Evidence for the Higgs Boson Decay to $\tau^+\tau^-$ in the Lepton + Hadrons Final State with ATLAS

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

von
Thomas Schwindt
aus
Mechernich

Bonn, 2014

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

1. Gutachter:	Prof. Dr. Norbert Wermes
2. Gutachter:	Prof. Dr. Klaus Desch
Tag der Promotion:	05.02.2015
Erscheinungsjahr:	2015

Contents

1 Introduction **1**

2 Theoretical Foundations **3**
2.1 The Standard Model of Particle Physics . . . 3
2.1.1 The Elementary Particles . . . 3
2.1.2 The Gauge Bosons . . . 4
2.1.3 Strong Interactions . . . 5
2.1.4 Electroweak Interactions . . . 6
2.1.5 The Role of the Mass . . . 7
2.1.6 The Higgs Mechanism . . . 9
2.2 Theoretical Predictions . . . 12
2.2.1 Basics of Scattering Theory . . . 12
2.2.2 Perturbative Calculations . . . 13
2.2.3 Running Couplings and Confinement . . . 13
2.3 Limitations of the Standard Model . . . 15
2.4 Physics of Proton-Proton-Interactions . . . 18
2.4.1 The Proton Structure . . . 18
2.4.2 Cross Section Calculation . . . 19
2.4.3 Observables in pp-Collisions . . . 20
2.5 Processes in pp-Collisions . . . 22
2.5.1 Dominant Processes . . . 22
2.5.2 QCD Initial-State-Radiation . . . 23
2.5.3 Higgs Boson Production . . . 23
2.6 Higgs Boson Decays . . . 25
2.6.1 Branching Ratios . . . 25
2.6.2 $\tau\tau$ Final States . . . 26
2.7 Monte Carlo Event Generation . . . 27
2.7.1 Parton Showers . . . 28
2.7.2 Generation of Unweighted Event Samples . . . 29

3 The ATLAS Experiment **31**
3.1 The Large Hadron Collider (LHC) . . . 31
3.2 The ATLAS Detector Systems . . . 32
3.2.1 The Inner Detector . . . 34
3.2.2 The Calorimeter . . . 35

3.2.3 The Muon Spectrometer . . . 36
3.3 Detector Simulation . . . 37
3.4 Reconstruction and Performance . . . 38
3.4.1 Data Reduction with Trigger Algorithms . . . 38
3.4.2 Basic Object Reconstruction . . . 38
3.4.3 High-Level Physics Objects . . . 39

4 Search Strategies for $H \to \tau_{\mathbf{lep}}\tau_{\mathbf{had}}$ Final States **45**
4.1 Properties of the $H \to \tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ Signal . . . 45
4.1.1 H+jets Final States . . . 45
4.2 Background Processes for $H \to \tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ Final States . . . 46
4.2.1 Irreducible Background: $Z \to \tau\tau$ . . . 46
4.2.2 Mis-Identified τ_h Objects . . . 48
4.2.3 Backgrounds from Complex Processes . . . 50
4.2.4 Monte Carlo Generator Implementations . . . 51
4.3 Extraction of Signal from Backgrounds . . . 53
4.3.1 Mass Reconstruction . . . 53
4.3.2 Analysis Observables . . . 56
4.3.3 Signal Extraction . . . 59

5 The $Z \to \tau\tau$ Background Process **63**
5.1 Measurement of the $Z \to \tau\tau$ Cross Section . . . 63
5.1.1 Event Selection . . . 63
5.1.2 Background Estimation . . . 66
5.1.3 Cross Section Calculation . . . 68
5.1.4 Systematic Uncertainties . . . 69
5.1.5 Results . . . 71
5.2 Modelling of $Z \to \tau\tau$ from $Z \to \mu\mu$ Data . . . 73
5.2.1 General Idea of the Embedding Method . . . 73
5.2.2 Event Selection . . . 76
5.2.3 Generation of $Z \to \tau\tau$ Decays . . . 76
5.2.4 Kinematic Filter . . . 76
5.2.5 Simulation of $Z \to \tau\tau$ Decays . . . 77
5.2.6 Replacement of Tracks and Cells . . . 78
5.2.7 Reconstruction of the Hybrid Event . . . 79
5.2.8 Estimation of Systematic Effects . . . 79
5.2.9 Validation of Embedded Data . . . 80
5.2.10 Corrections of Selection and Reconstruction Biases . . . 81
5.2.11 Applications and Outlook . . . 81

6 Cut-Based Analyses **83**
6.1 Search for SM $H \to \tau\tau$ Decays with 4.6 fb^{-1} of ATLAS Data . . . 83
6.1.1 Datasets and Event Selection . . . 83
6.1.2 Background Estimation . . . 85
6.1.3 Alternative OS-SS Estimation Method . . . 92
6.1.4 Systematic Uncertainties . . . 97
6.1.5 Results . . . 99

6.2 Optimised Analysis with 4.6+13.0 fb^{-1} of ATLAS Data . . . 101
6.2.1 Datasets and Event Selection . . . 101
6.2.2 Background Estimation . . . 103
6.2.3 Systematic Uncertainties . . . 111
6.2.4 Results . . . 112

7 MVA-Based Search for SM $H \to \tau\tau$ Decays with 20.3 fb^{-1} of ATLAS Data 115
7.1 Boosted Decision Trees . . . 116
7.1.1 Automated Generation of Decision Trees . . . 116
7.1.2 Machine Learning via Boosting . . . 116
7.2 Datasets and Event Selection . . . 118
7.3 Analysis Categorisation . . . 119
7.3.1 VBF Category . . . 119
7.3.2 Boosted Category . . . 120
7.3.3 1-jet and Rest Categories . . . 120
7.3.4 Additional Background Rejection . . . 120
7.4 Background Estimation . . . 120
7.4.1 VBF-Filtered Monte Carlo Samples . . . 120
7.4.2 Estimation of the r_{QCD} Extrapolation Factor . . . 121
7.4.3 Fake-Factor Estimation . . . 121
7.4.4 Correction Factors . . . 125
7.5 BDT Training Setup . . . 130
7.5.1 Signal and Background Composition . . . 130
7.5.2 n-Fold Cross-Evaluation . . . 130
7.5.3 Variable Selection and Parameter Optimisation . . . 131
7.6 Control Distributions . . . 135
7.7 Final Discriminant . . . 144
7.8 Systematic Uncertainties . . . 147
7.8.1 Background Modelling Uncertainties . . . 147
7.8.2 Experimental Systematic Uncertainties . . . 150
7.8.3 Theoretical Uncertainties . . . 151
7.9 Construction of the Fit Model for the Signal Extraction . . . 152
7.9.1 Control Regions and Normalisation Factors . . . 153
7.9.2 Statistical Uncertainties on the Background Model . . . 153
7.9.3 Validation of the Fit Model . . . 153
7.9.4 Systematic Uncertainties after the Final Fit . . . 161
7.10 Results . . . 164
7.10.1 Measurement of the Signal Strength . . . 164
7.10.2 Separation of the Higgs Boson Production Modes . . . 167
7.10.3 Statistical Significance of the Observed Excess . . . 167
7.10.4 Visualisation of the Observed Higgs Boson Mass . . . 168
7.10.5 Combination of $H \to \tau\tau$ Decay Channels . . . 170

8 Studies Towards a Mass Measurement 175
8.1 Compatibility of the Excess with Different Higgs Boson Mass Hypotheses . . . 175
8.2 Alternative Strategies for Mass Measurements . . . 178
8.2.1 Decorrelation of BDT and MMC . . . 178

8.2.2 Optimisation of the Variable Selection . . . 181
8.2.3 Optimisation of Training Parameters . . . 182
8.2.4 Mass-Fit Results . . . 182
8.2.5 Summary of the Mass Studies . . . 190

9 Conclusions **191**

A Search for SM $H \to \tau\tau$ Decays with 4.6 fb^{-1} of ATLAS Data **193**
A.1 Embedding Validation . . . 194
A.2 Validation of the Multijet Estimate . . . 198
A.3 Kinematic Distributions . . . 202

B Optimised Analysis with 4.6+13.0 fb^{-1} of ATLAS Data **207**
B.1 Embedding Validation . . . 208
B.2 Event Selection . . . 212

C MVA-Based Search for SM $H \to \tau\tau$ Decays with 20.3 fb^{-1} of ATLAS Data **213**
C.1 Background Estimation . . . 213
C.1.1 W+jets Re-weighting . . . 213
C.1.2 Background Summary . . . 215
C.2 Control Distributions . . . 218
C.3 Systematic Uncertainties . . . 220
C.3.1 Normalisation Uncertainties . . . 220
C.3.2 Systematic Uncertainties after the Final Fit . . . 226

D Studies Towards a Mass Measurement **231**
D.1 Control Distributions . . . 231

Bibliography **233**

CHAPTER 1

Introduction

More than 48 years after its first postulation by R. Brout, F. Englert, P. Higgs et al. in 1964 [1–4], 17 years after developing the conceptual design of the Large Hadron Collider (LHC) [5] dedicated to its search and only 2.5 years after the successful start of this machine, a new massive boson was observed by the ATLAS and CMS experiments [6, 7]. The mass of this new particle was determined to[1] $m = 125.5 \pm 0.2\,(\text{stat})\,^{+0.5}_{-0.6}\,(\text{sys})$ GeV[8], and all current observations favour the hypothesis of a spin-0 particle as postulated, while spin-1 can be excluded through theoretical arguments [9, 10] and spin-2 is disfavoured by dedicated measurements [11]. Following this discovery, the 2013 Nobel Prize was awarded to F. Englert and P. Higgs[2] "for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles, and which recently was confirmed through the discovery of the predicted fundamental particle" [12].

Although all observed couplings of this Higgs boson [8] are consistent with the predictions of the Standard Model of Particle Physics (SM), and also its mass agrees well with indirect measurements [13], one of its most important properties still remains to be confirmed in order to complete the predictions of the SM and to understand the origin of the masses of *all* elementary particles.

The existence of the Higgs boson is included in the SM, because the short range of the electroweak interaction could only be described by force carriers with large masses. Since the SM is based on a general field theory, which allows only massless force carriers, a new mechanism was necessary to explain the origin of the property, which we observe as their mass. A suitable explanation for this effect was found to be the spontaneous symmetry breaking of the vacuum, which requires the existence of a scalar vacuum field and describes the observed mass of a particle via the interaction of the particles with this field. P. Higgs was then the first to conclude that the excitation of this field must be observable as a new particle (a scalar boson), which was hence named after him.

While this interaction between force carriers (bosons) and the Higgs boson would already be sufficient to describe the large masses of the force carriers, the SM predicts that all masses of the elementary particles are generated by their coupling to this new field. Therefore, it includes additional so-called Yukawa couplings between all matter particles (fermions) and the Higgs field, which introduce the dif-

[1] Throughout this thesis, natural units with $\hbar = c = 1$ are used in order to simplify reading. Thus, e.g. masses are given in their energy equivalent $E = mc^2$.

[2] R. Brout, co-author of the publication with F. Englert, deceased in 2011.

ferent particle masses as additional parameters into the SM. While all decay channels contributing to the initial observation of the Higgs boson were dominated by its couplings to either the Z^0 or $W^\pm$ bosons, the next and most important step is now to observe a direct coupling between this new boson and a matter particle (fermion) as predicted in the SM.

Since the coupling strength of the Higgs boson to particles is proportional to their masses, good candidates to observe such couplings are the τ-leptons, which are the heavy partners of electrons with a more than three thousand times larger mass of $m_\tau = 1.78$ GeV [14]. Compared to the heavier quarks, these τ-leptons can result in clear signatures in proton-proton-collisions, so that $H \to \tau\tau$ decays provide a promising potential to confirm the direct coupling between the Higgs boson and matter particles.

In this thesis, the search for these decays is presented in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel, where one of the τ-leptons decays subsequently into a lighter lepton $\ell = e$ or μ and the other τ-lepton decays into one or more hadrons denoted as τ_{h}:

$$\tau^-_{\text{lep}} \hat{=} \ell^- \bar{\nu}_\ell \nu_\tau, \quad \tau^-_{\text{had}} \hat{=} \tau^-_{\text{h}} \nu_\tau$$

Following a general introduction of the Standard Model and the Higgs mechanism, an overview is given on the physics of proton-proton-collisions and their detection mechanisms used in the ATLAS experiment. After a description of the basic properties of the $H \to \tau_{\text{lep}}\tau_{\text{had}}$ signal and relevant background processes, general concepts and analysis techniques are discussed, which are later applied to extract signal from background in the collision data.

Starting with a first small dataset of data recorded with the ATLAS experiment, a first measurement of the $Z \to \tau\tau$ cross section is performed, which gives confidence in the simulation of the ATLAS detector response to $\tau_{\text{lep}}\tau_{\text{had}}$ final states. As a next step, the data-driven estimation of this dominant background process for the $H \to \tau\tau$ signal is described, which is achieved with an *embedding* method by replacing muons in $Z \to \mu\mu$ data events by simulated $\tau^+\tau^-$ final states.

With the τ-embedded data and the background estimation techniques developed in the context of the $Z \to \tau\tau$ cross section measurement, a cut-based analysis of $H \to \tau_{\text{lep}}\tau_{\text{had}}$ final states is then presented, which marks the first dedicated search for the SM Higgs boson in $\tau^+\tau^-$ final states performed on the dataset recorded by ATLAS in 2011. While this analysis is only sensitive to a six times larger signal than expected in the SM, several improvements are summarised, which lead to a significantly higher signal sensitivity of a second, optimised cut-based selection. Since this sensitivity is still not sufficient to observe a signal with a strength as predicted by the SM, the search is concluded with a multi-variate analysis based on Boosted Decision Trees. As its final result obtained from the dataset recorded by ATLAS in 2012, an excess of signal over the expected background from other SM processes is observed with a statistical significance of 3.2σ. This is evidence for $H \to \tau_{\text{lep}}\tau_{\text{had}}$ decays, which is consistent with the expected signal from the already discovered Higgs boson with a mass of $m_H = 125$ GeV[3]. This result, however, cannot provide an independent information on this mass. An additional outlook is therefore given on analysis strategies for future mass measurements, from which a better constraint on the Higgs boson mass can be expected.

[3] In combination with other $\tau^+\tau^-$ final states, this significance increases to 4.1σ, which confirms the evidence for $H \to \tau\tau$ decays.

CHAPTER 2

Theoretical Foundations

2.1 The Standard Model of Particle Physics

Since the first discovery of an elementary particle by J. J. Thomson [15] and the first indication of matter substructure found by E. Rutherford [16, 17], our understanding of the building blocks of nature has developed remarkably. During a century of discoveries of elementary particles, a single model has been established, which describes all constituents of matter and their microscopic interactions[1]. To date, this so-called *Standard Model* is one of the most successful theoretical models in physics. It was able to predict the existence of several new particles before their first direct observation and has been confirmed by a large number of experiments with high precision over the last decades. With the recent discovery of a new boson consistent with the predicted Higgs boson, the last missing piece has now been found.

2.1.1 The Elementary Particles

To our current knowledge, all elementary constituents of matter are fermions with spin-$\frac{1}{2}$. These particles carry three different types of charges, which subject them to the three fundamental interactions: the strong, the weak and the electro-magnetic interaction with the corresponding charges colour, weak isospin and electric charge[2]. The six quarks

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

exist with three different colour charges, while the six leptons are colourless:

$$\begin{pmatrix} e^- \\ \nu_e \end{pmatrix} \quad \begin{pmatrix} \mu^- \\ \nu_\mu \end{pmatrix} \quad \begin{pmatrix} \tau^- \\ \nu_\tau \end{pmatrix}$$

Both these fermion types are grouped into three doublets of the weak isospin, which are usually referred to as *families*. Each quark family contains an up-type and a down-type quark with electric charges $+2/3$ and $-1/3$ respectively. The lepton families consist of a charged lepton ℓ^- and a neutral neutrino ν_ℓ.

[1] As explained later in Sec. 2.1.5, the gravity cannot be included in the Standard Model and requires further extensions such as e.g. Supersymmetry

[2] The well-known electric charge Q/e is in fact not a fundamental charge in the Standard Model, but only the combination of the weak isospin and the hyper-charge Y given by $Q = T_3 + Y$ via the electroweak unification (cf. Sec 2.1.4).

For each of these fermions, a corresponding anti-fermion exists with similar properties but opposite electric charge. According to their different charges, the fermions are subject to the three fundamental interactions: neutrinos interact only weakly, whereas the charged leptons also take part in the electromagnetic interaction. The quarks carry all charges and are thus dominated by the strong interaction.

2.1.2 The Gauge Bosons

In general terms, the Standard Model is a relativistic quantum field theory, which describes the dynamics of a free fermionic field $\psi(x)$ with a Lagrangian density denoted as

$$\mathcal{L} = \bar{\psi}(x)(i\gamma^\mu \partial_\mu - m)\psi(x) \,, \tag{2.1}$$

where m denotes the mass of the fermion[3]. According to Noether's theorem, symmetries of the model under transformations in space time such as translation and rotation lead to the conservation of energy and momentum. Similarly, internal symmetries can be exploited in order to derive further conserved quantities and related interactions. As an example, a local gauge transformation of the general form

$$\psi(x) \rightarrow e^{i\theta(x)}\psi(x)$$

allows the field to have different phases at different points x. The free Lagrangian in Eqn. (2.1) is not invariant under this internal symmetry transformation, but it can be made invariant by introducing the covariant derivative D_μ, which includes a new field A_μ and a free parameter q:

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + iqA_\mu$$

The Lagrange density for free fermions can therefore only be invariant under a local gauge transformation if there is an additional interaction between the fermionic fields with the new vector field A_μ and a coupling parameter q, which leads to conserved charge currents. Depending on the structure of the gauge transformation, different fields, couplings and charges can be constructed, so that any interaction can in principle be imposed by choosing a corresponding symmetry. The electromagnetic interaction or *quantum electro-dynamics* (QED) requires only one additional vector field A_μ, for which an excitation is identified with the photon γ as its gauge boson. This single field corresponds to a generator $\theta(x)$ for the symmetry group $U(1)$, and the charge $q = e$ describes the coupling strength between a charged fermion and the photon. For the other fundamental interactions, however, the structure of generators and the resulting symmetry groups is more complicated: The strong interaction or *quantum chromo-dynamics* (QCD) with three different colour charges is mediated by eight gluon fields G_μ, which carry all combinations of different colour and anti-colour. Instead of a simple electro-magnetic interaction as explained above, a unified electro-weak interaction has been found to be realised in nature, which includes three additional weak fields W_μ. These couple to the weak isospin generated by $\mathrm{SU}(2)_T$ and mix with the Abelian field B_μ, which couples to the hyper-charge Y generated by $U(1)_Y$ as will be explained in Sec. 2.1.4. In total, the complete Standard Model can therefore be described for the free fermions by imposing a symmetry corresponding to the group

$$\mathrm{SU}(3)_C \otimes \mathrm{SU}(2)_T \otimes \mathrm{U}(1)_Y \,,$$

[3] Here, the γ^μ denotes the Dirac matrices and must not be confused with the photon γ, which is one of the gauge bosons itself.

Name	Field	Coupling	SU(3), SU(2), Y
Gluon	G_μ^a	g_S	(**8** , **1** , 0)
Weak Boson	W_μ^i	g	(**1** , **3** , 0)
Abelian Boson	B_μ	g'	(**1** , **1** , 0)

Table 2.1: The gauge bosons of the Standard Model and their multiplet structure within its symmetry groups (as presented in [18, 19]).

which automatically requires the existence of the three fundamental interactions via bosonic vector fields as a consequence of local gauge invariance. Quanta of these fields are identified as gauge bosons, which are the vector bosons listed in Tab. 2.1. By including the kinetic energy of the new fields in the Lagrange density, additional terms of the form

$$\mathcal{L}_A = -\frac{1}{4} A_{\mu\nu} A^{\mu\nu} \quad \text{with} \quad A_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

appear, which must also be gauge invariant. In the case of QED, the new field itself transforms according to $A_\mu \rightarrow \mathrm{A}_\mu - \frac{1}{e}\partial_\mu \theta(x)$, and the gauge invariance for the kinetic energy term is ensured due to the commutating derivatives $\partial_\mu \partial_\nu \theta(x) - \partial_\nu \partial_\mu \theta(x) = 0$. With multiple generators of non-Abelian groups, however, the transformation behaviour of the fields changes, which introduces self-interactions between gauge bosons of the symmetry groups $\mathrm{SU}(3)_C$ and $\mathrm{SU}(2)_T$, i.e. the gluons and the weak bosons.

2.1.3 Strong Interactions

The SU(3) symmetry of the strong interaction is generated by eight massless gluon fields G_μ^a ($a = 1, \ldots, 8$), which correspond to the generators t^a of the symmetry group and couple to the colour charges of the quark fields. These interactions between quarks and gluons are introduced via the covariant derivative

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + i g_S \left(t^a G_\mu^a \right)$$

with the coupling strength g_S. Gauge invariance in this non-Abelian group then requires the gluon fields to transform according to

$$G_\mu^a \rightarrow G_\mu^a - \frac{1}{g_S} \partial_\mu \theta_a(x) - f_{abc} G_\mu^b G_\mu^c$$

with the structure constants f_{abc} of SU(3), which connect the different colour charges via their commutator relation $[t^a, t^b] = i f_{abc} t_c$. The kinetic energy of the gluon fields

$$\mathcal{L}_{QCD} = -\frac{1}{4} G^{a\ \mu\nu} G_{\mu\nu}^a$$

is then added to the full Lagrangian with the term, where the field strength tensor

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_S f_{abc} G_\mu^b G_\nu^c$$

produces gluon self-interactions described with the SU(3) algebra of f_{abc}. This structure of the gluon fields is therefore interpreted as colour charge carried by gluons themselves, which has several phe-

nomenological implications such as the quark confinement and asymptotic freedom in QCD as discussed in Sec. 2.2.3.

2.1.4 Electroweak Interactions

The weak interaction has first been observed in nuclear β-decays, in which the electric charge of the decaying nucleus is always changed. This experimental fact hints at a deeper connection between electromagnetic and weak interactions, which is described in the Weinberg-Salam model of electroweak unification. Technically, this observation leads to the conclusion that the electric charge and the weak isospin operators do not commute. Therefore, the fundamental U(1) field cannot directly couple to the electric charge and thus cannot be the photon field A_μ. Instead, the fundamental Abelian field is denoted as B_μ and couples to the hyper-charge Y, which contributes to the electric charge according to the Gell-Mann–Nishijima relation:

$$Q = T_3 + \frac{Y}{2}$$

The generators σ_i ($i = 1, \ldots, 3$) of the SU(2) group require the three additional fields W^i_μ. The electroweak interaction with the fundamental $\mathrm{SU}(2)_T \otimes \mathrm{U}(1)_Y$ structure is then introduced into the Standard Model as in Eqn. (2.1) via the covariant derivative

$$\partial_\mu \to \partial_\mu \to D_\mu = \partial_\mu + ig\frac{\sigma^i}{2}W^i_\mu + ig'\frac{Y}{2}B_\mu$$

with the respective coupling strength g and g'. The physically observable *charged current* interactions, which change the electric charge of the fermion, are then mediated by the charge eigenstates

$$W^\pm_\mu = \frac{1}{\sqrt{2}}\left(W^1_\mu \mp iW^2_\mu\right)$$

and can be identified as $W^\pm$ bosons, which interact purely weakly. Their coupling has a so-called "V-A" (vector γ_μ minus axial vector $\gamma_\mu\gamma_5$) structure, which depends on the chirality of fermions. In fact, it is found to act only on left-handed particles, while it does not couple at all to their right-handed counterparts, as can experimentally be confirmed e.g. by observing a long lifetime of charged pions $\pi^\pm$. Therefore, it is necessary to distinguish the weak isospin of fermions according to their chirality: while left-handed fermions form doublets, the right-handed ones do not carry this charge and are thus singlets for the corresponding $\mathrm{SU}(2)_T$ symmetry. An important exception are right-handed neutrinos, which have neither weak isospin nor electric charge and are hence not observable and not included in the Standard Model.
The *neutral currents* of the electroweak interaction, which conserve the electric charge of the fermions, appear as mixed states of weak and electro-magnetic currents, being rotated by the Weinberg-angle θ_W:

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix}\begin{pmatrix} W^3_\mu \\ B_\mu \end{pmatrix}$$

The physically observable electroweak gauge bosons and their couplings are summarised in Tab. 2.2: the photon γ only couples to the electric charge $e = g\sin\theta_W = g'\cos\theta_W$ with a simple vector structure γ_μ. Due to the admixture of the weak boson W^3, which interacts only with left-handed fermions via a "V-A" coupling structure $\gamma_\mu\gamma_5$, the Z^0 boson instead shows a more complicated coupling structure interacting with both left- and charged right-handed fermions.

Name	Field	Coupling
$W^{\pm}$-Boson	$W^{\pm}_{\mu}$	$\propto g(\gamma_\mu - \gamma_\mu\gamma_5)$
Z^0-Boson	Z_μ	$\propto \frac{g}{\cos\theta_W}\left((T_3 + 2\sin\theta_W Q)\gamma_\mu - T_3\gamma_\mu\gamma_5\right)$
Photon (γ)	A_μ	$\propto eQ\gamma_\mu$

Table 2.2: The physically observable gauge bosons of the Standard Model and their couplings.

From the kinetic energy terms of the electroweak gauge fields

$$\mathcal{L}_{EW} = -\frac{1}{4}W^{i\ \mu\nu}W^{i}_{\mu\nu} - \frac{1}{4}B^{\mu\nu}B_{\mu\nu}\ ,$$

the Abelian U(1) group does not introduce any self-interaction via its field strength tensor $B_{\mu\nu}$. Since the SU(2) group is non-Abelian like SU(3) for the strong interaction, the corresponding tensor

$$W^{i}_{\mu\nu} = \partial_\mu W^{i}_{\nu} - \partial_\nu W^{i}_{\mu} - i\epsilon^{ijk}W^{j}_{\mu}W^{k}_{\nu}$$

contains the interactions between the weak fields described by the SU(2) algebra $\left[\sigma_i, \sigma_j\right] = 2i\epsilon_{ijk}\sigma_k$.

An additional important feature of electroweak interactions is the ability of its charged current to change quark flavours, while the neutral currents and the strong interaction always conserve this flavour in their interaction with quarks. Although changes within a family (e.g. $u \rightarrow W^+d$) can easily be described by the structure of the weak isospin, especially the observed flavour changes between different doublets (e.g. $s \rightarrow W^-u$) require a further extension of the model, introducing another set of parameters combined in the CKM quark-mixing matrix [20, 21]. This matrix describes the orthogonal combinations of down-type flavour eigenstates d_j to the weakly interacting eigenstates:

$$d'_i = \sum_j V^{\mathrm{CKM}}_{ij} d_j$$

While the diagonal elements of this unitary matrix are clearly dominant ($d'_i \approx d_i$) and thus favour changes within a doublet, the non-vanishing off-diagonal elements [14] describe the probabilities of transitions between families[4]. In summary, all fermions of the Standard Model can be grouped according to their multiplet structure within $\mathrm{SU(3)}_C \otimes \mathrm{SU(2)}_T \otimes \mathrm{U(1)}_Y$ as shown in Tab. 2.3.

2.1.5 The Role of the Mass

All three fermion families described in Sec. 2.1.1 and shown in Tab. 2.3 have identical properties when considering only the three fundamental interactions and their corresponding charges. Still, all elementary particles have a fourth distinctive property, which is their mass. Table 2.4 summarises the properties of the known fermions including their masses, which span a wide range of almost twelve orders of magnitude from the electron neutrino to the top quark. While these different masses have a large impact on the phenomenology observed in nature, to date their pattern cannot be explained by any underlying

[4] Similarly, the observed mixing of neutrino flavours can be described with such a matrix as summarised in [14], but will not be discussed further in this context.

Fermions	Generations			Q/e	T_3	SU(3),SU(2),Y
Quarks	$\begin{pmatrix} u \\ d' \end{pmatrix}_L$	$\begin{pmatrix} c \\ s' \end{pmatrix}_L$	$\begin{pmatrix} t \\ b' \end{pmatrix}_L$	$+2/3$ $-1/3$	$+1/2$ $-1/2$	$(\mathbf{3}, \mathbf{2}, +\frac{1}{6})$
	$u_R^\dagger$	$c_R^\dagger$	$t_R^\dagger$	$-2/3$	0	$(\bar{\mathbf{3}}, \mathbf{1}, -\frac{2}{3})$
	$d_R^\dagger$	$s_R^\dagger$	$b_R^\dagger$	$+1/3$	0	$(\bar{\mathbf{3}}, \mathbf{1}, +\frac{1}{3})$
Leptons	$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$	0 -1	$+1/2$ $-1/2$	$(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$
	$e_R^\dagger$	$\mu_R^\dagger$	$\tau_R^\dagger$	$+1$	0	$(\mathbf{1}, \mathbf{1}, +1)$

Table 2.3: The fermions of the Standard Model and their multiplet structure within its symmetry groups (as presented in [18, 19]).

Quarks	$\begin{pmatrix} u: & 2.3^{+0.7}_{-0.5} & \text{MeV} \\ d: & 4.8^{+0.7}_{-0.3} & \text{MeV} \end{pmatrix}$	$\begin{pmatrix} c: & 1.28\pm0.03 & \text{GeV} \\ s: & 95\pm5 & \text{MeV} \end{pmatrix}$	$\begin{pmatrix} t: & 173.2\pm0.9 & \text{GeV} \\ b: & 4.18\pm0.02 & \text{MeV} \end{pmatrix}$
Leptons	$\begin{pmatrix} e: & 0.511 & \text{MeV} \\ \nu_e: & < 2 & \text{eV} \end{pmatrix}$	$\begin{pmatrix} \mu: & 105.7 & \text{MeV} \\ \nu_\mu: & < 0.19 & \text{MeV} \end{pmatrix}$	$\begin{pmatrix} \tau: & 1.777 & \text{GeV} \\ \nu_\tau: & < 18.2 & \text{MeV} \end{pmatrix}$

Table 2.4: The elementary particles of the Standard Model and their masses as summarised in [14].

theory. However, due to the mass hierarchy, particles can decay very differently, and thus almost all stable matter known today is only made of the lightest up- and down quarks and electrons.

Although the gravitation would be an obvious candidate for a fourth fundamental interaction with the mass as its corresponding charge, the structure of this interaction requires its potential force carrier, the graviton, to be a spin-2 boson[5]. Due to this different structure, there is no simple way to combine the corresponding gravitational field with the other three vector fields of the Standard Model. Among others, the so-called Supersymmetry can provide such a unification by introducing an additional symmetry transformation between fermions and bosons [22], but no experimental evidence has been found until today.

Besides the unexplained mass pattern of the matter particles, the non-zero mass itself is an even more fundamental problem for the Standard Model: the chiral structure of the weak interaction forces a different behaviour of right- and left-handed fermions under gauge transformations

$$\psi_L \to \psi'_L = e^{i\alpha(x)\vec{\sigma}+i\beta(x)Y}\psi_L \qquad \psi_R \to \psi'_R = e^{i\beta(x)Y}\psi_R \,,$$

so that the general fermion mass terms of the form $m\bar{\psi}\psi$ in Eqn. (2.1) are no longer gauge invariant. Furthermore, any mass term for a gauge boson of the form $M_W^2 W_\mu W^\mu$ would also break the gauge invariance. Therefore, the theory even forbids the existence of mass as simple parameters and requires all particles - fermions and gauge bosons - to be massless. While this is found to be true for the photon and the gluons, the observed masses of the weak gauge bosons $m_W = 80.4$ GeV and $M_Z = 91.2$ GeV are found in clear contradiction to this expectation (as well as all large fermion masses). Thus, the Stan-

[5] As described by general relativity, gravity in fact does not directly couple to a particle's rest mass itself, but rather to its energy and momentum. This corresponds to the rest mass only in the non-relativistic Newtonian approximation.

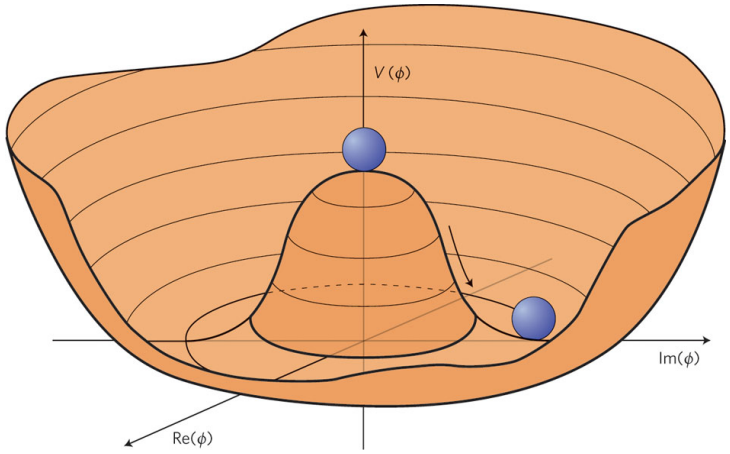


Figure 2.1: The effective potential of the Higgs-field, which leads to spontaneous symmetry breaking [25].

dard Model as described so far cannot explain all experimental observations and requires an additional mechanism to generate these large masses.

2.1.6 The Higgs Mechanism

The mechanism proposed by P. Higgs et al. [1–4, 23] applies ideas from superconductivity to elementary particle physics [24] and describes the mass of fermions and gauge bosons via their interaction with a scalar field, which is not generated by any charge, but must exist even in the vacuum. Due to the required interaction with the weak bosons, this so-called Higgs-field must carry weak isospin, so that its simplest form is a complex isospin doublet

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

The interaction of the electroweak $SU(2)_T \otimes U(1)_Y$ sector with this additional field ϕ is then described in the Standard Model Lagrange density with the additional term

$$\mathcal{L}_\phi = \left| \left(\partial_\mu + ig\frac{\vec{\sigma}}{2}\vec{W} + ig'\frac{Y}{2}B_\mu \right) \phi \right|^2 - V(\phi) , \tag{2.2}$$

where $V(\phi)$ describes the potential energy of the Higgs field. Restricting this to terms of the order $\propto \phi^4$, this potential obtains the form

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2 . \tag{2.3}$$

In order to ensure stability of the vacuum, the dominant parameter $\lambda > 0$ is chosen to be positive. With $\mu^2 > 0$, this additional field does not change the situation of the Standard Model, as its ground state with minimal $V(x)$ still corresponds to $< \phi >= 0$. When requiring $\mu^2 < 0$, however, the potential no longer has a single minimum as shown in Fig. 2.1. Instead, it develops a circle of minima with a non-vanishing vacuum expectation value $< \phi >= \sqrt{-\mu^2/\lambda} = \upsilon$. The ground state of such a potential is then chosen by nature as a specific point on this sphere, which spontaneously breaks the symmetry of the vacuum, since the potential is no longer symmetric around this point. As a consequence, the Higgs field does not vanish even in the ground state of the vacuum and can thus generate the masses of the elementary particles.

As only the weak bosons are massive, while the photon is found to be massless, this ground state as chosen by nature must be of the form

$$\phi_0 = \sqrt{\frac{1}{2}}\begin{pmatrix} 0 \\ \upsilon \end{pmatrix} \qquad \phi(x) = \sqrt{\frac{1}{2}}\begin{pmatrix} 0 \\ \upsilon + h(x) \end{pmatrix}$$

with a hyper-charge of $Y = +1$, while $\phi(x)$ can be written in terms of this ground state and $h(x)$ remains the only additional field in the model. The other three fields correspond to massless Goldstone bosons, but are absorbed by the longitudinal polarisation, which is an additional degree of freedom of the now massive bosons[6]. After inserting this choice into Eqn. (2.2), the result

$$\mathcal{L}_\phi = \underbrace{\left(\frac{1}{2}\upsilon g\right)^2}_{M_W^2} W_\mu^+ W^{-,\mu} + \frac{1}{8}\upsilon^2 \left(W_\mu^3, B_\mu\right)\begin{pmatrix} g^2 & -gg' \\ -gg' & g'^2 \end{pmatrix}\begin{pmatrix} W^{3,\mu} \\ B^\mu \end{pmatrix} - V(\phi_0)$$

includes an additional term for the W bosons of the form $M_W = 1/2\upsilon g$, which contains the vacuum expectation value υ of the Higgs field. Similarly, the mixed neutral fields

$$A_\mu = \frac{g' W_\mu^3 + g B_\mu}{\sqrt{g^2 + g'^2}} \quad \text{and} \quad Z_\mu = \frac{g W_\mu^3 - g' B_\mu}{\sqrt{g^2 + g'^2}}$$

obtain mass terms from this vacuum expectation value, so that in summary the following masses are predicted by this mechanism:

$$M_W = \frac{1}{2}\upsilon g\,, \quad M_Z = \frac{1}{2}\upsilon\sqrt{g^2 + g'^2}\,, \quad M_A = 0$$

For the Higgs field $h(x)$ itself, the potential $V(x)$ includes a corresponding term $\propto |h(x)|^2$, which can be identified as its mass term. The mass of the Higgs boson is thus predicted to be related to its self-coupling parameter[7] λ via $m_H = 2\upsilon^2\lambda$. The coupling strength of this Higgs boson to the pairs of gauge bosons is determined by

$$g_{HWW} = \frac{2M_W^2}{\upsilon}\,, \quad g_{HZZ} = \frac{2M_Z^2}{\upsilon}\,.$$

Generation of Fermion Masses

The Standard Model also introduces fermion masses via their interaction with the Higgs field similar to the mechanism generating the gauge boson masses. These are described by terms of the form

$$\mathcal{L}_{m_\ell} = -\frac{g_\ell}{\sqrt{2}}\left(\bar{\nu}', \bar{\ell}\right)_L \begin{pmatrix} 0 \\ \upsilon + h(x) \end{pmatrix} \ell_R + \text{hermitian conjugate.}$$

[6] The details of this specific choice and the absorption mechanism are described e.g. in [26].

[7] According to its definition in Eqn. (2.3), the parameter λ describes the quartic coupling $\propto |h(x)|^4$, which is in fact the self-coupling of the Higgs boson.

The coupling parameters g_ℓ are different for each type of fermion, and thus the resulting terms

$$\mathcal{L}_{m_\ell} = -\underbrace{\frac{g_\ell v}{\sqrt{2}}}_{m_\ell}\left(\bar{\ell}_L\ell_R + \bar{\ell}_R\ell_L\right) - \underbrace{\frac{g_\ell}{\sqrt{2}}}_{g_{H\ell\ell}}\left(\bar{\ell}_L\ell_R + \bar{\ell}_R\ell_L\right)h(x)$$

can again be interpreted as their masses and couplings being generated by the Higgs field. In the leptonic sector, the non-existing right-handed neutrinos imply $m_\nu = 0$ for the Standard Model[8]. In the quark sector, however, the generation of masses for up-type quarks requires the introduction of a new complex conjugated Higgs doublet $\phi_c = -i\sigma_2\phi*$ with opposite weak hyper charge $Y = -1$. After symmetry breaking, the choice of the ground state leads to the form

$$\phi_c(x) \to \sqrt{\frac{1}{2}}\begin{pmatrix} v + h(x) \\ 0 \end{pmatrix}.$$

Considering the up- and down-type weak eigenstates (u_i, d_i'), each family obtains an additional contribution to the Lagrangian

$$\mathcal{L}_{m_q} = -\frac{g_d^{ij}}{\sqrt{2}}\left(\bar{u}_i, \bar{d}_i'\right)_L\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}d_{j,R} - \frac{g_u^{ij}}{\sqrt{2}}\left(\bar{u}_i, \bar{d}_i'\right)_L\begin{pmatrix} v + h(x) \\ 0 \end{pmatrix}u_{j,R} + \text{h.c.}\,,$$

which again leads to the diagonal form with mass and coupling terms:

$$\mathcal{L}_{m_q} = -m_d^i\bar{d}_i d_i - m_u^i\bar{u}_i u_i - g^i_{Hd\bar{d}}\bar{d}_i d_i h(x) - g^i_{Hu\bar{u}}\bar{u}_i u_i h(x)\,.$$

As for each gauge boson, the resulting coupling strength $g_{Hf\bar{f}}$ of the Higgs boson to a fermion pair is proportional to its mass:

$$g_{Hf\bar{f}} = \frac{g_f}{\sqrt{2}} = \frac{m_f}{v}$$

From the vacuum expectation value v of the Higgs field, which can be measured e.g. via the Fermi constant G_F in $\mu^- \to e^-\bar{\nu}_e\nu_\mu$ decays

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

all couplings of the Higgs boson in the Standard Model are determined by the known fermion and gauge boson masses, while only the quartic coupling λ and thus the mass m_H of the Higgs boson are not predicted and can only be measured. This almost complete determination, however, is a consequence of the implementation of only one Higgs field. In nature, it is not necessary that gauge bosons and fermions or even up-type and down-type fermions interact with the same scalar field, so that only the measurement of all different types of couplings can confirm the specific implementation of the Higgs mechanism in the Standard Model.

[8] Experimental observations of neutrino oscillations show that neutrinos in fact have small masses $m_\nu > 0$, which are negligible in the context of this thesis. Several extensions of the Standard Model can provide consistent explanations of such masses, which are not discussed here.

2.2 Theoretical Predictions

After introducing the constituents of matter and the structure of their fundamental interactions in the Standard Model, the phenomenology of the corresponding processes must be derived from quantum mechanical calculations, for which the required techniques shall be described in the following. These calculations provide the basis for all theoretical predictions of specific interactions and can eventually be used to generate samples of events representing a process.

2.2.1 Basics of Scattering Theory

As opposed to classical mechanics, where the status of a system is well defined at any point of time for a given initial state, particle physics can only describe nature within quantum mechanics and thus can only calculate probabilities for an initial state to arrive at a given final state. In scattering theory, the basis for particle physics, interactions are described as transition from initial state $|i\rangle$ to final state $|f\rangle$ via the so-called S-matrix of probability amplitudes

$$\langle f|\hat{S}\,|i\rangle = \langle f|i\rangle + 2\pi\delta(E_i - E_f)\mathcal{M}\,,$$

where $\delta(E_i - E_f)$ requires the conservation of energy and momentum between the initial and final state four-momenta E_i and E_f. The interaction amplitude $\mathcal{M}$ describes the physics process and must therefore incorporate the interaction, which provides the transition from usually two initial state particles to n_F final state particles. The differential probability to arrive at a final state $|f\rangle \neq |i\rangle$ is given by

$$d\mathcal{P} = |\langle f|S\,|i\rangle|^2 \prod_{j=1}^{n_F} \frac{d\vec{p}_j}{(2\pi)^3}\,. \tag{2.4}$$

In order to calculate such a probability for a specific scattering process, an integration of the amplitude $|\mathcal{M}|^2$ in Eqn. (2.4) must be performed over the multi-dimensional phase space of all n_F final state particles. In a general scattering experiment, however, the number of colliding pairs is determined via the flux Φ of the colliding particles and the observation time T. Therefore, the interaction probability is usually substituted by the *cross section* σ, which is normalised accordingly:

$$d\sigma = \frac{1}{\Phi \cdot T} d\mathcal{P}$$

For a collider experiment, the flux is given via its *luminosity* L, which corresponds to the number of total collision events per unit area and unit time ($[L]$ =cm$^{-2}\cdot$ s-1). Integrated over the observation time, the product of luminosity and cross section is directly proportional to the number of interactions

$$N = \sigma \cdot \int_T L\,\mathrm{d}t\,.$$

Similar to the classical description of scattering, which can define a scattering probability as effective overlap of extended objects, the cross section is given in units of area:

$$[\sigma] = \mathrm{m}^2 = 10^{28}\ \mathrm{barn}$$

It is proportional to the interaction probability and thus the parameter of interest for most processes in collider experiments.

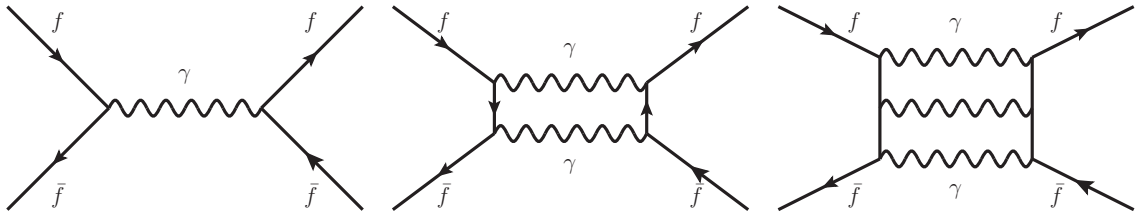


Figure 2.2: Feynman diagrams representing examples of leading- (a), next-to-leading (b) and next-to-next-to-leading order (c) matrix element contributions to the electromagnetic interaction of a fermion anti-fermion pair.

2.2.2 Perturbative Calculations

Due to the quantum mechanical nature of particles physics, the intermediate path between initial and final state is not defined, since it cannot be observed. Therefore, any integral of a matrix element $|\mathcal{M}|^2 = |\mathcal{M}_1 + \mathcal{M}_2 + \ldots|^2$ must include the superposition of all possible paths $\mathcal{M}_1, \mathcal{M}_2, \ldots$, which can provide the transition. Any calculation for a scattering process must thus be the integral over the sum of all possible interactions, which is usually impossible both analytically and numerically. In order to obtain an approximation of the true result, the calculation is usually performed in a perturbation series, which orders the different terms according to the size of their contributions. These terms are usually grouped by the number of couplings involved: for the basic coupling with two vertices, the corresponding term is proportional to $\sqrt{\alpha} \cdot \sqrt{\alpha}$ and is thus considered to be $O(\alpha^1)$ (cf. Fig. 2.2a), which is called *leading order* (LO) or born-level. A second interaction with two additional vertices hence enters the calculation as *next-to-leading order* (NLO) term with higher order $O(\alpha^2)$ (cf. Fig. 2.2b). As long as the coupling is small ($\alpha < 1$), terms with higher order interactions can usually be neglected, and the leading terms of the perturbation series up to the next-to-next-to-leading order (NNLO) provide a very good approximation of the result of a full calculation, while upper limits on contributions of the neglected terms are estimated and used as systematic uncertainty on the approximate result.

2.2.3 Running Couplings and Confinement

The basic interaction of two fermions via one of the three fundamental forces of the Standard Model is described via the exchange of a virtual gauge boson. However, as shown for the example of scattering two electrons (cf. Fig. 2.3), also higher order fermion loops can contribute to their interaction when fermion anti-fermion pairs are created from the vacuum and then annihilate again to the initial photon. As a quantum mechanical process, the interaction of two fermions via a photon is in fact a superposition of all possible loop corrections, so that the effective coupling between the electrons is not simply proportional to the coupling constant $\alpha = e^2/4\pi$, but must also be calculated in a perturbation series, where each additional loop is suppressed by a factor α.

Although the higher order correction terms depend on the momentum transfer Q^2 carried by the virtual photon, their calculation is logarithmically divergent, since the fermion momenta in the additional loops are not constrained. Therefore, these terms have to be renormalised by relating them to a measured coupling α at a fixed momentum transfer $Q^2 = \mu^2$, which can then absorb the divergent terms in a finite number. From this fixed point, the renormalisation group equation

$$Q^2 \frac{\partial \alpha}{\partial Q^2} = -b\alpha\left(1 + b'\alpha + b''\alpha^2 + O(\alpha^3)\right)$$

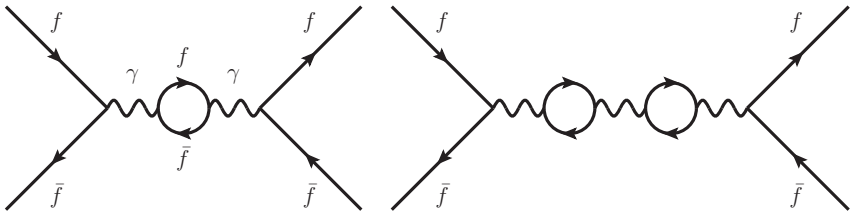


Figure 2.3: First order (a) and second order (b) loop corrections to the photon propagator in the electromagnetic interaction of a fermion anti-fermion pair.

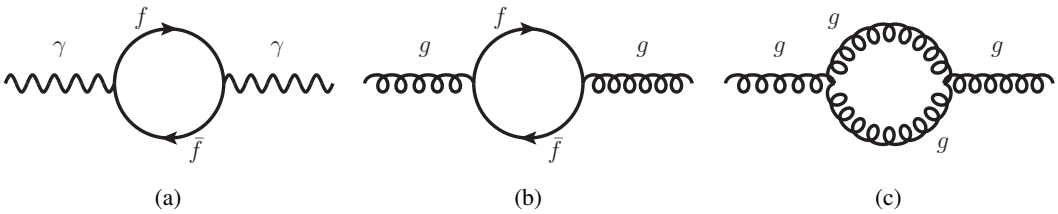


Figure 2.4: First order loop corrections to the photon- and gluon-mediated interaction. In addition to fermionic contributions to the photon (a) and gluon (b) propagators, the gluon self-interaction (c) leads to a different behaviour of the QCD as compared to QED.

describes the "running" of the coupling constant depending on the momentum transfer Q^2, while the coefficients b, b', b'' are determined by the effective couplings in higher order terms. A solution to this equation can only be given in a perturbative way by neglecting higher order terms and is usually restricted to the first order

$$\alpha(Q^2) = \frac{\alpha(\mu^2)}{1 + b\alpha(\mu^2)\log\left(Q^2/\mu^2\right)} \,.$$

The extrapolation from a given $\alpha(\mu^2)$ to any $\alpha(Q^2)$ is thus not exact, but is only possible with an uncertainty, which must be taken into account by estimating the size of the neglected terms.

Comparing the renormalisation group equation both for electro-magnetic (QED) and strong interactions (QCD), the coefficients can be calculated as

$$b_{\mathrm{QED}} = -\frac{1}{3\pi} < 0 \,, \qquad b_{\mathrm{QCD}} = \frac{33 - 2n_q}{12\pi} > 0 \,,$$

where n_q denotes the number of quark types with masses $m_q < Q^2$ smaller than the momentum transfer. Here, the fermion loops with quarks or leptons (cf. Fig. 2.4a) yield only negative contributions to the coefficient b_{QED}. Instead, the gluon self-interaction in QCD creates boson loops in addition to the quark loops (cf. Fig. 2.4b-2.4c). These provide positive contributions to the coefficient, which exceed the negative quark loop terms and thus lead to a positive $b_{\mathrm{QCD}} > 0$.

In QED, the negative coefficient causes the coupling α_{QED} to increase with higher momentum transfer Q^2. At experimentally accessible energies it stays small enough, so that QED can still be described via perturbative approximation of limited order. For the full electroweak interaction, the calculations are more complicated and involve both fermion and boson loops, but still provide negative coefficients for all couplings, which keep higher order corrections small and controllable in perturbation series.

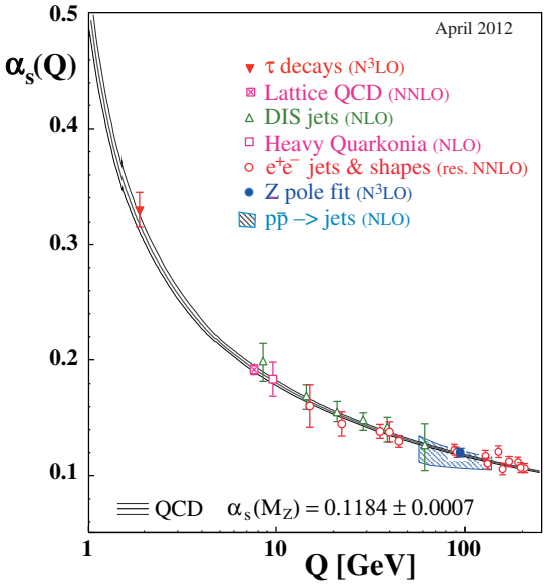


Figure 2.5: Best fit of the running coupling $\alpha_S(Q^2)$ to the data from different experiments [27].

For QCD, however, the positive coefficient b_{QCD} instead results in the opposite behaviour. The coupling α_{QCD}, often denoted as α_S, decreases only with higher momentum transfers, but becomes very large at low energies. This effect is shown in Fig. 2.5 as a summary of different measurements. Since high momentum transfers correspond to short range interactions and low momenta to long ranges, this behaviour is reflected in the observation that the strong interaction is only dominant at small distances within nuclei, but vanishes very quickly beyond. Quarks are thus usually confined in bound states and can never be observed as free particles due to a shielding effect: on separation of a quark pair, the energy of the strong field between their colour charges becomes sufficiently large to create a new quark pair, which shields the initial charges and fields by creating two new pairs with smaller field energies. The running of the coupling α_S from high to low energies therefore dominates the phenomenology of nature: at high energies, the strong interaction of coloured particles usually creates collimated cascades of quasi-free quarks and gluons. These continue to split under small angles until their energies become low enough, so that the strong coupling confines them into bound states, which are e.g. the nucleons n, p and other hadrons such as charged and neutral pions $\pi^{\pm}$, π^0. The resulting bundles of hadrons are called *jets* and are created in almost all processes involving strong interactions at high energies.

2.3 Limitations of the Standard Model

The Standard Model provides a very successful theoretical description of particle physics, and a global fit to all its measured parameters shows a remarkable agreement with the predictions [13] as demonstrated in Fig. 2.6.

Still, theoretical arguments indicate that this model cannot be valid up to very high energies, but must break down latest at the scale of the Planck mass $m_P \approx 10^{16}$ GeV. Especially the loop correction terms to the Higgs coupling, which are represented by the diagrams illustrated in Fig. 2.7, lead to quadratic divergencies. Unlike the logarithmic divergencies of other couplings, these cannot be renormalised, so that the integral over all momenta in the loops must be cut-off at a given scale Λ, which can be interpreted as an upper bound for the validity of the Standard Model.

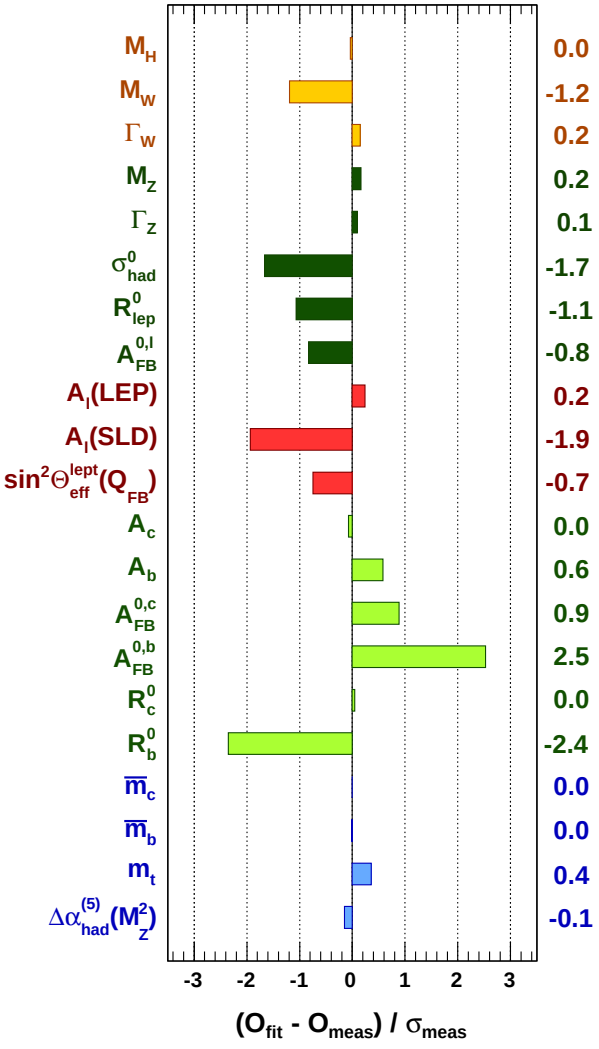


Figure 2.6: Deviations between direct experimental measurements and theoretical calculations of the SM in units of the experimental uncertainty [13].

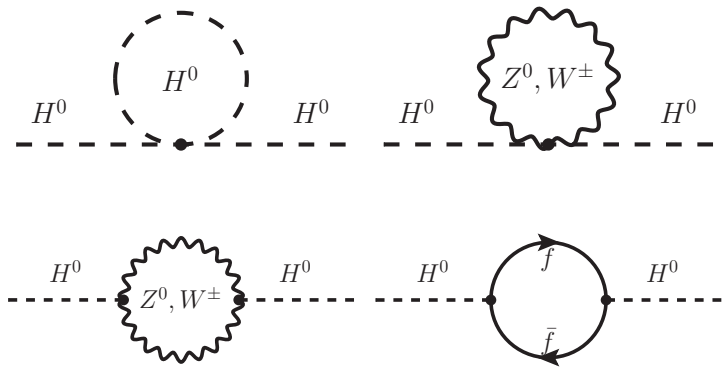


Figure 2.7: Bosonic and fermionic loop-corrections for the Higgs propagator.

With a given bare Higgs mass $m_{H,0}$, this scale defines the effective and observable Higgs mass m_H by absorbing all corresponding loop corrections up to the given momentum. This eventually leads to a correction term to the Higgs mass

$$m_H^2 = m_{H,0}^2 - \frac{3\Lambda^2}{8\pi^2 v^2}\left[4m_t^2 - m_H^2 - 2m_W^3 - m_Z^2\right] .$$

This term strongly depends on the cut-off scale Λ, since the quadratic mass difference is of the order of the vacuum expectation value v. Thus, the relatively small observable Higgs mass $m_H = O(100\ \text{GeV})$ would only be the result of the difference of two very large numbers, which must be *fine-tuned* such, that the large bare mass $m_{H,0}$ and the cut-off scale Λ are of similar order of magnitude.

Although it cannot be excluded that such a *fine-tuning* is realised in nature, already the existence of quadratic divergencies and introduction of a hard cut-off scale are no satisfying solutions from the theoretical point of view. Instead, contributions to the loop-corrections can also be cancelled by fermionic loops, which are introduced by Supersymmtery [22] as super-partners of the gauge bosons. In addition to solving this *hierarchy problem*, similar corrections for the gauge boson propagators would also be able to explain a unification of the three fundamental couplings at a large energy scale. Furthermore, certain super-partners fulfil all requirements for a dark matter particle candidate, which could explain approximately 24% of the energy content in our universe [28].

Despite all benefits of the (minimal) supersymmetric extension of the Standard Model (MSSM), it also introduces a large number of additional parameters to the theory. While already the pattern of the particle masses in the SM, the neutrino mixing angles and the quark mixing parameters in the CKM-matrix lack a satisfactory explanation as of today, even more mass parameters and mixing angles of Higgs doublets would have to be determined and explained for the MSSM, which would inevitably lead to more fundamental questions about the structure of our universe.

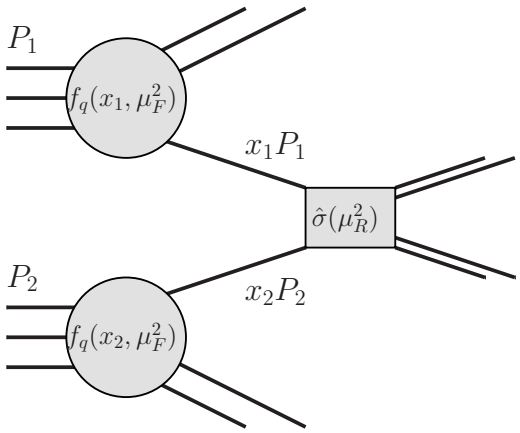


Figure 2.8: Example process in a pp-collision: From each of the two protons with momenta P_1, P_2 only one parton with momentum fraction $x_i P_i$ takes part in the interaction [34].

2.4 Physics of Proton-Proton-Interactions

Historically, experiments in high energy physics have always tried to observe elementary particles with large masses m, which must be produced in interactions at sufficiently high energies $E > mc^2$. After first observations in cosmic rays in 1912 by V. Hess [29], collider experiments have been developed successfully since the invention of the cyclotron in 1934 by E. Lawrence [30]. When the e^+e^--collider LEP reached its technical limit of $\sqrt{s} = 208$ GeV due to synchrotron radiation losses $\propto \left(E^4/m^4 r\right)$, the Large Hadron Collider (LHC) [31–33] was built at the "Centre Européen pour la Recherche Nucléaire" (CERN). This machine is designed to collide protons instead of electrons, which can be accelerated to higher energies while radiation losses are suppressed by a factor $(m_p/m_e)^4 \approx 11.4 \cdot 10^{12}$. Their collisions can now reach much higher centre-of-mass energies $\sqrt{s} \leq 14$ GeV, but the substructure of the protons leads to lower effective collision energies $\sqrt{\hat{s}}$ of their constituents with less well defined kinematic constraints and much more crowded events as compared to the clean environment in e^+e^--collisions. The processes in these pp-interactions are therefore dominated by QCD effects, which require dedicated phenomenological models, calculation techniques and observable definitions.

2.4.1 The Proton Structure

The collision of protons at high energies leads to inelastic scattering processes, which probe the substructure of the protons. These are compound particles and consist of three so-called valence quarks, which determine their main properties such as the electric charge. The strong interaction confines them into a bound state and therefore produces a cloud of gluons and quark-antiquark pairs inside the protons. These interactions, however, cannot be predicted with standard theoretical calculations, since the strong coupling α_S is large at low energy scales. The collision of two protons can thus only be described at very high energy and momentum transfers, which causes the running coupling α_S to be weak and the proton constituents to be quasi-free. Under these prerequisites, the interaction of two protons can be reduced to an interaction of two so-called partons, which are assumed to be free particles carrying momentum fractions $x_{1,2}$ of the colliding protons as shown in Fig. 2.8.

The probability to find a parton to be a specific quark, antiquark or gluon carrying a momentum fraction x of the proton is extracted from dedicated measurements [35] and is provided in parton-distribution-functions $f_p(x, Q^2)$. This probability depends additionally on the momentum transfer Q^2,

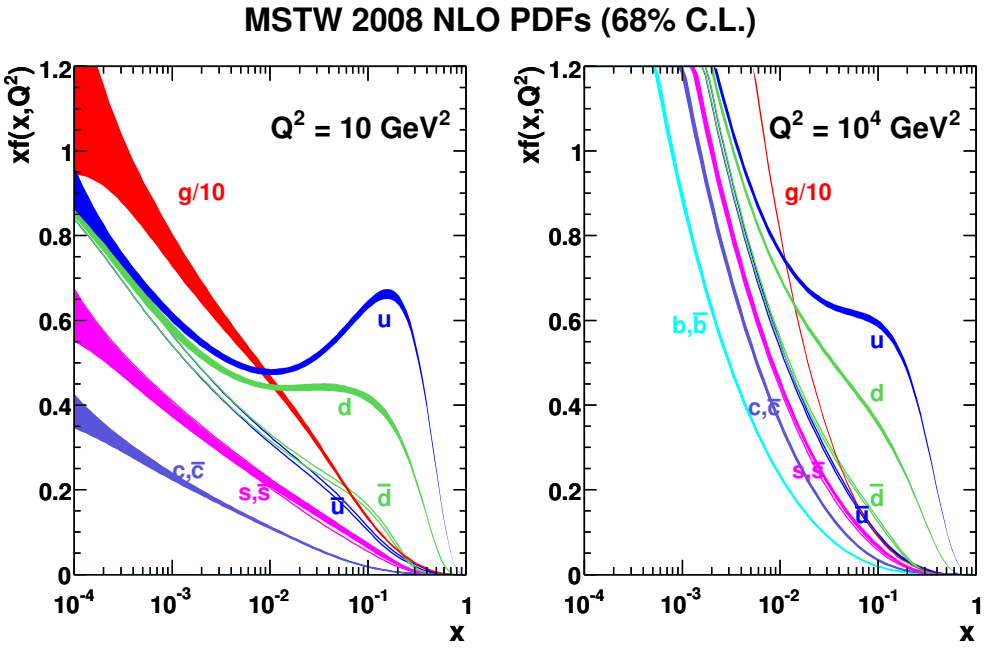


Figure 2.9: Parton-distribution-functions and uncertainties at 68% confidence level for momentum transfers of $Q^2 = 10\ \mathrm{GeV}^2$ (left) and $Q^2 = 10^4\ \mathrm{GeV}^2$ (right) [36].

which corresponds to the energy resolution probing the proton structure[9]. The measured distributions of the parton inside the proton (cf. Fig. 2.9) confirm that each of the valence quarks carries on average approximately one third of the proton momentum, while the interactions of protons are otherwise dominated by gluons. Other quark and antiquark types even with $m_q > m_p$ can also be found inside a proton, which are the so-called sea-quarks carrying much lower momentum fractions.

Besides the hard scattering interaction of the two partons, the remnants of the two protons can produce secondary interactions, which usually have significantly lower energies. This effect must also be modelled phenomenologically according to dedicated measurements and is usually referred to as multi-parton-interaction (MPI) or underlying-event (UE).

2.4.2 Cross Section Calculation

While the basic calculation of a cross section usually starts from a given initial state, pp-collisions provide a variety of initial states with different combinations of quarks and gluons with continuously distributed momenta $p_1 = x_1 P_1, p_2 = x_2 P_2$ resulting from the fractions x_1, x_2 of the proton momenta P_1, P_2. Cross sections for pp-collisions must therefore sum over all possible processes with initial states q, q' of different parton types ($q, q' \in g, u, d, s, c, b$) and momentum fractions x_1, x_2:

$$\sigma_{pp\to X}(P_1, P_2) = \sum_{q,q'} \int \mathrm{d}x_1 \mathrm{d}x_2 f_q(x_1, \mu_F^2) f_{q'}(x_2, \mu_F^2) \hat{\sigma}_{q,q'}(p_1, p_2, \alpha_S(\mu_R^2), Q^2/\mu_R^2)\,. \tag{2.5}$$

[9] Since these measurements are usually performed at lower energies and Q^2, an extrapolation to the energy scale of the LHC is necessary.

Here, the parton distribution functions are evaluated at the so-called *factorisation scale* $Q^2 = \mu_F^2$, so that all higher order QCD effects inside the proton with smaller momentum transfers $Q^2 < \mu_F^2$ are effectively absorbed by the phenomenological description of the proton structure. Similarly, all higher order effects with $Q^2 < \mu_R^2$ for the strong interaction of the partons are absorbed by evaluating the running coupling $\alpha_S(\mu_R^2)$ at the *renormalisation scale* μ_R^2. Both parameters are introduced in order to describe QCD processes with perturbative methods. The result of a full calculation including all higher orders would not depend on their choice, whereas the dependence of the perturbative approximation on these scales must be taken into account as systematic uncertainty of the calculation. As a common convention, these scales are usually chosen to be identical and to correspond to the centre-of-mass energy $\hat{s}$ of the parton-parton interaction

$$\mu^2 = \mu_R^2 = \mu_F^2 = \hat{s} = (p_1 + p_2)^2 \,.$$

The related uncertainty of the perturbative approximation is then evaluated by varying both scales together within a wide range between $\mu \to \mu \cdot 2$ and $\mu \to \mu/2$.

2.4.3 Observables in pp-Collisions

Due to the random momentum fractions $x_{1,2}$ carried by the two interacting partons in pp-collisions, the longitudinal boost of their centre-of-mass system is not determined a priori. Since the proton remnants mostly follow the beam direction and cannot be observed, they lead to missing momentum in the direction of the colliding beams. Therefore, only the transverse components of all observable kinematics are comparable quantities for all collision events, while all components in beam direction are distributed randomly due to the different boosts of the parton collisions. Momenta of all particles produced in a pp-collision are thus reduced to their transverse component

$$\vec{p}_T = \begin{pmatrix} p_x \\ p_y \end{pmatrix}, \quad p_T = \left|\vec{p}_T\right| = \sqrt{p_x^2 + p_y^2} \,,$$

where the z-axis is aligned with the beam direction. Although the total momentum conservation cannot be exploited due to the missing longitudinal components, the transverse momentum is still conserved in each collision, so that $\sum \vec{p}_T = \vec{0}$ can be assumed.

Due to the rotational symmetry in a collision, positions and distances of particles are given in spherical coordinates (ϕ, η), where ϕ is the angle in the x-y plane transverse to the beam. While the angle θ between the direction of the particle and the z-axis would transform with the longitudinal boost, the pseudo-rapidity η is used instead. It is defined as the logarithmic transformation

$$\eta = -\log\left(\tan\frac{\theta}{2}\right) \,,$$

where $\eta = 0$ corresponds to the plane perpendicular to the beam direction. For highly relativistic particles ($p \gg m$, $p \approx E$), this is equivalent to the rapidity

$$\eta \underbrace{\to}_{p \gg m} y = \frac{1}{2}\log\left(\frac{E + p_z}{E - p_z}\right) \,,$$

for which the distance Δy between two objects originating from the same pp-collision is invariant under longitudinal boosts of the colliding parton-parton system. While absolute directions (ϕ, η) are largely irrelevant due to the rotational symmetry in ϕ and the boost in the η direction, their relative distances

$(\Delta\phi, \Delta\eta) = (|\phi_2 - \phi_1|, |\eta_2 - \eta_1|)$ or the three-dimensional distance

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$$

are important invariant observables, which are comparable between different collision events. In cases where invisible particles are produced, their η-component cannot be observed, and only $\Delta\phi = |\phi_p - \phi_{\text{miss}}|$ can provide information about correlations between invisible and visible particles. With this angle, a so-called *transverse mass* is defined, which can include the momenta of invisible particles:

$$m_T = \sqrt{2 \cdot p_T \cdot p_{\text{T}}^{\text{miss}} (1 - \cos\Delta\phi)}$$

2.5 Processes in pp-Collisions

Before focusing on the production of Higgs bosons, the probabilities for all other processes in pp-collisions must be considered, which are relevant for the background composition in search for a potential signal from Higgs boson decays.

2.5.1 Dominant Processes

In a pp-collision, the two partons usually interact via the strong force, which usually creates two or more jets of hadrons. These so-called *multijet* events are thus by far the dominant process. In very rare cases, however, the partons can also produce electroweak bosons (cf. Fig. 2.10), which can decay into leptons with large transverse momenta. In an experiment, such leptons can be identified very well, so that in general most electroweak interactions can easily be extracted from the large number of multijet events. Similarly, the production of a Higgs boson can be identified and separated from this background if it subsequently decays into a leptonic final state. The background for such a signal is therefore dominated by other electroweak processes involving Z^0 and $W^\pm$ bosons, which are still produced with much higher rates than Higgs bosons.

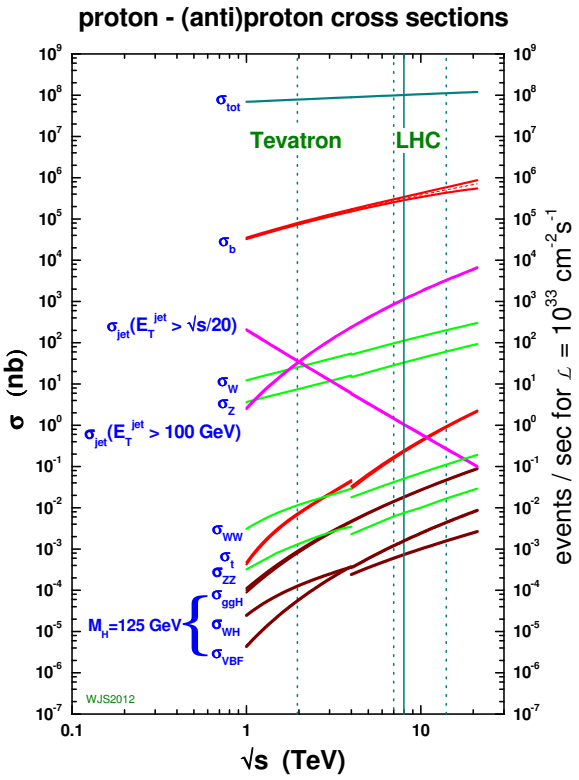


Figure 2.10: Production cross sections for different processes in pp-collisions and their dependence on the centre-of-mass energy $\sqrt{s}$ [36, 37]. The solid vertical line indicates the cross sections for $\sqrt{s} = 8$ TeV relevant for the LHC as operated in 2012.

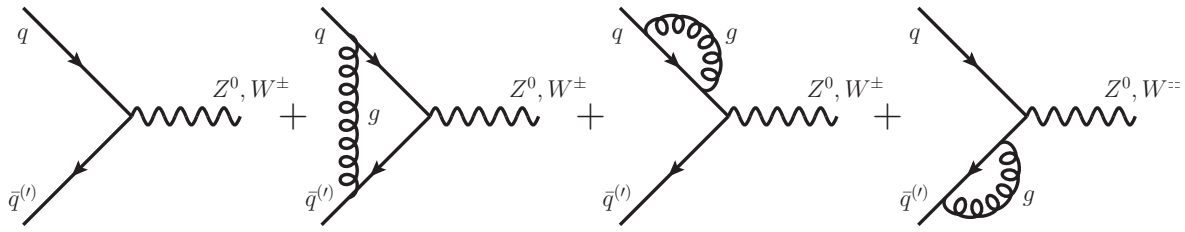


Figure 2.11: Next-to-leading order loop-corrections for the vector boson production process $q\bar{q}^{(\prime)} \to V + g$.

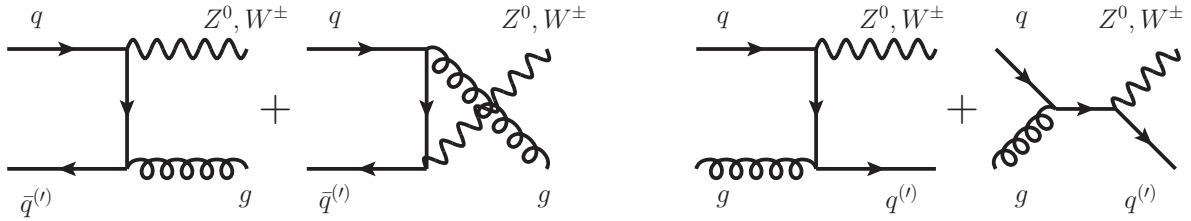


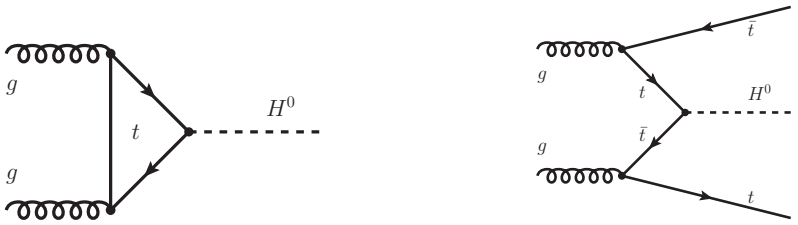
Figure 2.12: Leading-order matrix element representations of the vector boson ($V = W^\pm, Z^0$) production process with one additional parton: $q\bar{q}^{(\prime)} \to V + g$ (left) and $qg \to V + q^{(\prime)}$ (right).

2.5.2 QCD Initial-State-Radiation

In the collision of the two hard partons, Z^0, $W^\pm$ or Higgs bosons are produced via their electroweak interaction or their interaction with the Higgs field. Additionally, these partons, are subject to the strong interaction, which substantially affects their scattering processes. Besides virtual higher order corrections from QCD loops (cf. Fig. 2.11), also real emissions of quarks or gluons can occur as initial state radiation, which are similarly suppressed by the running coupling $\alpha_S < 1$ at high energies. Such additional parton emissions shown in Fig. 2.12 produce jets of hadrons and thus result in final states including both leptons and jets with large transverse momenta. Since the transverse momentum conservation must still hold in these events, these jets require a recoil of the Z^0, $W^\pm$ or H bosons and can therefore significantly affect the kinematics and topologies of their subsequent decays.

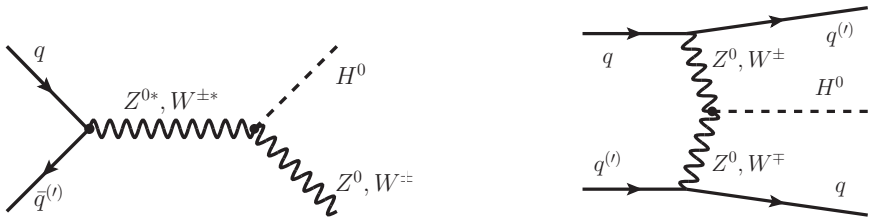
2.5.3 Higgs Boson Production

Considering the production of Higgs bosons in pp-collisions, the simplest mechanism would be the direct coupling of the interacting partons with the Higgs field. In most cases, however, these partons are gluons or the light valence quarks with only a very weak coupling to the Higgs boson. Although the gluons are massless and thus have no direct coupling, they can interact with the Higgs boson via a fermion loop, which is dominated by contributions from top quarks. The most probable Higgs production mechanism at the LHC is thus the gluon-fusion (ggH) depicted in Fig. 2.13a. Due to the properties of the strong interaction, the initial-state-radiation of gluons from the loop can create additional jets, which must then recoil against the Higgs boson. With large transverse momenta $p_T(H)$, special "boosted" topologies of the subsequent Higgs decays can occur, which provide distinct signatures for experimental searches. A similar process can create two quasi-free quarks in association with the Higgs boson (ttH) as shown in Fig. 2.13b. As compared to ggH, this production mechanism is strongly suppressed due to the required centre-of-mass energy for two free top quarks in addition to the Higgs boson ($\sqrt{\hat{s}} > 2m_t + m_H$). Therefore, also the bottom-associated production (bbH) contributes with a comparable production cross section despite the weaker g_{Hbb} coupling.



(a) (b)

Figure 2.13: Production modes for the Higgs boson via the g_{Htt} coupling: (a) gluon-fusion ggH and (b) top-associated production.



(a) (b)

Figure 2.14: Production modes for the Higgs boson via the g_{HVV} coupling: (a) Higgs-Strahlung and (b) vector-boson-fusion qqH.

