]	$\tau \tau$ misidentified τ , FF for QCD multi-jet (W) with 1(3)-prong per category
lh fake IsoFactor 1(3)prong stat [category]	$\tau \tau$ misidentified τ , IF with 1(3)-prong per category
lh fake RQCD 1(3)prong stat [category]	$\tau \tau$ misidentified τ , $R_{\text{multi-jet}}$ with 1(3)-prong per category
lh fake closure 1(3)prong stat [category]	$\tau \tau$ misidentified τ , closure test with 1(3)-prong per category
lh fake mc subtr 1(3)prong stat [category]	$\tau \tau$ misidentified τ , background subtraction with 1(3)-prong per category
hh fake ff composition higheta(ss)	$\tau_h \tau_h$ misidentified τ , FF for high values of η and same-sign events
hh fake ff param	$\tau_h \tau_h$ misidentified τ , FF parametrization
hh fake ff stat 1(3)p (l)nm	$\tau_h \tau_h$ misidentified τ , FF statistic for 1(3)-prong (loose) not medium ID

Table A.5: NPs of the experimental in the different event selection for the $H \rightarrow \tau\tau$ analysis based on the data set recorded in 2015 and 2018. NPs with similar naming and source are condensed in one row using () for the alternative NPs.

Wilson coefficient	Operator	Wilson coefficient	Operator
$c_{H\Box}$	$(H^\dagger H)\Box(H^\dagger H)$	c_{uG}	$(\bar{q}_p\sigma^{\mu\nu}T^A u_r)\tilde{H}G_{\mu\nu}^A$
c_{HDD}	$(H^\dagger D^\mu H)^*(H^\dagger D_\mu H)$	c_{uW}	$(\bar{q}_p\sigma^{\mu\nu}u_r)\tau^I\tilde{H}W_{\mu\nu}^I$
c_{HG}	$H^\dagger H G_{\mu\nu}^A G^{A\mu\nu}$	c_{uB}	$(\bar{q}_p\sigma^{\mu\nu}u_r)\tilde{H}B_{\mu\nu}$
c_{HB}	$H^\dagger H B_{\mu\nu}B^{\mu\nu}$	c'_{ll}	$(\bar{l}_p\gamma_\mu l_t)(\bar{l}_r\gamma^\mu l_s)$
c_{HW}	$H^\dagger H W_{\mu\nu}^I W^{I\mu\nu}$	$c_{qq}^{(1)}$	$(\bar{q}_p\gamma_\mu q_t)(\bar{q}_r\gamma^\mu q_s)$
c_{HWB}	$H^\dagger\tau^I H W_{\mu\nu}^I B^{\mu\nu}$	$c_{qq}^{(3)}$	$(\bar{q}_p\gamma_\mu\tau^I q_r)(\bar{q}_s\gamma^\mu\tau^I q_t)$
c_{eH}	$(H^\dagger H)(\bar{l}_p e_r H)$	c_{qq}	$(\bar{q}_p\gamma_\mu q_t)(\bar{q}_r\gamma^\mu q_s)$
c_{uH}	$(H^\dagger H)(\bar{q}_p u_r \tilde{H})$	$c_{qq}^{(31)}$	$(\bar{q}_p\gamma_\mu\tau^I q_t)(\bar{q}_r\gamma^\mu\tau^I q_s)$
c_{dH}	$(H^\dagger H)(\bar{q}_p d_r \tilde{H})$	c_{uu}	$(\bar{u}_p\gamma_\mu u_r)(\bar{u}_s\gamma^\mu u_t)$
$c_{Hl}^{(1)}$	$(H^\dagger i\overleftrightarrow{D}_\mu H)(\bar{l}_p\gamma^\mu l_r)$	$c_{uu}^{(1)}$	$(\bar{u}_p\gamma_\mu u_t)(\bar{u}_r\gamma^\mu u_s)$
$c_{Hl}^{(3)}$	$(H^\dagger i\overleftrightarrow{D}_\mu^I H)(\bar{l}_p\tau^I\gamma^\mu l_r)$	$c_{qu}^{(1)}$	$(\bar{q}_p\gamma_\mu q_t)(\bar{u}_r\gamma^\mu u_s)$
c_{He}	$(H^\dagger i\overleftrightarrow{D}_\mu H)(\bar{e}_p\gamma^\mu e_r)$	$c_{ud}^{(8)}$	$(\bar{u}_p\gamma_\mu T^A u_r)(\bar{d}_s\gamma^\mu T^A d_t)$
$c_{Hq}^{(1)}$	$(H^\dagger i\overleftrightarrow{D}_\mu H)(\bar{q}_p\gamma^\mu q_r)$	$c_{qu}^{(8)}$	$(\bar{q}_p\gamma_\mu T^A q_r)(\bar{u}_s\gamma^\mu T^A u_t)$
$c_{Hq}^{(3)}$	$(H^\dagger i\overleftrightarrow{D}_\mu^I H)(\bar{q}_p\tau^I\gamma^\mu q_r)$	$c_{qd}^{(8)}$	$(\bar{q}_p\gamma_\mu T^A q_r)(\bar{d}_s\gamma^\mu T^A d_t)$
c_{Hu}	$(H^\dagger i\overleftrightarrow{D}_\mu H)(\bar{u}_p\gamma^\mu u_r)$	c_W	$\epsilon^{IJK}W_\mu^{I\nu}W_\nu^{J\rho}W_\rho^{K\mu}$
c_{Hd}	$(H^\dagger i\overleftrightarrow{D}_\mu H)(\bar{d}_p\gamma^\mu d_r)$	c_G	$f^{ABC}G_\mu^{A\nu}G_\nu^{B\rho}G_\rho^{C\mu}$

Table A.6: Overview of the considered Wilson coefficients and their corresponding operator of dimension six [30].

Parameter	Definition	Eigenvalue	Fit Parameter
$c_{Hq}^{(3)}$	$c_{Hq}^{(3)}$	1900	✓
$c_{HW,HB,HWB,HDD,uW,uB}^{[6]}$	1	$-0.27c_{HW} - 0.84c_{HB} + 0.47c_{HWB} - 0.02c_{uW} - 0.05c_{uB}$	245000 ✓
	2	$-0.96c_{HW} + 0.19c_{HB} - 0.20c_{HWB} + 0.02c_{uB}$	33 ✓
	3	$-0.08c_{HW} + 0.50c_{HB} + 0.86c_{HWB} + 0.07c_{HDD} + 0.03c_{uW} + 0.06c_{uB}$	4 ✓
	4	$0.03c_{HWB} - 0.85c_{HDD} + 0.32c_{uW} + 0.43c_{uB}$	0.017
	5	$-0.01c_{HW} + 0.07c_{HB} + 0.05c_{HWB} - 0.44c_{HDD} - 0.86c_{uW} - 0.23c_{uB}$	0.0077
	6	$-0.01c_{HW} + 0.06c_{HB} + 0.04c_{HWB} - 0.29c_{HDD} + 0.39c_{uW} - 0.87c_{uB}$	0.0025
$c_{HG,uG,uH,top}^{[6]}$	1	$+0.999c_{HG} + 0.038c_{uG}$	176000 ✓
	2	$-0.03c_{HG} + 0.73c_{uG} - 0.03c_{qq}^{(1)} - 0.23c_{qq} - 0.05c_{qq}^{(3)} - 0.54c_{qq}^{(31)} - 0.02c_{uu} - 0.24c_{uu}^{(1)} - 0.04c_{ud}^{(8)} - 0.01c_{qu}^{(1)} - 0.15c_{qu}^{(8)} - 0.04c_{qd}^{(8)} - 0.18c_G + 0.06c_{uH}$	20 ✓
	3	$-0.03c_{HG} + 0.67c_{uG} + 0.04c_{qq}^{(1)} + 0.25c_{qq} + 0.05c_{qq}^{(3)} + 0.55c_{qq}^{(31)} + 0.02c_{uu} + 0.26c_{uu}^{(1)} + 0.03c_{ud}^{(8)} + 0.01c_{qu}^{(1)} + 0.16c_{qu}^{(8)} + 0.03c_{qd}^{(8)} + 0.29c_G + 0.1c_{uH}$	1.3 ✓
	4	$+0.11c_{uG} + 0.01c_{qq} - 0.018c_{qq}^{(3)} + 0.029c_{qq}^{(31)} + 0.012c_{uu}^{(1)} - 0.993c_{uH}$	0.14
	5	$+0.02c_{qq} - 1.0c_{qq}^{(3)} + 0.06c_{qq}^{(31)} + 0.03c_{uu}^{(1)} + 0.02c_{qu}^{(8)} + 0.02c_{uH}$	0.02
	6	$+0.07c_{uG} - 0.02c_{qq}^{(1)} + 0.07c_{qq} + 0.03c_{qq}^{(3)} + 0.32c_{qq}^{(31)} + 0.06c_{uu}^{(1)} + 0.04c_{ud}^{(8)} + 0.08c_{qu}^{(8)} + 0.04c_{qd}^{(8)} - 0.94c_G + 0.02c_{uH}$	0.0092
$c_{Hl^{(1)},He}^{[1]}$	$+0.78c_{Hl}^{(1)} - 0.62c_{He}$	2.6	✓
$c_{Hl^{(1)},He}^{[2]}$	$+0.62c_{Hl}^{(1)} + 0.78c_{He}$	0.056	
$c_{Hu,Hd,Hq^{(1)}}^{[1]}$	$-0.87c_{Hu} + 0.26c_{Hd} + 0.42c_{Hq}^{(1)}$	59	✓
$c_{Hu,Hd,Hq^{(1)}}^{[2]}$	$+0.41c_{Hu} - 0.09c_{Hd} + 0.91c_{Hq}^{(1)}$	0.10	
$c_{Hu,Hd,Hq^{(1)}}^{[3]}$	$-0.28c_{Hu} - 0.96c_{Hd} + 0.03c_{Hq}^{(1)}$	0.0018	
$c_{Hl^{(3)},ll'}^{[1]}$	$0.87c_{Hl}^{(3)} - 0.50c_{ll}'$	27	✓
$c_{Hl^{(3)},ll'}^{[2]}$	$0.50c_{Hl}^{(3)} + 0.87c_{ll}'$	0.33	

Table A.7: Eigenvectors and eigenvalues for the rotated subspace. Eigenvectors with a eigenvalue below 0.0001 are not shown, such as $c_{HG,uG,uH,top}^{[7]}$ and $c_{HG,uG,uH,top}^{[17]}$. The last column marks the parameter which enter as PoIs in the fit model [30].