Instead of the fermionic g_{Htt} (g_{Hbb}) coupling, two further production mechanisms involve the g_{HVV} coupling to the heavy vector bosons, which are shown in Fig. 2.14. In the vector-boson-associated production (VH), a Higgs boson is radiated off a virtual Z^0 or $W^{\pm}$ boson (cf. Fig 2.14a), so that the final state includes both a Higgs and a weak boson. This so-called Higgs-Strahlung is similar to the production of vector bosons, but is heavily suppressed due to the required virtuality of the vector boson ($m_{V^\star} > m_H + m_V$). A very different mechanism is the so-called vector-boson-fusion (VBF) depicted in Fig. 2.14b, which is in fact a scattering of the two interacting quarks ($qq \rightarrow qqH$), as opposed to the annihilation of quarks or gluons in the other processes. This scattering is mediated by two vector bosons, which are radiated off each quark and then annihilate in the production of the Higgs boson. Although this process is suppressed by the additional weak coupling, the virtuality requirement for the vector bosons is less strong than for the VH production. Furthermore, the average momentum fraction x_q carried by the valence quarks is much higher than for gluons or sea-quarks required by the initial states of the other production mechanisms, so that the fraction of parton collisions with sufficiently high centre-of-mass energy $\sqrt{\hat{s}}$ is larger. This qqH production creates a very distinct VBF event topology, in which the two quarks from the colliding protons are scattered with relatively low momentum transfer and form two highly energetic jets with a large separation in η.

Comparing the production cross sections for $m_H = 125\,\text{GeV}$ in Fig. 2.10 and Tab. 2.5, the ggH production is the dominant process followed by the qqH production via VBF and the Higgs-Strahlung ($\sigma(VH) = \sigma(WH + ZH)$), which are of similar size. The specific experimental VBF signature, however, can better distinguish this mechanism from other electroweak background processes, so that the qqH production is of equal importance as the ggH process.

Process	σ in pb	Systematic Uncertainty in %	
		QCD scale	PDF $\oplus$ α_S
ggH	19.27	$^{+7.2}_{-7.8}$	$^{+7.5}_{-6.9}$
qqH	1.578	$^{+0.2}_{-0.2}$	$^{+2.6}_{-2.8}$
WH	0.705	$^{+1.0}_{-1.0}$	$^{+2.3}_{-2.3}$
ZH	0.415	$^{+3.1}_{-3.1}$	$^{+2.5}_{-2.5}$
bbH	0.211	$^{+13.0}_{-23.0}$	$^{+6.4}_{-6.4}$
ttH	0.129	$^{+3.8}_{-9.3}$	$^{+8.1}_{-8.1}$

Table 2.5: Cross sections for the different Higgs boson production processes and their relative systematic uncertainties (in %) from the choice of the QCD scale, the choice of parton-distribution-function (PDF) and the uncertainty on α_S [59].

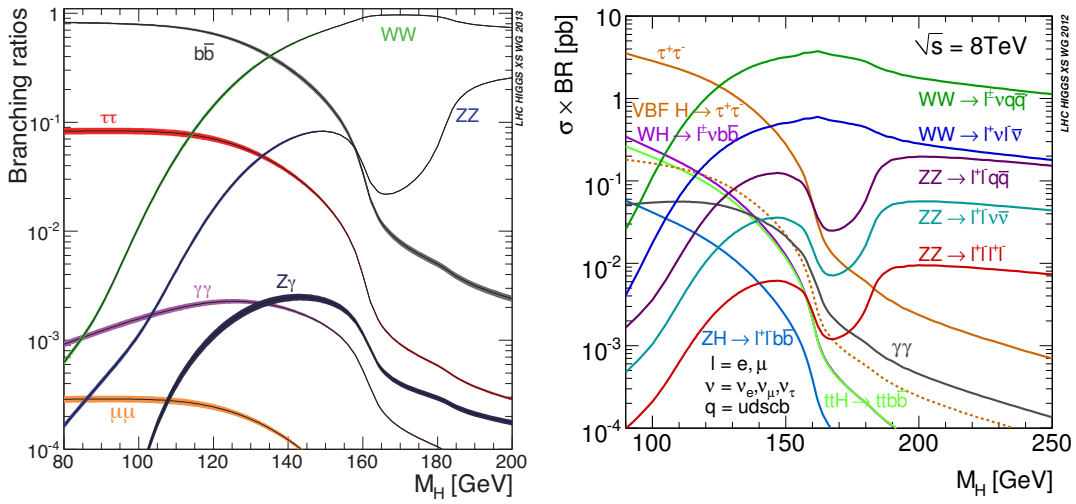
The calculations for these cross sections include NNLO QCD and NLO electroweak corrections for the ggH [38–49] and the VH [50–52] process, and full NLO QCD and electroweak corrections [53–55] with an approximation of NNLO QCD [56] for the qqH production. The results are summarised in [57–59].

2.6 Higgs Boson Decays

According to the prediction of the Standard Model, the Higgs boson couples to all elementary particles with a coupling strength proportional to their rest mass. It can thus decay into a variety of final states, so that the observable number of Higgs bosons in a specific final state is given by the corresponding branching ratio.

2.6.1 Branching Ratios

Depending on the mass of the Higgs boson, the Standard Model predicts different branching ratios for its decay into pairs of fermions or bosons, which are illustrated in Fig. 2.15a. Especially the kinematic threshold of $m_H > 2m_X$ for its decay $H \to XX$ into a pair of particles with masses m_X dominates its branching ratios and has a strong impact on possible observations. For a low Higgs boson mass, decays into fermionic final states ($b\bar{b}$, $\tau\tau$) are dominant, while the probability for decays into $WW^\star$ and $ZZ^\star$ pairs with virtual bosons increases with higher masses and becomes dominant when the Higgs mass approaches $2m_W$ ($2m_Z$) and real W (Z) bosons can be produced. For a mass of $m_H = 125$ GeV, the expected branching ratios are listed in Tab. 2.6, in which the decay channel $H \to b\bar{b}$ is dominant (58%), followed by $H \to WW$ (22%). Considering the corresponding background processes, however, it is very difficult to find a statistically significant excess of inclusive $b\bar{b}$ final states in the abundant multijet background from pp-collisions. Therefore, only the Higgs-Strahlung processes $WH \to \ell\nu b\bar{b}$ and $ZH \to \ell^+\ell^- b\bar{b}$ are accessible via the leptonic decays of the vector bosons. Similarly, a large fraction of W^+W^- subsequently decaying into the fully hadronic final states ($WW \to q\bar{q}'q\bar{q}'$) are difficult to disentangle from the multijet background, so that Higgs searches are mainly performed in the leptonic final states $H \to WW \to \ell^+\nu\ell^-\bar{\nu}$ (2.3% with $\ell \in \{e, \mu, \tau\}$) and $H \to WW \to \ell\nu q\bar{q}'$ (9.4% with $\ell \in \{e, \mu, \tau\}$), for which the branching ratios are included in Fig. 2.15b. The decay of the Higgs boson



(a) (b)

Figure 2.15: Branching ratio for the different decay channels of the Higgs boson in dependence of its mass m_H (a) and cross section times branching ratio for the experimentally accessible final states (b) [59].

Decay Channel	$b\bar{b}$	$c\bar{c}$	$\tau\tau$	gg	$\gamma\gamma$	$Z\gamma$	WW	ZZ
Branching Ratio (in %)	57.7	2.9	6.3	8.6	0.23	0.15	21.5	2.6
Relative Uncertainty (in %)	$^{+3.2}_{-3.3}$	$^{+12.2}_{-12.2}$	$^{+5.7}_{-5.7}$	$^{+10.2}_{-10.0}$	$^{+5.0}_{-4.9}$	$^{+9.0}_{-8.8}$	$^{+4.3}_{-4.2}$	$^{+4.3}_{-4.2}$

Table 2.6: Branching fractions and relative uncertainties (in %) of a Higgs boson with $m_H = 125$ GeV into the different decay channels [59].

into $\tau\tau$ final states (6.3%) has thus one of the largest products of cross section times branching ratio as compared to the other experimentally accessible final states, and is hence the fermionic decay channel with the largest discovery potential.

2.6.2 $\tau\tau$ Final States

From the experimental point of view, the decay of a Higgs boson into a $\tau^+\tau^-$ pair only denotes an intermediate state, which usually cannot be observed directly. Since τ-leptons have a mean lifetime of $\tau_\tau = (290.6 \pm 1.0)$ fs, in Higgs boson decays they usually decay within a distance[10] of 3 mm to the interaction point, so that only their subsequent decay products can be identified as final state. The mass of the τ-lepton $m_\tau = (1776.82 \pm 0.16)$ MeV is large enough for it to decay not only into a lighter lepton $\ell = e$ or μ, but as well into one or more light hadrons (cf. Tab. 2.7), which are usually referred to as a τ_h object. The notation is thus chosen to be

$$\tau^-_{\text{lep}} \hat{=} \ell^- \bar{\nu}_\ell \nu_\tau \,, \quad \tau^-_{\text{had}} \hat{=} \tau^-_h \nu_\tau$$

[10] For average momenta of the τ-leptons from Higgs boson decays with $p_\tau = m_H/2 \approx 60$ GeV, the mean flight distance is estimated via $L = \beta\gamma c\tau_\tau = p_\tau/m_\tau c\tau_\tau \approx 3$ mm.

τ^- Decay Mode	Branching Ratio
$\mu^-\bar{\nu}_\mu\nu_\tau$	$(17.41 \pm 0.04)\%$
$e^-\bar{\nu}_e\nu_\tau$	$(17.83 \pm 0.04)\%$
$\pi^-\nu_\tau$	$(10.83 \pm 0.06)\%$
$\pi^-\pi^0\nu_\tau$	$(25.95 \pm 0.09)\%$
$\pi^-\pi^0\pi^0\nu_\tau$ (ex. K_0)	$(9.36 \pm 0.11)\%$
$\pi^-\pi^+\pi^-\nu_\tau$	$(9.31 \pm 0.06)\%$
$\pi^-\pi^+\pi^-\pi^0\nu_\tau$	$(4.62 \pm 0.06)\%$

Table 2.7: Dominant decay modes of τ^--leptons [14].

in order to distinguish between the τ-lepton $\tau^\pm_{\text{had}}$ ($\tau^\pm_{\text{lep}}$) decaying hadronically (leptonically) and its visible decay product $\tau^\pm_{\text{h}}$ ($\ell^\pm = e^\pm$ or $\mu^\pm$). The combination of two independent decays can therefore lead to a variety of final states, which include two or more unobservable neutrinos. While these decay combinations are common for the irreducible $Z \to \tau^+\tau^-$ background, the composition of all other Standard Model backgrounds depends strongly on the final state under consideration. Especially in pp-collisions, this leads to three separate analyses in terms of event selection and background estimation:

While fully leptonic $\tau^+\tau^-$ final states ($\ell\ell 4\nu$ with $\ell\ell = ee, e\mu, \mu\mu$) have the smallest branching fraction of $\tau^+\tau^-$ pairs ($\approx 12\%$), they can be selected cleanly with only a very small contribution from multijet backgrounds. Dominant backgrounds are the electroweak $Z \to \ell^+\ell^-$ process for ee and $\mu\mu$ final states, while the top pair production involving two $W \to \ell\nu$ decays yields a relevant contribution of $e\mu$ events.

A larger fraction of the $\tau^+\tau^-$ decays ($\approx 42\%$) includes all fully hadronic final states ($\tau^+_{\text{h}}\tau^-_{\text{h}}2\nu$), which are more difficult to select from the huge number of multijet events produced in pp-collisions. Although the electroweak τ decays into hadrons can experimentally be distinguished from jets originating from purely strong interactions, even a small fraction of the abundant multijet events contributes a significant amount of di-jet events being mis-identified as two hadronic τ decays. These create a large background for Higgs signatures in the fully hadronic final state.

The remaining fraction of $\tau^+\tau^-$ final states ($\approx 46\%$) contains both a hadronic and a leptonic τ decay ($\ell^\pm\tau^\mp_{\text{h}}3\nu$), which represents the largest fraction due to the possible combinations of charges. It has a slightly less clean signature than the $\ell^+\ell^-4\nu$ final states, since only one electron or muon can be selected. However, three times more Higgs signal events can be expected due to the branching fraction. While this decay channel thus yields the largest significance, its background composition is the most complex one, as it involves large contributions from very different production mechanisms.

2.7 Monte Carlo Event Generation

In order to arrive at theoretical predictions of the different processes in pp-collisions, the phase space integral in Eqn. (2.4) must be calculated for the matrix elements corresponding to the desired interaction of the two initial state partons leading to a certain final state. Since this integral is usually multi-dimensional and cannot be solved analytically, the numerical Monte Carlo integration technique is applied, which scans the phase space by evaluating the matrix elements at a large number of random points. Each of these points corresponds to a single collision event representing the specific process, so that the calculation technique automatically provides a sample of events containing all kinematics of the involved particles. Their creation usually starts from the two colliding partons p_1, p_2 with momentum

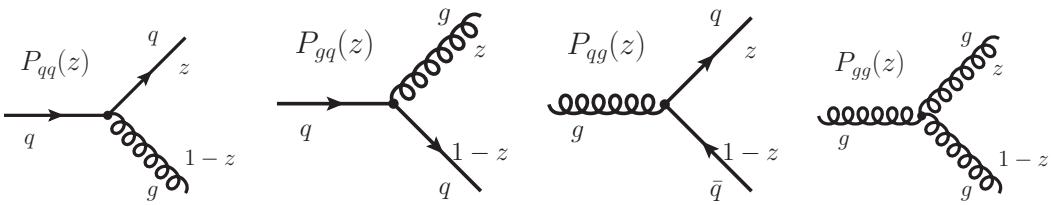


Figure 2.16: Quark- and gluon-splittings as described by Altarelli-Parisi functions [60].

fractions x_{p_1} and x_{p_2}, which are weighted according to the probability given by the parton-distribution-functions for the colliding protons. With constraints on the particle content and phase space boundaries of the desired final state, events are randomly created based on energy-momentum conservation. From the sum over all evaluated matrix element values for the created events, the cross section for the generated process is then calculated, so that the obtained result corresponds to the overall probability for the generated process to occur in a proton-proton-collision.

2.7.1 Parton Showers

The perturbative calculation of a leading order process with a limited number of loop corrections usually provides a good approximation, when the coupling is small ($\alpha < 1$) and higher order terms can be neglected due to vanishing contributions. For the strong interaction, however, any process evolving the partons towards lower energies to form stable hadrons instead would require perturbation series with $\alpha_S \geq 1$. Higher order terms can no longer be neglected in this case, since their contributions grow with the order n: $\alpha_S^{n+1} > \alpha_S^n$. In addition, the increasingly strong coupling for low momentum transfers leads to divergent phase space integrals, which result in infinite probabilities for soft and collinear emissions of partons. The theoretical prediction of the low energy regime of the strong interaction must therefore rely on a phenomenological model, which is provided by *parton shower* algorithms.

Since the perturbative calculation at high momentum transfers Q^2 and low $\alpha_S(Q^2) \ll 1$ must be combined with the phenomenological model of the showering, the phase space for the integration is usually restricted to large momentum transfers by applying kinematic cuts on the partons. The remaining part of the phase space, which would include the divergent part of the integration for soft and collinear emissions, is instead covered by the shower algorithm. This relies on a model of parton splitting probabilities: the so-called *Altarelli-Parisi* functions [60] $P_{qq}(z)$, $P_{gq}(z)$, $P_{qg}(z)$ and $P_{gg}(z)$ (cf. Fig. 2.16) describe the probability P_{ij} that a parton ($i = q, g$) splits into one parton ($j = q, g$) with momentum fraction z and a second one with $1-z$. Together with the *Sudakov-Form-Factors* [61], which provide the probability of a parton to propagate without splitting during a time interval $[t_0, t]$, the partons and their splittings are then combined to parton shower cascades as shown in Fig. 2.17.

In order to describe the formation of jets as observed in nature, the phenomenological model of the shower requires an ordering of the splittings e.g. according to their corresponding momentum transfers[11] $Q_0^2 > Q_1^2 > \ldots > Q_T^2$. In addition, an energy threshold Q_T^2 must be defined, below which no further splitting is performed and fragmentation algorithms form stable hadrons from the final partons of the shower [62, 63].

[11] These showers can also be ordered according to the transverse momenta or the angles of the parton emission.

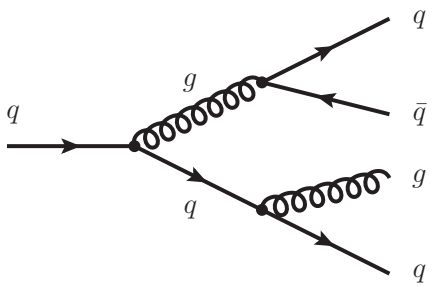


Figure 2.17: Example for a parton shower cascade involving quark- and gluon-splittings.

While these algorithms are usually applied to parton emissions as obtained from matrix element calculations, also an inverse showering algorithm [64] can be applied to the initial state partons. With this inverse shower, QCD initial-state-radiation as shown in Fig. 2.12 can also be provided by this parton shower, which can be used instead of the additional matrix elements in the phase space integral.

2.7.2 Generation of Unweighted Event Samples

Although the Monte Carlo technique produces samples of events for a given process, from which all kinematical distributions can be obtained, each event enters the calculation with a corresponding probability for its production. The composition of the generated sample therefore does not correspond to a set of events as collected in real collisions, but is only valid when each event is used in combination with its calculated weight. For practical reasons such as the later simulation of experimental effects in a detector, the generated events are usually unweighted by randomly rejecting unlikely configurations according to their weights. This is performed via the generation of another random number between zero and the maximum possible weight for each event. If the matrix element weight of an event is larger than the random number, the event is kept as an unweighted representative of its kinematic configuration, so that events with large weights and probabilities are accepted more frequently, while events with low weights and probabilities are discarded in most cases. The resulting sample is thus produced with a correct composition of events, which are expected from the generated process in real collision events.

CHAPTER 3

The ATLAS Experiment

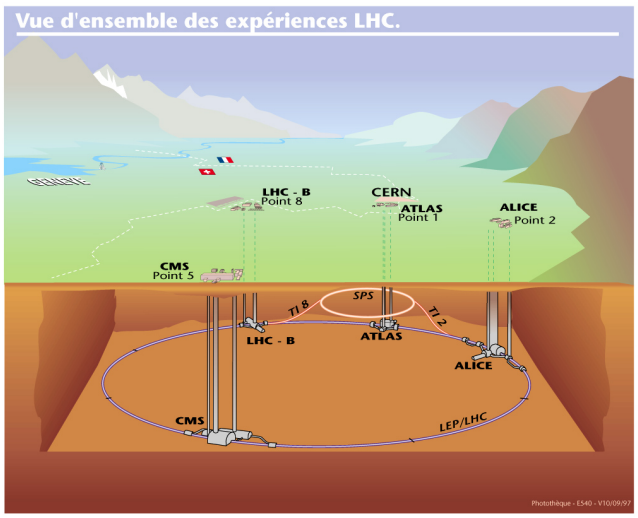


Figure 3.1: Overall view of the LHC experiments [65].

The latest direct searches for the Higgs boson in $\tau\tau$ final states are performed with two multi-purpose experiments: **A** **T**oroidal **LHC** **A**pparatu**S** (ATLAS) and the **C**ompact **M**uon **S**pectrometer (CMS). Both are highly sophisticated machines consisting of multiple components to detect different types of particles after their production in proton-proton-collisions delivered by the Large Hadron Collider. The results presented within this thesis are obtained from data recorded with the ATLAS detector, for which an overview of the experimental setup, the detector components and the particle reconstruction algorithms and techniques is given in the following.

3.1 The Large Hadron Collider (LHC)

The proton-proton-collider LHC at CERN is installed in a circular tunnel with a circumference of 27 km, which is situated approximately 100 m below ground level near Geneva at the border between France and Switzerland (cf. Fig. 3.1) and was originally built for the e^+e^--collider LEP operating since 1989.

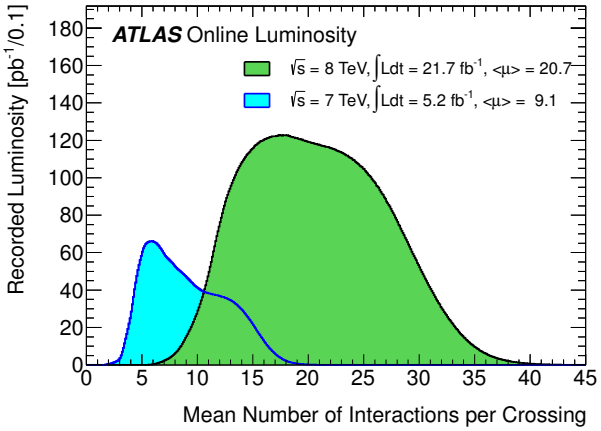


Figure 3.2: Distribution of the mean number of simultaneous interactions per bunch crossing for the LHC runs in 2011 (blue) and 2012 (green) [68].

Over the full length of the tunnel, a superconducting system of 1232 dipole and 392 quadrupole magnets bends two proton beams into circular orbits and focuses them for collisions at four interaction points, where the main experiments ALICE, ATLAS, CMS and LHCb are located[1]. The colliding protons are first stripped from hydrogen gas and accelerated in one linear and three circular pre-accelerators (PBS, PS, SPS) until they are finally filled in bunches into the main LHC storage-ring. During the operation in 2012 (2011, 2010) [66, 67], each beam usually consisted of 1374 (1331, 348) bunches containing up to 1.7 (1.2, 0.9) $\cdot 10^{11}$ protons each, which were accelerated to an energy of 4.0 (3.5, 3.5) TeV and thus brought to collision at a centre-of-mass energy of $\sqrt{s}$ =8 (7, 7) TeV. A crossing of the bunches at the interaction points occurred every 50 (50, 150) ns, leading to collisions of several protons from each bunch. With these beam parameters used at the LHC in 2012 (2011, 2010), maximum instantaneous luminosities of L = 7.7 (3.6, 0.2) $\cdot 10^{33}$ cm^{-2} s^{-1} could be reached, where each bunch crossing produced up to 40 (20, 5) simultaneous pp-collisions. Figure 3.2 shows the distribution of the number of interactions per bunch crossing, which is referred to as *in-time pile-up*, and effectively increases the total number of collisions in order to observe a sufficient number of events from processes with low cross sections such as the Higgs production. During the first run period, the total number of collisions in a single interaction point of the LHC corresponded to an integrated luminosity of $\int$L dt=23.1 fb^{-1} (5.5 fb^{-1}, 47 pb^{-1}) of which the ATLAS detector recorded $\int$L dt=21.3 fb^{-1} (5.1 fb^{-1}, 45 pb^{-1}) [68].

3.2 The ATLAS Detector Systems

In order to detect particles from collision events efficiently and measure their kinematics precisely, the layout of the ATLAS detector as shown in Fig. 3.3 follows the typical structure of particle physics experiments with a barrel-like, cylindrically symmetric shape with end-caps at each side, covering a pseudo-rapidity range of $|\eta| < 4.9$ corresponding to a polar angle of $1° < \theta < 179°$ to the beam pipe. Closest to the interaction point, the *inner detector* is placed inside a homogeneous magnetic field, which allows to reconstruct the tracks of charged particles and measure their momenta via curvature. The extrapolation of all tracks to their origins provides measurements of the primary interaction vertex or secondary decay

[1] In addition to the two multi-purpose experiments ATLAS and CMS, the ALICE experiment is designed to analyse the quark-gluon plasma created in heavy-ion collisions, while LHCb is optimised to study CP-violation effects in specific processes.

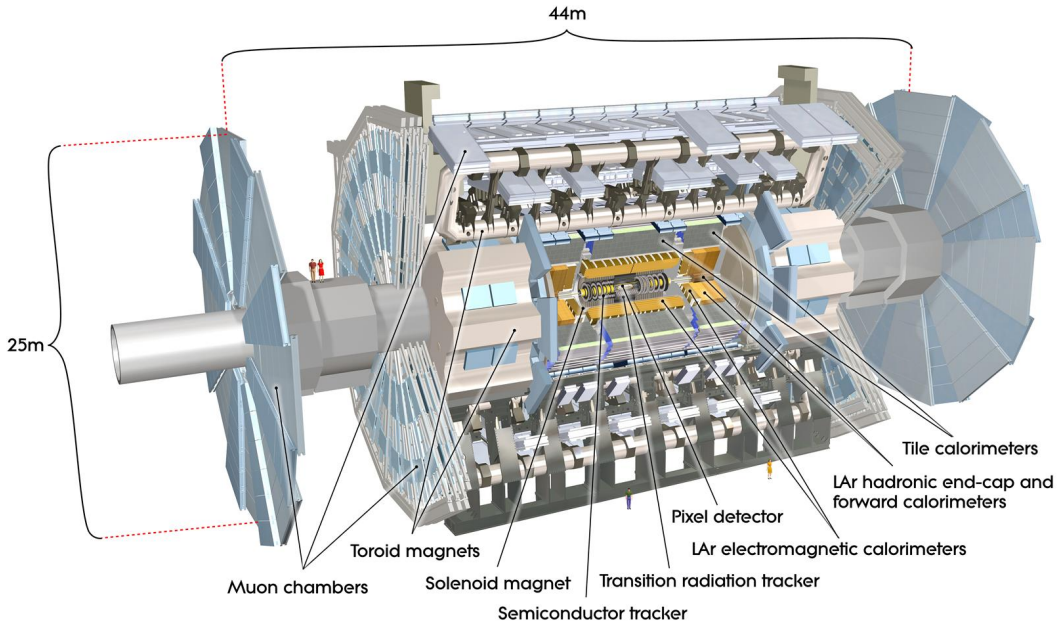


Figure 3.3: Cut-away view of the ATLAS detector [69].

vertices in a collision event. Very thin layers are desired to keep the interactions of the particles with the inner detector components at a minimum, so that their energies can be measured precisely in the *Calorimeter*, which is divided into two parts. These are separately optimised to absorb the total energy of a particle either via an electromagnetic or hadronic showering process. While electrons and photons are usually stopped within a limited amount of material of the **E**lectromagnetic **CAL**orimeter, especially highly energetic jets of hadrons require a sufficiently large **H**adron **CAL**orimeter, which must contain the whole hadronic shower and completely absorb its energy. Due to this technical constraint, the solenoid creating the magnetic field for the track momentum measurement is placed between the inner detector and the ECAL. With this layout in favour of a large calorimeter, the solenoid adds to the material of the inner detector and thus causes energy loss of particles before the calorimetric measurement, which must be compensated by calibration schemes.

The momentum measurement of muons, which usually pass through the calorimeters as minimum-ionising particles with little energy loss, is improved with additional tracking detectors in a second magnetic field outside the HCAL. In ATLAS, this is generated by eight additional coils, which are assembled radially around the beam axis in the barrel and each endcap. These create a toroidal field, which is orthogonal both to the beam axis and to the magnetic field in the inner detector. Muon tracks are thus bent in ϕ direction in the solenoidal and in z direction (η direction) by the toroidal field and must therefore be combined in a reconstruction from these two measurements.

While a detailed description of the detector and its performance can be found in [69], a short summary of the main components shall be presented in the following sections.

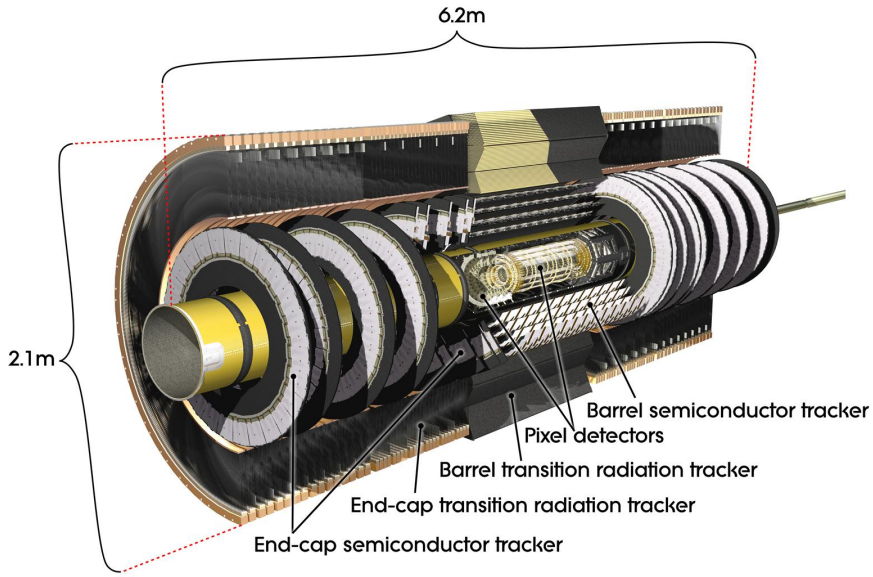


Figure 3.4: Cut-away image of the ATLAS inner detector [70].

3.2.1 The Inner Detector

The ATLAS inner detector (ID) consists of three sub-systems (cf. Figs. 3.4-3.5), which are placed inside a homogeneous magnetic field with a strength of $B = 2$ T:

Closest to the beam pipe, with a distance of only 5-12 cm, three layers of silicon sensors with pixels of the size $r\phi \times z = 50 \times 400\ \mu\text{m}^2$ form the pixel detector, which is designed as a barrel for the central region and as disks for the endcaps, covering the forward regions up to $|\eta| < 2.5$. With approximately 80.4 million readout channels, it can provide up to three position measurements per track with a spatial resolution of $10 \times 115\ \mu\text{m}^2$ (in $r\phi \times z$). Due to their proximity to the interaction point, these can yield a very precise measurement of the primary interaction vertex and allow the reconstruction of secondary decay vertices e.g. of b-hadrons within jets or of τ-leptons, which decay into multiple charged particles.

Enclosing the pixel detector, the *Semi Conductor Tracker* (SCT) consists of a barrel of four double-layers and two endcaps with nine disks of silicon strip detectors. In each double-layer, these strips with a length of 6.4 cm and a pitch of 80 μm are arranged orthogonally in the $r\phi \times z$-plane, so that both the $r\phi$ and the z-coordinate can be measured with high precision, while a small stereo angle of 40 mrad between these two layer components provides an improved measurement of the three dimensional space point on the track. With this technique, typically eight SCT hits per track corresponding to four space points are measured with an intrinsic resolution of $17 \times 580\ \mu\text{m}^2$ (in $r\phi \times z$), which are read out via approximately 6.3 million channels.

Outside the SCT, a barrel and two endcaps of straw tubes with a diameter of 4 mm are arranged along the z direction (r direction in the endcaps) and provide an average of 36 additional measurements of $r\phi$-positions for tracks within $|\eta| < 2.0$. Inside these tubes, a mixture of xenon, carbon dioxide and oxygen is ionised by crossing particles, and the created charges are separated by an applied high voltage. From the constant drift velocity, *drift circles* can be reconstructed, which allow for an intrinsic

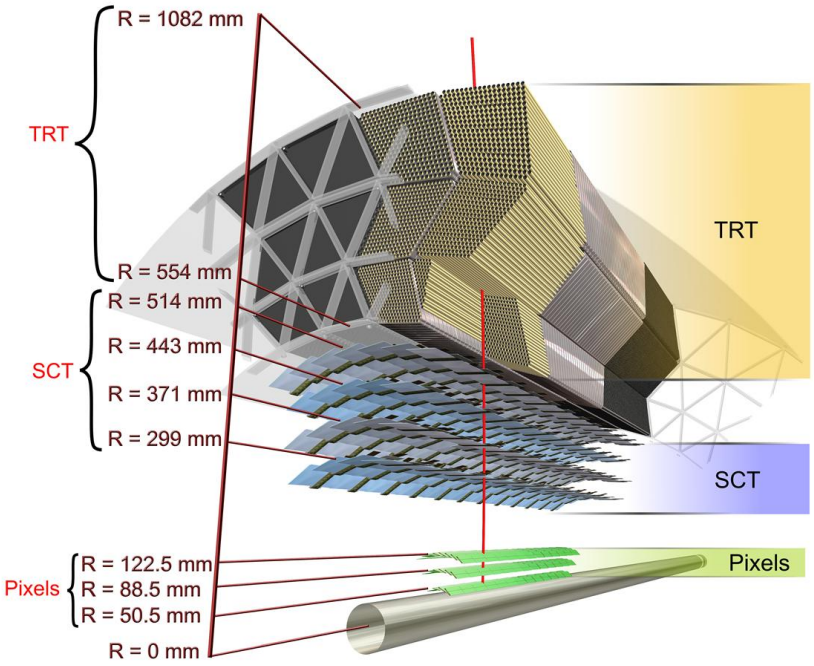


Figure 3.5: Schematic view on the components of the ATLAS inner detector [69].

resolution of 130 μm (in $r\phi$) per straw. Besides the space point measurement of each straw hit, the signal-strength is exploited by the particle-identification: especially electrons crossing the radiator material surrounding the straws create transition-radiation photons, which are registered as *high-threshold* hits and can thus distinguish electrons from other ionising particles such as charged pions. Based on the important information gained by exploiting this effect, this outermost part of the inner detector is named *Transition-Radiation Tracker* (TRT).

3.2.2 The Calorimeter

Since the energy of electromagnetically and strongly interacting particles can only be absorbed with different types of particle showers, the ATLAS calorimeter is designed as a combination of different sampling calorimeters (cf. Fig. 3.6): the inner ECAL with a radiation length of $> 22X_0$ ($> 24X_0$) in the barrel (endcaps) and the outer HCAL with an interaction length of approximately 10λ.

The ECAL consists of a barrel ($|\eta| < 1.375$) and two endcap components ($1.375 < |\eta| < 2.5$ and $2.5 < |\eta| < 3.2$) of lead absorbers, which induce the electromagnetic showers, and liquid argon acting as active material to measure the deposited energy of the particles with a resolution of $\sigma(E)/E = 10\%/\sqrt{E/\,\mathrm{GeV}} \oplus 0.17\%$. An accordion-shape of the absorbers and kapton-electrodes is chosen to avoid azimuthal cracks and to ensure complete ϕ-symmetry, while their arrangement in three layers and an additional pre-sampler ($|\eta| < 1.8$) with segmentations of varying granularity provide a good resolution of the shower shape.

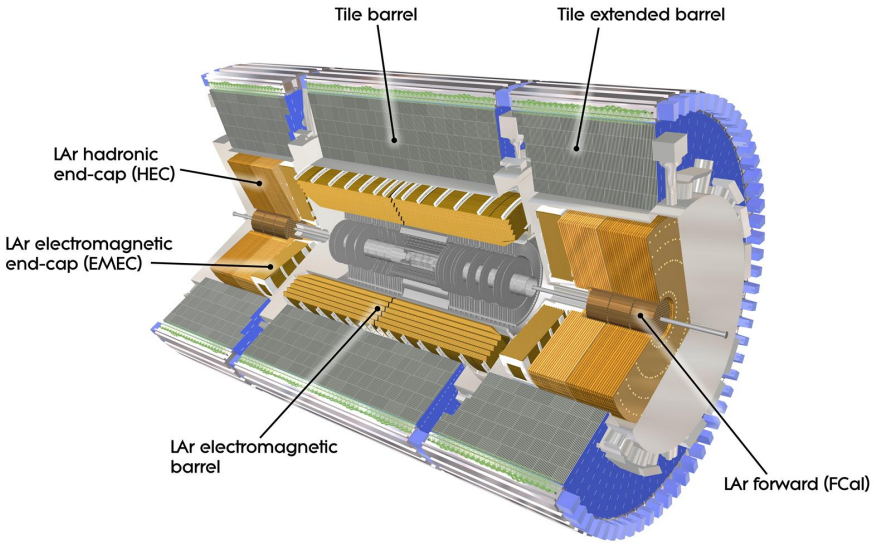


Figure 3.6: The ATLAS calorimeter system [69].

In the barrel region ($|\eta| < 1.5$), the HCAL consists of three layers, which initiate hadronic showers in iron absorbers and measure the corresponding energy with plastic scintillators providing a resolution of $\sigma(E)/E = 50\%/\sqrt{E/\text{ GeV}} \oplus 10\%$. Due to radiation hardness, liquid argon (LAr) is preferred as active material in the endcaps ($1.5 < |\eta| < 3.2$) with high activity, while copper is used as absorber.

With a coverage up to $|\eta| < 3.2$, the calorimeter can measure the energies of particles emitted at angles larger than 4.7 degrees with respect to the beam axis. However, especially the measurement of the missing transverse momentum and of jets originating from the vector-boson-fusion process requires a better coverage of the forward region. This is achieved with the forward calorimeter (FCAL), which reaches up to $3.2 < |\eta| < 4.9$. Situated very close to the beam pipe and exposed to very high radiation doses, LAr is used as active material throughout the FCAL. Its inner layer is dedicated to initiate electromagnetic showers using copper as absorber, while hadronic showers are produced in tungsten material of the outer two layers. In combination, these FCAL components provide energy measurements with a resolution of $\sigma(E)/E = 100\%/\sqrt{E/\text{ GeV}} \oplus 10\%$.

3.2.3 The Muon Spectrometer

As muons usually pass the calorimeters with only a minimal loss of energy and momentum, the muon spectrometer (MS) is the outermost and largest component of the ATLAS detector and dominates its overall size as shown in Fig. 3.7. Within the central range ($|\eta| < 2.0$), it provides space point measurements from *Monitored Drift Tubes* (MDTs), which are substituted by *Cathode Strip Chambers* (CSCs) in the outer region ($2.0<|\eta|<2.7$) working as multi-wire proportional chambers. On average, these allow momentum measurements for muons with a resolution of $\sigma(p_T)/p_T = 10\%/(p_T/\text{ GeV})$. For the purpose of triggering, *Resistive Plate Chambers* (RPCs) and *Thin Gap Chambers* (TGCs) are used to gain fast detection signals in the central and outer region respectively.

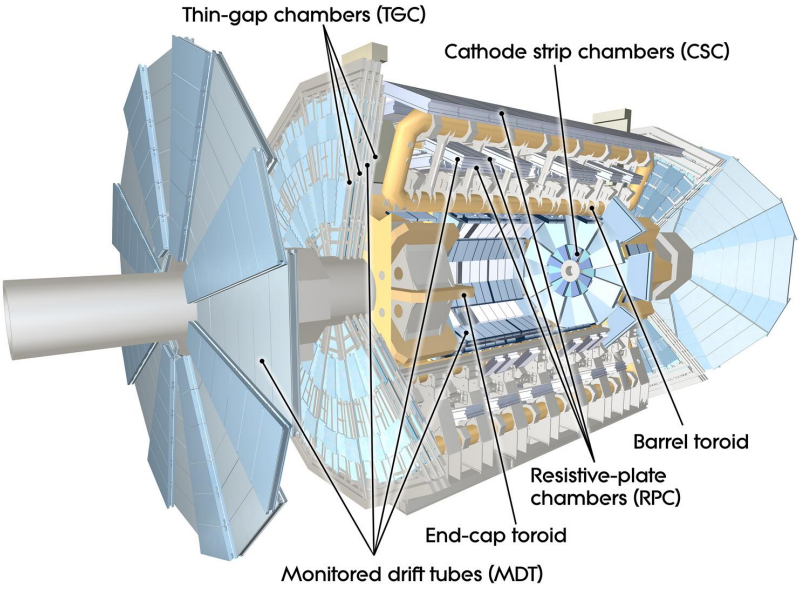


Figure 3.7: The ATLAS muon system [69].

3.3 Detector Simulation

For any physics analysis, events from signal processes must be selected very efficiently from the data and thus require a good understanding of the detector response both for signal and all background processes with similar signatures. Although certain background processes can eventually be estimated with data-driven methods by applying dedicated selections and extrapolation methods, predominantly the signal and irreducible background contributions are conceptually impossible to be determined from data only. Therefore, an essential part of high energy physics analysis consists of the event simulation, which allows the creation of large event samples for almost any physics process on the basis of probability calculations and random numbers. While the size of these samples and thus the statistical uncertainties are mainly limited by computing time and disk space, especially the precision and accuracy of the underlying calculations as well as the quality of the modelled detector response have to be taken into account as systematic uncertainties on predictions based on such simulations.

Generated events obtained via the Monte Carlo technique as described in Sec. 2.7 are hence passed to the detector simulation, which is based on a GEANT4 [71] model of all components and support materials of the ATLAS detector [72]. For each event, the simulation software then computes the interaction of all generated particles with the detector material and the corresponding response of each detector component as expected from a similar collision event in the real detector. Besides noise effects in the detector electronics, which are easily generated by random numbers, also the simulation of pile-up effects must be included in this step. The generated hard scattering event is thus overlaid with a given number of pp-collisions occurring at the same time (the so-called *in-time pile-up*), which are usually multijet events with relatively low transverse momenta of the jets. The corresponding number of simultaneous interactions is randomly drawn from a distribution, which is modelled according to the conditions during data taking, while further out-of-time effects on electronic signals in detector components are included by accounting for the 50 (150) ns timing between the bunch crossings. Since the underlying distribution of simultaneous interactions cannot correspond perfectly between the simulated events and the varying conditions of all data taking runs, the Monte Carlo events are eventually

re-weighted, so that their pile-up effects can match the expected effects in the data more accurately.

3.4 Reconstruction and Performance

In order to convert the detector response to a collision event into physics objects, which can be used in a search for a specific process, all measured silicon detector hits, energy depositions in the calorimeter cells and muon track segments must be converted into tracks and calorimeter clusters. With dedicated reconstruction algorithms, these low-level objects are then assigned to different particle types or jets. While these basic particles can still overlap and thus double-count certain energy depositions, identification algorithms assess the quality of the object, based on which the analysis particles are uniquely defined with an overlap removal. Finally, the vector sum of all reconstructed objects and all remaining cells un-assigned to any object is calculated to reconstruct the *missing transverse energy* $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$. The naming of this observable is chosen to refer to the underlying measurement of energy depositions, for which directions can be assigned via the detector geometry. Their vector sum $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ yields an approximation of the physics observable *missing transverse momentum* $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}} \approx \vec{E}_{\mathrm{T}}^{\mathrm{miss}}$, which usually represents the vector sum of the neutrino momenta.

3.4.1 Data Reduction with Trigger Algorithms

Due to the large amount of data per event and the high interaction rate of the LHC (one event per 50 ns), the processing and storage of all data events is technically impossible. Instead, the amount of data must be reduced by selecting only the "interesting" collision events out of the large number of multijet events produced in pp-collisions. Therefore, events are only recorded by ATLAS and stored for further analysis if they are found to contain relevant features such as electrons or muons, hadronic τ decays or other special jet-like objects with sufficiently high momentum, or a large imbalance of measured energies contradicting the conservation of transverse momentum. Each event is analysed for such features with dedicated and very fast *trigger* algorithms and only selected for further processing with more sophisticated but slower reconstruction algorithms if an "interesting" feature is found by any trigger. Starting from a collision rate of 20 MHz, the event rate is consecutively reduced over three trigger stages to approximately 100 Hz, at which the large amount of data per event can be stored. These trigger algorithms are optimised for speed and high efficiency and are described in detail in [70, 73]. During the data taking periods, especially the transverse momentum thresholds of the accepted trigger objects must be adapted to the number of simultaneous interactions, so that the event rate is kept sufficiently low to ensure the further processing of the data.

3.4.2 Basic Object Reconstruction

Starting from the large amount of electronic signals, which are recorded with all the detector components for a single collision event, all knowledge about the particle interaction in the detector can be used to reduce the information to a limited number of objects, which can be handled in a physics analysis later on. The first step is the reconstruction of tracks from the large number of hits in the inner detector. These can later be associated to muon segments as well as clustered energy depositions in the calorimeter.

Building Tracks out of Silicon Hits

The reconstruction of tracks in ATLAS is seeded by clusters of hits in the pixel detector and the first layer of the SCT. The obtained directions of these seed tracks are then extrapolated throughout the

SCT, so that complete track candidates can be found. These are required to have a sufficiently high momentum of $p_T > 0.5$ GeV to traverse all layers of the SCT when bent in the homogeneous magnetic field. The clusters assigned to the candidates are then fitted, while outlier clusters are removed. Clusters originally included by more than one track are either assigned uniquely or considered as shared between two tracks. In addition to requirements on the fit quality, the number of hits of a candidate is exploited to reject fake tracks originating from noise. The remaining tracks with good quality are then matched to calibrated drift circles in the TRT and re-fitted with the full tracking information. TRT drift circles, which lead to significantly worse fit results, are considered as outliers, while the well matching TRT hits provide further information for the particle identification, enabling to distinguish especially electrons based on their characteristic high threshold hits. The corresponding performance of the track and vertex reconstruction in ATLAS is documented in detail in [74].

Clustering of Calorimeter Cells

Due to the high granularity of the ATLAS calorimeters, most particles produced in the collision event deposit their energy in several calorimeter cells. In order to collect the energy from each particle, these cells can be grouped by two different clustering algorithms [75]:

The **sliding-window** algorithm starts from so-called calorimeter *towers*, which are groups of adjacent cells towering in radial direction. These towers thus contain the full longitudinal shower, while they provide a segmentation in the (η, ϕ)-space. Within a window of pre-defined size (e.g. 3×7 towers in units of 0.025×0.025 in $\eta \times \phi$), the sum of all deposited energy is calculated and compared to the energy sum obtained by sliding this window over the neighbouring towers. If this sum of transverse energy reaches a certain threshold (e.g. $E_T > 2.5$ GeV) and if the window with the maximum energy is found to be a local maximum by sliding over all towers in the neighbouring (η, ϕ)-space, a cluster is built by grouping the towers of the corresponding window.

Instead of building towers, the **topological clustering** starts from the individual cells in the calorimeter. Since each cell is subject to electronic noise, only those above a certain signal-to-noise threshold t_{s} are considered as seeds for the clustering. Neighbouring cells are then iteratively and exclusively added to the seeds, which are ordered according to their signal-to-noise ratio: neighbouring cells above a medium threshold t_{n} are used to extend the cluster, while all adjacent cells of the extended cluster are added if their signal-to-noise ratio exceeds a low threshold t_{c}. With different threshold values (t_s, t_n, t_c) for the hadron $(4, 2, 0)$ and electromagnetic calorimeter $(6, 3, 3)$, the constructed clusters then serve as inputs for the separate particle reconstruction algorithms defining the high level objects.

3.4.3 High-Level Physics Objects

Starting from cells and tracks, candidates for high-level physics objects are reconstructed with dedicated algorithms [70]. While the corresponding requirements for a reconstruction of these candidates are very loose, dedicated identification criteria are subsequently applied in order to assess the object quality and to reject mis-identified objects, which are often referred to as *fake* objects. Depending on the physics analysis, different quality working points with given identification efficiency and fake rejection can be required to chose the optimal object definitions for an analysis.

Muon Reconstruction and Identification

In the final states produced by pp-collisions, muons can be created at different stages. Especially hadronic showers can contain e.g. $\pi^{\pm} \rightarrow \mu^{\pm}\nu$ decays in the calorimeters, so that the muons can only be observed in the outer part of the detector. Muons from $\tau^{-} \rightarrow \mu^{-}\bar{\nu}_{\mu}\nu_{\tau}$ decays instead occur within a radius of a few millimetres around the interaction point and create a track both in the inner detector (ID) and in the muon spectrometer (MS), similar to prompt muons directly originating from the very short-lived Z^0 or $W^{\pm}$ bosons. From the various existing reconstruction algorithms within the ATLAS reconstruction software, muons are usually obtained with the *Staco* (statistical combination) algorithm, which requires a combination of two well matching tracks in both detector components. Due to the limited acceptance of the track reconstruction, the combined muons are restricted to the range of $|\eta| < 2.5$. Additionally, supporting structures for the calorimeter lead to significant inefficiencies in the central barrel region around $|\eta| \approx 0$. While combined muons from the *Muid* (μ-identification) algorithm are reconstructed by a re-fit of the complete track from hit-level, *Staco combined* muons are simply built as the statistical combination

$$\mathbf{P}_{\text{comb.}} = \left(\mathbf{C}_{\text{ID}}^{-1} + \mathbf{C}_{\text{MS}}^{-1}\right)^{-1} \left(\mathbf{C}_{\text{ID}}^{-1}\mathbf{P}_{\text{ID}} + \mathbf{C}_{\text{MS}}^{-1}\mathbf{P}_{\text{MS}}\right) .$$

Here, the separate tracks are denoted by the parameter vectors $\mathbf{P}_{\text{ID}}$ and $\mathbf{P}_{\text{MS}}$ with their corresponding co-variance matrices $\mathbf{C}_{\text{ID}}$ and $\mathbf{C}_{\text{MS}}$. The quality of their match is evaluated with a χ^2-test, which is constructed via

$$\chi^2_{\text{match}} = \left(\mathbf{T}_{\text{ID}}^{-1} - \mathbf{T}_{\text{MS}}^{-1}\right)^{T} \left(\mathbf{C}_{\text{ID}}^{-1} + \mathbf{C}_{\text{MS}}^{-1}\right)^{-1} \left(\mathbf{T}_{\text{ID}}^{-1} - \mathbf{T}_{\text{MS}}^{-1}\right) .$$

Besides quality criteria on the matching, further requirements on the hits in the inner detector are imposed in order to ensure a precise measurement of the momentum as well as to reject fake muon candidates, which do not originate from the primary interaction. The resulting performance of the muon reconstruction and identification in ATLAS over the data taking period is documented in [76, 77].

Electron Reconstruction and Identification

Starting from sliding-window clusters of the size 3×5 in the electromagnetic calorimeter, electron candidates are reconstructed within a range of $|\eta| < 2.47$ if the cluster energy is sufficiently large ($E_T > 3$ GeV) and a matching track is found in the inner detector within $\Delta\eta < 0.05$ and $-0.05 < \Delta\phi < 0.10$[2]. In case of multiple tracks fulfilling the matching criteria, the closest in ΔR is associated to the cluster and considered as the electron track. For each seed candidate, the corresponding cluster is afterwards rebuilt with windows of 3×7 (5×5) in the electromagnetic barrel (endcaps). Its energy is then corrected according to the expected energy loss in the inner detector material, its deposited energy in neighbouring towers and its leakage into the adjacent hadron calorimeter. In order to reject problematic candidates from the calorimeter transition regions between the barrel and the endcaps, electrons within $1.37 < |\eta| < 1.52$ are usually excluded from physics analyses.

The reconstructed candidates are then assessed with different quality criteria, such as the hadronic leakage fraction, the lateral shower width, track matching quality and impact parameter with respect to the reconstructed interaction vertex as well as required hits in the inner detector including high-threshold TRT hits. The different criteria are grouped into three classes `loose`, `medium` and `tight`,

[2] Due to the bending of the track in the magnetic field, a direction of $\Delta\phi$ can be defined. In order to account for possible bremsstrahlung loss, the larger threshold in the positive $\Delta\phi$ direction allows for a matching of tracks, which lose momentum and are thus subject to a stronger bending as compared to the direction of flight of the emitted bremsstrahlung photon. This direction must be opposite for e^+ and e^- candidates.

which are documented in detail in [78], including measurements of the corresponding reconstruction and identification efficiencies.

Jet Reconstruction Algorithms and Calibration

Due to the nature of the strong interaction, highly energetic quarks and gluons from hard scattering processes produce collimated particle showers, which dominantly consist of hadrons. In order to reconstruct their energy and direction, all corresponding energy depositions in the calorimeter must be collected and assigned to different jets, where especially the separation of two jets with small distance must be ensured. The anti-k_t algorithm [79, 80] works as the inverse of a parton shower by re-combining particles instead of splitting them. It orders all momentum weighted distances

$$d_{ij} = \min\left\{p_T(i)^{-2}, p_T(j)^{-2}\right\} \frac{\Delta R(i,j)^2}{D^2}$$

for all combinations of two particles i and j together with all distances $d_{iB} = p_T(i)^{-2}$ of each single particle to the beam. As long as the smallest of these values is a two-particle distance, these particles are combined into a new object, and all distances are recalculated. After iterating this procedure and combining more and more particles, an object k is eventually identified as jet, if its value d_{kB} is smaller than any d_{ij} in the ordered list of distances. The size of this jet is then determined by the parameter D, since particles are only grouped together if

$$\min\left\{p_T(i)^{-2}, p_T(j)^{-2}\right\} \frac{\Delta R(i,j)^2}{D^2} < \min\left\{p_T(i)^{-2}, p_T(j)^{-2}\right\} \quad \Rightarrow \quad \Delta R(i,j)^2 < D^2 \,.$$

While this algorithm is motivated on the level of particles in the parton shower, the reconstruction of jets in the detector has to rely on energy depositions in the calorimeter, which are clustered into particle-like objects. In the ATLAS reconstruction, the anti-k_t algorithm is therefore applied to the topological clusters of calorimeter cells, and so-called *AntiKt4Topo* jets are built with a size-parameter of $D = 0.4$. Although the corresponding jet energy and direction is automatically provided by the algorithm as the sum of the combined clusters, this measured hadronic shower energy cannot contain the complete energy of the initial parton from the hard scattering process. Instead, neutrinos produced in pion or neutron decays usually escape the measurement, so that the directly reconstructed energy is systematically lower than the true value as obtained e.g. on particle level in simulated events. Therefore, a local calibration scheme [81] for the separate clusters as well as the jet energy scale (JES) for the reconstructed jets is applied to correct for energy losses and calibrate the energy measurement [82]. This calibration is largely performed with di-jet events, which are expected to have balanced in transverse momenta. While this performs especially well in the central η region of the detector, the measured energies of jets in the forward regions ($|\eta| > 3.2$) are corrected with an *η-intercalibration* scheme [83, 84].

With an increasing number of simultaneous pp-interactions per bunch crossing, jets cannot only be produced in the primary interaction, but can also originate from secondary collisions at lower parton energies. Contamination from these pile-up effects to the reconstructed jets are identified by analysing all tracks inside each jet and calculating the *jet-vertex fraction* (JVF) as the scalar sum of track transverse momenta associated to the primary vertex over the total scalar sum of track transverse momenta inside the jet. A cut on this jet-vertex fraction is then applied as quality criterion for jets inside the acceptance of the tracker $|\eta| < 2.4$, which ensures that jets with dominant energy contributions from pile-up are rejected. Outside this acceptance, jets are restricted to a region of $|\eta| < 4.5$ in order to ensure that all

reconstructed jets are completely contained in the forward calorimeter and do not leak a substantial fraction of their energy into the uncovered region close to the beam line.

Identification of b-Hadron Decays Inside Jets

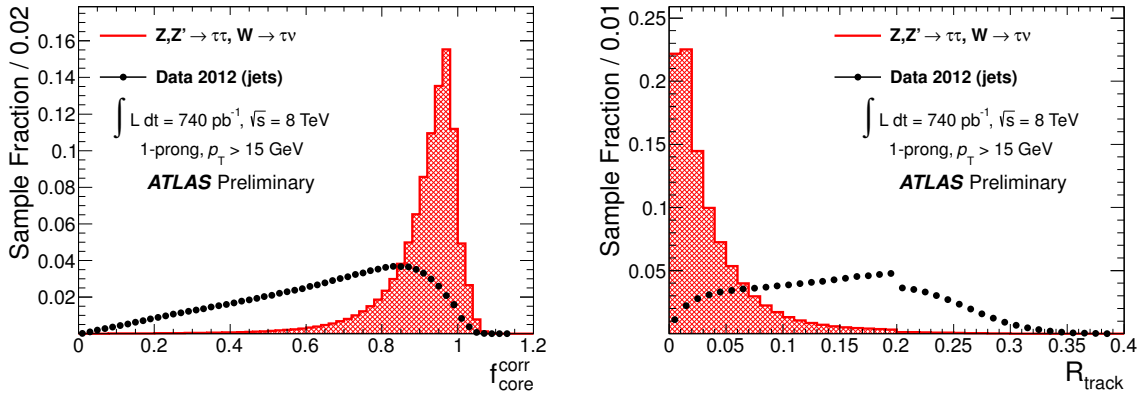
During the hadronisation process, b-quarks are confined in bound states, which are known as b-hadrons and have a relatively long lifetime of ≈ 1.5 ps. They can thus typically travel a few millimetres before they decay into lighter hadrons and in some cases also into leptons. Since their hadronisation always requires the interaction with other quarks, these decays mostly occur within jets of other light hadrons and are therefore reconstructed with the standard jet algorithm. The additional decay signature of the b-hadron then allows the identification of a *b-jet* via a secondary vertex, which can be reconstructed from the tracks originating from the in-flight decay of the long-lived hadron [85]. An alternative algorithm to provide such a *b-tagging* does not reconstruct this vertex, but provides an identification of b-jets via the association of their tracks to the primary vertex, which is expected to perform less well for the tracks originating from the b-hadron decay [86].

Identification of Hadronic Decays of τ-Leptons

Hadronic τ decays produce a variety of different final states, which in most cases consist of an odd number of charged pions and a low number of neutral pions. While neutral pions decay immediately into $\gamma\gamma$ pairs and thus are observed only in the electromagnetic calorimeter, each hadronic τ decay contains at least one charged pion, which typically creates a track in the inner detector and a particle shower reaching into the hadron calorimeter. Since single particles cannot be resolved very well in the calorimeter, all these different final states $\tau_{\text{had}} = \tau^{\pm} \rightarrow (2n+1)\pi^{\pm} + m\pi^0 + \nu$ (or with Kaons $K^{\pm,0}$ instead of $\pi^{\pm,0}$) are grouped into the class of visible τ_{h} objects. These are usually divided into *1-prong* and *3-prong* τ_{h} candidates according to the number of charged decay products creating tracks. In the detector, any τ_{h} object resembles a jet, so that their reconstruction starts from the default *AntiKt4Topo* jets as seeds. These seeds are required to have a sufficiently large transverse momentum of $p_T > 10$ GeV and to be found within the acceptance of the track reconstruction $|\eta| < 2.5$. For each such candidate, a dedicated vertex association (TJVA) [87] is performed, since the τ_{h} candidate is not necessarily originating from the primary interaction point. The corresponding topological clusters of the seed are considered as τ decay particles (pions), and the τ_{h} direction is re-calculated as the barycenter of the corresponding cluster four-vectors[3]. Since the particle composition of these τ_{h} candidates is very different from the average jet, which includes a larger fraction of hadronic energy and especially neutrons, the measured energy of the τ_{h} candidate is then calibrated with a dedicated τ-energy-scale (TES) [88]. Tracks within an inner cone of $\Delta R \leq 0.2$, which fulfil $p_T > 1$ GeV and quality criteria on hits and impact parameters, are then associated to the τ_{h} object. These provide the distinction between 1-prong and 3-prong τ_{h} candidates, while other track multiplicities are rejected to suppress fake candidates from parton-induced jets or other objects.

For each reconstructed τ_{h} candidate, a set of discrimination variables is calculated, which are then used for the identification of real hadronic τ decays and distinguish them from the large number of fake candidates. Since these fake τ_{h} objects are mostly jets, which usually produce wider showers with more charged particles and a larger hadronic fraction, especially shower shape variables and track properties can provide a good separation between hadronic τ decays and jets from other QCD processes. During the first data taking period in 2010, only three such variables were used for a cut-based identification

[3] The masses of the clusters and thus of the corresponding decay particles are neglected as well as the mass of the visible hadronic τ_{h} object.



(a) (b)

Figure 3.8: Normalised distributions of the jet discriminating variables core fraction f_{core} (a) and track radius R_{track} (b) for τ_{h} candidates in simulated Z, $Z' \to \tau\tau$ and $W \to \tau\nu_\tau$ signal samples and in a jet background sample selected from 2012 data [90]. The merging of candidates with and without tracks in the outer cone ($0.2 < \Delta R \leq 0.4$) creates the discontinuity at $\Delta R_{\text{track}} = 0.2$.

parametrised in $E_T(\tau_{\text{h}})$ [89]: the energy weighted cell radius R_{EM} in the electromagnetic calorimeter, the momentum weighted track radius R_{tracks} and the fraction of momentum carried by the leading track $p_T^{\text{lead. track}}/E_T$. Their dependence on $E_T(\tau_{\text{h}})$ takes into account that τ decays are more collimated with a higher boost [89].

Due to the increasing cell energy contribution from pile-up effects, the identification method has been developed over the later periods by reducing the calorimeter radius for all calculations from $\Delta R \leq 0.4$ to $\Delta R \leq 0.2$ by considering a larger set of discrimination variables [90]: the core fraction f_{core} of energy in an inner cone $\Delta R < 0.1$ as compared to $\Delta R < 0.2$ (cf. Fig. 3.8a), the fraction of momentum carried by the tracks within $\Delta R < 0.2$ as compared to $\Delta R < 0.4$ and the momentum weighted track radius R_{track} (cf. Fig. 3.8b) are used both for 1-prong and 3-prong candidates. Further variables such as the number of tracks in the isolation annulus $0.2 < \Delta R \leq 0.4$ for 1-prong τ_{h} or the significance of the transverse flight path of the τ between its production in the associated primary vertex and its reconstructed decay vertex in 3-prong candidates are used for the separation of jets and hadronic τ decays. For the identification, these variables are combined via boosted decision trees (BDTs, cf. Sec. 7.1), which provide a single discriminant quantifying the probability that the reconstructed object is a real hadronic τ decay. On the obtained output distribution, three different cut values are chosen to define the working points `loose`, `medium` and `tight` with corresponding target efficiencies of 70%, 60% and 40% (65%, 55% and 35%) for 1-prong (3-prong) candidates.

While the dominant contribution of fake-τ_{h} candidates originates from jets, also electrons can be reconstructed as τ_{h} candidates and in most cases fulfil all identification criteria for 1-prong candidates. Therefore, a dedicated electron veto is implemented, which provides an additional separation between hadronic τ decays and electrons. It relies on several shower-shape related variables such as the fraction of electromagnetic energy in the candidate as well as the fraction of high-threshold TRT hits in the associated tracks.

Overlap Removal

All described reconstruction and identification algorithms work independently in the ATLAS software. As a result, it cannot be excluded that the same particle or jet is successfully reconstructed and identified by more than one algorithm and is thus considered as two different objects. Hence, a geometrical overlap removal is applied for close-by objects within $\Delta R < 0.2$, which are thus assumed to be identified more than once. Motivated by the interactions of the different particle types in the detector, the following hierarchy of object types is employed:

- Combined muons originate from within the inner detector and can thus deposit a small fraction of their energy in the calorimeters. Overlapping electrons, jets or τ_{h} candidates can potentially be reconstructed from these clusters and are removed for this reason.
- Calorimeter clusters from electrons are likely to be reconstructed as jets and τ_{h} candidates, so that these objects are removed when overlapping with an identified electron.
- The reconstruction of hadronic τ decays is seeded by jets, so that by construction every τ_{h} object is also a jet. All jets overlapping with identified τ_{h} candidates are therefore removed.

Missing Transverse Momentum

In order to reconstruct the missing transverse momentum in a collision event, the momenta of all particles must be summed. Since a momentum measurement from tracks can only be obtained for charged particles, an energy balance in the calorimeter is considered instead of the corresponding momentum balance. The energy is therefore summed over all cells, while the directions of the cell vectors are provided by the cell positions. Due to the different energy scales, each cell contribution associated to an object must enter the sum with a calibration correction, so that e.g. the missing energy contribution from hadronic showers is compensated. As in the overlap removal, the cells must be uniquely associated to the corresponding physics objects in the event [91, 92]. Especially for τ_{h} objects, this association via the τ_{h} identification is of great importance, since the energy calibration of the associated cells depends on whether it is considered as a τ_{h} candidate or as a jet. All cells, which are not matched to any reconstructed high-level object, enter the energy balance with a pile-up correction [93]: the unmatched energy depositions are corrected by the soft term vertex fraction, which is calculated from the scalar sum of unmatched track momenta and denotes its fraction associated to the primary vertex. With all described corrections, the missing transverse momentum $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ corresponds to the vectorial sum over all momenta of particles, which are absorbed in the calorimeter. The momenta of all muons measured in the spectrometer outside the calorimeter are then subtracted, so that an equivalent of the sum over all neutrinos is provided. Fake missing momentum in an event is thus reduced to a minimum and can only be created by particles escaping the calorimeter at high $|\eta|$ very close to the beam pipe or by resolution effects e.g. from the energy measurements of jets.

CHAPTER 4

Search Strategies for $H \to \tau_{\text{lep}}\tau_{\text{had}}$ Final States

While the ATLAS collaboration studies all three decay channels of $H \to \tau\tau$ events ($\tau_{\text{lep}}\tau_{\text{lep}}$, $\tau_{\text{lep}}\tau_{\text{had}}$, $\tau_{\text{had}}\tau_{\text{had}}$), all analyses described in the following focus on the $\tau_{\text{lep}}\tau_{\text{had}}$ final states. These require a sufficiently clean selection of the $\tau_{\text{lep}}\tau_{\text{had}}$ signal events, but must reject contributions from all other processes resulting in a lepton $\ell = e, \mu$ and a hadronic object. These processes are described, and strategies are discussed in the following to estimate the remaining contributions to the $H \to \tau_{\text{lep}}\tau_{\text{had}}$ signal selection. In order to obtain final results from the corresponding signal and background expectations and observed collision data, a statistical analysis must be performed, which is explained here in the context of the profile likelihood method [94].

4.1 Properties of the $H \to \tau_{\text{lep}}\tau_{\text{had}}$ Signal

Considering the general features of the signal process $H \to \tau_{\text{lep}}\tau_{\text{had}}$, the selection of such events must require a light lepton $\ell = e, \mu$ and a hadronic τ decay with opposite charge. In absence of highly energetic jet radiation, the Higgs boson is produced with no significant transverse momentum, so that the τ-leptons as its decay products are aligned in a back-to-back topology with $\Delta\phi(\tau_{\text{lep}}, \tau_{\text{had}}) \approx \pi$. Since these τ-leptons are highly relativistic ($m_H/2 \gg m_\tau$), their subsequent decay products are collimated, and thus both the light lepton ℓ and the hadronic τ decay (τ_{h}) follow the direction of their mother particles ($\Delta\phi(\ell, \tau_{\text{h}}) \approx \pi$). Similarly, the undetectable neutrinos follow the initial directions of the τ-leptons, keeping their back-to-back topology $\Delta\phi(\nu_\ell\bar{\nu}_\tau, \nu_\tau) \approx \pi$ as depicted in Fig. 4.1a. Their contributions to the missing transverse momentum thus balance to a large extent, so that the full reconstruction of this event and the invariant mass becomes almost impossible. However, as the leptonic τ decay is accompanied by two neutrinos, its average fraction of invisible momentum is larger than for the hadronic counterpart. This slight bias favours small angular separation $\Delta\phi(\ell, p_{\text{T}}^{\text{miss}}) < \Delta\phi(\tau_{\text{h}}, p_{\text{T}}^{\text{miss}})$ leading to small transverse masses $m_T(\ell)$.

4.1.1 H+jets Final States

When recoiling in the transverse direction against additionally produced jets from initial-state-radiation, the topology of the Higgs boson decay can change, so that the τ-leptons as well as their visible decay products enclose an azimuthal angle $\Delta\phi(\tau_{\text{lep}}, \tau_{\text{had}}) \approx \Delta\phi(\ell, \tau_{\text{h}}) < \pi$. In such a topology, the invisible

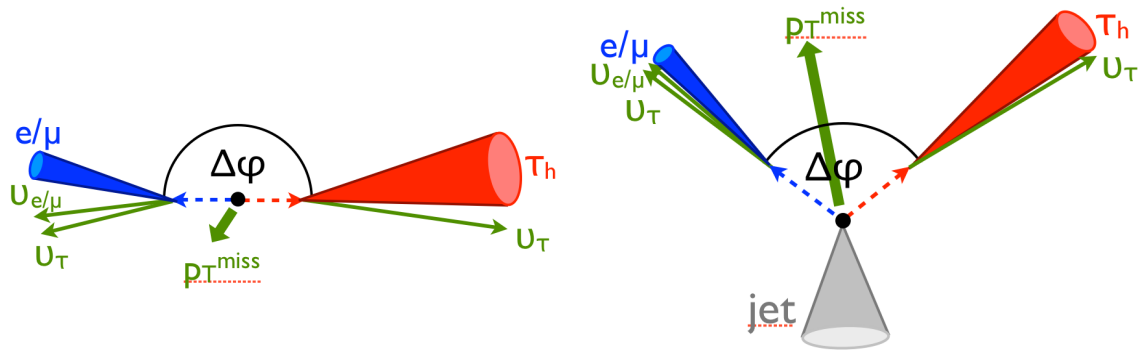


Figure 4.1: Schematic view on the event topology of $H \to \tau_{\text{lep}}\tau_{\text{had}}$ decays (a) in a back-to-back topology with $\Delta\phi(\ell, \tau_{\text{h}}) \approx \pi$ and (b) in a boosted topology with a recoiling jet and $\Delta\phi(\ell, \tau_{\text{h}}) < \pi$.

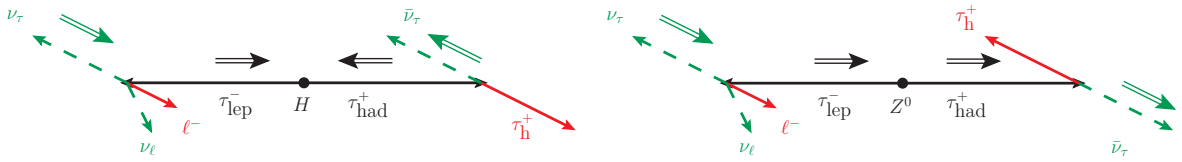
momenta of the neutrinos cannot balance, but instead add up to a significant missing transverse momentum with a direction that must lie between the visible decay products (cf. Fig. 4.1b).

In the special case of the VBF production, a transverse recoil can also be caused by the two quark jets. However, since their production is less affected by soft QCD radiation, the transverse momentum of the qqH system should vanish if no additional jet is produced in the interaction. The two VBF-jets are well separated in the η direction, and the invariant mass $m(j_1, j_2)$ of this di-jet system must thus be large. While the whole qqH system can be boosted parallel to the beam axis, in most cases the Higgs decay products will lie between the jets.

4.2 Background Processes for $H \to \tau_{\text{lep}}\tau_{\text{had}}$ Final States

4.2.1 Irreducible Background: $Z \to \tau\tau$

With a similar signature to $H \to \tau\tau$ but much higher in rate, the $Z \to \tau\tau$ process yields the dominant background contribution, which is irreducible and can mainly be distinguished by reconstructing the invariant mass of the $\tau\tau$ system. Although this mass is expected to show a significant difference between $m_Z = 91.2$ GeV ($\Gamma_Z = 2.5$ GeV) and the expected $m_H \approx 125$ GeV ($\Gamma_H < 0.01$ GeV), it can only be obtained from the correct vector sum of all decay products. Apart from experimental resolution of all measured kinematics, especially the undetectable neutrinos affect this reconstruction severely, since they smear out the kinematic differences in the τ-lepton momenta, leading to almost indistinguishable distributions of visible transverse momenta for ℓ or τ_{h}. Furthermore, the missing momentum can only reflect the vector sum of the neutrino momenta, which causes a significant overlap between H and Z in their reconstructable mass spectrum.



(a) $H \to \tau\tau$ (b) $Z \to \tau\tau$

Figure 4.2: Example for spin configurations of $\tau\tau$ pairs originating from the decay of a Higgs boson (a) and a Z boson (b). For simplicity, the subsequent τ decays are sketched in the rest frame of the corresponding τ-lepton. In a collision event, these obtain large transverse momenta $p_T(\tau) > m_\tau$, so that their decay products are strongly boosted and obtain smaller or larger transverse momenta according to their direction of flight relative to the boost.

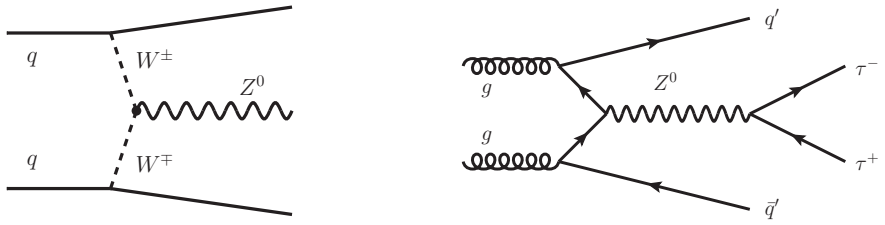
Spin Effects

Besides the different masses of the Z^0 and the Higgs boson, their spins give rise to small differences in the differential momentum distributions. For a $\tau\tau$ pair from a Higgs decay with spin-0, both τ-leptons must have identical helicity[1] (cf. Fig. 4.2a), while their helicities must be opposite in $Z \to \tau\tau$ decays with spin-1 (cf. Fig. 4.2b). The neutrino ν_τ and anti-neutrino $\bar{\nu}_\tau$ from their sub-sequent decays, however, must have opposite helicity due to the opposite chirality of ν_τ and $\bar{\nu}_\tau$. For a $Z \to \tau\tau$ decay, the τ- and ν-helicity are thus either identical or opposite for both τ decays. Instead, for the decay of a Higgs boson the helicity must always be identical for τ and ν in one of the sub-sequent decays and opposite for the other. Since the τ-leptons are strongly boosted in their direction of flight transverse to the beam axis, the case of opposite helicity leads to a low transverse momentum of the neutrino in the laboratory system and a large momentum of the visible decay products. The case of identical helicities rather leads to a large neutrino momentum and thus lower visible momenta. Therefore, the spin correlation of the two τ decays differs between Higgs and Z^0 boson decays, which manifests itself in different correlations of the visible momenta. For Higgs boson decays, the momentum difference $|p_T(\tau_h) - p_T(\ell)|$ of the visible decay products tends to be larger, whereas Z^0 decays produce configurations, where these transverse momenta are either both large or both small and hence tend towards smaller differences. In a similar way, this correlation can be exploited via the ratio $p_T(\tau_h)/p_T(\ell)$. In addition, a more sophisticated analysis of the missing transverse momentum p_T^{miss} can reconstruct the visible momentum fraction of each τ decay in an event and thus provide similar observables and correlations, which are also sensitive to these spin correlation effects.

Vector-Boson-Fusion Topology

As a further and more powerful suppression of $Z \to \tau\tau$ background, the VBF topology can be exploited by selecting two jets with large separation $\Delta\eta$, which essentially increases the ratio of Higgs boson to Z^0 boson events [95, 96]. While the cross section of the dominant $gg \to H \to \tau\tau$ production ($\sigma_{ggH} \approx 1.2$ pb) is almost three orders of magnitude smaller than for $Z \to \tau\tau$ ($\sigma_Z \approx 1.2 \times 10^3$ pb), the ratio is almost equal when considering only the qqH process ($\sigma_{qqH} \approx 0.1$ pb) and the corresponding $qqZ \to \tau\tau$ ($\sigma_{qqZ} \approx 0.1$ pb [97]) as shown in Fig. 4.3a. The total $Z \to \tau\tau$ background for a VBF selection, however, is still dominated by the contribution from the standard gluon- or quark-induced Z +2-jets production

[1] The helicity is defined as the projection of the spin direction on the flight direction of a particle: $h = \frac{\vec{\sigma}\cdot\vec{p}}{|\sigma p|}$.



(a) (b)

Figure 4.3: Example processes for (a) the vector boson fusion production $qq \to qqZ$ and (b) the gluon-induced $gg \to qqZ$ production of Z +2-jets.

processes similar to Fig. 4.3b ($\sigma_Z \times \epsilon_{\text{VBF}} \approx$ 3-11 pb)[2]. Nonetheless, the selection of a VBF topology can lead to a significantly cleaner selection of qqH events[3] decaying to $\tau\tau$ final states as compared to event selections optimised for events from ggH production.

4.2.2 Mis-Identified τ_h Objects

Considering the possibility that any hadronic object can incorrectly be identified as a τ_{h} object, the simple requirement of oppositely charged (ℓ, τ_{h}) pairs can still accept processes with ℓ+jet final states. Furthermore, also the lighter leptons $\ell = e, \mu$ can either be missed completely by the detector acceptance, or they can also be mis-identified as τ_{h} objects. This variety of backgrounds can be classified by their underlying physics processes and their corresponding mis-identification mechanisms. Since they are largely reducible, their composition is not only dominated by production cross sections, but to a large extent by the performance of the experimental object identification. Further considerations on their topology shall motivate the definition of alternative selections, which can be used to control these background contributions or even provide estimates from the recorded collision data.

Multijet Production

Although the requirement of a light lepton ℓ with large transverse momenta can strongly reduce the contribution from multijet events, the large cross section of this process can still produce a substantial number of background events, where the light lepton can originate from a hadron decay inside a jet. In order to suppress such contributions from jets, the leptons are usually required to be isolated in the inner detector and in the calorimeter. At the same time, the selected τ_{h} decay must be a mis-identified jet, which in most cases is gluon-induced ($gg \to gg$, $qg \to qg$). Since these gluon-induced jets usually have no electric net charge, also the mis-identified τ_{h} objects must be neutral when averaging over all events, so that the charge correlation of this background is very weak. Hence, it can be assumed that the number n_{OS} of $(\ell^{\pm}, \tau_{\text{h}}^{\mp})$ pairs with opposite-sign charges ($q_\ell \cdot q_{\tau_{\text{h}}} < 0$) is of the same order as n_{SS} of $(\ell^{\pm}, \tau_{\text{h}}^{\pm})$ pairs with same-sign charges ($q_\ell \cdot q_{\tau_{\text{h}}} > 0$) and can thus be estimated by comparing opposite-sign and same-sign selections. Certain multijet events from other production mechanisms ($qq^{(\prime)} \to qq^{(\prime)}$), however, potentially bias this ratio $r_{\text{QCD}} = n_{\text{OS}}/n_{\text{SS}} \geq 1$, which therefore should be determined from the data. Alternatively, the number of background events of this type can be estimated from a control selection

[2] The numbers for $\sigma \times \epsilon_{\text{VBF}} \approx 3$ (11) pb are obtained from generator filters applying $m(q_1, q_2) > 400$ (200) GeVand $\Delta\eta(q_1, q_2) >$ 4.0 (2.0).

[3] In addition, contributions from ggH events to a VBF selection can make up to $\approx$ 30 % of the total signal yield.

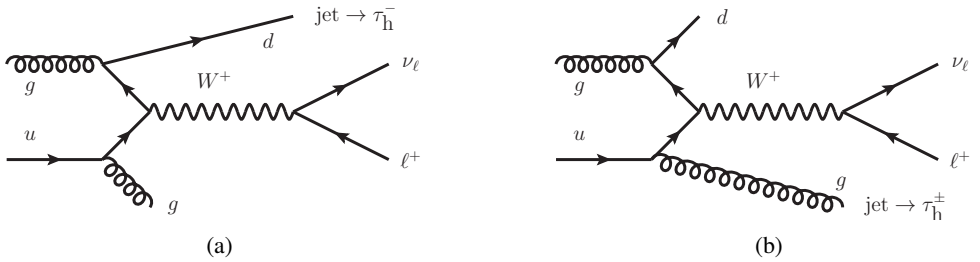


Figure 4.4: Example processes for mis-identification of (a) quark-induced and (b) gluon-induced jets as τ_h objects in W+jets.

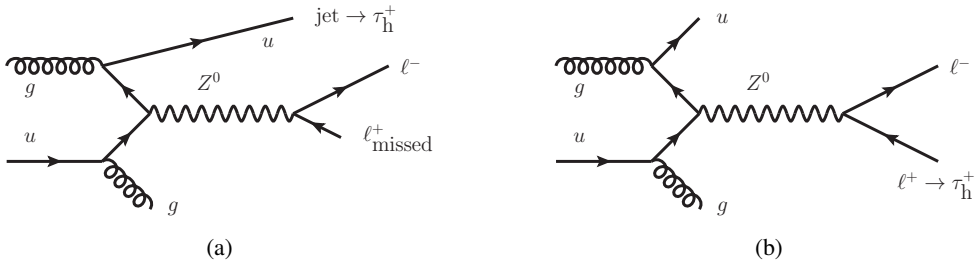


Figure 4.5: Example processes for the mis-identification of (a) jets and (b) leptons as τ_h objects in Z+jets.

of τ_h objects with weaker identification criteria, which can then be extrapolated via a measurement of their mis-identification probability.

W+jets Production

With an about ten times larger cross section than Z+jets production, also the W+jets process yields a sizeable background contribution for (ℓ, τ_h) final states, in which a light lepton originates directly from the decay of the W boson or from a subsequent decay via[4] $W^\pm \to \tau\nu \to \ell\nu_\ell\nu_\tau\bar{\nu}_\tau$. In these cases, a jet is mis-identified as τ_h, which is quark-induced in most W+jets events (cf. Fig. 4.4a). As opposed to other types of events with gluon-induced jets as shown in Fig. 4.4b, their charge is correlated via the production of the $W^\pm$ boson, so that these events cannot be assumed to be symmetric in opposite-sign and same-sign charges. Instead, even the τ_h mis-identification probability for $(\ell^\pm, \tau_h^\mp)$ and $(\ell^\pm, \tau_h^\pm)$ can be different, since also the fraction of quark- and gluon-induced jets depends on the charge. As compared to multijet events, though, the missing transverse momentum in these events is large and dominated by the neutrino produced in the decay of the W boson. Therefore, the azimuthal angle $\Delta\phi(\ell, p_T^{miss})$ as well as the transverse mass $m_T(\ell)$ are large. The latter peaks near the mass of the W boson of $m_W = 80.4$ GeV and is thus well separated from the expected signal distribution.

$Z \to \ell\ell(+j\to\tau_h)$ Background

When one of the two leptons produced in $Z \to \ell\ell$ decays is not observed due to the analysis acceptance or identification inefficiencies, Z+jets events as shown in Fig. 4.5a can provide a similar background

[4] The case of hadronic τ decays in $W^\pm \to \tau\nu \to \tau_h\nu_\tau\bar{\nu}_\tau$ is not addressed here, since the τ_h would be identified correctly, and instead only the lepton would have to originate from a jet, which is far more unlikely.

contribution as W+jets, but with a smaller cross section. In this case, however, there is no charge correlation due to the neutral Z^0 boson and the random charge of the missing lepton ℓ. Therefore, this background is completely symmetric ($n_{OS} = n_{SS}$) and can be estimated very precisely from same-sign events in the data. Although the lepton is missed as analysis object, it can still be included in the energy balance via the independent $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction algorithm[5], so that the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ is expected to be small in these events. The transverse mass $m_T(\ell)$ of these events is thus significantly smaller than in decays of W bosons and thus less well separated from the signal.

$Z \to \ell\ell(\ell \to \tau_{\mathrm{h}})$ Background

In addition to jets, also light leptons can be mis-identified as τ_{h} objects: besides fake-τ_{h} candidates from electrons (cf. Sec. 3.4.3), also muons can be reconstructed as τ_{h} if they loose a significant fraction of their energy in the calorimeters and can thus not be identified as high momentum tracks in the muon chambers. Due to these $\ell \to \tau_{\mathrm{h}}$ mis-identification mechanisms, events with $Z \to \ell\ell$ decays can also contribute to $\ell\tau_{\mathrm{h}}$ final states (cf. Fig. 4.5b). Since the charge-flip probability and the resulting contributions from $(\ell^{\pm}, \tau_{\mathrm{h}}^{\pm})$ pairs are negligible, no same-sign $\ell^{\pm}\tau_{\mathrm{h}}^{\pm}$ events are expected from this background. In contrast to all other processes with mis-identified τ_{h}, which originate from non-resonant lepton-jet topologies, this $\ell \to \tau_{\mathrm{h}}$ background is the decay product of the Z resonance with a well defined mass. While only small missing transverse energy is expected in those events due to mis-calibration of the $\ell \to \tau_{\mathrm{h}}$ objects, the transverse momenta of the visible decay products are higher than for real $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ final states. Therefore, the invariant $\tau\tau$ mass as reconstructed under the wrong $\ell \to \tau_{\mathrm{h}}$ assumption can obtain higher values than for $Z \to \tau\tau$ events, which are thus closer to the expected signal mass.

4.2.3 Backgrounds from Complex Processes

Two further classes of processes give rise to more complex background events (cf. Fig. 4.6), which can contain both leptons and either real hadronic τ decays or mis-identified τ_{h} objects in ℓ+jet final states. Since their cross sections are well-known and small enough, a separate estimation of these cases is found to be unnecessary.

Top-Quark Production

Although produced via the strong interaction, top quarks can only decay via the weak $t \to W^+b$ process, so that processes with top quarks always involve subsequent decays of W bosons and can therefore yield contributions to almost any leptonic final state. While the process with the largest cross section is the $t\bar{t}$ pair production (cf. Fig. 4.6a), also the production of single top quarks must be considered. Usually all types of events with top quarks are characterised by large missing transverse momentum due to neutrinos from the subsequent W boson decays and a large number of jets due to the production via the strong interaction. Experimentally, the jets originating from bottom quarks can be identified by the long lifetime of b-hadrons produced during their hadronisation, so that contributions from top quark backgrounds can be both suppressed and estimated via the selection of b-tagged jets.

[5] The missing lepton can e.g. be outside the analysis acceptance, but still within the detector acceptance. It can also be mis-identified as jet (e) or considered as in-flight decay within a jet (μ). In most cases, it therefore enters the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction, while any mis-measurement is mostly limited to calibration effects.

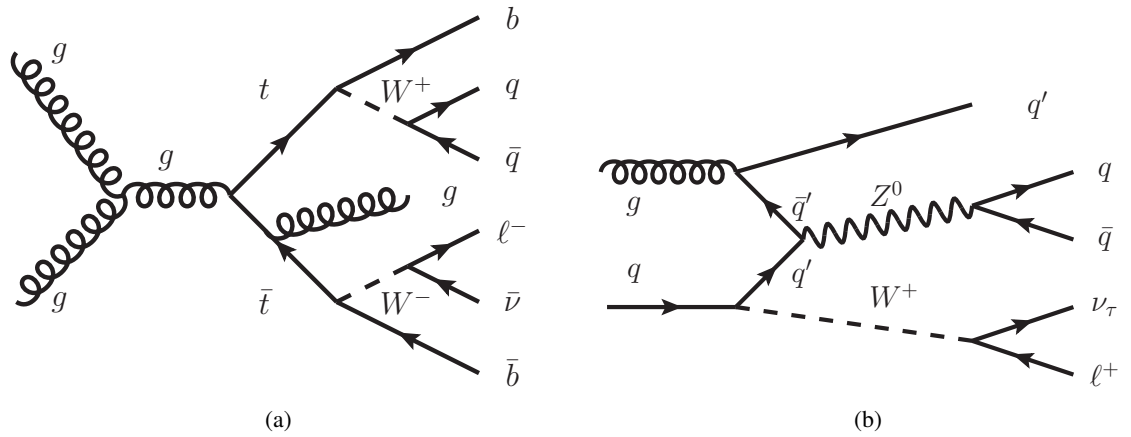


Figure 4.6: Examples for ℓ+jet final states produced via top quark pair production (a) and di-boson production (b). Since the decays of $W^{\pm}$ and Z^0 bosons can produce τ-leptons, these processes can also result in real $\ell\tau_{\mathrm{h}}$ final states with additional jets.

Di-Boson Production

Similar to the production of two W bosons from $t\bar{t}$ pairs, two vector bosons can also be produced directly in other electroweak interactions (cf. Fig. 4.6b). This production mechanism can result not only in WW events, but also in WZ and ZZ pairs, which can produce a large variety of final states, including two real τ decays as well as direct lepton production or mis-identified τ_{h} decays. Almost identical final states are also expected from $H \to WW$ and $H \to ZZ$ decays, but their contributions to the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ selection are found to be negligible due to their small cross sections.

4.2.4 Monte Carlo Generator Implementations

Although the properties of certain background processes allow a data-driven estimation of their contributions to the selection of $H \to \tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ events, the use of Monte Carlo event simulation cannot be avoided in most cases. Therefore, different generator implementations are used to predict the yields of the various background contributions, which are optimised to provide reliable estimates of the different production mechanisms.

Leading-Order Generators with Parton Shower

For most electroweak processes, the matrix element calculation can provide all particles and their kinematics in the final state, which corresponds to the stable particles produced in nature. Due to the low-energy behaviour of QCD, however, the generation of jets is usually provided by parton shower algorithms such as PYTHIA [98] and HERWIG[6] [100, 101], which mainly differ in their fragmentation into hadrons implemented via string dynamics [62] or clustering methods [63]. For simple processes such as the direct production of one or two vector bosons, these generators can also provide the full event kinematics via the simple leading order matrix elements and can thus be run standalone. Only the obtained cross sections need to be corrected in order to account for higher order effects (NNLO), which can e.g. be calculated with FEWZ [102–104].

[6] In combination with HERWIG, the underlying event is always modelled by JIMMY [99].

Matrix-Element and Parton Shower Matching

The differential jet distributions as generated by the parton shower algorithms cannot describe the production of highly energetic jets e.g. in Z+jets and W+jets processes well enough, so that this part of the phase space must be covered by more accurate matrix element calculations. These are usually provided by the ALPGEN [105] generator, which calculates up to five hard parton emissions from the superposition of all corresponding leading-order matrix elements. The generated parton final states are then passed to HERWIG or PYTHIA, which evolve each parton with a shower towards hadronisation. In order to avoid double counting of phase space regions, the MLM matching scheme [106, 107] is applied to prevent a parton splitting in the shower algorithm if it is already included in the phase space of the matrix element calculation. The inclusive cross section is afterwards normalised to the result of NNLO calculations [103]. Since higher order electroweak corrections are not included in these processes, the VBF production of Z bosons is not accounted for in the ALPGEN samples and is instead obtained separately from the SHERPA [108] generator.

Next-to-Leading-Order Generators

While the differential kinematic distributions of the Z+jets or W+jets production processes are not significantly affected by missing virtual loop-corrections at LO[7] (cf. Fig. 2.11), more QCD-dominated final states such as $t\bar{t}$ and ggH and more complex electroweak processes such as di-boson production or vector boson fusion require NLO calculations, which include the virtual corrections. Therefore, different next-to-leading-order generators are used: MC@NLO [110] interfaced to HERWIG and POWHEG [111, 112] interfaced to PYTHIA. The cross sections for the $t\bar{t}$ and di-boson production are obtained from NNLO [113–116] and NLO [117] calculations respectively. Single-top production is provided by AcerMC [118] and normalised by NNLO cross sections [119–121], while the loop-induced $gg \to WW$ production is obtained from the dedicated generator gg2WW [122].

For the ggH process, the differential distribution of the Higgs boson transverse momentum $p_T(H)$ is calculated with the HqT program [123], which accounts for quark mass effects on the QCD ISR parton emission. The events produced with standard Monte Carlo generators as described in Sec. 2.7 are re-weighted, so that their differential $p_T(H)$ distribution corresponds to the HqT calculation.

Event Post-Processing

The decays of τ-leptons are not implemented in the standard generators, so that these particles are passed to the dedicated TAUOLA [124] program. This substitutes the τ-leptons by decay products, which are randomly chosen according to the implemented library of branching fractions as obtained from a large number of measurements from different experiments. The kinematic configuration of each τ decay is required to fulfil constraints from the matrix element calculation including spin effects. As additional post-processing step, QED radiation effects for all stable particles are included by the PHOTOS [125] program, so that the generated final states correspond exactly to the expected events produced in the vacuum[8].

[7] The calculated cross sections are corrected with constant k-factors in order to correspond to NNLO accuracy [109].

[8] An exception to the standard post-processing with TAUOLA and PHOTOS is the SHERPA generator, which implements its own parton shower and τ decay library as well as an algorithm to include QED radiation effects.

Parton-Distribution-Functions

As parton-distribution-functions, the set CT10 [126] is used for the MC@NLO, POWHEG and SHERPA samples. ALPGEN samples are generated with CTEQ6L1 [127] instead, while samples produced by PYTHIA or HERWIG standalone use either CTEQ6L1 or MRST2007 [128]. Systematic uncertainties are usually obtained by generating samples with different sets of distributions e.g. for CTEQ [129].

4.3 Extraction of Signal from Backgrounds

In order to extract the small number of expected $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$ events from the vast majority of background events, observables need to be defined, which can provide a sufficient separation between signal and background. A selection of events based on a combination of these variables should then define a phase space region with small background contamination but high signal acceptance. An excess of observed $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$ signal over the backgrounds expected from other Standard Model processes can then be interpreted within a statistical analysis, which must be employed to obtain the significance of the excess.

4.3.1 Mass Reconstruction

As the best separating variable between Higgs and Z boson decays, the invariant mass of both particles is a key observable when searching for $H \rightarrow \tau\tau$. Due to the sub-sequent τ decays, undetected neutrinos create missing transverse momentum, which complicates the situation and severely affects the mass resolution.

Visible and Effective Mass

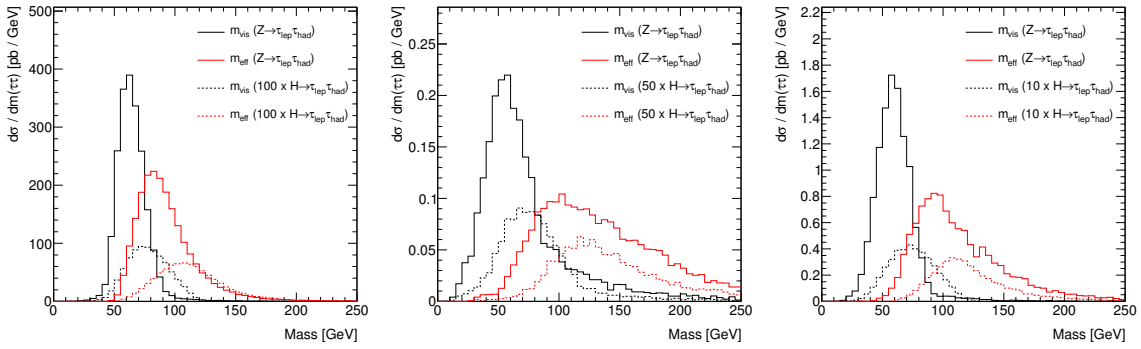
A simple reconstruction of the mass of the visible decay products ℓ and τ_{h} with four-vectors $\mathbf{p}_\ell$ and $\mathbf{p}_{\tau_{\text{h}}}$

$$m^2_{\text{vis}} = m(\ell, \tau_{\text{h}})^2 = (\mathbf{p}_\ell + \mathbf{p}_{\tau_{\text{h}}})^2 < \left(\mathbf{p}_{\tau_{\text{lep}}} + \mathbf{p}_{\tau_{\text{had}}}\right)^2 = m^2(\tau, \tau)$$

misses a significant fraction of the corresponding total four-momenta $\mathbf{p}_{\tau_{\text{lep}}}$ and $\mathbf{p}_{\tau_{\text{had}}}$ of the τ-leptons. This results in a mass peak, which is shifted with respect to the nominal value of the invariant mass $m(\tau, \tau)$, broader than the expected resolution and thus yields only a weak separation between the Higgs and Z boson mass. Instead, the effective mass defined as

$$m^2_{\text{eff}} = m^2(\ell, \tau_{\text{h}}, p_{\text{T}}^{\text{miss}}) = (\mathbf{p}_\ell + \mathbf{p}_{\tau_{\text{h}}} + \mathbf{p}_{T,\text{miss}})^2 < (\mathbf{p}_\ell + \mathbf{p}_{\tau_{\text{h}}} + \mathbf{p}_{\text{miss}})^2 = \left(\mathbf{p}_{\tau_{\text{lep}}} + \mathbf{p}_{\tau_{\text{had}}}\right)^2 = m^2(\tau, \tau)$$

can account for the missing transverse momentum $\mathbf{p}_{T,\text{miss}} = (|\vec{p}_{\text{T}}^{\,\text{miss}}|, \vec{p}_{\text{T}}^{\,\text{miss}})^T$, but its longitudinal component cannot be observed and is still missing from this mass definition. In order to reconstruct this longitudinal component, the large boost of the τ decays must be exploited, which causes the directions of neutrinos and visible decay products to be very similar to the original directions of the τ-leptons.



(a) (b) (c)

Figure 4.7: Distributions for the visible mass m_{vis} (black) and the effective mass (red) m_{eff} for (a) an inclusive selection dominated by back-to-back topologies of the two τ decays, (b) a selection of boosted $H \to \tau\tau$ decays and (c) a di-jet selection dominated by VBF topologies. These are shown for the $Z \to \tau_{\text{lep}}\tau_{\text{had}}$ process as solid lines and for $H \to \tau_{\text{lep}}\tau_{\text{had}}$ as dashed lines, of which the latter process is scaled by a factor given in the legend.

Collinear Approximation

With the so-called collinear approximation [130], these directions of the visible and invisible decay products[9] are assumed to be identical ($\Delta R(\tau_{\text{h}}, \nu_\tau) = 0$ and $\Delta R(\ell, \nu_{\tau+\ell}) = 0$), so that the missing transverse momentum can be projected onto the two directions given by the visible decay products as shown in Fig. 4.8. With geometrical calculations, the azimuthal angles $\Delta\phi$ between $\vec{p}_T(\ell)$, $\vec{p}_T(\tau_{\text{h}})$ and $\vec{p}_{\text{T}}^{\,\text{miss}}$ in the transverse plane can thus be converted into the visible energy fractions

$$x_{\text{lep}} = p_T(\ell)/p_T(\tau_{\text{lep}}) \,, \qquad x_{\text{had}} = p_T(\tau_{\text{h}})/p_T(\tau_{\text{had}}) \,.$$

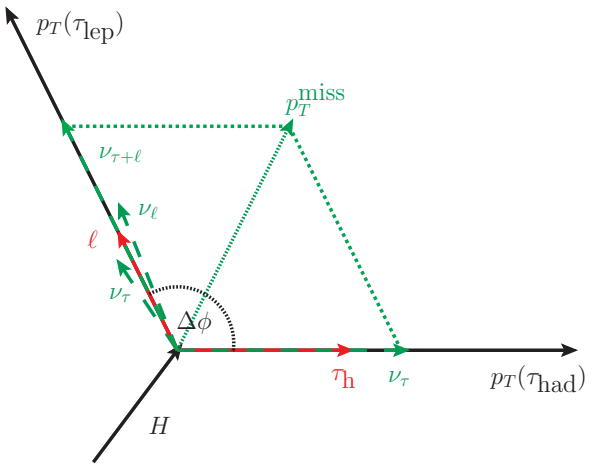


Figure 4.8: Sketch of the momenta of the $H \to \tau_{\text{lep}}\tau_{\text{had}}$ decay products in the plane transverse to the beam axis. In the collinear approximation, the directions between the τ-leptons and their decay products are identical, and the projection of the missing transverse momentum $p_{\text{T}}^{\text{miss}}$ provides the neutrino momenta ($\nu_{\ell+\tau}$ and ν_τ).

[9] For simplicity, the two neutrino momenta in the leptonic decay are added to an effective neutrino momentum $\vec{p}_{\nu_{\tau+\ell}} = \vec{p}_{\nu_\tau} + \vec{p}_{\nu_\ell}$, so that $\nu_{\tau+\ell}$ is used as a notation for this vectorial sum.

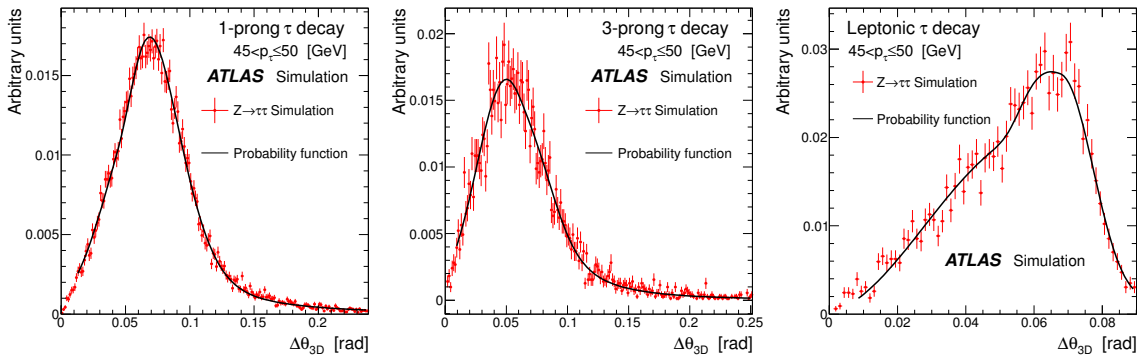
Since the momenta of all involved particles are much larger than the rest mass of the τ-leptons, the invariant mass can be approximated by $m^2 \approx p_\tau p_\tau \, (1 - \cos \Delta\phi(\ell, \tau_\mathrm{h}))$, which results in the definition

$$m_\mathrm{coll} = \frac{m_\mathrm{vis}}{\sqrt{x_\ell x_{\tau_\mathrm{h}}}} \approx m(\tau, \tau) \; .$$

The calculation of this *collinear mass*, however, has to rely on the correct reconstruction of $\vec{p}_\mathrm{T}^{\,\mathrm{miss}}$, while the experimental measurement of $\vec{E}_\mathrm{T}^{\mathrm{miss}}$ is always subject to resolution effects. In experimental conditions, the momentum fractions can thus exceed their physically defined range $0 < x_\ell, x_{\tau_\mathrm{h}} < 1$, which yield unphysical values of the mass. Especially in back-to-back topologies, the distribution of $p_\mathrm{T}^{\mathrm{miss}}$ on the visible momenta is therefore not well defined, and can lead to large tails in the mass distributions resulting in an enhanced overlap of the Higgs boson signal with the upper tail of the Z mass peak.

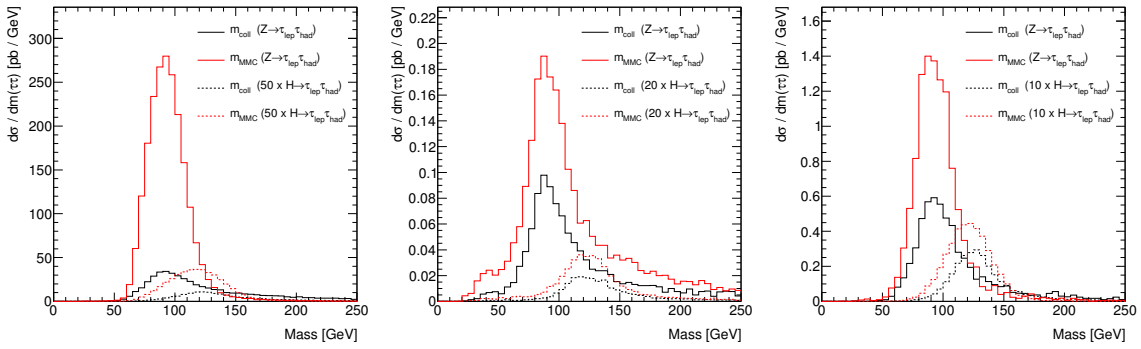
Missing Mass Calculator

Instead of approximating the τ decay angles $\Delta R(\tau_\mathrm{h}, \nu_\tau) = \Delta R(\ell, \nu_{\tau+\ell}) = 0$, the effective distributions of these angles can be obtained and parametrised from simulated events (cf. Fig. 4.9). When interpreting these as probability distributions, a kinematic scan can identify the most likely configuration of decay angles for each data event, which corresponds to the most probable mass given the kinematics of the visible decay products ℓ, τ_h and $p_\mathrm{T}^{\mathrm{miss}}$. This technique is implemented in the Missing Mass Calculator (MMC) [131], which additionally accounts for experimental resolution effects on the missing transverse momentum in the kinematic scan. As compared to the collinear mass in Fig. 4.10, the upper tail of the MMC mass distribution for $Z \to \tau\tau$ events shows a smaller overlap with a potential $H \to \tau\tau$ signal. Especially for events with low $E_\mathrm{T}^{\mathrm{miss}}$, the requirement of $0 < x_\ell, x_{\tau_\mathrm{h}} < 1$ for the collinear approximation leads to a significant loss of efficiency due to $E_\mathrm{T}^{\mathrm{miss}}$ resolution effects, while similar losses within the kinematic scans of the MMC are at the order of 5%. The reconstructed MMC mass $m_\mathrm{MMC}(\tau, \tau)$ thus provides the best separation between Higgs and Z boson decays into $\tau\tau$ final states.



(a) (b) (c)

Figure 4.9: Angular distance distributions between neutrinos and visible decay products for τ-leptons with generated momentum $45 < p \leq 50$ GeV obtained from simulated $Z \to \tau\tau$ events for (a) 1-prong, (b) 3-prong and (c) leptonic decays. The solid black line shows the functions used in the calculations of global event probabilities [132].



(a) (b) (c)

Figure 4.10: Distributions for the collinear mass m_{coll} (black) and the MMC mass (red) m_{MMC} for (a) an inclusive selection dominated by back-to-back topologies of the two τ decays, (b) a selection of boosted $H \to \tau\tau$ decays and (c) a di-jet selection dominated by VBF topologies. These are shown for the $Z \to \tau_{\text{lep}}\tau_{\text{had}}$ process as solid lines and for $H \to \tau_{\text{lep}}\tau_{\text{had}}$ as dashed lines, while the latter process is scaled by a factor given in the legend. For the collinear approximation, the momentum fractions are required to be in the physical range of $0 < x_\ell, x_{\tau_{\text{h}}} < 1$, which especially suppresses events with back-to-back topologies due to $E_{\text{T}}^{\text{miss}}$ resolution and distribution effects discussed in the text.

4.3.2 Analysis Observables

Besides the mass $m(\tau,\tau)$, the observables $p_T(\tau_{\text{h}}) - p_T(\ell)$ and $p_T(\ell)/p_T(\tau)$ can provide a weak separation between Higgs and Z boson decays, since they are sensitive to spin correlations as described in Sec. 4.2.1. Similar information is contained in the correlations of the visible momentum fractions, so that also $x_{\text{lep}}/x_{\text{had}}$ and $\Delta x = x_{\text{lep}} - x_{\text{had}}$ can be considered as variables to identify $H \to \tau\tau$ decays via spin correlation effects. For other backgrounds, however, which originate from other processes and not from real decays of τ-leptons, these observables are expected to show even larger differences as compared to both $H \to \tau\tau$ and $Z \to \tau\tau$. Especially for W+jets, the prompt lepton produced in the decay of the W boson obtains a significantly higher momentum than a lepton produced in a $H \to \tau_{\text{lep}}\tau_{\text{had}}$ decay with two additional neutrinos. Therefore, already the visible kinematics can be exploited to distinguish signal from background events. When considering the missing momentum at the same time, which is carried by the neutrinos, further observables with even larger separation power can be constructed.

Topology of ℓ, τ_{h} and $\vec{p}_{\text{T}}^{\,\text{miss}}$

An important aspect of the $\tau\tau$ decay topology is that the missing transverse momentum originates from the subsequent decays of the τ-leptons, and that their directions are thus related. This is exploited with the transverse mass

$$m_T(\ell) = \sqrt{2 \cdot p_T(\ell) \cdot E_{\text{T}}^{\text{miss}} \left(1 - \cos\Delta\phi(\ell, E_{\text{T}}^{\text{miss}})\right)},$$

which is usually large for W+jets events with back-to-back topologies of the lepton ℓ and the neutrino reconstructed as missing transverse energy $E_{\text{T}}^{\text{miss}}$. Similarly, the transverse mass $m_T(\tau_{\text{h}})$ can be calculated with the τ_{h} object and becomes large when the $E_{\text{T}}^{\text{miss}}$ is dominantly created by an energy mis-measurement of a jet.

While these observables can already reject a large part of the W+jets and multijet backgrounds, es-

pecially the more signal-like events can be investigated in more detail in order to define additional suppression criteria. Considering a $\tau\tau$ pair with an enclosed opening angle $\Delta\phi(\ell,\tau) = |\phi(\ell) - \phi(\tau)| < \pi$, the resulting $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ direction lies between the two visible objects:

$$\phi(\ell) < \phi(E_{\mathrm{T}}^{\mathrm{miss}}) < \phi(\tau_{\mathrm{h}}) \qquad \text{or} \qquad \phi(\ell) > \phi(E_{\mathrm{T}}^{\mathrm{miss}}) > \phi(\tau_{\mathrm{h}}) \,.$$

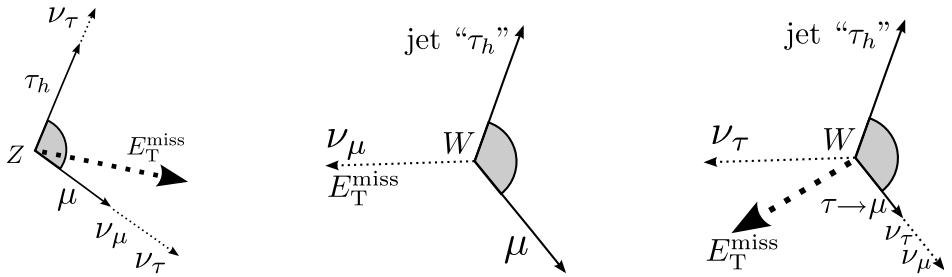
In W+jets, $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ is instead dominated by the neutrino from the W boson decay, which is usually back-to-back with the lepton with $\Delta\phi(\ell,\nu) = |\phi(\ell) - \phi(\nu)| \approx \pi$. As shown in Fig. 4.11, the resulting direction of $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ is unlikely to be found between the directions of the visible $\tau\tau$ decay products when the τ_{h} is a mis-identified jet with $\Delta\phi(\ell,\mathrm{jet}) = |\phi(\ell) - \phi(\mathrm{jet})| < \pi$. The variable

$$\sum \cos\Delta\phi = \cos\left(\phi(\ell) - \phi(E_{\mathrm{T}}^{\mathrm{miss}})\right) + \cos\left(\phi(E_{\mathrm{T}}^{\mathrm{miss}}) - \phi(\tau_{\mathrm{h}})\right)$$

thus tends towards negative values, when the direction $\phi(E_{\mathrm{T}}^{\mathrm{miss}})$ is inconsistent with a $\tau\tau$ final state. Similarly, the variable

$$\sum \Delta\phi = |\phi(\tau_{\mathrm{h}}) - \phi(E_{\mathrm{T}}^{\mathrm{miss}})| + |\phi(\ell) - \phi(E_{\mathrm{T}}^{\mathrm{miss}})|$$

is expected to be smaller $\sum \Delta\phi < \pi$ for $\tau\tau$ decays than for other backgrounds.



(a) $Z \to \tau\tau \to \mu\tau_{\mathrm{h}}$ (b) $W \to \mu\nu$ (c) $W \to \tau\nu \to \mu\nu\nu\nu$

Figure 4.11: Drawings of representative transverse plane orientations of $W^{\pm}$ and Z^0 decay products and the $E_{\mathrm{T}}^{\mathrm{miss}}$. The shaded angles indicate the angle $\Delta\phi < \pi$ between the lepton and the τ_{h} candidate. In (a), the Z boson is depicted to have non-zero transverse momentum, which must be balanced on the left by other jet activity [133].

As a more sophisticated variable, the $E_{\mathrm{T}}^{\mathrm{miss}}\phi$-centrality

$$C^{\phi}_{\ell,\tau_{\mathrm{h}}}(E_{\mathrm{T}}^{\mathrm{miss}}) = \frac{\psi_\ell(E_{\mathrm{T}}^{\mathrm{miss}}) - \psi_{\tau_{\mathrm{h}}}(E_{\mathrm{T}}^{\mathrm{miss}})}{\sqrt{\psi_\ell(E_{\mathrm{T}}^{\mathrm{miss}})^2 + \psi_{\tau_{\mathrm{h}}}(E_{\mathrm{T}}^{\mathrm{miss}})^2}}$$

gives the relative position of the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ direction with respect to the visible τ decay products. After transforming the azimuthal angles ϕ into the system $\psi_\ell(E_{\mathrm{T}}^{\mathrm{miss}}), \psi_{\tau_{\mathrm{h}}}(E_{\mathrm{T}}^{\mathrm{miss}})$ via

$$\psi_\ell(E_{\mathrm{T}}^{\mathrm{miss}}) = \sin\left(\phi(\ell) - \phi(E_{\mathrm{T}}^{\mathrm{miss}})\right) / \sin\left(\phi(\ell) - \phi(\tau_{\mathrm{h}})\right) \,,$$

$$\psi_{\tau_{\mathrm{h}}}(E_{\mathrm{T}}^{\mathrm{miss}}) = \sin\left(\phi(\tau_{\mathrm{h}}) - \phi(E_{\mathrm{T}}^{\mathrm{miss}})\right) / \sin\left(\phi(\ell) - \phi(\tau_{\mathrm{h}})\right) \,,$$

this centrality $C^{\phi}_{\ell,\tau_{\mathrm{h}}}(E_{\mathrm{T}}^{\mathrm{miss}})$ assumes its maximum of $\sqrt{2}$ for the situation of a perfectly central $\phi(E_{\mathrm{T}}^{\mathrm{miss}})$ direction between the visible τ decay products.

Additional Jet Activity

Besides the topology of the boson decay, further distinction between electroweak production mechanisms and $t\bar{t}$ or multijet production is provided by the jet activity in the collision events. Therefore, the observable

$$\sum p_T = \sum_{j\in\{\ell,\tau_\mathrm{h},\mathrm{jets}\}} |p_T^j|$$

is expected to yield lower scalar sums of object transverse momenta in the production of Higgs, Z or W bosons than for backgrounds produced via the strong interaction such as $t\bar{t}$. Likewise, the transverse momentum of the reconstructed $\tau\tau$ system

$$p_T(\tau\tau) = p_T(H) \quad \text{with} \quad \vec{p}_T(H) = \vec{p}_T(\ell) + \vec{p}_T(\tau_\mathrm{h}) + \vec{E}_\mathrm{T}^\mathrm{miss}$$

is sensitive to the topology of additional jet radiation in the event. Although this direct recoil from the jet activity does not have a large separation power on its own, especially its correlation with other observables provides a good distinction between signal and background events. While the boost of the Higgs boson leads to a rather well defined angular separation $\Delta R(\tau,\tau) \approx \Delta R(\ell,\tau_\mathrm{h})$, this is not the case for multijet or $t\bar{t}$ events or other backgrounds with mis-identified τ_h decays, where the kinematics of ℓ and τ_h are not directly related via a common mother particle. Furthermore, the requirement of such a boost significantly improves the mass resolution in Higgs and Z boson decays, since the missing transverse momentum can better be distributed on the neutrino directions if the τ decays are boosted more strongly and the angular separation $\Delta\phi(\ell,\tau_\mathrm{h})$ is small.

In case of VBF-like events, the absolute value of the vectorial sum

$$p_T^\mathrm{tot} = |\vec{p}_T(\ell) + \vec{p}_T(\tau_\mathrm{h}) + \vec{p}_\mathrm{T}^\mathrm{miss} + \vec{p}_T(q_1) + \vec{p}_T(q_2)|$$

should vanish for the electroweak quark scattering in qqH or qqZ events, while jet activity from QCD ISR effects in most cases leads to a non-vanishing transverse momentum of the system of objects entering p_T^tot. Additionally, the two jets in background events prefer large azimuthal separation $\Delta\phi(q_1,q_2) \approx \pi$, since they are dominated by highly energetic di-jet events with softer ℓ and τ_h objects, whereas the quark-jets in the qqH signal process have no such preference in $\Delta\phi(q_1,q_2)$.

Observables for VBF-like Events

While the previous variables for qqH events only contain information in the transverse plane, the longitudinal correlation of the two jets is provided by the difference $\Delta\eta(j_1,j_2) = \eta(j_1) - \eta(j_2)$ of their η components or their product $\eta(j_1) \times \eta(j_2)$, which is expected to be negative, when both quarks are scattered into opposite detector hemispheres. In addition to their large separation in η, the momenta of the jets from VBF production are usually larger than for average jets from ISR, multijet or $t\bar{t}$ production, which results in large invariant di-jet masses $m(j_1,j_2)$. As another topological feature of real VBF events, the direction $\eta(H)$ of the Higgs boson must be between the directions of the jets (i.e. $\eta(j_1) < \eta(H) < \eta(j_2)$ without loss of generality). Although the true direction $\eta(H)$ cannot be reconstructed directly due to the unknown longitudinal neutrino momenta[10], further background suppression can be obtained from the visible decay products ℓ and τ_h either by requiring this η-centrality as binary decision or by defining a

[10] Within the collinear approximation or from a kinematic scan as performed by the MMC, a better estimation of this direction is possible and could be used in the future, but this additional complication was not considered in the context of this thesis.

continuous η-centrality

$$C^{\eta}_{j_1,j_2}(x) = \exp\left(-\frac{(\eta(x)-\bar{\eta})^2}{w_\eta^2}\right) \quad \text{with} \quad \bar{\eta} = \frac{\eta(j_1)+\eta(j_2)}{2} \quad \text{and} \quad w_\eta = \frac{|\eta(j_1)-\eta(j_2)|}{2} \,.$$

This observable peaks at the value 1 for perfect centrality of $\eta(x)$ between $\eta(j_1), \eta(j_2)$ and drops below $1/e$ when $\eta(x)$ is outside the expected range $[\eta(j_1), \eta(j_2)]$.

4.3.3 Signal Extraction

With a given selection of $\ell\tau_\mathrm{h}$ final states and an observed number of events selected from pp-collision data, the statistical compatibility of the observation with signal and background hypotheses must be tested with refined techniques (cf. [94]) in order to interpret the result. In case of an excess of signal over the background expectation, the null hypothesis H_0 corresponds to the assumption that only background and no signal exists, which is attempted to be rejected with the observed data. In the different situation with no observed excess, the exclusion of a signal can be tested, for which the role of the hypothesis is flipped: the null hypothesis assumes signal and background and should be rejected to exclude the signal process.

Significance and p-Value

The level of agreement of observed data with a hypothesis H is quantified by the probability that statistical fluctuations of any data obtained from a true H (*pseudo-experiment*) would result in equal or worse compatibility with the expectation from H than the actual observation. This probability is usually based on the number of expected events in a binned distribution with strong signal and background separation. In order to discover a new signal, it is a standard procedure to test the background-only hypothesis and thus to calculate the so-called p_b- or p_0-value as the probability that the observed data are consistent with the expectation from H_0 without signal. If this is found to be sufficiently low, the null hypothesis can be rejected and an observed excess is confirmed. The significance Z of this excess is then calculated from the p-value by interpreting it as the upper-tail probability (cf. 4.12a) of a Gaussian distributed variable, which is observed to be Z standard deviations (σ) above its mean value (m) as shown in Fig. 4.12a. With the Gaussian quantile $\Phi^{-1}(p)$ denoting the inverse of the cumulative Gaussian distribution, this significance is therefore given by

$$Z = \Phi^{-1}(p)(1-p) \quad \text{with} \quad \Phi^{-1}(p) = m + \sigma\sqrt{2}\,\mathrm{erf}^{-1}(2p-1) \,.$$

For the observed signal, the compatibility with the expected signal-plus-background hypothesis can be verified with the corresponding p_{s+b}- or p_1-value, which can either confirm or reject that the observed signal is in agreement with the signal expectation. Following the convention of particle physics, the discovery of a new signal process such as expected from the Higgs boson requires a significance of at least 5σ or a p-value of $p \leq 2.87\times10^{-7}$, while an excess with at least 3σ or $p \leq 1.35\times10^{-3}$ is considered as evidence for a new process. In the absence of a significant excess, the expected signal can be excluded by rejecting the null hypothesis of signal and background, which usually requires a confidence level of 95% corresponding to $p_{s+b} \leq 0.05$ (i.e. $Z \geq 1.64$). If a signal can neither be excluded nor produces a sufficiently large excess, the sensitivity of an analysis can still be quantified by excluding the same signal with a larger signal strength scaled by the parameter μ. In this situation, a scan for the minimal signal strength μ is performed, which can be excluded at 95% confidence level with the observed data.

Construction of a Likelihood Function

For observed data distributed over N bins of a histogram $\mathbf{n} = (n_1, ..., n_N)$, the statistical compatibility of the values n_i must be tested against their corresponding expectation values

$$E[n_i] = \mu s_i(\theta) + b_i(\theta) \,,$$

which are constructed from the expected number of background events b_i and an admixture of s_i expected signal events scaled by the signal strength parameter μ. These expected numbers are subject to a set of systematic parameters $\theta = (\theta_1, ..., \theta_M)$ with assumed Gaussian widths $\Delta\theta_k$ for each θ_k. Although the correct parameter set $\tilde{\theta}$ corresponding to the true values in data cannot be reconstructed a priori, the optimal set $\hat{\theta} \approx \tilde{\theta}$ can be estimated by maximising a likelihood function. From the Poisson probability of observing n_i events where $E[n_i]$ are expected, such a likelihood function can be constructed as the product of the individual bins:

$$L(\mu, \theta) = \prod_{j=1}^{N} \frac{(\mu s_i(\theta) + b_i(\theta))_i^n}{n_i!} \cdot \prod_{k=1}^{k=M} G(\theta_k) \,.$$

Additional Gaussian terms[11] ensure the given constraints on the parameters θ_k, which are transformed, so that the central values $\theta_{k,0}$ are set to 0 and the corresponding uncertainties $\Delta\theta_k$ are normalised to 1:

$$G(\theta_k) = \frac{1}{\sqrt{2\pi(\Delta\theta)^2}} \exp\left(\frac{(\theta_k - \theta_{k,0})^2}{2(\Delta\theta_k)^2}\right) \,.$$

The dependence of the background and signal components on these so-called *nuisance parameters* is then implemented via the corresponding variations

$$b_i(\theta) = \sum_{j \in \text{ processes}} b_{i,j} + \sum_{k=1}^{k=M} \theta_k \sigma_{i,j,k} \qquad \text{and} \qquad s_i(\theta) = \sum_{j \in \text{ processes}} s_{i,j} + \sum_{k=1}^{k=M} \theta_k \sigma_{i,j,k} \,,$$

where the systematic uncertainties $\sigma_{i,j,k}$ can be different for each process j, and in case of shape uncertainties also for each bin i. The values of all θ_k and their corresponding uncertainty $\Delta\theta_k$ can then be fitted to the observed data $\mathbf{n}$, so that the resulting set $\hat{\theta}$ maximises the likelihood, and thus gives the best statistical agreement between expectation $\mathbf{E}[\mathbf{n}]$ and observation $\mathbf{n}$.

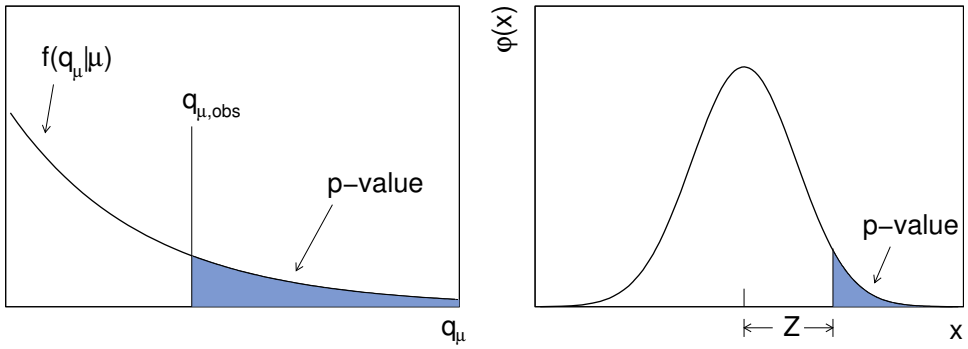
Profile Likelihood

The calculation of a p-value and a significance from an experiment requires the definition of a test statistic q, for which the p-value can be obtained as the upper-tail probability

$$p = \int_{q_{\text{obs}}}^{\infty} f(q) \mathrm{d}q$$

from an observed value q_{obs}. This integration is performed over all values $q \geq q_{\text{obs}}$ corresponding to equal or worse compatibility of a pseudo-experiment (q) as compared to the actual observation (q_{obs}) and thus requires the knowledge of the density function $f(q)$ describing the probability to observe the value

[11] In special cases such as statistical uncertainties of a background expectation, Poisson terms can also be used in order to implement the underlying probability distribution correctly.



(a) (b)

Figure 4.12: The p-value of the observed q as the upper-tail probability of the distribution $f(q|\mu)$ and (b) the relation between p-value and significance via its interpretation in terms of a Gaussian distributed variable x [94].

q in a single random experiment (cf. Fig. 4.12a). While any such distribution can usually be obtained for any definition of q by computing a large number of such randomised pseudo-experiments and extracting the corresponding values of the test statistic q, the amount of required computing time is unreasonably large for most applications. Instead, a more elegant and less computing intensive method is to choose q in a way that its probability distribution $f(q)$ is known beforehand. Such a special choice is the test statistic $q_\mu = -2 \ln \lambda(\mu)$, which is based on the profile likelihood ratio [94] defined in dependence of the signal strength μ to be tested:

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

It is built as the ratio of the conditional likelihood $L(\mu, \hat{\hat{\theta}})$, which is maximised by the parameter set $\hat{\hat{\theta}}$ for the condition of testing a fixed μ, and the maximum likelihood $L(\hat{\mu}, \hat{\theta})$ with its corresponding estimators $\hat{\theta}$ and $\hat{\mu}$, which are found as results from an unconditional maximisation considering μ as a free parameter. According to Wald [134], this ratio can be simplified by the approximation

$$q_\mu = \frac{(\mu - \hat{\mu})^2}{\sigma^2} + O(\frac{1}{\sqrt{N}})$$

with the variance $\sigma^2 = (\delta\mu)^2$ of the signal strength. As long as the sample size N of the observed data is large, this approximation holds and leads to probability densities $f(q_\mu|\mu)$ for given values of μ, which are χ^2 distributed in $\sqrt{q_\mu}$ [135]. Hence, the p-value can be obtained via

$$p_\mu = 1 - F(q_\mu|\mu) = 1 - \Phi(\sqrt{q_\mu}) \, .$$

The significance for a potential discovery can thus be simplified to $Z_0 = \Phi^{-1}(1 - p_0) = \sqrt{q_0}$ for an observed excess in the data compared to the background-only ($\mu = 0$) hypothesis. For a signal exclusion, an upper limit on the signal strength parameter can be obtained numerically from the value μ, for which $p_\mu = 0.05$ ($Z_\mu = \sqrt{q_\mu} = 1.64$) corresponds to the confidence level of 95%.

Expected and Observed Significance and Exclusion Limits

The p-value and significance as discussed so far are only defined for an observed dataset and are naturally subject to statistical fluctuations in the data. These quantities, however, cannot independently quantify the sensitivity of an analysis without considering the actual observation and should thus not be used for any further development or optimisation. Therefore, an additional quantification of the *expected significance* is desired, which is valid for an analysis regardless of fluctuations in the data. For this purpose, the so-called *Asimov* data are generated directly from the expectation values $E[n_i]$, which are constructed from the signal and background expectations with a signal strength of $\mu = 1$ ($\mu = 0$) for an observation (exclusion). Although these expectation values are usually non-integer numbers, the factorial terms in the Poisson likelihood function cancel in the profile likelihood ratio, so that an expected significance and exclusion limit are equally well defined as their observed counterparts. They can be interpreted as the results expected from "perfect" data, which do not fluctuate around the expectation values. The difference between observed and expected values hence corresponds to the statistical fluctuation of the observed data.

In practice, two likelihood functions must thus be constructed from the binned histograms: one with the observed data and one with the Asimov data. With standard numerical methods, these are usually maximised first with an unconstrained signal strength parameter μ and afterwards with a constrained value (e.g. $\mu = 0$). The ratios of the maximised function values can then be converted directly to the corresponding expected and observed p_μ-values and significances Z_μ or exclusion limits on μ.

At the same time, this procedure yields a set of nuisance parameters $\hat{\theta}$ and the signal strength $\hat{\mu}$, which correspond to the best fit to the observed data. The Asimov data, however, cannot reflect the true values of the estimators θ, so that the maximisation must always result in $\theta_i = 0\ \forall i$. The usual procedure to obtain expected significances or exclusion limits is therefore to perform the likelihood maximisation with the observed data first and to generate the Asimov data with μ =1(0) only afterwards from the nuisance parameter values $\hat{\theta}$ resulting from the fit to the observed data.

Technical Implementation

The data analysis is performed by producing event histograms with the ROOT framework [136]. The likelihood function is then constructed with the help of the HistFactory tool, which is based on the RooFit [137] and RooStats [138] toolkits and performs the minimisation with MINUIT [139].

CHAPTER 5

The $Z \to \tau\tau$ Background Process

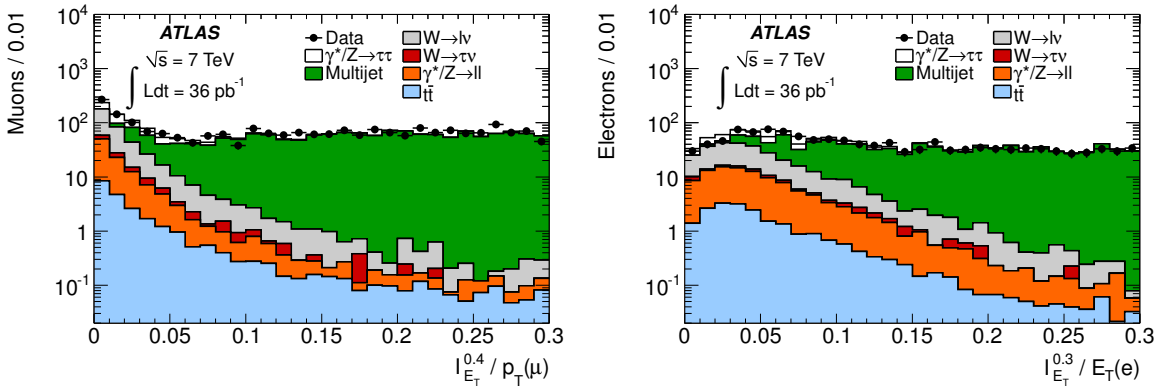
In search for a $H \to \tau\tau$ signal, the dominant background contribution consists of events with similar $\tau\tau$ final states produced in $Z \to \tau\tau$ decays. Hence, the modelling of these events is an essential ingredient for the extraction of the Higgs signal from the collision data. Although this background can be estimated from simulated Monte Carlo events, their modelling must be validated against $Z \to \tau\tau$ events selected from the data. Especially the detector simulation of the τ decays is of special interest in this validation, since this is also relevant for the simulation of the signal process.
In order to perform this validation, also other backgrounds contributions to $\tau\tau$ final states must be estimated. The relevant techniques can thus be developed already for the $Z \to \tau\tau$ analysis and can later be used to estimate the same backgrounds in the context of the $H \to \tau\tau$ search. At the same time, the $Z \to \tau\tau$ process is not only the dominant background, but in terms of detector performance also provides an important handle to control the modelling of $\tau\tau$ final states and all other backgrounds events relevant both for $H \to \tau\tau$ and $Z \to \tau\tau$ final states.

5.1 Measurement of the $Z \to \tau\tau$ Cross Section

As a first step towards the search for $H \to \tau\tau$ decays, the much larger cross section of the $Z \to \tau\tau$ process allows the first observation of $\tau\tau$ final states already with the first collision data. With a dataset corresponding to a total integrated luminosity of about 36 pb^{-1}, which was delivered by the LHC in 2010, the ATLAS collaboration was able to measure the production cross section times branching ratio of this process [140].

5.1.1 Event Selection

For this analysis, events are selected with either single-muon or single-electron triggers, which require $E_T(e) > 15$ GeV for $\mu\tau_h$ final states or $p_T(\mu) > 10 - 13$ GeV for $\mu\tau_{\text{h}}$ depending on the increasing instantaneous luminosity towards the end of the data taking period. The efficiencies of these triggers are obtained with "tag-and-probe" methods from $Z \to \mu\mu$ and $Z \to ee$ events respectively. For muons, the trigger is measured to be 95% efficient in the end-caps, while the coverage around $|\eta| \approx 0$ results in a reduction to 80% in the barrel region. The corresponding electron efficiency is determined to be 96% for low transverse energies $E_T(e) = 16 - 20$ GeV, while the trigger is found to be 99% efficient for electrons with $E_T(e) > 20$ GeV.



(a) (b)

Figure 5.1: Isolation variables (a) $I_{ET}^{0.4}/p_{\mathrm{T}}$ for muon and (b) $I_{ET}^{0.3}/E_{\mathrm{T}}$ for electron candidates, after selecting one hadronic τ candidate and one lepton with opposite signs in $\mu\tau_h$ and $e\tau_h$ final states respectively. The multijet background is estimated from data according to the method described in Sec. 5.1.2; all other processes are estimated using MC simulations [140].

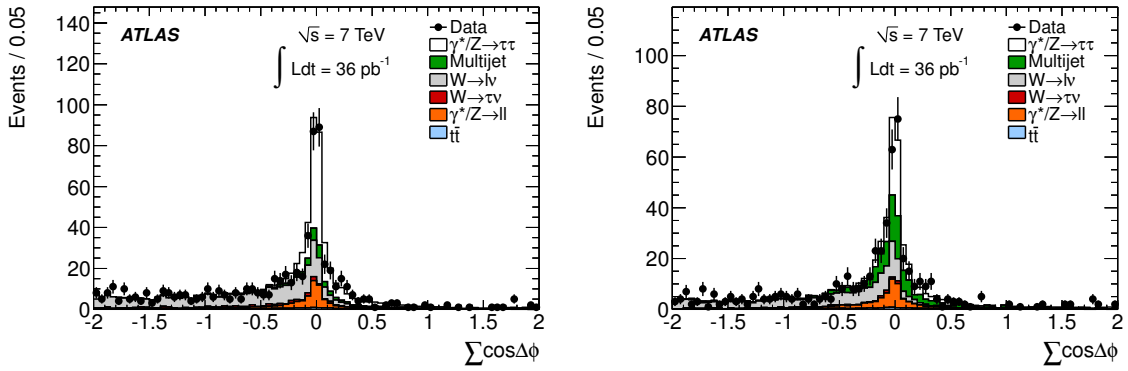
Triggered events are only considered for the analysis if at least one vertex with a minimum of three associated tracks is reconstructed, and standard calorimeter timing requirements can exclude that noise or out-of-time events from cosmic-ray or non-collision backgrounds produce analysis object candidates or contributions to $E_{\mathrm{T}}^{\mathrm{miss}}$. Electron candidates are selected with $E_T > 16$ GeV with `tight` quality requirements leading to a reconstruction and identification efficiency of 73%. Combined muon candidates with $p_T > 15$ GeV and $|\eta| < 2.4$ are selected with a standard selection ensuring good track quality for muons[1] [141] resulting in an efficiency of 92%.

Additionally, isolation criteria are applied for the leptons in order to suppress multijet backgrounds. The corresponding isolation calorimeter energies I_{ET}^{D} and track momenta I_{PT}^{D} are obtained by summing over all cell energies and track momenta in a cone with $\Delta R \leq D$ around the lepton while excluding the associated tracks and clusters. For muons, the fraction of isolation momentum within $\Delta R < 0.4$ is required to be $I_{PT}^{0.4}/p_T < 0.06$ and the corresponding energy fraction $I_{ET}^{0.4}/p_T < 0.06$. For electrons, the same momentum isolation requirement $I_{PT}^{0.4}/E_T < 0.06$ is applied, while the fraction of isolation energy in a smaller cone of $\Delta R < 0.3$ is allowed to be slightly larger with $I_{ET}^{0.3}/E_T < 0.10$. As shown in Fig. 5.1, these selections suppress a large fraction of events from multijet production, while they result in efficiencies of $75 - 98\%$ for muons and $60 - 95\%$ for electrons, depending on their transverse momenta.

Hadronic τ_h candidates with a calibrated transverse energy of $E_T(\tau_{\mathrm{h}}) > 20$ GeV are selected within a range of $|\eta| < 2.47$ excluding the calorimeter transition region ($1.37 < |\eta| < 1.52$). Their identification is based on a cut-based selection [89] and leads to an efficiency of 40% (30%) for real $\tau_{\mathrm{h}}^{1\mathrm{prong}}$ ($\tau_{\mathrm{h}}^{3\mathrm{prong}}$) decays, while the mis-identification probability of jets from multijet production is restricted to 2% (6%) [142]. An improved rejection of fake $\tau_{\mathrm{h}}^{1\mathrm{prong}}$ candidates from electrons is employed, which leaves a remaining mis-identification probability of about 2% for electrons [142].

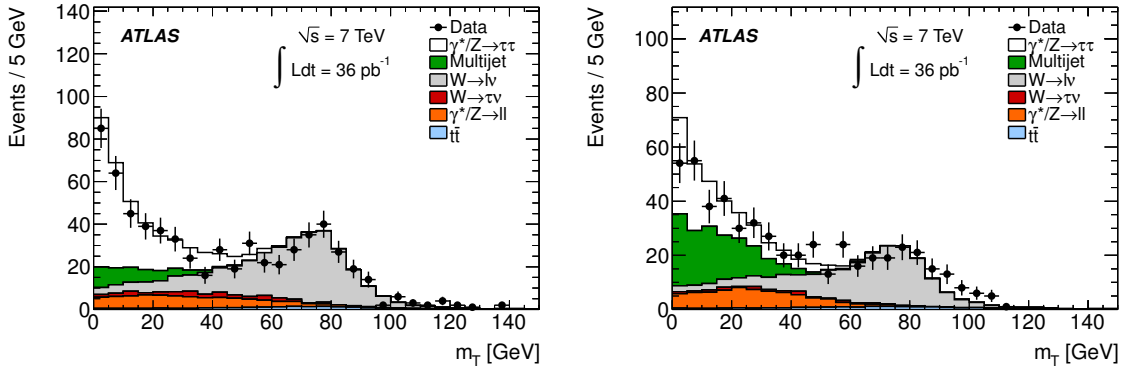
While the background from multijet production is already sufficiently reduced by the lepton isolation

[1] The longitudinal impact parameter of the muon track is additionally required to be less than $|z_0| < 10$ mm with respect to the primary vertex.



(a) $\mu\tau_h$ final state

(b) $e\tau_h$ final state



(c) $\mu\tau_h$ final state

(d) $e\tau_h$ final state

Figure 5.2: Distributions of $\sum \cos\Delta\phi$ shown for the (a) $\mu\tau_h$ and (b) $e\tau_h$ final state and distributions of the transverse mass $m_T(\ell)$ for the (c) $\mu\tau_h$ and (d) $e\tau_h$ final states. All distributions are shown after the object selection for the given final state and after requiring exactly one lepton and a τ_{h} candidate with opposite sign charge. The multijet background is estimated from data according to the method described in Sec. 5.1.2; all other processes are estimated using MC simulations [140].

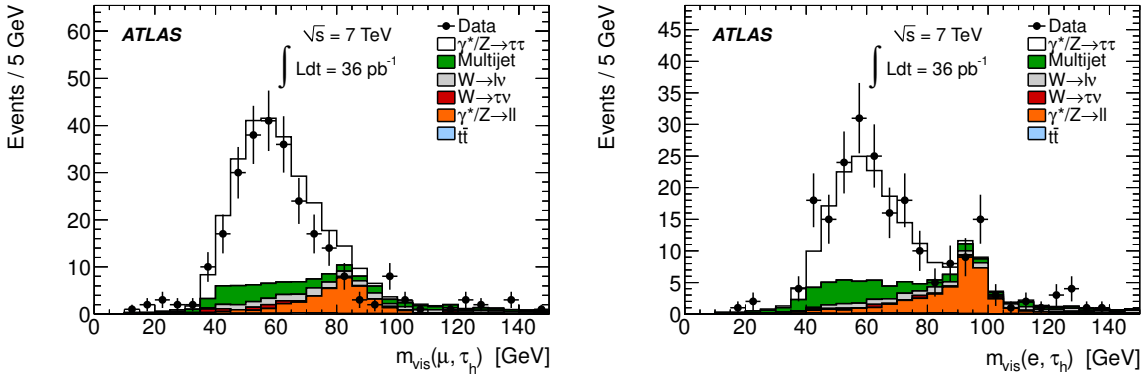
and the requirement of exactly one identified τ_h, further suppression of $Z \to \ell\ell$ background is provided by a veto of events with more than one electron or muon candidate. In order to account for selection inefficiencies due to reconstruction quality requirements, looser criteria are applied to the additional lepton candidates for the veto. As can be seen in Fig. 5.2, further rejection of W+jets background events is obtained by cuts on the transverse mass:

$$m_T = \sqrt{2 \cdot p_T(\ell) \cdot E_{\mathrm{T}}^{\mathrm{miss}} \cdot \left(1 - \cos\Delta\phi\left(\ell, E_{\mathrm{T}}^{\mathrm{miss}}\right)\right)} < 50\ \mathrm{GeV}$$

The event topology of lepton, hadronic τ_{h} candidate and $E_{\mathrm{T}}^{\mathrm{miss}}$ is exploited with an additional selection applied to the observable

$$\sum \cos\Delta\phi = \cos\left(\phi(\ell) - \phi(E_{\mathrm{T}}^{\mathrm{miss}})\right) + \cos\left(\phi(E_{\mathrm{T}}^{\mathrm{miss}}) - \phi(\tau_{\mathrm{h}})\right) > -0.15\ .$$

For the final selection of a sample of $Z \to \tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ events with high purity, only a mass window of 35 GeV $< m_{\mathrm{vis}} <$ 75 GeV is considered, which effectively provides further reduction of $Z \to \ell\ell$



(a) $\mu\tau_h$ final state

(b) $e\tau_h$ final state

Figure 5.3: Distributions of the visible mass $m(\ell, \tau_\mathrm{h})$ for the (a) $\mu\tau_h$ and (b) $e\tau_h$ final states. The distributions are shown after the full event selection, except for the visible mass window requirement. The multijet background is estimated from data according to the method described in Sec. 5.1.2; all other processes are estimated using MC simulations [140].

background events peaking at masses above this threshold and multijet backgrounds with very high or very low mass $m(\ell, \tau_\mathrm{h})$ as shown in Fig. 5.3.

5.1.2 Background Estimation

The background contributions of Z+jets, W+jets and di-boson as well as top-quark production, which enter this event selection, are largely estimated from simulated Monte Carlo samples (cf. Tab. 5.1). The number of multijet events, however, cannot be estimated from simulation, since its contribution to the analysis selection strongly depends on the very small lepton mis-identification probabilities for jets, which are usually not modelled well enough in the detector simulation. Also in terms of computing time and data storage, such a simulation of multijet events cannot be performed efficiently, because a very large number of simulated events would be required, which would contribute only a small fraction of events with mis-identified leptons to the analysis. Therefore, the estimation of the multijet background is the most important ingredient of a data-driven background estimation for $\ell\tau_\mathrm{h}$ final states and has to be determined from a side-band selection in data. In this analysis, the estimation is performed via an extrapolation of contributions from same-sign charge events, which otherwise pass all criteria of the final $Z \to \tau_\mathrm{lep}\tau_\mathrm{had}$ selection. This same-sign selection still includes significant contamination from other processes, which has to be subtracted in order to obtain a correct estimation of the remaining multijet component. For the dominant contribution from W+jets, a correction for mis-modelling effects of jet properties is found to be necessary, which must be propagated into the multijet estimation.

Normalisation Correction for the $W \to \ell\nu$ Background

Although the general features of $W \to \ell\nu$ events can be modelled well by Monte Carlo generators and the detector simulation, especially the τ_h mis-identification probability for jets cannot be predicted well enough, since it depends strongly on phenomenological shower shapes of jets. Instead of relying completely on the simulation of mis-identified $j \to \tau_\mathrm{h}$ candidates, a correction of the predicted W+jets background is therefore derived from a dedicated control region. By inverting both suppression cuts on m_T and $\sum \cos\Delta\phi$, an enriched selection of W+jets events in data is obtained, which is assumed to be

Process	Generator	PDF	$\sigma\times$ BR$\times\epsilon_f$/pb
$Z \to \ell\ell$ ($m_{\ell\ell} > 60$ GeV)	PYTHIA	MRST2007	0.99×10^3
$W \to \ell\nu$	PYTHIA	MRST2007	10.46×10^3
$t\bar{t}$ ($\geq 1\ell$)	MC@NLO + HERWIG	MRST2007	91.5
WW,WZ,ZZ	HERWIG	MRST2007	11.0, 3.4, 1.0

Table 5.1: List of Monte Carlo samples, their generators, used parton-distribution-function and calculated cross sections multiplied by the relevant branching ratios (BR) and filtering efficiencies ϵ_f used for the $Z \to \tau\tau$ analysis. For leptonic decays, the cross section is given per lepton channel.

free of multijet contamination due to the requirement of large $m_{\mathrm{T}} > 50$ GeV. As a difference between opposite-sign charges $\ell^{\pm}\tau_{\mathrm{h}}^{\mp}$ and same-sign $\ell^{\pm}\tau_{\mathrm{h}}^{\pm}$ is expected from the different fractions of quark- and gluon-induced jets, four such control regions are defined: $e\tau_{\mathrm{h}}$ (OS), $\mu\tau_{\mathrm{h}}$ (OS), $e\tau_{\mathrm{h}}$ (SS) and $\mu\tau_{\mathrm{h}}$ (SS).

A comparison of data and simulation in this control region shows a significant overestimation of the W+jets background prediction from simulation. Since this effect is found not to depend on the transverse momentum of the mis-identified τ_{h} candidate, a simple normalisation correction is derived, which extracts correction factors from the event numbers N^{WCR} in each W+jets control region (WCR) according to

$$f_W = \frac{N^{\mathrm{WCR}}_{\mathrm{Data}} - N^{\mathrm{WCR}}_{Z\to\tau\tau} - N^{\mathrm{WCR}}_{Z\to\ell\ell} - N^{\mathrm{WCR}}_{\text{di-boson}} N^{\mathrm{WCR}}_{\mathrm{top}}}{N^{\mathrm{WCR}}_{W\to\ell\nu}} \,. \tag{5.1}$$

The derived factors as presented in Tab. 5.2 are applied to the corresponding W+jets predictions in order to correct the mis-modelled simulation to the event yield observed in data.

	$\mu\tau_h$	$e\tau_h$
opposite-sign	0.73 ± 0.06	0.63 ± 0.07
same-sign	0.94 ± 0.13	0.83 ± 0.15

Table 5.2: The correction factors f_W derived from the opposite-sign and same-sign charge W control regions [140].

Multijet Background from Sideband Extrapolation

When attempting to estimate the contribution $N^{\mathrm{OS}}_{\mathrm{Multijet}}$ of multijet events in the $Z \to \tau\tau$ signal region with opposite-sign charges from the corresponding number $N^{\mathrm{SS}}_{\mathrm{Multijet}}$ of events with same-sign charges, the ratio

$$r_{\mathrm{QCD}} = \frac{N^{\mathrm{OS}}_{\mathrm{Multijet}}}{N^{\mathrm{SS}}_{\mathrm{Multijet}}} \approx 1$$

is expected to be close to unity from considerations on the underlying physics processes. However, since the contribution from certain processes can cause charge asymmetries and thus deviations of this ratio $r_{\mathrm{QCD}} \gtrapprox 1$, in this analysis the corresponding value is extracted from additional control regions with inverted isolation criteria. For these regions, which are again divided into opposite-sign and same-sign charge selections, the number of multijet events is obtained from data after subtracting residual contributions of other backgrounds:

$$N_{\mathrm{Multijet}} = N_{\mathrm{Data}} - f_W \cdot N_{W\to\ell\nu} - N_{Z\to\ell\ell} - N_{t\bar{t}} - N_{\text{Di-boson}} \,. \tag{5.2}$$

The ratio of these multijet event numbers in the anti-isolated regions is assumed to be identical to the ratio in the isolated signal region, so that the extrapolation factor

$$r_{\mathrm{QCD}} \underbrace{=}_{\text{assumption}} r_{\mathrm{QCD}}^{\text{anti-isol}} = N_{\mathrm{Multijet}}^{\text{anti-isol,OS}} / N_{\mathrm{Multijet}}^{\text{anti-isol,SS}}$$

can be applied to the number $N_{\mathrm{Multijet}}^{\mathrm{SS}}$ of multijet events in the isolated region with same-sign charges. As in the anti-isolated regions, this number is obtained by subtracting the residual background contributions via Eqn.(5.2), which leads to the predicted number of background events in the isolated signal region with opposite-sign charges

$$N_{\mathrm{Multijet}}^{\mathrm{OS}} = r_{\mathrm{QCD}} \cdot N_{\mathrm{Multijet}}^{\mathrm{SS}} \underbrace{=}_{\text{assumption}} \frac{N_{\mathrm{Multijet}}^{\text{anti-isol,OS}}}{N_{\mathrm{Multijet}}^{\text{anti-isol,SS}}} \cdot N_{\mathrm{Multijet}}^{\mathrm{SS}} \,.$$

From the observed data events in the corresponding regions, the following numbers are obtained:

$$r_{\mathrm{QCD}} = \begin{cases} 1.07 \pm 0.07\ (\mathrm{stat}) \pm 0.07\ (\mathrm{syst}) \text{ for } e\tau_{\mathrm{h}} \text{ final states} \\ 1.07 \pm 0.04\ (\mathrm{stat}) \pm 0.04\ (\mathrm{syst}) \text{ for } \mu\tau_{\mathrm{h}} \text{ final states} \end{cases}$$

These ratios are expected to be independent not only of the lepton isolation, but also of other kinematic distributions such as the di-τ mass. Therefore, this background estimation of $N_{\mathrm{Multijet}}^{\mathrm{OS}}$ from $N_{\mathrm{Multijet}}^{\mathrm{OS}}$ can be applied separately for each bin of the distribution, which effectively extracts both shape and normalisation of the multijet background from the isolated side-band region with same-sign charges.

Background Summary

The composition of background contributions and $Z \to \tau\tau$ signal for the described analysis selection are summarised in Table 5.3. In the $\mu\tau_h$ channel[2], an expected signal of 186 ± 2 events and a total sum of 49 ± 6 background events are estimated, which is in statistical agreement with 213 events observed in the data. Similarly, 151 observed events in the $e\tau_h$ channel are consistent with the expected 98 ± 6 signal and 37 ± 6 background events. All quoted uncertainties include only statistical uncertainties related to a limited number of generated Monte Carlo events or data events in control regions.

5.1.3 Cross Section Calculation

The cross section measurement from a dataset corresponding to an integrated luminosity $\mathcal{L} = \int L\mathrm{d}t$ with an observed number N_{obs} of data events and an expected number N_{bkg} needs to account for acceptance and efficiency effects according to

$$\sigma(pp \to Z \to \tau\tau) \times \mathrm{BR}(\tau\tau \to \ell\tau_{\mathrm{h}} + 3\nu) = \frac{N_{\mathrm{obs}} - N_{\mathrm{bkg}}}{A_Z \cdot C_Z \cdot \mathcal{L}} \,. \tag{5.3}$$

This yields the cross section times branching ratio $\sigma \times \mathrm{BR}$ of the final state under consideration and corrects the observed number of events with the reconstruction efficiency C_Z and the acceptance A_Z, which is determined by the event selection applied for the measurement. While C_Z is specific to the detector components and the employed reconstruction algorithms, the acceptance A_Z is independent of

[2] In order to simplify the notation, channels are denoted only by their visible τ decay products, while the additionally produced neutrinos are suppressed from here on.

	$\mu\tau_h$	$e\tau_h$
$\gamma^*/Z \to \ell\ell$	11.1 ± 0.5	6.9 ± 0.4
$W \to \ell\nu$	9.3 ± 0.7	4.8 ± 0.4
$W \to \tau\nu$	3.6 ± 0.8	1.5 ± 0.4
$t\bar{t}$	1.3 ± 0.1	1.02 ± 0.08
Diboson	0.28 ± 0.02	0.18 ± 0.01
Multijet	24 ± 6	23 ± 6
$\gamma^*/Z \to \tau\tau$	186 ± 2	98 ± 1
Total expected events	235 ± 6	135 ± 6
N_{obs}	213	151

Table 5.3: Number of events expected for the different background processes and observed in the data after the full selection. The background estimates have been obtained as described in Sec. 5.1.2. The quoted uncertainties are statistical only [140].

the specific detector response entering the measurement and is only driven by the detector geometry. The corresponding extrapolation from the accepted to the full phase space can thus be determined from Monte Carlo events without detector simulation or even from analytical calculations for the specific process, in which the phase space can be restricted to the geometrical detector acceptance. Without this acceptance correction, a fiducial cross section can be defined via

$$\sigma^{\text{fid}}(pp \to Z \to \tau\tau) \times \text{BR}(\tau\tau \to \ell\tau_{\text{h}} + 3\nu) = \frac{N_{\text{obs}} - N_{\text{bkg}}}{C_Z \cdot \mathcal{L}} \,, \tag{5.4}$$

which mainly contains experimental uncertainties and is valid in a fiducial phase space without any theory-based extrapolation and associated uncertainties. In this analysis, the selection entering A_Z contains the W+jets suppression cuts and the mass window as well as the momentum thresholds and excluded η regions as summarised in Tab. 5.4.

	$\mu\tau_h$ channel	$e\tau_h$ channel
Lepton	$p_{\text{T}} > 15$ GeV, $\lvert\eta\rvert < 2.4$	$E_{\text{T}} > 16$ GeV, $\lvert\eta\rvert < 2.47$, excluding $1.37 < \lvert\eta\rvert < 1.52$
Tau	$p_{\text{T}} > 20$ GeV, $\lvert\eta\rvert < 2.47$, excluding $1.37 < \lvert\eta\rvert < 1.52$	
Event	$\Sigma \cos\Delta\phi > -0.15, m_{\text{T}} < 50$ GeV, m_{vis} within [35, 75] GeV	

Table 5.4: Definition of the fiducial phase space for the acceptance correction with A_Z [140].

5.1.4 Systematic Uncertainties

Besides the statistical uncertainty of the observed data, systematic uncertainties must be evaluated for the background estimates, which largely rely either on the detector simulation or on data-driven background estimation techniques. According to dedicated performance measurements, momentum and energy scales and resolutions as well as object identification efficiencies are thus varied within their measured uncertainties, and the resulting effects on the analysis are estimated. For τ_{h} candidates, the

energy scale uncertainty is assumed to be fully correlated with the jet energy scale, while the effects of all energy and momentum scale uncertainties are propagated to the calculated $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ and thus the related variables $m_T(\ell)$ and $\sum \cos \Delta\phi$.

For the backgrounds estimated from data, all these variations are propagated through the complete background estimation procedure with residual contributions from simulation, so that their effects on the data-driven estimates are taken into account. In addition, a possible systematic effect introduced by the side-band extrapolation is estimated by varying the isolation criteria used to define the anti-isolated control regions.

For the theory-based extrapolation via the acceptance A_Z, the systematic uncertainty on the modelling of the $Z \to \tau\tau$ signal production is estimated by substituting the default Monte Carlo sample generated with PYTHIA by a similar sample from MC@NLO + HERWIG. Different sets of parton distributions provided for CTEQ6.6 [129] are used to derive corresponding systematic uncertainties in the context of this analysis. All estimated uncertainties are summarised in Tab. 5.5, in which the dominant uncertainties can be identified to be related to the τ_{h} identification efficiency as well as to the different energy scales for electrons, τ_{h} candidates and jets.