REFERENCES

- [1] UA1 Collaboration. “Experimental Observation of Lepton Pairs of Invariant Mass Around $95 \text{ GeV}/c^2$ at the CERN SPS Collider”. In: *Phys. Rev. Lett.* 126B (1983), p. 398. DOI: [10.1016/0370-2693\(83\)90188-0](https://doi.org/10.1016/0370-2693(83)90188-0).
- [2] UA1 Collaboration. “Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540 \text{ GeV}$ ”. In: *Phys. Rev. Lett.* 122B (1983), p. 611. DOI: [10.1016/0370-2693\(83\)91177-2](https://doi.org/10.1016/0370-2693(83)91177-2).
- [3] CDF Collaboration. “Observation of top quark production in $\bar{p}p$ collisions”. In: *Phys. Rev. Lett.* 74 (1995), p. 534. DOI: [10.1103/PhysRevLett.74.2626](https://doi.org/10.1103/PhysRevLett.74.2626). arXiv: [hep-ex/9503002](https://arxiv.org/abs/hep-ex/9503002) [hep-ex].
- [4] UA2 Collaboration. “Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN anti-p p Collider”. In: *Phys. Rev. Lett.* 129B (1983), p. 130. DOI: [10.1016/0370-2693\(83\)90744-X](https://doi.org/10.1016/0370-2693(83)90744-X).
- [5] UA2 Collaboration. “Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider”. In: *Phys. Rev. Lett.* 122B (1983), p. 476. DOI: [10.1016/0370-2693\(83\)91605-2](https://doi.org/10.1016/0370-2693(83)91605-2).
- [6] D0. “Observation of the Top Quark”. In: *Phys. Rev. Lett.* 74.14 (1995), pp. 26322637. ISSN: 1079-7114. DOI: [10.1103/physrevlett.74.2632](https://doi.org/10.1103/physrevlett.74.2632).
- [7] K. Kodama et al. “Observation of tau neutrino interactions”. In: *Physics Letters B* 504.3 (2001), pp. 218224. ISSN: 0370-2693. DOI: [10.1016/S0370-2693\(01\)00307-0](https://doi.org/10.1016/S0370-2693(01)00307-0). URL: [http://dx.doi.org/10.1016/S0370-2693\(01\)00307-0](http://dx.doi.org/10.1016/S0370-2693(01)00307-0).
- [8] ATLAS Collaboration. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Phys. Rev. Lett.* B716 (2012), pp. 1–29. DOI: [10.1016/j.physletb.2012.08.020](https://doi.org/10.1016/j.physletb.2012.08.020).
- [9] CMS Collaboration. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In: *Phys. Rev. Lett.* B716 (2012), pp. 30–61. DOI: [10.1016/j.physletb.2012.08.021](https://doi.org/10.1016/j.physletb.2012.08.021).
- [10] L. Evans and P. Bryant. “LHC Machine”. In: *JINST* 3 (2008), S08001. DOI: [10.1088/1748-0221/3/08/S08001](https://doi.org/10.1088/1748-0221/3/08/S08001).
- [11] ATLAS Collaboration. “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *JINST* 3 (2008), S08003. DOI: [10.1088/1748-0221/3/08/S08003](https://doi.org/10.1088/1748-0221/3/08/S08003).
- [12] CMS Collaboration. “The CMS Experiment at the CERN LHC”. In: *JINST* 3 (2008), S08004. DOI: [10.1088/1748-0221/3/08/S08004](https://doi.org/10.1088/1748-0221/3/08/S08004).
- [13] P. W. Higgs. “Broken symmetries, massless particles and gauge fields”. In: *Phys. Lett.* 12 (1964), p. 125. DOI: [10.1016/0031-9163\(64\)91136-9](https://doi.org/10.1016/0031-9163(64)91136-9).
- [14] P. W. Higgs. “Spontaneous Symmetry Breakdown without Massless Bosons”. In: *Phys. Rev.* 145 (1966), p. 1156. DOI: [10.1103/PhysRev.145.1156](https://doi.org/10.1103/PhysRev.145.1156).

-
- [15] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. “Global Conservation Laws and Massless Particles”. In: *Phys. Rev. Lett.* 13 (1964), p. 585. DOI: [10.1103/PhysRevLett.13.585](#).
- [16] F. Englert and R. Brout. “Broken Symmetry and the Mass of Gauge Vector Mesons”. In: *Phys. Rev. Lett.* 13 (1964), p. 321. DOI: [10.1103/PhysRevLett.13.321](#).
- [17] P. W. Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. In: *Phys. Rev. Lett.* 13 (1964), p. 508. DOI: [10.1103/PhysRevLett.13.508](#).
- [18] T. W. B. Kibble. “Symmetry breaking in nonAbelian gauge theories”. In: *Phys. Rev.* 155 (1967), p. 1554. DOI: [10.1103/PhysRev.155.1554](#).
- [19] ATLAS Collaboration and CMS Collaboration. “Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s}=7$ and 8 TeV with the ATLAS and CMS Experiments”. In: *Phys. Rev. Lett.* 114.19 (2015). ISSN: 1079-7114. DOI: [10.1103/physrevlett.114.191803](#).
- [20] ATLAS Collaboration and CMS Collaboration. *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*. Tech. rep. Geneva: CERN, 2015. URL: <http://cds.cern.ch/record/2053103>.
- [21] ATLAS Collaboration. “Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector”. In: *Eur. Phys. J. C* 75.10 (2015), p. 476. DOI: [10.1140/epjc/s10052-015-3685-1](#). arXiv: [1506.05669 \[hep-ex\]](#).
- [22] CMS Collaboration. “Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV”. In: *Phys. Rev. D* 92.1 (2015), p. 012004. DOI: [10.1103/PhysRevD.92.012004](#). arXiv: [1411.3441 \[hep-ex\]](#).
- [23] Vera C. Rubin, W. Kent Ford, and Norbert Thonnard. “Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 ($R=4\text{kpc}$) to UGC 2885 ($R=122\text{kpc}$)”. In: *The Astrophysical Journal* 238 (1980), pp. 471–487.
- [24] Jose M. Diego et al. “A free-form lensing grid solution for A1689 with new multiple images”. In: *Monthly Notices of the Royal Astronomical Society* 446.1 (2014), pp. 683–704. ISSN: 0035-8711. DOI: [10.1093/mnras/stu2064](#).
- [25] Planck Collaboration. “Planck 2018 results”. In: *Astronomy Astrophysics* 641 (2020), A6. ISSN: 1432-0746. DOI: [10.1051/0004-6361/201833910](#).
- [26] ATLAS Collaboration. “Cross-section measurements of the Higgs boson decaying into a pair of τ -leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Physical Review D* 99.7 (2019). ISSN: 2470-0029. DOI: [10.1103/physrevd.99.072001](#).
- [27] LHC Higgs Cross Section Working Group. “Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector”. In: (2016). DOI: [10.23731/CYRM-2017-002](#). arXiv: [1610.07922 \[hep-ph\]](#).

REFERENCES

- [28] Georges Aad et al. “Measurements of Higgs boson production cross-sections in the $H \rightarrow \tau^+ \tau^-$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: (2022). arXiv: [2201.08269](https://arxiv.org/abs/2201.08269) [hep-ex].
- [29] ATLAS Collaboration. “Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment”. In: *Phys. Rev. D* 101 (1 2020), p. 012002. DOI: [10.1103/PhysRevD.101.012002](https://doi.org/10.1103/PhysRevD.101.012002).
- [30] ATLAS Collaboration. *Interpretations of the combined measurement of Higgs boson production and decay*. Tech. rep. Geneva: CERN, 2020. URL: <http://cds.cern.ch/record/2743067>.
- [31] ATLAS Collaboration. *Combined measurements of Higgs boson production and decay using up to 139 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment*. Tech. rep. Geneva: CERN, 2021. URL: <http://cds.cern.ch/record/2789544>.
- [32] F. Halzen and A. D. Martin. *Quarks and leptons: an introductory course in modern particle physics*. Wiley Professional, 2010. ISBN: 9780471887416.
- [33] M. E. Peskin and D. V. Schroeder. *An Introduction To Quantum Field Theory (Frontiers in Physics)*. Westview Press, 1995. ISBN: 0201503972.
- [34] D. Griffiths. *Introduction to Elementary Particles*. Wiley Professional, 2008. ISBN: 3527406018.
- [35] D. Ebert. *Eichtheorien Grundlage Der Elementarteilchenphysik*. Wiley Professional, 1989. ISBN: 3527278192.
- [36] H. Fritzsch, M. Gell-Mann, and H. Leutwyler. “Advantages of the color octet gluon picture”. In: *Phys. Rev. Lett.* 47.4 (1973), pp. 365–368. ISSN: 0370-2693. DOI: [https://doi.org/10.1016/0370-2693\(73\)90625-4](https://doi.org/10.1016/0370-2693(73)90625-4).
- [37] D. J. Gross and F. Wilczek. “Ultraviolet Behavior of Nonabelian Gauge Theories”. In: *Phys. Rev. Lett.* 30 (1973), p. 1343. DOI: [10.1103/PhysRevLett.30.1343](https://doi.org/10.1103/PhysRevLett.30.1343).
- [38] H. D. Politzer. “Reliable Perturbative Results for Strong Interactions?” In: *Phys. Rev. Lett.* 30 (1973), p. 1346. DOI: [10.1103/PhysRevLett.30.1346](https://doi.org/10.1103/PhysRevLett.30.1346).
- [39] R. P. Feynman. “Space-Time Approach to Quantum Electrodynamics”. In: *Phys. Rev.* 76 (6 1949), pp. 769–789. DOI: [10.1103/PhysRev.76.769](https://doi.org/10.1103/PhysRev.76.769).
- [40] S. Tomonaga. “On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields*”. In: *Progress of Theoretical Physics* 1.2 (Aug. 1946), pp. 27–42. ISSN: 0033-068X. DOI: [10.1143/PTP.1.27](https://doi.org/10.1143/PTP.1.27).
- [41] J. Schwinger. “CXXIV. A note on the quantum dynamical principle”. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 44.357 (1953), pp. 1171–1179. DOI: [10.1080/14786441008520377](https://doi.org/10.1080/14786441008520377).
- [42] G. V. Chibisov. “Astrophysical upper limits on the photon rest mass”. In: *Physics-Uspekh* 19.7 (1976), pp. 624–626. DOI: [10.1070/pu1976v019n07abeh005277](https://doi.org/10.1070/pu1976v019n07abeh005277).

-
- [43] V. V. Ezhela, S. B. Lugovsky, and O. V. Zenin. “Hadronic part of the muon g-2 estimated on the $\sigma_{total}^{2003}(e+e^- \rightarrow \text{hadrons})$ evaluated data compilation”. In: (2003). arXiv: [hep-ph/0312114 \[hep-ph\]](#).
 - [44] P. Schmüser. *Feynman-Graphen und Eichtheorien für Experimentalphysiker (German Edition)*. Springer, 1994. ISBN: 3540584862.
 - [45] LHCb Collaboration. “Observation of the resonant character of the $Z(4430)^-$ state”. In: *Phys. Rev. Lett.* 112.22 (2014), p. 222002. DOI: [10.1103/PhysRevLett.112.222002](#). arXiv: [1404.1903 \[hep-ex\]](#).
 - [46] LHCb Collaboration. “Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \rightarrow J/\psi K^- p$ Decays”. In: *Phys. Rev. Lett.* 115 (2015), p. 072001. DOI: [10.1103/PhysRevLett.115.072001](#). arXiv: [1507.03414 \[hep-ex\]](#).
 - [47] Particle Data Group. “Review of Particle Physics”. In: *Phys. Rev. D* 98 (3 2018), p. 030001. DOI: [10.1103/PhysRevD.98.030001](#).
 - [48] T. D. Lee and C. Yang. “Question of Parity Conservation in Weak Interactions”. In: *Phys. Rev.* 104 (1956), p. 254. DOI: [10.1103/PhysRev.104.254](#).
 - [49] C. S. Wu et al. “Experimental Test of Parity Conservation in Beta Decay”. In: *Phys. Rev. Lett.* 105 (1957), p. 1413. DOI: [10.1103/PhysRev.105.1413](#).
 - [50] S. L. Glashow. “Partial Symmetries of Weak Interactions”. In: *Nucl. Phys.* 22 (1961), p. 579. DOI: [10.1016/0029-5582\(61\)90469-2](#).
 - [51] S. Weinberg. “A Model of Leptons”. In: *Phys. Rev. Lett.* 19 (1967), p. 1264. DOI: [10.1103/PhysRevLett.19.1264](#).
 - [52] A. Salam and J. C. Ward. “Weak and electromagnetic interactions”. In: *Nuovo Cim.* 11 (1959), p. 568. DOI: [10.1007/BF02726525](#).
 - [53] M. Gell-Mann. “The interpretation of the new particles as displaced charge multiplets”. In: *Nuovo Cim.* 4.S2 (1956), p. 19. DOI: [10.1007/BF02748000](#).
 - [54] K. Nishijima. “Charge Independence Theory of V Particles”. In: *Prog. Theor. Phys.* 13.3 (1955), p. 285. DOI: [10.1143/PTP.13.285](#).
 - [55] Z. Maki, M. Nakagawa, and S. Sakata. “Remarks on the unified model of elementary particles”. In: (1962), p. 663. DOI: [10.1143/PTP.28.870](#).
 - [56] K2K Collaboration. “Measurement of Neutrino Oscillation by the K2K Experiment”. In: *Phys. Rev. D* 74 (2006), p. 072003. DOI: [10.1103/PhysRevD.74.072003](#). arXiv: [hep-ex/0606032 \[hep-ex\]](#).
 - [57] B. Pontecorvo. “Neutrino Experiments and the Problem of Conservation of Leptonic Charge”. In: *Sov. Phys. JETP* 26 (1968), p. 984. URL: <https://www.osti.gov/biblio/4814050>.
 - [58] M. Kobayashi and T. Maskawa. “CP Violation in the Renormalizable Theory of Weak Interaction”. In: *Prog. Theor. Phys.* 49 (1973), p. 652. DOI: [10.1143/PTP.49.652](#).