Systematic uncertainty	$\mu\tau_h$ channel	$e\tau_h$ channel
Muon efficiency	3.8%	–
Muon resolution & energy scale	0.2%	–
Electron efficiency, resolution &		
Charge mis-identification	–	9.6%
τ_h identification efficiency	8.6%	8.6%
τ_h mis-identification	1.1%	0.7%
Energy scale (e/τ/jets/$E_{\mathrm{T}}^{\mathrm{miss}}$)	10%	11%
Multijet estimation method	0.8%	2%
W+jets normalization factor	0.1%	0.2%
Object quality selection criteria	1.9%	1.9%
Pile-up description in simulation	0.4%	0.4%
Theoretical cross section	0.2%	0.1%
A_Z systematics	3%	3%
Total Systematic uncertainty	15%	17%
Statistical uncertainty	9.8%	12%
Luminosity	3.4%	3.4%

Table 5.5: Relative statistical and systematic uncertainties in % on the total cross section measurement. The electron and muon efficiency terms include the lepton trigger, reconstruction, identification and isolation uncertainties [140].

5.1.5 Results

From the observed events and the estimated background contributions in combination with the correction factors and integrated luminosities as summarised in Tab. 5.6, the cross section for the $Z \to \tau\tau$ production process is calculated separately for the $\mu\tau_h$ and the $e\tau_h$ channel. Besides the fiducial cross sections shown in Tab. 5.7, the total cross sections are estimated as

$$\sigma(pp \to Z \to \tau\tau) = 0.86 \pm 0.08\,(\text{stat}) \pm 0.12\,(\text{syst}) \pm 0.03\,(\text{lumi})\ \text{nb} \quad (\mu\tau_h)\,,$$

$$\sigma(pp \to Z \to \tau\tau) = 1.14 \pm 0.14\,(\text{stat}) \pm 0.20\,(\text{syst}) \pm 0.04\,(\text{lumi})\ \text{nb} \quad (e\tau_h)\,.$$

These separate values are in agreement with the combination of all $\tau\tau$ final states including $e\mu(+4\nu)$ and $\mu\mu(+4\nu)$:

$$\sigma(pp \to Z \to \tau\tau) = 0.97 \pm 0.4\,(\text{stat}) \pm 0.3\,(\text{syst}) \pm 0.2\,(\text{lumi})\ \text{nb} \quad (\mu\tau_h \oplus e\tau_h \oplus e\mu \oplus \mu\mu)\,.$$

Further comparisons to the theoretical expectation of $\sigma = 0.95 \pm 0.05$ nb as well as other cross section measurements of $Z \to ee$ and $Z \to \mu\mu$ [76] show a similar agreement (cf. Fig. 5.4). Although these other final states provide a more precise measurement of the Z production cross section, this $Z \to \tau\tau$ analysis can confirm that both the simulated $Z \to \tau\tau$ signal and the data-driven background estimates for $\tau_{\text{lep}}\tau_{\text{had}}$ final states provide a good description of the data. It is thus used as the basis for further physics measurements involving τ-leptons, and provides a first important contribution to the search for $H \to \tau\tau$ decays.

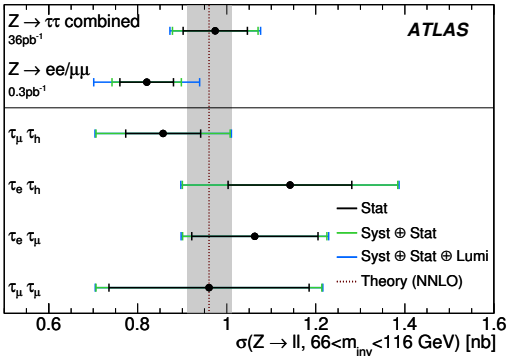


Figure 5.4: The individual cross section measurements by final state and the combined result [140]. The $Z \to \ell\ell$ combined cross section measured by ATLAS in the $Z \to \mu\mu$ and $Z \to ee$ final states is also shown for comparison. The gray band indicates the uncertainty on the NNLO cross section prediction.

	$\mu\tau_h$	$e\tau_h$
N_{obs}	213	151
$N_{\text{obs}} - N_{\text{bkg}}$	$164 \pm 16 \pm 4$	$114 \pm 14 \pm 3$
A_Z	0.117 ± 0.004	0.101 ± 0.003
C_Z	0.20 ± 0.03	0.12 ± 0.02
BR	0.2250 ± 0.0009	0.2313 ± 0.0009 [143]
$\mathcal{L}$	35.5 ± 1.2 pb^{-1}	35.7 ± 1.2 pb^{-1}

Table 5.6: The components of the $Z \to \tau\tau$ cross section calculations for the two final states [140]. For $N_{\text{obs}} - N_{\text{bkg}}$, the first uncertainty is statistical and the second systematic. For all other values the total error is given.

Final State	Fiducial cross section (pb)
$\mu\tau_h$	$23 \pm 2 \pm 3 \pm 1$
$e\tau_h$	$27 \pm 3 \pm 5 \pm 1$
Final State	Total cross section ([66, 116] GeV) (nb)
$\mu\tau_h$	$0.86 \pm 0.08 \pm 0.12 \pm 0.03$
$e\tau_h$	$1.14 \pm 0.14 \pm 0.20 \pm 0.04$

Table 5.7: The production cross section times branching fraction for the $Z \to \tau\tau$ process as measured in each of the two final states and their combination [140]. For the fiducial cross sections, the measurements include also the branching fraction of the τ to its decay products. The first error is statistical, the second systematic and the third corresponds to the uncertainty from the luminosity measurement.

5.2 Modelling of $Z \to \tau\tau$ from $Z \to \mu\mu$ Data

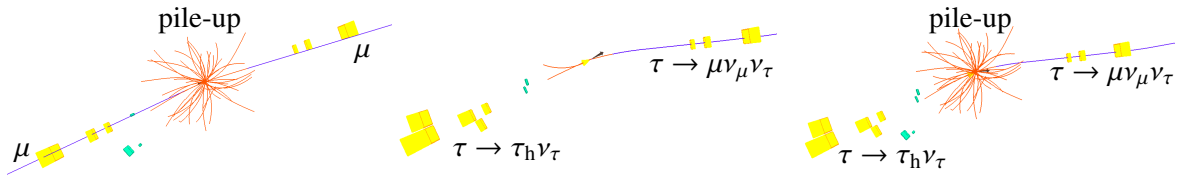
Although the measurement of the $Z \to \tau\tau$ cross section mainly validates the simulation of τ decays for the $H \to \tau\tau$ search, many important observables such as the MMC mass, the transverse mass or the di-jet mass rely on other event features such as the $E_{\mathrm{T}}^{\mathrm{miss}}$ and additional jet activity. These are known to suffer from mis-modelling effects both in the Monte Carlo event generation[3] and in the detector simulation, so that a data-driven estimate of this background is strongly preferred over a purely simulation based estimation. Although the $Z \to \tau\tau$ analysis already performs a selection of these events from data, any such selection implies a conceptual problem for the $H \to \tau\tau$ search: while especially the tail of the Z mass distribution is of main interest as dominant background to Higgs boson decays, exactly this region cannot be extracted from data without a significant signal contamination from $H \to \tau\tau$ final states. When assuming the existence of $H \to \tau\tau$ decays, this will inevitably lead to a mixture of signal and background in the phase space of interest, which makes the direct data-driven estimation of $\tau\tau$ backgrounds basically impossible. In order to avoid problems of mis-modelings either in perturbative QCD calculations or with detector response and reconstruction in simulated events, a different selection of $Z \to \mu\mu$ data is used instead. These events can effectively be extracted signal-free, since the $H \to \mu\mu$ branching ratio[4] is smaller by a factor $m_\mu^2/m_\tau^2 \approx 3.6 \times 10^{-3}$. With a technical trick, these events are converted into $Z \to \tau\tau$ final states by removing the muon objects from the data and replacing them with two simulated τ decays. After this method has been developed on Monte Carlo simulations already before the start of the data taking period at the LHC [144, 145], several improvements of the muon cell subtraction procedure, validation techniques and the evaluation systematic uncertainties of the method were developed during the first application of this technique to real $Z \to \mu\mu$ events collected in pp-collisions with ATLAS.

5.2.1 General Idea of the Embedding Method

The correct modelling of the complex $\tau\tau$ final states is essential for a robust estimation of the corresponding background contributions. Reconstructed objects such as jets and $E_{\mathrm{T}}^{\mathrm{miss}}$, which can potentially suffer from mis-modelling in various ways, should therefore either be validated carefully against or directly taken from data. Since any mis-modelling in simulated events would require corrections to be derived in control regions, the data-driven approach is followed with the embedding technique. This starts from real data events and only introduces relatively small modifications by replacing $\mu\mu$ final states with $\tau\tau$ decays. Many important experimental effects such as pile-up and underlying event contributions to soft energy depositions in the calorimeters are thus directly obtained from the data and enter the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction without any uncertainty related to simulation effects. In addition, all jets from initial-state-radiation as well as the corresponding recoiling momentum $p_T(Z)$ of the Z boson are unchanged by the technique. As a consequence, neither theoretical uncertainties nor the jet energy scale need to be considered for these events. Nevertheless, the products of the τ decays cannot be extracted from data and therefore must be produced in simulations with identical kinematics corresponding to the initial $\mu\mu$ final states. Since these $\tau\tau$ final states must be embedded into a data environment later on, their simulation should be free of pile-up or noise effects, so that only little modelling issues are expected for their clean τ decay products. The clean removal of the original muons from the selected data events, however, introduces the largest modification for these events by the embedding procedure, because the

[3] The additional jet activity can also affect the transverse momentum $p_T(Z)$ of the Z boson and thus also $p_{\mathrm{T}}^{\mathrm{miss}}$ from the neutrinos, which are produced in the subsequent decays.

[4] Also the $H \to \tau\tau \to \mu\mu$ decay channel contributes similar final states, which are suppressed by a branching ratio of 3% and have a significantly lower invariant mass, so that a possible contamination is expected to be negligible.

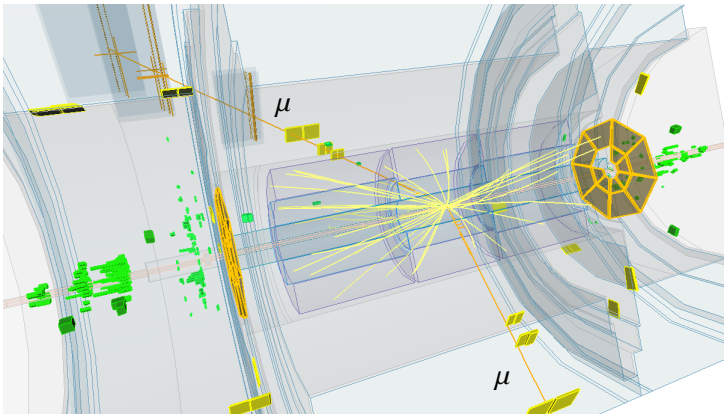


(a) (b) (c)

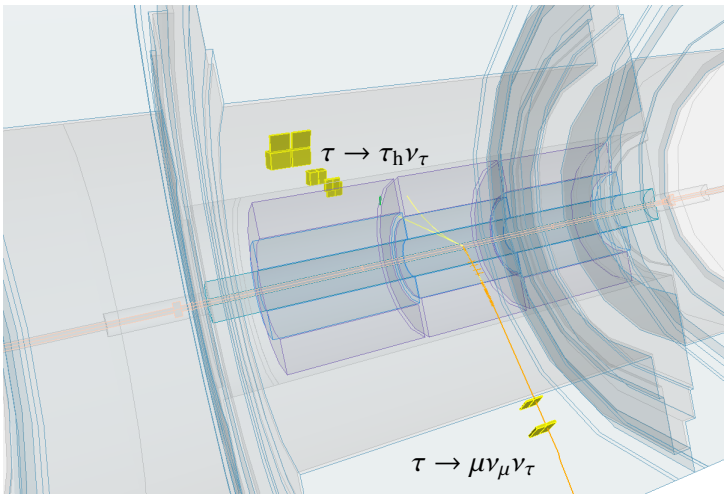
Figure 5.5: The embedding method applied to an example event as seen in the (r, ϕ)-plane. Starting from a $Z \to \mu\mu$ decay selected from data (a), a $Z \to \tau\tau \to \mu\tau_h 3\nu$ decay with identical kinematics but no pile-up (b) is simulated and then merged with the original data event (c), from which the muons have been removed beforehand.

energy loss of the muons deposits energy clusters in the calorimeters, which must be removed.

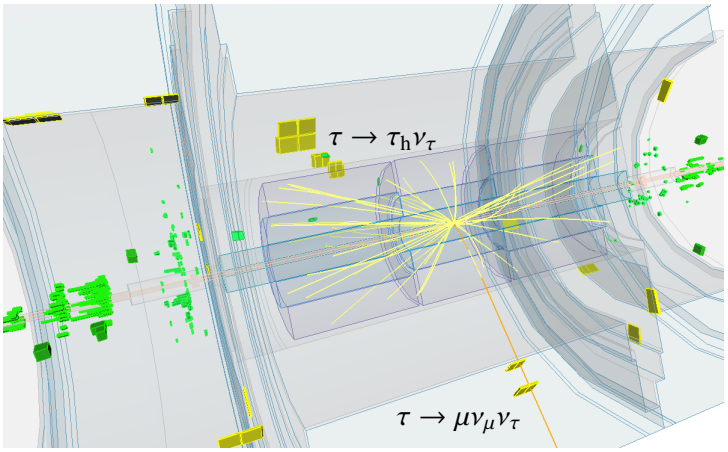
Although the general idea of replacing objects within a set of selected events as visualised in Fig. 5.5-5.6 is conceptually simple, the whole method depends strongly on the reconstruction, which is used to convert either data or simulated detector responses into objects. The corresponding algorithms must therefore be compatible with a partial re-reconstruction of objects from an artificially reduced set of event information. Especially the removal of muon signatures as well as issues with alignment between the real detector and its software model disfavour an implementation of a low-level embedding implementation, so that the replacement is performed with reconstructed tracks instead of starting from the basic silicon hit information. The final hybrid $Z \to \tau\tau$ events then provide a dataset, which is obtained in a data-driven way, but with certain inefficiencies due to muon trigger, acceptance and reconstruction effects. The absolute normalisation of this dataset is therefore not extracted directly from the $Z \to \mu\mu$ data, but rather obtained either from a $Z \to \tau\tau$ control region or from a normalisation factor calculated with the help of a simulated Monte Carlo sample.



(a)



(b)



(c)

Figure 5.6: The embedding method applied to an example event in a schematic view of the ATLAS detector. Starting from a $Z \to \mu\mu$ decay selected from data (a), a $Z \to \tau\tau \to \mu\tau_h 3\nu$ decay with identical kinematics but no pile-up (b) is simulated and then merged with the original data event (c), from which the muons have been removed beforehand.

5.2.2 Event Selection

The selection of $Z \to \mu\mu$ data events is performed with a combination of a single muon trigger with a threshold of $p_T(\mu) > 24$ GeV and a di-muon trigger requiring $p_T(\mu) > 18$ (8) GeV for the leading (sub-leading) muon. In order to allow a proper replacement of all objects, the selected muons must be combined with the Staco-algorithm from MS and ID measurements with standard track quality criteria. The final event selection requires at least two such muons with $p_T(\mu) > 20$ (15) GeV, $|\eta(\mu)| < 2.5$ and a track isolation momentum of $I_{PT}^{0.2}/p_T < 0.2$. A selected pair of muons must then have opposite-sign charges, and both muons must originate from the same reconstructed vertex[5]. Low-mass events far away from the Z pole are then suppressed by an invariant mass cut of $m(\mu,\mu) > 40$ (60) GeV in 2012 (2011) data, while in the rare cases, in which more than one such pair is found in an event, the combination closest to the expected mass m_Z is taken as the kinematic basis for the embedding procedure.

5.2.3 Generation of $Z \to \tau\tau$ Decays

From each selected data event, the kinematics of the $Z \to \mu\mu$ decay are extracted and converted into a corresponding $Z \to \tau\tau$ decay as input for the new event simulation. Therefore, the four-momentum of the Z boson is reconstructed, and the momenta of the two muons are first boosted into its rest frame. The muon momenta are then rescaled via

$$p(\tau) = \sqrt{E(\mu)^2 - m_\tau^2}$$

to account for the effect of the rest mass m_τ of the τ-leptons on the final kinematics, before the modified τ decay kinematics are boosted back into the Z rest frame.

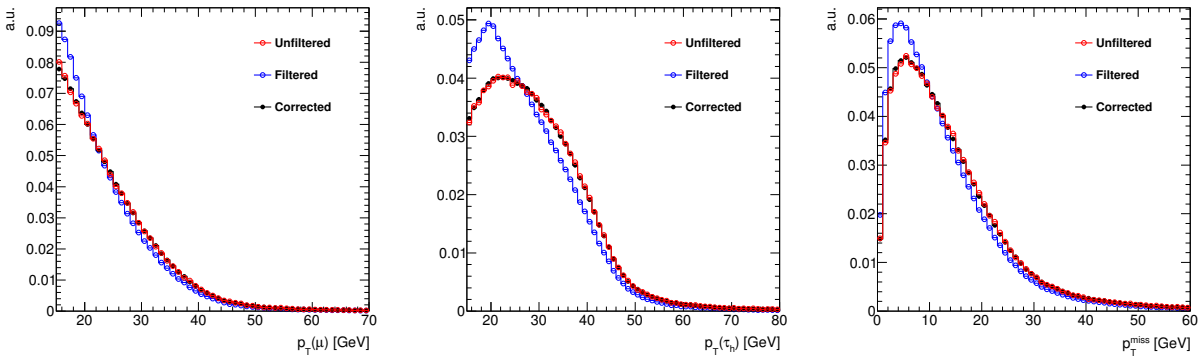
For each $Z \to \mu\mu$ data event, these rescaled $Z \to \tau\tau$ kinematics are stored in HepEvt format (as proposed in [146]) and then processed by the TAUOLA program, which generates the decays of the τ-leptons on a random basis for a given $\tau\tau$ decay channel (e.g. $\tau_{\text{lep}}\tau_{\text{had}}$). While the determination of the spin of each Z boson would require knowledge about its production from quarks, which cannot be determined from data events, TAUOLA assumes an average Z polarisation of $\langle h_z \rangle = 0$ by randomly assigning a helicity of $h_z = +1, -1$ to each Z boson and propagating this to the spin correlation of the two τ-leptons and their sub-sequent decays. Although the correct polarisation in Z production is found to be $\langle h_z \rangle = 0.15$, the effect of this simplified assumption is found to be negligible in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel[6].

5.2.4 Kinematic Filter

In the two- or three-body-decays of the τ-leptons, the visible decay products usually obtain relatively low transverse momenta, so that about $60-70\%$ of the generated decays cannot be expected to enter the later analysis due to the applied momentum thresholds. This kinematic feature of the $\tau^+\tau^-$ final states would effectively reduce the number of embedded $Z \to \mu\mu$ events from data only due to the topologies of randomly generated τ decays, which are otherwise independent from the Z kinematics. The latest embedding implementation therefore includes a kinematic filter, which can recover a large fraction of

[5] For both tracks, the same vertex must have the highest matching probability, which is then considered as the position of the $Z \to \mu\mu$ decay in the detector.

[6] As a new method, the TauSpinner [147] implementation in TAUOLA assigns an improved polarisation, which can calculate the probability for the Z spin direction from the accessible $Z \to \tau\tau$ kinematics. A corresponding re-weighting scheme has been tested in the $\tau_{\text{had}}\tau_{\text{had}}$ channel, where spin effects are expected to be the largest. The resulting correction of kinematic distributions was found to be small, so that only negligible effects on $\tau_{\text{lep}}\tau_{\text{had}}$ are expected.



(a) (b) (c)

Figure 5.7: Effect of the kinematic filter on the transverse momentum of the lepton (a), the hadronic τ_h (b) and the missing transverse momentum (c) as the vector sum of the neutrino momenta. The filter biases these kinematics distributions, which are then corrected back by applying the corresponding weights from the event-by-event filter efficiency.

these events by requiring sufficiently high transverse momenta for the visible τ decay products:

$$p_T(e) > 18 \text{ GeV} , \qquad p_T(\mu) > 15 \text{ GeV} , \qquad p_T(\tau_h) > 15 \text{ GeV} .$$

For each $Z \to \mu\mu$ event, 1000 different τ decays are generated with TAUOLA, while the first $\tau\tau$ final state fulfilling the filter requirements is kept for further processing[7]. For the overall dataset, this kinematic preselection introduces a bias, which favours certain helicity and especially low p_T^{miss} configurations (cf. Fig. 5.7), so that a correction is necessary to restore the original kinematic distributions in the embedded data. Therefore, the fraction of accepted $\tau\tau$ decay configurations is used as an event weight, which corresponds to the filter efficiency for the specific $Z \to \tau\tau$ event configuration. This filter weight is later applied to the embedded event, so that $Z \to \tau\tau$ events, which are unlikely to be accepted and thus enhanced more strongly, are weighted down accordingly. Since the weight corresponds to the probability that a random $\tau_{lep}\tau_{had}$ configuration for the specific $Z \to \tau\tau$ event is accepted by the filter[8], the bias is corrected as shown in Fig. 5.7, while the statistical uncertainty for the later analysis is reduced by a factor of approximately 1.7 due to an almost three times larger number of weighted events.

5.2.5 Simulation of $Z \to \tau\tau$ Decays

After the $Z \to \tau_{lep}\tau_{had}$ decay kinematics are generated, a full detector simulation and object reconstruction is performed for each event. For this simulation, special care needs to be taken that certain effects are not double counted when embedding the objects from the $Z \to \tau\tau$ decay into the original $Z \to \mu\mu$ data event. Starting from the position of the reconstructed $Z \to \mu\mu$ decay, the corresponding $Z \to \tau\tau$ event is placed into the detector at the reconstructed interaction point from the original event. The default smearing of the simulated interaction point, which is necessary for Monte Carlo generated events,

[7] In rare cases such as configurations, where already the initial τ-leptons have too low transverse momenta, none of the 1000 decay configurations can be accepted by the filter. The last generated decay is then kept for further processing, although it is expected to fail the later analysis selection.

[8] Events without any accepted decay configuration are not important, since they cannot enter the final analysis selection. Their weight is nonetheless set to 0.

must therefore be disabled, since the position is already determined within the resolution of the data. As pile-up effects are already present in the original data events, their simulation must as well be excluded from the default procedure for Monte Carlo events. Similarly, calorimeter noise is present in the original data event, so that no additional noise must be generated for the cells in the detector simulation. Only for the reconstruction of tracks, hit-level noise in the tracking system must be included in order to obtain comparable tracking performance as in other simulations and in the data[9]. With an otherwise similar setup as used in the standard Monte Carlo production, this $Z \to \tau\tau$ simulation produces clean events, which only contain the detector response and reconstructed objects from the decay products of the τ-leptons.

5.2.6 Replacement of Tracks and Cells

In order to merge each cleanly simulated $Z \to \tau\tau$ event with its corresponding $Z \to \mu\mu$ data event, only the inner detector and muon spectrometer tracks of the original muons are identified and removed from the data event, while all other tracks from the underlying event or secondary vertices remain untouched in the data. From the clean $Z \to \tau\tau$ event, all tracks resulting from the τ decays are then copied into the original data event. Although this procedure cannot account for limitations of the two-track resolution, which would result in a merging of overlapping tracks from the τ decay and pile-up or underlying event, such cases are expected to occur very rarely due to the requirements on the track quality and isolation both in the $Z \to \mu\mu$ data before embedding and in the hybrid $Z \to \tau\tau$ sample when used in the final analysis.

While the removal of the original muon tracks is technically straightforward, also their energy deposition in the calorimeter cells must be determined and subtracted before any contribution from the simulated τ decays can be added. This deposition of muon calorimeter energy is thus determined for each event from another clean simulation of $Z \to \mu\mu$ with identical kinematics, which produces a set of cell energies corresponding to the calorimeter response for the identical $Z \to \mu\mu$ event as expected from simulation. For each calorimeter cell in the data, only this simulated amount of muon calorimeter energy is subtracted. It is clear though that the energy loss of muons in the calorimeter occurs as a statistical process and that the amount as well as the shape of the energy deposit can vary statistically and on an event-by-event basis between a muon in data and its random detector simulation. Still, this statistical variance is expected to be small enough to vanish in calorimeter noise effects during clustering, while extreme cases with large discrepancies (i.e. the subtracted energy is very different from the actual energy deposit - either in amount or shape) will be rejected by isolation and shower shape requirements in the final analysis.

After this removal of muon cell energy, the contribution from the τ decays is added to the data. The energy of each cell in the clean $Z \to \tau\tau$ event is therefore added to the corresponding cell in the data event, while all the other cells in detector regions without $Z \to \tau\tau$ activity remain untouched. Compared to earlier embedding implementations as described in [144, 145], this method does not depend any further on ΔR-cones, which were originally defined around the muon and τ-lepton axes, so that all calorimeter energy in a small cone around the muon was removed from the data event, and only the energy in a larger cone around the τ-lepton[10] was copied from the simulated $Z \to \tau\tau$ event into the data. The improved method now does not artificially restrict the copied cells to pre-defined cones, but instead

[9] If the embedding procedure operated on hit-level, noise in the tracking system would not need to be simulated. However, additional problems might arise when removing hits from muon tracks, which can remove noise as well and might require compensation.

[10] Although the cone radii were different, both cone axes were chosen identically, since the direction of muon and τ-lepton is identical by definition.

includes all cells with energy depositions while relying on the fact that empty cells simulated without noise contain an energy equivalent of 0 and effectively do not contribute[11]. Hence, problems are avoided when energy is deposited outside the larger cone and thus not embedded into the data, which can occur in cases of large separations $\Delta R(\tau, \tau_{\mathrm{h}})$ or $\Delta R(\tau, \ell)$ and could potentially even create biases for certain helicity configurations of the τ-leptons. While the previous implementation of the muon subtraction removed all energy in a small cone, this inevitably affected energy contributions from noise and pile-up as well and did not involve any estimate of the actual energy deposition from the muon. With the new method, such an estimate is provided by an additional clean simulation, which now allows a well motivated procedure to evaluate systematic effects due to the subtraction.

5.2.7 Reconstruction of the Hybrid Event

Each set of tracks and calorimeter cells for an event, which corresponds to a simulated $Z \to \tau\tau$ decay with a data environment obtained from $Z \to \mu\mu$, is finally re-reconstructed with the standard ATLAS software, so that all analysis objects are affected by reconstruction effects as expected for real $Z \to \tau\tau$ data events. Although starting only from tracks instead of hits, the vertex finding algorithms are re-run, so that the difference between the track momenta in $Z \to \tau\tau$ and $Z \to \mu\mu$ events is taken into account for the vertex reconstruction and track association. All calorimeter clusters are re-built from the modified set of cell energies, and all electron, muon and τ_{h} objects are newly reconstructed including track, isolation and shower shape variables responsible for their identification. From all these objects and all remaining clusters without association to any object, $E_{\mathrm{T}}^{\mathrm{miss}}$ is finally re-built, so that the full embedding procedure results in hybrid $Z \to \tau\tau$ events. These are essentially obtained from data and only have to rely on simulation for the pure τ decay products, which conceptually cannot be avoided. Especially the re-reconstruction of hadronic τ decays and $E_{\mathrm{T}}^{\mathrm{miss}}$ is of great importance, since these depend strongly on the pile-up environment in data and cannot be estimated with simpler methods such as adding the $p_{\mathrm{T}}^{\mathrm{miss}}$ from neutrinos or copying the already reconstructed τ_{h} objects from the cleanly simulated $Z \to \tau\tau$ event.

5.2.8 Estimation of Systematic Effects

Although the estimation of the $Z \to \tau\tau$ background with the embedding method is largely data driven and jet- or $E_{\mathrm{T}}^{\mathrm{miss}}$-related systematic uncertainties are therefore obsolete for the embedded events, the method to replace the objects can still introduce systematic effects, which need to be evaluated. All simulated objects are subject to systematic uncertainties from efficiencies, energy scales and resolutions as the default Monte Carlo samples. Additional uncertainties arise for the embedded events due to the event selection and the muon cell energy subtraction as implemented in the embedding procedure.

The selection of $Z \to \mu\mu$ events introduces two separate systematic effects, which cannot easily be disentangled. The background contamination of the selected $Z \to \mu\mu$ events is dominated by mis-identified leptons from multijet and W+jets events, which can be reduced with tighter isolation criteria applied to the original muons. This isolation, however, affects the environment of the selected muons and thus propagates into the final hybrid events, resulting in a bias towards better isolation of embedded objects as compared to the expectation for real $Z \to \tau\tau$ final states. Therefore, the isolation criteria are varied in order to estimate the effects related both to background contamination and isolation bias of embedded objects, where one variation does not require any isolation criterion and thus includes a

[11] Technically, cells are even compactified in the data model and are thus not even processed when empty.

larger background contamination but less isolation bias. The other applies the tighter selection of

$$I_{PT}^{0.4}/p_T < 0.06\,, \qquad I_{ET}^{0.2}/p_T < 0.04$$

on the original muons before embedding.

The impact of the muon cell subtraction method on the embedded events is crucial especially for the energy scale of the embedded τ_h objects, while it can also affect the isolation or shower shape variables of the other objects as well as $E_\mathrm{T}^\mathrm{miss}$. Since the amount of energy subtracted in the procedure is determined from simulation, it is not clear whether this interaction of muons in the calorimeter is modelled correctly or whether it might introduce a systematic effect due to mis-modelling of the energy depositions. The subtracted amount of energy is therefore varied up and down by scaling each cell contribution ±20%, which is chosen as a very conservative estimate of a possible mis-modelling of muon interaction in the calorimeter. Although this variation does not account for possible shape differences in the simulated energy depositions, the size of the variation is expected to provide a conservative envelope of the possible systematic effects on the embedded $Z \to \tau\tau$ events.

5.2.9 Validation of Embedded Data

In order to validate the modelling of the hybrid $Z \to \tau\tau$ events obtained from embedding the τ decays into $Z \to \mu\mu$ data, several aspects need to be considered.

As a simple technical modification of the method, the removed muons in data can be replaced by simulated muons, which are then embedded within the same framework. Besides validating its technical implementation, comparisons of kinematic distributions before and after applying this muon-embedding variation are able to show that the subtraction of the muon cell energy performs provides reasonable results and that the track reconstruction does not introduce any significant bias into the spectra.

The validation of the embedded dataset by comparison to Monte Carlo simulated datasets is another important aspect of the validation, which can prove that basic kinematic distributions such as decay angles or basic transverse momenta are modelled correctly and reproduced with the embedding method. The transverse momentum of τ_h candidates, however, can depend on the pile-up conditions, since any additional energy in the calorimeter can wrongly increase the reconstructed $p_T(\tau_\mathrm{h})$. While this basic kinematic contribution is thus not expected to be modelled well by simulation, the lepton transverse momenta $p_T(\ell)$ are in general more reliable. Although isolation requirements are usually defined relative to $p_T(\ell)$ and can lead to biases for leptons as well, these are expected to be smaller than for τ_h candidates. Similar comparisons of other observables such as jet kinematics or $E_\mathrm{T}^\mathrm{miss}$ can either provide further confidence in the modelling or point towards differences to Monte Carlo simulation, but cannot be interpreted as embedding validation. Since they are not expected to be modelled better in simulated events than in embedded data, they can at most show the importance of the data-driven estimate and the level of improvement over simulated datasets.

The final and most important aspect of the validation is the comparison of all relevant kinematic distributions obtained from a τ-embedded dataset with the data in signal and control regions[12], which is always performed automatically for each analysis.

[12] Although a validation cannot be done in presence of signal, the final discriminant distribution usually contains bins with low or negligible signal expectation, which can build confidence in the modelling.

5.2.10 Corrections of Selection and Reconstruction Biases

Due to the local inefficiencies of the muon system, which are caused by the detector design and support structures, the embedded $Z \to \tau\tau$ events are subject to certain structures in the (η, ϕ) distributions of the τ decay products and in some cases can even have $p_T(\mu)$ dependent efficiencies from the muon triggers. The absolute normalisation of the τ-embedded sample, however, is usually obtained from the data in a control region or from a simulated sample and its corresponding theoretical cross section, so that only the $p_T(\mu)$-dependent effects can influence relevant event properties such as the mass. Although these are expected to be very small, the muon trigger, reconstruction and identification efficiencies are taken into account in order to correct for any effect on the embedded τ-leptons and their decays, which could bias certain analysis distributions (cf. [148]). Additional trigger-related corrections are necessary, since the technical implementation of the trigger response in the ATLAS software prevents a proper merging of the simulated trigger objects in the $Z \to \tau\tau$ event with the data. The trigger information of the decay products in the τ-embedded hybrid event can thus not be used in the analysis, but must be implemented via correction factors, which parametrise the trigger efficiencies in dependence of the object transverse momentum and (η, ϕ) position. In combination with the absolute normalisation of the τ-embedded data, effects from this missing trigger simulation are expected to be of negligible impact for all relevant analysis observables.

5.2.11 Applications and Outlook

Since the beginning of the data-taking period at the LHC, the embedding method has been developed and applied successfully for several analyses starting with a search for a neutral MSSM Higgs boson [149]. At the same time, modifications have been developed by the ATLAS collaboration e.g. for the τ-embedding of only one reconstructed muon in μ+jets events, which has been used for the background estimation in search for a charged (MSSM) Higgs boson [150]. The validation of the method is usually performed in the context of a specific analysis and can thus be found in this thesis for the presented analyses in Sections 6.1.2, 6.2.2, A.1 and B.1. While already the first development of the method was performed on simulated $Z \to \mu\mu$ samples and validated with simulated $Z \to \tau\tau$ events before the first data were recorded, these additional checks have been abandoned due to various technical reasons. For the future, a general and more detailed validation is foreseen [151], which will involve such simulation closures and allow a more precise evaluation of systematic effects to the method.

CHAPTER 6

Cut-Based Analyses

During the data taking periods between 2010 and 2012, the $H \to \tau\tau$ analysis has been developed from the $Z \to \tau\tau$ selection. Employing the developed and validated background estimation techniques, it has been optimised in several steps towards sensitivity to a signal from the SM Higgs boson. First exclusion limits for minimal supersymmetric models (MSSM) with enhanced $g_{H\tau\tau}$ couplings could already be established with the first 36 pb^{-1} as documented in [149]. A similar analysis strategy was then applied to a larger dataset of 1.06 fb^{-1} [132], which was also interpreted in the context of a SM Higgs boson. Although no direct sensitivity could be expected, this analysis excluded such a signal with a strength of $\mu \geq 12$ at 95% confidence level. While the selections of later analyses searching for MSSM Higgs bosons in $\tau^+\tau^-$ final states [152] focused on other features of the MSSM phenomenology, the sensitivity to a SM Higgs boson could be increased significantly by developing a dedicated search for SM production mechanisms, which especially exploits the VBF topology from the qqH process. The corresponding analysis is based on simple cuts on kinematic observables and marks an important milestones on the route towards an observation of $H \to \tau\tau$ decays. In the following, it is presented for the $\tau_{\text{lep}}\tau_{\text{had}}$ final states.

6.1 Search for SM $H \to \tau\tau$ Decays with 4.6 fb^{-1} of ATLAS Data

The first analysis of $\tau\tau$ final states dedicated to a search for the SM Higgs boson [153] was performed on the dataset collected by ATLAS in 2011, which corresponds to an integrated luminosity of 4.6 fb^{-1} at a centre-of-mass energy of $\sqrt{s} = 7$ TeV. While previous analyses such as the search for a neutral MSSM Higgs boson [149] or the $Z \to \tau\tau$ cross section measurement [140] defined only a single set of selection cuts for $\tau_\ell\tau_{\text{h}}$ final states, this analysis includes several categories, which are separately optimised for different event topologies and production mechanisms.

6.1.1 Datasets and Event Selection

Data events with $e\tau_{\text{h}}$ ($\mu\tau_{\text{h}}$) final states are selected with a single muon trigger requiring $p_T(\mu) > 18$ GeV or with a single electron trigger, for which the thresholds on the transverse momentum varied from $E_T(e) > 20$ GeV to $E_T(e) > 22$ GeV depending on the increasing instantaneous luminosity of the LHC over the 2011 data taking period.

In each event, electron candidates with $E_T(e) > 25$ GeV are selected with `tight` quality requirements on shower shapes and tracks and isolation criteria for track momenta $I_{PT}^{0.4}/E_T < 0.06$ and calorimeter

energy $I_{ET}^{0.2}/E_T < 0.08$.

Combined muon candidates with good quality ID tracks are selected with $p_T(\mu) > 20$ GeV if their isolation energy and momentum fulfil $I_{ET}^{0.2}/p_T < 0.04$ and $I_{PT}^{0.4}/p_T < 0.06$.

Hadronic τ decays are selected with a transverse momentum threshold of $p_T(\tau_\mathrm{h}) > 20$ GeV. They are identified with the Boosted Decision Tree method [154], for which an optimised working point `medium` with 50% efficiency yields a background mis-identification rate below 1%.

Jets with a transverse momentum of $p_T(j) > 25$ GeV are selected if their jet-vertex-fraction is sufficiently large (JVF > 0.75) or if they are located outside the coverage of the tracking system with $|\eta(j)| > 2.4$.

An event is selected in this analysis, when exactly one light lepton $\ell = e, \mu$ and one τ_h decay with opposite-sign charges are reconstructed. In order to suppress backgrounds with more than one light lepton ($Z \to \ell\ell$+jets, $t\bar{t}$) more efficiently, the transverse momentum thresholds for the second lepton candidate are lowered to $p_T(e) > 15$ GeV or $p_T(\mu) > 10$ GeV, and only a `medium` electron quality and no muon isolation are required. For a further suppression of the W+jets background, a cut on the transverse lepton mass is applied, so that only events with $m_\mathrm{T}(\ell) < 30$ GeV are considered in an inclusive preselection. This preselected phase space is then divided into seven analysis categories, which are summarised in the following.

$H+2$-jet (VBF) Category

The first and most important selection of this analysis defines a category, which exploits the VBF jet topology of the qqH production mechanism. Therefore, only events with at least two jets are selected, of which the two leading[1] ones must be in opposite detector hemispheres ($\eta(j_1) \times \eta(j_2) < 0$) , widely separated by $\Delta\eta(j_1, j_2) > 3.0$ and must have a large invariant mass $m(j_1, j_2) > 300$ GeV. In addition, both visible τ decay objects $\ell = e, \mu$ and τ_h are required to be central in η between the two leading jets. Further suppression of multijet and Z+jets backgrounds is achieved by imposing a cut of $E_\mathrm{T}^\mathrm{miss} > 20$ GeV on the events, which at the same time improves the resolution of the MMC mass by rejecting events with balanced neutrino momenta $p_\mathrm{T}^\mathrm{miss}(\nu_\tau) \lesseqgtr p_\mathrm{T}^\mathrm{miss}(\nu_{\ell+\tau})$.

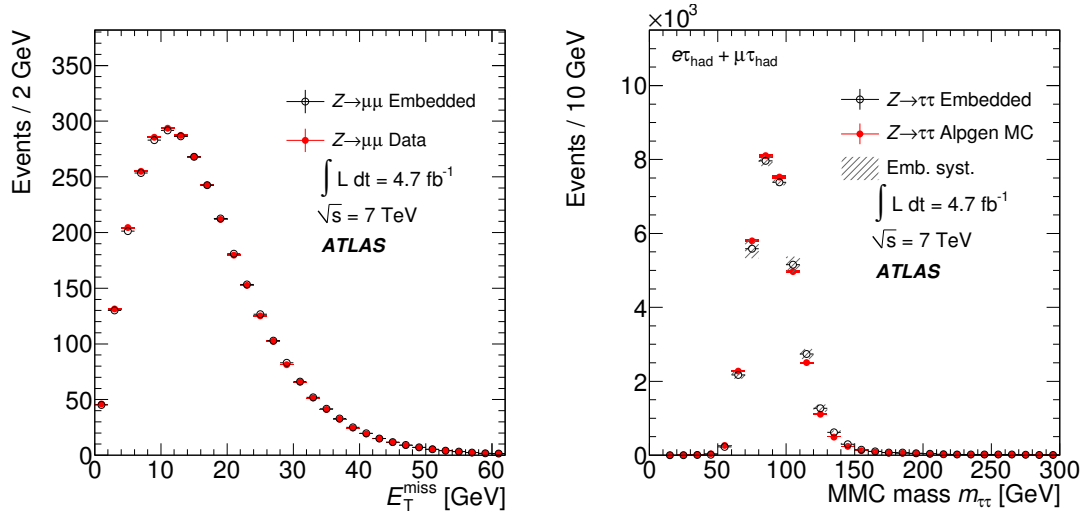
$H+1$-jet Categories

Events that fail the VBF selection of the $H+2$-jet category, but still contain at least one jet and fulfil $E_\mathrm{T}^\mathrm{miss} > 20$ GeV, are considered in the $H+1$-jet categories, which are dominated by signal from the ggH production. The additional jet ensures that the $Z \to \tau\tau$ decay products are not back-to-back, which improves the MMC mass reconstruction from $E_\mathrm{T}^\mathrm{miss}$ and thus the mass resolution. While the $H+2$-jet (VBF) category merges $e\tau_\mathrm{h}$ and $\mu\tau_\mathrm{h}$ due to the limited number of expected events, two $H+1$-jet categories separate these final states.

$H+0$-jet Categories

In absence of any jet with $p_T(j) > 25$ GeV, events enter the $H+0$-jet categories separated by their final state $e\tau_\mathrm{h}$ or $\mu\tau_\mathrm{h}$. A further separation of high $E_\mathrm{T}^\mathrm{miss} > 20$ GeV categories provides a selection of events with unbalanced neutrino momenta and thus a better mass resolution. At the same time, the $E_\mathrm{T}^\mathrm{miss} < 20$ GeV categories additionally recover all remaining events with balanced momenta $p_\mathrm{T}^\mathrm{miss}(\nu_\tau) \approx p_\mathrm{T}^\mathrm{miss}(\nu_{\ell+\tau})$, which represent the largest fraction of the expected signal from ggH production.

[1] The jets are considered to be ordered in $p_T(j)$, so that in this context "leading" means the two jets with the largest transverse momenta $p_T(j_1) > p_T(j_2) > p_T(j_3)$.



(a) $E_{\mathrm{T}}^{\mathrm{miss}}$ ($Z \to \mu\mu$) (b) MMC mass ($Z \to \tau\tau$)

Figure 6.1: (a) $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions for the $Z \to \mu\mu$ data before and after muon embedding and (b) MMC mass distributions for the τ-embedded $Z \to \mu\mu$ data and simulated $Z \to \tau\tau$ events [153]. For (a), only statistical uncertainties are shown; (b) also includes systematic uncertainties associated with the embedding procedure as discussed in Sec. 5.2.8.

6.1.2 Background Estimation

The background contribution from the dominant $Z \to \tau\tau$ process is determined from a sample of embedded data as described in Sec. 5.2. Other backgrounds are largely taken from simulated Monte Carlo samples, which are listed in Tab. 6.1. Further corrections and normalisations for these backgrounds as well as the estimation technique for contribution from the multijet background processes are described in the following.

Embedding Normalisation and Validation

As acceptances and efficiencies are not determined completely from the embedding method itself, the absolute normalisation of the embedded sample is performed with the help of a simulated Monte Carlo sample, which is used to extrapolate the theoretical $Z \to \tau\tau$ cross section into the inclusively selected phase space before categorisation. Therefore, the sample of embedded data events only provides the cut efficiencies and shapes of kinematic distributions after the inclusive selection. The efficiency of this selection $\varepsilon_{\mathrm{emb.}}$ is instead corrected to the value $\varepsilon_{\mathrm{MC}}$ obtained from the simulated sample, so that the embedded sample is effectively normalised to the theoretical cross section $\sigma_{Z\to\tau\tau}$ via correction factors:

$$\sigma_{\mathrm{emb.}}(\ell,\tau_h) = \sigma_{Z\to\tau\tau} \cdot C_f(\ell,\tau_h) \qquad C_f(\ell,\tau_h) = \frac{\varepsilon_{\mathrm{MC}}(\ell,\tau_h)}{\varepsilon_{\mathrm{emb.}}(\ell,\tau_h)} \,. \tag{6.1}$$

These factors are determined from the selection efficiencies ε in embedded data and in the Monte Carlo simulation to be $C_f(e,\tau_{\mathrm{h}}) = 0.305$ and $C_f(\mu,\tau_{\mathrm{h}}) = 0.176$. They differ for the two final states, since the detector acceptance is different for electrons and muons. The application of these factors ensures the normalisation to yield the same expected event number after the inclusive selection of one lepton and one τ_{h} candidate. The statistical error of the normalisation at this basic selection step is negligible

Process	Generator	PDF	$\sigma\times$ BR$\times\epsilon_f$/pb
$W \to \ell$+jets ($\ell = e, \mu, \tau$)	ALPGEN+HERWIG	CTEQ6L1	10.46×10^3
$Z/\gamma^* \to \ell\ell$+jets ($m(\ell,\ell) > 40$ GeV)	ALPGEN+HERWIG	CTEQ6L1	1.07×10^3
$Z/\gamma^* \to \ell\ell$+jets (10< $m_(\ell,\ell) < 40$ GeV)	ALPGEN+HERWIG	CTEQ6L1	3.89×10^3
$t\bar{t}$-production	MC@NLO	CT10	164.6
Single-top: t-, s-, Wt-channel	AcerMC+PYTHIA	CTEQ6L1	20.8, 1.5, 15.6
WW ($\geq 2\ell$), WZ($\geq 1\ell$), ZZ ($\geq 2\ell$)	MC@NLO	CT10	4.9, 4.4, 1.1
$ggH(\to \tau_{\text{lep}}\tau_{\text{had}})$	POWHEG+PYTHIA	MRST2007	0.45
$qqH(\to \tau_{\text{lep}}\tau_{\text{had}})$	POWHEG+PYTHIA	MRST2007	0.035
$WH((\to \tau_{\text{lep}}\tau_{\text{had}})$	PYTHIA	MRST2007	0.016
$ZH(\to \tau_{\text{lep}}\tau_{\text{had}})$	PYTHIA	MRST2007	9.2×10^{-3}

Table 6.1: Summary of Monte Carlo samples and cross sections multiplied by the relevant branching ratios (BR) and filtering efficiencies ϵ_f used in the $\sqrt{s}$ = 7 TeV analysis. In the case of the decays to leptons, the cross section is given per lepton channel. Numbers for $t\bar{t}$, single-top and di-boson processes do not include fully hadronic final states. Cross sections for the signal processes are listed for m_H=125 GeV, while samples are generated for different mass points in the range $m_H \in$ [100 GeV, 150 GeV] in steps of 5 GeV.

compared to the statistical error of the samples after the final selection is applied.

The results of this background estimation technique are scrutinised in a detailed validation, which starts with an embedding of simulated $Z \to \mu\mu$ decays into the selected data events. Figure 6.1a demonstrates that no significant bias is introduced in the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution by the cell subtraction and track replacement technique. In comparison to the Monte Carlo simulation in Fig. 6.1b, the reconstructed MMC mass distribution obtained from the embedded data is found to be in good agreement. This confirms that the basic kinematics of the $\tau^+\tau^-$ decays are described well by the embedded $Z \to \tau\tau$ decays, while especially the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution and jet kinematics are expected to be modelled better with the embedding technique than in the Monte Carlo simulation. Further comparisons of the $Z \to \tau\tau$ and $Z \to \mu\mu$ embedding are shown in Fig. A.1-A.4.

Normalisation Correction for the $W \to \ell\nu$ Background

Due to the known problems with mis-identification probabilities of τ_{h} candidates in simulated events, correction factors for the dominant W+jets process are derived from data as described in Sec. 5.1.2. This mis-modelling is expected to be different for quark- and gluon-induced jets, which appear with different probabilities depending on the requirement of additional jets in an event. The correction factors must therefore be derived separately for each category, which is achieved by defining an orthogonal control region with 70 GeV $< m_{\mathrm{T}} <$ 120 GeV for each signal region. In these control regions, the W+jets contribution is enriched, while especially the contamination from top production processes is suppressed by the upper threshold. An additional selection of $E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV is applied in order to suppress a residual multijet contamination. The W+jets component is then normalised to the observed data after the residual contamination of other backgrounds obtained from simulation or embedding is subtracted. The extracted correction factors are summarised in Tab. 6.2, which shows that the values are compatible between H +0-jet and H +1-jet selections, but show significantly higher values for the H +2-jet (VBF) category with two well-separated jets. The numbers for the H +0-jet categories are applied both for the low- and high-$E_{\mathrm{T}}^{\mathrm{miss}}$ categories, since the effect of multijet contamination in a low-$E_{\mathrm{T}}^{\mathrm{miss}}$ control region is expected to be larger than any correlation between the τ_{h} mis-identification probability and the $E_{\mathrm{T}}^{\mathrm{miss}}$ in W+jets events.

	H+0-jet		H+1-jet		H+2-jet VBF
	$e\tau_h$	$\mu\tau_h$	$e\tau_h$	$\mu\tau_h$	$e\tau_h$+$\mu\tau_h$
OS	0.574 ± 0.009	0.537 ± 0.008	0.585 ± 0.014	0.544 ± 0.013	0.967 ± 0.120
SS	0.814 ± 0.026	0.681 ± 0.024	0.770 ± 0.029	0.713 ± 0.026	1.179 ± 0.195

Table 6.2: Correction factors f_W for the W+jets normalisation and their statistical uncertainties for the different analysis categories.

Multijet Background from Sideband Extrapolation

Corresponding to the description in Sec. 5.1.2, the multijet background estimation is based on the extraction of normalisation and shape from the sideband region with same-sign charges (SS) of $(\ell^{\pm}, \tau_{\mathrm{h}}^{\pm})$ and their extrapolation into the signal region with opposite-sign charges (OS) $(\ell^{\pm}, \tau_{\mathrm{h}}^{\mp})$ via Eqn. (5.2). Two central assumptions are necessary in order to allow the application of this background estimation method: the kinematic distributions must be identical between opposite-sign and same-sign events, and

the ratio r_{QCD} does not depend on the lepton isolation. These assumptions are both validated in Sec. A.2.

The extrapolation factors r_{QCD} entering the estimate are expected to depend both on the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the jet selection and are thus obtained from dedicated control regions for the different signal categories, for which the isolation and charge requirements are inverted. The extracted factors are summarised in Tab. 6.3 and show significant differences between the categories as expected from the corresponding selection of multijet topologies.

	0jet		≥1jet		VBF
	$e\tau_h$	$\mu\tau_h$	$e\tau_h$	$\mu\tau_h$	$e\tau_h$+$\mu\tau_h$
r_{QCD} (high $E_{\mathrm{T}}^{\mathrm{miss}}$)	1.060 ± 0.044	1.138 ± 0.019	0.999 ± 0.034	1.139 ± 0.012	1.215 ± 0.061
r_{QCD} (low $E_{\mathrm{T}}^{\mathrm{miss}}$)	1.057 ± 0.024	1.145 ± 0.009	-	-	-

Table 6.3: Extrapolation factors r_{QCD} and their statistical uncertainties from the different anti-isolated control regions corresponding to the signal region definitions.

Summary of Signal and Background Expectations

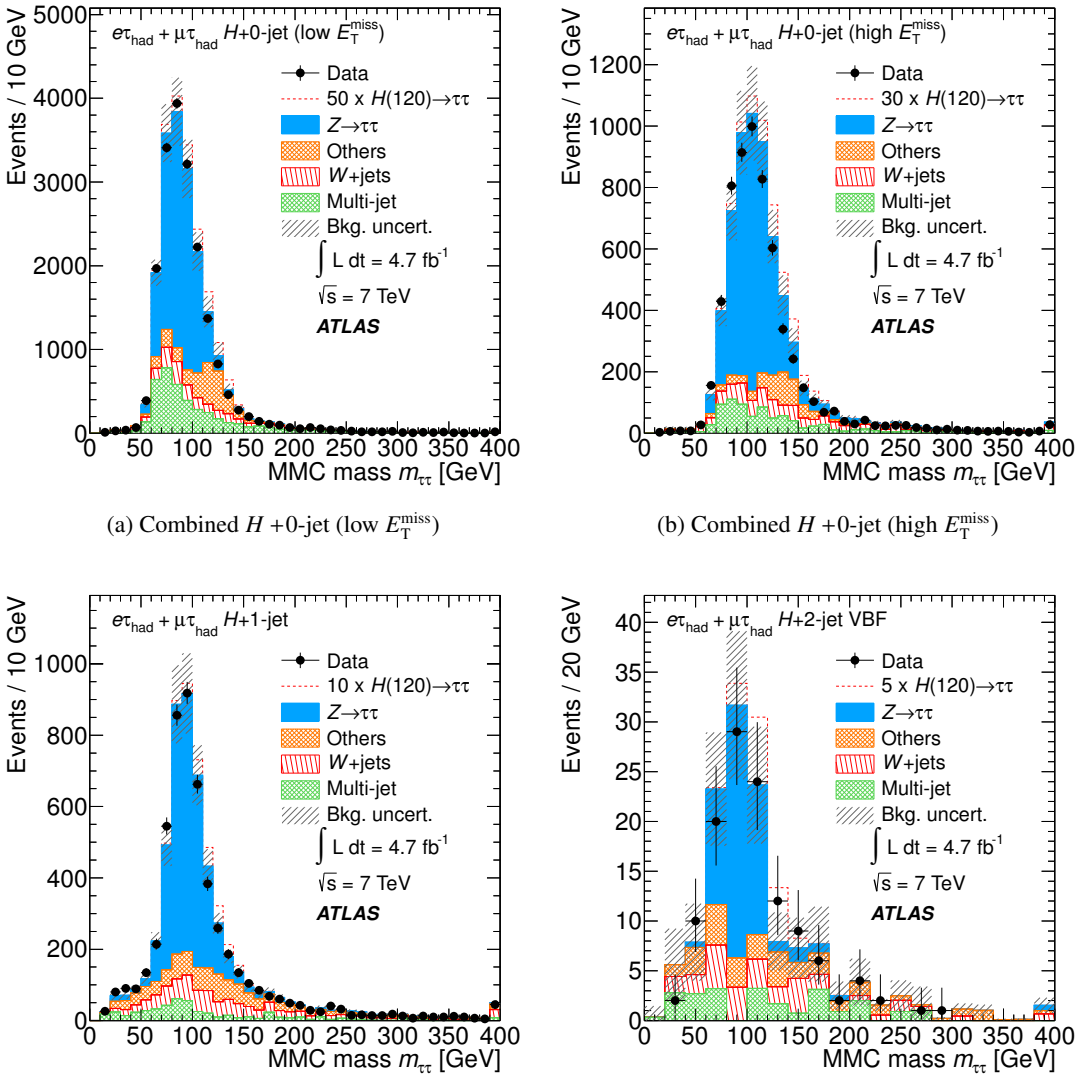
The signal and background expectations for the different analysis categories are largely obtained from simulated Monte Carlo samples, while the $Z \to \tau\tau$ background is estimated from τ-embedded $Z \to \mu\mu$ data events. The estimated contributions from simulated W+jets events are scaled with correction factors, which adjust their normalisation to the data observed in the dedicated control regions. The estimated numbers of signal and background events are summarised in Tab. 6.4-6.5 together with the multijet estimates obtained from the sideband subtraction method. The corresponding MMC mass distributions for all categories in Fig. 6.2 show no significant excess of the data over the estimated backgrounds, which is expected due to the small signal expectation from a SM Higgs boson. Figures A.8-A.11 show further kinematic distributions for the data and the estimated backgrounds at preselection level (cf. Sec. 6.1.1).

	low $E_{\mathrm{T}}^{\mathrm{miss}}$		high $E_{\mathrm{T}}^{\mathrm{miss}}$	
	$e\tau_h$	$\mu\tau_h$	$e\tau_h$	$\mu\tau_h$
ggF	$11.2 \pm 0.6 \pm^{1.3}_{1.5}$	$17.5 \pm 0.7 \pm^{2.0}_{2.4}$	$7.4 \pm 0.5 \pm^{0.7}_{0.7}$	$10.2 \pm 0.6 \pm^{1.0}_{0.7}$
VBF	$0.08 \pm 0.01 \pm^{0.02}_{0.03}$	$0.11 \pm 0.01 \pm^{0.02}_{0.03}$	$0.09 \pm 0.01 \pm^{0.02}_{0.01}$	$0.14 \pm 0.01 \pm^{0.02}_{0.02}$
WH	$0.05 \pm 0.01 \pm^{0.01}_{0.01}$	$0.04 \pm 0.01 \pm^{0.03}_{0.01}$	$0.04 \pm 0.01 \pm^{0.01}_{0.00}$	$0.05 \pm 0.01 \pm^{0.01}_{0.01}$
ZH	$0.02 \pm 0.01 \pm^{0.01}_{0.00}$	$0.06 \pm 0.01 \pm^{0.01}_{0.01}$	$0.04 \pm 0.01 \pm^{0.01}_{0.01}$	$0.03 \pm 0.01 \pm^{0.01}_{0.00}$
$Z \to \tau\tau$	$3718 \pm 58 \pm^{539}_{539}$	$7374 \pm 62 \pm^{1025}_{1023}$	$1511 \pm 37 \pm^{246}_{259}$	$2846 \pm 38 \pm^{462}_{456}$
$W \to \ell\nu$ + jets	$800 \pm 28 \pm^{28}_{104}$	$935 \pm 35 \pm^{66}_{26}$	$356 \pm 18 \pm^{29}_{83}$	$382 \pm 20 \pm^{31}_{83}$
$Z \to \ell\ell$ + jets	$1375 \pm 29 \pm^{594}_{596}$	$1033 \pm 27 \pm^{184}_{176}$	$337 \pm 13 \pm^{186}_{180}$	$217 \pm 11 \pm^{55}_{54}$
(Single) Top	$1.3 \pm 0.2 \pm^{0.4}_{0.2}$	$1.7 \pm 0.3 \pm^{0.3}_{0.5}$	$2.1 \pm 0.3 \pm^{0.5}_{0.7}$	$2.9 \pm 0.4 \pm^{0.7}_{0.4}$
Di-boson	$10.1 \pm 0.5 \pm^{0.7}_{0.7}$	$13.5 \pm 0.6 \pm^{0.8}_{5.9}$	$6.6 \pm 0.8 \pm^{0.8}_{0.8}$	$8.6 \pm 0.6 \pm^{0.6}_{0.7}$
QCD (estimate)	$2772 \pm 66 \pm^{94}_{92}$	$1372 \pm 64 \pm^{71}_{61}$	$543 \pm 32 \pm^{117}_{36}$	$289 \pm 30 \pm^{41}_{46}$
Total	$8677 \pm 97 \pm^{808}_{816}$	$10729 \pm 100 \pm^{1046}_{1040}$	$2756 \pm 53 \pm^{331}_{328}$	$3745 \pm 54 \pm^{468}_{469}$
Observed	8363	10911	2545	3570

Table 6.4: Observed data events and expected numbers of events obtained from Monte Carlo simulation in the H +0-jet categories. For the W+jets samples, the correction factors from Tab. 6.2 are applied to the Monte Carlo simulation. The multijet estimate is obtained from the sideband subtraction method with the extrapolation factors in Tab. 6.3. The quoted uncertainties are separated into their statistical component and the quadratic sum of all systematic effects described in Sec. 6.1.4.

	≥1jet		VBF
	$e\tau_h$	$\mu\tau_h$	$e\tau_h$+$\mu\tau_h$
ggF	$8.40 \pm 0.41 \pm^{1.06}_{0.84}$	$11.23 \pm 0.48 \pm^{1.19}_{1.44}$	$0.92 \pm 0.13 \pm^{0.26}_{0.19}$
VBF	$1.58 \pm 0.04 \pm^{0.10}_{0.11}$	$1.93 \pm 0.04 \pm^{0.10}_{0.12}$	$2.20 \pm 0.05 \pm^{0.24}_{0.23}$
WH	$0.71 \pm 0.05 \pm^{0.05}_{0.06}$	$0.87 \pm 0.06 \pm^{0.11}_{0.07}$	$0.02 \pm 0.01 \pm^{0.00}_{0.01}$
ZH	$0.42 \pm 0.03 \pm^{0.06}_{0.07}$	$0.46 \pm 0.03 \pm^{0.06}_{0.06}$	$0.00 \pm 0.00 \pm^{0.00}_{0.00}$
$Z \to \tau\tau$	$1289 \pm 34 \pm^{194}_{189}$	$1825 \pm 31 \pm^{268}_{268}$	$57 \pm 6 \pm^{8}_{8}$
$W \to \ell\nu$ + jets	$398 \pm 11 \pm^{19}_{30}$	$415 \pm 14 \pm^{17}_{24}$	$24.6 \pm 3.2 \pm^{7.0}_{10.3}$
$Z \to \ell\ell$ + jets	$360 \pm 12 \pm^{111}_{101}$	$95 \pm 7 \pm^{30}_{24}$	$11.2 \pm 1.8 \pm^{5.6}_{1.8}$
(Single) Top	$211 \pm 3 \pm^{17}_{17}$	$211 \pm 3 \pm^{15}_{14}$	$14.3 \pm 0.9 \pm^{0.9}_{1.1}$
Di-boson	$18.8 \pm 1.2 \pm^{2.6}_{1.5}$	$20.3 \pm 1.1 \pm^{2.6}_{1.5}$	$1.7 \pm 0.5 \pm^{0.3}_{0.3}$
QCD (estimate)	$353 \pm 33 \pm^{100}_{31}$	$157 \pm 26 \pm^{60}_{12}$	$22.7 \pm 8.2 \pm^{19.9}_{5.5}$
Total	$2630 \pm 50 \pm^{246}_{219}$	$2724 \pm 43 \pm^{277}_{271}$	$131 \pm 11 \pm^{23}_{14}$
Observed	2610	2711	122

Table 6.5: Observed data events and expected numbers of events obtained from Monte Carlo simulation in the H +1-jet and H +2-jet (VBF) categories. For the W+jets samples, the correction factors from Tab. 6.2 are applied to the Monte Carlo simulation. The multijet estimate is obtained from the sideband subtraction method with the extrapolation factors in Tab. 6.3. The quoted uncertainties are separated into their statistical component and the quadratic sum of all systematic effects described in Sec. 6.1.4.



(a) Combined H +0-jet (low $E_{\mathrm{T}}^{\mathrm{miss}}$)

(b) Combined H +0-jet (high $E_{\mathrm{T}}^{\mathrm{miss}}$)

(c) Combined H +1-jet

(d) H +2-jet VBF

Figure 6.2: MMC mass distributions of the selected events in the analysis categories [153]. The corresponding $e\tau_{\mathrm{h}}$ and $\mu\tau_{\mathrm{h}}$ final states are shown combined for the H +0-jet and H +1-jet categories, while they are considered separately in the background estimation and data analysis. The selected events in data are shown together with the prediction from a simulated Higgs boson signal ($m_H = 120$ GeV) stacked above the estimated background contributions, for which an uncertainty band represents the quadratic sum of all statistical and systematic uncertainties. For illustration, the signal contributions have been scaled by factors given in the legends.

6.1.3 Alternative OS-SS Estimation Method

In this specific analysis, a modification of the sideband extrapolation method is applied in parallel, which reduces the contributions from simulation to the final background composition. Instead of subtracting residual electroweak contamination from the data in the same-sign sideband, in this method the full same-sign data N^{SS}_{Data} are considered as background component in the corresponding opposite-sign region, while for all other processes $p \in \{Z \to \tau\tau, Z \to \ell\ell, \text{di-boson}, \text{top}\}$, only the differences

$$N_p^{\text{OS-SS}} = N_p^{\text{OS}} - N_p^{\text{SS}}$$

are included as contributions to the expected event yield in the signal region. Under the assumption of a perfect charge symmetry of $r_{\text{QCD}} = 1$, this method is equivalent to the sideband subtraction described in Sec. 5.1.2, since the total expected background

$$N^{\text{OS}}_{\text{Background}} = \sum_p N_p^{\text{OS}} + N^{\text{OS}}_{\text{Multijet}} = \sum_p N_p^{\text{OS}} + \left(N^{\text{SS}}_{\text{Data}} - \sum_p N_p^{SS} \right)$$

is only regrouped by combining the two sums:

$$N^{\text{OS}}_{\text{Background}} = \sum_p \left(N_p^{\text{OS}} - N_p^{\text{SS}} \right) + N^{\text{SS}}_{\text{Data}}$$

In terms of statistical analysis and presentation, the advantage of this method and its differently grouped background components is that smaller contributions $N_p^{\text{OS-SS}} \geq 0$ enter the final estimate, while a larger fraction of the backgrounds is obtained directly from the data. Furthermore, downward fluctuations of $N^{\text{SS}}_{\text{Data}}$ can lead to negative estimates of the multijet component $N^{\text{OS}}_{\text{Multijet}}$ from the sideband subtraction method, whereas the event numbers N_p^{OS} and N_p^{SS} are based on large simulated samples and their difference $N_p^{\text{OS-SS}} \geq 0$ is usually positive.

Instead of measuring the extrapolation factor r_{QCD} from the data, however, this method has to rely on the central assumption of $r_{\text{QCD}} = 1$. The error of this procedure must thus be quantified in a multijet enriched control region with no isolation criteria, a `loose` τ_{h} identification working point and $E_{\text{T}}^{\text{miss}} < 15$ GeV. A value of $r_{\text{QCD}} = 1.10 \pm 0.01$ is extracted, which is consistent with the values obtained from the dedicated control regions in the sideband extrapolation method. The difference to unity can still not be included directly in this OS-SS subtraction method, but must instead be included as a systematic uncertainty of 13% on the full contribution N^{SS}_{Data}. The correction for the W+jets simulation is applied as a scaling of the expected same-sign charge contribution:

$$N^{\text{OS-SS}}_{W \to \ell\nu} = k_W \cdot N^{\text{SS}}_{W \to \ell\nu}$$

According to Eqn. (5.1), this factor is equivalent to the following combination of the two separate correction factors f_W^{OS} and f_W^{SS}:

$$k_W = \frac{N^{\text{OS}}_{W \to \ell\nu} - N^{\text{SS}}_{W \to \ell\nu}}{N^{\text{SS}}_{W \to \ell\nu}} = \frac{f_W^{\text{OS}} N^{\text{WCR,OS}}_{W \to \ell\nu}}{f_W^{\text{SS}} N^{\text{WCR,SS}}} - 1$$

Instead of deriving both values separately, only the combined correction k_W is extracted from W+jets enriched control regions (WCR) with opposite-sign (OS) and same-sign (SS) charges, which correspond

to the signal regions with a selection of $m_{\mathrm{T}}(\ell) > 50$ GeV:

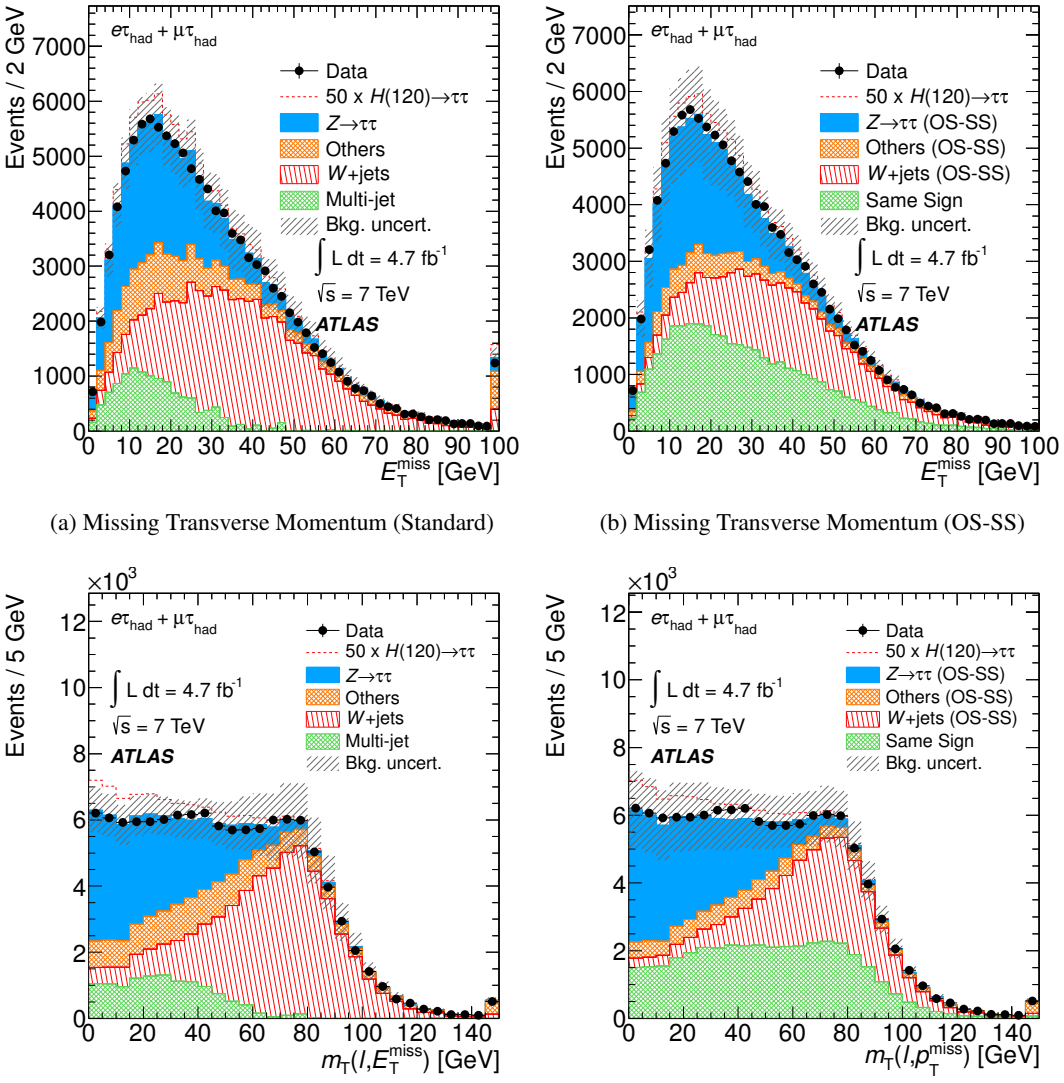
$$k_W = \frac{N_{\mathrm{Data}}^{\mathrm{WCR,OS}} - \sum_{p\in\mathcal{P}} N_p^{\mathrm{WCR,OS}}}{N_{\mathrm{Data}}^{\mathrm{WCR,SS}} - \sum_{p\in\mathcal{P}} N_p^{\mathrm{WCR,SS}}} - 1 \quad \text{with} \quad \mathcal{P} = (Z \to \tau\tau, Z \to \ell\ell, \text{di-boson}, \text{top})$$

As opposed to the corrections f_W, the k_W factors for each signal region are obtained primarily from data and do not have to rely on the Monte Carlo predictions for W+jets. The ratio $r_W = k_W + 1$ thus represents the opposite-sign to same-sign ratio as found in W+jets data, while the f_W depend on the simulation. For the final estimation of the background in the signal regions with opposite-sign charge requirement, however, their application has to rely on the Monte Carlo prediction of $N_{W\to\ell\nu}^{\mathrm{SS}}$ for the same-sign regions. Since the τ_{h} mis-identification probability is again expected to be mis-modelled in these predictions, a further correction with the corresponding f_W^{SS} is applied, so that the final W+jets estimate is derived as

$$N_{W\to\ell\nu}^{\text{OS-SS}} = k_W \cdot f_W^{\mathrm{SS}} \cdot N_{W\to\ell\nu}^{\mathrm{SS}} \,.$$

In contrast to the standard sideband extrapolation, which extracts the two different shapes $N_{W\to\ell\nu}^{\mathrm{OS}}$ and $N_{W\to\ell\nu}^{\mathrm{SS}}$ separately from the Monte Carlo simulation, this OS-SS modification only extracts the same-sign shape and extrapolates it to the opposite-sign via the factor r_W. Although both methods would provide identical yields if the definition of control regions were identical and the multijet background were complete charge symmetric ($r_{\mathrm{QCD}} = 1$), the OS-SS modification has to rely on the additional assumption that the shape of the kinematic distribution is identical for opposite-sign and same-sign W+jets events. This is found to be true in Monte Carlo simulations, where shape validations in the signal regions agree reasonably well.

As demonstrated in Fig. 6.3, the total background estimate from this OS-SS method compares well with the corresponding result of the standard sideband subtraction, while only the simulated OS-SS contributions from the different processes are smaller in comparison to the contribution from same-sign data. Similarly, the background estimates for the MMC distributions in the final analysis categories shown in Fig. 6.4 are compatible with the estimations of the standard method in Fig. 6.2. The predicted event yields obtained with this method are summarised in Tab. 6.6.



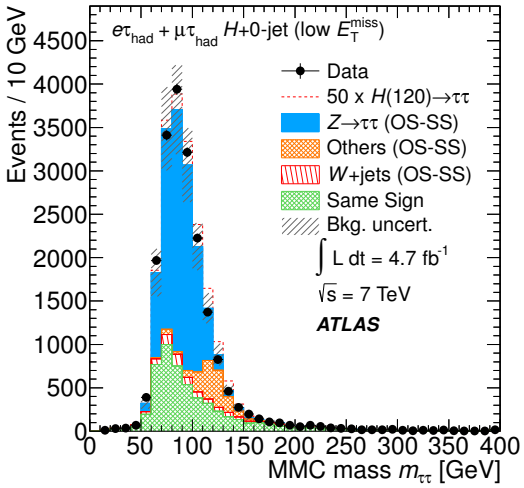
(a) Missing Transverse Momentum (Standard)

(b) Missing Transverse Momentum (OS-SS)

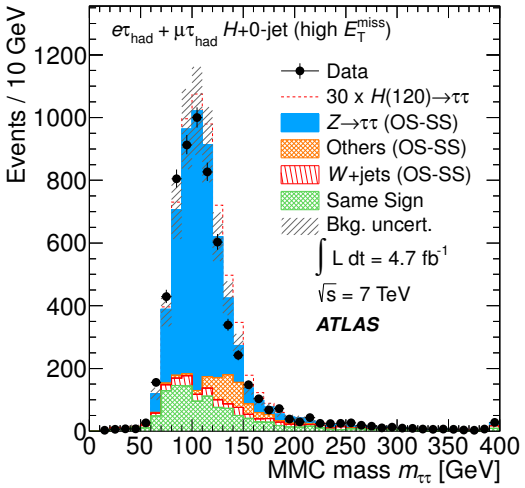
(c) Transverse Mass (Standard)

(d) Transverse Mass (OS-SS)

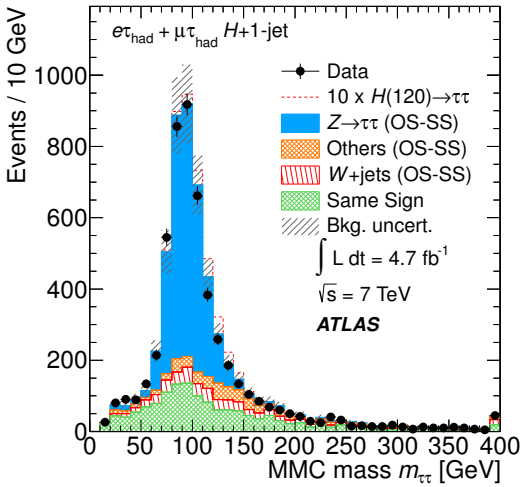
Figure 6.3: Missing transverse momentum and transverse mass distribution shown at preselection both for the standard sideband subtraction and the OS-SS background estimation method [153].



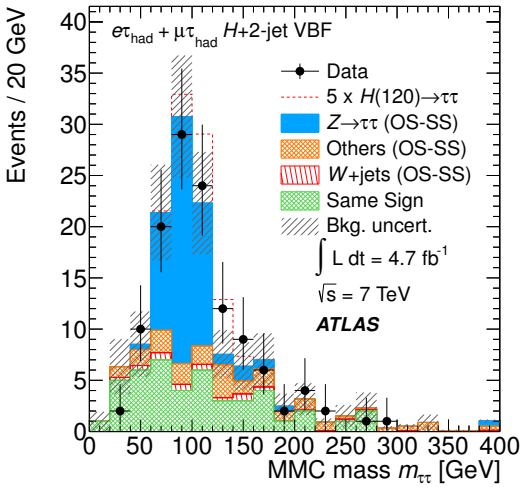
(a) Combined H +0-jet (low E_T^{miss})



(b) Combined H +0-jet (high E_T^{miss})



(c) Combined H +1-jet



(d) H +2-jet VBF

Figure 6.4: MMC mass distributions of the selected events in the analysis categories [153]. The corresponding $e\tau_h$ and $\mu\tau_h$ categories for the H +0-jet and H +1-jet categories are shown combined here, while in the data analysis they are considered separately. The selected events in data are shown together with the predicted Higgs boson signal (m_H = 120 GeV) stacked above the background contributions, which are estimated with the OS-SS method. For illustration, the signal contributions have been scaled by factors given in the legends.