REFERENCES

- [59] N. Cabibbo. “Unitary Symmetry and Leptonic Decays”. In: *Phys. Rev. Lett.* 10 (1963), p. 531. DOI: [10.1103/PhysRevLett.10.531](https://doi.org/10.1103/PhysRevLett.10.531).
- [60] J. H. Christenson et al. “Evidence for the 2π Decay of the K_2^0 Meson”. In: *Phys. Rev. Lett.* 13 (1964), p. 138. DOI: [10.1103/PhysRevLett.13.138](https://doi.org/10.1103/PhysRevLett.13.138).
- [61] Y. Nambu. “Quasiparticles and Gauge Invariance in the Theory of Superconductivity”. In: *Phys. Rev.* 117 (1960), p. 648. DOI: [10.1103/PhysRev.117.648](https://doi.org/10.1103/PhysRev.117.648).
- [62] J. Goldstone. “Field Theories with Superconductor Solutions”. In: *Nuovo Cim.* 19 (1961), p. 154. DOI: [10.1007/BF02812722](https://doi.org/10.1007/BF02812722).
- [63] J. Goldstone, A. Salam, and S. Weinberg. “Broken Symmetries”. In: *Phys. Rev.* 127 (3 1962), pp. 965–970. DOI: [10.1103/PhysRev.127.965](https://doi.org/10.1103/PhysRev.127.965).
- [64] PLUTO Collaboration. “Jet analysis of the $\gamma(9.46)$ decay into charged hadrons”. In: *Physics Letters B* 82.3 (1979), pp. 449–455. ISSN: 0370-2693. DOI: [10.1016/0370-2693\(79\)90265-X](https://doi.org/10.1016/0370-2693(79)90265-X).
- [65] TASSO Collaboration. “Evidence for planar events in e^+e^- annihilation at high energies”. In: *Physics Letters B* 86.2 (1979), pp. 243–249. ISSN: 0370-2693. DOI: [https://doi.org/10.1016/0370-2693\(79\)90830-X](https://doi.org/10.1016/0370-2693(79)90830-X).
- [66] MARK-J Collaboration. “Discovery of Three-Jet Events and a Test of Quantum Chromodynamics at PETRA”. In: *Phys. Rev. Lett.* 43 (12 1979), pp. 830–833. DOI: [10.1103/PhysRevLett.43.830](https://doi.org/10.1103/PhysRevLett.43.830).
- [67] PLUTO Collaboration. “Evidence for gluon bremsstrahlung in e^+e^- annihilations at high energies”. In: *Physics Letters B* 86.3 (1979), pp. 418–425. ISSN: 0370-2693. DOI: [https://doi.org/10.1016/0370-2693\(79\)90869-4](https://doi.org/10.1016/0370-2693(79)90869-4).
- [68] JADE Collaboration. “Observation of planar three-jet events in e^+e^- annihilation and evidence for gluon bremsstrahlung”. In: *Physics Letters B* 91.1 (1980), pp. 142–147. ISSN: 0370-2693. DOI: [https://doi.org/10.1016/0370-2693\(80\)90680-2](https://doi.org/10.1016/0370-2693(80)90680-2).
- [69] S. W. Herb et al. “Observation of a Dimuon Resonance at 9.5 GeV in 400 GeV Proton-Nucleus Collisions”. In: *Phys. Rev. Lett.* 39 (5 1977), pp. 252–255. DOI: [10.1103/PhysRevLett.39.252](https://doi.org/10.1103/PhysRevLett.39.252).
- [70] ATLAS Collaboration. *Standard Model Summary Plots June 2021*. Tech. rep. Geneva: CERN, 2021. URL: <http://cds.cern.ch/record/2777014>.
- [71] Lance J. Dixon. “Ultraviolet Behavior of $\mathcal{N} = 8$ Supergravity”. In: *Subnucl. Ser.* 47 (2011). Ed. by Antonino Zichichi, pp. 1–39. DOI: [10.1142/9789814374125_0001](https://doi.org/10.1142/9789814374125_0001). arXiv: [1005.2703 \[hep-th\]](https://arxiv.org/abs/1005.2703).
- [72] Adam G. Riess et al. “Observational evidence from supernovae for an accelerating universe and a cosmological constant”. In: *Astron. J.* 116 (1998), pp. 1009–1038. DOI: [10.1086/300499](https://doi.org/10.1086/300499). arXiv: [astro-ph/9805201](https://arxiv.org/abs/astro-ph/9805201).
- [73] S. Perlmutter et al. “Measurements of Ω and Λ from 42 high redshift supernovae”. In: *Astrophys. J.* 517 (1999), pp. 565–586. DOI: [10.1086/307221](https://doi.org/10.1086/307221). arXiv: [astro-ph/9812133](https://arxiv.org/abs/astro-ph/9812133).

-
- [74] Super-Kamiokande Collaboration. “Evidence for oscillation of atmospheric neutrinos”. In: *Phys. Rev. Lett.* 81 (1998), pp. 1562–1567. DOI: [10.1103/PhysRevLett.81.1562](https://doi.org/10.1103/PhysRevLett.81.1562). arXiv: [hep-ex/9807003](https://arxiv.org/abs/hep-ex/9807003).
- [75] SNO Collaboration. “Measurement of the Rate of $\nu_e + d \rightarrow p + p + e$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory”. In: *Physical Review Letters* 87.7 (2001). ISSN: 1079-7114. DOI: [10.1103/physrevlett.87.071301](https://doi.org/10.1103/physrevlett.87.071301).
- [76] SNO Collaboration. “Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory”. In: *Physical Review Letters* 89.1 (2002). ISSN: 1079-7114. DOI: [10.1103/physrevlett.89.011301](https://doi.org/10.1103/physrevlett.89.011301).
- [77] B. Grzadkowski et al. “Dimension-six terms in the Standard Model Lagrangian”. In: *JHEP* 2010.10 (2010). ISSN: 1029-8479. DOI: [10.1007/jhep10\(2010\)085](https://doi.org/10.1007/jhep10(2010)085).
- [78] Gian Francesco Giudice et al. “The strongly-interacting light Higgs”. In: *JHEP* 2007.06 (2007), pp. 045045. ISSN: 1029-8479. DOI: [10.1088/1126-6708/2007/06/045](https://doi.org/10.1088/1126-6708/2007/06/045).
- [79] Roberto Contino et al. “Effective Lagrangian for a light Higgs-like scalar”. In: *JHEP* 2013.7 (2013). ISSN: 1029-8479. DOI: [10.1007/jhep07\(2013\)035](https://doi.org/10.1007/jhep07(2013)035).
- [80] Steven Weinberg. “Baryon- and Lepton-Nonconserving Processes”. In: *Phys. Rev. Lett.* 43 (21 1979), pp. 1566–1570. DOI: [10.1103/PhysRevLett.43.1566](https://doi.org/10.1103/PhysRevLett.43.1566).
- [81] Brian Henning et al. “2, 84, 30, 993, 560, 15456, 11962, 261485, ...: Higher dimension operators in the SM EFT”. In: *JHEP* 08 (2017). [Erratum: *JHEP* 09, 019 (2019)], p. 016. DOI: [10.1007/JHEP08\(2017\)016](https://doi.org/10.1007/JHEP08(2017)016). arXiv: [1512.03433 \[hep-ph\]](https://arxiv.org/abs/1512.03433).
- [82] Andrew Kobach. “Baryon number, lepton number, and operator dimension in the Standard Model”. In: *Physics Letters B* 758 (2016), pp. 455–457. ISSN: 0370-2693. URL: <https://www.sciencedirect.com/science/article/pii/S0370269316301976>.
- [83] A. de Gouvêa, J. Herrero-García, and A. Kobach. “Neutrino masses, grand unification, and baryon number violation”. In: *Phys. Rev. D* 90 (1 2014), p. 016011. DOI: [10.1103/PhysRevD.90.016011](https://doi.org/10.1103/PhysRevD.90.016011).
- [84] Elizabeth E. Jenkins, Aneesh V. Manohar, and Michael Trott. “Renormalization Group Evolution of the Standard Model Dimension Six Operators I: Formalism and lambda Dependence”. In: *JHEP* 10 (2013), p. 087. DOI: [10.1007/JHEP10\(2013\)087](https://doi.org/10.1007/JHEP10(2013)087). arXiv: [1308.2627 \[hep-ph\]](https://arxiv.org/abs/1308.2627).
- [85] Elizabeth E. Jenkins, Aneesh V. Manohar, and Michael Trott. “Renormalization Group Evolution of the Standard Model Dimension Six Operators II: Yukawa Dependence”. In: *JHEP* 01 (2014), p. 035. DOI: [10.1007/JHEP01\(2014\)035](https://doi.org/10.1007/JHEP01(2014)035). arXiv: [1310.4838 \[hep-ph\]](https://arxiv.org/abs/1310.4838).

- [86] Rodrigo Alonso et al. “Renormalization Group Evolution of the Standard Model Dimension Six Operators III: Gauge Coupling Dependence and Phenomenology”. In: *JHEP* 04 (2014), p. 159. DOI: [10.1007/JHEP04\(2014\)159](https://doi.org/10.1007/JHEP04(2014)159). arXiv: [1312.2014](https://arxiv.org/abs/1312.2014) [hep-ph].
- [87] ATLAS Collaboration. “Measurement of the Higgs boson mass in the $H \rightarrow ZZ^* \rightarrow 4l$ and $H \rightarrow \gamma\gamma$ channels with $\sqrt{s}=13$ TeV pp collisions using the ATLAS detector”. In: *Phys. Rev. Lett.* 784 (2018), pp. 345–366. ISSN: 0370-2693. DOI: <https://doi.org/10.1016/j.physletb.2018.07.050>.
- [88] CMS Collaboration. “A measurement of the Higgs boson mass in the diphoton decay channel”. In: *Phys. Rev. Lett.* 805 (2020), p. 135425. ISSN: 0370-2693. DOI: <https://doi.org/10.1016/j.physletb.2020.135425>.
- [89] Particle Data Group and Zyla, P. A. and Barnett, R. M. and Beringer, J. “Review of Particle Physics”. In: *Progress of Theoretical and Experimental Physics* 2020.8 (2020). ISSN: 2050-3911. DOI: [10.1093/ptep/ptaa104](https://doi.org/10.1093/ptep/ptaa104).
- [90] ATLAS Collaboration. “Combined measurement of the total and differential cross sections in the $H \rightarrow \gamma\gamma$ and the $H \rightarrow ZZ^* \rightarrow 4l$ decay channels at $\sqrt{s}=13$ TeV with the ATLAS detector”. In: (2019). URL: <http://cds.cern.ch/record/2682844>.
- [91] G. D’Agostini. “A multidimensional unfolding method based on Bayes’ theorem”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 362.2 (1995), pp. 487–498. ISSN: 0168-9002. DOI: [https://doi.org/10.1016/0168-9002\(95\)00274-X](https://doi.org/10.1016/0168-9002(95)00274-X).
- [92] G. Cowan. “A SURVEY OF UNFOLDING METHODS FOR PARTICLE PHYSICS”. In: *Psychology Journal* (2002). URL: https://www.researchgate.net/publication/255672407_A_SURVEY_OF_UNFOLDING_METHODS_FOR_PARTICLE_PHYSICS.
- [93] “SVD approach to data unfolding”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 372.3 (1996), pp. 469–481. ISSN: 0168-9002. DOI: [https://doi.org/10.1016/0168-9002\(95\)01478-0](https://doi.org/10.1016/0168-9002(95)01478-0).
- [94] M. Hostettler and G. Papotti. “Luminosity Lifetime at the LHC in 2012 Proton Physics Operation”. In: 2013. URL: <http://JACoW.org/IPAC2013/papers/tupfi029.pdf>.
- [95] “Some LHC milestones.... Quelques dates clés du LHC...” In: BUL-NA-2008-132. 38/2008. 38/2008 (2008). Cover article, p. 1. URL: <http://cds.cern.ch/record/1125888>.
- [96] ALICE Collaboration. “The ALICE experiment at the CERN LHC”. In: *Journal of Instrumentation* 3.08 (2008), S08002–S08002. DOI: [10.1088/1748-0221/3/08/s08002](https://doi.org/10.1088/1748-0221/3/08/s08002).

-
- [97] LHCb Collaboration. “The LHCb Detector at the LHC”. In: *Journal of Instrumentation* 3.08 (2008), S08005–S08005. DOI: [10.1088/1748-0221/3/08/s08005](https://doi.org/10.1088/1748-0221/3/08/s08005).
- [98] ATLAS Collaboration. *ATLAS detector and physics performance: Technical Design Report, 1*. Technical design report. ATLAS. Geneva: CERN, 1999. URL: <https://cds.cern.ch/record/391176>.
- [99] ATLAS Collaboration. *ATLAS detector and physics performance: Technical Design Report, 2*. Technical design report. ATLAS. Geneva: CERN, 1999. URL: <https://cds.cern.ch/record/391177>.
- [100] Joao Pequeno. “Computer generated image of the whole ATLAS detector”. In: (2008). URL: <https://cds.cern.ch/record/1095924>.
- [101] ATLAS Collaboration. *ATLAS inner detector: Technical Design Report, 2*. Technical design report. ATLAS. Geneva: CERN, 1997. URL: <https://cds.cern.ch/record/331064>.
- [102] ATLAS Collaboration. *ATLAS Insertable B-Layer Technical Design Report*. Tech. rep. 2010. URL: <https://cds.cern.ch/record/1291633>.
- [103] ATLAS Collaboration. *ATLAS calorimeter performance: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1996. URL: <https://cds.cern.ch/record/331059>.
- [104] ATLAS Collaboration. *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1996. URL: <https://cds.cern.ch/record/331061>.
- [105] ATLAS Collaboration. *ATLAS tile calorimeter: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1996. URL: <https://cds.cern.ch/record/331062>.
- [106] ATLAS Collaboration. *ATLAS muon spectrometer: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1997. URL: <https://cds.cern.ch/record/331068>.
- [107] A. R. Martinez. “The Run-2 ATLAS Trigger System”. In: *J. Phys. Conf. Ser.* 762.1 (2016), p. 012003. DOI: [10.1088/1742-6596/762/1/012003](https://doi.org/10.1088/1742-6596/762/1/012003).
- [108] J. C. Collins, Davison E. Soper, and G. F. Sterman. “Factorization of Hard Processes in QCD”. In: *Adv. Ser. Direct. High Energy Phys.* 5 (1989), pp. 1–91. DOI: [10.1142/9789814503266_0001](https://doi.org/10.1142/9789814503266_0001). arXiv: [hep-ph/0409313](https://arxiv.org/abs/hep-ph/0409313).
- [109] M. Asai. “Geant4 - a simulation toolkit”. In: *Trans. Amer. Nucl. Soc.* 95 (2006), p. 757.
- [110] A. Bhatti and D. Lincoln. “Jet Physics at the Tevatron”. In: *Annual Review of Nuclear and Particle Science* 60.1 (2010), pp. 267297. ISSN: 1545-4134. DOI: [10.1146/annurev.nucl.012809.104430](https://doi.org/10.1146/annurev.nucl.012809.104430).
- [111] V.N. Gribov and L.N. Lipatov. “Deep inelastic electron scattering in perturbation theory”. In: *Phys. Rev. Lett.* 37.1 (1971), pp. 78–80. ISSN: 0370-2693. DOI: [https://doi.org/10.1016/0370-2693\(71\)90576-4](https://doi.org/10.1016/0370-2693(71)90576-4).

REFERENCES

- [112] G. Altarelli and G. Parisi. “Asymptotic freedom in parton language”. In: *Nucl. Phys.* 126.2 (1977), pp. 298–318. ISSN: 0550-3213. DOI: [https://doi.org/10.1016/0550-3213\(77\)90384-4](https://doi.org/10.1016/0550-3213(77)90384-4).
- [113] Y. L. Dokshitzer. “Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics.” In: *Sov. Phys. JETP* 46 (1977), pp. 641–653.
- [114] J. C. Collins. “Sudakov form-factors”. In: *Adv. Ser. Direct. High Energy Phys.* 5 (1989), pp. 573–614. DOI: [10.1142/9789814503266_0006](https://doi.org/10.1142/9789814503266_0006). arXiv: [hep-ph/0312336](https://arxiv.org/abs/hep-ph/0312336).
- [115] R. D. Field and D. Wolfram. “A QCD model for e^+e^- annihilation”. In: *Nucl. Phys.* 213.1 (1983), pp. 65–84. ISSN: 0550-3213. DOI: [https://doi.org/10.1016/0550-3213\(83\)90175-X](https://doi.org/10.1016/0550-3213(83)90175-X).
- [116] J.-C. Winter, F. Krauss, and G. Soff. “A modified cluster-hadronisation model”. In: *Eur. Phys. J. C* 36.3 (2004), pp. 381–395. ISSN: 1434-6052. DOI: [10.1140/epjc/s2004-01960-8](https://doi.org/10.1140/epjc/s2004-01960-8).
- [117] T. Gleisberg et al. “Event generation with SHERPA 1.1”. In: *JHEP* 2009.02 (2009), pp. 007–007. DOI: [10.1088/1126-6708/2009/02/007](https://doi.org/10.1088/1126-6708/2009/02/007).
- [118] B. Andersson et al. “Parton fragmentation and string dynamics”. In: *Phys. Rep.* 97.2 (1983), pp. 31–145. ISSN: 0370-1573. DOI: [https://doi.org/10.1016/0370-1573\(83\)90080-7](https://doi.org/10.1016/0370-1573(83)90080-7).
- [119] T. Sjöstrand et al. “An introduction to PYTHIA 8.2”. In: *Comp. Phy. Com.* 191 (2015), pp. 159–177. ISSN: 0010-4655. DOI: <https://doi.org/10.1016/j.cpc.2015.01.024>.
- [120] ATLAS Collaboration. *ATLAS public twiki*. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2> (visited on 08/27/2017).
- [121] K. Hamilton et al. “NNLOPS simulation of Higgs boson production”. In: *JHEP* 2013.10 (2013). ISSN: 1029-8479. DOI: [10.1007/jhep10\(2013\)222](https://doi.org/10.1007/jhep10(2013)222).
- [122] K. Hamilton, P. Nason, and G. Zanderighi. “Finite quark-mass effects in the NNLOPS POWHEG+MiNLO Higgs generator”. In: *JHEP* 2015 (2015). ISSN: 1029-8479. DOI: [10.1007/jhep05\(2015\)140](https://doi.org/10.1007/jhep05(2015)140).
- [123] S. Alioli et al. “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”. In: *JHEP* 2010.6 (2010). ISSN: 1029-8479. DOI: [10.1007/jhep06\(2010\)043](https://doi.org/10.1007/jhep06(2010)043).
- [124] P. Nason. “A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms”. In: *JHEP* 2004.11 (2004), pp. 040040. ISSN: 1029-8479. DOI: [10.1088/1126-6708/2004/11/040](https://doi.org/10.1088/1126-6708/2004/11/040).
- [125] S. Frixione, P. Nason, and C. Oleari. “Matching NLO QCD computations with parton shower simulations: the POWHEG method”. In: *JHEP* 2007.11 (2007), pp. 070070. ISSN: 1029-8479. DOI: [10.1088/1126-6708/2007/11/070](https://doi.org/10.1088/1126-6708/2007/11/070).

-
- [126] Keith Hamilton, Paolo Nason, and Giulia Zanderighi. “MINLO: Multi-Scale Improved NLO”. In: *JHEP* 10 (2012), p. 155. DOI: [10.1007/JHEP10\(2012\)155](https://doi.org/10.1007/JHEP10(2012)155). arXiv: [1206.3572 \[hep-ph\]](https://arxiv.org/abs/1206.3572).
- [127] John M. Campbell et al. “NLO Higgs Boson Production Plus One and Two Jets Using the POWHEG BOX, MadGraph4 and MCFM”. In: *JHEP* 07 (2012), p. 092. DOI: [10.1007/JHEP07\(2012\)092](https://doi.org/10.1007/JHEP07(2012)092). arXiv: [1202.5475 \[hep-ph\]](https://arxiv.org/abs/1202.5475).
- [128] Keith Hamilton et al. “Merging H/W/Z + 0 and 1 jet at NLO with no merging scale: a path to parton shower + NNLO matching”. In: *JHEP* 05 (2013), p. 082. DOI: [10.1007/JHEP05\(2013\)082](https://doi.org/10.1007/JHEP05(2013)082). arXiv: [1212.4504 \[hep-ph\]](https://arxiv.org/abs/1212.4504).
- [129] Stefano Catani and Massimiliano Grazzini. “An NNLO subtraction formalism in hadron collisions and its application to Higgs boson production at the LHC”. In: *Phys. Rev. Lett.* 98 (2007), p. 222002. DOI: [10.1103/PhysRevLett.98.222002](https://doi.org/10.1103/PhysRevLett.98.222002). arXiv: [hep-ph/0703012 \[hep-ph\]](https://arxiv.org/abs/hep-ph/0703012).
- [130] E. Bagnaschi et al. “Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM”. In: *JHEP* 2012.2 (2012). ISSN: 1029-8479. DOI: [10.1007/jhep02\(2012\)088](https://doi.org/10.1007/jhep02(2012)088).
- [131] T. Sjöstrand, S. Mrenna, and P. Skands. “A brief introduction to PYTHIA 8.1”. In: *Computer Physics Communications* 178.11 (2008), pp. 852867. ISSN: 0010-4655. DOI: [10.1016/j.cpc.2008.01.036](https://doi.org/10.1016/j.cpc.2008.01.036).
- [132] J. Butterworth et al. “PDF4LHC recommendations for LHC Run II”. In: *Journal of Physics G: Nuclear and Particle Physics* 43.2 (2016), p. 023001. ISSN: 1361-6471. DOI: [10.1088/0954-3899/43/2/023001](https://doi.org/10.1088/0954-3899/43/2/023001).
- [133] ATLAS Collaboration. “Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector”. In: *JHEP* 2014.9 (2014). ISSN: 1029-8479. DOI: [10.1007/jhep09\(2014\)145](https://doi.org/10.1007/jhep09(2014)145).
- [134] J. Pumplin et al. “New Generation of Parton Distributions with Uncertainties from Global QCD Analysis”. In: *JHEP* 2002.07 (2002), pp. 012012. ISSN: 1029-8479. DOI: [10.1088/1126-6708/2002/07/012](https://doi.org/10.1088/1126-6708/2002/07/012).
- [135] N. Davidson, T. Przedzinski, and Z. Was. “PHOTOS interface in C++: Technical and Physics Documentation”. In: *Comp. Phys. Com.* 199 (2016), pp. 86–101. DOI: [10.1016/j.cpc.2015.09.013](https://doi.org/10.1016/j.cpc.2015.09.013). arXiv: [1011.0937 \[hep-ph\]](https://arxiv.org/abs/1011.0937).
- [136] 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”. In: *JHEP* 2014.7 (2014). ISSN: 1029-8479. DOI: [10.1007/jhep07\(2014\)079](https://doi.org/10.1007/jhep07(2014)079).
- [137] R. D. Ball et al. “Parton distributions with LHC data”. In: *Nucl. Phys.* 867.2 (2013), pp. 244289. ISSN: 0550-3213. DOI: [10.1016/j.nuclphysb.2012.10.003](https://doi.org/10.1016/j.nuclphysb.2012.10.003).
- [138] ATLAS Collaboration. *ATLAS Pythia 8 tunes to 7 TeV data*. Tech. rep. Geneva: CERN, 2014. URL: <https://cds.cern.ch/record/1966419>.
- [139] R. D. Ball et al. “Parton distributions for the LHC run II”. In: *JHEP* 2015.4 (2015). ISSN: 1029-8479. DOI: [10.1007/jhep04\(2015\)040](https://doi.org/10.1007/jhep04(2015)040).