	H +0-jet (low $E_{\mathrm{T}}^{\mathrm{miss}}$)		H +0-jet (high $E_{\mathrm{T}}^{\mathrm{miss}}$)	
	Electron	Muon	Electron	Muon
ggH signal	11 ± 1 ± 2	17 ± 1 ± 4	7.1 ± 0.8 ± 1.5	9.8 ± 0.9 ± 2.1
VBF H signal	0.08 ± 0.02 ± 0.12	0.11 ± 0.03 ± 0.03	0.09 ± 0.02 ± 0.02	0.14 ± 0.03 ± 0.03
VH signal	0.07 ± 0.02 ± 0.05	0.10 ± 0.03 ± 0.01	0.08 ± 0.0 2± 0.01	0.08 ± 0.02 ± 0.01
$N_{\mathrm{Data}}^{\mathrm{SS}}$	$(3.3 \pm 0.2 \pm 0.7) \cdot 10^3$	$(2.0 \pm 0.1 \pm 0.4) \cdot 10^3$	$(0.69 \pm 0.06 \pm 0.14) \cdot 10^3$	$(0.47 \pm 0.04 \pm 0.09) \cdot 10^3$
$N_{W\to\ell\nu}^{\mathrm{OS\text{-}SS}}$	$(0.33 \pm 0.02 \pm 0.04) \cdot 10^3$	$(0.50 \pm 0.02 \pm 0.07) \cdot 10^3$	$(0.15 \pm 0.01 \pm 0.02) \cdot 10^3$	$(0.18 \pm 0.01 \pm 0.03) \cdot 10^3$
$N_{Z\to\tau\tau}^{\mathrm{OS\text{-}SS}}$	$(3.70 \pm 0.06 \pm 0.61) \cdot 10^3$	$(7.29 \pm 0.06 \pm 1.21) \cdot 10^3$	$(1.49 \pm 0.04 \pm 0.23) \cdot 10^3$	$(2.80 \pm 0.04 \pm 0.40) \cdot 10^3$
$N_{\mathrm{other}}^{\mathrm{OS\text{-}SS}}$	$(0.97 \pm 0.04 \pm 0.22) \cdot 10^3$	$(0.59 \pm 0.04 \pm 0.14) \cdot 10^3$	$(0.27 \pm 0.02 \pm 0.08) \cdot 10^3$	$(0.14 \pm 0.02 \pm 0.04) \cdot 10^3$
Total background	$(8.2 \pm 0.2 \pm 0.8) \cdot 10^3$	$(10.4 \pm 0.2 \pm 1.2) \cdot 10^3$	$(2.59 \pm 0.07 \pm 0.26) \cdot 10^3$	$(3.59 \pm 0.06 \pm 0.43) \cdot 10^3$
Observed data	8363	10911	2545	3570
Default estimate	$(8.7 \pm 0.1 \pm 0.8) \cdot 10^3$	$(10.7 \pm 0.1 \pm 1.0) \cdot 10^3$	$(2.76 \pm 0.05 \pm 0.33) \cdot 10^3$	$(3.75 \pm 0.05 \pm 0.47) \cdot 10^3$

	H +1-jet		H +2-jet VBF
	Electron	Muon	Electron + Muon
ggH signal	8.1 ± 0.7 ± 1.6	10.8 ± 0.8 ± 2.2	0.9 ± 0.2 ± 0.3
VBFH signal	1.6 ± 0.1 ± 0.1	1.9 ± 0.1 ± 0.1	2.2 ± 0.1 ± 0.2
VH signal	1.1 ± 0.1 ± 0.1	1.4 ± 0.1 ± 0.1	0.02 ± 0.01 ± 0.01
$N_{\mathrm{Data}}^{\mathrm{SS}}$	$(0.93 \pm 0.07 \pm 0.19) \cdot 10^3$	$(0.49 \pm 0.04 \pm 0.10) \cdot 10^3$	45 ± 7 ± 9
$N_{W\to\ell\nu}^{\mathrm{OS\text{-}SS}}$	$(0.25 \pm 0.01 \pm 0.03) \cdot 10^3$	$(0.26 \pm 0.01 \pm 0.03) \cdot 10^3$	5 ± 1 ± 2
$N_{Z\to\tau\tau}^{\mathrm{OS\text{-}SS}}$	$(1.23 \pm 0.03 \pm 0.17) \cdot 10^3$	$(1.76 \pm 0.03 \pm 0.25) \cdot 10^3$	54 ± 6 ± 8
$N_{\mathrm{other}}^{\mathrm{OS\text{-}SS}}$	$(0.28 \pm 0.02 \pm 0.04) \cdot 10^3$	$(0.24 \pm 0.01 \pm 0.03) \cdot 10^3$	20 ± 3 ± 5
Total background	$(2.69 \pm 0.08 \pm 0.26) \cdot 10^3$	$(2.75 \pm 0.05 \pm 0.27) \cdot 10^3$	124 ± 10 ± 13
Observed data	2610	2711	122
Default estimate	$(2.63 \pm 0.05 \pm 0.25) \cdot 10^3$	$(2.72 \pm 0.04 \pm 0.28) \cdot 10^3$	131 ± 11 ± 23

Table 6.6: Predicted number of signal events (for $m_H = 120$ GeV) and predicted backgrounds obtained from the alternative OS-SS method as described in Sec. 6.1.3, together with the observed number of events in data for the analysis categories [153]. The total background yield predicted by the default estimation method is given as well for comparison. The listed uncertainties are statistical and systematic.

6.1.4 Systematic Uncertainties

The expectations of signal and background contributions in the analysis categories are estimated from simulated Monte Carlo samples or with data-driven extrapolation techniques, which are subject to systematic uncertainties. These must be estimated in order to interpret the observed result correctly in the context of a statistical analysis. Their estimation is based on the variation of the source of the systematic uncertainty, for which the effect is then evaluated per category and per signal and background component. If not mentioned otherwise, these effects are thus treated fully correlated across categories. While theoretical uncertainties mainly affect the event yields of the signal and background components, important detector related systematic uncertainties such as the energy scales for electrons, τ_{h} candidates and jets influence the shape of the MMC distribution and are thus treated as shape systematics in the final fit.

Theoretical Uncertainties

Systematic uncertainties arising from the calculation of production cross sections are related to the QCD scale and the parton distribution functions used to generate the Monte Carlo samples. Due to the various production mechanisms for the processes, these uncertainties are treated as uncorrelated between different processes. The Higgs production cross sections, branching ratios and their uncertainties are compiled in [57–59]. For the signal cross section, the QCD scale uncertainty depends on m_H and is obtained to be as large as 8–25% for the ggH production process depending on the number of jets [155, 156], while it is much smaller at the order of 1% for the VBF and VH production.

A corresponding uncertainty of 5% is assigned for the di-boson production mechanism, whereas 1% is assigned for the inclusive cross section of the single vector boson [157]. For the latter process, however, further variations of the parton shower algorithm and corresponding choices of scales and matching parameters are found to result in an additional uncertainty of 12.5% for the $Z \to \tau\tau$ process. This uncertainty must be applied to the estimate obtained with the embedding technique, since it is normalised according to the theoretical cross section. For the $t\bar{t}$ and single top-quark production, QCD scale uncertainties are obtained in a range of 3–6% [121, 158, 159].

Uncertainties related to the choice of the parton distribution function are assumed to be correlated for all predominantly gluon-initiated processes such as ggH and $t\bar{t}$, and a value of 8% is extracted from corresponding variations. For the predominantly quark-initiated processes as VBF, VH, single vector boson and di-boson production, a value of 4% is assigned [36, 126, 160, 161]. Further uncertainties related to hadronisation effects are estimated by varying ISR and UE effects within the default PYTHIA parton shower implementation according to [162] and by replacing this default parton shower implementation with the corresponding HERWIG algorithm, which yields a total uncertainty of approximately 10% for the expected number of $H \to \tau\tau$ events from the qqH production mechanism selected in the H +2-jet VBF category.

Detector-Related Uncertainties

The simulated efficiencies of the different triggers for electrons and muons in the Monte Carlo samples are corrected to the values obtained from dedicated performance measurements and are varied within the measured uncertainties. Similarly, reconstruction and identification efficiencies of electrons, muons and hadronic τ decays are corrected in all simulations, and corresponding variations within uncertainties are performed in order to evaluate the systematic effects on the analysis, which are typically found to be of the order of 1–2%. For the hadronic τ decay identification efficiency, however, which is estimated from $W \to \tau\nu$ and $Z \to \tau\tau$ enriched data samples, an uncertainty of 4% is obtained [154]. In addition to the

efficiencies, which usually affect the overall event yields of a background component, also momentum scales and resolutions of electrons and muons are corrected and varied according to performance measurements. Besides these leptons, especially the energy scale of the τ_{h} candidates and jets can affect the MMC mass distribution. The corresponding systematic uncertainties are evaluated based on the single hadron response in the calorimeters [83] and are treated as fully correlated between τ_{h} candidates and jets. Therefore, both the τ_{h} and jet energy scale are varied up and down in parallel within their measured uncertainties, which are in the range of 2.5–12% for the jet energy and of 2–5% for the τ_{h} energy depending on the object p_T and η. The corresponding variations of all transverse momenta in an event are taken into account as contributions to the momentum balance and are thus directly propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. Their effect on the MMC distribution for the final fit is extracted bin-by-bin as shape uncertainty. Additional $E_{\mathrm{T}}^{\mathrm{miss}}$ contributions from unmatched cells are considered as independent, so that uncertainties related to pile-up noise and low energy clusters in the calorimeters are treated as uncorrelated. Further pile-up effects especially on the tracking performance are estimated by re-weighting the Monte Carlo events according to the simultaneous interactions added during their simulation, so that the average number of simultaneous interactions can effectively be varied within the measured uncertainty of the instantaneous luminosity during the data taking periods. The integrated luminosity is measured over all periods, and its uncertainty of 3.9% [163] is included for the normalisation of all simulated backgrounds, which are normalised according to the calculated cross section and are not corrected to observed event yields in the data.

Background Modelling Uncertainties

The modelling of the $Z \to \tau\tau$ background is performed with τ-embedded $Z \to \mu\mu$ data as described in Sec. 5.2. Corresponding uncertainties are obtained by propagating variations of the $Z \to \mu\mu$ event selection and the muon energy subtraction procedure through the τ-embedding procedure (cf. Sec. 5.2.8). The uncertainty on the QCD estimate is based on the statistical uncertainty of r_{QCD} given by the numbers of events in the anti-isolated control regions. In addition, the W+jets correction factors f_W are varied within their statistical uncertainty for each category, and the effects are propagated to the multijet estimate as an additional systematic uncertainty. Since the variation of f_W^{OS} and f_W^{SS} has opposite effects on the background estimate in the signal region, their uncertainties are anti-correlated, so that the maximal effect of the W+jets contribution is constructed as a conservative estimate. The resulting method-specific uncertainties are summarised in Tab. 6.7.

	0jet				≥1jet		VBF
	low $E_{\mathrm{T}}^{\mathrm{miss}}$		high $E_{\mathrm{T}}^{\mathrm{miss}}$				
	$e\tau_h$	$\mu\tau_h$	$e\tau_h$	$\mu\tau_h$	$e\tau_h$	$\mu\tau_h$	$e\tau_h$+$\mu\tau_h$
W+jets stat. f_W	1.6%	1.5%	1.6%	1.5%	2.5%	2.43%	12.4%
QCD stat. r_{QCD}	2.3%	0.8%	4.2%	1.6%	3.4%	1.0%	5.0%
QCD syst. f_W	0.3%	0.8%	0.6%	1.6%	2.4%	5.7%	12.6%

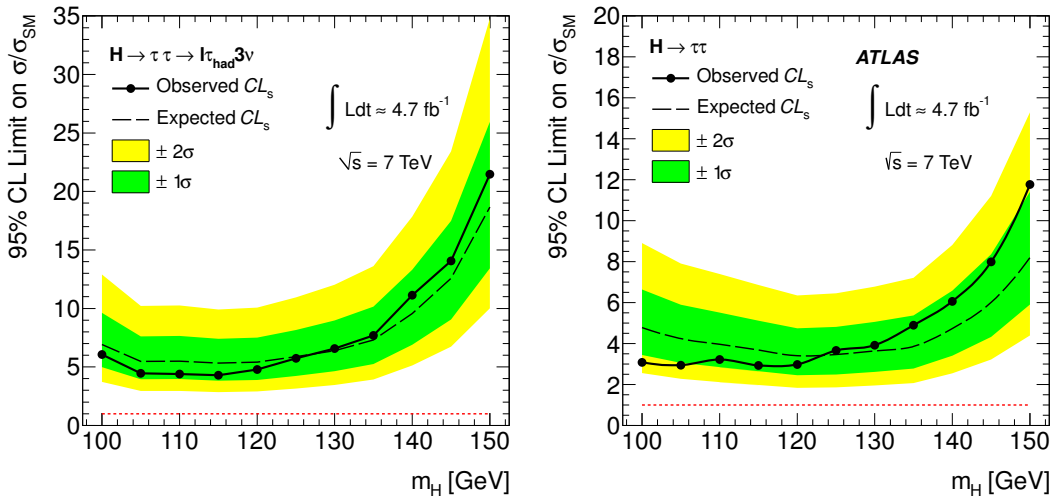
Table 6.7: Systematic uncertainties on the data-driven background estimates obtained from the statistical uncertainties in the control regions and the propagation of the varied W+jets contributions to the sideband subtraction method for the multijet estimate.

6.1.5 Results

With the procedure described in Sec. 4.3.3, exclusion limits can be set on the signal strength parameter μ with a confidence level of 95%. These are shown in Fig. 6.5 with their dependence of the Higgs boson mass m_H and are normalised to the expected cross section times branching ratio $\sigma_{\mathrm{SM}}(H \to \tau_{\mathrm{lep}}\tau_{\mathrm{had}})$ ($\sigma_{\mathrm{SM}}(H \to \tau\tau)$) in the Standard Model. The expected limits represent the exclusion on the signal obtained from the Asimov data, while their bands correspond to the expected statistical fluctuations in real data. The observed limits are in statistical agreement with this expectation, since no significant excess over the SM background is present in the data.

For the $H \to \tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channel, the observed (expected) limit of 5.8 (5.9) times the predicted Standard Model cross section for $m_H = 125$ GeV is consistently provided with both background estimation methods. In addition, a combination of the OS-SS method with other channels is performed, resulting in an exclusion limit of 3.7 (3.5) on$\mu = \sigma/\sigma_{\mathrm{SM}}$ for the combined $H \to \tau\tau$ signal.

This result of the first dedicated search for a SM Higgs boson in the $\tau^+\tau^-$ decay channel is a major milestone towards the confirmation of its fermion couplings and significantly improves the previous result obtained with a dataset corresponding to 1.06 fb^{-1} [132]. This previous analysis was optimised for potential signals of MSSM Higgs bosons and considered only a single inclusive category for the $H \to \tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channel. It was thus limited by systematic uncertainties rather than the statistical uncertainties from the available dataset. The categorisation of events now yields an improvement of the expected exclusion limit for a SM-like signal in the $H \to \tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channel from $\mu < 12$ to an exclusion limit of $\mu < 6$. This is achieved with the separate optimisation of the categories for different event topologies, of which the VBF category dominates the final sensitivity of the analysis.



(a) $H \to \tau_{\text{lep}}\tau_{\text{had}}$ (b) Combined

Figure 6.5: Observed (solid) and expected (dashed) 95% confidence level upper limits on the signal strength $\mu = \sigma/\sigma_{\text{SM}}$ as a function of the Higgs boson mass. Expected limits are obtained from the Asimov data created for the scenario with no signal ($\mu = 0$). The bands around the dashed line indicate the expected statistical fluctuations of the observed limit. Results are given for the estimate with the sideband subtraction method (a) and for the combination of the $H \to \tau_{\text{lep}}\tau_{\text{had}}$ OS-SS estimate with the results from the other channels $H \to \tau_{\text{lep}}\tau_{\text{lep}}$ and $H \to \tau_{\text{lep}}\tau_{\text{had}}$ (b) [149].

6.2 Optimised Analysis with 4.6+13.0 fb^{-1} of ATLAS Data

During the LHC run period in 2012, an intermediate dataset corresponding to an integrated luminosity of 13.0 fb^{-1} at a centre-of-mass energy of $\sqrt{s} = 8$ TeV was recorded with ATLAS. For the analysis of these data, an optimised event selection is applied, which directly exploits the transverse momenta of the Higgs bosons and adds further background suppression by introducing cuts on more variables. This improved selection is also adapted for the previous dataset (4.6 fb^{-1} at $\sqrt{s} = 7$ TeV), which in combination with the new data results in a large increase of sensitivity due to enhanced background suppression and reduced statistical uncertainty [164].

6.2.1 Datasets and Event Selection

Channel	Trigger	Trigger p_{T} Threshold (GeV)	Offline p_{T} Threshold (GeV)
		7 TeV	
$H \to \tau_{\mathrm{lep}}\tau_{\mathrm{had}}$	single electron	$p_{\mathrm{T}}^{e} > 20 - 22$	$p_{\mathrm{T}}^{e} > 25$
		–	$p_{\mathrm{T}}^{\tau_{\mathrm{h}}} > 20$
	single muon	$p_{\mathrm{T}}^{\mu} > 18$	$p_{\mathrm{T}}^{\mu} > 25$
		–	$p_{\mathrm{T}}^{\tau_{\mathrm{h}}} > 20$
	combined $e + \tau_{\mathrm{h}}$	$p_{\mathrm{T}}^{e} > 15$	$17 < p_{\mathrm{T}}^{e} < 25$
		$p_{\mathrm{T}}^{\tau_{\mathrm{h}}} > 16 - 20$	$p_{\mathrm{T}}^{\tau_{\mathrm{h}}} > 25$
$H \to \tau_{\mathrm{lep}}\tau_{\mathrm{had}}$	single electron	$p_{\mathrm{T}}^{e} > 24$	$p_{\mathrm{T}}^{e} > 26$
		–	$p_{\mathrm{T}}^{\tau_{\mathrm{h}}} > 20$
	single muon	$p_{\mathrm{T}}^{\mu} > 24$	$p_{\mathrm{T}}^{\mu} > 26$
		–	$p_{\mathrm{T}}^{\tau_{\mathrm{h}}} > 20$
	combined $e + \tau_{\mathrm{h}}$	$p_{\mathrm{T}}^{e} > 18$	$20 < p_{\mathrm{T}}^{e} < 26$
		$p_{\mathrm{T}}^{\tau_{\mathrm{h}}} > 20$	$p_{\mathrm{T}}^{\tau_{\mathrm{h}}} > 25$
	combined $\mu + \tau_{\mathrm{h}}$	$p_{\mathrm{T}}^{\mu} > 15$	$17 < p_{\mathrm{T}}^{\mu} < 26$
		$p_{\mathrm{T}}^{\tau_{\mathrm{h}}} > 20$	$p_{\mathrm{T}}^{\tau_{\mathrm{h}}} > 25$

Table 6.8: The triggers used in the various channels, along with the corresponding trigger and offline p_{T} thresholds on the reconstructed objects [164]. Ranges indicate that the threshold varied during the data-taking between the given values.

With a similar selection of basic objects as described in Sec. 6.1.1, this analysis includes combined $\ell\tau_{\mathrm{h}}$ triggers in addition to the previously used single-lepton triggers, which contribute events with lower transverse momenta $p_T(\ell)$ and $p_T(\tau_{\mathrm{h}})$ and thus increase the size of the available dataset. For the 7 TeV data, events selected by a combined $e\tau_{\mathrm{h}}$ trigger are added, while for the 8 TeV data events from an $e\tau_{\mathrm{h}}$ and a $\mu\tau_{\mathrm{h}}$ trigger are included in the analysis. In order to avoid an overlap between the different data streams, the corresponding trigger and analysis thresholds on the object transverse momenta are chosen exclusively as documented in Tab. 6.8.

For the suppression of backgrounds with more than one light lepton such as $Z \to \ell\ell$ or top-quark production, these thresholds are lowered to 10 GeV (15 GeV) for any additional muon (electron), and isolation and quality criteria are loosened. Events fulfilling all these preselection criteria are then grouped into six categories, which are developed from the categorisation of the previous analysis. The thresholds for the VBF selection are tightened, and the low-$E_{\mathrm{T}}^{\mathrm{miss}}$ H +0-jet categories are omitted, while an additional *Boosted* category is introduced in order to exploit the features of the specific $\tau^+\tau^-$ decay topology

in events with large $p_T(H)$. A detailed overview of the improved categorisation can be found in Tab. B.1, which is summarised in the following.

VBF Category

For the 8 TeV analysis, $e\tau_h$ and $\mu\tau_h$ events are selected with $p_T(\tau_h) > 30$ GeV, $p_T(\ell) > 26$ GeV, $E_T^{miss} > 20$ GeV and at least two jets with transverse momenta of $p_T(j) > 40$ (30) GeV. These two leading jets must be found in opposite detector hemispheres ($\eta(j_1) \times \eta(j_2) < 0$), with a large separation of $\Delta\eta(j_1, j_2) > 3.0$ in pseudorapidity and a large invariant mass of $m(j_1, j_2) > 500$ GeV. Both the lepton and τ_h candidate must be central in η between the two leading jets, and the transverse momentum of the qqH system is required to be small with $p_T^{tot} < 40$ GeV. For the 7 TeV analysis, the additional requirement on $p_T(\ell)$ is omitted, while the threshold of $p_T(j) > 40$ GeV is also applied for the sub-leading jet, and the qqH transverse momentum cut is tightened to $p_T^{tot} < 30$ GeV.

Boosted Category

Events with $e\tau_h$ or $\mu\tau_h$ final states failing the VBF selection are selected in the Boosted category, if the reconstructed momentum of the Higgs boson is large $p_T(H) > 100$ GeV. In order to select topologies with well reconstructable invariant masses, a sufficiently large $E_T^{miss} > 20$ GeV is required to result in visible momentum fractions in the ranges $0 < x_\ell < 1$ and $0.2 < x_{\tau_h} < 1.2$. For the 8 TeV analysis, an additional threshold of $p_T(\tau_h) > 30$ GeV is applied.

1-jet Categories

Failing both the VBF and Boosted selections, events with $E_T^{miss} > 20$ GeV and at least one jet with $p_T(j) > 30$ (25) GeV enter the 1-jet categories of the 8 (7) TeV analysis, which are separated into $e\tau_h$ and $\mu\tau_h$ final states.

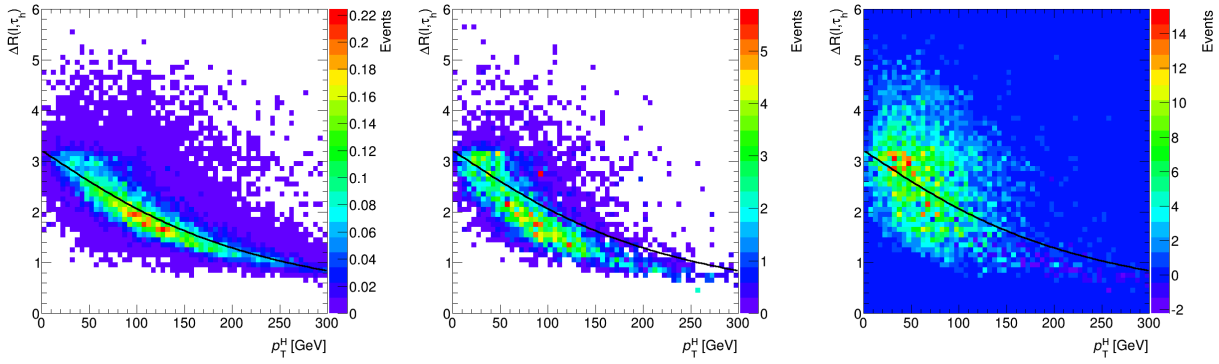
0-jet Categories

Without any jet with $p_T(j) > 30$ (25) GeV, events enter the $e\tau_h$ or the $\mu\tau_h$ 0-jet category of the 8 (7) TeV analysis if sufficient $E_T^{miss} > 20$ GeV is reconstructed in the event.

Category-Dependent Background Rejection

While this categorisation creates mutually exclusive selections containing all preselected events with $E_T^{miss} > 20$ GeV, additional requirements as summarised in Tab. 6.8 are applied for each category in order to provide further suppression of backgrounds:

- The dominant W+jets contribution is reduced by accepting only events with $m_T < 50$ GeV and $\sum \Delta\phi < 3.5$.
- In the Boosted category, the distribution of $\sum \Delta\phi$ changes significantly, since ℓ, τ_h and E_T^{miss} become more collinear with higher boosts. While the corresponding cut is thus tightened to $\sum \Delta\phi < 1.6$ for the 7 TeV analysis, it is completely omitted for the 8 TeV Boosted selection.
- In the 0-jet categories, the selection is even tightened by requiring $m_T < 30$ GeV and applying an additional cut on the difference of visible transverse momenta $p_T(\tau_h) - p_T(\ell) > 0$, which



(a) $H \to \tau\tau$ Signal (m_H = 125 GeV) (b) $Z \to \tau\tau$ Background (c) Other Backgrounds

Figure 6.6: Two-dimensional distribution of $\Delta R(\ell, \tau_h)$ and $p_T(H)$ for (a) the $H \to \tau\tau$ signal with m_H = 125 GeV, (b) the $Z \to \tau\tau$ background from τ-embedding and (c) all other background processes, shown for a loose VBF selection as described in Sec. 7.3. The solid black line indicates the fitted Landau function to the signal profile, which is used to obtain ΔR_{pred} as a function of $p_T(H)$.

is especially effective against W+jets events with large $p_T(\ell)$ and lower $p_T(j \to \tau_h)$ of a mis-identified jet.

- The correlation between the transverse momentum $p_T(H)$ of the Higgs boson and the angular separation $\Delta R(\ell, \tau_h)$ of the sub-sequent τ decays is exploited by comparing $\Delta R(\ell, \tau_h)$ to a predicted ΔR_{pred}, which is extracted from signal simulations as a function of $p_T(H)$. This is obtained by fitting a Landau function to the signal profile of the most probable $\Delta R(\ell, \tau_h)$ as a function of $p_T(H)$. The difference $\Delta(\Delta R)$=$|\Delta R(\ell, \tau_h) - \Delta R_{\text{pred}}(p_T(H))|$ is then used to further suppress non-$\tau\tau$ backgrounds.

- The background contribution from top quark production processes is significant in the VBF and Boosted categories and is therefore suppressed by a so-called *b-veto*, which rejects events with b-tagged jets. For their identification, a working point with 60%-70% efficiency [165, 166] for b-quark induced jets is chosen, for which a mis-identification probability of 0.1%-0.5% is measured [167].

6.2.2 Background Estimation

The estimation of backgrounds is largely based on simulated Monte Carlo samples, which are summarised in Tab. 6.9 with their corresponding cross sections for the 8 TeV analysis[2]. In order to reduce the statistical uncertainty of the Z+jets prediction, in the VBF category the default samples are replaced by dedicated samples produced with a filter applied to jets on particle level: an invariant mass $m(q_1, q_2) > 200$ GeV and a separation of $\Delta\eta(q_1, q_2) > 2.0$ of the two leading jets is required for these events before being passed to the detector simulation. This filter effectively rejects the large number of events with other topologies, which would otherwise be passed to the CPU-intensive detector simulation. The filtered samples therefore contain only Z+jets events with VBF-like topologies, which are mainly selected by the corresponding category.

[2] The Monte Carlo samples for the 7 TeV analysis remain unchanged.

Process	Generator	PDF	$\sigma\times$ BR$\times\epsilon_f$/pb
$W \to \ell$+jets ($\ell = e, \mu, \tau$)	ALPGEN+HERWIG	CTEQ6L1	12.22×10^3
$Z/\gamma^* \to \ell\ell$+jets (10< $m_(\ell, \ell)$ <60 GeV)	ALPGEN+HERWIG	CTEQ6L1	4.35×10^3
$Z/\gamma^* \to \ell\ell$+jets ($m(\ell, \ell)$ >60 GeV)	ALPGEN+HERWIG	CTEQ6L1	1.15×10^3
$Z/\gamma^* \to \ell\ell$+jets ($m(q_1, q_2) > 200$ GeV, $\Delta\eta(q_1, q_2) > 2.0$)	ALPGEN+HERWIG	CTEQ6L1	140.4
$t\bar{t}$-production	MC@NLO	CT10	238.1
Single-top: t-channel	AcerMC+PYTHIA	CTEQ6L1	28.4
Single-top: s-, Wt-channel	MC@NLO	CT10	1.8, 22.4
WW, WZ, ZZ ($\geq 1\ell$)	MC@NLO	CT10	20.8, 6.8, 1.6
$ggH(\to \tau_{\text{lep}}\tau_{\text{had}})$	POWHEG+PYTHIA	CT10	0.56
$qqH(\to \tau_{\text{lep}}\tau_{\text{had}})$	POWHEG+PYTHIA	CT10	0.045
$WH(\to \tau_{\text{lep}}\tau_{\text{had}})$	PYTHIA	CTEQ6L1	0.020
$ZH(\to \tau_{\text{lep}}\tau_{\text{had}})$	PYTHIA	CTEQ6L1	0.011

Table 6.9: Summary of Monte Carlo samples and cross sections multiplied by the relevant branching ratios (BR), k-factors and filtering efficiencies ϵ_f used in the $\sqrt{s}$ = 8 TeV analysis. In the case of decays to leptons, the cross section is given per lepton channel. Numbers for $t\bar{t}$, single-top and di-boson processes do not include fully hadronic final states. Cross sections for the signal processes are listed for m_H=125 GeV, while samples are generated for different mass points in the range $m_H \in [100\text{ GeV}, 150\text{ GeV}]$ in steps of 5 GeV.

In addition to the optimisation of the event selection with improved signal region definitions, also the data-driven background estimation methods described in Sec. 6.1.2 are consolidated, and a new so-called *fake-factor* method is established.

Improved Sideband Subtraction

After maintaining two different variations of the sideband subtraction in the previous analysis, an improved method is developed, which combines the benefits of both alternatives into a single combination of opposite-sign and same-sign contributions of background processes. In order to reduce the size of systematic effects, the full same-sign data $N^{\mathrm{SS}}_{\mathrm{Data}}$ from the control region are used in each analysis category, while for all other simulated or embedded backgrounds, only the difference between yields with opposite-sign and same-sign charges are included in the final estimate. Instead of applying the simple OS-SS method, the improved estimate is built from OS $- r_{\mathrm{QCD}} \cdot$ SS terms for each electroweak process. As a result, the assumption $r_{\mathrm{QCD}} = 1$ is no longer necessary and in contradiction with the values $r_{\mathrm{QCD}} > 1$, which are obtained from the multijet enriched control regions (cf. Tab. 6.3). The regrouping of the sideband extrapolation and the subtraction of simulated components is therefore starting from the composition

$$N^{\mathrm{OS}}_{\mathrm{Background}} = \sum_{p\in\mathcal{P}} N^{\mathrm{OS}}_{p} + N^{\mathrm{OS}}_{\mathrm{Multijet}} = \sum_{p\in\mathcal{P}} f^{\mathrm{OS}}_{p} \cdot N^{\mathrm{OS}}_{p} + r_{\mathrm{QCD}} \cdot \left(N^{\mathrm{SS}}_{\mathrm{Data}} - \sum_{p\in\mathcal{P}} f^{\mathrm{SS}}_{p} \cdot N^{SS}_{p} \right)$$

and results in the sum of the full same-sign data $N^{\mathrm{SS}}_{\mathrm{Data}}$ and the OS $- r_{\mathrm{QCD}} \cdot$ SS background contributions for all simulated or embedded backgrounds:

$$N^{\mathrm{OS}}_{\mathrm{Background}} = \sum_{p\in\mathcal{P}} \left(f^{\mathrm{OS}}_{p} \cdot N^{\mathrm{OS}}_{p} - r_{\mathrm{QCD}} \cdot f^{\mathrm{OS}}_{p} \cdot N^{\mathrm{SS}}_{p} \right) + r_{\mathrm{QCD}} \cdot N^{\mathrm{SS}}_{\mathrm{Data}}$$

$$\text{with} \quad \mathcal{P} = \{Z \to \tau\tau, Z \to \ell\ell_{(+j\to\tau_{\mathrm{h}})}, Z \to \ell\ell_{(\ell\to\tau_{\mathrm{h}})}, W \to \ell\nu, \text{di-boson}, \text{top}\}$$

The factors $f^{\mathrm{OS,SS}}_{p}$ are included as general corrections of event yields to data, which are potentially mis-modelled in simulations.

Correction Factors from Control Regions

For the W+jets correction, the factors f^{OS}_{W} and f^{SS}_{W} are obtained from control regions with large transverse mass, similar to the procedure described in Sec. 5.1.2. In this analysis, a control region is defined for each category by requiring $m_{\mathrm{T}}(\ell) > 70$ GeV, while no additional suppression cuts on x_{ℓ},$x_{\tau_{\mathrm{h}}}$, $\sum \Delta\Phi$ or $\Delta(\Delta R)$ are applied[3].

As for the W+jets events, correction factors for the jet $\to \tau_{\mathrm{h}}$ mis-identification probability are obtained for $Z \to \ell\ell_{(+j\to\tau_{\mathrm{h}}}$ events from dedicated control regions, which are defined by requiring a second isolated lepton with the same flavour as the first but with opposite charge ($e^{\pm}e^{\mp}$ or $\mu^{\pm}\mu^{\mp}$). Since the default charge correlation between the jet$\to\tau_{\mathrm{h}}$ candidate and the identified lepton is not meaningful in this case and the charge of the missed lepton is random, only one factor

$$f_{Z(+j\to\tau_{\mathrm{h}})} = f^{\mathrm{OS}}_{Z(+j\to\tau_{\mathrm{h}})} = f^{\mathrm{SS}}_{Z(+j\to\tau_{\mathrm{h}})}$$

[3] For the Boosted category, the requirement is relaxed to $p_T(H) > 50$ GeV, which largely improves the statistical uncertainty of the correction, but does not affect the $p_T(j \to \tau_{\mathrm{h}})$ distribution or the corresponding correction factor significantly.

is extracted from each control region. In addition, no significant differences between the simulated $N^{\text{OS}}_{Z(+j\to\tau_\text{h})}$ and $N^{\text{SS}}_{Z(+j\to\tau_\text{h})}$ are observed, so that the OS-r_{QCD}·SS estimate for this background is reduced to the term

$$f_{Z(+j\to\tau_\text{h})} \cdot N^{\text{OS}}_{Z(+j\to\tau_\text{h})} \cdot (1 - r_{\text{QCD}}) \,,$$

which is negative and thus subtracts an overestimated contribution from the same-sign data.

For top-quark production processes, correction factors are extracted from two different sets of 2-jet control regions with $p_T(j) > 30\,\text{GeV}$ and $m_\text{T} > 50\,\text{GeV}$: the first definition relies on the kinematic selection of $\sum p_T(j)$, while the second simply inverts the b-veto, which is applied as a suppression in the signal region definition. For both selections, consistent factors $f^{\text{OS}}_{\text{Top}}$ and $f^{\text{SS}}_{\text{Top}}$ are obtained, which are applied as corrections for the corresponding background contributions from top-quark production processes.

Fit of the r_{QCD} Extrapolation Factor

While in the previous analysis, the extrapolation factors r_{QCD} are derived from different control regions with inverted isolation requirements, the improved selection of $E^{\text{miss}}_\text{T} > 20\,\text{GeV}$ for all signal regions allows to define a single orthogonal multijet enriched control region with $E^{\text{miss}}_\text{T} < 15\,\text{GeV}$ and $m_\text{T} < 30\,\text{GeV}$. In this region, the dependence of r_{QCD} on the isolation criteria is now taken into account as systematic uncertainties by varying the isolation requirements and fitting the obtained ratios of selected opposite-sign events over same-sign events. After subtracting residual electroweak contributions from the data in the control regions, the OS/SS ratios r_{QCD} of the measured multijet contributions are then extracted as functions of the isolation variables. The linear fits of these dependencies and their extrapolations to the analysis requirements yield the values:

$$r_{\text{QCD}} = \begin{cases} 1.00 \pm 0.05\ (\text{stat} \oplus \text{syst}) \text{ for } e\tau_\text{h} \text{ final states} \\ 1.11 \pm 0.08\ (\text{stat} \oplus \text{syst}) \text{ for } \mu\tau_\text{h} \text{ final states} \end{cases}$$

Fake-Factor Estimate

For the tightened selection of the VBF category, an alternative background estimation method is implemented in this analysis, since the number $N^{\text{OS}}_{\text{Data}}$ of data events in the same-sign control region is too small to provide a sufficient estimate for a mass shape represented by a histogram with several bins. Instead of relying on the charge correlation of the mis-identified τ_h candidates, this method extrapolates their the mis-identification probability, which is extracted in form of so-called *fake-factors* $f^{\text{fake}}(p_T)$ for jet→τ_h mis-identifications in dependence of their transverse momentum $p_T(j) \to p_T(\tau_\text{h})$. A new control region is therefore defined via inverting the τ_h identification requirement, so that the extrapolation is performed from a so-called anti-τ_h region to the signal region with τ_h candidates fulfilling the standard identification criterion (id-τ_h):

$$N^{\text{id-}\tau_\text{h}}_{j\to\tau_\text{h}}(p_T) = f^{\text{fake}}(p_T) \cdot \left(N^{\text{anti-}\tau_\text{h}}_{\text{Data}}(p_T) - \sum_{p\in\mathcal{P}} N^{\text{anti-}\tau_\text{h}}_{p}(p_T) \right) \tag{6.2}$$

As in the standard sideband extrapolation, the contributions $N^{\text{anti-}\tau_\text{h}}_{p}$ from other processes $p \in \mathcal{P}$ are subtracted in the anti-τ_h region, so that the estimate $N^{\text{id-}\tau_\text{h}}_{j\to\tau_\text{h}}(p_T)$ covers all processes, which are not subtracted. While the standard same-sign sideband method can only estimate the multijet contribution in a data-driven way, the fake-factor method can include other backgrounds with fake-τ_h objects as well. Its implementation in this analysis therefore includes the W+jets process and provides a combined

estimate with the multijet contribution. The subtracted processes in the anti-τ_{h} region are therefore limited to the set

$$\mathcal{P} = (Z \to \tau\tau, Z \to \ell\ell_{(+j\to\tau_{\mathrm{h}})}, Z \to \ell\ell_{(\ell\to\tau_{\mathrm{h}})}, \text{di-boson, top}) \,,$$

which excludes W+jets. The mis-identification probability $f^{\mathrm{fake}}(p_T)$, however, is expected to be different for multijet and W+jets events, since the gluon-initiated jets are less likely to be identified as τ_{h} than quark-initiated jets from ISR in the W boson production: $f^{\mathrm{fake}}_{\mathrm{Multijet}} < f^{\mathrm{fake}}_{W\to\ell\nu}$. These fake-factors are thus extracted from separate control regions with high transverse mass $m_{\mathrm{T}} > 70$ GeV ($f^{\mathrm{fake}}_{W\to\ell\nu}$) or loose lepton isolation ($f^{\mathrm{fake}}_{\mathrm{Multijet}}$). They are split into 1-prong and 3-prong τ_{h} candidates and binned in $p_T(j \to \tau_{\mathrm{h}})$, since the jet$\to\tau_{\mathrm{h}}$ mis-identification probability is found to depend on these properties of the jets. In order to account for their difference, the extrapolation factor f^{fake} is formed as the weighted average of the two factors

$$f^{\mathrm{fake}} = (1 - r_W) \cdot f^{\mathrm{fake}}_{\mathrm{Multijet}} + r_W \cdot f^{\mathrm{fake}}_{W\to\ell\nu} \,,$$

for which the ratio r_W denotes the fraction of W+jets events expected in the anti-τ_{h} region:

$$r_W = \frac{N_{W\to\ell\nu}}{N_{\mathrm{Multijet}} + N_{W\to\ell\nu}}$$

This ratio depends on the event selection and must therefore be obtained for the signal region. It is extracted from the data in the high-m_{T} control region, which is then extrapolated into the signal region by applying the corresponding ratio extracted from simulated Monte Carlo events:

$$N_{W\to\ell\nu} = N^{\text{high-}m_{\mathrm{T}}}_{\mathrm{Data}} \cdot \frac{N_{W,\mathrm{MC}}}{N^{\mathrm{high}m_{\mathrm{T}}}_{W,\mathrm{MC}}}$$

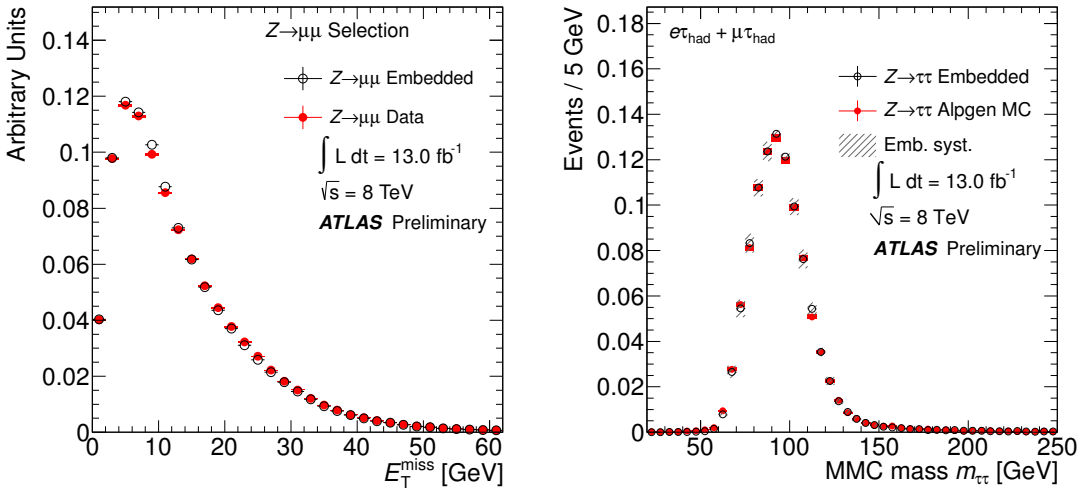
For the multijet estimate, this number is taken directly from the data in the signal region, after subtracting this data-driven $N_{W\to\ell\nu}$ and all other contributions estimated from embedding or simulation via

$$N_{\mathrm{Multijet}} = N_{\mathrm{Data}} - N_{W\to\ell\nu} - \sum_{p\in\mathcal{P}} N_{\mathrm{p}} \,.$$

Embedding Normalisation and Validation

Instead of normalising the event yield from the embedded samples to the corresponding number obtained from Monte Carlo simulation, a new $Z \to \tau\tau$ enriched control region with less than 0.1% expected signal contamination is defined with 40 GeV $< m_{\mathrm{vis}} <$ 70 GeV. In this region, the $Z \to \tau\tau$ contribution provided by the τ-embedded events is normalised to data, so that the embedded sample becomes independent of the theoretical cross section as opposed to the previous analysis. In order to correct for the efficiencies of the different triggers used in the analysis (single and combined, $e\tau_{\mathrm{h}}$ and $\mu\tau_{\mathrm{h}}$), four normalisation factors are extracted separately from control regions, which are defined according to the kinematic selection for the triggers as summarised in Tab. 6.8.

As for the validation of the embedding technique applied to the 7 TeV data (cf. Sec. 6.1.2), the corresponding checks are also performed with the 8 TeV data: Figure 6.7a demonstrates that the embedding technique preserves the $E^{\mathrm{miss}}_{\mathrm{T}}$ distribution when embedding the data with simulated $Z \to \mu\mu$ decays. In comparison to the ALPGEN+HERWIG Monte Carlo simulation in Fig. 6.7b, the MMC mass distribution is again found to be in good agreement. Further comparisons of the $Z \to \tau\tau$ and $Z \to \mu\mu$ embedding are shown in Fig. B.1-B.4.



(a) $E_{\mathrm{T}}^{\mathrm{miss}}$ ($Z \to \mu\mu$) (b) MMC mass ($Z \to \tau\tau$)

Figure 6.7: (a) $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions for the $Z \to \mu\mu$ data before and after muon embedding and (b) MMC mass distribution for the τ-embedded $Z \to \mu\mu$ data and simulated $Z \to \tau\tau$ events after preselection requirements. For (a), only statistical uncertainties are shown, while (b) also includes systematic uncertainties associated with the embedding procedure [164].

Background Summary

In the H +0-jet, H +1-jet and Boosted categories, the background expectations are obtained from the same-sign data and residual OS-r_{QCD}·SS contributions estimated from simulated Monte Carlo samples, while the $Z \to \tau\tau$ background is estimated from τ-embedded $Z \to \mu\mu$ data events. The estimated contributions from simulated W+jets, $Z \to \ell\ell(+j \to \tau_{\mathrm{h}})$ and top quark production are scaled with correction factors, which adjust their normalisation to the data observed in the dedicated control regions.

For the VBF category, the tightened di-jet selection cuts cause the standard background estimates to break down due to low numbers both of τ-embedded $Z \to \mu\mu$ data events and same-sign data events in the control region for the multijet estimate. Therefore, the $Z \to \tau\tau$ background in the VBF category is estimated with Monte Carlo samples, for which the simulated di-jet observables are carefully validated both in $Z \to \tau\tau$ against the τ-embedded $Z \to \mu\mu$ data and in $Z \to \ell\ell$ against corresponding di-jet events selected from the data. The multijet estimate is performed with the fake-factor method, which also includes the W+jets background. Similar to the other categories, the remaining opposite-sign estimates for the other background processes are obtained from simulated Monte Carlo samples after data-driven normalisation corrections extracted from control regions.

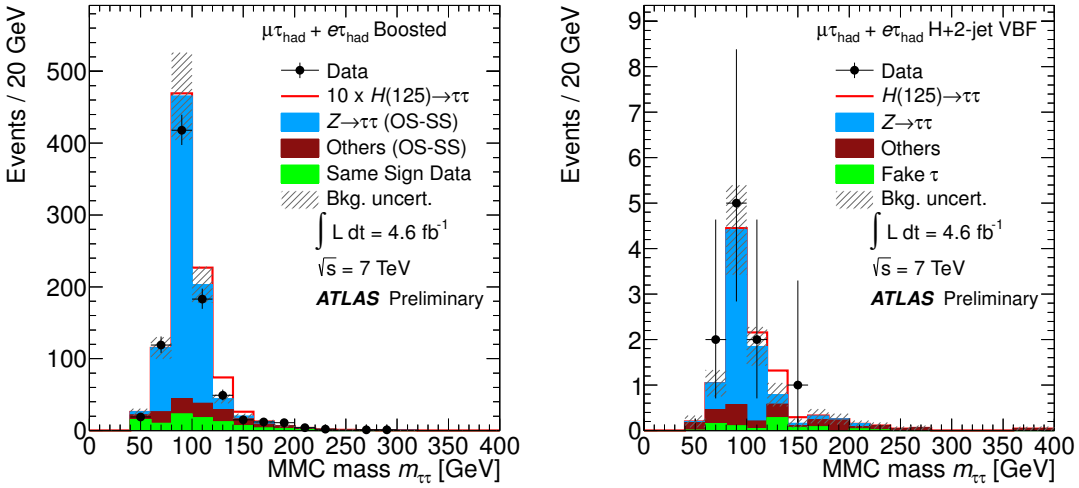
The estimated numbers of signal and background events for the most sensitive VBF and Boosted categories are summarised in Tab. 6.10 for the 7 TeV analysis and in Tab. 6.11 for the 8 TeV analysis. The corresponding MMC mass distributions for these categories are shown in Fig. 6.8 and Fig. 6.9. No significant excess of the data over the estimated backgrounds is observed, which is expected from the signal of a SM Higgs boson.

Process	Events	
	Boosted	VBF
$gg \to H$ (125 GeV)	4.1 ± 0.1 ± 1.0	0.17 ± 0.03 ± 0.06
VBF H (125 GeV)	1.52 ± 0.03 ± 0.13	0.87 ± 0.02 ± 0.15
VH (125 GeV)	0.86 ± 0.04 ± 0.08	<0.001
$Z \to \tau\tau^{\dagger}$	(0.70 ± 0.02 ± 0.10)×10^3	6.5 ± 0.6 ± 1.5
Di-boson †	8.4 ± 0.7 ± 0.8	0.12 ± 0.06 ± 0.03
$Z \to \ell\ell^{\dagger}$	3.7 ± 1.3 ± 1.0	0.8 ± 0.3 ± 1.0
Top †	52 ± 2 ± 9	1.2 ± 0.3 ± 0.1
W boson + jets (OS-SS)	41 ± 7 ± 8	–
Same sign data	90 ± 10 ± 5	–
Fake-τ_{had} backgrounds	–	0.8 ± 0.2 ± 0.4
Total background	(0.90 ± 0.02 ± 0.10)×10^3	9.5 ± 0.8 ± 1.9
Observed data	834	10

Table 6.10: Number of events in the Boosted and VBF categories of the 7 TeV analysis combined for the $e\tau_{\text{h}}$ and $\mu\tau_{\text{h}}$ channels [164]. The first uncertainty is statistical and the second systematic. For the backgrounds marked with a dagger (†), the values in the Boosted column indicate the (OS-SS) component.

Process	Events	
	Boosted	VBF
$gg \to H$ (125 GeV)	20.3 ± 0.7 ± 5.1	0.5 ± 0.1 ± 0.3
VBF H (125 GeV)	5.3 ± 0.2 ± 0.3	2.5 ± 0.2 ± 0.4
VH (125 GeV)	2.7 ± 0.2 ± 0.2	<0.001
$Z \to \tau\tau^{\dagger}$	(1.78 ± 0.03 ± 0.11)×10^3	17 ± 2 ± 6
Di-boson †	12.2 ± 0.9 ± 1.0	0.6 ± 0.3 ± 0.4
$Z \to \ell\ell^{\dagger}$	18 ± 9 ± 4	1.7 ± 0.5 ± 1.2
Top †	111 ± 8 ± 33	2.0 ± 0.7 ± 1.0
W boson + jets (OS-SS)	(0.27 ± 0.06 ± 0.04)×10^3	–
Same sign data	(0.34 ± 0.02 ± 0.01)×10^3	–
Fake-τ_{had} backgrounds	–	7.6 ± 0.7 ± 3.8
Total background	(2.53 ± 0.07 ± 0.13)×10^3	29 ± 2 ± 7
Observed data	2602	29

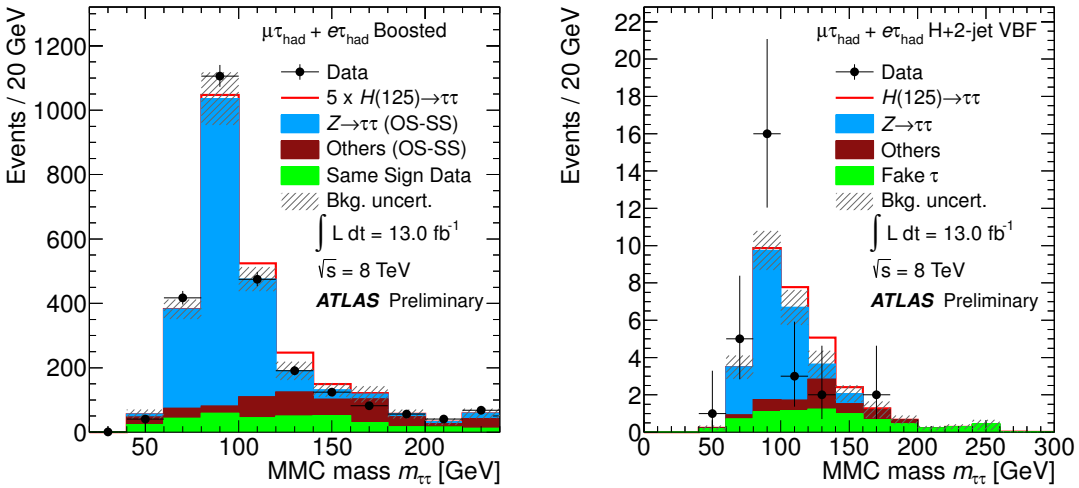
Table 6.11: Number of events in the Boosted and VBF categories of the 8 TeV analysis combined for the $e\tau_{\text{h}}$ and $\mu\tau_{\text{h}}$ channels [164]. The first uncertainty is statistical and the second systematic. For the backgrounds marked with a dagger (†), the values in the Boosted column indicate the (OS-SS) component.



(a) Boosted category

(b) VBF category

Figure 6.8: MMC mass distributions of the selected events in the Boosted and VBF categories of the $H \to \tau_{\text{lep}}\tau_{\text{had}}$ channel for the 7 TeV analysis [164]. The selected events in data are shown together with the predicted Higgs boson signal ($m_H = 125$ GeV) stacked above the background contributions. For illustration, the signal contributions in the Boosted category have been scaled by a factor 10. The last bin in the histograms contains the overflow.



(a) Boosted category

(b) VBF category

Figure 6.9: MMC mass distributions of the selected events in the Boosted and VBF categories of the $H \to \tau_{\text{lep}}\tau_{\text{had}}$ channel for the 8 TeV analysis [164]. The selected events in data are shown together with the predicted Higgs boson signal ($m_H = 125$ GeV) stacked above the background contributions. For illustration, the signal contributions in the Boosted category have been scaled by a factor 5. The last bin in the histograms contains the overflow.

6.2.3 Systematic Uncertainties

The treatment of systematic uncertainties in this analysis differs only slightly from the description in Sec. 6.1.4, so that the following summary focuses on the main changes with respect to the previous analysis.

Detector-Related Uncertainties: Improvements in the determination of the energy scale uncertainties for jets and τ_{h} candidates include *in-situ* calibration procedures in addition to the results from single hadron response measurements. The τ_{had} energy scale is further validated with the data in $Z \to \tau\tau$ enriched control samples. Due to the new and more independent methods to determine the uncertainties, their corresponding effects on the visible kinematics and $E_{\mathrm{T}}^{\mathrm{miss}}$ are treated as uncorrelated between jets and identified τ_{h} objects. Although it cannot be excluded that a certain level of correlation is present, this choice for the treatment is expected to be more conservative than assuming a full correlation as before.

Background Modelling Uncertainties: For the improved sideband subtraction technique, the systematic uncertainty on the expected event yield is evaluated by varying the measured values for the correction factors within their statistical uncertainties, while the separate uncertainties for opposite-sign and same-sign factors are anti-correlated in order to obtain conservative estimates for the OS$-r_{\mathrm{QCD}}\cdot$SS background yield in the signal regions. The extrapolation factor r_{QCD} is varied within its total uncertainty, which includes systematic effects estimated by applying different anti-isolation criteria.

For the fake-factor estimate, a conservative systematic uncertainty of 40% is applied to the estimated event yield. This value is obtained by varying the measured fraction of W+jets events r_W within the interval [0, 1], which effectively represents the difference between the fake-factors measured for multijet and for W+jets events.

Summary of Dominant Systematic Uncertainties: Table 6.12 shows a summary of the dominant systematic uncertainties for the $Z \to \tau\tau$ background component and the expected signal. The given ranges account for the different systematic effects in the individual analysis categories.

$Z \to \tau\tau$	
Embedding	2–4% (S)
Tau Energy Scale	4–15% (S)
Tau Identification	4–5%
Trigger Efficiency	2–5%
Normalisation	4% (non-VBF), 16% (VBF)
Signal	
Jet Energy Scale	3–9% (S)
Tau Energy Scale	2–9% (S)
Tau Identification	4–5%
Theory	18–23%

Table 6.12: Summary of $Z \to \tau\tau$ background and signal systematic uncertainties by channel [164]. The quoted ranges refer specifically to the 8 TeV dataset, but they are similar for the 7 TeV dataset. Uncertainties indicated with (S) are also applied bin-by-bin and therefore affect the shape of the final distributions. Signal systematic uncertainties are derived from the sum of all signal production modes.

6.2.4 Results

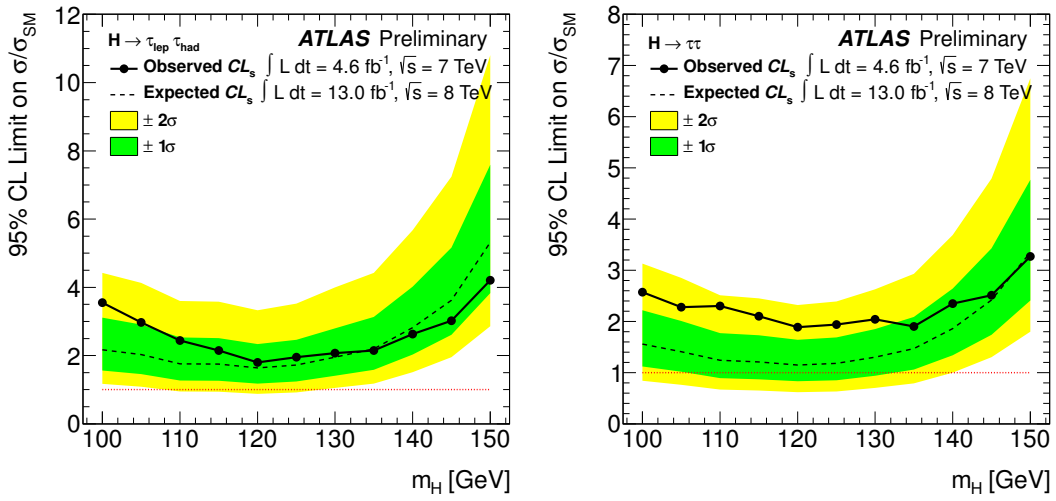
Exclusion limits on the signal strength parameter μ are set as described in Sec. 4.3.3 by rejecting the corresponding signal (H_1) hypothesis with a signal scaled by a strength parameter of μ. The obtained limits are shown in Fig. 6.10 for the $H \to \tau_{\text{lep}}\tau_{\text{had}}$ channel as well as for the combination with the $H \to \tau_{\text{lep}}\tau_{\text{lep}}$ and $H \to \tau_{\text{had}}\tau_{\text{had}}$ channels. For a mass of $m_H = 125$ GeV, the analysis of the $H \to \tau_{\text{lep}}\tau_{\text{had}}$ channel yields an observed (expected) exclusion limit of $\mu = \sigma/\sigma_{\text{SM}} < 2.0$ ($\mu < 1.7$), which is improved to $\mu < 1.9$ ($\mu < 1.2$) by the channel combination.

Although the observed data provide a weaker exclusion than expected from the Asimov data created for the signal scenario ($\mu = 1$), the probability of a statistical fluctuation to create a similar or even larger excess is found to be $p_0 = 13.5\%$, which is well consistent with the SM background-only hypothesis (H_0). This probability corresponds to a local observed significance of $Z_{\text{obs}} = 1.1\,\sigma$ with a best fit value of $\hat{\mu} = 0.7 \pm 0.7$ for the measured signal strength. The excess of data above the background-only expectation ($\mu = 0$) is thus even smaller than expected from a SM Higgs boson signal with $m_H = 125$ GeV, for which a significance of $Z_{\text{exp}} = 1.7\,\sigma$ is obtained from the Asimov data.

Optimisation Improvements

As a result of the optimisation, the sensitivity of the analysis has been improved significantly from an exclusion limit of $\mu < 5.9$ to $\mu < 3.0$ in the $H \to \tau_{\text{lep}}\tau_{\text{had}}$ channel for 7 TeV data only. Besides the gain from the tightened VBF selection with a higher signal to background ratio, which was only possible to achieve due to the new fake-factor background estimate, an important contribution to this improvement is provided by the newly included Boosted category. Its selection of events with high-$p_T(H)$ better exploits the recoil of the Higgs boson against additional jet radiation than the $H+1$-jet category, which is still included separately with almost no loss of sensitivity compared to the previous analysis. The new category now contains an improved ratio of signal events with high $p_T(H)$ to backgrounds with correspondingly high $p_T(Z)$ or $p_T(\ell + \tau_{\text{h}} + E_{\text{T}}^{\text{miss}})$, while a better mass resolution of these boosted events provides a better separation between signal and background in the sensitive region of the mass distribution.

A more conceptual improvement of this analysis is the implementation of a new background suppression cut, which depends on two variables at the same time. By defining the selection cut on $\Delta(\Delta R)$ of the visible τ decay products in dependence on the transverse momentum $p_T(H)$ of the $\tau\tau$ system, the correlation of these two variables is exploited, since it is found to be different for events with real $\tau\tau$ and background events with mis-identified τ_{h} candidates. The two-dimensional cut can therefore reject background events with a significantly higher signal efficiency than any one-dimensional selection of ΔR or $p_T(H)$. While more such correlations are certainly present in the observables considered in this analysis, their investigation and the definition of further two- or more-dimensional selection cuts cannot be performed with manual methods. Instead, multi-variate methods have been developed for such applications in high energy physics. These scan the multi-dimensional phase space very efficiently and automatically provide an optimised classification of events as signal- or background-like by exploiting their different correlations. Since even the optimised sensitivity of the presented cut-based analysis is not sufficient to expect an evidence or exclusion of $H \to \tau\tau$ in the near future, a multi-variate strategy is followed for the final analysis of the full dataset and is described in the next chapter. In parallel, also the cut-based analysis is maintained for the full dataset and is documented in [148, 168].



(a) $H \to \tau_{\text{lep}}\tau_{\text{had}}$ (b) Combined

Figure 6.10: Observed (solid) and expected (dashed) upper limits at 95% confidence level on the signal strength $\mu = \sigma/\sigma_{\text{SM}}$ as a function of the Higgs boson mass [164]. Expected limits are obtained from the Asimov data created with $\mu = 0$. The bands around the dashed line indicate the expected statistical fluctuations of the observed limit. Results are given for the $H \to \tau_{\text{lep}}\tau_{\text{had}}$ channel (a) and for the combination with the $H \to \tau_{\text{lep}}\tau_{\text{lep}}$ and $H \to \tau_{\text{had}}\tau_{\text{had}}$ channels (b).

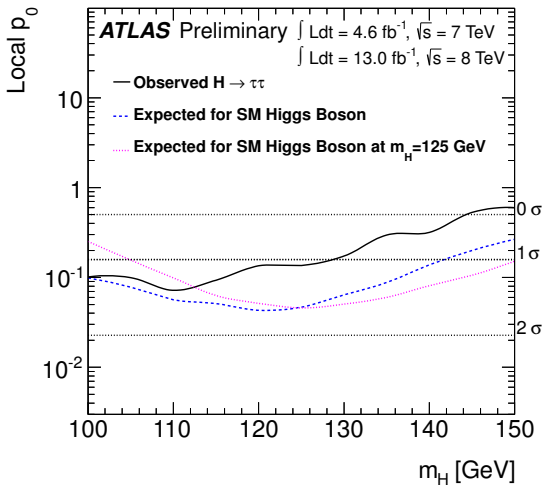


Figure 6.11: The observed (solid) and expected (dashed) p_0-value as a function of the Higgs boson mass [164]. The dotted line shows the injected p_0, which is calculated from Asimov data created by including a SM Higgs boson signal with m_H =125 GeV and strength $\mu = 1$ to the background expectation.

CHAPTER 7

MVA-Based Search for SM $H \to \tau\tau$ Decays with 20.3 fb^{-1} of ATLAS Data

The search for the SM Higgs boson as presented so far mainly relied on event selections, which applied simple one-dimensional cuts on kinematic event observables. Despite a significant improvement of the signal exclusion achieved with more refined selection cuts, a simple extrapolation[1] of their expected signal sensitivity does not reach the important threshold of 3σ even with the full dataset of 20.3 fb^{-1} recorded with ATLAS in 2012 at $\sqrt{s} = 8$ TeV.

In order to further increase the sensitivity to a SM Higgs boson, the analysis strategy has been changed to rely on multi-variate methods, which automatically exploit correlations of analysis observables and can thus improve the signal and background discrimination and the expected sensitivity. This automated process, however, must be *trained* on samples of signal and background events, from which a set of observables must be selected for the analysis. The choice of these observables must be optimised, while a detailed validation must be performed in order to ensure an accurate modelling of all background components and the signal.

From the various existing multivariate analysis methods, the Boosted Decision Trees (BDTs) [169] are chosen for this analysis, since they have been proven to provide good separation performance with a robust classifier and to be applicable for the situation of several backgrounds estimated with different data-driven techniques. As a technical implementation of this method, the Toolkit for Multivariate Data Analysis (TMVA) [170] is used, which is included in ROOT.

With the background estimation techniques developed in the cut-based analysis, a new BDT-based analysis is designed, which maintains a general categorisation, but relies on the separation power of the BDT output to define signal dominated regions of the phase space. Although results on the data recorded in 2011 at $\sqrt{s} = 7$ TeV have been published before, and a complete analysis of the complete Run-1 dataset of (4.6 + 20.3) fb^{-1} is envisaged for the future, the results presented here are restricted to the full dataset from the data taking period in 2012 at $\sqrt{s} = 8$ TeV [171], which corresponds to an integrated luminosity of 20.3 fb^{-1}.

[1] Such an extrapolation can e.g. be based on the reduction of statistical uncertainties $\propto \sqrt{(4.6 + 20.3)/(4.6 + 13.0)}$.

7.1 Boosted Decision Trees

Boosted Decision Trees (BDTs) denote an ensemble of binary trees, which are automatically generated from decision cuts on a defined set of observables. They are usually built from samples of background and signal events and can afterwards provide a corresponding classification of a signal-likeliness for any event based on the values of its observables. Starting from a first generated tree, the mis-classified events are weighted up in the training samples, so that the following tree can *learn* from the mistakes of the previous and is thus *boosted* by the weighting procedure. The generation of the trees is then iterated in order to obtain a large number of different classifiers, which are combined as a weighted average into a single estimator. For each event, this can be interpreted as an information on how similar it is to the signal or background events provided for the training based on the decisions of all generated trees.

7.1.1 Automated Generation of Decision Trees

From a given set of N input observables, the training algorithm selects the one with the best separation between signal and background and determines the best cut-value according to the so-called Gini-Index[2] $G(p) = p \cdot (1 - p)$. This has a maximum for a perfect mix of signal and background with purity $p = 0.5$, and it is minimal for an optimal separation with perfect purity $p = 1$. The optimal cut is stored as a binary decision as the first node of the tree and creates sub-samples of signal-like and background-like events. From both resulting branches, the tree is then grown by defining further and similarly optimised cuts on other input observables until either a limited depth N_{max} of the tree or a minimal number of training events in a leaf node is reached. As compared to a single signal region with N cuts on the selected observables, the final tree consists of several signal enriched regions, which in general are N_{max}-dimensional hypercubes of the N-dimensional phase space of input observables.

For $N_{\text{max}} = N$, the most signal-like leaf node can be identical to a single cut-based signal region as could be obtained from a similar optimisation method. As shown in Fig. 7.1, however, the decision tree recovers additional sensitivity also from other leaf nodes with high signal purity, which are e.g. considered as background-like in only one of the decisions. These can still be considered as signal regions, since all subsequent decisions will still be chosen to obtain the best signal separation. With this step-by-step optimisation, the decision on the second observable depends on the selection applied to the first, so that correlations of the observables are taken into account.

7.1.2 Machine Learning via Boosting

While already the first tree can provide an improved sensitivity as compared to a simple cut-based approach, the mis-classification of events can additionally be exploited to achieve a learning effect. Therefore, each mis-classified event (i.e. signal in background dominated region or vice versa) is weighted up with respect to the other correctly classified events, so that the next tree is then generated from training samples modified by a re-weighting scheme, which depends on the performance of the first one. Since the mis-classified events are thus considered with enhanced importance during the generation of the new tree, it can result in different decisions and cut-values. This *boosting* is repeated for a large number of such trees $N_{\text{trees}} = O(100 - 1000)$, so that the learning effect is focused on the separation of those events, which are very similar in the provided signal and background events.

[2] Other separation criteria such as the cross entropy $-p \cdot \ln p - (1 - p) \cdot \ln(1 - p)$ or the statistical significance $S / \sqrt{S + B}$ have been found to provide similar performance.

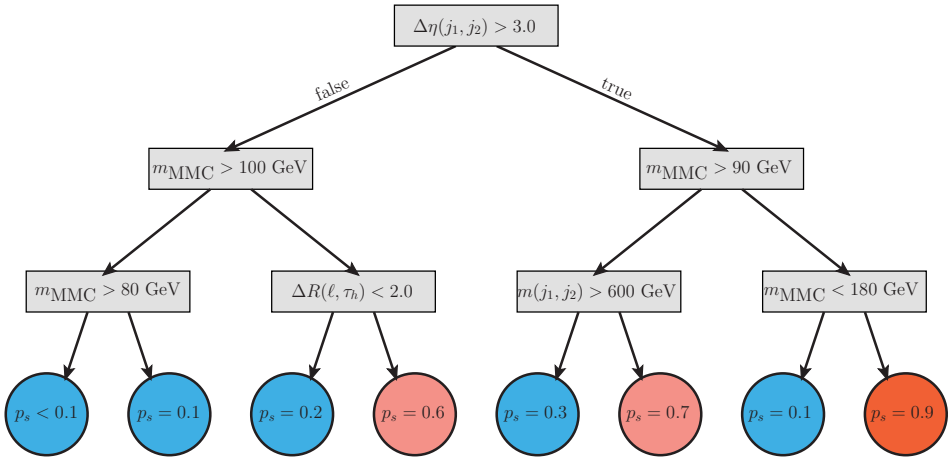


Figure 7.1: Simplified example of a single decision tree with $N_{\mathrm{max}} = 3$. In addition to events in the leaf node with the highest signal purity $p_s = 0.9$, also events in other nodes with $p_s > 0.5$ are considered signal-like.

For each event, which can be represented as a vector $\vec{x}$ of its input observables, the signal purity $p_i(\vec{x})$ of the leaf node that selects $\vec{x}$ depends on the number i of the corresponding tree. This purity can be interpreted as the probability that a randomly selected event in this leaf node is a signal event and can thus be considered as the signal-likeliness of any event, which is selected by this node.

With a simple transformation to the interval $[-1, +1]$, this purity is used as the classifier[3]

$$h_i(\vec{x}) = 2 \cdot p_i^s(\vec{x}) - 1$$

of tree i for event $\vec{x}$. From the weighted sum of all classifiers over all trained trees, the so-called BDT score $y_{\mathrm{boost}}(\vec{x})$ can be extracted as a final discriminant for any data or simulated event:

$$y_{\mathrm{boost}}(\vec{x}) = \frac{1}{N_{\mathrm{trees}}} \sum_{i \in \mathrm{trees}} w_i \cdot h_i(\vec{x})$$

Here, the w_i correspond to the weights applied in the boosting process to the mis-classified events, which can be chosen in different ways.

Adaptive Boost

The boost weights w_i can in general be extracted from the minimisation of a *loss-function*

$$L\left(y_{\mathrm{boost}}(\vec{x}), y_{\mathrm{true}}(\vec{x})\right) \ ,$$

which represents the difference between the boosted $y_{\mathrm{boost}}(\vec{x})$ and the true $y_{\mathrm{true}}(\vec{x}) = +1(-1)$ of an event provided by the signal (background) training sample. The choice of such a loss-function with $L\left(y_{\mathrm{boost}}, y_{\mathrm{true}}\right)) \to 0$ for $y_{\mathrm{boost}} \to y_{\mathrm{true}}$ determines the boosting process by providing the weights w_i for the calculation of the score $y_{\mathrm{boost}}(\vec{x})$ and for the re-weighting of mis-classified events in the training samples. The most popular function is the exponential loss $L\left(y(\vec{x}), h(\vec{x})\right) = \exp\left(-y(\vec{x}) \cdot h(\vec{x})\right)$, which re-

[3] Other implementations simply assign $h_i^s(\vec{x}) = +1$ (-1) for an event classified as signal (background) in tree i. For this analysis, the TMVA option UseYesNoLeaf=False is used, which enables the purity weighted classification instead of the binary decision of signal or background.

sults in the standard *AdaBoost* (*Adaptive Boost*) [172] weighting method. Due to the simple exponential function, weights of the form

$$w_i = \beta \cdot \ln(\alpha_i) \quad \text{with} \quad \alpha_i = \frac{1-m}{m}$$

can be obtained directly from the mis-classification probability m of the previous tree, where β is a free parameter to control the learning rate.

Gradient Boost

While the AdaBoost allows to define the weights with a simple algorithm due to a special choice of $L(y_{\text{boost}}(\vec{x}), y_{\text{true}}(\vec{x}))$, other more complicated weighting schemes can be obtained for any loss-function by performing the minimisation with a gradient method. With different choices of this function, this *GradientBoost* [173] method can provide a more robust behaviour as compared to the AdaBoost, which in certain cases is found to have instabilities in the boosting performance as well as unwanted features in its final discriminant[4]. For this analysis, the TMVA implementation [170] of the GradientBoost with a binomial log-likelihood loss of the form

$$L(y(\vec{x}), h(\vec{x})) = \ln\left(1 + e^{-2y(\vec{x})h(\vec{x})}\right)$$

is used, which is found to be more robust with a performance comparable to AdaBoost. Instead of the parameter β, the learning rate of the gradient boosting can be controlled with a *shrinkage* parameter, which can affect both the robustness and the performance of the trained BDT.

Overtraining Effects

While the performance of a BDT is usually driven by its learning process from the training samples, it is important to note that this learning must be focused on the distinction of signal and background based on their different physics features. As soon as the training focuses on statistical fluctuations in the provided event samples, any performance gain on the training sample will result in a loss of performance for statistically independent event samples as well as for the data. In any application, the BDT scores must therefore be calculated for statistically independent samples (*testing*), which are expected to provide a good modelling of the data. Extensive tests have shown that especially very large (deep) trees are likely to suffer from strong *overtraining* effects, but that this degradation is significantly smaller when using a large number of shallow trees (*weak learners*). For each specific application, however, the exact setup for the boosting can be optimised to obtain the best performance on testing samples as well as the data.

7.2 Datasets and Event Selection

Data events with $e\tau_{\text{h}}$ ($\mu\tau_{\text{h}}$) final states are selected from data with a single muon (electron) trigger and a combined $\mu\tau_{\text{h}}$ ($e\tau_{\text{h}}$) trigger.

Depending on the trigger recording the event, different transverse momentum thresholds are applied to the analysis objects, which are summarised in Tab. 7.1. As for the cut-based analysis, electrons and muons are selected with standard `tight` quality criteria for the final $\ell\tau_{\text{h}}$ selection, which are loosened

[4] Especially for the signal extraction from the BDT score, both the ranges of the scores for signal and background and their separation cannot be controlled very well.

for the suppression of mis-identified $\ell \to \tau_{\mathrm{h}}$ candidates and multi-lepton final states. Fake leptons dominantly contributed by multijet events are suppressed by isolation requirements of $I_{ET}^{0.2}/p_T < 0.06$ on the surrounding calorimeter energy and $I_{PT}^{0.4}/p_T < 0.06$ on the tracks within $\Delta R < 0.4$ of the electrons or muons.

Candidates of τ_{h} objects are identified with a BDT method, for which a working point with `medium` signal efficiency of 55–60% and a typical mis-identification probability of 1–2% is chosen. These candidates are not considered for the analysis if they overlap with a `loose` muon with $p_T(\mu) > 4$ GeV or a `loose` electron with $p_T(e) > 15$ GeV. In addition to the electrons used in the analysis, this overlap removal is extended to candidates in the calorimeter transition region ($1.37 < |\eta(e)| < 1.52$), which must fulfil `medium` requirements in order to remove overlapping τ_{h} candidates.
If more than one electron or muon is found fulfilling $p_T(e, \mu) > 15$ GeV and the reduced quality criteria, the event is rejected as potential contribution from a $Z \to \ell\ell$+jets or top-quark production process. In addition, the dominant part of the W+jets contribution is rejected by selecting only events with low transverse mass $m_T(\ell) < 70$ GeV. All preselected events with exactly one identified τ_{h} candidate and an oppositely charged electron (muon), which fulfil these selection criteria and the transverse momentum thresholds, are separated further into four categories depending on the jet topology and trigger requirement.

Trigger	Trigger p_T Threshold (GeV)	Analysis p_T Threshold (GeV)
single electron	$p_T(e) > 24$	$p_T(e) > 26$
	–	$p_T(\tau_{\mathrm{had}}) > 20$
single muon	$p_T(\mu) > 24$	$p_T(\mu) > 26$
	–	$p_T(\tau_{\mathrm{had}}) > 20$
combined $e + \tau_{\mathrm{had}}$	$p_T(e) > 18$	$20 < p_T(e) < 26$
	$p_T(\tau_{\mathrm{had}}) > 20$	$p_T(\tau_{\mathrm{had}}) > 25$
combined $\mu + \tau_{\mathrm{had}}$	$p_T(\mu) > 15$	$17 < p_T(\mu) < 26$
	$p_T(\tau_{\mathrm{had}}) > 20$	$p_T(\tau_{\mathrm{had}}) > 25$

Table 7.1: The triggers used to select $\ell\tau_{\mathrm{h}}$ final states from the pp-collision data with the corresponding p_{T} thresholds of the trigger and the applied selection to the reconstructed objects.

7.3 Analysis Categorisation

As in the cut-based analysis, a basic event categorisation is maintained for this multi-variate analysis, so that dedicated BDTs can be trained on the different signal topologies. For each category, the set of input observables is later chosen appropriately, and the boosting parameters are adjusted separately, which results in a better performance as compared to a single BDT trained with all possible signal topologies at the same time.

7.3.1 VBF Category

Events recorded by the single lepton trigger are selected in the VBF category if they contain at least two jets, of which the leading one is required to have a large transverse momentum $p_T(j) > 50$ GeV as well as a large separation $\Delta\eta(j_1, j_2) > 3.0$ from the sub-leading jet.

7.3.2 Boosted Category

Failing the VBF selection, events triggered by a single lepton enter the Boosted category if the reconstructed momentum of the Higgs boson is large: $p_T(H) > 100$ GeV.

7.3.3 1-jet and Rest Categories

Events fulfilling neither the VBF nor the boost requirement are considered in the 1-jet category if at least one jet ($p_T(j) > 30$ GeV) is found. Without any jet, an event enters the Rest category. Since the VBF and the Boosted category veto events selected with the combined triggers, the 1-jet and Rest categories include these events with lower transverse momenta of the lepton or τ_{h} candidate, which are either VBF-like or have a boosted Higgs boson. Although the final analysis is performed only on the Boosted and the VBF categories, control distributions from the Rest and 1-jet categories are included for completeness. These provide an important validation of all background models in a larger part of the phase space and thus give further confidence in the robustness of the analysis.

7.3.4 Additional Background Rejection

In addition to the inclusive categorisation, which covers all preselected events, further selection criteria are applied and completely remove events from the analysis:

- In the VBF and Boosted categories, a b-jet veto is applied, which removes events with a b-tagged jet ($p_T(j) > 30$ GeV).
- For the VBF selection, only events with $m_{\text{vis}} > 40$ GeV are considered, while the low-mass region is excluded. This is dominated by contributions from fake backgrounds, which are found to be mis-modelled due to threshold effects in the data-driven background estimate and are thus not considered in the final event selection.