REFERENCES

- [140] C. Anastasiou et al. “Higgs Boson Gluon-Fusion Production in QCD at Three Loops”. In: *Phys. Rev. Lett.* 114.21 (2015). ISSN: 1079-7114. DOI: [10.1103/physrevlett.114.212001](#).
- [141] C. Anastasiou et al. “High precision determination of the gluon fusion Higgs boson cross-section at the LHC”. In: *JHEP* 05 (2016), p. 058. DOI: [10.1007/JHEP05\(2016\)058](#). arXiv: [1602.00695 \[hep-ph\]](#).
- [142] S. Actis et al. “NLO electroweak corrections to Higgs boson production at hadron colliders”. In: *Phys. Rev. Lett.* 670.1 (2008), pp. 1217. ISSN: 0370-2693. DOI: [10.1016/j.physletb.2008.10.018](#).
- [143] C. Anastasiou, R. Boughezal, and F. Petriello. “Mixed QCD-electroweak corrections to Higgs boson production in gluon fusion”. In: *JHEP* 2009.04 (2009), pp. 003003. ISSN: 1029-8479. DOI: [10.1088/1126-6708/2009/04/003](#).
- [144] M. Ciccolini, A. Denner, and S. Dittmaier. “Strong and electroweak corrections to the production of Higgs + 2jets via weak interactions at the LHC”. In: *Phys. Rev. Lett.* 99 (2007), p. 161803. DOI: [10.1103/PhysRevLett.99.161803](#). arXiv: [0707.0381 \[hep-ph\]](#).
- [145] M. Ciccolini, A. Denner, and S. Dittmaier. “Electroweak and QCD corrections to Higgs production via vector-boson fusion at the CERN LHC”. In: *Physical Review D* 77.1 (2008). ISSN: 1550-2368. DOI: [10.1103/physrevd.77.013002](#).
- [146] P. Bolzoni et al. “Higgs Boson Production via Vector-Boson Fusion at Next-to-Next-to-Leading Order in QCD”. In: *Phys. Rev. Lett.* 105.1 (2010). ISSN: 1079-7114. DOI: [10.1103/physrevlett.105.011801](#).
- [147] O. Brein, A. Djouadi, and R. Harlander. “NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders”. In: *Physics Letters B* 579.1-2 (2004), pp. 149156. ISSN: 0370-2693. DOI: [10.1016/j.physletb.2003.10.112](#).
- [148] L. Altenkamp et al. “Gluon-induced Higgs-strahlung at next-to-leading order QCD”. In: *JHEP* 2013.2 (2013). ISSN: 1029-8479. DOI: [10.1007/jhep02\(2013\)078](#).
- [149] A. Denner et al. “Electroweak corrections to Higgs-strahlung off W/Z bosons at the Tevatron and the LHC with Hawk”. In: *JHEP* 2012.3 (2012). ISSN: 1029-8479. DOI: [10.1007/jhep03\(2012\)075](#).
- [150] W. Beenakker et al. “Higgs Radiation Off Top Quarks at the Tevatron and the LHC”. In: *Phys. Rev. Lett.* 87.20 (2001). ISSN: 1079-7114. DOI: [10.1103/physrevlett.87.201805](#).
- [151] W. Beenakker et al. “NLO QCD corrections to production in hadron collisions”. In: *Nucl. Phys.* 653.1-2 (2003), pp. 151203. ISSN: 0550-3213. DOI: [10.1016/s0550-3213\(03\)00044-0](#).
- [152] S. Dawson et al. “Next-to-leading order QCD corrections to $pp \rightarrow t\bar{t}h$ at the CERN Large Hadron Collider”. In: *Phys. Rev. D* 67 (7 2003), p. 071503. DOI: [10.1103/PhysRevD.67.071503](#).

-
- [153] S. Dawson et al. “Associated Higgs boson production with top quarks at the CERN Large Hadron Collider: NLO QCD corrections”. In: *Phys. Rev. D* 68 (3 2003), p. 034022. DOI: [10.1103/PhysRevD.68.034022](https://doi.org/10.1103/PhysRevD.68.034022).
- [154] Y. Zhang et al. “QCD NLO and EW NLO corrections to ttH production with top quark decays at hadron collider”. In: *Phys. Rev. Lett.* 738 (2014), pp. 15. ISSN: 0370-2693. DOI: [10.1016/j.physletb.2014.09.022](https://doi.org/10.1016/j.physletb.2014.09.022).
- [155] S. Frixione et al. “Electroweak and QCD corrections to top-pair hadroproduction in association with heavy bosons”. In: *JHEP* 06 (2015), p. 184. DOI: [10.1007/JHEP06\(2015\)184](https://doi.org/10.1007/JHEP06(2015)184). arXiv: [1504.03446 \[hep-ph\]](https://arxiv.org/abs/1504.03446).
- [156] Enrico Bothmann et al. “Event Generation with Sherpa 2.2”. In: *SciPost Phys.* 7.3 (2019), p. 034. DOI: [10.21468/SciPostPhys.7.3.034](https://doi.org/10.21468/SciPostPhys.7.3.034). arXiv: [1905.09127 \[hep-ph\]](https://arxiv.org/abs/1905.09127).
- [157] T. Gleisberg and S. Höche. “Comix, a new matrix element generator”. In: *JHEP* 2008.12 (2008), pp. 039039. ISSN: 1029-8479. DOI: [10.1088/1126-6708/2008/12/039](https://doi.org/10.1088/1126-6708/2008/12/039).
- [158] Federico Buccioni et al. “OpenLoops 2”. In: *Eur. Phys. J. C* 79.10 (2019), p. 866. DOI: [10.1140/epjc/s10052-019-7306-2](https://doi.org/10.1140/epjc/s10052-019-7306-2). arXiv: [1907.13071 \[hep-ph\]](https://arxiv.org/abs/1907.13071).
- [159] F. Cascioli, P. Maierhöfer, and S. Pozzorini. “Scattering Amplitudes with Open Loops”. In: *Phys. Rev. Lett.* 108 (11 2012), p. 111601. DOI: [10.1103/PhysRevLett.108.111601](https://doi.org/10.1103/PhysRevLett.108.111601).
- [160] Fabio Cascioli, Philipp Maierhöfer, and Stefano Pozzorini. “Scattering Amplitudes with Open Loops”. In: *Phys. Rev. Lett.* 108 (2012), p. 111601. DOI: [10.1103/PhysRevLett.108.111601](https://doi.org/10.1103/PhysRevLett.108.111601). arXiv: [1111.5206 \[hep-ph\]](https://arxiv.org/abs/1111.5206).
- [161] Ansgar Denner, Stefan Dittmaier, and Lars Hofer. “Collier: A fortran-based complex one-loop library in extended regularizations”. In: *Computer Physics Communications* 212 (2017), pp. 220238. ISSN: 0010-4655. DOI: [10.1016/j.cpc.2016.10.013](https://doi.org/10.1016/j.cpc.2016.10.013).
- [162] S. Schumann and F. Krauss. “A parton shower algorithm based on Catani-Seymour dipole factorisation”. In: *JHEP* 2008.03 (2008), pp. 038038. ISSN: 1029-8479. DOI: [10.1088/1126-6708/2008/03/038](https://doi.org/10.1088/1126-6708/2008/03/038).
- [163] Stefan Höche et al. “QCD matrix elements and truncated showers”. In: *JHEP* 05 (2009), p. 053. DOI: [10.1088/1126-6708/2009/05/053](https://doi.org/10.1088/1126-6708/2009/05/053). arXiv: [0903.1219 \[hep-ph\]](https://arxiv.org/abs/0903.1219).
- [164] NNPDF Collaboration, R.D. Ball et al. “Parton distributions for the LHC Run II”. In: *JHEP* 04 (2015), p. 040. DOI: [10.1007/JHEP04\(2015\)040](https://doi.org/10.1007/JHEP04(2015)040). arXiv: [1410.8849 \[hep-ph\]](https://arxiv.org/abs/1410.8849).
- [165] Stefan Höche et al. “A critical appraisal of NLO+PS matching methods”. In: *JHEP* 09 (2012), p. 049. DOI: [10.1007/JHEP09\(2012\)049](https://doi.org/10.1007/JHEP09(2012)049). arXiv: [1111.1220 \[hep-ph\]](https://arxiv.org/abs/1111.1220).

REFERENCES

- [166] Stefan Höche et al. “QCD matrix elements + parton showers. The NLO case”. In: *JHEP* 04 (2013), p. 027. DOI: [10.1007/JHEP04\(2013\)027](https://doi.org/10.1007/JHEP04(2013)027). arXiv: [1207.5030](https://arxiv.org/abs/1207.5030) [hep-ph].
- [167] Charalampos Anastasiou et al. “High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO”. In: *Phys. Rev. D* 69 (2004), p. 094008. DOI: [10.1103/PhysRevD.69.094008](https://doi.org/10.1103/PhysRevD.69.094008). arXiv: [hep-ph/0312266](https://arxiv.org/abs/hep-ph/0312266).
- [168] F. Cascioli, P. Maierhöfer, and S. Pozzorini. “Scattering Amplitudes with Open Loops”. In: *Phys. Rev. Lett.* 108.11 (2012). ISSN: 1079-7114. DOI: [10.1103/physrevlett.108.111601](https://doi.org/10.1103/physrevlett.108.111601).
- [169] S. Höche et al. “QCD matrix elements + parton showers. The NLO case”. In: *JHEP* 2013.4 (2013). ISSN: 1029-8479. DOI: [10.1007/jhep04\(2013\)027](https://doi.org/10.1007/jhep04(2013)027).
- [170] S. Alioli, S.-O. Moch, and P. Uwer. “Hadronic top-quark pair-production with one jet and parton showering”. In: *JHEP* 2012.1 (2012). ISSN: 1029-8479. DOI: [10.1007/jhep01\(2012\)137](https://doi.org/10.1007/jhep01(2012)137).
- [171] S. Alioli et al. “NLO single-top production matched with shower in POWHEG: s- and t-channel contributions”. In: *JHEP* 09 (2009), p. 111. DOI: [10.1088/1126-6708/2009/09/111](https://doi.org/10.1088/1126-6708/2009/09/111). arXiv: [0907.4076](https://arxiv.org/abs/0907.4076) [hep-ph].
- [172] E. Re. “Single-top Wt-channel production matched with parton showers using the POWHEG method”. In: *Eur. Phys. J. C* 71.2 (2011). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-011-1547-z](https://doi.org/10.1140/epjc/s10052-011-1547-z).
- [173] H.-L. Lai et al. “New parton distributions for collider physics”. In: *Phys. Rev. D* 82 (2010), p. 074024. DOI: [10.1103/PhysRevD.82.074024](https://doi.org/10.1103/PhysRevD.82.074024). arXiv: [1007.2241](https://arxiv.org/abs/1007.2241) [hep-ph].
- [174] P. Artoisenet et al. “Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations”. In: *JHEP* 2013.3 (2013). ISSN: 1029-8479. DOI: [10.1007/jhep03\(2013\)015](https://doi.org/10.1007/jhep03(2013)015).
- [175] T. Sjöstrand, S. Mrenna, and P. Skands. “PYTHIA 6.4 physics and manual”. In: *JHEP* 2006.05 (2006), pp. 026026. ISSN: 1029-8479. DOI: [10.1088/1126-6708/2006/05/026](https://doi.org/10.1088/1126-6708/2006/05/026).
- [176] P. Z. Skands. “Tuning Monte Carlo generators: The Perugia tunes”. In: *Physical Review D* 82.7 (2010). ISSN: 1550-2368. DOI: [10.1103/physrevd.82.074018](https://doi.org/10.1103/physrevd.82.074018). URL: <http://dx.doi.org/10.1103/PhysRevD.82.074018>.
- [177] ATLAS Collaboration. “The ATLAS Simulation Infrastructure”. In: *Eur. Phys. J. C* 70.3 (2010), pp. 823874. ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-010-1429-9](https://doi.org/10.1140/epjc/s10052-010-1429-9).
- [178] A. D. Martin et al. “Parton distributions for the LHC”. In: *Eur. Phys. J. C* 63.2 (2009), pp. 189285. ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-009-1072-5](https://doi.org/10.1140/epjc/s10052-009-1072-5).
- [179] ATLAS Collaboration. *Summary of ATLAS Pythia 8 tunes*. Tech. rep. Geneva: CERN, 2012. URL: <https://cds.cern.ch/record/1474107>.