7.4 Background Estimation

Backgrounds in this analysis are estimated with the techniques described in Sec. 6.2.2. The OS$-r_{\text{QCD}}\cdot$SS method is thus applied at preselection level as well as for the 1-jet and Rest categories, while fake backgrounds in the Boosted and VBF categories are estimated with the fake-factor method.

7.4.1 VBF-Filtered Monte Carlo Samples

The simulated Monte Carlo samples, which are used for the background estimation, are summarised in Tab. 7.2 together with their corresponding cross sections. As for the previous cut-based analysis (cf. Sec. 6.2.1), dedicated Z+jets samples with filters applied to jets on particle-level are included in order to improve the statistical uncertainty related to the low selection efficiency for VBF-like topologies. In addition to the previously used samples filtered with $m(q_1, q_2) > 200$ GeV and $\Delta\eta(q_1, q_2) > 2.0$, new samples are created with a tighter filter selecting $m(q_1, q_2) > 400$ GeV and $\Delta\eta(q_1, q_2) > 4.0$. Due to a looser VBF selection in the categorisation, both these must be merged together with the non-filtered samples into a hybrid dataset, so that the full phase space can be covered. This merging is performed by successively excluding the phase space in the non-filtered (looser-filtered) sample, which is covered by the (tighter) filtered sample. Since the contribution of soft jets to events with VBF-like topologies is expected to be small, filtered samples are only considered if they contain at least two hard partons in

the matrix element calculation, whereas the default unfiltered samples are used for lower parton multiplicities.

While contributions from W+jets are estimated in a data-driven way for the final analysis, the training of BDTs is performed with simulated samples, for which the number of events with VBF-like topology is also enhanced with a filtered sample requiring $(q_1, q_2) > 400$ GeV and $\Delta\eta(q_1, q_2) > 3.0$. Since the τ_h candidate is expected to be a jet, the corresponding filter requirements are only used to enhance the phase space with three hard partons originating from the matrix element calculation. Lower parton multiplicities are again covered with the default samples. The merging of the filtered with the inclusive samples is performed in a similar way as for Z+jets.

7.4.2 Estimation of the r_{QCD} Extrapolation Factor

As in the optimised cut-based analysis, the OS/SS ratio r_{QCD} is again extracted from a multijet enriched control region ($E_{\mathrm{T}}^{\mathrm{miss}} < 15$ GeV, $m_{\mathrm{T}}(\ell) < 30$ GeV), but its dependence on the lepton isolation is taken into account with a linear fit. For the control region, the τ_h identification criteria are relaxed to the `loose` working point, and the calorimeter isolation of the muons (electrons) is inverted to $I_{ET}^{0.2}/p_T > 0.06$ ($I_{ET}^{0.2}/E_T > 0.06$). The threshold of this isolation cut is varied, and the dependence of the OS/SS ratio is fitted with a linear function as shown in Fig. 7.2. The value of r_{QCD} is then estimated via the extrapolation to the isolation selection of the signal region. From the fitted range $0.06 < I_{ET}^{0.2}/p_T < 0.38$ ($0.06 < I_{ET}^{0.2}/E_T < 0.25$), the final values of

$$r_{\mathrm{QCD}} = \begin{cases} 1.00 \pm 0.05\ (\mathrm{stat}) \pm 0.12\ (\mathrm{syst}) \text{ for } e\tau_h \text{ final states} \\ 1.10 \pm 0.06\ (\mathrm{stat}) \pm 0.13\ (\mathrm{syst}) \text{ for } \mu\tau_h \text{ final states} \end{cases}$$

are extracted. As compared to the corresponding measurement in Sec. 6.2.2, significantly larger systematic uncertainties are obtained by varying the fit range[5], the track isolation requirements and the τ_h identification working point. For the combined categories, an average value of 1.05 ± 0.15 is applied, for which the systematic uncertainty is chosen conservatively.

7.4.3 Fake-Factor Estimation

In the Boosted and VBF categories, the contribution from backgrounds with mis-identified τ_h candidates is estimated with the fake-factor method as described in Sec. 6.2.2. The corresponding fake-factors and their uncertainties are obtained separately for the Boosted and VBF selection from W+jets and multijet enriched control regions defined by inverting the `medium` τ_h identification criterion. Figure 7.3 shows the dependence of these factors f^{fake} on the transverse momentum of the $p_T(j{\to}\tau_h)$ candidate separated into the significantly different 1-prong and 3-prong τ_h candidates. In addition to the statistical uncertainties of the separate factors $f^{\mathrm{fake}}_{\mathrm{Multijet}}(p_T)$ and $f^{\mathrm{fake}}_{\mathrm{W+jets}}(p_T)$ arising from the data in the control regions, a systematic uncertainty is assigned to their weighted mean:

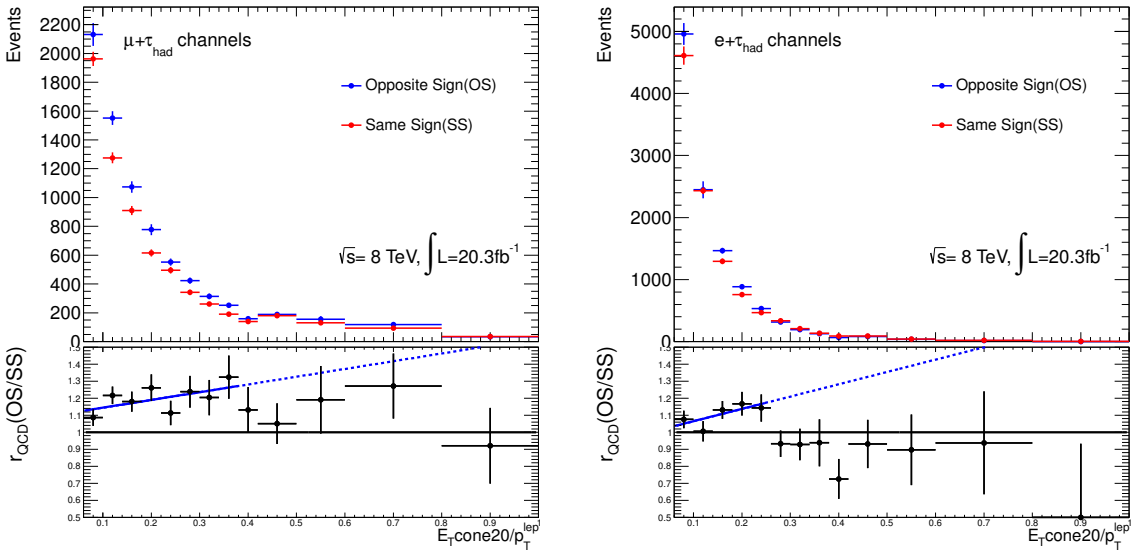
$$f^{\mathrm{fake}} = (1 - r_W) \cdot f^{\mathrm{fake}}_{\mathrm{Multijet}} + r_W \cdot f^{\mathrm{fake}}_{W\to\ell\nu}$$

This is estimated by varying the ratio r_W between the values 0 and 1, which correspond to the cases that the fakes in the signal region are either completely dominated by gluon-initiated jets from multijet

[5] Although the assumption of a linear behaviour is not well motivated by physics, any such extrapolation is expected to be dominated by the value close to the threshold. Therefore, the variation of the range for the linear fit gives a sufficient estimate of the systematic uncertainty, while the variation of the function is found to be unnecessarily complicated.

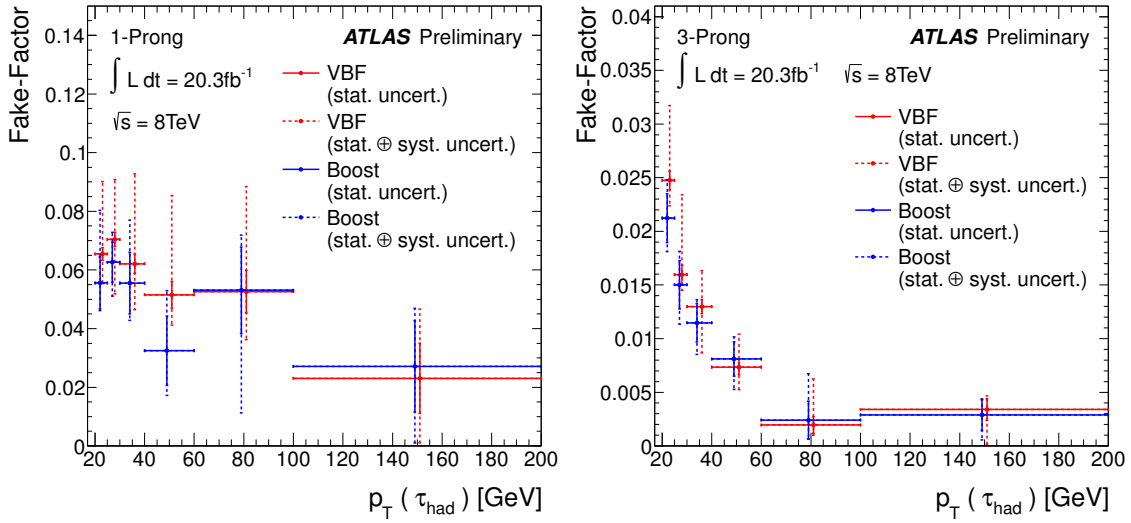
Process	Generator	PDF	$\sigma\times$ BR$\times\epsilon_f$/pb
$W \to \ell$+jets ($\ell = e, \mu, \tau$)	ALPGEN+HERWIG	CTEQ6L1	12.22×10^3
$W \to \ell$+jets ($m(q_1, q_2) > 400$ GeV, $\Delta\eta(q_1, q_2) > 3.0$)	ALPGEN+PYTHIA	CTEQ6L1	35.4
$Z/\gamma^* \to \ell\ell$+jets (10< $m_{\ell\ell}$ <60 GeV)	ALPGEN+PYTHIA	CTEQ6L1	4.35×10^3
$Z/\gamma^* \to \ell\ell$+jets ($m_{\ell\ell}$ >60 GeV)	ALPGEN+PYTHIA	CTEQ6L1	1.15×10^3
$Z/\gamma^* \to \ell\ell$+jets ($m(q_1, q_2) > 200$ GeV, $\Delta\eta(q_1, q_2) > 2.0$)	ALPGEN+HERWIG	CTEQ6L1	11.2
$Z/\gamma^* \to \ell\ell$+jets ($m(q_1, q_2) > 400$ GeV, $\Delta\eta(q_1, q_2) > 4.0$)	ALPGEN+HERWIG	CTEQ6L1	2.9
$qqZ \to \ell\ell$ (VBF production)	SHERPA	CTEQ6L1	0.36
$t\bar{t}$-production	MC@NLO	CT10	238.1
Single top: t-channel	AcerMC+PYTHIA	CTEQ6L1	28.4
Single top: s-, Wt-channel	MC@NLO	CT10	1.8, 22.4
WW, $gg \to WW$	ALPGEN+HERWIG, gg2WW+PYTHIA	CTEQ6L1, CT10	29.2, 0.2
WZ, ZZ	HERWIG	CTEQ6L1	6.8, 1.6
$ggH(\to \tau_{\text{lep}}\tau_{\text{had}})$	POWHEG+PYTHIA	CT10	0.56
$qqH(\to \tau_{\text{lep}}\tau_{\text{had}})$	POWHEG+PYTHIA	CT10	0.045
$WH(\to \tau_{\text{lep}}\tau_{\text{had}})$	PYTHIA	CTEQ6L1	0.020
$ZH(\to \tau_{\text{lep}}\tau_{\text{had}})$	PYTHIA	CTEQ6L1	0.011

Table 7.2: Summary of Monte Carlo samples and cross sections multiplied by the relevant branching ratios (BR), k-factors and filtering efficiencies ϵ_f used in the multi-variate analysis. In the case of the decays to leptons the cross section is given per lepton channel. Numbers for $t\bar{t}$, single-t, and di-boson processes do not include fully hadronic final states. Cross sections for the signal processes are listed for m_H=125 GeV, while samples are generated for different mass points in the range $m_H \in [100\text{ GeV}, 150\text{ GeV}]$ in steps of 5 GeV.



(a) (b)

Figure 7.2: Linear fits of r_{QCD} dependence on (a) $I_{ET}^{0.2}/p_T$ for $\mu\tau_{\mathrm{h}}$ and (b) $I_{ET}^{0.2}/E_T$ for $e\tau_{\mathrm{h}}$ events. The bottom plots show the ratio of OS and SS events for each bin. The blue lines show a linear fit performed in the range indicated by the solid line. The χ^2/ndof in the fit range is 5.515/6.000 = 0.919 for $\mu\tau_{\mathrm{h}}$ (2.187/3.000 = 0.730 for $e\tau_{\mathrm{h}}$) [148].



(a) (b)

Figure 7.3: Fake-factors used to derive estimates for multijet and W+jets backgrounds [171]. Factors are plotted as a function of the p_{T} of the τ_{h} candidate for the VBF and Boosted categories for 1-prong τ_{h} candidates (a) and 3-prong τ_{h} candidates (b).

production (r_W=0) or completely quark-initiated as dominant in W+jets ($r_W = 1$).

Correction of the $E_{\mathrm{T}}^{\mathrm{miss}}$ Direction

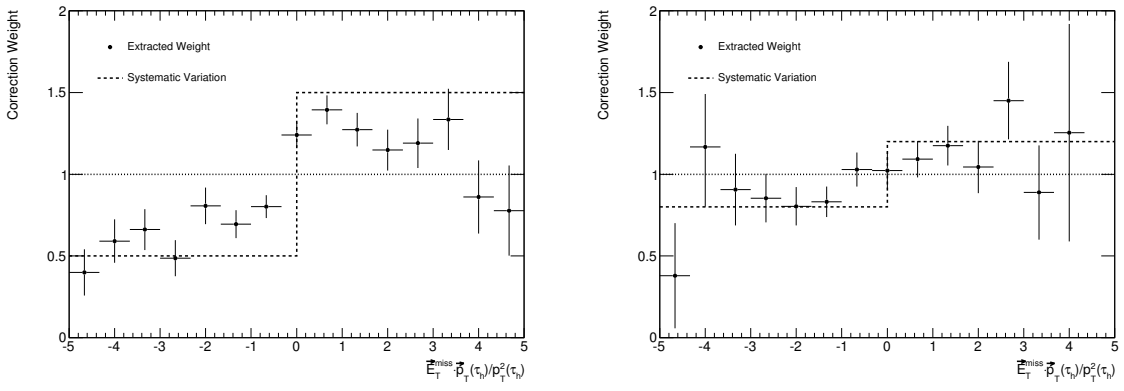
Due to the technical implementation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction used for this analysis [93], the measured contribution from an anti-τ_{h} candidate is not calibrated correctly, but its energy enters the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation with a calibration of a hadronic jet, while an identified $j \to \tau_{\mathrm{h}}$ candidate is assumed to be a real τ decay and thus obtains a smaller energy correction. Therefore, especially the direction of $E_{\mathrm{T}}^{\mathrm{miss}}$ is not modelled correctly when using anti-τ_{h} events, so that this fake-factor estimate requires an additional correction to compensate for this effect. The modelling is found to be improved significantly after re-weighting the events according to the dimensionless observable

$$\frac{\vec{E}_{\mathrm{T}}^{\mathrm{miss}}}{\vec{p}_T(\tau_{\mathrm{h}})} = \frac{\vec{E}_{\mathrm{T}}^{\mathrm{miss}} \cdot \hat{\vec{p}}_T(\tau_{\mathrm{h}})}{|\vec{p}_T(\tau_{\mathrm{h}})|} \,,$$

which is the projection of the missing transverse energy on the direction of the τ_{h} candidate normalised to its transverse momentum. The corresponding correction factors shown in Fig. 7.4 are obtained separately for the Boosted and the VBF category from their corresponding W+jets control regions. As an estimate of systematic uncertainties arising from this re-weighting, conservative variations are produced by either applying no such weights or a stronger weighting scheme. The effect of these systematic variations on the final BDT score distribution is shown in Fig. 7.21.
As an additional consequence of the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction, the distribution of the p_T^{tot} observable is strongly biased towards low values for identified τ_{h} candidates, which cannot be reproduced by the fake estimate with anti-τ_{h} candidates (cf. Fig. 7.5a). For this reason, the distribution is truncated with

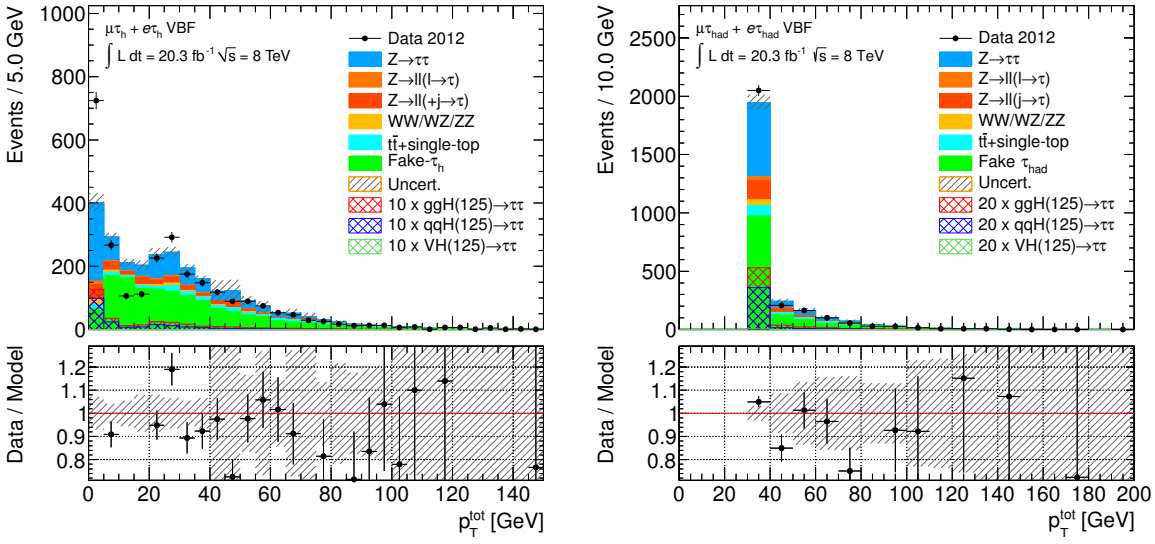
$$p_T^{\mathrm{tot}} \to \max\left(30 \text{ GeV}, p_T^{\mathrm{tot}}\right) \,.$$

This absorbs the mis-modelled region below 30 GeV as shown in Fig. 7.5b, so that it cannot be exploited by the BDT training.



(a) Boosted Category (b) VBF Category

Figure 7.4: Weights for the Fake Background estimate, as a function of the $E_{\mathrm{T}}^{\mathrm{miss}}$ projection on the τ_{h} direction for (a) the Boosted and (b) the VBF category.



(a) Mis-Modelled (b) Truncated at $p_T^{\text{tot}} > 30$ GeV

Figure 7.5: Distribution of the observable p_T^{tot} (a) mis-modelled due to the $E_{\text{T}}^{\text{miss}}$ calibration of anti-τ_{h} objects and (b) truncated at a threshold of $p_T^{\text{tot}} > 30$ GeV.

7.4.4 Correction Factors

For each category, additional correction factors f_p as for the cut-based analysis are extracted from dedicated control regions for contributions from W+jets, $Z \to \ell\ell(+j{\to}\tau_{\text{h}})$ and top-quark production. The extracted factors are summarised in Tab. 7.3, which correspond to the categories presented in Tab. 7.4.

	Rest	1-jet	Boosted	VBF
W+jets (OS)	0.80±0.01	0.97±0.01	(0.43±0.04)	(0.92±0.07)
W+jets (SS)	0.96±0.01	1.03±0.02	(0.39±0.04)	(0.95±0.09)
top (OS)	0.87±0.01	0.87±0.01	0.84±0.02	0.71±0.03
top (SS)	1.07±0.04	1.07±0.04	(0.81±0.05)	(0.77±0.11)
$Z \to \ell\ell$ (j$\to \tau_{\text{h}}$)	0.80±0.01	0.82±0.01	0.75±0.04	1.03±0.09
$Z \to \ell\ell$ (j$\to \tau_{\text{h}}$)	anti-τ_{h}:		0.91±0.01	1.16±0.03

Table 7.3: Normalisation factors used for each background and each category.

Corrections from top control regions are obtained by requiring ≥ 1 b-tagged jet and $m_T(\ell) > 50$ GeV for all categories except for Rest, which does not allow the presence of any (b-tagged) jet and is thus corrected with the factor from the 1-jet category. As for W+jets, these corrections are derived both for the opposite-sign and same-sign charge selection ($f_{\text{Top}}^{\text{OS}}$ and $f_{\text{top}}^{\text{SS}}$).

For the $Z \to \ell\ell(+j{\to}\tau_{\text{h}})$ process, the control region is defined by requiring two oppositely charged same-flavour leptons ($e^+e^-\tau_{\text{h}}$ or $\mu^+\mu^-\tau_{\text{h}}$) with 61 GeV $< m(\ell^+,\ell^-) <$ 121 GeV. Since no charge correlation of the τ_{h} candidate is expected in this $\ell^+\ell^-\tau_{\text{h}}$ selection, only one factor per category is necessary to

correct both the opposite-sign and the subtracted same-sign contributions. Additional correction factors $f^{\text{anti-}\tau_\text{h}}_{Z\to\ell\ell(+\text{jet}\to\tau_\text{h})}$ are obtained for the fake-factor estimate in the Boosted and the VBF selection from two similar control regions with inverted τ_h identification criteria. These factors are necessary to correct the contributions from $Z \to \ell\ell$, which must be subtracted from the anti-τ_h data before applying the fake-factor extrapolation to the signal region[6].

As compared to the previous analyses, especially the W+jets corrections differ significantly, which are derived for the new ALPGEN+PYTHIA Monte Carlo simulation here. This modified generator combination uses a different parton showering algorithm to generate the jets, which therefore have different $j\to\tau_\text{h}$ mis-identification probabilities. While the ALPGEN+HERWIG Monte Carlo in general predicted a higher $j\to\tau_\text{h}$ probability for the previous analyses than observed in data and thus needed smaller correction factors $f_W \approx 0.7$, the new modelling shows better agreement with the data, so that only weaker corrections $f_W \approx 0.9$ are necessary. While these are only required for the OS$-r_\text{QCD}\cdot$SS estimate of the Rest and 1-jet categories, similar numbers are obtained for the Boosted and VBF categories for comparison and in order to maintain an alternative background estimate to the fake-factor method[7]. Same-sign factors f^SS_top for the Boosted and VBF categories are also kept for the same reason.

Background Summary

As in the cut-based analysis, the multijet background in the Rest and 1-jet categories is estimated from the same-sign data in combination with the OS$-r_\text{QCD}\cdot$SS contributions from the other backgrounds, which are obtained from simulation and normalised to the data in the dedicated control region via correction factors. For the simulated W+jets events, a re-weighting procedure as described in Sec. C.1.1 is applied. This corrects angular correlations of the selected lepton-jet($\to\tau_\text{h}$) pairs, which have been found to be badly modelled in the new ALPGEN+PYTHIA samples as compared to the data.

In the Boosted and VBF category, the combined contribution of events with mis-identified τ_h objects from multijet and W+jets is estimated with the fake-factor method. Also for this estimate, residual contributions from simulated backgrounds are normalised to data in dedicated anti-τ_h control regions and subtracted from the observed data in these regions.

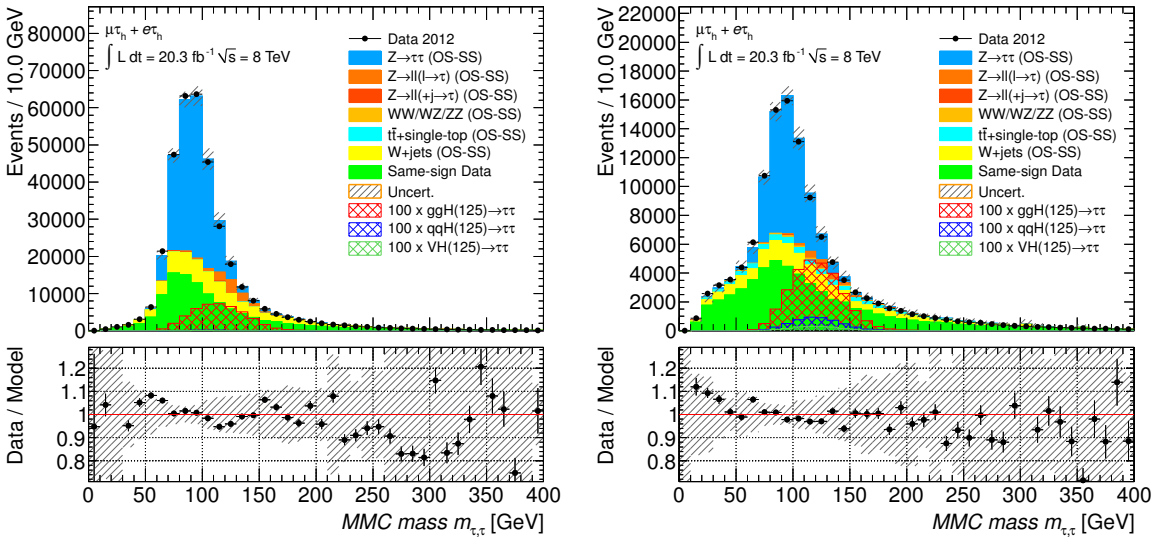
The $Z \to \tau\tau$ background for all categories is again obtained from the τ-embedded $Z \to \mu\mu$ data, which are normalised to data in four control regions ($40 < m_\text{vis} < 70$ GeV) corresponding to the four different trigger selections[8].

The estimated numbers of signal and background events for the most sensitive VBF and Boosted categories are summarised in Tab. 7.5. The corresponding numbers of opposite-sign and same-sign yields are included in Tab. C.1 and Tab. C.2 for the Rest and 1-jet category. The corresponding MMC mass distributions are shown in Fig. 7.6 for all four categories.

[6] Although the corresponding $j\to\tau_\text{h}$ candidates are similar to the contribution from W+jets, the $Z \to \ell\ell(+j\to\tau_\text{h})$ background is not included during the extraction of fake-factors and must thus be subtracted according to the prediction of the (corrected) simulation. An important improvement of the method to account for this fake background is foreseen for the future.

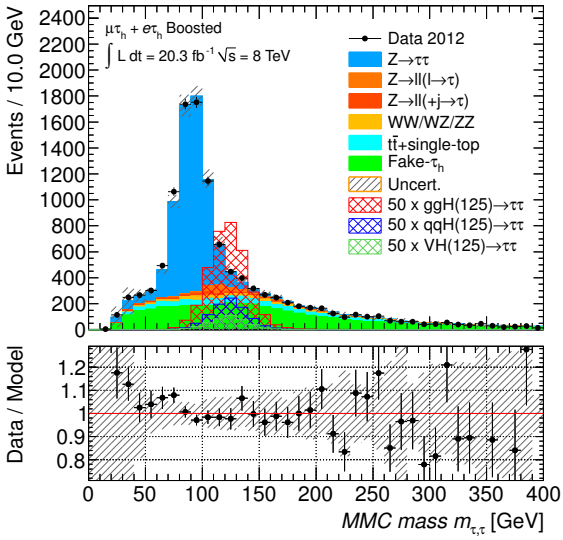
[7] The W+jets corrections are applied in the BDT training, which is based on a variation of the OS-SS estimate.

[8] The overall normalisation is later obtained from the final fit (cf. Sec. 7.9.1), so that these control regions mainly define the relative contributions from the different triggers.

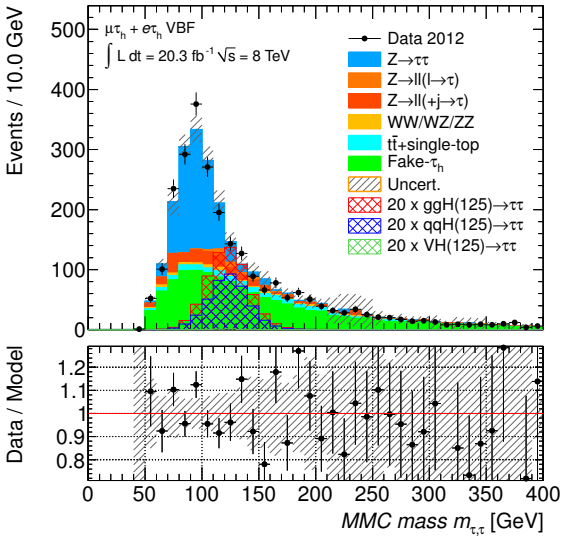


(a) Rest Category

(b) 1-jet Category



(c) Boosted Category



(d) VBF Category

Figure 7.6: Distributions of the MMC mass for the different categories. The background components are estimated with the techniques described in Sec. 7.4.

Region	VBF	Boosted	1-jet	Rest
Signal Region (SR)	≥ 2 jets ($p_{\mathrm{T}}(j) > 50/30$ GeV) $\Delta\eta_{j_1 j_2} > 3.0$ Single Lepton Trigger ▹ b-jet veto ▹ $m_{\mathrm{vis}} > 40$ GeV	Not VBF $p_{\mathrm{T}}(H) > 100$ GeV Single Lepton Trigger ▹ b-jet veto	Not VBF, Boost ≥ 1 jet ($p_{\mathrm{T}}(j) > 30$ GeV)	Not Boost 0 jets ($p_{\mathrm{T}}(j) > 30$ GeV)
$Z \to \tau\tau$	As SR, but $m_T(\ell) < 40$ GeV and $m_{\mathrm{MMC}}(\tau,\tau) < 110$ GeV	As SR, but $m_T(\ell) < 40$ GeV and $m_{\mathrm{MMC}}(\tau,\tau) < 110$ GeV	As SR, but $m_T(\ell) < 40$ GeV and $m_{\mathrm{MMC}}(\tau,\tau) < 110$ GeV	As SR, but $m_T(\ell) < 40$ GeV and $m_{\mathrm{MMC}}(\tau,\tau) < 110$ GeV
W+jets OS & SS	As SR, but $m_T > 70$ GeV	As SR, but $m_T > 70$ GeV	As SR, but $m_T > 70$ GeV	As for SR, but $m_T > 70$ GeV
Top	As SR, but ≥ 1 b-tagged jet and $m_T > 50$ GeV	As for SR, but ≥ 1 b-tagged jet and $m_T > 50$ GeV	≥ 1 b-tagged jet ($p_{\mathrm{T}} > 30$ GeV) and $m_T > 50$ GeV (OS & SS)	
$Z \to \ell\ell(j{\to}\tau_{\mathrm{h}})$	As SR, but 2 OS same-flavor leptons* and $61 < m_{ll} < 121$ GeV	As for SR, but 2 OS same-flavor leptons* and $61 < m_{ll} < 121$ GeV	2 OS same-flavor leptons* and $61 < m(\ell,\ell) < 121$ GeV	

*: Standard kinematic and identification requirements, but isolation requirements only applied on the leading lepton.

Table 7.4: Summary of signal and control regions used in the multi-variate analysis. All regions are defined on top of the preselection requirements described in Sec. 7.2. Signal region requirements are defined to cover the full preselected phase space, so that events failing one signal region selection are recovered in a sub-sequent category. Events failing additional selections marked with • are excluded and not recovered in any other signal region.

Process/Category	VBF	Boosted
$ZH(H \to \tau\tau)$ (m_H = 125 GeV)	$0.11 \pm 0.01^{+0.03}_{-0.03}$	$2.9 \pm 0.1^{+0.3}_{-0.3}$
$WH(H \to \tau\tau)$ (m_H = 125 GeV)	$0.22 \pm 0.02^{+0.05}_{-0.06}$	$5.8 \pm 0.1^{+0.6}_{-0.6}$
$VBF(H \to \tau\tau)$ (m_H = 125 GeV)	$19.9 \pm 0.2^{+2.3}_{-2.3}$	$10.3 \pm 0.1^{+0.8}_{-0.8}$
$gg \to H \to \tau\tau$ (m_H = 125 GeV)	$11.6 \pm 0.4^{+5.4}_{-5.1}$	$49.8 \pm 0.9^{+21.4}_{-19.2}$
$Z \to \tau\tau$	$818 \pm 16^{+49}_{-45}$	$5690 \pm 54^{+265}_{-240}$
$Z \to \ell\ell(\ell \to \tau_{\mathrm{h}})$	$40 \pm 3^{+24}_{-8}$	$233 \pm 13^{+51}_{-47}$
$Z \to \ell\ell(+j \to \tau_{\mathrm{h}})$	$229 \pm 16^{+59}_{-55}$	$281 \pm 20^{+39}_{-40}$
Di-boson	$65 \pm 4^{+10}_{-13}$	$436 \pm 12^{+40}_{-41}$
Top	$163 \pm 8^{+19}_{-19}$	$986 \pm 20^{+116}_{-110}$
Fake-Factor Estimate	$1326 \pm 10^{+519}_{-263}$	$4280 \pm 18^{+1853}_{-1065}$
Total Background	$2642 \pm 27^{+525}_{-274}$	$11906 \pm 66^{+1877}_{-1100}$
Data	2676	12051

Table 7.5: Predicted event yields for the VBF and Boosted categories. The quoted uncertainties correspond to the statistical and systematic uncertainties respectively (before the global fit). The backgrounds are estimated as described in Sec. 7.4.

7.5 BDT Training Setup

For this analysis, BDTs are trained and optimised separately for each of the categories defined in Sec. 7.3. After outlining the general method of the generation of trees with event weight boosting in Sec. 7.1, the detailed training setup of this analysis is described in the following.

7.5.1 Signal and Background Composition

For the generation of BDTs, sets of training and testing events labelled as signal or background are provided to TMVA after the category selection. In order to gain the best training performance, the separate contributions from different signal and background processes must be as close as possible to the expectation in data[9]. For the signal processes, the events are thus weighted according to their cross section. For the background processes, however, this is not sufficient, since the estimate of the multijet contribution requires a side-band subtraction technique. Although such a subtraction can be implemented by providing events with negative weights, their treatment in the training is not well defined[10]. Instead, the OS$-r_{\mathrm{QCD}}\cdot$SS estimate is implemented in the training by providing the same-sign data weighted with r_{QCD}, to which only simulated opposite-sign events are added. In doing so, the weights of all events from process p are scaled with the factor

$$w_p = \frac{f_p^{\mathrm{OS}} N_p^{\mathrm{OS}} - r_{\mathrm{QCD}} \cdot f_p^{\mathrm{SS}} N_p^{\mathrm{SS}}}{N_p^{\mathrm{OS}}}$$

to the corrected OS$-r_{\mathrm{QCD}}\cdot$SS yield of the corresponding background. Due to more difficulties with subtractions in the fake-factor estimate, this method is applied also for the Boosted and VBF category, for which the background estimation technique is thus different between training and the final analysis.

While in general the standard samples for all background processes are used for training, the $Z \to \ell\ell(j{\to}\tau_{\mathrm{h}})$ process is left out, since its contribution is expected to be charge symmetric and the OS$-r_{\mathrm{QCD}}\cdot$SS yield to be negligible. Instead of training on embedded $Z \to \tau\tau$ data, the Monte Carlo samples are used in the training, which provide a larger amount of training events than available from $Z \to \mu\mu$ data. At the same time, training on simulated $Z \to \tau\tau$ is a more conservative choice, since the signal samples naturally have to be obtained from simulation as well. The training can thus focus on physics differences between the simulated signal and background processes and cannot exploit a possible mis-modelling of the simulated signal, which would distinguish it by other means from the data-driven $Z \to \tau\tau$ estimate obtained with the embedding technique.

For the signal, only the simulated $H \to \tau\tau$ templates with a mass of $m_H = 125$ GeV are included in the training. In order to focus on the dominant production mechanisms of the Higgs boson in each of the categories, only the ggH signal is used for the training of the BDT in the Boosted category, and only the qqH signal is included in the BDT training for the VBF category.

7.5.2 n-Fold Cross-Evaluation

For the optimisation of the observable selection and learning rate, the training is performed on a randomly selected part of the provided events. The performance of the trained BDT is then immediately

[9] During the training, the sum of all background contributions is automatically normalised to the sum of the signal.

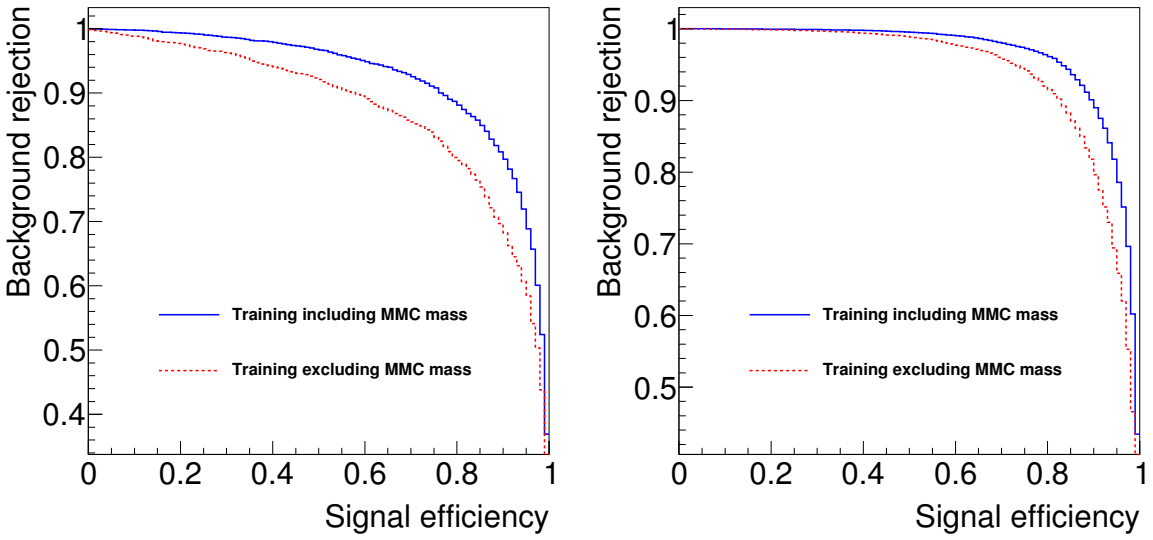
[10] It is debatable whether a background event with a negative weight should be able to "reduce" the background in a signal-like node. Since this essential difference of otherwise identical events is expected to degrade the performance due to overtraining effects, negative event weights are avoided in this analysis, although tests have not shown significant performance differences.

evaluated on the other statistically independent part of the samples to avoid an optimisation on overtraining effects. Events are therefore split based on their event numbers (even for training, odd for testing). This procedure inevitably reduces the number of available events both for the training and for the testing in the final analysis. For the training, however, large samples are desired to avoid overtraining effects and to obtain a strong separation between signal and background. On the other hand, the statistical uncertainties of the testing samples result in significant contributions to the systematic uncertainties of the analysis, which can be reduced by sufficiently large testing samples. In order to exploit all available events both within the training and testing procedure, a so-called cross-evaluation technique is applied, for which n BDTs with identical setup are trained after splitting each sample into n statistically independent subsets. In each of the different trainings, $n-1$ of these subsets are used for the training, so that the fraction $(n-1)/n$ of the training events is maximised for large n. Although each BDT can be used to test the remaining n-th subset with a fraction of $1/n$ of the available events, the final analysis can combine the testing samples of all n BDTs, so that the full sample statistics enter the final signal and background prediction. Due to the separation of the n sub-samples, each event can always be tested with a BDT, which did not include it during training. Hence, the n different BDTs with an identical training setup can provide the required separation of testing and training sub-sample with a maximal exploitation of the available number of events both for training and testing.

In this analysis, a two-fold ($n = 2$) cross-evaluation is performed, for which two BDTs are trained per category: the first on events with even numbers, the second on events with odd numbers. This mainly doubles the number of available testing events for the final analysis, while the training performance is not improved, since still only 50% of the available sample is used for each training. For the signal samples, however, the statistical uncertainty of the final prediction has been found to be sufficiently small, so that a reduction of testing events does not significantly affect the systematic uncertainty of the signal prediction. In order to improve the training performance, the signal training samples have thus been increased to 75%, whereas only 25% of the available events can be tested per BDT. With this setup, only 50% of the available signal events are used for the final prediction of the signal expectation, while each BDT is trained on a correspondingly larger signal sample.

7.5.3 Variable Selection and Parameter Optimisation

The performance of a BDT in terms of signal extraction depends strongly on the set of input observables, which are used to split the nodes during the tree generation. In general, the training algorithms will automatically identify observables with strong separation power between signal and background events in order to split the nodes during the tree generation, whereas weaker observables are neglected in the process. The performance should therefore always increase with a larger number of input observables. The separation between signal and background is only expected to saturate if the observables are strongly correlated, but it should never degrade by extending the list of inputs. In practice, however, the number of event observables provided for the BDT training should be limited, since the background and signal modelling must be controlled well for each observable and all correlations in order to guarantee a similarly well modelled BDT score to extract an expected signal from the data. In addition to these technical limitations on the required validation effort, instabilities in the training performance can occur, which can degrade the separation e.g. due to overtraining effects, although an additional observable is added to the list of inputs. Similarly, the exclusion of an observable can increase the performance, so that a small set of powerful observables is the desired basis for any BDT-based analysis.



(a) Boosted Category (b) VBF Category

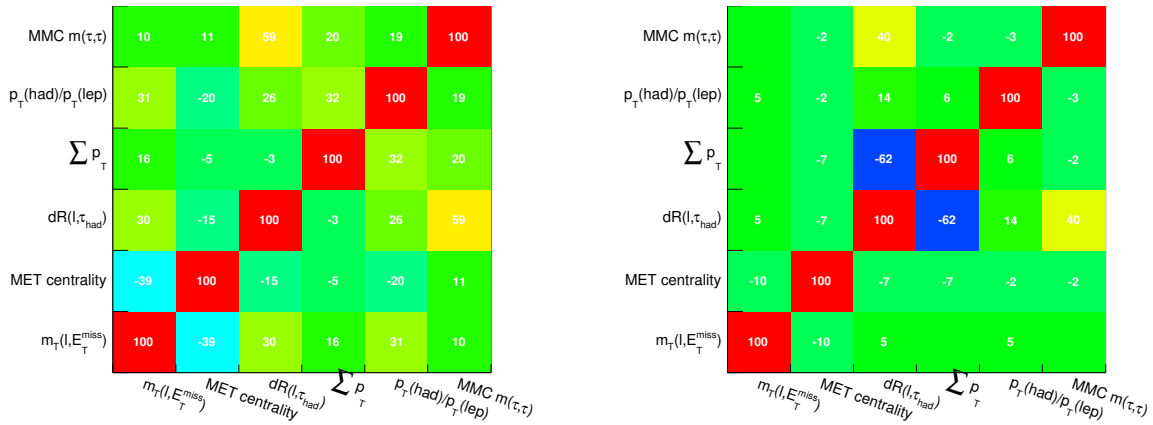
Figure 7.7: Receiver-operating-characteristic (ROC) of the final BDTs trained in the (a) Boosted and (b) VBF category, which include the MMC mass as an input observable. For comparison, the ROC is shown for a similar training setup, which excludes the MMC mass from the input observables.

Automated Ranking of Observables

Starting from a large set of observables, the impact of each on the separation power of the BDT is evaluated, so that the weakly discriminating observables can be dropped from the set of inputs. The usage of the observables to split nodes during the generation of each tree is automatically registered, which eventually provides a *ranking* of observables from the training procedure. This ranking, however, cannot directly reflect the impact of a specific observable on the BDT performance, since the presence of other observables can have unwanted effects: if A and B are very similar observables, and any cut on either A or B exploits the same physics difference between signal and background, A and B might be used alternately to generate the trees and might thus both be ranked low. If instead only A is used as an input observable, and B is not available for the training, the ranking of A can improve significantly.

In order to evaluate the impact of a specific input observable on a BDT, a different BDT must be trained without this specific input, and its effect on the performance can be estimated by comparing the separation power between the two BDTs. Weak observables can then be identified based on the criterion that their impact on the training performance is small. Starting from a basic set of N observables, which have been found to rank high in different training setups, each of the observables is dropped separately from the training setup, and N BDTs are trained with the remaining $N-1$ observables. Their performance is then estimated from the integral of their *receiver-operating-characteristic* (ROC), which describes their background rejection as a function of signal acceptance. Examples for these curves are shown in Fig. 7.7, which are obtained from a scan of cut-values on the BDT score. The integral of such a curve is an indicator of the performance in terms of background rejection. An observable is thus considered to be obsolete if its removal from the list of inputs does not significantly reduce the ROC-integral of the BDT.

Also for this method of observable ranking, strong correlations between weak observables can influ-



(a) Background Boosted

(b) Signal Boosted

Figure 7.8: Linear correlation coefficients of the input observables for the training of the BDT in the Boosted category estimated for (a) background and (b) signal contributions.

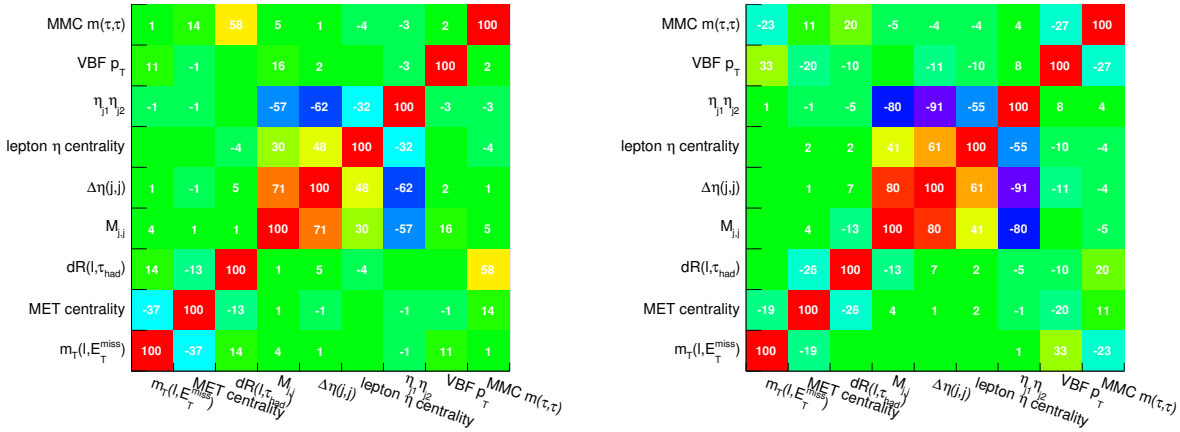
ence the outcome of this test. In a situation of two strongly correlated observables A and B, either A or B can be dropped from the list of inputs without significant performance degradation of the BDT. Dropping both at the same time, however, can lead to a significant loss of performance. Therefore, the tests are also repeated after dropping pairs of observables in order exclude such effects.

Final Set of Observables

With this method identifying weakly discriminating observables, final sets of remaining training inputs with high separation power are obtained for each category and are summarised in Tab. 7.6. The linear correlation coefficients of these observables (cf. Figs. 7.8-7.9) show significant differences between correlations in background and signal processes, which are eventually exploited by the BDTs. In the Boosted category, the largest of these differences is found for the correlation between $\Delta R(\ell, \tau_\mathrm{h})$ and $\sum |p_T|$, which is -62% for the signal and almost absent for the sum of backgrounds. The di-jet observables in the VBF category show many small such differences, while especially the correlation of $\Delta R(\ell, \tau_\mathrm{h})$ with the MMC mass is stronger for the backgrounds (58%) as compared to the signal (20%).

From an additional scan of the training performance in dependence of training parameters (i.e. the number of trees N_trees, their maximum depth $N_\mathrm{nodes}^\mathrm{max}$ and the `Shrinkage`)[11], the values providing optimal performance with the selected observables are identified and are summarised in Tab. 7.7. These optimised parameters are then used to train the final analysis BDTs on the selected sets of input observables.

[11] Each node was required to contain $N_\mathrm{events}^\mathrm{min} \geq 150$ events in order to ensure the statistical significance of their separation power. The number of cuts $N_\mathrm{cuts} = 20$ was kept constant, since no effect was observed from variations within $N_\mathrm{cuts} \in [20, 50]$, while large numbers resulted in very large computing times for the BDT training.



(a) Background VBF

(b) Signal VBF

Figure 7.9: Linear correlation coefficients of the input observables for the training of the BDT in the VBF category estimated for (a) background and (b) signal contributions.

Variable	Rest	1-jet	Boosted	VBF
$\Delta R_{\tau\ell}$	•	•	•	•
$m_T(\ell)$	•	•	•	•
$C^{\phi}_{\ell,\tau_\mathrm{h}}(E_\mathrm{T}^\mathrm{miss})$	•	•	•	•
$m_\mathrm{MMC}(\tau,\tau)$	•	•	•	•
$\sum \lvert p_T \rvert$	•	•	•	
$p_T(\tau_\mathrm{h})/p_T(\ell)$	•	•	•	
$m(j_1, j_2)$				•
$\eta(j_1) \times \eta(j_2)$				•
$\lvert\eta(j_2) - \eta(j_1)\rvert$				•
$C^{\eta}_{j_1,j_2}(\ell)$				•
p_T^tot				•

Table 7.6: Discriminating observables used for the training of the BDTs. The filled circles identify which observables are used in each category.

Parameter	Rest	1-jet	Boosted	VBF
N_trees	800	800	600	400
$N^\mathrm{max}_\mathrm{nodes}$	4	4	4	4
`Shrinkage`	1.0	1.0	0.3	0.2

Table 7.7: Optimal training parameters used for the analysis BDTs.

7.6 Control Distributions

With the defined set of observables and parameters to train the different BDTs per category, the modelling of all observables and their correlations must be validated. Therefore, comparisons of the background model against the observed data are performed for all relevant input observables in each category. Figures 7.10-7.11 show the comparisons for all input observables of the BDT trained in the Boosted category. Good agreement between model and data is observed within the uncertainties of the background model. These include all systematic uncertainties discussed in Sec. 7.8 except for the very conservative estimate of the fake-factor method, which is later constrained strongly by the final fit. Similar comparisons are included in Fig. C.2 and Fig. C.3 for the Rest and 1-jet categories, which provide additional confidence especially for the $Z \to \tau\tau$ background modelling in the full accessible $\tau\tau$ phase space. These two categories, however, are not considered any further from now on, since they are excluded from the final fit[12].

Although these one-dimensional comparisons validate the background model independently for each observable, their correlations are of equal importance and are thus investigated additionally. As the number of higher order correlations is too large to validate, these checks are restricted to all two-dimensional correlations, which are provided here in two different versions (cf. Fig. 7.12). Starting from two-dimensional correlation histograms, Fig. 7.13 allows a comparison between data and background model in terms of mean values. Each sub-figure shows the mean and the root mean square (RMS) values of observable Y in bins of observable X. This effectively reduces the second dimension of each distribution to the two characteristic values and thus allows direct comparisons between data and the background model for the correlations between all pairs of observables. Similar to the one-dimensional validation, good agreement within uncertainties is observed between data and model. As an additional bin-wise check, Fig. 7.14 contains the two-dimensional distribution of the significance obtained by relating the difference between data d_{ij} and background model b_{ij} in bin (X_i, Y_j) to the corresponding statistical uncertainties $\sqrt{d_i j}$ and σ_{ij} :

$$\mathrm{sig}_{ij} = \frac{d_{ij} - b_{ij}}{\sqrt{d_{ij} + \sigma_{ij}^2}}$$

Also in this version of the validation, the modelled correlations of the backgrounds are found to be consistent with the observation in the data. The number of bins with differences larger than 2σ is consistent with the assumption of statistical fluctuations in the data, and no mis-modelling is visible within uncertainties.

The same two-dimensional distributions as described for the VBF category are provided for the Boosted category in Figs. 7.15-7.17 and demonstrate a similar level of agreement between the background model and the observed data.

[12] This decision has been made in order to simplify the fit model, while including them in the final fit would not significantly improve the total signal sensitivity of the analysis.

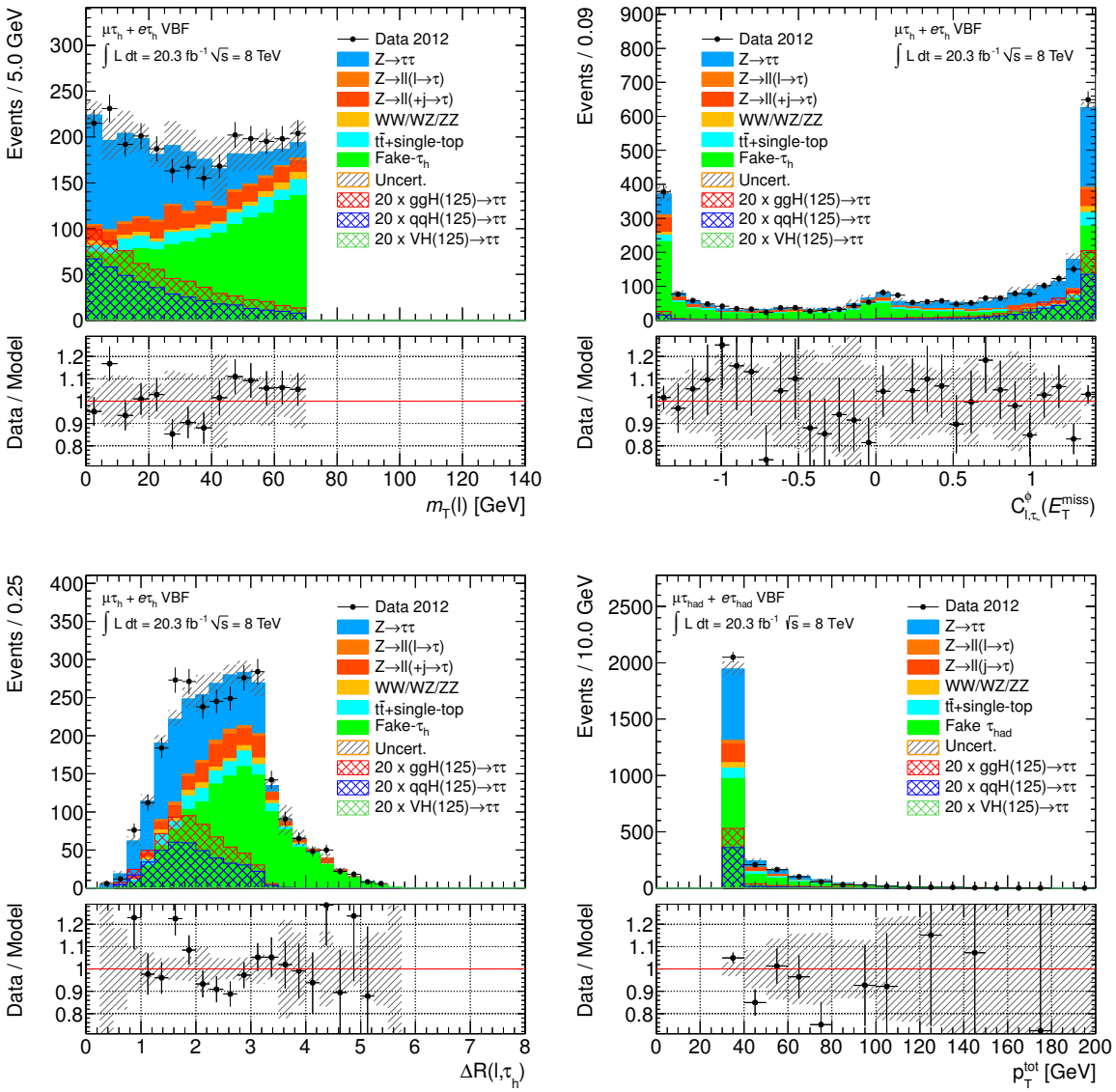


Figure 7.10: Distributions of the $\tau\tau$-related input observables for the BDT training in the VBF category. The background components are estimated with the techniques described in Sec. 7.4. The total uncertainty on the background model excludes the fake-factor systematic and is obtained before the global fit by adding all systematic uncertainties in quadrature to the statistical uncertainty per bin.

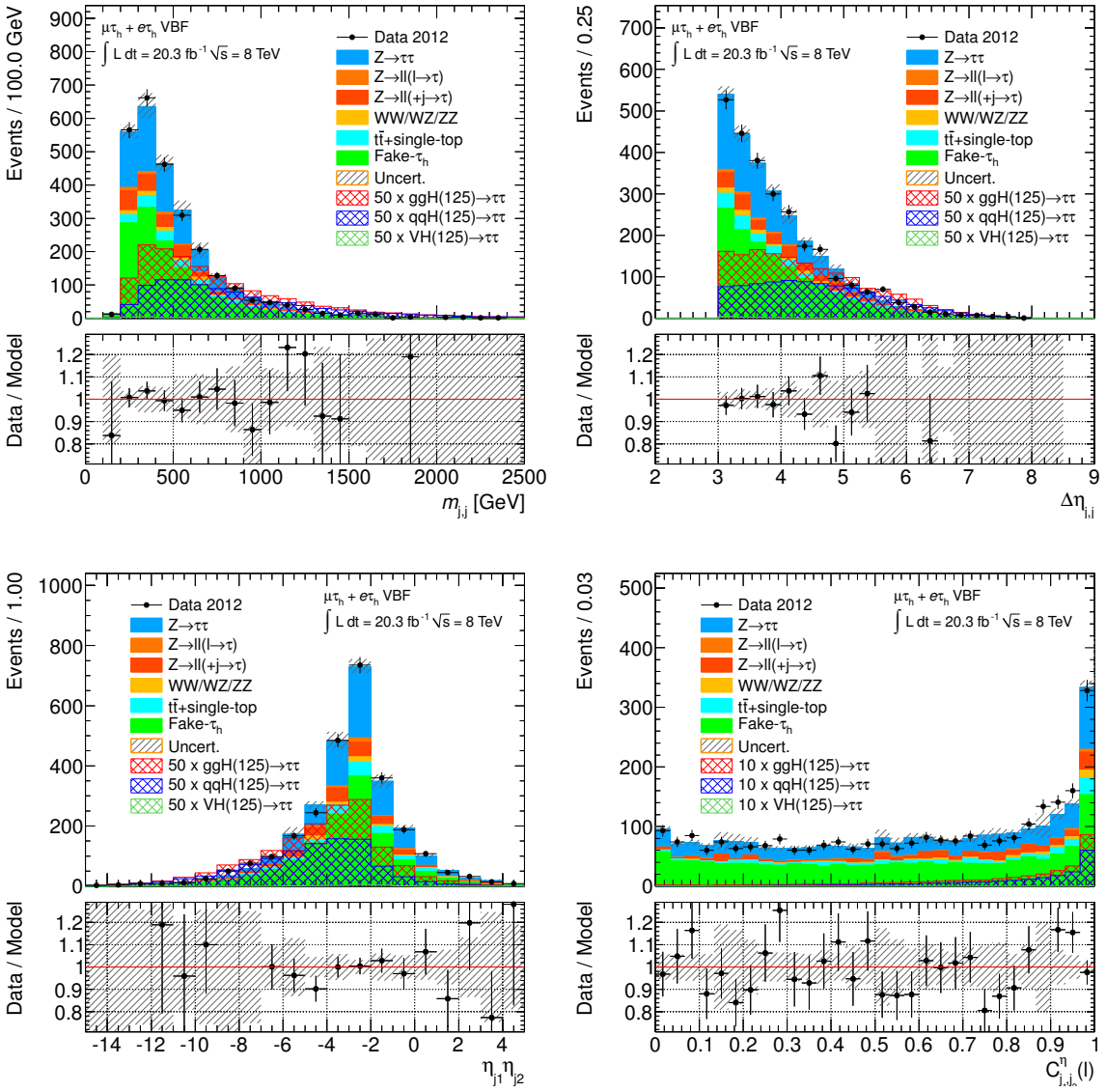


Figure 7.11: Distributions of the jet-related input observables for the BDT training in the VBF category. The background components are estimated with the techniques described in Sec. 7.4. The total uncertainty on the background model excludes the fake-factor systematic and is obtained before the global fit by adding all systematic uncertainties in quadrature to the statistical uncertainty per bin.

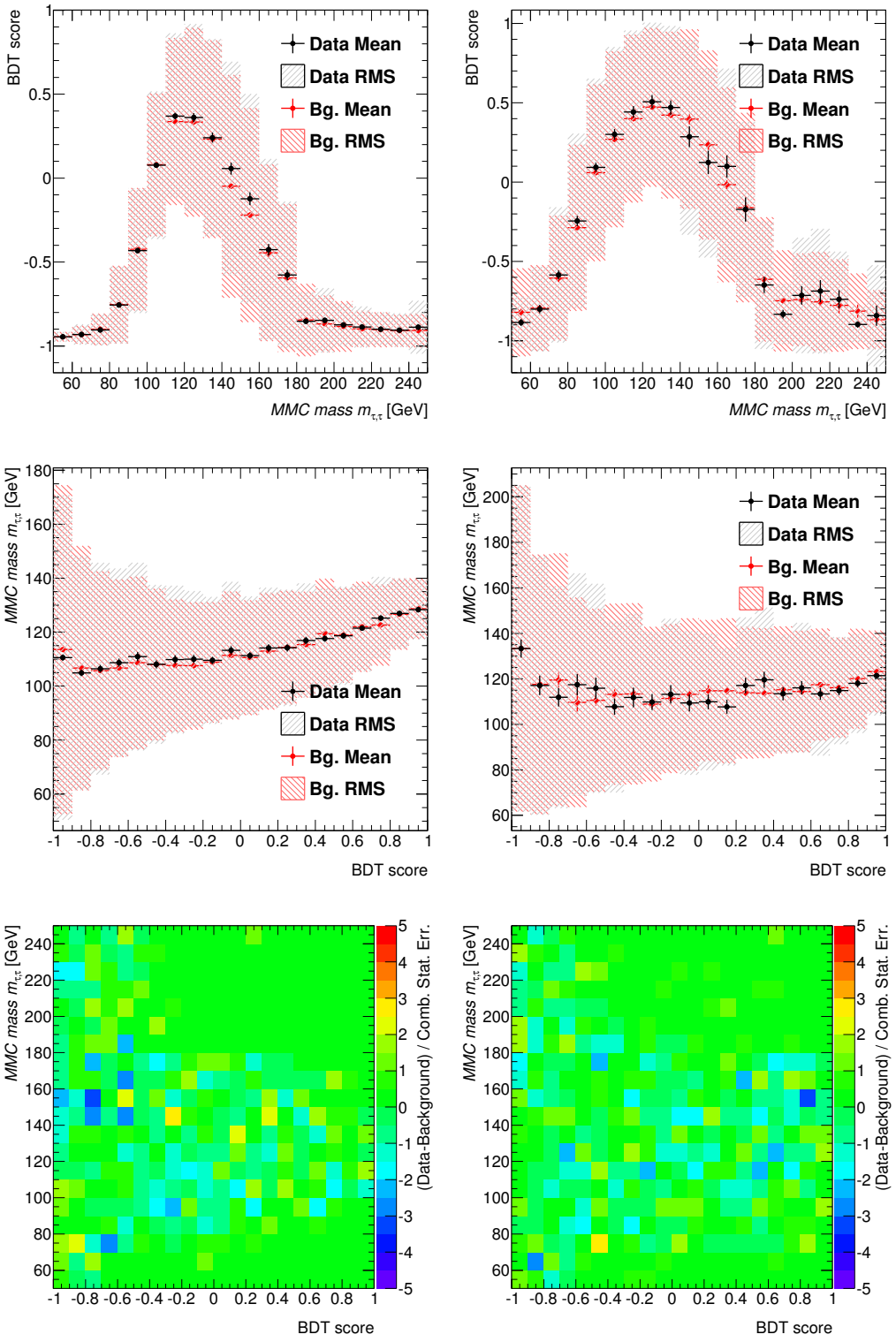


Figure 7.12: Correlations of the BDT score and the MMC mass compared between data and background in the Boosted (left) and VBF category (right): the mean value ⟨BDT score⟩ as a function of the MMC mass (top), the mean value $\langle m_{\mathrm{MMC}} \rangle$ as a function of the BDT score (centre) and the deviation between data and background normalised to the total statistical uncertainty (bottom).

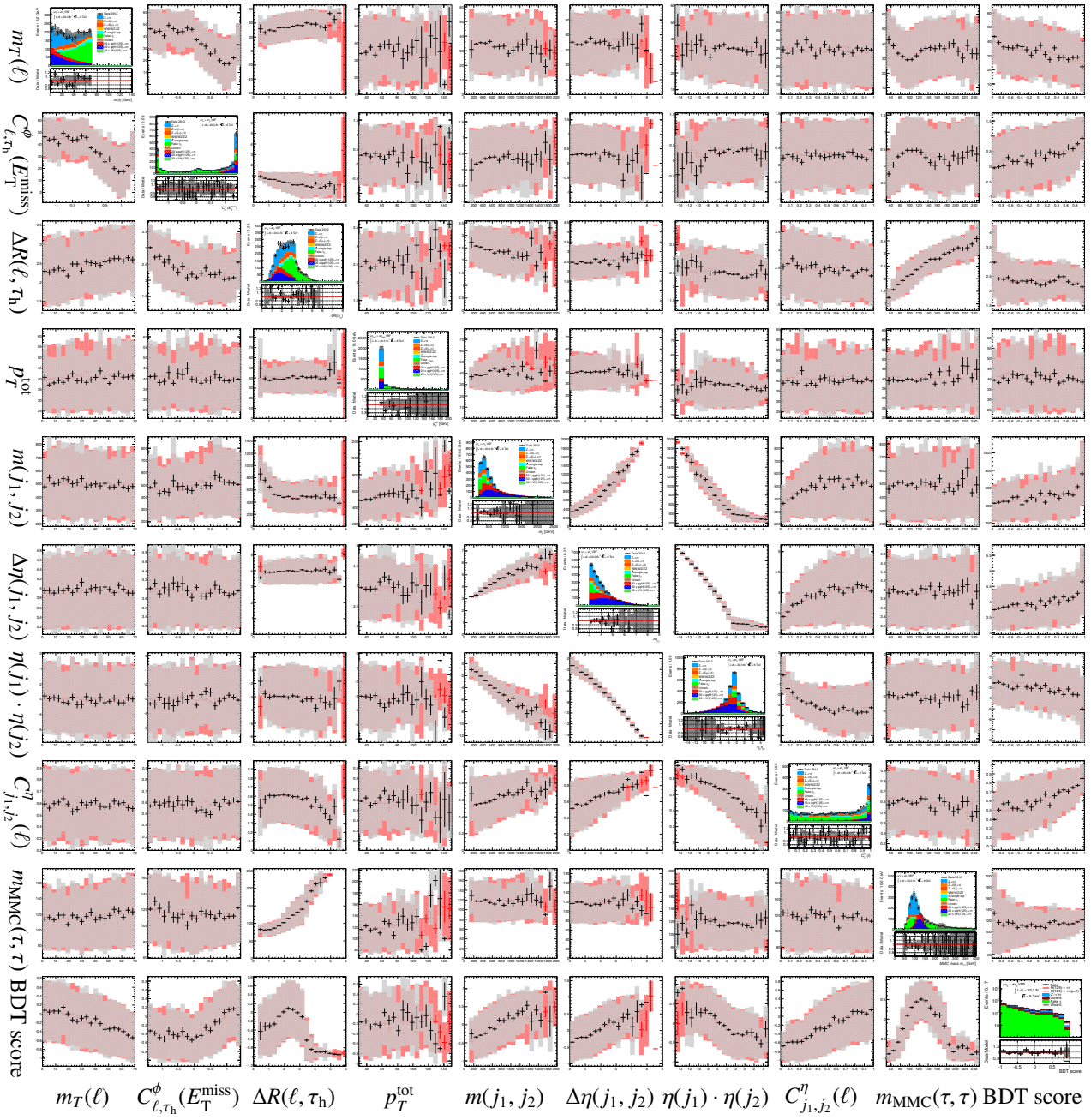


Figure 7.13: Correlations of the input observables for the BDT training in the VBF category compared between data (black) and background (red) in terms of mean values as described in the text. The total background is estimated with the techniques described in Sec. 7.4.

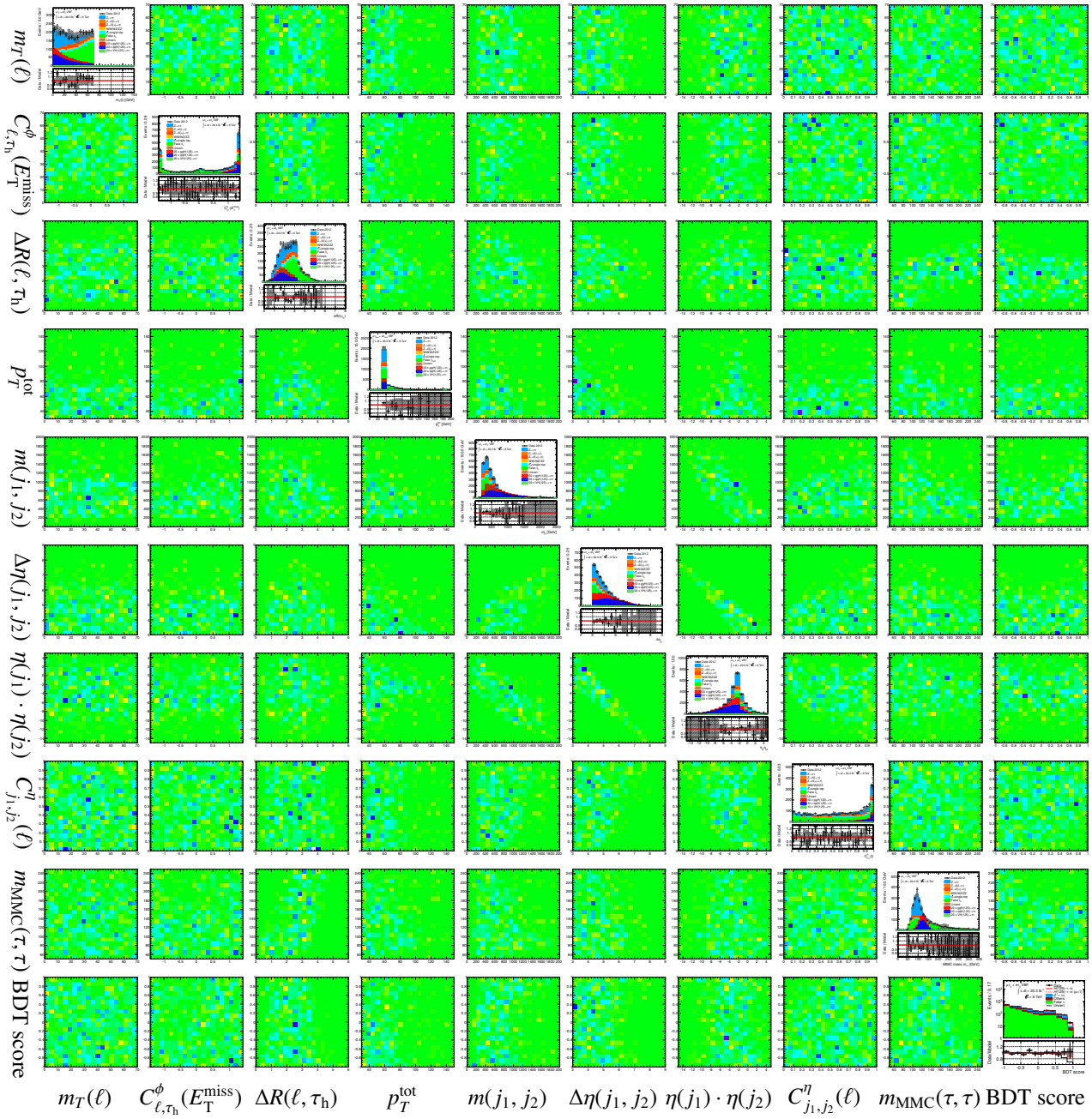


Figure 7.14: Correlations of the input observables for the BDT training in the VBF category in terms of deviations between data and background. The total background is estimated with the techniques described in Sec. 7.4. Only statistical uncertainties are considered for this comparison.

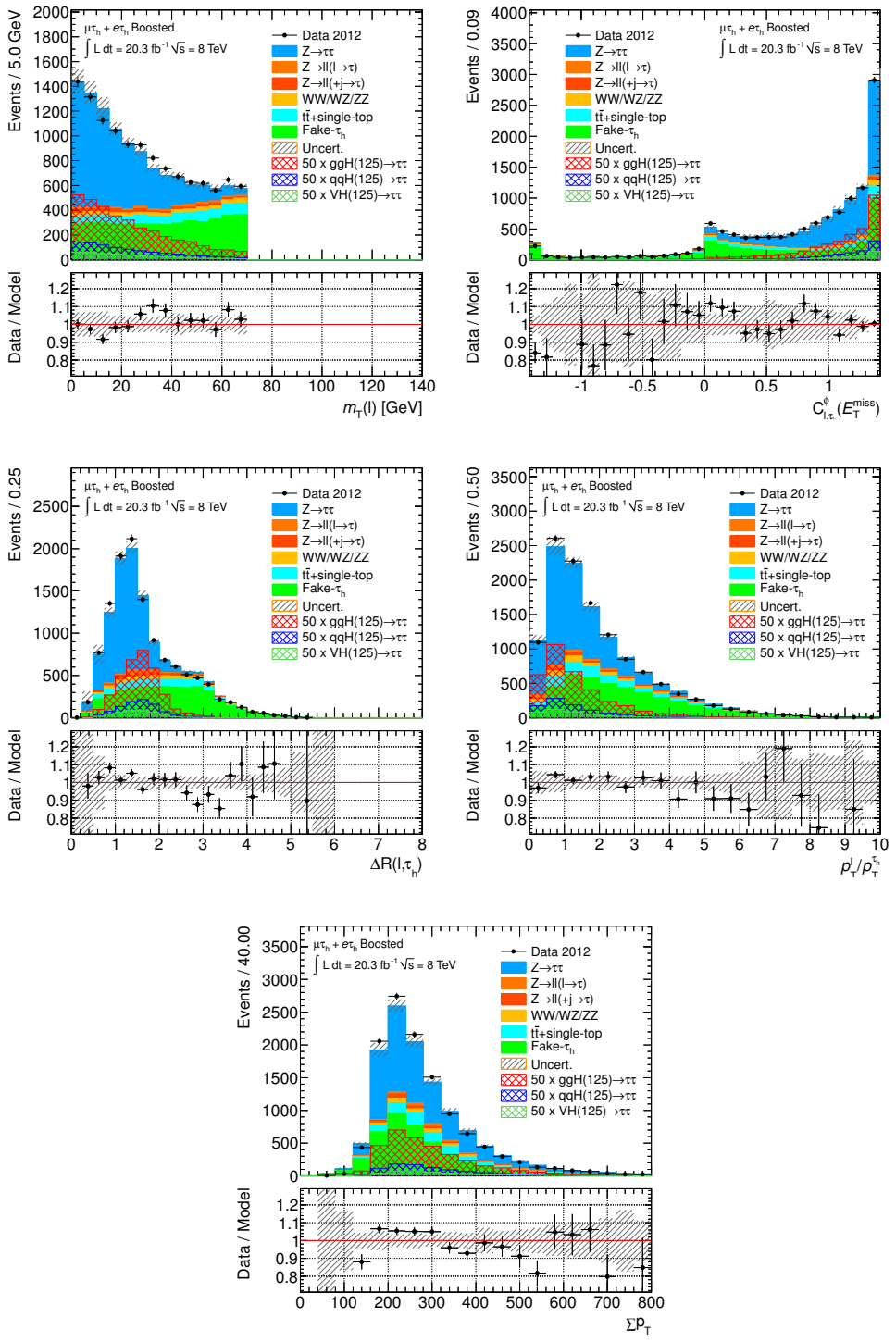


Figure 7.15: Distributions of the input observables for the BDT training in the Boosted category. The background components are estimated with the techniques described in Sec. 7.4. The total uncertainty on the background model excludes the fake-factor systematic and is obtained before the global fit by adding all systematic uncertainties in quadrature to the statistical uncertainty per bin.

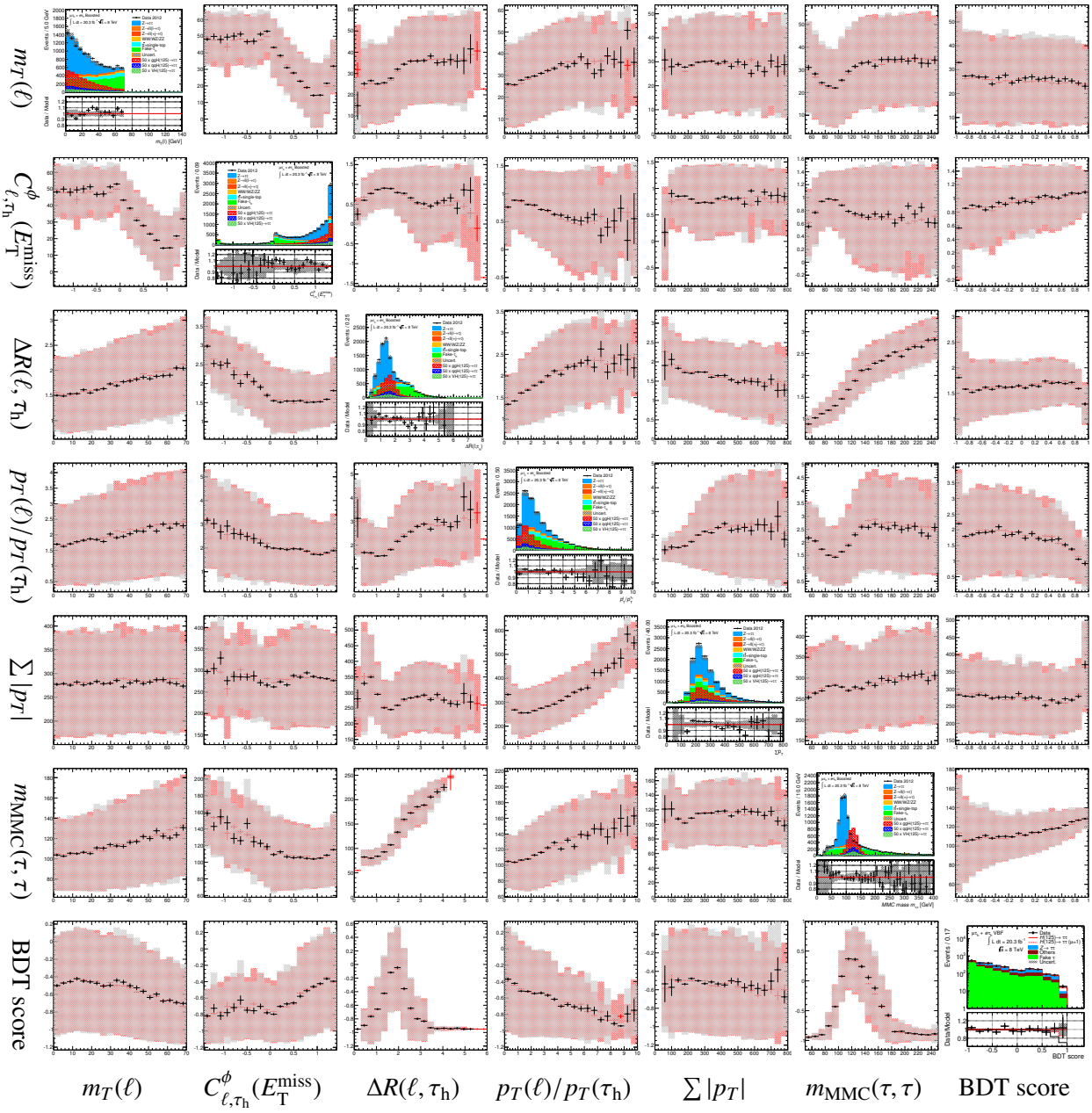


Figure 7.16: Correlations of the input observables for the BDT training in the Boosted category compared between data (black) and background (red) in terms of mean values as described in the text. The total background is estimated with the techniques described in Sec. 7.4.

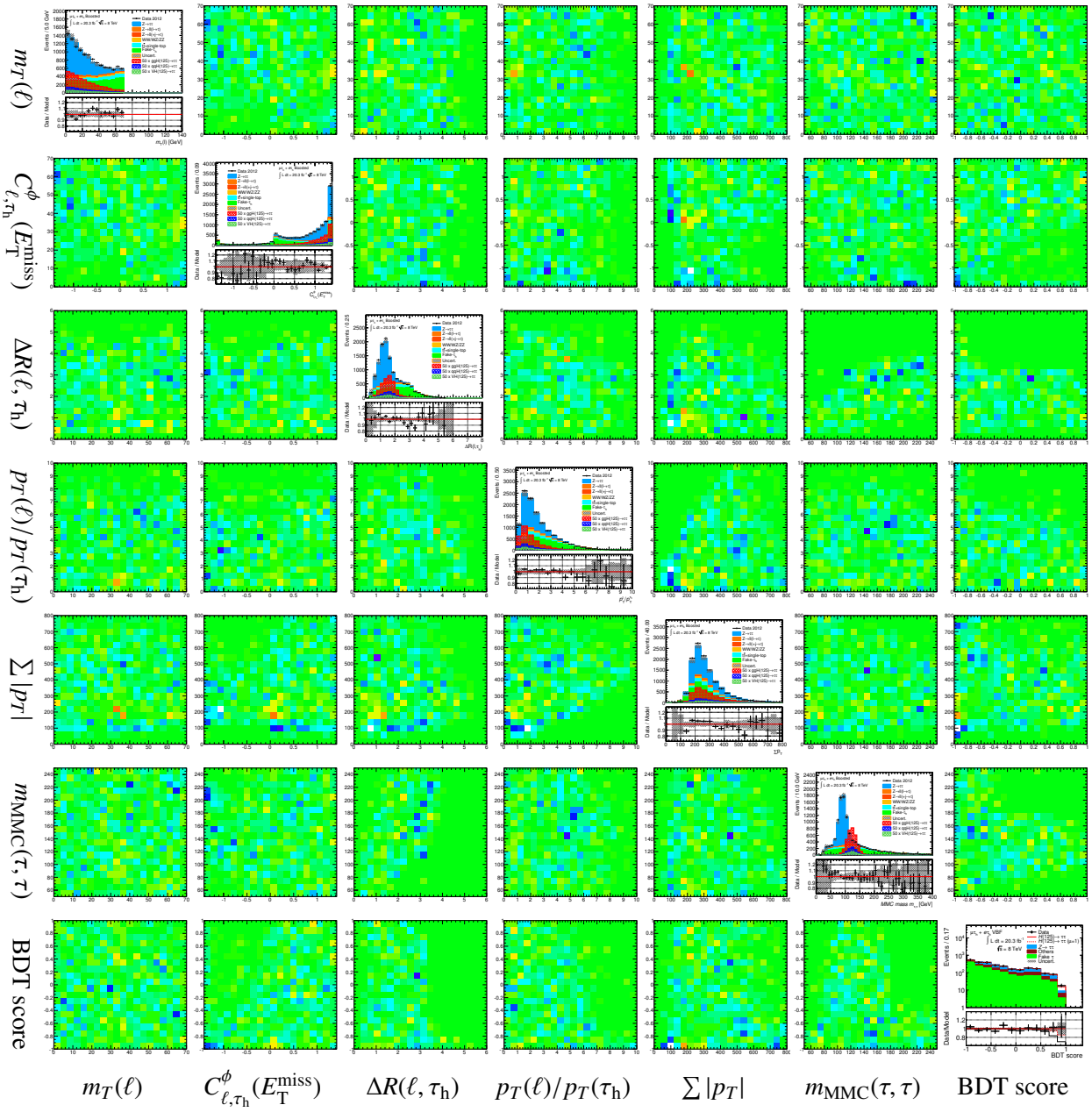


Figure 7.17: Correlations of the input observables for the BDT training in the Boosted category in terms of deviations between data and background. The total background is estimated with the techniques described in Sec. 7.4. Only statistical uncertainties are considered for this comparison.

7.7 Final Discriminant

After finding optimal sets of input observables for the BDT training and after validating their independent modelling as well as the corresponding two-dimensional correlations, the BDT scores are calculated for each event according to the prescription of the cross-validation. In order to optimise the sensitivity of the final binned fit to their score distributions, the bin edges for the final histogram are defined according to the following criteria:

- Each bin must contain a sufficient number of background events, so that the statistical uncertainty is limited and does not improve the expected sensitivity due to downwards fluctuation of the background prediction.
- The drop of expected sensitivity in the sub-leading bins should be limited, so that a sufficient number of bins contains a significant signal contribution.
- Since the normalisation of the $Z \to \tau\tau$ background from embedding shall later be left unconstrained and obtained from the fit (cf. Sec. 7.9.1), it must be adjusted simultaneously with the large systematic uncertainty on the fake-τ_{h} background. Therefore, a shape difference between the two contributions must be represented in the binning, so that the two uncertainties are not degenerate in the final fit.

Starting with the highest bin and proceeding in steps of 0.001 for the VBF category and 0.01 for the Boosted category, a dedicated re-binning algorithm defines a new bin edge if

- each dominant background component yields a contribution in the new bin, while at least 10 embedded events are required for the $Z \to \tau\tau$ component,
- the total background expectation does not drop with respect to the previous bin,

and at least one of the following criteria is fulfilled:

- the expected sensitivity $S/\sqrt{B}$ in the new bin does not drop by more than 20% as compared to the previous bin (if any), and the bin edge corresponds to a step size of 0.01.
- the total background expectation increases by more than 20%.
- the fraction of $Z \to \tau\tau$ background to the fake-τ_{h} component changes by more than 20%.

With this algorithm, new bin edges are defined as

$$[-1.00, -0.90, -0.55, -0.20, 0.15, 0.40, 0.60, 0.70, 0.75, 0.80, 0.85, 0.90, 1.00] \quad \text{for Boosted,}$$

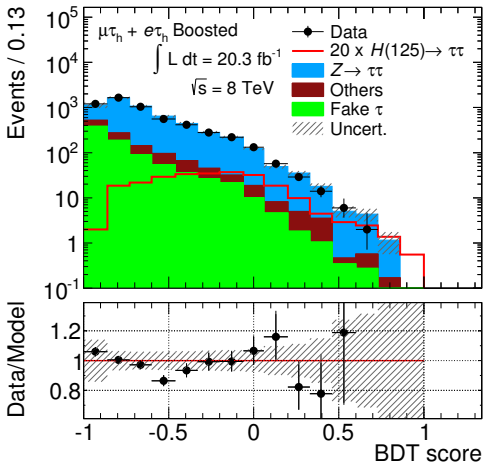
$$[-1.00, -0.90, -0.70, -0.35, 0.00, 0.30, 0.50, 0.65, 0.80, 0.90, 0.95, 0.987, 1.00] \quad \text{for VBF,}$$

which are applied directly to the BDT scores obtained from the training. Since their original distribution does not carry any physical meaning and is only a measure for background- or signal-likeliness of an event, these new bins are then re-mapped to equidistant intervals. This mainly improves the graphical presentation of the distributions, but does not affect the physical meaning, so that the re-binned and re-mapped scores are considered as the final discriminant in the following without further reference to the original distribution.

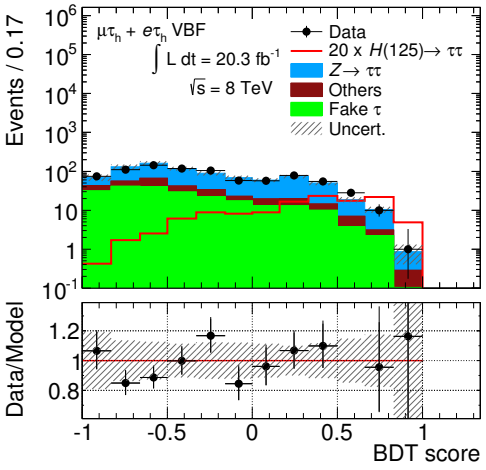
With these re-defined BDT scores, the next step in the validation chain is a check of the final BDT score distribution in the different control regions (cf. Tab. 7.4), which could reveal the mis-modelling

of more complicated correlations among the training observables. In each of these regions, a specific background component is enhanced, so that the modelling of this background can be validated in terms of the final discriminant. As shown in Fig. 7.18, good agreement between each background component and the data is observed within uncertainties[13]. The modelling of the backgrounds is therefore considered to be valid for all observables and correlations exploited by the trained BDTs, so that a reliable result can be expected from a final fit of signal and background distributions to the data observed in the signal regions.

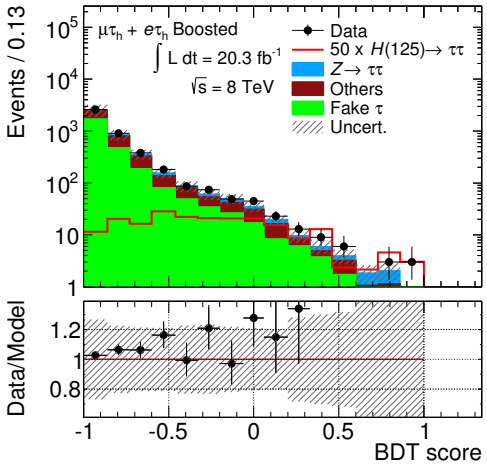
[13] A slope in the W+jets control region for the Boosted category (cf. Fig. 7.18c) is covered by the systematic uncertainties of the fake-τ_h estimate and is consistent with the pull of Fake_boost_RW (cf. Figs. 7.20a, 7.27) obtained from the final fit in the signal region.



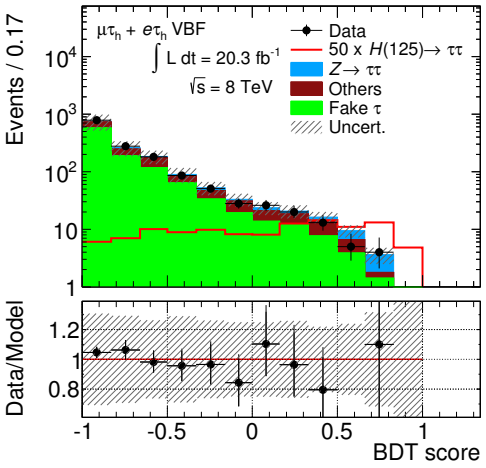
(a) Boosted $Z \to \tau\tau$ CR



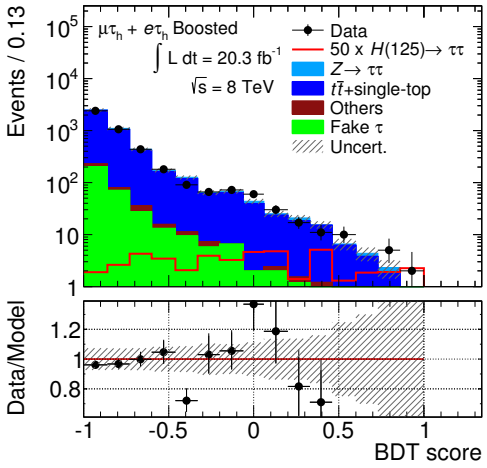
(b) VBF $Z \to \tau\tau$ CR



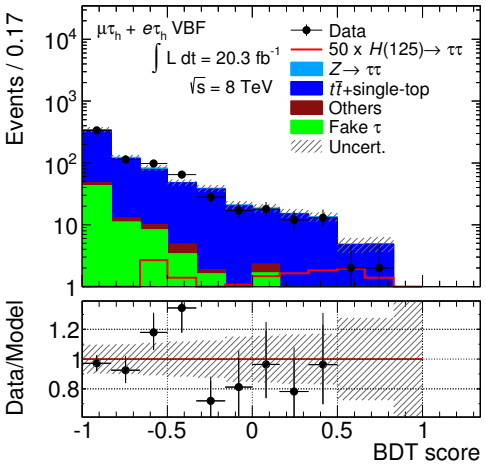
(c) Boosted W+jets CR



(d) VBF W+jets CR



(e) Boosted Top CR



(f) VBF Top CR

Figure 7.18: Distributions of the BDT scores in the different control regions. The background components are estimated with the techniques described in Sec. 7.4. Their uncertainties include the systematic uncertainties without constraints from the final fit.

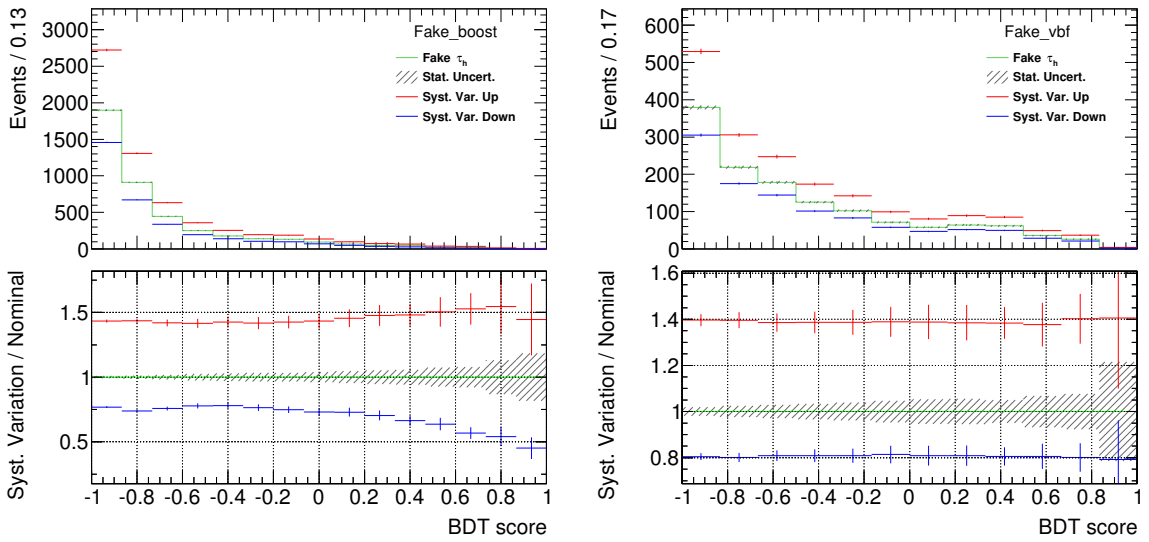
7.8 Systematic Uncertainties

In order to account for systematic mis-modelling effects in the predicted background and signal yields, uncertainties related to the theoretical calculations, the detector simulation and the techniques applied for the background estimation are evaluated, and corresponding systematic variations are performed.

7.8.1 Background Modelling Uncertainties

Systematic uncertainties originating from the isolation requirement (`EMB_ISOL`) and muon cell subtraction (`EMB_MFS`) within the τ-embedding procedure are estimated (cf. Sec. 5.2.8) and are treated as correlated between the $Z \rightarrow \tau\tau$ contributions in all categories.

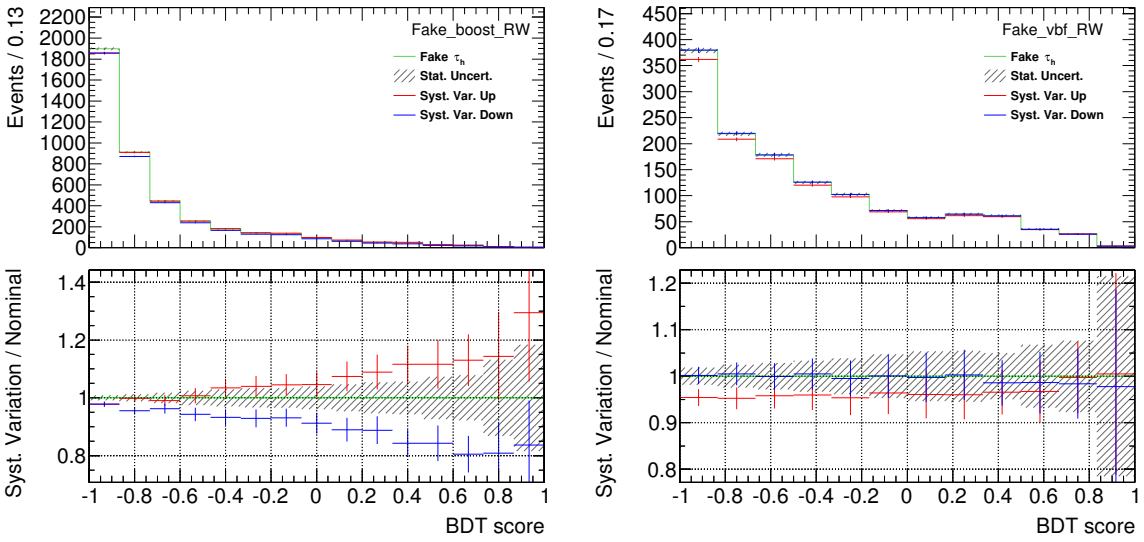
As described in Sec. 7.4.3, the systematic uncertainties on the fake-τ_{h} background contribution are estimated by varying the fake-factor f_{fake} within the measured statistical uncertainties and the relative contribution r_W from W+jets at the same time. An uncertainty of 100% is applied to the measured r_W, which results in BDT shape dependent variations of 20–50% after propagation to the background estimate (cf. Fig. 7.19). The corresponding systematics are treated as uncorrelated between all categories entering the fit (`Fake_boost`, `Fake_vbf`)[14], since the true central values are not expected to be identical. In addition, uncertainties on the correction of the $\vec{E}_{\text{T}}^{\text{miss}}$direction in the anti-$\tau_{\text{h}}$ events (`Fake_boost_RW`, `Fake_vbf_RW`) are estimated (cf. Sec. 7.4.3). The corresponding shape variations on the BDT score distribution are shown in Fig. 7.20.



(a) Boosted Category

(b) VBF Category

Figure 7.19: BDT score dependent shape and normalisation systematic uncertainties for the fake-factor based background estimation for (a) the Boosted and (b) the VBF category. The green line shows the nominal distribution, for which the statistical uncertainty is represented by the shaded band. The red (blue) line depicts the systematic upwards (downwards) variation.



(a) Boosted Category

(b) VBF Category

Figure 7.20: The effect of the $E_{\mathrm{T}}^{\mathrm{miss}}$ correction on the BDT score distribution of the fake-τ_{h} estimate in (a) the Boosted and (b) the VBF category. The green line shows the nominal distribution, for which the statistical uncertainty is represented by the shaded band. The red (blue) line depicts the systematic upwards (downwards) variation.

Validation of the Fake-Factor Systematic Uncertainty

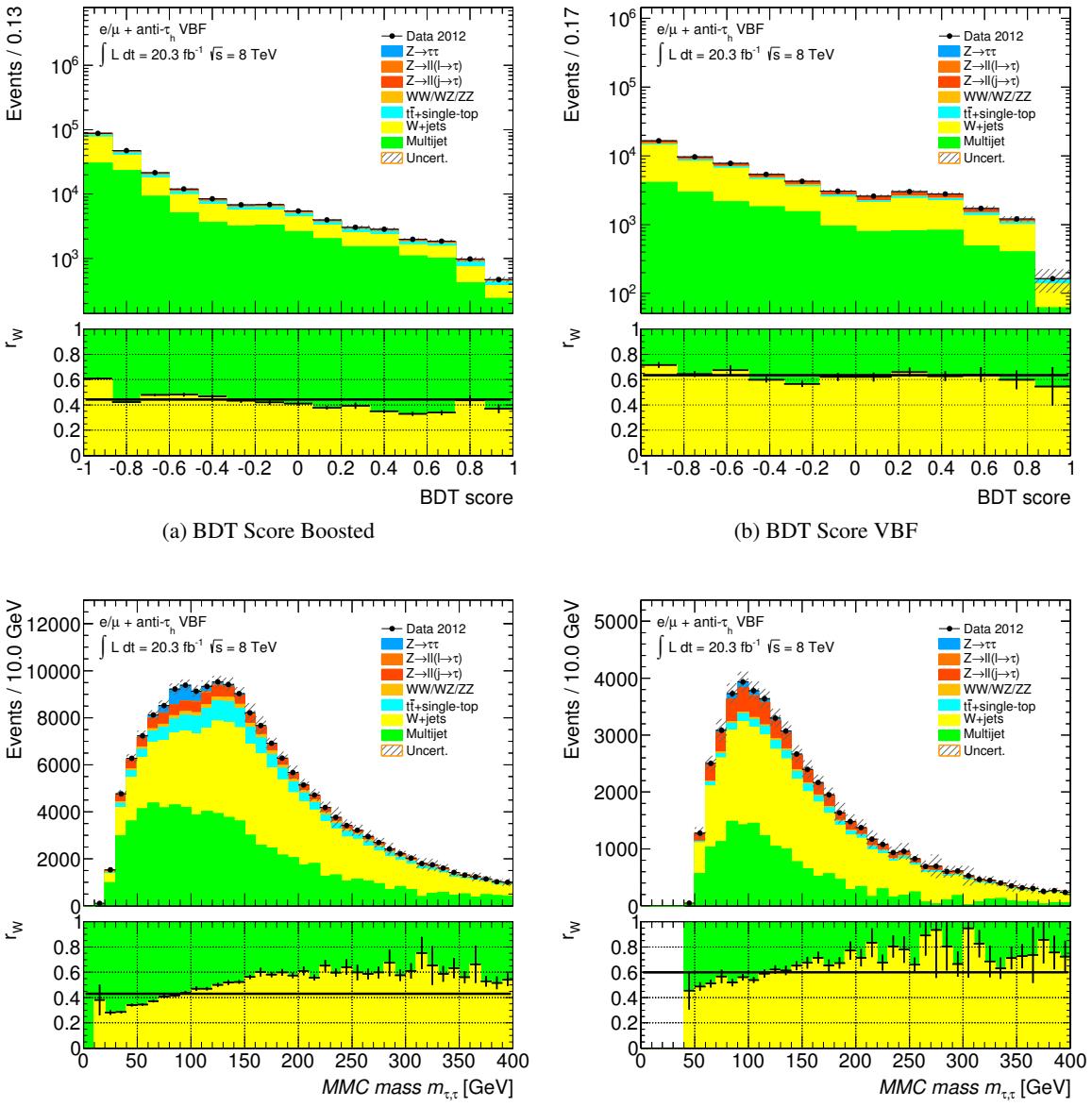
Despite the conservative uncertainty applied to the value of r_W, the fake-factor method still relies on the assumption that the ratio r_W is constant over the full BDT range. Otherwise, a shape uncertainty would be invalid in case of a very different r_W at low BDT score values as compared to the high BDT score region: if the fit of the background model to data effectively results in a constraint on r_W from the low BDT score region, the fitted value might not apply for high scores where the signal is expected. Therefore, the dependence of r_W on the BDT score is tested in the anti-τ_{h} control regions for the Boosted and VBF selection. In Fig. 7.21, the ratio $r_W = N_{W+\mathrm{jets}}/(N_{\mathrm{Multijet}} + N_{W+\mathrm{jets}})$ is shown for the assumption that the multijet contribution is given by the difference between the data and all simulated backgrounds in the anti-τ_{h} control regions. This ratio is found to be consistent with the assumption of a constant r_W within the constrained uncertainties obtained from the final fit to the data (cf. Tab. C.10). The systematic variation of the fake-factor estimate via r_W is thus applicable as a shape uncertainty and can be treated as correlated between the different bins of the BDT score distributions[15].

Uncertainties on the Z+jets Estimate

For the modelling of the Z+jets contribution in the VBF category, two additional systematics are considered in order to account for possible mis-modelling effects. Since the default Monte Carlo samples generated with ALPGEN+PYTHIA are merged with filtered ones from ALPGEN+HERWIG, different correction factors for the $j \to \tau_{\mathrm{h}}$ mis-identification are required (cf. Tab. 6.2, 7.3). Using the

[14] For the two top control regions, two additional uncorrelated systematics (Fake_btop, Fake_vtop) are included.

[15] Instead of using a shape systematic, this uncertainty would otherwise have to be implemented independently for each bin, so that fitted constraints from one bin cannot affect the background expectation in other bins.



(a) BDT Score Boosted

(b) BDT Score VBF

(c) MMC Mass Boosted

(d) MMC Mass VBF

Figure 7.21: BDT score and MMC mass distribution for anti-τ events in the Boosted and VBF category. The difference between data and all other background predictions in these control regions is taken as estimate for the multijet contribution. The lower plots show the ratio of $N_{W+\text{jets}}$ to the sum of $N_{W+\text{jets}} + N_{\text{Multijet}}$, which corresponds to r_W. The black line corresponds to the assumption of a constant r_W. This assumption is found to be valid within systematic uncertainties for the BDT score distributions, while a clear trend is observed for the MMC mass.

ALPGEN+PYTHIA correction factors, the weight of events generated with ALPGEN+HERWIG are scaled accordingly, while the difference between both is extracted as systematic uncertainty (Zll_HERWIG) for the $Z \to \ell\ell(+j{\to}\tau_{\mathrm{h}})$ contributions. In addition, a mis-modelling of the simulated $\Delta\eta(j_1, j_2)$ distribution has been observed in comparison to data in $Z \to \mu\mu$ and $Z \to ee$ control regions, which is corrected by a corresponding re-weighting scheme. An effect of 10% on the efficiency of standard VBF selections has been observed, which is considered as systematic uncertainty (Zlt_DETAJJ) for the $Z \to \ell\ell$ ($\ell{\to}\tau_{\mathrm{h}}$) contributions[16].

7.8.2 Experimental Systematic Uncertainties

The following experimental uncertainties are considered:

- Reconstruction, identification, isolation and trigger efficiencies: the τ_{h} identification efficiency as obtained from simulated events is corrected to values measured from data using tag and probe methods [90]. Systematic effects are derived by varying the corrections within their uncertainties (TAU_ID). Similarly, all other efficiencies related to the electron and muon analysis objects, which are taken from simulation, are corrected to the values of corresponding tag and probe measurements and varied within the extracted uncertainties (EL_EFF, MU_EFF). For the embedded $Z \to \tau\tau$ events, these analysis objects are also produced in a simulation, and similar corrections and uncertainties are therefore applied as well. These are, however, treated as uncorrelated with the uncertainties of the standard simulated objects in order to account for the different environment and the missing trigger information in the embedded samples (EL_EFF_Emb, MU_EFF_Emb).

- Lepton energy/momentum resolution: the electron energy and the muon momentum are smeared according to the resolutions measured in data. The resulting uncertainties (EL_RES, MU_SCALE) are found to have only negligible impact on the final result.

- $\ell{\to}\tau_{\mathrm{h}}$ mis-identification: according to the measured $e{\to}\tau_{\mathrm{h}}$ mis-identification probability, the corresponding uncertainty is considered as a systematic for the Z+jets ($e{\to}\tau_{\mathrm{h}}$) contribution (TAU_EFAKE). Since no corresponding measurement of the almost negligible $\mu{\to}\tau_{\mathrm{h}}$ mis-identification probability has been performed so far, a conservative estimate of 15% is applied as systematic for the Z+jets ($\mu{\to}\tau_{\mathrm{h}}$) contribution.

- b-tagging efficiency: the b-tagging efficiencies are corrected to match those measured in data, using a sample of jets containing muons [166]. Systematic uncertainties on these corrections are considered separately for the efficiencies on b-jets, c-jets and jets initiated by light flavour quarks (BTag_BEFF, BTag_CEFF, BTag_LEFF).

- τ_{h} energy scale (TES): the energy and momenta of the τ_{h} candidates are varied according to the TES uncertainties described in [88]. For each variation and each event, the introduced difference on the energy of the τ_{h} candidate is propagated to the calculated $E_{\mathrm{T}}^{\mathrm{miss}}$ in order to account for their correlation. The corresponding systematic uncertainties are treated as uncorrelated between background components with real hadronic decays of τ-leptons (TES_TRUE) and others with fake-τ_{h} contributions (TES_FAKE), which as jets or electrons are not expected to have the same true TES value.

[16] For contributions with a mis-identified jet ($j{\to}\tau_{\mathrm{h}}$), the situation is different, since this $\Delta\eta(j_1, j_2)$ distribution does not directly refer to the distance between two jets but to the sum of the distances between three jets: $\Delta\eta(j_1, \tau_{\mathrm{h}}) + \Delta\eta(\tau_{\mathrm{h}}, j_2)$.

- Jet energy: the jet energy scale (JES) uncertainties arise from several independent sources [174] and are thus treated as uncorrelated systematics.
  - Uncertainties related to the detector simulation, the physics modelling and the data statistics for the measurement (`JES_Detector`, `JES_Modelling`, `JES_Statistical`).
  - Uncertainties related to the physics modelling and method including data statistics for the η-intercalibration (`JES_Eta_Modelling`, `JES_Eta_StatMethod`), which are especially important for jets with high $|\eta|$.
  - Effects due to a different response of gluon- and quark-induced jets in the simulated samples as compared to the data are included as flavour-related uncertainties, which are split into components of the flavour composition and modelled response (`JES_FlavComp_G`, `JES_FlavComp_Q`, `JES_FlavResp`). The composition is treated as uncorrelated between background samples with dominantly quark-initiated (`Q`: ggH, W+jets, Z+jets) and gluon-initiated (`G`: qqH, VH, top and di-boson) jets, while a different energy scale response is considered for b-jets (`JES_Flavb`).
  - Pile-up modelling uncertainties on JES related to the number of primary vertices (`JES_NPV`) for-time pile-up and the average expected $\langle\mu\rangle$ (`JES_Mu`) for out-of-time pile-up effects. In addition, uncertainties related to the jet area correction[17] (`JES_PileRho`) are treated separately for the different samples according to their dominant production channel via gluon-gluon- (`GG`: ggH, top), quark-gluon- (`QG`: W+jets, Z+jets) or quark-quark-interaction (`QQ`: qqH, VH, di-boson).
- Jet energy resolution: the uncertainty related to the resolution effects in jet energy measurements (`JER`) is obtained by smearing each jet in an event within the uncertainty of the resolution as determined by the in-situ measurement described in [175].
- Jet-vertex-fraction: systematic effects arising from the jet-vertex-fraction requirement for the jets (`JVF`) are estimated by varying the cut-value within the measured uncertainties of the corresponding efficiency measurements.
- Pile-up rescaling: the modelling of the simulated pile-up effects is improved by a re-weighting of the simulated events. Corresponding systematic uncertainties are estimated by scaling the underlying number for $\langle\mu\rangle$ by $\pm 3\%$ (`PU_RESCALE`). These predominantly affect the tracking and vertexing performance.
- $E_{\mathrm{T}}^{\mathrm{miss}}$: in addition to the re-calculation of $E_{\mathrm{T}}^{\mathrm{miss}}$ accounting for its correlations with TES and JES, the soft $E_{\mathrm{T}}^{\mathrm{miss}}$ contributions are taken into account, which originate from cells that could not be matched to reconstructed objects. Their energy scale and resolution uncertainties (`MET_SCALESOFT`, `MET_RESOSOFT`) are varied independently according to the performance measurements [93].
- Luminosity: an uncertainty on the integrated luminosity of $\pm 2.8\%$ is obtained from a preliminary calibration following the procedure described in [67].

7.8.3 Theoretical Uncertainties

For all signal and background samples, which rely on the simulation of Monte Carlo generated events, uncertainties on the event modelling are estimated by varying parameters and implementations espe-

[17] The average pile-up energy contribution ρ to a jet is usually corrected via its area A by subtracting a term $\propto \rho \times A$ from the measured energy.

cially relevant for QCD calculations. These provide systematic differences related to the QCD scale and the choice of parton distribution functions, the modelling of the $p_T(H)$ distribution related to the recoil of the Higgs boson against parton radiation and the underlying event as implemented in the Monte Carlo event generators.

For the QCD scale uncertainties, prescriptions from [58] and [155] are followed, which estimate higher order contributions of QCD loops relevant for the jet requirements applied in the categorisation. As a result, the corresponding uncertainties in the categories are determined to range between 5–29%, although the inclusive uncertainty for the *ggH* process is only 8%. In order to account for migrations between categories, the (anti-)correlation of the different 1-jet and 2-jet uncertainties (`QCDscale_ggH1in` and `QCDscale_ggH2in`) is taken into account. For other processes such as the *qqH* and *VH* production or the production of one or more vector bosons (`QCDscale_qqH`, `QCDscale_VH`, `QCDscale_V` and `QCDscale_VV`), only smaller effects of 1–5% are obtained.

An additional uncertainty is assigned to the radiation of a third jet in the *ggH* production. While the VBF categorisation only requires an η-separation of the two leading jets, which is already included as a normalisation uncertainty, the presence of a third jet can have an impact on the event kinematics and thus affect the BDT score distribution. Therefore, a shape uncertainty (`QCDscale_ggH3in`) is derived for the BDT score distribution of the *ggH* signal entering the VBF category [148], which is shown in Fig. 7.23e. It is obtained by following a recipe from [59] as an extension of [155].

Systematic uncertainties due to the choice of the parton distribution function are evaluated by comparing the different sets MSTW [36], NNPDF [176] and CT10 [126]. With MCFM [117], the obtained uncertainties have been verified to be consistent with the defined categories, so that 8% can be assigned for the *ggH* signal component (`pdf_Higgs_gg`) and 3% for the contributions from *qqH* and *VH* production processes (`pdf_Higgs_qq`) in all categories. For backgrounds, the corresponding uncertainty (4%) is only relevant for the production of one or two vector bosons, which are produced via *qq*-interaction (`pdf_qq`), while the uncertainty for the *gg*-induced $t\bar{t}$ production is later eliminated in the fit by a normalisation to data.

Another theoretical uncertainty is assigned due to the QCD radiation from the quark loop in the *ggH* production mechanism (cf. Sec. 4.2.4), which is related to the effect of quark masses in the matrix element calculation and parton shower matching scheme and especially affects the transverse momentum $p_T(H)$ of the Higgs boson [161]. While the default *ggH* simulation in ATLAS is obtained from POWHEG, the MC@NLO implementation produces a different differential spectrum for the $d\sigma/dp_T^H$ distribution. The difference between both distributions is used to derive a systematic effect related to this uncertainty (`Gen_Qmass_ggH`) and results in a normalisation uncertainty of 29% (18%) on the *ggH* signal yield in the Boosted (VBF) category.

As underlying event effects, uncertainties of 6% and 30% for the signal yields from *qqH* and *ggH* (`UE_qq`, `UE_qq`) production in the VBF category have been obtained by comparing the Perugia 2011C underlying event tuning [162] with the AUET2B tune [177].

7.9 Construction of the Fit Model for the Signal Extraction

For all sources of systematic uncertainties considered in this analysis, their impact on the category selection efficiencies and on the shape of the BDT score distribution is evaluated. In the final binned fit of the BDT score, the resulting variations are then connected to nuisance parameters θ (cf. Sec. 4.3.3), which are included in the final likelihood with gaussian constraint terms.

7.9.1 Control Regions and Normalisation Factors

In addition to the experimental and theoretical uncertainties, five unconstrained[18] normalisation factors are introduced for the $Z \to \tau\tau$, $Z \to \ell\ell(j{\to}\tau_{\mathrm{h}})$ and top backgrounds in order to obtain the normalisation of these backgrounds directly from the fit, so that their correlations with other systematic uncertainties are automatically taken into account. The $Z \to \tau\tau$ background is dominant in the signal regions and can thus be constrained well by the fit of the BDT score to the data. Since its normalisation is expected to be identical for the Boosted and VBF category, it is connected with the common factor `norm_Ztt`. The other two background components are too small to be constrained simultaneously from the same BDT score distribution. Therefore, the corresponding $Z \to \ell\ell(j{\to}\tau_{\mathrm{h}})$ and top control regions as shown in Fig. 7.22 are included with all systematic uncertainties in the fit, in which the event yields of the dominant background are simultaneously fitted to the data. Since these yields are subject to the different τ_{h} mis-identification probabilities in the different categories, the fits are performed with the separate normalisation factors `norm_Zll_boost` and `norm_Zll_vbf` (`norm_Top_boost` and `norm_Top_vbf`) for the two $Z \to \ell\ell(j{\to}\tau_{\mathrm{h}})$ (Top) control regions. In order to ensure the stability of the final fit, the previously obtained pre-fit normalisations as summarised in Tab. 7.3 are used as initial values for these unconstrained factors.

7.9.2 Statistical Uncertainties on the Background Model

While the fitted constraints on these normalisation factors mainly represent the statistical uncertainty of the data in the corresponding control regions, also the background estimates are obtained either from a limited number of simulated events or data events in separate control regions. The resulting statistical uncertainties are taken into account as systematic uncertainties in the fit model by introducing Poisson distributed nuisance parameters for those bins, in which the combined statistical uncertainty of the backgrounds exceeds 5% of the total background sum. This procedure introduces five additional parameters `stat_vbf_bin` (`stat_boost_bin`) for the highest bins of the BDT score distribution in the VBF (Boosted) signal region.

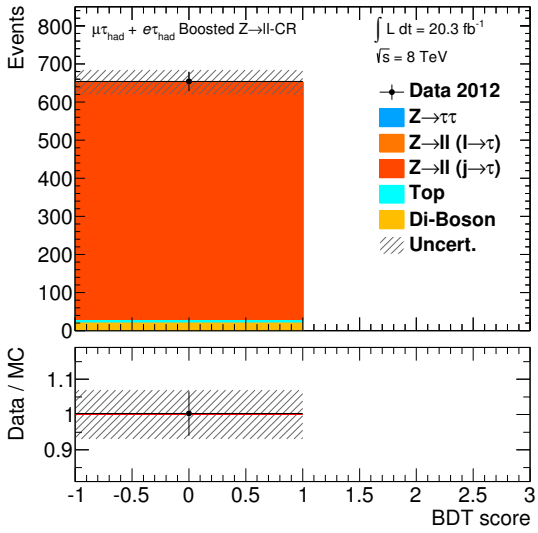
7.9.3 Validation of the Fit Model

The validation of the fit model is then performed in a *blinded* way: the full model is fitted to the data distribution, although the observed signal strength $\hat{\mu}$ and significance are kept hidden, so that the stability of the fit to the data can be tested without introducing any bias by comparing the observed result to the expected strength of $\hat{\mu} = 1$. From this fit, the behaviour of the nuisance parameters θ_i is extracted, which represent the implemented systematic uncertainties in the fit model.

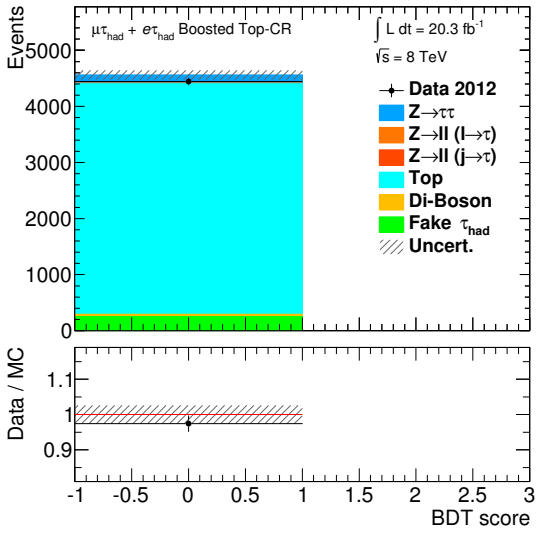
Treatment of Shape Uncertainties

For most systematic uncertainties, simple normalisation variations are included, which affect all bins of the BDT score distributions and are summarised in Tabs. C.3-C.8. These constant uncertainties are only considered if their effect on the corresponding signal or background sample is larger than 0.5%. In case that a systematic variation results in a significant variation of the BDT score distribution for a background or signal component in one of the signal regions, the resulting uncertainty is assigned bin-by-bin, so that the corresponding nuisance parameter is connected to a shape uncertainty. Since

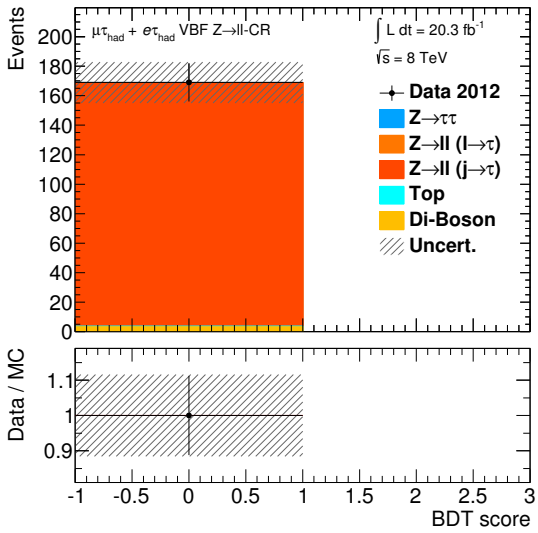
[18] In this context, "unconstrained" means that no Gaussian constraint term is included for the normalisation factors in the final likelihood function.



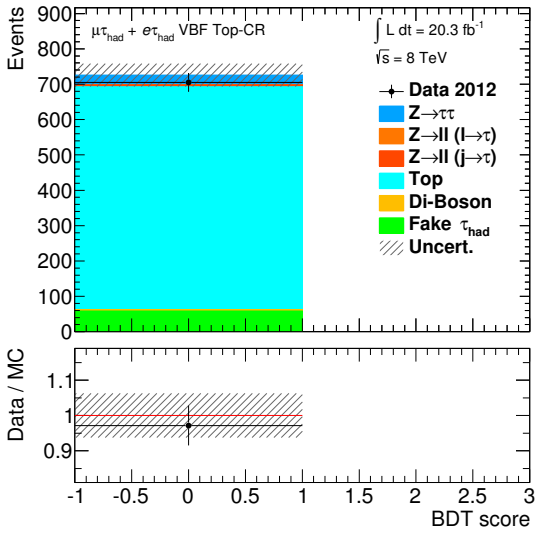
(a) Boosted $Z \to \ell\ell(+j{\to}\tau_{\mathrm{h}})$ CR



(b) Boosted Top CR



(c) VBF $Z \to \ell\ell(+j{\to}\tau_{\mathrm{h}})$ CR



(d) VBF Top CR

Figure 7.22: Control regions for the $Z \to \ell\ell(+j{\to}\tau_{\mathrm{h}})$ and top-quark backgrounds for the Boosted and VBF category, from which the free normalisations of these background components are constrained in the global fit.

the evaluation of the systematic variations is mostly performed on simulated samples with only a limited number of events, especially the estimated shape uncertainties suffer from statistical effects. While certain variations are obtained by re-weighting an identical set of events and thus represent a purely systematic uncertainty, other variations affecting the event kinematics (e.g. JES, TES, $E_{\mathrm{T}}^{\mathrm{miss}}$) can result in a different set of selected events, so that their systematic effect is usually entangled with statistical uncertainties in the final BDT score distribution. These statistical effects, however, should not be included in any systematic shape variation, since they effectively impose a correlation of such fluctuations between the bins. The fit would otherwise be able to constrain a statistical fluctuation similar to a systematic uncertainty, which can lead to wrong results. Since the statistical uncertainties of the default background estimate are already included as additional and uncorrelated nuisance parameters for each bin, their double-counting with each systematic uncertainty must be avoided. Therefore, additional *pruning* and *smoothing* algorithms are applied in several steps to all evaluated shape variations, which are expected to be entangled with statistical fluctuations. After extracting the variations of a systematic to the overall normalisation of a signal or background component, the significance of the remaining shape differences to the nominal histogram are then tested on their statistical probability and their significance with respect to the total uncertainty on the background estimate as described in the following steps:

- *Statistics Pruning*: the remaining shape differences to the nominal histogram are only considered in the fit model if the Kolmogorov-Smirnov (KS) [178, 179] probability either for the upwards or for the downwards variation is found to be smaller than 95% when compared to the nominal distribution[19]. In the other cases, in which a probability of $P(\mathrm{KS}) > 95\%$ is obtained for both directions of the variation, the additional shapes are discarded, since they are completely dominated by statistical fluctuations. All remaining systematic effects, which can possibly be hidden within these shape variations, are then expected to be small enough, so that they are already covered by the separate treatment of the background statistic uncertainty.

- *Shape Smoothing*: even if a systematic shape variation is found to be statistically significant when compared to the nominal histogram, it can still contain entangled fluctuations, which in certain cases can lead to wrong estimates of the systematic effects in the tail of the BDT score distribution. Therefore, a standard *smoothing* algorithm [180] is applied to each shape variation in order to enforce the assumption of smooth and continuous systematic effects on the analysis. Since the BDT score distributions themselves are not necessarily smooth, the algorithm is applied only to the bin-by-bin ratios built by dividing the varied histograms by the nominal. The smoothed ratios are then multiplied back to the nominal histogram, so that its original distribution is preserved and only the resulting variations are changed by the procedure. From these, statistical effects are then disentangled as far as possible, while any remaining fluctuation of the nominal is again expected to be small compared to the separated treatment of statistical uncertainties in the fit model.

- *Significance Pruning*: after reducing statistical effects from systematic variations, an additional pruning procedure is applied in order to remove background shape variations, which are negligible compared to the total statistical uncertainty of the background estimate. Therefore, systematic variations are only considered in combination with a shape if the maximal difference between upwards and downwards variation (u_i,d_i) of the background component in at least one bin i is

[19] The KS-test estimates the probability that the varied histogram is drawn from the same underlying distribution as the nominal histogram. For this test, the statistical uncertainty of the nominal histogram is ignored, since it is strongly correlated with the uncertainty of the varied histograms, and a double-counting should thus be avoided.

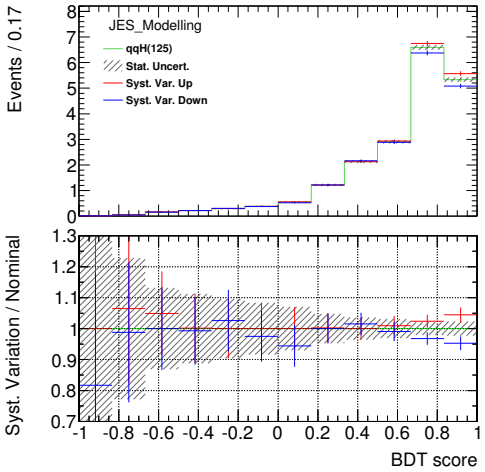
larger than 10% of the total statistical uncertainty σ_i^{tot} of the background sum in the same bin:

$$\max_i\{S_i\} = \max_i \left\{(u_i - d_i)/\sigma_i^{\text{tot}}\right\} > 0.1$$

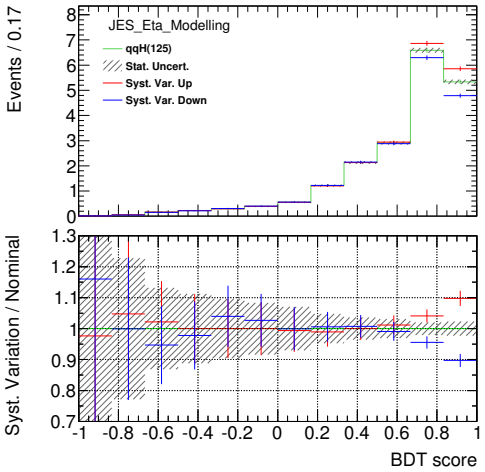
Otherwise, both shape variations are neglected for the systematic uncertainty and the background samples under consideration, and only the constant factors between variations and the nominal value are taken into account in the fit model. Similarly, a shape variation is neglected if the effect of the systematic is found to be smaller than 0.5% for all bins.

These procedures are found to significantly improve the stability of the fit to the data, which otherwise does not converge in a global minimum of the likelihood function and creates unphysical constraints on certain nuisance parameters. The exact pruning thresholds, the choice of the KS-test over a χ^2-test and the number of smoothing iterations, however, are found to be irrelevant for the fit result, which reaches a stable minimum after removing double-counted statistical effects and insignificant variations.

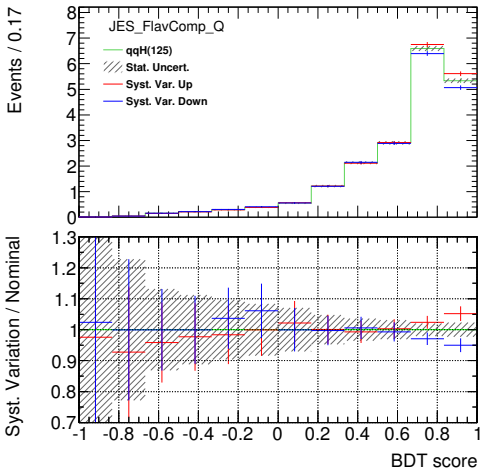
The most relevant shape uncertainties for this analysis are shown in Fig. 7.23 (Fig. 7.25) for the Higgs boson signal ($m_H = 125$ GeV) and in Fig. 7.24 (Fig. 7.26) for the background components in the VBF (Boosted) category.



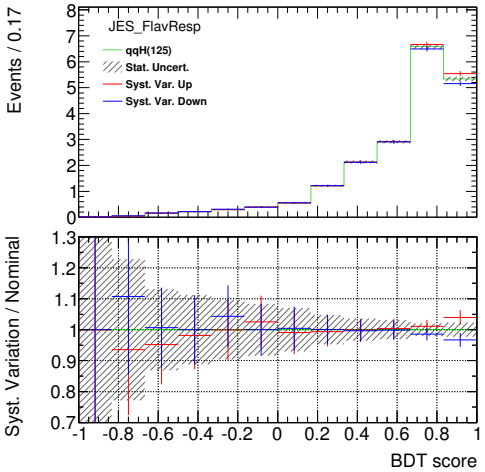
(a) qqH JES_Modelling



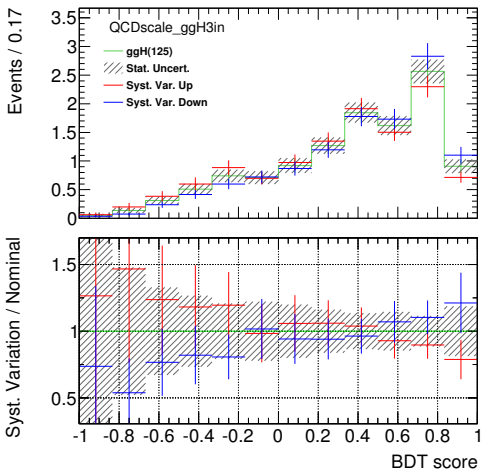
(b) qqH JES_Eta_Modelling



(c) qqH JES_FlavComp

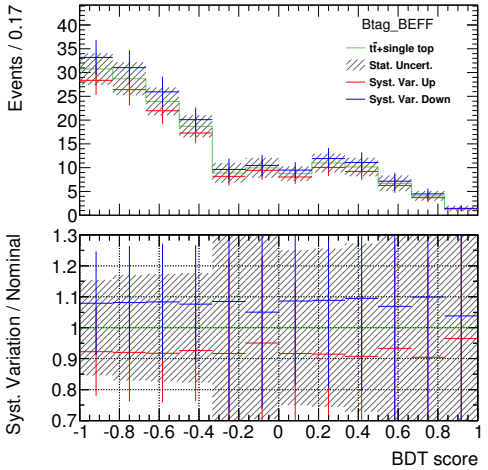


(d) qqH JES_FlavResp

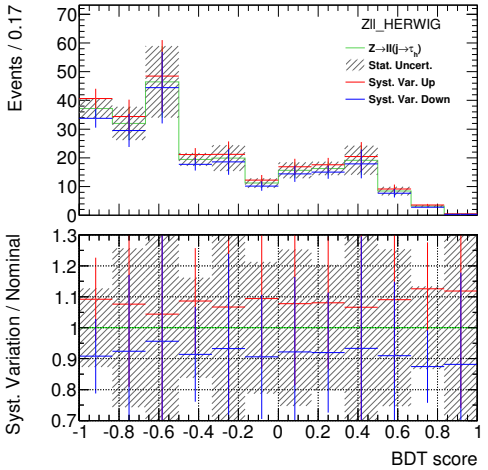


(e) ggH QCDscale_ggH3in

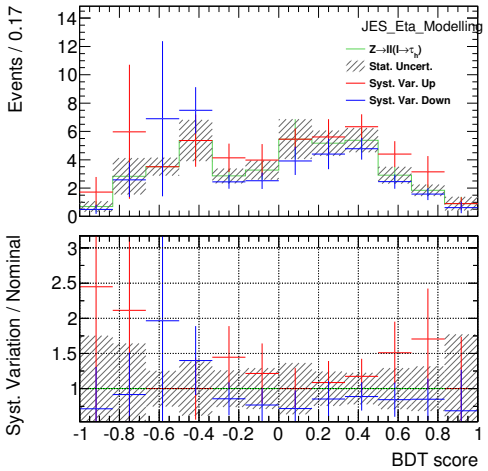
Figure 7.23: Relevant systematic uncertainties on the BDT score shape for the qqH and ggH ($m_H = 125$ GeV) signal contributions in the VBF category. The green line shows the nominal distribution, for which the statistical uncertainty is represented by the shaded band. The red (blue) line depicts the systematic upwards (downwards) variation.



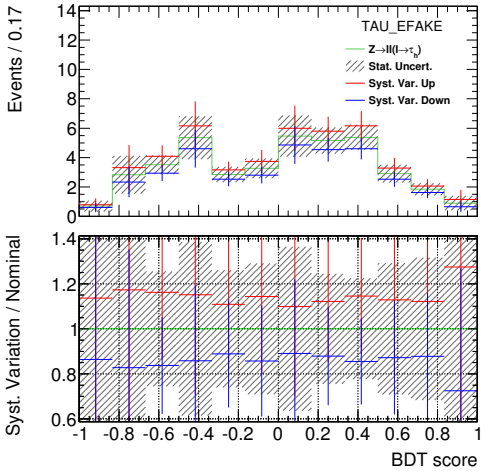
(a) Top Btag_BEFF



(b) $Z \to \ell\ell(+j \to \tau_{\mathrm{h}})$ Zll_HERWIG

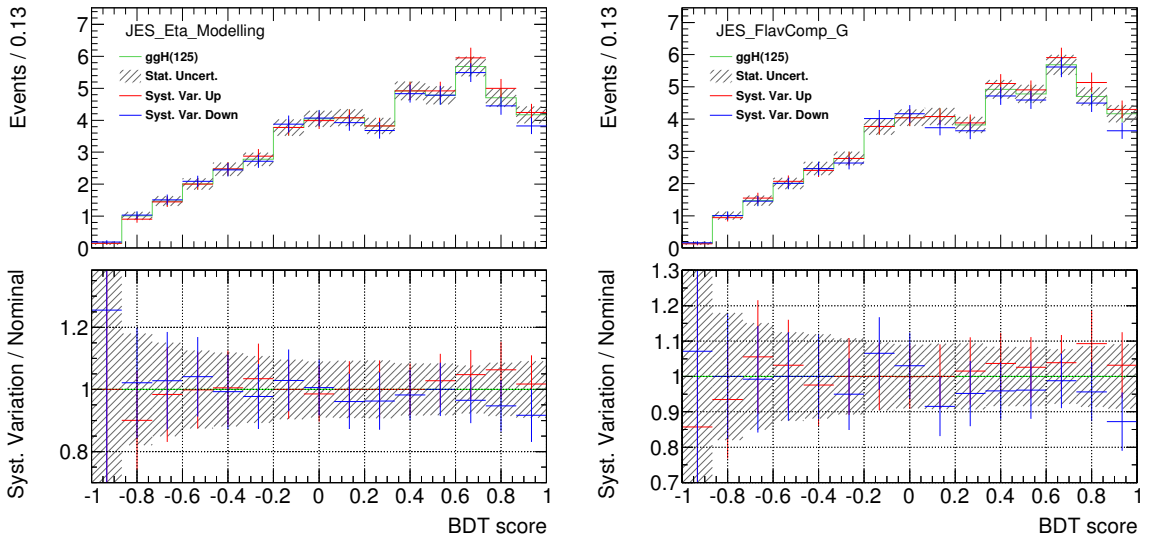


(c) $Z \to \ell\ell(\ell \to \tau_{\mathrm{h}})$ JES_Eta_Modelling



(d) $Z \to \ell\ell(\ell \to \tau_{\mathrm{h}})$ TAU_EFAKE

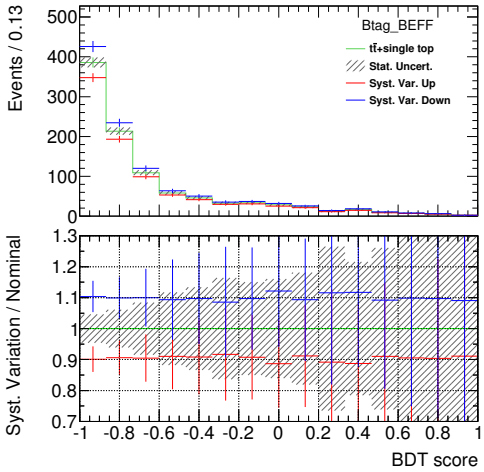
Figure 7.24: Relevant systematic uncertainties on the BDT score shape for background components in the VBF category. The green line shows the nominal distribution, for which the statistical uncertainty is represented by the shaded band. The red (blue) line depicts the systematic upwards (downwards) variation.



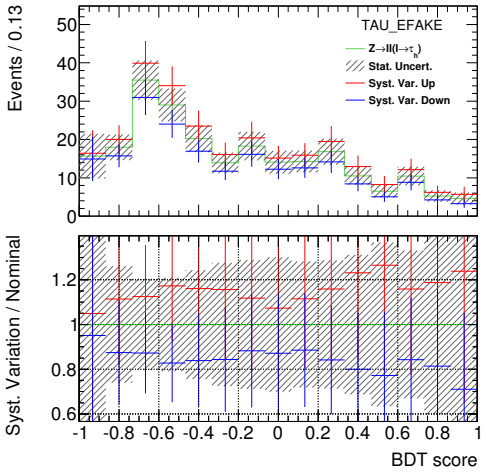
(a) ggH JES_Eta_Modelling

(b) ggH JES_FlavComp

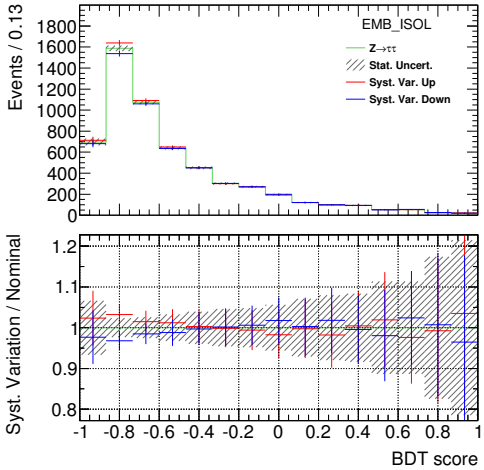
Figure 7.25: Relevant systematic uncertainties on the BDT score shape for the ggH ($m_H = 125$ GeV) signal in the Boosted category. The green line shows the nominal distribution, for which the statistical uncertainty is represented by the shaded band. The red (blue) line depicts the systematic upwards (downwards) variation.



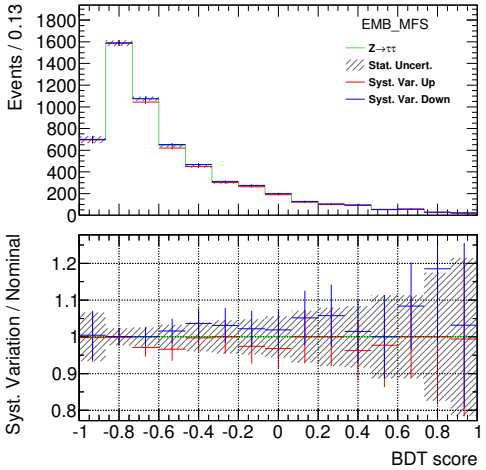
(a) Top Btag_BEFF



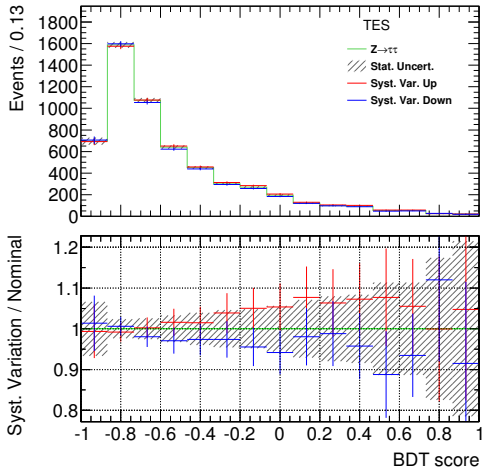
(b) $Z \to \ell\ell(\ell \to \tau_{\mathrm{h}})$ TAU_EFAKE



(c) $Z \to \tau\tau$ EMB_ISOL



(d) $Z \to \tau\tau$ EMB_MFS



(e) $Z \to \tau\tau$ TES

Figure 7.26: Relevant systematic uncertainties on the BDT score shape for background components in the Boosted category. The shaded bands represent the statistical uncertainties of the nominal distributions.

7.9.4 Systematic Uncertainties after the Final Fit

Tables C.9-C.11 summarise all systematics and their nuisance parameters with their fitted values $\hat{\theta}$ and uncertainties $\Delta\hat{\theta}$, which are compared to the initial values of $\theta_0 = 0$ and $\Delta\theta = 1$. In order to visualise these results in combination with the importance of the different uncertainties for this analysis, the corresponding nuisance parameters are ordered by their impact on the fitted value of the signal strength $\hat{\mu}$ in Fig. 7.27 and Tabs. C.12-C.13. This impact is evaluated by varying each nuisance parameter by $\pm 1\sigma$ of its post-fit (pre-fit) uncertainty and thus fixing it to $\hat{\theta}_i \pm \Delta\hat{\theta}_i$ ($0 \pm \Delta\theta = \pm 1$). After performing the complete fit again, the new result $\hat{\mu}(\theta_i = \hat{\theta}_i \pm \Delta\hat{\theta}_i)$ is then compared to the unconstrained signal strength $\hat{\mu}$ and its uncertainty $\Delta\hat{\mu}$, so that the relative uncertainty $\Delta\hat{\mu}(\Delta\hat{\theta}_i)$ can be evaluated.

In this validation, most of the fitted nuisance parameters are found to remain close to their initial values, and their fitted uncertainties are not significantly constrained by this fit. For the parameters, which are significantly affected by the fit, the observed behaviour is well understood and discussed in the following.

Fake-Factor Systematics

The estimate of the fake-τ_h backgrounds involves the extrapolation of the ratio r_W from the high-m_T control region to the signal region, for which highly conservative uncertainties are assigned separately for the two signal regions (`Fake_boost`, `Fake_vbf`). This large uncertainty of varying the measured r_W within $[0, 1]$ is therefore strongly constrained in the final fit, while the fitted value is slightly pulled from the initial value. This corresponds to a slightly higher r_W than extrapolated from the control region, for which the final fit can constrain the uncertainty from background dominated regions of the BDT score distributions. In addition, a nuisance parameter for the ad-hoc systematic uncertainty assigned for the re-weighting procedure of the $E_\mathrm{T}^\mathrm{miss}$ direction (`Fake_boost_RW`) is pulled towards a higher value. The corresponding weights are obtained from a control distribution and are subject to statistical fluctuations in the data, so that a pull within the assigned uncertainty is expected especially for the Boosted category, where the statistical uncertainty of the data is smaller for the fit in the fake-τ_h dominated region as compared to the VBF category.

Embedding Systematics

Since the embedding procedure suffers from muon selection and cell subtraction effects, which cannot be evaluated perfectly, the corresponding systematic uncertainties (`EMB_ISOL`, `EMB_MFS`) are conservatively estimated in order to cover the expected central value, while the chosen central value is not necessarily a reliable choice. From the final fit of the BDT score distributions to the data, the $Z \to \tau\tau$ dominated regions constrain these uncertainties and prefer higher values for the corresponding nuisance parameters, which can be interpreted as a tighter isolation selection of the muons as well as a higher amount of cell energy subtracted from the simulated muons. For future embedding productions, these results could therefore be used to tune the default selection and cell subtraction parameters, so that the nominal sample can be improved, and the systematic uncertainties can be reduced.

Systematics Related to τ_h Candidates

This fit to the data in an $\ell\tau_\mathrm{h}$ final state is sensitive to the τ_h energy scale (TES) for real τ_h objects. Since this TES nuisance parameter (`TES_TRUE`) is de-correlated from the energy scale uncertainty for fake-τ_h (`TES_FAKE`), its sensitivity is mainly related to the real τ_h objects from the embedded $Z \to \tau\tau$ background component with free normalisation and thus mainly to the corresponding shape uncertainty

shown in Fig. 7.26e. Although this does not have enough statistical power to constrain this systematic uncertainty via its nuisance parameter, the fit prefers a slightly lower value for the TES. This is found to be consistent with several studies in $Z \to \tau\tau$ analyses as well as the analysis of the $H \to \tau_{\text{had}}\tau_{\text{had}}$ channel [148]. In addition, the mis-identification probability of electrons as τ_{h} candidates (`TAU_EFAKE`) prefers a lower value in the fit to the data. Since the applied central value for this probability is obtained from $Z \to ee$ events, for which no further jets are required, the transverse momentum range of the Z bosons and subsequently of the mis-identified τ_{h} candidates is expected to be lower than with this analysis selection. The measured central value does not correspond to the fitted value in this analysis. With a dedicated measurement for boosted $Z \to ee$ events or an improved binning of the correction factors in $p_T(e{\to}\tau_{\text{h}})$, this small tension within systematic uncertainties could be reduced.

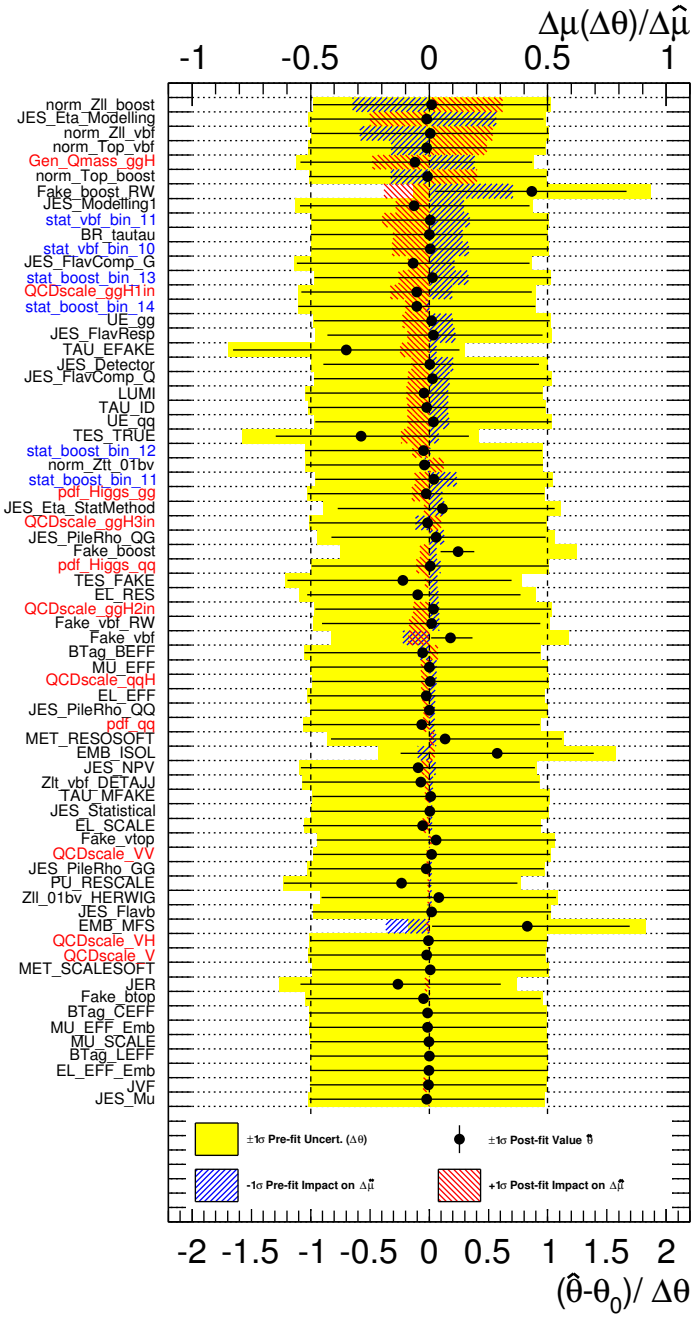
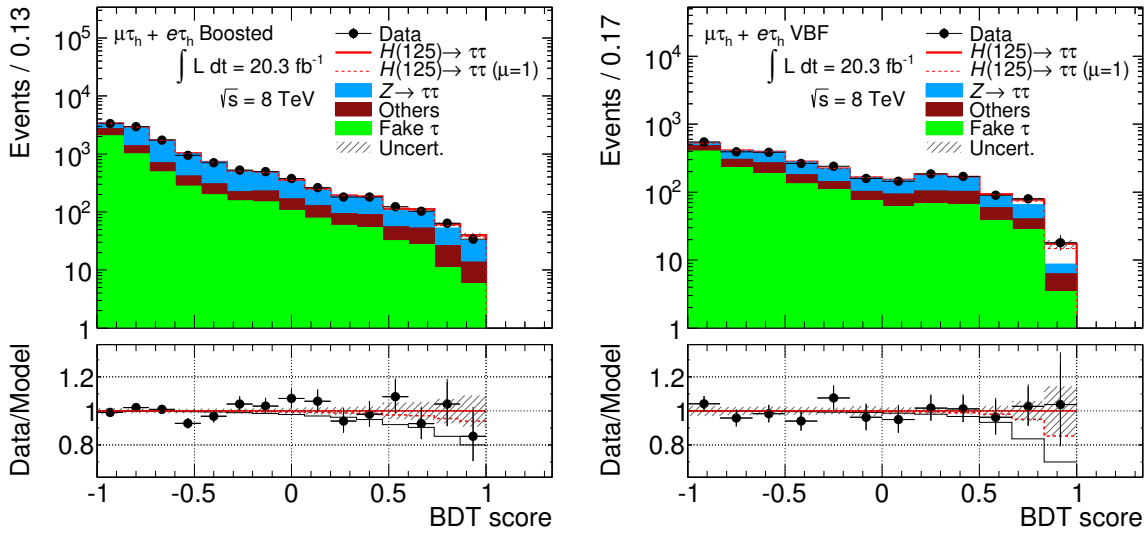


Figure 7.27: Fitted values and uncertainties of the nuisance parameters for the systematics related to theory (red), experimental effects (black) and statistics of Monte Carlo samples and data-driven estimates (blue), which are ordered by their post-fit impact on the signal strength parameter $\hat{\mu}$. The most important pulls and constraints on nuisance parameters are discussed in the text.

7.10 Results

After the thorough validation of the control distributions, of the final distribution in control regions and of the full fit model as described in the previous sections, the result of this search for the Higgs boson is obtained from the BDT score distributions in the signal regions. In Fig. 7.28, these distributions are shown for the fitted signal and background distributions in the Boosted and VBF categories. In the highest bins of these distributions, a clear excess of the observed data over the expected background contribution is visible (cf. Tab. 7.8), which is consistent with a signal of a Higgs boson with a mass of $m_H = 125$ GeV.



(a) Boosted Category

(b) VBF Category

Figure 7.28: Final distributions of the BDT scores for (a) the Boosted and (b) the VBF category after the final fit.

7.10.1 Measurement of the Signal Strength

As the final result from the fit, the best agreement of signal and background is observed with a fitted signal strength of $\hat{\mu} = 1.44^{+0.64}_{-0.52}$. As compared to the expected signal strength of $\mu = 1$ for the Standard Model Higgs boson with $m_H = 125$ GeV, the observed value is only slightly larger but well compatible within uncertainties. These can be split into its components by fixing nuisance parameters to their fitted values and re-evaluating $\Delta\hat{\mu}$ and can thus be summarised as a measurement of:

$$\hat{\mu} = 1.44^{+0.40}_{-0.40}(\text{stat})\ ^{+0.45}_{-0.26}(\text{syst})\ ^{+0.22}_{-0.22}(\text{theo})$$

The contributions of the most important components to these systematic uncertainties are listed in Tab. 7.9. As already visible from the nuisance parameter ranking in Fig. 7.27, the normalisations of the $Z \to \ell\ell(j{\to}\tau_\text{h})$ and top background estimates have a large impact on the uncertainty of the fitted signal strength, since their contribution in the highest BDT score bins is still sizeable. Due to a limited number of data events in the dedicated control regions, the corresponding constraints from the fit are weaker than for the normalisation of the $Z \to \tau\tau$ background, which is dominant in the low BDT score bins of the signal region. An inclusion of the corresponding components from $j{\to}\tau_\text{h}$ mis-identification into the

Process/Category	VBF			Boosted		
BDT score bin edges	0.5-0.667	0.667-0.833	0.833-1.0	0.6-0.733	0.733-0.867	0.867-1.0
ggF	2.2 ± 0.9	3.5 ± 1.5	1.2 ± 0.6	7.7 ± 2.9	6.3 ± 2.3	5.5 ± 2.1
VBF	4.1 ± 1.2	9.2 ± 2.7	7.5 ± 2.2	1.7 ± 0.5	1.5 ± 0.5	1.3 ± 0.4
WH	< 0.05	< 0.05	< 0.05	0.95 ± 0.29	0.85 ± 0.26	0.81 ± 0.25
ZH	< 0.05	< 0.05	< 0.05	0.42 ± 0.13	0.47 ± 0.14	0.41 ± 0.12
$Z \to \tau\tau$	28.6 ± 1.4	25.0 ± 1.6	2.41 ± 0.35	48.3 ± 3.4	26.1 ± 2.7	18.4 ± 2.0
Fake	37.7 ± 1.8	27.9 ± 2.1	3.5 ± 0.5	27 ± 4	10.8 ± 1.8	5.8 ± 1.4
Top	6.5 ± 0.7	4.1 ± 0.8	1.5 ± 0.4	7.0 ± 0.9	5.7 ± 0.8	2.23 ± 0.33
Di-boson	2.9 ± 0.4	3.0 ± 0.5	0.23 ± 0.04	4.8 ± 0.5	4.0 ± 0.5	1.69 ± 0.23
$Z \to \ell\ell(+j \to \tau_{\mathrm{h}})$	8.7 ± 1.7	3.3 ± 0.5	0.40 ± 0.10	3.8 ± 0.5	0.71 ± 0.07	< 0.05
$Z \to \ell\ell(\ell \to \tau_{\mathrm{h}})$	2.8 ± 1.2	1.9 ± 1.2	0.7 ± 0.6	9.4 ± 1.9	4.9 ± 1.1	3.8 ± 1.2
Total Background	87.2 ± 2.7	65 ± 5	8.7 ± 2.5	101 ± 6	52 ± 4	32 ± 4
Total Signal	6.3 ± 1.8	12.7 ± 3.5	8.7 ± 2.4	10.7 ± 3.3	9.2 ± 2.8	8.0 ± 2.5
S/B	0.07	0.20	1.0	0.11	0.18	0.25
Data	90	80	18	103	64	34

Table 7.8: Predicted event yields for $m_H = 125$ GeV in the three highest bins of the BDT score distributions [171]. The background normalisations, signal normalisation and uncertainties reflect the preferred values from the global fit. The uncertainties on the total background and total signal reflect the full statistical and systematic uncertainties, while the individual background components contain the full systematic uncertainty only.