-
- [180] Philip Ilten. “Tau Decays in Pythia 8”. In: *Nuclear Physics B - Proceedings Supplements* 253-255 (2014). The Twelfth International Workshop on Tau-Lepton Physics (TAU2012), pp. 77–80. ISSN: 0920-5632. DOI: <https://doi.org/10.1016/j.nuclphysbps.2014.09.019>.
- [181] P. Golonka et al. “The tauola-photos-F environment for the TAUOLA and PHOTOS packages, release II”. In: *Computer Physics Communications* 174.10 (2006), pp. 818–835. ISSN: 0010-4655. DOI: <https://doi.org/10.1016/j.cpc.2005.12.018>.
- [182] Janina Anna Krzysiak. *Validation of Tau Lepton Simulation in the Sherpa Generator and a Generator-Level Study of tauBDT Input Variables in the Sherpa and Powheg+Pythia8 Generators*. Tech. rep. ATL-COM-PHYS-2019-032. Geneva: CERN, 2019. URL: <https://cds.cern.ch/record/2654798>.
- [183] S. Frixione, G. Ridolfi, and P. Nason. “A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction”. In: *Journal of High Energy Physics* 2007.09 (2007), pp. 126126. ISSN: 1029-8479. DOI: [10.1088/1126-6708/2007/09/126](https://doi.org/10.1088/1126-6708/2007/09/126).
- [184] H.B. Hartanto et al. “Higgs boson production in association with top quarks in the POWHEG BOX”. In: *Physical Review D* 91.9 (2015). ISSN: 1550-2368. DOI: [10.1103/physrevd.91.094003](https://doi.org/10.1103/physrevd.91.094003).
- [185] D. J. Lange. “The EvtGen particle decay simulation package”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 462.1 (2001), pp. 152–155. ISSN: 0168-9002. DOI: [10.1016/S0168-9002\(01\)00089-4](https://doi.org/10.1016/S0168-9002(01)00089-4).
- [186] ATLAS Collaboration. “Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2”. In: *Eur. Phys. J. C* 77.10 (2017). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-017-5225-7](https://doi.org/10.1140/epjc/s10052-017-5225-7).
- [187] T. Cornelissen et al. “The new ATLAS track reconstruction (NEWT)”. In: *Journal of Physics: Conference Series* 119.3 (2008), p. 032014. DOI: [10.1088/1742-6596/119/3/032014](https://doi.org/10.1088/1742-6596/119/3/032014).
- [188] R. Frühwirth. “Application of Kalman filtering to track and vertex fitting”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 262.2 (1987), pp. 444–450. ISSN: 0168-9002. DOI: [10.1016/0168-9002\(87\)90887-4](https://doi.org/10.1016/0168-9002(87)90887-4).
- [189] D. Wicke. “A New Algorithm for Solving Tracking Ambiguities”. In: (1998). URL: <https://cds.cern.ch/record/2625731>.
- [190] ATLAS Collaboration. *Vertex Reconstruction Performance of the ATLAS Detector at $\sqrt{s} = 13$ TeV*. Tech. rep. Geneva: CERN, 2015. URL: <https://cds.cern.ch/record/2037717>.

REFERENCES

- [191] G. Piacquadio, K. Prokofiev, and A. Wildauer. “Primary vertex reconstruction in the ATLAS experiment at LHC”. In: *Journal of Physics: Conference Series* 119.3 (2008), p. 032033. DOI: [10.1088/1742-6596/119/3/032033](https://doi.org/10.1088/1742-6596/119/3/032033).
- [192] W. Lampl et al. *Calorimeter Clustering Algorithms: Description and Performance*. Tech. rep. Geneva: CERN, 2008. URL: <https://cds.cern.ch/record/1099735>.
- [193] ATLAS Collaboration. “Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC protonproton collision data at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J. C* 79.8 (2019). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-019-7140-6](https://doi.org/10.1140/epjc/s10052-019-7140-6).
- [194] T. G. Cornelissen et al. “The global χ^2 track fitter in ATLAS”. In: *Journal of Physics: Conference Series* 119.3 (2008), p. 032013. DOI: [10.1088/1742-6596/119/3/032013](https://doi.org/10.1088/1742-6596/119/3/032013).
- [195] ATLAS Collaboration. *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*. Tech. rep. Geneva: CERN, 2012. URL: <https://cds.cern.ch/record/1449796>.
- [196] ATLAS Collaboration. “Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data”. In: *Eur. Phys. J. C* 74.10 (2014). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-014-3071-4](https://doi.org/10.1140/epjc/s10052-014-3071-4).
- [197] ATLAS Collaboration. “Electron and photon energy calibration with the ATLAS detector using 2015-2016 LHC proton-proton collision data”. In: *Journal of Instrumentation* 14.03 (2019), P03017–P03017. DOI: [10.1088/1748-0221/14/03/p03017](https://doi.org/10.1088/1748-0221/14/03/p03017).
- [198] ATLAS Collaboration. “Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1”. In: *Eur. Phys. J. C* C77 (2017), p. 02934. DOI: [10.1140/epjc/s10052-017-5004-5](https://doi.org/10.1140/epjc/s10052-017-5004-5). arXiv: [1603.02934 \[hep-ex\]](https://arxiv.org/abs/1603.02934).
- [199] M. Cacciari and G. P. Salam. “Pileup subtraction using jet areas”. In: *Phys. Rev. Lett.* 659.1-2 (2008), pp. 119126. ISSN: 0370-2693. DOI: [10.1016/j.physletb.2007.09.077](https://doi.org/10.1016/j.physletb.2007.09.077).
- [200] ATLAS Collaboration. “Muon reconstruction performance of the ATLAS detector in protonproton collision data at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J. C* 76.5 (2016). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-016-4120-y](https://doi.org/10.1140/epjc/s10052-016-4120-y).
- [201] 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*. Tech. rep. Geneva: CERN, 2015. URL: <https://cds.cern.ch/record/2044564>.
- [202] ATLAS Collaboration. “Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13$ TeV”. In: *ATLAS-CONF-2017-029* (2017). URL: <https://cds.cern.ch/record/2261772>.

- [203] ATLAS Collaboration. “Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector”. In: *Eur. Phys. J. C* 76.11 (2016), p. 581. DOI: [10.1140/epjc/s10052-016-4395-z](https://doi.org/10.1140/epjc/s10052-016-4395-z). arXiv: [1510.03823](https://arxiv.org/abs/1510.03823) [hep-ex].
- [204] ATLAS Collaboration. “Jet reconstruction and performance using particle flow with the ATLAS Detector”. In: *Eur. Phys. J. C* 77.7 (2017). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-017-5031-2](https://doi.org/10.1140/epjc/s10052-017-5031-2).
- [205] M. Cacciari, G. P. Salam, and Soyez G. “The anti-ktjet clustering algorithm”. In: *JHEP* 2008.04 (2008), pp. 063–063. DOI: [10.1088/1126-6708/2008/04/063](https://doi.org/10.1088/1126-6708/2008/04/063).
- [206] M. Cacciari, G. P. Salam, and G. Soyez. “FastJet user manual”. In: *Eur. Phys. J. C* 72.3 (2012). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-012-1896-2](https://doi.org/10.1140/epjc/s10052-012-1896-2).
- [207] ATLAS Collaboration. “Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Physical Review D* 96.7 (2017). ISSN: 2470-0029. DOI: [10.1103/physrevd.96.072002](https://doi.org/10.1103/physrevd.96.072002).
- [208] ATLAS Collaboration. “Performance of b-jet identification in the ATLAS experiment”. In: *Journal of Instrumentation* 11.04 (2016), P04008P04008. ISSN: 1748-0221. DOI: [10.1088/1748-0221/11/04/p04008](https://doi.org/10.1088/1748-0221/11/04/p04008).
- [209] ATLAS Collaboration. *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*. Tech. rep. Geneva: CERN, 2016. URL: <https://cds.cern.ch/record/2160731>.
- [210] ATLAS Collaboration. “ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J. C* 79.11 (2019). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-019-7450-8](https://doi.org/10.1140/epjc/s10052-019-7450-8).
- [211] ATLAS Collaboration. *Identification of Jets Containing b-Hadrons with Recurrent Neural Networks at the ATLAS Experiment*. Tech. rep. Geneva: CERN, 2017. URL: <https://cds.cern.ch/record/2255226>.
- [212] ATLAS Collaboration. *Secondary vertex finding for jet flavour identification with the ATLAS detector*. Tech. rep. Geneva: CERN, 2017. URL: <https://cds.cern.ch/record/2270366>.
- [213] G. Piacquadio and C. Weiser. “A new inclusive secondary vertex algorithm for b-jet tagging in ATLAS”. In: *Journal of Physics: Conference Series* 119.3 (2008), p. 032032. DOI: [10.1088/1742-6596/119/3/032032](https://doi.org/10.1088/1742-6596/119/3/032032).
- [214] ATLAS Collaboration. *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*. Tech. rep. Geneva: CERN, 2015. URL: <https://cds.cern.ch/record/2064383>.
- [215] ATLAS Collaboration. *Local Hadronic Calibration*. Tech. rep. Geneva: CERN, 2008. URL: <https://cds.cern.ch/record/1112035>.

- [216] ATLAS Collaboration. “Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at $\sqrt{s} = 8$ TeV”. In: *Eur. Phys. J. C* 75.7 (2015). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-015-3500-z](https://doi.org/10.1140/epjc/s10052-015-3500-z).
- [217] ATLAS Collaboration. “Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data”. In: (2016). URL: <https://cds.cern.ch/record/2157687>.
- [218] ATLAS Collaboration. “Electron efficiency measurements with the ATLAS detector using 2012 LHC protonproton collision data”. In: *Eur. Phys. J. C* 77.3 (2017). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-017-4756-2](https://doi.org/10.1140/epjc/s10052-017-4756-2).
- [219] ATLAS Collaboration. “Performance of missing transverse momentum reconstruction with the ATLAS detector using protonproton collisions at $\sqrt{s}=13$ TeV”. In: *Eur. Phys. J. C* 78.11 (2018). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-018-6288-9](https://doi.org/10.1140/epjc/s10052-018-6288-9).
- [220] ATLAS Collaboration. E_T^{miss} performance in the ATLAS detector using 2015-2016 LHC p-p collisions. Tech. rep. Geneva: CERN, 2018. URL: <https://cds.cern.ch/record/2625233>.
- [221] R.K. Ellis et al. “Higgs decay to $\tau^+\tau^-$ A possible signature of intermediate mass Higgs bosons at high energy hadron colliders”. In: *Nucl. Phys.* 297.2 (1988), pp. 221–243. ISSN: 0550-3213. DOI: [10.1016/0550-3213\(88\)90019-3](https://doi.org/10.1016/0550-3213(88)90019-3).
- [222] A. Loesle. “Measurement of the Higgs-Boson production cross-section and test of CP invariance in Vector-Boson fusion production with $H \rightarrow \tau\tau$ decays in the fully-leptonic final state with the ATLAS detector at the LHC”. In: (2020). DOI: <https://doi.org/10.6094/UNIFR/167201>.
- [223] A. Elagin et al. “A new mass reconstruction technique for resonances decaying to di-tau”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 654.1 (2011), pp. 481489. ISSN: 0168-9002. DOI: [10.1016/j.nima.2011.07.009](https://doi.org/10.1016/j.nima.2011.07.009).
- [224] D. Sammel. “Measurement of Higgs-boson production cross-sections and test of CP invariance in vector-boson fusion production in the $H \rightarrow \tau \text{ lep } \tau \text{ had}$ decay mode with the ATLAS detector”. In: (2020). DOI: [10.6094/UNIFR/194328](https://doi.org/10.6094/UNIFR/194328).
- [225] G. Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: *Eur. Phys. J. C* 71.2 (2011). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-011-1554-0](https://doi.org/10.1140/epjc/s10052-011-1554-0).
- [226] K. Cranmer et al. *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*. Tech. rep. New York: New York U., 2012. URL: <https://cds.cern.ch/record/1456844>.
- [227] R. Barlow and C. Beeston. “Fitting using finite Monte Carlo samples”. In: *Computer Physics Communications* 77.2 (1993), pp. 219–228. ISSN: 0010-4655. DOI: [10.1016/0010-4655\(93\)90005-W](https://doi.org/10.1016/0010-4655(93)90005-W).

-
- [228] A. L. Read. “Presentation of search results: the CLs technique”. In: 28.10 (2002), pp. 2693–2704. DOI: [10.1088/0954-3899/28/10/313](https://doi.org/10.1088/0954-3899/28/10/313).
- [229] ATLAS Collaboration. *The ATLAS Tau Trigger in Run 2*. Tech. rep. Geneva: CERN, 2017. URL: <https://cds.cern.ch/record/2273236>.
- [230] R. Nisius. “On the combination of correlated estimates of a physics observable”. In: *Eur. Phys. J. C* 74.8 (2014), p. 3004. DOI: [10.1140/epjc/s10052-014-3004-2](https://doi.org/10.1140/epjc/s10052-014-3004-2).
- [231] R. Nisius. “BLUE: Combining correlated estimates of physics observables within ROOT using the Best Linear Unbiased Estimate method”. In: *SoftwareX* 11 (2020), p. 100468. ISSN: 2352-7110. DOI: <https://doi.org/10.1016/j.softx.2020.100468>.
- [232] Schleicher, K. *private communication*. 2021.
- [233] ATLAS Collaboration. “Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector”. In: *JHEP* 04 (2015), p. 117. DOI: [10.1007/JHEP04\(2015\)117](https://doi.org/10.1007/JHEP04(2015)117). arXiv: [1501.04943 \[hep-ex\]](https://arxiv.org/abs/1501.04943).
- [234] ATLAS Collaboration. “Performance of electron and photon triggers in ATLAS during LHC Run 2”. In: *Eur. Phys. J. C* 80.1 (2020), p. 47. DOI: [10.1140/epjc/s10052-019-7500-2](https://doi.org/10.1140/epjc/s10052-019-7500-2). arXiv: [1909.00761 \[hep-ex\]](https://arxiv.org/abs/1909.00761).
- [235] ATLAS Collaboration. “Performance of the ATLAS muon triggers in Run 2”. In: *JINST* 15.09 (2020), P09015. DOI: [10.1088/1748-0221/15/09/p09015](https://doi.org/10.1088/1748-0221/15/09/p09015). arXiv: [2004.13447 \[physics.ins-det\]](https://arxiv.org/abs/2004.13447).
- [236] ATLAS Collaboration. *The ATLAS Tau Trigger in Run 2*. Tech. rep. Geneva: CERN, 2017. URL: <http://cds.cern.ch/record/2274201>.
- [237] J. Bellm et al. “Herwig 7.0/Herwig++ 3.0 release note”. In: *Eur. Phys. J. C* 76.4 (2016), p. 196. DOI: [10.1140/epjc/s10052-016-4018-8](https://doi.org/10.1140/epjc/s10052-016-4018-8). arXiv: [1512.01178 \[hep-ph\]](https://arxiv.org/abs/1512.01178).
- [238] ATLAS Collaboration. *Analysis of $t\bar{t}H$ and $t\bar{t}W$ production in multilepton final states with the ATLAS detector*. Tech. rep. Geneva: CERN, 2019. URL: <https://cds.cern.ch/record/2693930>.
- [239] J. Therhaag. “TMVA: Toolkit for multivariate data analysis”. In: *AIP Conf. Proc.* 1504.1 (2012). Ed. by T. E. Simons and G. Maroulis, pp. 1013–1016. DOI: [10.1063/1.4771869](https://doi.org/10.1063/1.4771869).
- [240] The ATLAS collaboration. “Modelling $Z \rightarrow \tau\tau$ processes in ATLAS with τ -embedded $Z \rightarrow \mu\mu$ data”. In: *Journal of Instrumentation* 10.09 (2015), P09018P09018. ISSN: 1748-0221. DOI: [10.1088/1748-0221/10/09/p09018](https://doi.org/10.1088/1748-0221/10/09/p09018). URL: <http://dx.doi.org/10.1088/1748-0221/10/09/P09018>.
- [241] ATLAS Collaboration. *Estimation of non-prompt and fake lepton backgrounds in final states with top quarks produced in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*. Tech. rep. Geneva: CERN, 2014. URL: <https://cds.cern.ch/record/1951336>.

REFERENCES

- [242] *Les Houches 2017: Physics at TeV Colliders Standard Model Working Group Report*. 2018. arXiv: [1803.07977 \[hep-ph\]](#).
- [243] I. W. Stewart and F. J. Tackmann. “Theory uncertainties for Higgs and other searches using jet bins”. In: *Phys. Rev. D* 85 (2012), p. 034011. DOI: [10.1103/PhysRevD.85.034011](#). arXiv: [1107.2117 \[hep-ph\]](#).
- [244] R. Frederix and S. Frixione. “Merging meets matching in MC@NLO”. In: *JHEP* 12 (2012), p. 061. DOI: [10.1007/JHEP12\(2012\)061](#). arXiv: [1209.6215 \[hep-ph\]](#).
- [245] S. Gangal and F. J. Tackmann. “Next-to-leading-order uncertainties in Higgs+2 jets from gluon fusion”. In: *Phys. Rev. D* 87 (2013), p. 093008. DOI: [10.1103/PhysRevD.87.093008](#). arXiv: [1302.5437 \[hep-ph\]](#).
- [246] J. K. Anders and M. D’Onofrio. *V+Jets theoretical uncertainties estimation via a parameterisation method*. Tech. rep. Geneva: CERN, 2016. URL: <https://cds.cern.ch/record/2125718>.
- [247] E. Bothmann, M. Schönherr, and S. Schumann. “Reweighting QCD matrix-element and parton-shower calculations”. In: *Eur. Phys. J. C* 76.11 (2016), p. 590. DOI: [10.1140/epjc/s10052-016-4430-0](#). arXiv: [1606.08753 \[hep-ph\]](#).
- [248] S. Dulat et al. “New parton distribution functions from a global analysis of quantum chromodynamics”. In: *Phys. Rev. D* 93.3 (2016), p. 033006. DOI: [10.1103/PhysRevD.93.033006](#). arXiv: [1506.07443 \[hep-ph\]](#).
- [249] L.A. Harland-Lang et al. “Parton distributions in the LHC era: MMHT 2014 PDFs”. In: *Eur. Phys. J. C* 75.5 (2015), p. 204. DOI: [10.1140/epjc/s10052-015-3397-6](#). arXiv: [1412.3989 \[hep-ph\]](#).
- [250] M. Bähr and et. all. “Herwig++ physics and manual”. In: *Eur. Phys. J. C* 58 (2008), p. 639. DOI: [10.1140/epjc/s10052-008-0798-9](#). arXiv: [0803.0883 \[hep-ph\]](#).
- [251] ATLAS Collaboration. *Studies on top-quark Monte Carlo modelling with Sherpa and MG5_aMC@NLO*. ATL-PHYS-PUB-2017-007. 2017. URL: <https://cds.cern.ch/record/2261938>.
- [252] ATLAS Collaboration. “Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data”. In: *JINST* 14.12 (2019), P12006. DOI: [10.1088/1748-0221/14/12/P12006](#). arXiv: [1908.00005 \[hep-ex\]](#).
- [253] ATLAS Collaboration. “Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Eur. Phys. J. C* 81 (2020), p. 689. DOI: [10.1140/epjc/s10052-021-09402-3](#). arXiv: [2007.02645 \[hep-ex\]](#).
- [254] ATLAS Collaboration. *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*. Tech. rep. Geneva: CERN, 2019. URL: <https://cds.cern.ch/record/2677054>.

-
- [255] G. Avoni and et. all. “The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS”. In: *JINST* 13.07 (2018), P07017. DOI: [10.1088/1748-0221/13/07/P07017](https://doi.org/10.1088/1748-0221/13/07/P07017).
- [256] *RooStats*. 2007. URL: <https://twiki.cern.ch/twiki/bin/view/RooStats/WebHome> (visited on 09/29/2020).
- [257] *Proceedings of the 1974 CERN school of computing*. 1974. URL: <http://cds.cern.ch/record/186223/files/CERN-74-23.pdf> (visited on 01/04/2021).
- [258] *TRexFitter*. 2020. URL: <https://trexfitter-docs.web.cern.ch/trexfitter-docs/> (visited on 10/19/2020).
- [259] ATLAS Collaboration. “Measurements of Higgs boson properties in the diphoton decay channel with 36 fb^{-1} of pp collision data at $\sqrt{s} = 13\text{ TeV}$ with the ATLAS detector”. In: *Phys. Rev. D* 98 (2018), p. 052005. DOI: [10.1103/PhysRevD.98.052005](https://doi.org/10.1103/PhysRevD.98.052005). arXiv: [1802.04146](https://arxiv.org/abs/1802.04146) [hep-ex].
- [260] ATLAS Collaboration. *Measurement of the properties of Higgs boson production at $\sqrt{s} = 13\text{ TeV}$ in the $H \rightarrow \gamma\gamma$ channel using 139 fb^{-1} of pp collision data with the ATLAS experiment*. Tech. rep. Geneva: CERN, 2020. URL: <https://cds.cern.ch/record/2725727>.
- [261] ATLAS Collaboration. “Measurement of the Higgs boson coupling properties in the $H \rightarrow ZZ^* \rightarrow 4\ell$ decay channel at $\sqrt{s} = 13\text{ TeV}$ with the ATLAS detector”. In: *JHEP* 03 (2018), p. 095. DOI: [10.1007/JHEP03\(2018\)095](https://doi.org/10.1007/JHEP03(2018)095). arXiv: [1712.02304](https://arxiv.org/abs/1712.02304) [hep-ex].
- [262] ATLAS Collaboration. “Higgs boson production cross-section measurements and their EFT interpretation in the 4ℓ decay channel at $\sqrt{s} = 13\text{ TeV}$ with the ATLAS detector”. In: *Eur. Phys. J. C* 80.10 (2020), p. 957. DOI: [10.1140/epjc/s10052-020-8227-9](https://doi.org/10.1140/epjc/s10052-020-8227-9). arXiv: [2004.03447](https://arxiv.org/abs/2004.03447) [hep-ex].
- [263] ATLAS Collaboration. “Observation of $H \rightarrow b\bar{b}$ decays and VH production with the ATLAS detector”. In: *Phys. Lett. B* 786 (2018), pp. 59–86. DOI: [10.1016/j.physletb.2018.09.013](https://doi.org/10.1016/j.physletb.2018.09.013). arXiv: [1808.08238](https://arxiv.org/abs/1808.08238) [hep-ex].
- [264] ATLAS Collaboration. “Measurement of VH , $H \rightarrow b\bar{b}$ production as a function of the vector-boson transverse momentum in 13 TeV pp collisions with the ATLAS detector”. In: *JHEP* 05 (2019), p. 141. DOI: [10.1007/JHEP05\(2019\)141](https://doi.org/10.1007/JHEP05(2019)141). arXiv: [1903.04618](https://arxiv.org/abs/1903.04618) [hep-ex].
- [265] ATLAS Collaboration. “Measurements of WH and ZH production in the $H \rightarrow b\bar{b}$ decay channel in pp collisions at 13 TeV with the ATLAS detector”. In: *Eur. Phys. J. C* 81.2 (2021), p. 178. DOI: [10.1140/epjc/s10052-020-08677-2](https://doi.org/10.1140/epjc/s10052-020-08677-2). arXiv: [2007.02873](https://arxiv.org/abs/2007.02873) [hep-ex].

- [266] ATLAS Collaboration. “Measurement of the associated production of a Higgs boson decaying into b-quarks with a vector boson at high transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Physics Letters B* 816 (2021), p. 136204. ISSN: 0370-2693. DOI: <https://doi.org/10.1016/j.physletb.2021.136204>. URL: <https://www.sciencedirect.com/science/article/pii/S0370269321001441>.
- [267] ATLAS Collaboration. *Combination of measurements of Higgs boson production in association with a W or Z boson in the $b\bar{b}$ decay channel with the ATLAS experiment at $\sqrt{s} = 13$ TeV*. Tech. rep. Geneva: CERN, 2021. URL: <https://cds.cern.ch/record/2782535>.
- [268] ATLAS Collaboration. “Search for the Dimuon Decay of the Higgs Boson in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector”. In: *Phys. Rev. Lett.* 119 (5 2017), p. 051802. DOI: [10.1103/PhysRevLett.119.051802](https://doi.org/10.1103/PhysRevLett.119.051802).
- [269] ATLAS Collaboration. “A search for the dimuon decay of the Standard Model Higgs boson with the ATLAS detector”. In: *Phys. Lett. B* 812 (2021), p. 135980. DOI: [10.1016/j.physletb.2020.135980](https://doi.org/10.1016/j.physletb.2020.135980). arXiv: [2007.07830 \[hep-ex\]](https://arxiv.org/abs/2007.07830).
- [270] ATLAS Collaboration. “Measurements of gluon-gluon fusion and vector-boson fusion Higgs boson production cross-sections in the $H \rightarrow WW^* \rightarrow e\nu\mu\nu$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Phys. Lett. B* 789 (2019), pp. 508–529. DOI: [10.1016/j.physletb.2018.11.064](https://doi.org/10.1016/j.physletb.2018.11.064). arXiv: [1808.09054 \[hep-ex\]](https://arxiv.org/abs/1808.09054).
- [271] ATLAS Collaboration. *Measurements of gluon fusion and vector-boson-fusion production of the Higgs boson in $H \rightarrow WW^* \rightarrow e\nu\mu\nu$ decays using pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Tech. rep. Geneva: CERN, 2021. URL: <https://cds.cern.ch/record/2759651>.
- [272] ATLAS Collaboration. *Measurements of Higgs boson production cross-sections in the $H \rightarrow \tau^+\tau^-$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Tech. rep. Geneva: CERN, 2021. URL: <https://cds.cern.ch/record/2779179>.
- [273] ATLAS Collaboration. “Search for Higgs bosons produced via vector-boson fusion and decaying into bottom quark pairs in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector”. In: *Phys. Rev. D* 98.5 (2018), p. 052003. DOI: [10.1103/PhysRevD.98.052003](https://doi.org/10.1103/PhysRevD.98.052003). arXiv: [1807.08639 \[hep-ex\]](https://arxiv.org/abs/1807.08639).
- [274] ATLAS Collaboration. “Measurements of Higgs bosons decaying to bottom quarks from vector boson fusion production with the ATLAS experiment at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J. C* 81.6 (2021), p. 537. DOI: [10.1140/epjc/s10052-021-09192-8](https://doi.org/10.1140/epjc/s10052-021-09192-8). arXiv: [2011.08280 \[hep-ex\]](https://arxiv.org/abs/2011.08280).
- [275] ATLAS Collaboration. “Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector”. In: *Phys. Rev. D* 97.7 (2018), p. 072003. DOI: [10.1103/PhysRevD.97.072003](https://doi.org/10.1103/PhysRevD.97.072003). arXiv: [1712.08891 \[hep-ex\]](https://arxiv.org/abs/1712.08891).

-
- [276] 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”. In: *Phys. Rev. D* 97 (7 2018), p. 072016. DOI: [10.1103/PhysRevD.97.072016](https://doi.org/10.1103/PhysRevD.97.072016).
- [277] ATLAS Collaboration. “Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector”. In: *Phys. Rev. D* 97 (7 2018), p. 072003. DOI: [10.1103/PhysRevD.97.072003](https://doi.org/10.1103/PhysRevD.97.072003).
- [278] ATLAS Collaboration. “Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector”. In: *Physics Letters B* 784 (2018), pp. 173–191. ISSN: 0370-2693. DOI: <https://doi.org/10.1016/j.physletb.2018.07.035>.
- [279] ATLAS Collaboration. *Measurement of the Higgs boson decaying to b-quarks produced in association with a top-quark pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Tech. rep. Geneva: CERN, 2020. URL: <https://cds.cern.ch/record/2743685>.
- [280] ATLAS Collaboration. “Search for invisible Higgs boson decays in vector boson fusion at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Physics Letters B* 793 (2019), pp. 499–519. ISSN: 0370-2693. DOI: <https://doi.org/10.1016/j.physletb.2019.04.024>.
- [281] ATLAS Collaboration. “Search for an invisibly decaying Higgs boson or dark matter candidates produced in association with a Z boson in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Physics Letters B* 776 (2018), pp. 318–337. ISSN: 0370-2693. DOI: <https://doi.org/10.1016/j.physletb.2017.11.049>.
- [282] ATLAS Collaboration. “Search for dark matter in events with a hadronically decaying vector boson and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *JHEP* 10 (2018), p. 180. DOI: [10.1007/JHEP10\(2018\)180](https://doi.org/10.1007/JHEP10(2018)180). arXiv: [1807.11471](https://arxiv.org/abs/1807.11471) [hep-ex].
- [283] ATLAS Collaboration. “Combination of Searches for Invisible Higgs Boson Decays with the ATLAS Experiment”. In: *Phys. Rev. Lett.* 122 (23 2019), p. 231801. DOI: [10.1103/PhysRevLett.122.231801](https://doi.org/10.1103/PhysRevLett.122.231801).
- [284] ATLAS Collaboration. *Search for invisible Higgs boson decays with vector boson fusion signatures with the ATLAS detector using an integrated luminosity of 139 fb^{-1}* . Tech. rep. Geneva: CERN, 2020. URL: <https://cds.cern.ch/record/2715447>.
- [285] ATLAS Collaboration. “Constraints on off-shell Higgs boson production and the Higgs boson total width in $ZZ \rightarrow 4l$ and $ZZ \rightarrow 2l2\nu$ final states with the ATLAS detector”. In: *Physics Letters B* 786 (2018), pp. 223–244. ISSN: 0370-2693. DOI: <https://doi.org/10.1016/j.physletb.2018.09.048>.
- [286] ATLAS Collaboration. *Methodology for EFT interpretation of Higgs boson Simplified Template Cross-section results in ATLAS*. Tech. rep. Geneva: CERN, 2019. URL: <http://cds.cern.ch/record/2694284>.

REFERENCES

- [287] I. Brivio and M. Trott. “The Standard Model as an Effective Field Theory”. In: *Phys. Rept.* 793 (2019), pp. 1–98. DOI: [10.1016/j.physrep.2018.11.002](https://doi.org/10.1016/j.physrep.2018.11.002). arXiv: [1706.08945](https://arxiv.org/abs/1706.08945) [hep-ph].
- [288] I. Brivio, Y. Jiang, and M. Trott. “The SMEFTsim package, theory and tools”. In: *JHEP* 12 (2017), p. 070. DOI: [10.1007/JHEP12\(2017\)070](https://doi.org/10.1007/JHEP12(2017)070). arXiv: [1709.06492](https://arxiv.org/abs/1709.06492) [hep-ph].
- [289] I. Brivio. “SMEFTsim 3.0 — a practical guide”. In: *JHEP* 04 (2021), p. 073. DOI: [10.1007/JHEP04\(2021\)073](https://doi.org/10.1007/JHEP04(2021)073). arXiv: [2012.11343](https://arxiv.org/abs/2012.11343) [hep-ph].
- [290] C. Degrande et al. *Standard Model Effective Theory at Next-to-Leading-Order in QCD*. URL: <http://feynrules.irmp.ucl.ac.be/wiki/SMEFTatNLO>.
- [291] S. Dawson and P. P. Giardino. “Electroweak corrections to Higgs boson decays to $\gamma\gamma$ and W^+W^- in standard model EFT”. In: *Phys. Rev. D* 98 (9 2018), p. 095005. DOI: [10.1103/PhysRevD.98.095005](https://doi.org/10.1103/PhysRevD.98.095005).

ACKNOWLEDGEMENTS

Finally, I want to thank everybody, who supported me during the last years. Many of them contributed in different ways to this thesis either during my daily work or within the ATLAS collaboration. The success of this thesis would not have been possible without their support. In addition to the people mentioned here, I am also thankful for many other people, who supported me throughout the years, such as my family, my partner and my friends. However, I want to keep this part directly work-related.

Zunächst möchte ich mich bei dir, Markus, bedanken. Bereits während meiner Bachelorarbeit hast du mir die Möglichkeit gegeben, einen ersten Einblick in die Forschung und die ATLAS Kollaboration zu erlangen. Seitdem hast du mir immer unterstützend zur Seite gestanden, um mir bei Problemen und Herausforderungen weiter zu helfen, vor allem in Bezug auf meine Master- und Doktorarbeit, aber auch bei weiterführenden Fragestellungen.

Liebe Chris, ein besonderer Dank geht an dich für die Unterstützung bei kleineren, aber auch großen organisatorischen Problemen während der gesamten Zeit. Ich werde mich immer dankbar an deinen Einsatz und deine Hilfe bei der bürokratischen Arbeit erinnern.

Ich bin dankbar für die Hilfe meiner direkten Betreuer, Elias, Kathrin, David und Valerie, welche für mich bei physikalischen, aber auch organisatorischen und beruflichen Fragen immer ein offenes Ohr hatten. Durch euch lernte ich, eigenständig zu arbeiten, aber auch, wie wichtig es ist, einen Rückhalt während der schwierigeren Phasen der Promotion zu haben.

Ein großer Dank geht an das ganze Analyse-Team, welches mich während meiner Promotion begleitet hat. Ich bin sehr dankbar für die beständige Unterstützung und Hilfe, auch zu arbeit-suntypischen Zeiten und trotz teils großer Entfernungen. Hier möchte ich besonders meine Betreuer, aber auch die Personen aus den Arbeitsgruppen in Bonn, Göttingen und Freiburg hervorheben. Auch möchte ich Katharina und Serhat für die Arbeit an unserer gemeinsamen Analyse danken. Ich werde mich besonders gerne an unsere kreativen und ausschweifenden Diskussionen erinnern.

Ich möchte auch meiner Arbeitsgruppe in Freiburg einen Dank aussprechen. Während der ganzen Zeit bot diese eine angenehme Arbeitsatmosphäre und die Möglichkeit, sich sowohl beruflich als auch privat auszutauschen. Gerade die Kaffeepausen und Gespräche, ob im Institut oder online während der Pandemie, ergänzten die Arbeit wertvoll. Hier möchte ich Benjamin, Benoit und David hervorheben, die sich über einen langen Zeitraum ein Büro mit mir geteilt haben und immer ein offenes Ohr für Fragen hatten, aber auch immer wieder für notwendige Ablenkungen gesorgt haben. Auch möchte ich mich speziell bei Dirk, Alena, Benjamin und David bedanken, welche mir nicht nur während der Arbeitszeit, sondern auch während der Freizeit einen wichtigen Rückhalt gaben. Auch gilt mein Dank allen voran Anton, Felix, Stefan und Michael, die es mir durch ihre tägliche Arbeit in der IT ermöglicht haben, meine Analysen durchzuführen.

Zusätzlich möchte ich mich bei allen Mitgliedern des Graduiertenkollegs bedanken. Insbesondere bei all jenen, die die diversen Veranstaltungen ermöglicht haben.

Zuletzt möchten ich noch all jenen danken, welche mir beim Verfassen dieser Arbeit durch Korrekturen geholfen haben. Hier sind vor allem Markus, David und Christina hervorzuheben, aber auch Benedict und Jacques.