Source of Uncertainty	Uncertainty on μ
$Z \to \ell\ell$ normalisation in Boosted	0.20
Jet energy scale (η modelling)	0.18
$Z \to \ell\ell$ normalisation in VBF	0.17
Top normalisation in VBF	0.16
Jet energy scale (others combined)	0.14
Quark mass effect in ggH	0.13
Top normalisation (Boosted)	0.13
QCD scale (combined)	0.11
Fake-τ_{h} in Boosted: $E_{\text{T}}^{\text{miss}}$ re-weighting	0.11
BR($H \to \tau\tau$)	0.09
UE (ggH)	0.06
$e \to \tau_{\text{h}}$ mis-identification	0.06
τ_{h} energy scale	0.06
Luminosity	0.05
τ_{h} identification	0.05
UE (qqH)	0.05
Fake-τ_{h} in VBF: r_W	0.05
$Z \to \tau\tau$ normalisation	0.04
Fake-τ_{h} in VBF: MET	0.04
Fake-τ_{h} in Boosted: r_W	0.03

Table 7.9: The important sources of uncertainty on the measured signal strength parameter μ, given as absolute uncertainties on μ.

fake-factor estimate would therefore be able to improve the result of this analysis. As an experimental systematic uncertainty, the jet energy scale is dominant, while the component related to the calibration of jets with large η has the largest effect on the result. This is expected for the VBF category, which is based on a selection of two jets with a large separation in η and exploits their invariant mass via the BDT score. The theory systematics are dominated by the effect of quark masses on the jet radiation in the ggH production and by the uncertainties on the QCD scale and the branching ratio BR($H \to \tau\tau$). The conservative systematic variation of the fake-τ_{h} estimate is reasonably constrained by the fit, so that their impact on the final result of this analysis is small in comparison to the other sources of uncertainties. Only the $E_{\text{T}}^{\text{miss}}$ re-weighting in the Boosted category has a sizeable effect on μ, which is a result of the large shape variation of the BDT score in Fig. 7.20a as compared to the same effect in the VBF category (cf. Fig. 7.20b).

7.10.2 Separation of the Higgs Boson Production Modes

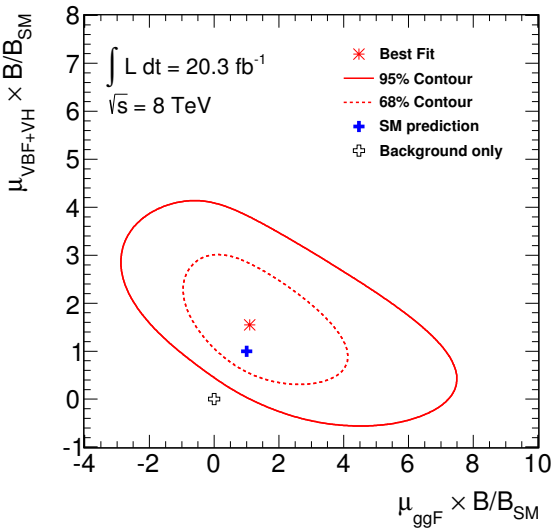


Figure 7.29: Likelihood contours in the ($\mu_{\mathrm{ggF}} \times B/B_{\mathrm{SM}}$, $\mu_{\mathrm{VBF}+VH} \times B/B_{\mathrm{SM}}$) plane for confidence levels of 68% and 95% shown as dashed and solid lines for $m_H = 125$ GeV. The SM and the background-only expectation are shown by a filled plus and an open plus symbol, respectively. The best fit to the data is shown for the case that the signal strength parameters μ_{ggF} and $\mu_{\mathrm{VBF}+VH}$ are unconstrained.

Since the analysis is based on two categories, which are sensitive to different signal production modes, this result can be split into two separate strength parameters μ_{ggF} and $\mu_{\mathrm{VBF}+VH}$ [181]. The best fit with two independent signals is obtained for the observed values of $\hat{\mu}_{\mathrm{ggF}} = 1.10^{+2.55}_{-1.70}$ and $\hat{\mu}_{\mathrm{VBF+VH}} = 1.55^{+1.28}_{-1.08}$. As shown in Fig. 7.29, the scan of these parameters provides likelihood contours for this best fit, which exclude the background-only point ($\mu_{\mathrm{ggF}} = 0$, $\mu_{\mathrm{VBF+VH}} = 0$) at 95% confidence level and are well consistent with the expectation for the Standard Model ($\mu_{\mathrm{ggF}} = 1$, $\mu_{\mathrm{VBF+VH}} = 1$). In this representation, both the small excess and the stronger constraint of the observed signal strength $\hat{\mu}$ above the expectation ($\mu = 1$) are found to be driven by the VBF category.

7.10.3 Statistical Significance of the Observed Excess

Interpreting the fit result in terms of the profile likelihood as described in Sec. 4.3.3, Fig. 7.30 shows the χ^2-distribution of the test statistic $q_{\mu=0}$ for the background-only case $f(q_{\mu=0}\|B)$ as well as for the case with a signal $f(q_{\mu=0}\|S + B)$. Integrating the first distribution from the observed value $q_{\mathrm{obs}} = 9.96$, the probability for the observed data to arise from the background-only hypothesis is found to be $p_0 = 8.0 \cdot 10^{-4}$, which corresponds to an exclusion of this hypothesis with a significance of 3.2σ. Corresponding to the observed signal strength, this observed significance is slightly larger than obtained from the Asimov data ($\mu = 1$), from which a significance of 2.3σ is expected. From the integral of the second distribution in Fig. 7.30, a probability of $p_1 = 0.21$ is obtained for the signal hypothesis, so that the observed excess is found to be compatible with a signal from a Standard Model Higgs boson with $m_H = 125$ GeV within 0.8σ.

With the observed significance of 3.2σ, this is direct evidence for a Standard Model $H \to \tau\tau$ signal in the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ decay channel. It is found to be consistent with the signal expectation from the Higgs boson with mass $m_H \approx 125$ GeV, which has already been observed in the bosonic decay channels.

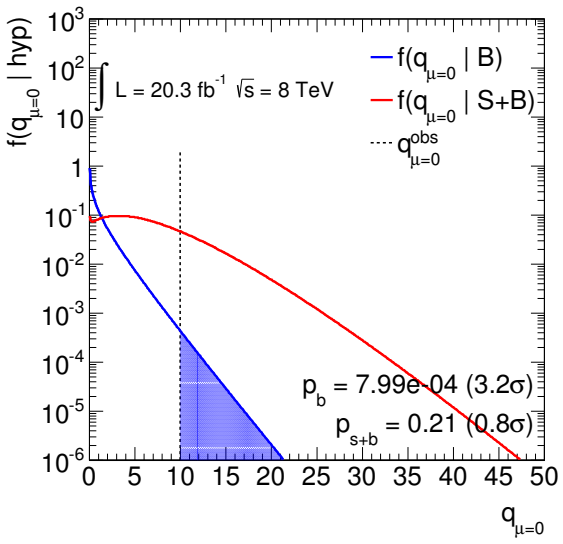
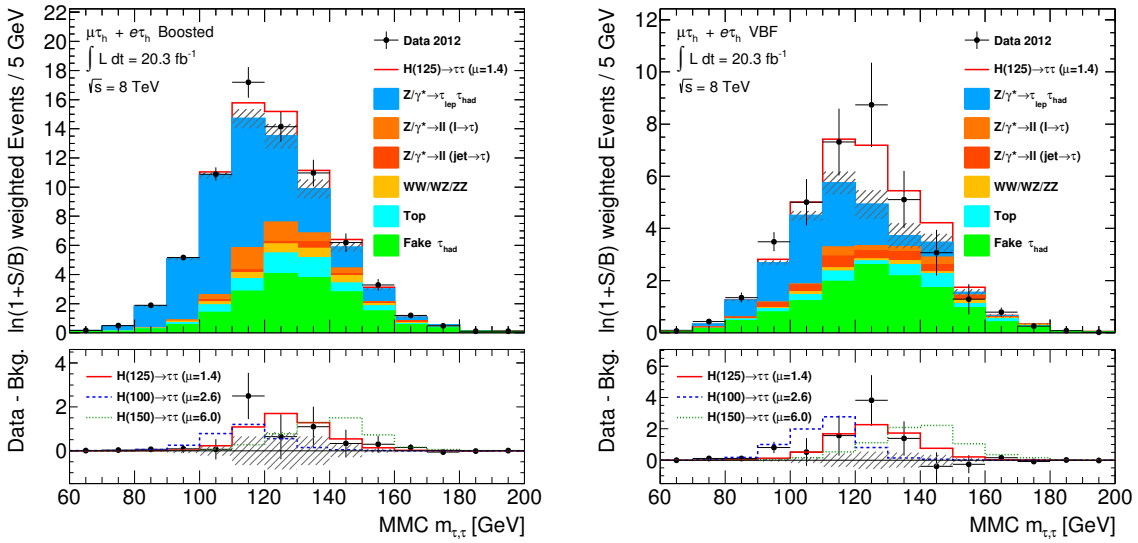


Figure 7.30: Distribution of the $q_{\mu=0}$ for the $H \to \tau_{\text{lep}}\tau_{\text{had}}$ background only (blue) and signal+background hypothesis (red) in the asymptotic approximation. The observed value (dotted line) obtained on data is also shown. The probability of the background only hypothesis (p_{b}) and the probability of the signal plus background hypothesis ($p_{\text{s+b}}$) are computed as the integral of $f(q_{\mu=0} \mid \text{hyp})$ over $q_{\mu=0} > q^{\text{obs}}_{\mu=0}$.

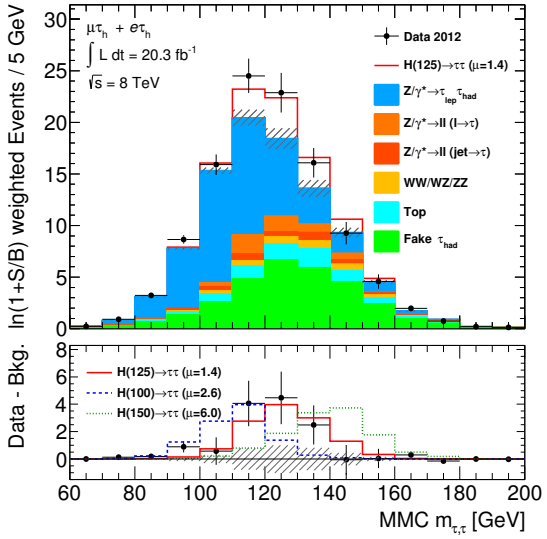
7.10.4 Visualisation of the Observed Higgs Boson Mass

In order to visualise the observation also in an invariant $\tau\tau$-mass distribution, Fig. 7.31 shows the MMC mass in the signal regions (cf. Fig 7.6) after applying a significance-based weighting. For this procedure, each event is weighted with the Poisson significance $\ln(1+S/B)$ of its corresponding BDT score bin, so that the weighted mass distribution is dominated by the very signal-like events. With this additional mass information, the discrepancy between the fitted signal templates $m_H = 100, 150$ GeV and the observed data becomes obvious, and the agreement of the data with $m_H = 125$ GeV can be confirmed.



(a) Boosted Category

(b) VBF Category



(c) Combined (Boosted+VBF)

Figure 7.31: Distributions for the MMC mass in the (a) Boosted and (b) VBF category and their combination (c), obtained by weighting events by the significance $\ln(1+S/B)$ of their assigned BDT score bin. Systematic uncertainties are shown with constraints from the global fit.

7.10.5 Combination of $H \to \tau\tau$ Decay Channels

Searches for $H \to \tau\tau$ decays in $\tau_{\text{lep}}\tau_{\text{lep}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ final states are performed in parallel, which are based on similar multi-variate analyses as presented here and both perform the signal extraction from BDT score distributions in a Boosted and a VBF category (cf. Fig 7.32). These are combined with the $\tau_{\text{lep}}\tau_{\text{had}}$ distributions into a single plot of significance-ordered BDT score bins shown in Fig. 7.33a, for which the upper edges of each bin are set to its S/B ratio[20]. In this arrangement, the three bins with the highest significance correspond to the most sensitive VBF-BDT score bins contributed from each $\tau^+\tau^-$ decay channel. This demonstrates that the excess of signal over background is consistently observed in all sensitive BDT score bins of the $\tau^+\tau^-$ decay channels.

The $\ln(1+S/B)$ weighted MMC distributions are combined in Fig. 7.33b, which confirms that all three signals observed in different $\tau^+\tau^-$ final states are compatible with the expected Higgs boson mass of $m_H = 125$ GeV.

The fitted signal strength parameter from the combination is measured as

$$\hat{\mu} = 1.43^{+0.31}_{-0.29}(\text{stat})^{+0.31}_{-0.26}(\text{syst})^{+0.26}_{-0.15}(\text{theo})$$

for $m_H = 125$ GeV and is found to be in good agreement with the results from the separate channels as compared in Fig. 7.35.

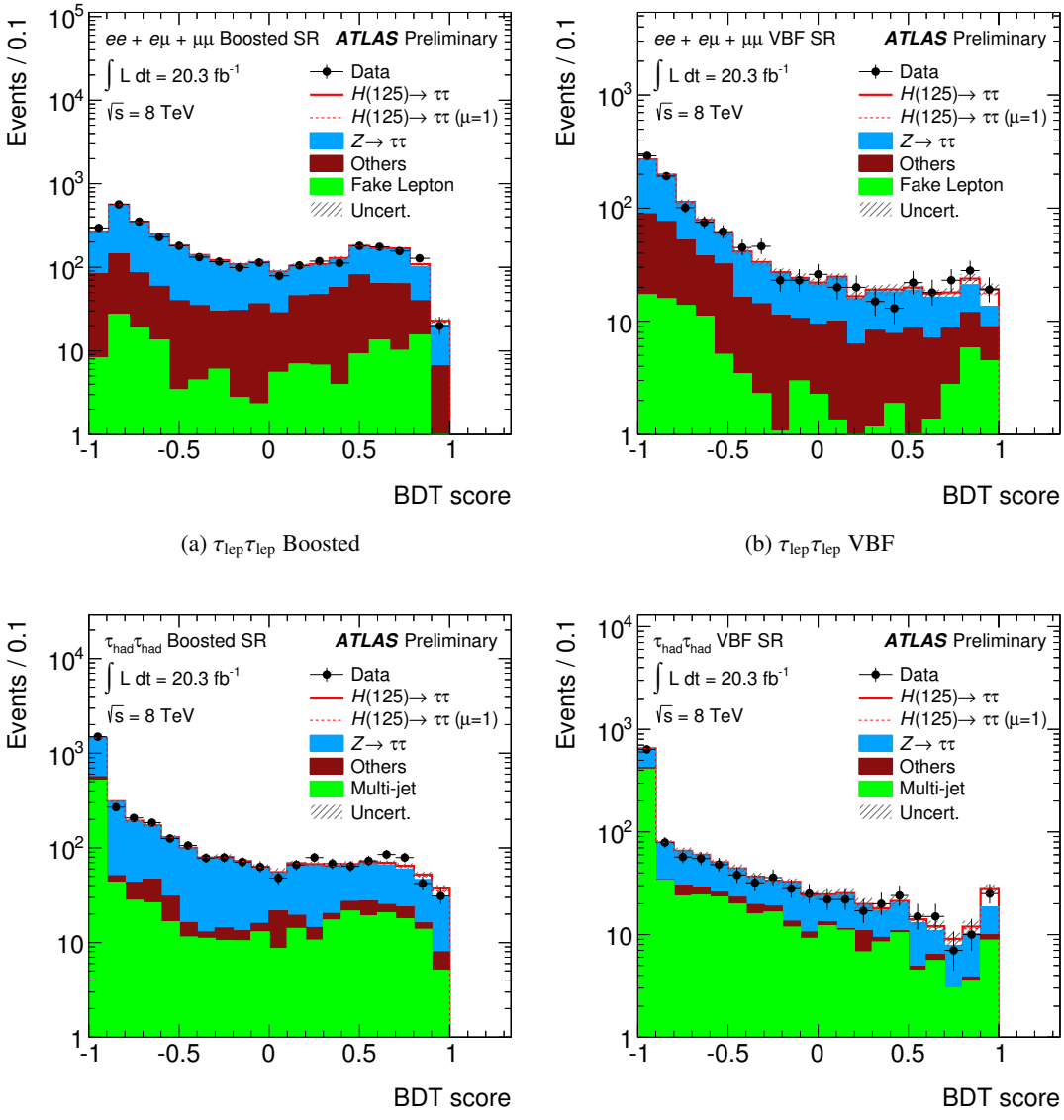
By splitting the signal into the different production modes, values of

$$\mu_{\text{ggF}} = 1.1^{+1.3}_{-1.0} \quad \text{and} \quad \mu_{\text{VBF}+VH} = 1.6^{+0.8}_{-0.7}$$

are obtained from the combination (cf. Fig. 7.34a), which again show that the VBF categories dominate both the sensitivity and the small excess above the expected signal strength.

Interpreting the combined result in terms of the profile likelihood (cf. Fig. 7.34b), a significance of 3.2σ is expected from the Asimov data. Combining the data in all $H \to \tau\tau$ final states, the observed excess over the background-only hypothesis is found to have a significance of 4.1σ, which confirms the evidence from the $H \to \tau_{\text{lep}}\tau_{\text{had}}$ channel. It is compatible within 0.9σ with a signal from a Standard Model Higgs boson with $m_H = 125$ GeV.

[20] All bins with small S/B ratios are merged into the lowest bin of the significance ordered distribution.



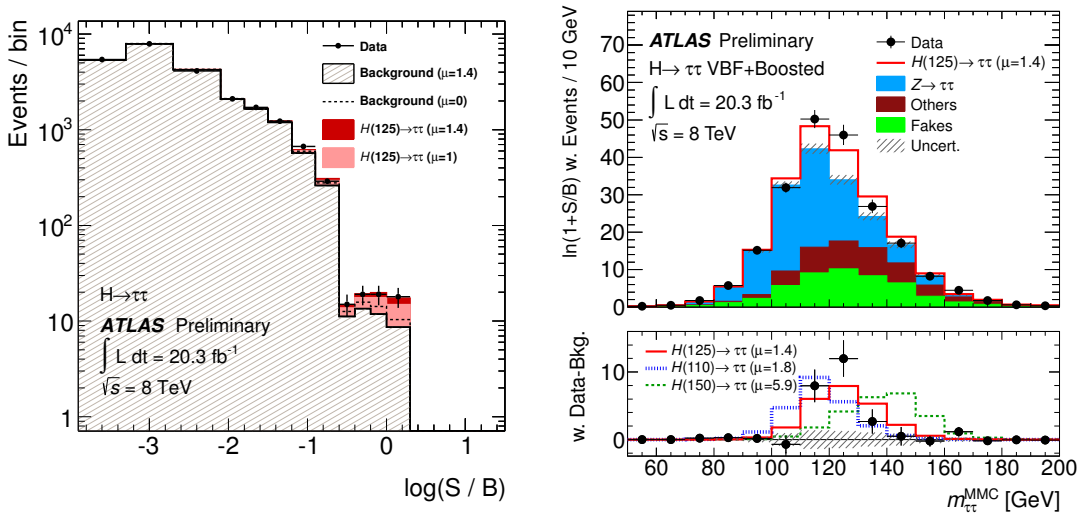
(a) $\tau_{\text{lep}}\tau_{\text{lep}}$ Boosted

(b) $\tau_{\text{lep}}\tau_{\text{lep}}$ VBF

(c) $\tau_{\text{had}}\tau_{\text{had}}$ Boosted

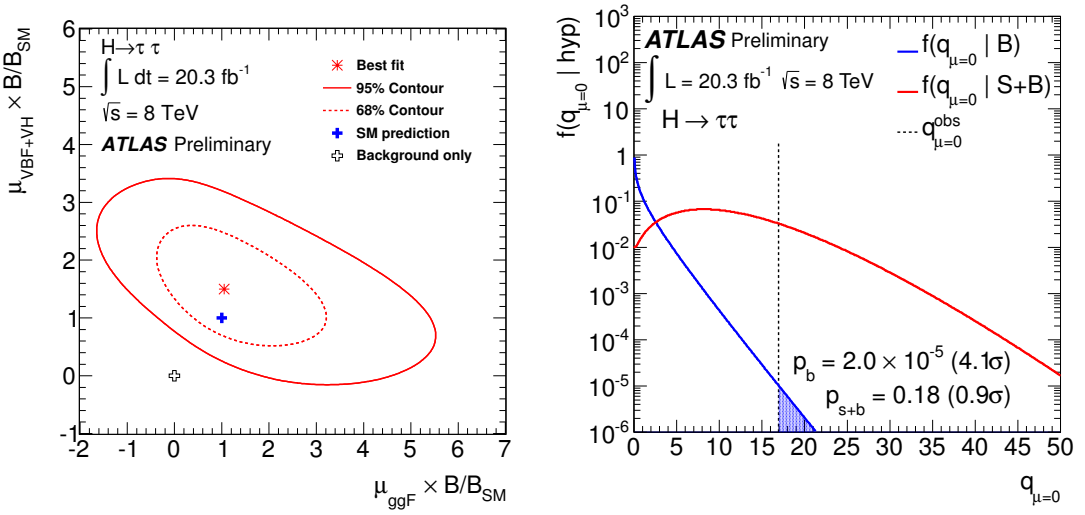
(d) $\tau_{\text{had}}\tau_{\text{had}}$ VBF

Figure 7.32: Final BDT score distributions of the Boosted and VBF categories obtained from the $\tau_{\text{lep}}\tau_{\text{lep}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ final states [171].



(a) (b)

Figure 7.33: (a) Summary of final binned BDT score distributions of all channels ($\tau_{\text{lep}}\tau_{\text{lep}}$, $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$), where the upper edges of the bins are set to their corresponding S/B ratio. (b) Sum of $\ln(1+S/B)$ weighted MMC mass distributions of all channels [171].



(a) (b)

Figure 7.34: (a) Contour for the separate signal strengths μ_{ggF} vs. $\mu_{\text{VBF}+VH}$ and (b) distribution of the test statistics q_0 for the combined fit of all channels in the asymptotic approximation [171].

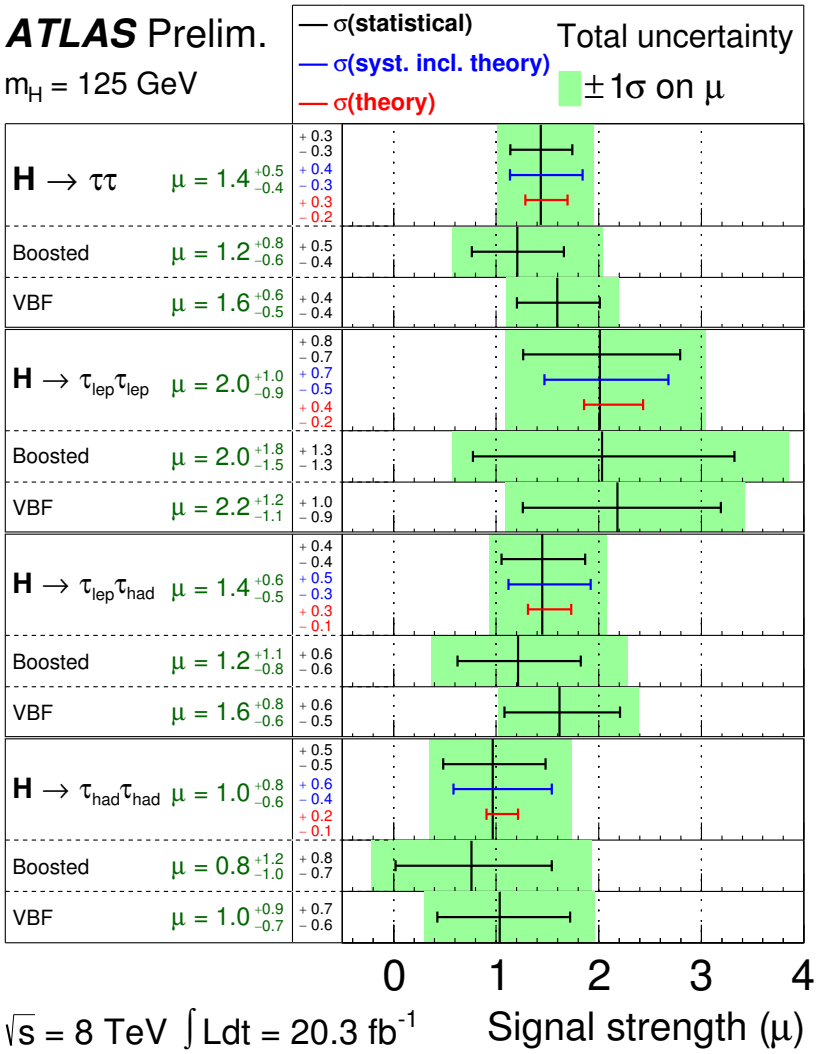


Figure 7.35: The best-fit value for the signal strength μ in the individual channels and the combination. The total $\pm 1\sigma$ uncertainty is indicated by the shaded green band with the individual contributions from the statistical uncertainty (top, black), the total (experimental and theoretical) systematic uncertainty (middle, blue) and the theory uncertainty (bottom, red) on the signal cross section (from QCD scale, PDF and branching ratios), which is shown by the error bars and printed in the central column [171].

CHAPTER 8

Studies Towards a Mass Measurement

The mass distribution is the most important observable in searches for $H \to \tau\tau$ decays, since it can provide a separation of the $H \to \tau\tau$ from the dominant $Z \to \tau\tau$ background. The observation of a resonance in the mass spectrum can thus be interpreted more easily than an excess in a BDT score bin, for which the phase space definition is less clear. The mass therefore plays a special role in the analysis and the scrutinisation of the result.

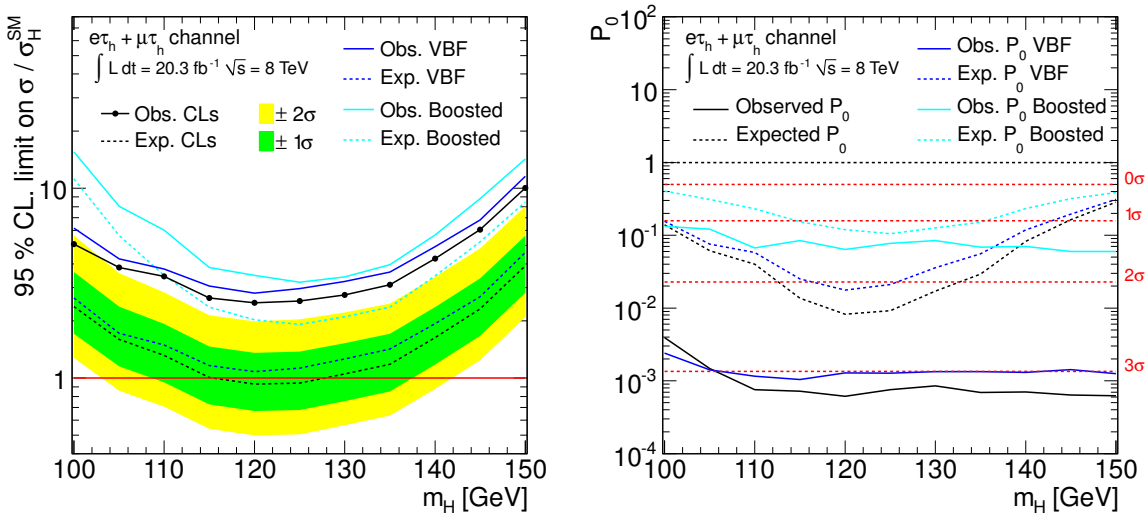
The weighted mass distributions as presented in Fig. 7.31, however, show strongly biased shapes with only a weak separation between signal and background, which is a result of the BDT training setup focusing exclusively on signal events with a mass of $m_H = 125$ GeV. A solid analysis of the statistical agreement in these distributions would require a new fit to provide a mass measurement. Unfortunately, such a fit of the weighted distributions would not be able to constrain systematic uncertainties and background normalisations sufficiently, since variations of the signal strength and background normalisations have very similar effects due to their small shape differences. Therefore, the mass of the observed excess is first investigated within the existing fit model, before alternative analysis strategies are presented, which are designed to extract the signal from unbiased mass distributions rather than from the BDT score.

8.1 Compatibility of the Excess with Different Higgs Boson Mass Hypotheses

Although the observed excess is found to be compatible with the expectation from $m_H = 125$ GeV, its interpretation in the context of signals with a different mass can provide additional information, which can support this observation. As an attempt to exclude such signals with different masses, the corresponding limits are calculated and presented in Fig. 8.1a in the region of 100 GeV $< m_H <$ 150 GeV. These, however, show a broad excess of observed exclusion over the expected limit, so that the observed mass cannot be constrained further with this method. Similarly, the dependence of the local significance and the p_0-value on m_H in Fig. 8.1b is found to be almost independent of m_H, so that no mass point can be uniquely identified to provide the best fit to the data.

Instead, the excess over the background-only expectation $p(q_{\mu=0}|B)$ can be interpreted as excess over the signal expectation $p(q_{\mu=0}|S + B)$, so that more information on the observed mass can be extracted from the p_1-value. As demonstrated in Fig. 8.2a, the best compatibility is obtained for $m_H = 125$ GeV, while for masses $m_H <$ 115 GeV and 130 GeV $< m_H$ the excess over the signal expectation is above the

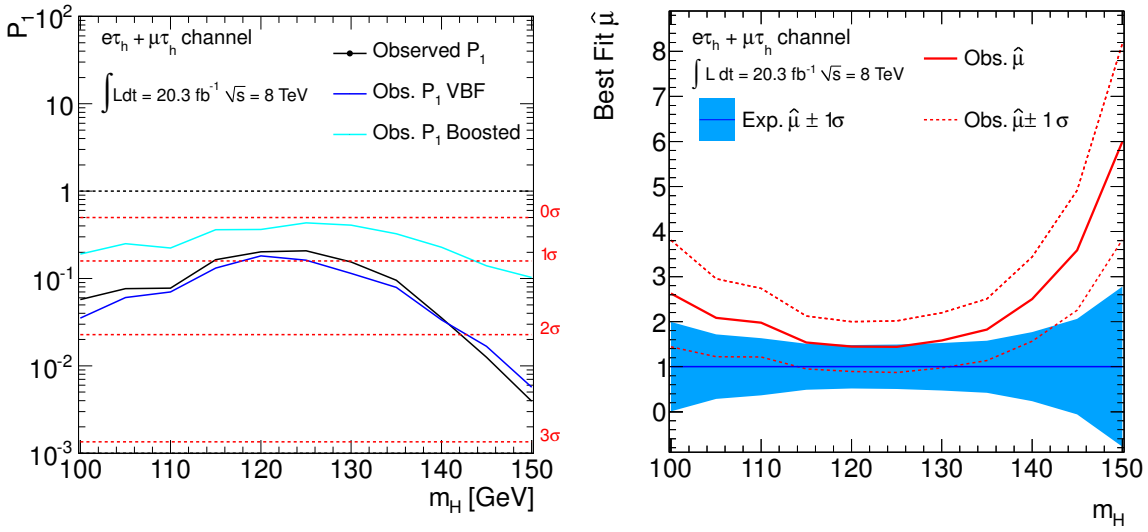
1σ-level. Nevertheless, only masses with 145 GeV $< m_H$ show a larger discrepancy above the 2σ-level. The same behaviour is reflected in the signal strength $\hat{\mu}$ in Fig. 8.2b, which is fitted with different signal mass templates: lower masses ($m_H < 115$ GeV) and higher masses (130 GeV $< m_H$) prefer a larger signal strength, which is less consistent with the expected $\mu = 1$ than the fitted value for $m_H = 125$ GeV.



(a) (b)

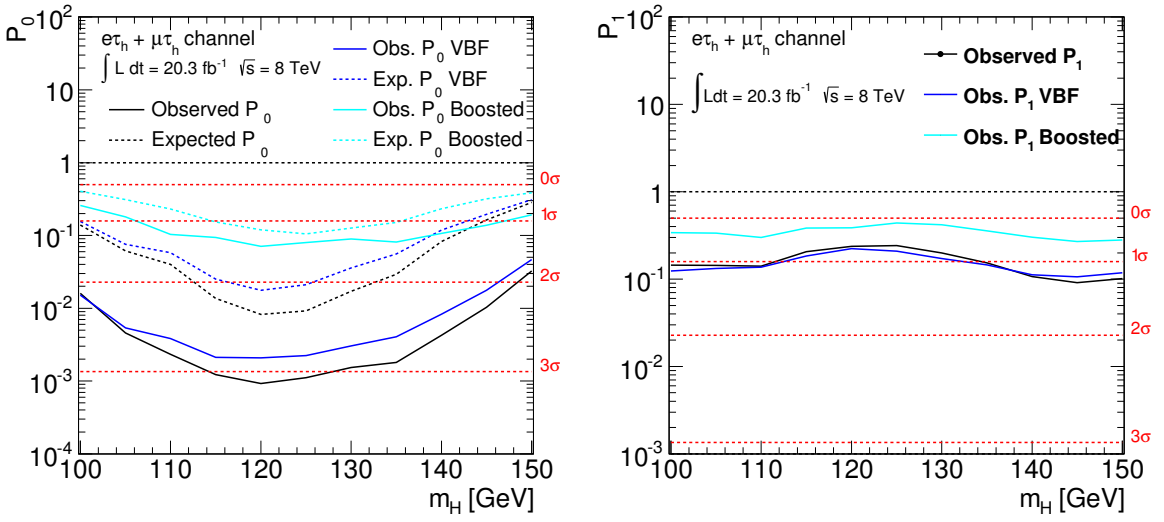
Figure 8.1: (a) Expected and observed exclusion limit on the cross section of a Standard Model Higgs boson and (b) expected and observed significance of the signal excess in dependence of the mass m_H of the Higgs boson signal expectation.

This weak mass dependence of the fit is a direct consequence of the analysis strategy: by including the MMC mass as an input observable for the BDT, which is then trained on the signal template for $m_H = 125$ GeV only, the information on the signal mass is largely lost. The final fit of the BDT score distributions is thus mainly sensitive to an excess in the highest bin, which corresponds to a mass window around the trained mass ($m_H = 125$ GeV). When fitting a signal template with $m_H = 150$ GeV, its small contribution in the selected mass window dominates the fitted signal strength, which prefers the signal contribution to be scaled by $\hat{\mu} = 6.0 \pm 2.2$. Although this factor increases the dominant signal contribution at $m_H = 150$ GeV as well, this high-mass region is considered less signal-like by the BDT and is therefore dominated by background contributions. The fit can then compensate for any signal contamination by adjusting the backgrounds and thus converges to the large fitted signal strength with a slightly worse fit quality. Only a fit with a signal strength constrained to the Standard Model expectation $\mu = 1$ would thus result in the expected behaviour and could provide a more reasonable mass measurement under the assumption that the observed excess over the background-only hypothesis is indeed a Standard Model Higgs boson. Although this procedure is not common practice, the result on the p_0 and p_1 distributions is visualised in Fig. 8.3: the mass dependence of the observed p_0 distribution in Fig. 8.3a is more pronounced around $m_H = 125$ GeV, while the compatibility p_1 of the data with the signal hypothesis is improved for other masses $m_H \gtrsim 125$ GeV (cf. Fig.8.3b) due to the constrained signal strength. From the $\chi^2 = -2 \ln \Delta L$ distributions shown in Fig. 8.4, the constrained fit allows the estimation of the mass to $m_H = (120^{+11}_{-7})$ GeV, while the same distribution from the unconstrained fit cannot provide any reasonable estimate on the mass.



(a) (b)

Figure 8.2: (a) Compatibility $p(s + b)$ of the observed data and (b) the expected and observed signal strength $\hat{\mu}$ (right) as interpreted with the expected signal from a Higgs boson with mass 100 GeV $< m_H <$ 150 GeV.



(a) (b)

Figure 8.3: (a) Expected and observed significance and (b) the compatibility $p(s + b)$ of the observed data for the constrained fits ($\mu = 1$) as interpreted with the expected signal from a Higgs boson with mass 100 GeV $< m_H <$ 150 GeV.

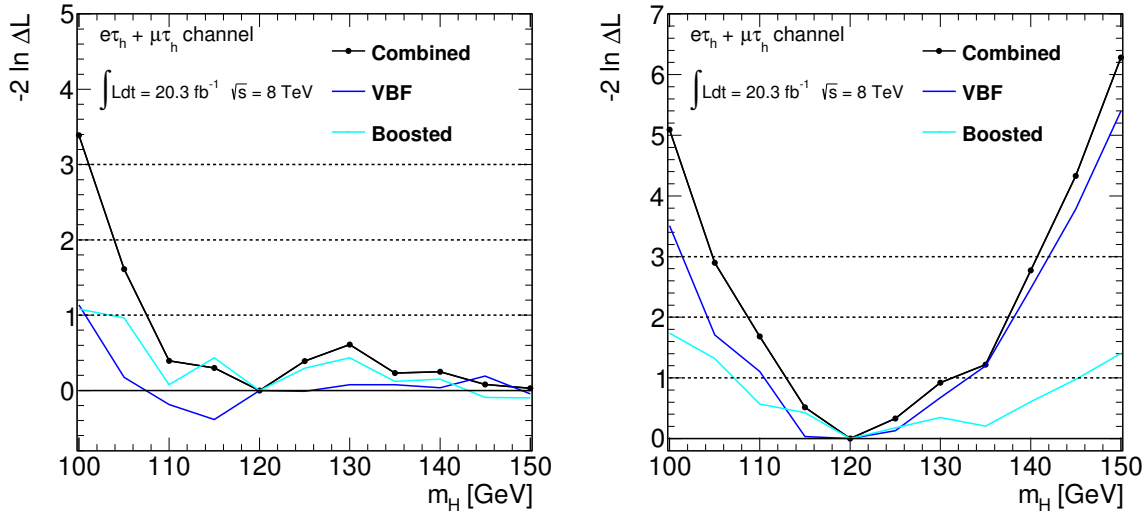


Figure 8.4: Distribution of $\chi^2 = -2 \ln \Delta L$ in dependence of the Higgs boson mass m_H (a) for the unconstrained fit and (b) for the constrained fit ($\mu = 1$).

8.2 Alternative Strategies for Mass Measurements

The BDT score distributions of the default analysis are found to be highly correlated with the MMC mass provided in the training, so that the BDT-weighted MMC mass distributions are strongly biased. This becomes especially obvious for the $Z \to \tau\tau$ background, which peaks near the expected signal mass $m_H = 125$ GeV rather than at the original mass $m_Z = 91$ GeV of the Z boson. In order to obtain a mass measurement from MMC distributions, methods to de-correlate the BDTs from the mass are investigated in the following. Different analysis strategies are presented, which are based on these re-designed BDTs. In these analyses, the expected signal sensitivities are extracted from the MMC mass distribution, which automatically provides a better sensitivity to the mass of the Higgs boson. Since any bias from the observation in the data should be avoided, these improvement studies are performed only on the expected background and signal contributions, while the data in the sensitive regions are blinded. They are replaced by Asimov data, which are constructed by adding the signal expectation for $m_H = 125$ GeV to the expected background.

8.2.1 Decorrelation of BDT and MMC

A very simple strategy to de-correlate the BDT from the MMC mass is to exclude this observable from the list of input observables for the training. With the newly trained BDTs, new signal regions are defined by cuts on the BDT score, in which the MMC distributions are then fitted as shown in Fig. 8.5. By scanning this cut-value and calculating the expected significance, the optimal cuts are determined to be 0.96 for the VBF category and 0.60 for the Boosted category. As a result, however, the optimal sensitivity obtained from the MMC distributions after these cuts is expected to be 1.5σ for $m_H = 125$ GeV as compared to 2.3σ from the default analysis by fitting the BDT score trained on the mass. Although both mass distributions for the Boosted and VBF category show a better separation of signal and background than the distributions weighted according to the sensitivity of the default BDT

(cf. Fig. 7.31), the combined expected significance is largely reduced from 2.3σ to 1.5σ (1.3σ to 0.8σ for Boosted, 2.0σ to 1.3σ for VBF). While the mass information is exploited by both the BDT-fit and the mass-fit approach, the default BDT can use this information in a more sophisticated way, and thus yield a higher sensitivity.

The selection of observables, however, which is found to be optimal for the default BDT, is not necessarily the best choice when excluding the MMC from the training. Since certain observables such as the distance $\Delta R(\ell, \tau_{\mathrm{h}})$ are strongly correlated with the mass, a dedicated optimisation of the training is therefore performed in the following in order to identify the best performing set of observables for the mass-fit approach.

As another alternative to the simple BDT-cut and MMC-fit approach, the BDT training is also split into two parts: instead of training a single BDT to separate the signal from all backgrounds at the same time, two BDTs are trained with different backgrounds. While the first of the two is trained against all fake-τ_{h} backgrounds excluding the dominant $Z \to \tau\tau$ contribution, the second BDT is only trained against the resonant $Z \to \tau\tau$ background. With this method, the first BDT can be specialised on the separation of signal and fake-τ_{h} backgrounds, since it is not combined with the dominant background from a well defined mass resonance. The Z/H-separation can then be driven by other differences in spin-sensitive observables as well as jet-related observables sensitive to the VBF production mechanism. As in the single-BDT analysis, the cut-values can be scanned on both these BDT scores, so that signal regions for the mass-fit are defined in a similar way with this two-BDT approach.

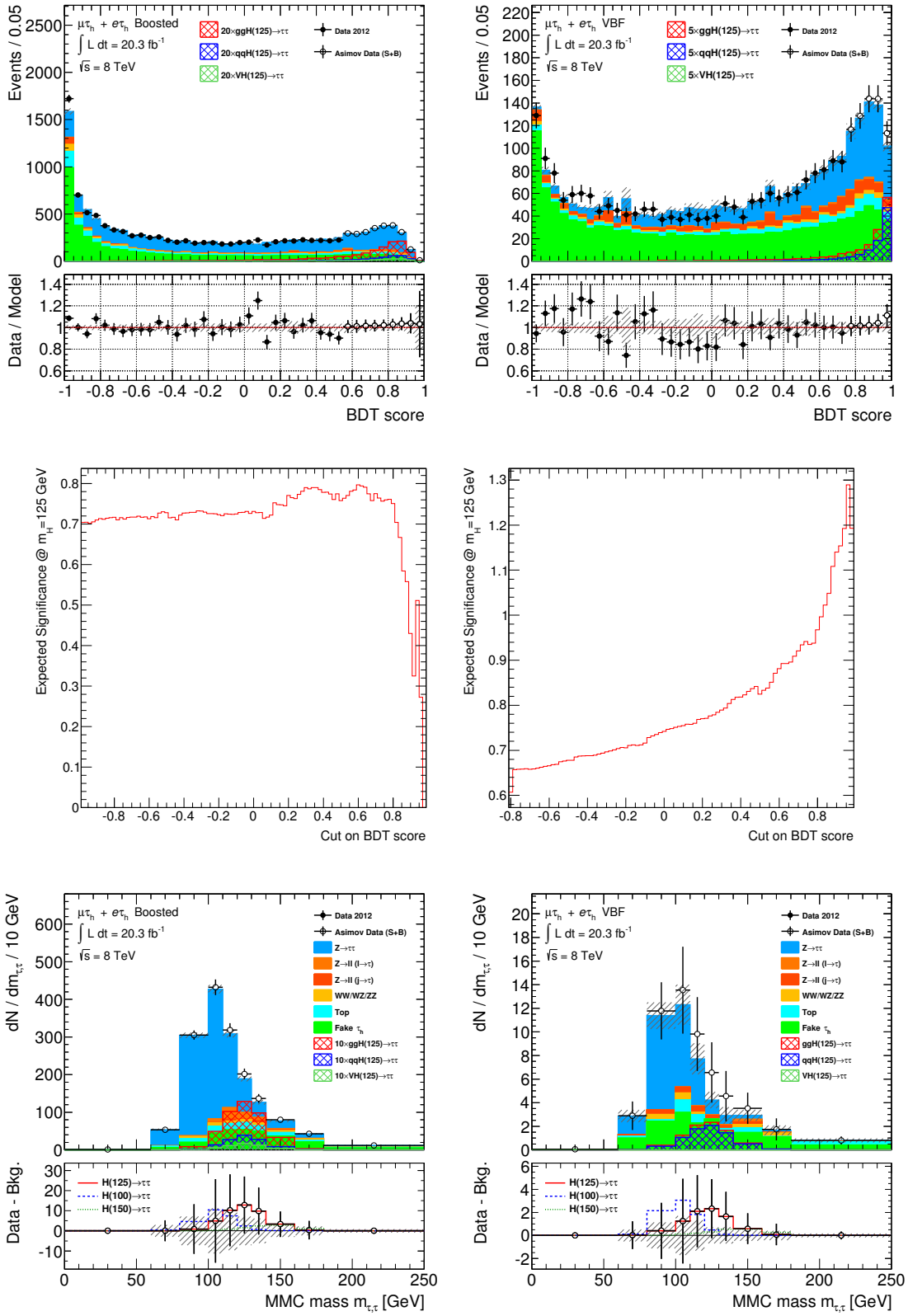


Figure 8.5: Distributions of the default BDT scores trained without MMC (top), scans of the expected significance versus the cut-value (middle) and MMC distributions after the optimal cut (bottom) for the Boosted (left) and VBF category (right). The uncertainty band only represents the statistical uncertainty of the background estimates.

8.2.2 Optimisation of the Variable Selection

The BDTs for the default analysis, which include the MMC mass as a training observable, have been optimised by comparing the ROC-integral obtained from trainings with different sets of observables. Without including the mass information implicitly into the BDT, however, this method cannot provide good results for an analysis that relies on fitting the MMC mass after a cut on the BDT score. In order to improve the weak performance of the default BDT trained without the MMC mass, a new optimisation is therefore performed by comparing the best statistical significance, which can be obtained from the MMC distribution after cutting on the BDT score from the tested training setup[1].

Starting from a large list of observables excluding the MMC mass, a BDT with four-fold cross-evaluation[2] is trained similarly as described for the default analysis. By scanning the cut-value on the resulting BDT score, MMC distributions for signal and background are created, and for each distribution the Poisson significance

$$s = \sum_{\text{bins}} \ln\left(1 + \frac{S_i}{B_i}\right)$$

is calculated from the signal (background) events S_i (B_i), while the MMC binning is adapted such, that the signal and background expectation in each bin is larger than one event. The best significance from all MMC distributions obtained by cutting on the BDT score is stored as a measure of its performance.

As for the optimisation of the default BDTs, the importance of each input observable v is then tested by dropping it from the list L_N of training observables and evaluating the difference in performance of $s(L_N \backslash \{v\})$ obtained with the reduced set as compared to the full training $s(L)$ with all observables $v \in L_N$. In order to quantify an effect of a variable on the training performance, the difference

$$d_0(v) = \frac{s(L_N) - s(L_N \backslash v)}{s(L_N)}$$

is defined, which compares the significance s obtained with the full list L_N to the reduced list $L_N \backslash \{v\}$. In addition to this simple difference, also two further differences are calculated in order to account for possible correlations of the observables:

$$d_1(v) = \frac{s(L_N \backslash \{v\}) - \langle s(L_{N-1}) \rangle}{\langle s(L_{N-1}} \qquad d_2(v) = \frac{\langle s(L_{N-1} \backslash \{v\} \rangle - \langle s(L_{N-2}) \rangle}{\langle s(L_{N-2}) \rangle}$$

These values compare the performances based on average significances: a positive $d_1(v)$ indicates that dropping observable v results in a better performance than the average $\langle s(L_{N-1}) \rangle$ over all trees, which are trained after dropping each observable from the list L_N. Similarly, the second value $d_2(v)$ compares the average performance $\langle s(L_{N-1} \backslash \{v\} \rangle$ of all trees trained after dropping v and one additional observable to the average performance $\langle s(L_{N-2}) \rangle$ of dropping any combination of two observables. Only if an observable shows a consistently low importance for the BDT performance

$$d_1(v) > 3\% \quad \text{and} \quad d_2(v) > 3\% \,,$$

and the overall drop of significance $d_0(v) < 5\%$ is not too large, it is removed from the training list L. The procedure is then repeated with the reduced list $L' := L \backslash \{v\}$ until all remaining observables show a sufficiently high importance.

[1] The background and signal compositions for the training are identical to the setup applied for the default BDTs.

[2] In fact, for each set of input observables a set of four BDTs is trained with identical setup but on different sets of events. For simplicity, this is always referred to as a single BDT training (setup).

For the training of the two separate BDTs, the same optimisation procedure is applied, while the only difference in the training setup is the selection of samples: the *resonance-BDT* is trained only against the $Z \to \tau\tau$ background process, while for the *fake-BDT* all other processes except for $Z \to \tau\tau$ are included as background. The automated optimisation procedure is performed separately on the two different BDT types, since a simultaneous evaluation of different BDTs on all signal and background samples is not yet implemented in the TMVA and analysis framework[3]. The resulting observable lists are shown in Tab. 8.1 for the Boosted category. For the VBF category, this optimisation results in the fact that a large number of observables is found to be important for both BDT types at the same time, and that as a consequence more correlations enter the training. In order to prevent effects of mis-modelled correlations especially in the fake-BDT, the observable lists for the fake- and resonance-training are separated by hand into $\tau\tau$-related and jet-related observables, so that a set of more robust and de-correlated fake- and resonance-BDT is created from the automatically optimised versions. Table 8.2 summarises these observables used in the VBF category, which contains both the optimised list of observables and the de-correlated lists. This corresponds to a $\tau\tau$-BDT separating between fake-backgrounds and $\tau\tau$ final states and a VBF-BDT separating between $Z \to \tau\tau$ and $H \to \tau\tau$ decays based only on the di-jet kinematics.

8.2.3 Optimisation of Training Parameters

In addition to the observable selection, the final shape of the BDT score distribution is important for the performance of the analysis. While the fit to the BDT score distribution requires only a good separation of signal and background in the highest bins, the situation is very different if the MMC mass is fitted after cutting on this BDT score, which can potentially be correlated with the mass. Therefore, after identifying the optimised lists of training observables, a further optimisation of BDT score shape via the training parameters is performed for each new BDT with the following algorithm:

Starting from the default set of parameters $N_{\text{Trees}} = 1000$, `Depth`$= 4$ and `shrinkage`$= 0.3$, a grid of BDTs with the steps $\Delta N_{\text{Trees}} = \pm 200$, Δ`Depth`$= \pm 1$ and Δ`shrinkage`$= \pm 0.20$ is created around the starting point, which corresponds to 26 points on the surfaces of a three-dimensional cube in the parameter space. The point of the best BDT performance is then selected, and a new cube is created around this point, for which many points can be re-used from the first cube, depending on the direction of the parameter change. This procedure is iterated until the centre of the cube is the point of the best BDT performance, in which case the procedure is repeated with reduced step-sizes of $\Delta N_{\text{Trees}} = \pm 100(50)$ and Δ`shrinkage`$= \pm 0.10(0.05)$. The endpoints of this step-wise grid-scan are afterwards used as training parameters for the corresponding BDTs in the final analyses and are summarised in Tab. 8.3[4].

8.2.4 Mass-Fit Results

With the new BDTs, the different analyses are performed similarly to the default analysis, but with a final fit to the resulting MMC distribution obtained by cutting on the BDT score instead of fitting the BDT score distribution directly. Otherwise, all systematics and the fit model are treated in the same way as described for the default result. With this full systematic treatment, the cut-value on the BDT score is again scanned by calculating the expected sensitivity, so that the final cuts do not necessarily coincide with the values, which yielded the optimal performance in the training optimisation. This is

[3] Especially for the evaluation of the significance from the final MMC distribution, a simultaneous optimisation would be desirable.

[4] As for the default BDTs, the number of cuts $N_{\text{cuts}} = 20$ was kept constant, since no effect was observed from variations within $N_{\text{cuts}} \in [20, 50]$, while large numbers resulted in very large computing times for the BDT training. Each node was required to contain $N^{\text{min}}_{\text{events}}/N^{\text{total}}_{\text{events}} \geq 5\%$ events in order to ensure statistical significance of their separation power.

Variable	$Z \to \tau\tau \oplus$ Fake	Fake	$Z \to \tau\tau$
$m_T(\ell)$	•	•	•
$C^{\phi}_{\ell,\tau_h}(E_T^{miss})$	•	•	•
$p_T(\ell)/p_T(\tau_h)$	•	•	•
$\sum p_T$	•	•	•
$m(\tau_h, \ell, j_1)$	•	•	•

Table 8.1: Discriminating observables used for the training of the alternative BDTs for the Boosted category. The filled circles identify the observables used in each training setup.

Variable	$Z \to \tau\tau \oplus$ Fake	Fake	$Z \to \tau\tau$	Fake de-corr.	$Z \to \tau\tau$ de-corr.
$m_T(\ell)$		•	•	•	
$m_T(\tau_h)$		•		•	
$\sum \Delta\phi$	•	•		•	
$C^{\phi}_{\ell,\tau_h}(E_T^{miss})$			•		
$\Delta R_{\tau\ell}$	•				
$p_T(\ell)/p_T(\tau_h)$	•				
$p_T(\tau_h) - p_T(\ell)$	•	•		•	
$x_{\tau_h} \cdot x_\ell$	•		•		
$x_{\tau_h} - x_\ell$	•	•		•	
$p_T(H)$		•		•	
p_T^{tot}	•	•	•		•
$\Delta\phi(j_1, j_2)$	•	•	•		•
$m(j_1, j_2)$	•		•		•
$\eta(j_1) \times \eta(j_2)$	•		•		•
$\lvert\eta(j_2) - \eta_(j_1)\rvert$	•		•		•
$C^{\eta}_{j_1,j_2}(\ell) \cdot C^{\eta}_{j_1,j_2}(\tau_h)$	•		•		•

Table 8.2: Discriminating observables used for the training of the alternative BDTs for the VBF category. The filled circles identify the observables used in each training setup.

	Boosted			VBF		
Parameter	$Z \to \tau\tau \oplus$ Fake	Fake	$Z \to \tau\tau$	$Z \to \tau\tau \oplus$ Fake	Fake	$Z \to \tau\tau$
N_{trees}	1000	800	1200	1200	200	800
N^{max}_{nodes}	3	4	3	5	4	4
`Shrinkage`	0.20	0.30	0.10	0.50	0.35	0.30

Table 8.3: Optimal training parameters used for the alternative analysis BDTs.

especially important for the two-BDT approach, for which the optimisation of resonance-BDT and fake-BDT was performed separately by neglecting their effects on the other background components or the resulting product of the two signal efficiencies. An additional criterion for the selection of cut-values is the robustness of the result, which requires that dominant background contributions are not depleted, leading to unreasonably high significances due to statistical fluctuations in the background model. The final results with optimal but robust performance for the different analysis strategies are presented in the following.

Single-BDT Approach

After the optimisation of the BDT training observables and parameters, the BDTs trained simultaneously against $Z \to \tau\tau$ and fake backgrounds show only small changes in the shape of their score distributions in Fig. 8.6 as compared to the default BDTs trained without the MMC mass in Fig. 8.5. The scan of the expected significance for different cut-values on the BDT scores, however, demonstrates that the performance of the BDTs in combination with an MMC-fit is indeed improved due to the optimisation of their observable selection and training parameters. With the optimal cut-values on the BDT score at 0.26 for Boosted and at 0.96 for VBF, the separation of signal and background is significantly improved in the final MMC distributions in Fig. 8.6, which results in increased expected significances of 1.0σ for the Boosted category and 1.6σ for the VBF category. In combination, a sensitivity of 1.8σ is expected for $m_H = 125$ GeV, which is compared to the default BDTs without MMC in Fig. 8.7. As illustrated by the injection of a corresponding signal with $m_H = 125$ GeV, also the sensitivity to the Higgs mass m_H is improved as compared to the default analysis, so that the best fit to an observed signal can be identified with a higher precision. Unfortunately, the overall expected sensitivity can still not reach the performance of the default result from the BDT-fit (2.3σ), so that the overall performance of this single-BDT strategy is found to provide no real improvement of the analysis, but only a gain in mass separation for a less significant excess of signal over the background expectation.

Two-BDT Approach

For the two-BDT approach, the trained fake- and resonance-BDT score distributions are shown in Fig. 8.8. The expected significances from the MMC-fit are obtained by scanning a two-dimensional grid of cut-values on the fake- and resonance-BDT (cf. Fig. 8.9) and fitting the resulting MMC distribution for each point in the same way as in the single-BDT approach. From this scan for the Boosted category, an optimal expected significance of 1.1σ is found with cuts at $(0.46, -0.18)$, which result in the MMC mass distribution in Fig. 8.10a. Similarly, a significance of 1.9σ is obtained for the VBF category with cuts at $(0.70, 0.90)$ (cf. Fig. 8.10a). The combination of these categories yields an expected significance of 2.1σ for a signal from $H \to \tau\tau$ decays with $m_H = 125$ GeV.

While this result is reasonably close to the result of the default analysis, which expects a significance of 2.3σ, the improvement of this alternative method becomes obvious when comparing the dependence of the sensitivity on the mass m_H as illustrated in Fig. 8.11a: with a similar overall sensitivity as the default analysis, the mass of an injected signal with $m_H = 125$ GeV can be recovered from the best fit, so that the sensitivity to the Higgs mass is significantly increased. Although other mass points can still not to excluded with this two-BDT analysis, the difference of the fitted likelihood $-2\ln\Delta L$ in Fig. 8.11b can be interpreted as a measurement of the mass m_H. With a simple extrapolation from the given mass points, a measured value of $m_H = 125^{+11}_{-9}$ GeV can be extracted from the injected signal.

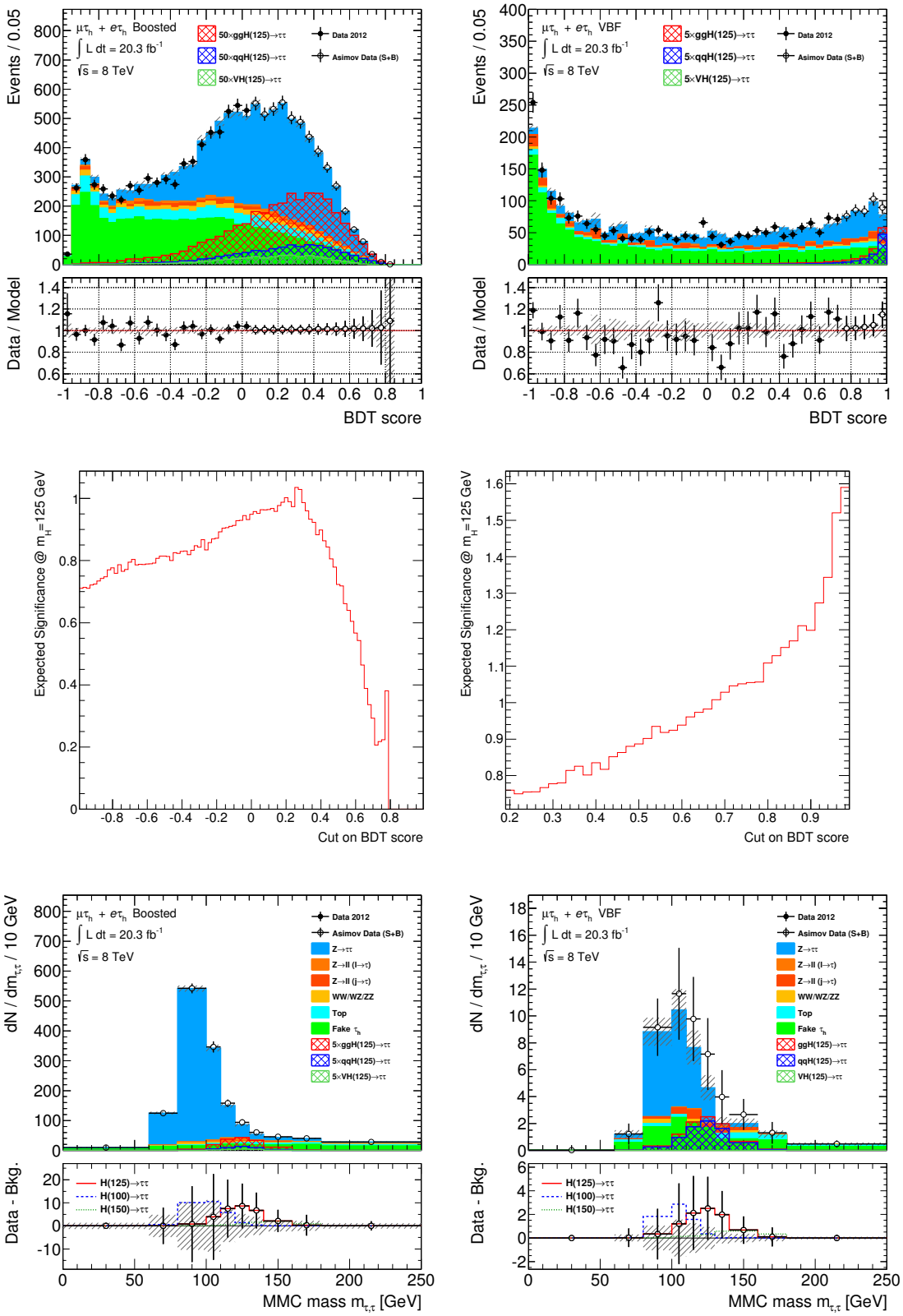


Figure 8.6: Distributions of the optimised BDT scores (top), scans of the expected significance of the MMC-fit versus the cut-value (middle) and MMC distributions after the optimal cut (bottom) for the Boosted (left) and VBF category (right). The uncertainty band represents only the statistical uncertainty of the background estimates.

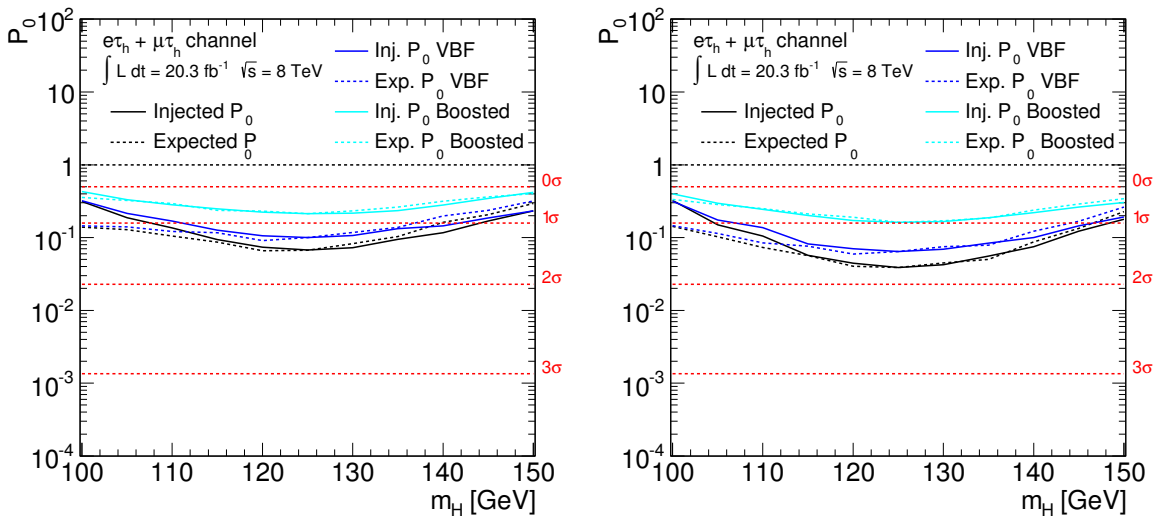
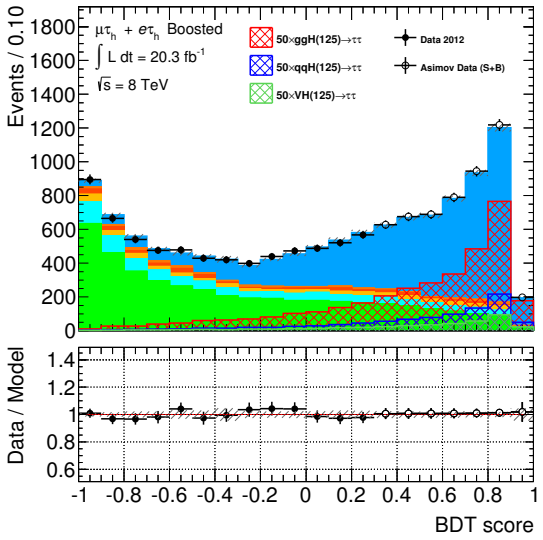
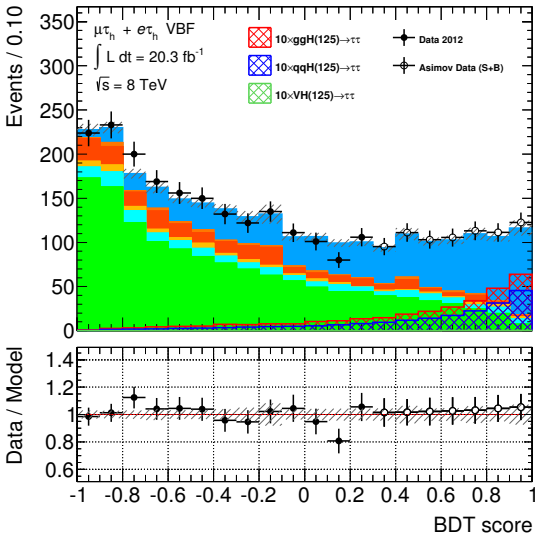


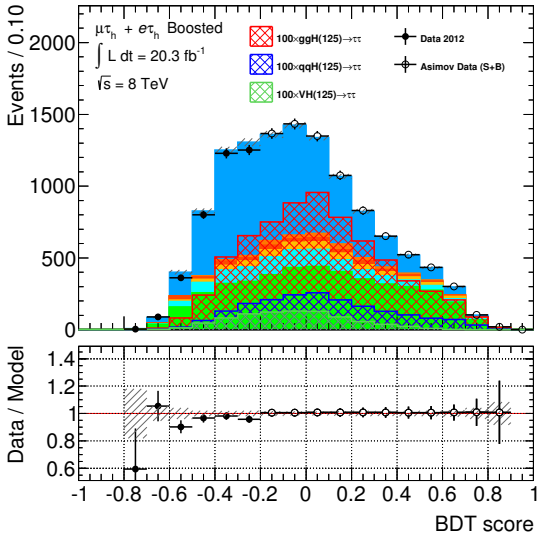
Figure 8.7: Expected and injected ($m_H = 125$ GeV) significance of the signal from a Standard Model Higgs boson in dependence of the mass m_H as obtained from fits to the MMC mass after cutting on the default BDTs trained without the MMC mass (left) and after cutting on the optimised BDTs (right). The uncertainty band only represents the statistical uncertainty of the background estimates.



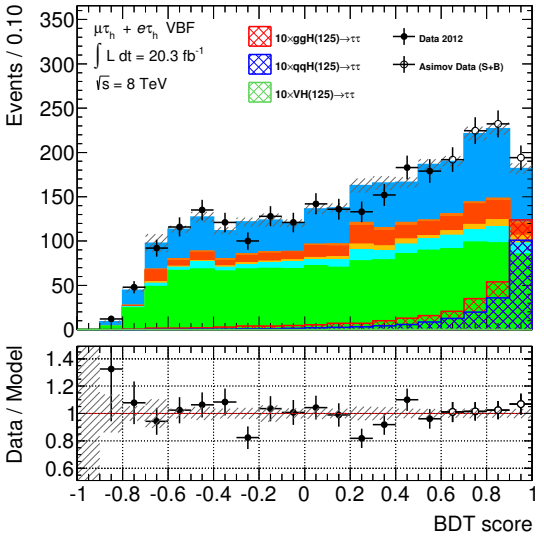
(a) Fake-BDT Boosted



(b) Fake-BDT VBF

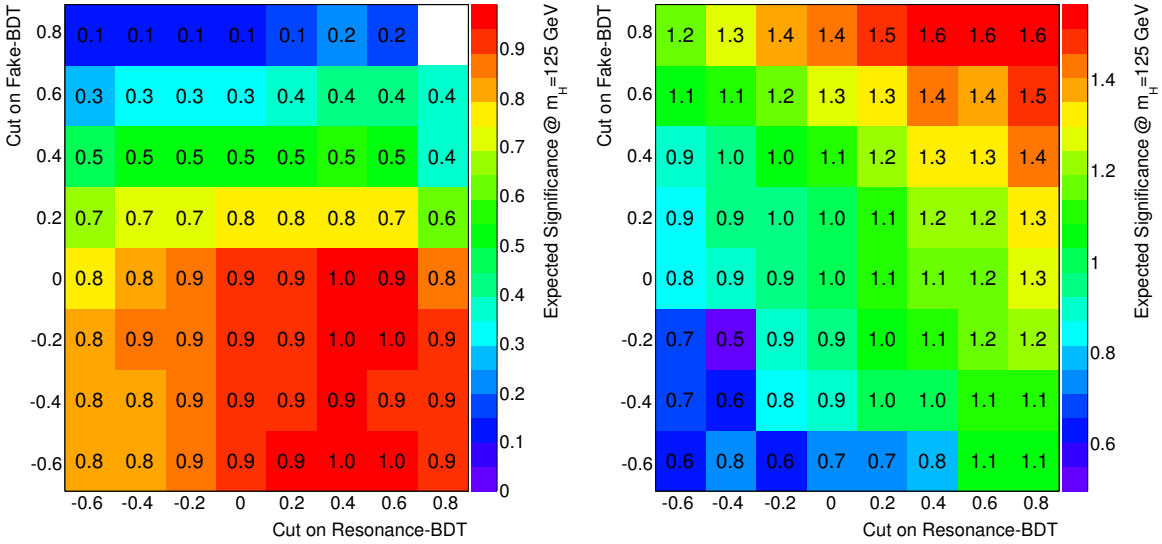


(c) Resonance-BDT Boosted



(d) Resonance-BDT VBF

Figure 8.8: Distributions of the fake-BDT and the resonance-BDT score for the Boosted and VBF category. The uncertainty band only represents the statistical uncertainty of the background estimates.



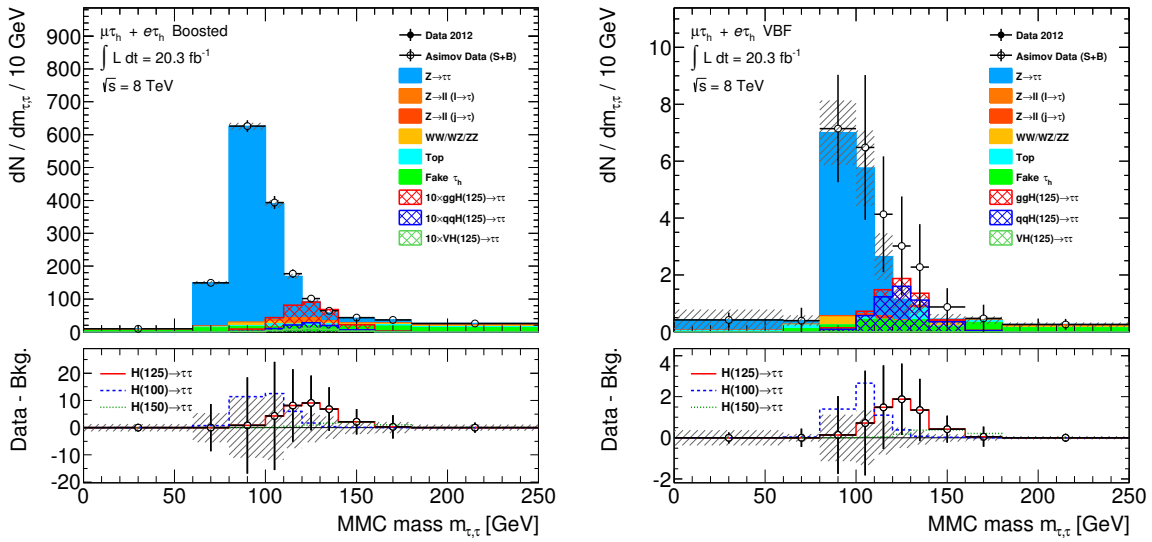
(a) Coarse Scan Boosted

(b) Coarse Scan VBF

(c) Fine Scan Boosted

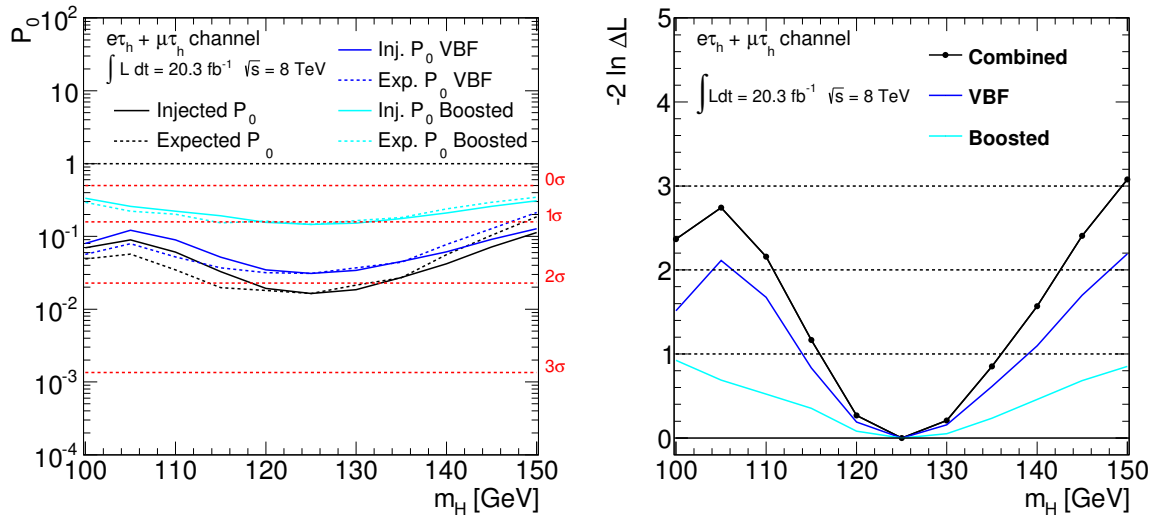
(d) Fine Scan VBF

Figure 8.9: Coarse and fine scans of the expected significance of the MMC-fit versus the cut-values on the fake- and resonance-BDT scores for the Boosted and VBF category.



(a) Boosted Category (b) VBF Category

Figure 8.10: Distributions of the MMC mass after the optimised cuts in (a) the Boosted and (b) the VBF category. The uncertainty band only represents the statistical uncertainty of the background estimates.



(a) (b)

Figure 8.11: (a) Expected and pseudo-observed significance and (b) distribution of $\chi^2 = -2\ln\Delta L$ distribution for the fit with different mass points as obtained from the two-BDT analysis for an injected signal with $m_H = 125$ GeV.

Category	Mass-Fit Results			BDT-Fit Result
Strategy	Default BDT excl. MMC	Opt. Single-BDT	Opt. Two-BDT	Default BDT
Boosted	0.8σ	1.0σ	1.1σ	1.3σ
VBF	1.3σ	1.6σ	1.9σ	2.0σ
Combined	1.5σ	1.8σ	2.1σ	2.3σ

Table 8.4: Expected sensitivities for the different alternative mass-fit analysis strategies in comparison to the results from the default BDT-fit analysis.

8.2.5 Summary of the Mass Studies

While the default BDT-fit analysis shows a high sensitivity due to the training on the MMC mass as an input observable, it cannot provide a direct mass measurement, since other Higgs boson masses cannot be excluded. When constraining the signal strength to $\mu = 1$, however, the mass can be constrained to

$$m_H = 120^{+11}_{-7}\ \text{GeV}$$

from the $\tau_{\text{lep}}\tau_{\text{had}}$ channel only. In order to obtain a measurement directly from the mass distribution, alternative analysis strategies are presented, which de-correlate the BDTs from the MMC mass. With an optimised selection of training observables and parameters, expected sensitivities are extracted from fits to the mass distributions, which are compared to the default result in Tab. 8.4. With an injected signal of $m_H = 125$ GeV, the best performing alternative analysis of $\tau_{\text{lep}}\tau_{\text{had}}$ final states constrains the pseudo-observed mass to

$$m_H = 125^{+11}_{-9}\ \text{GeV}\ .$$

Although more careful checks of the background modelling as well as the fit model of these alternative analyses are necessary before unblinding the observed data, the results of these mass-fit studies show that a more sophisticated analysis strategy can yield a comparable result to the default analysis with a better sensitivity to the mass of the $\tau\tau$ resonance. Especially for the VBF category, this study demonstrates that the dedicated training of BDTs separated for special classes of backgrounds can significantly improve the performance as compared to a single BDT. While the topologies of all backgrounds in the Boosted category are rather similar due to the requirement of large $p_T(H)$, this method can also provide improvements for the background suppression in the Rest and 1-jet category, which are not studied here. The resulting clean selections of $\tau^+\tau^-$ final states with an unbiased mass distribution can eventually be used as a basis for a mass measurement, but also for further studies on the spin correlation of the τ-leptons such as spin- or CP-measurements of the observed resonance.

CHAPTER 9

Conclusions

After the development of analysis strategies and background estimation techniques with Monte Carlo simulations documented in [144, 145], the search for $H \to \tau\tau$ final states has been performed with the ATLAS data collected during the first run of the LHC and is presented in this thesis.

Starting with the cross section measurement of the most important $Z \to \tau\tau$ background process in the first dataset recorded with ATLAS in 2010 [140], the modelling of $\tau\tau$ final states in simulated events could be validated. A result of

$$\sigma = 0.97 \pm 0.4\,(\text{stat}) \pm 0.3\,(\text{syst}) \pm 0.2\,(\text{lumi})\ \text{nb}$$

has been obtained, which is consistent with the expectation from theory and measurements in other $Z \to \ell\ell$ final states.

For the $H \to \tau\tau$ search, the prediction of the $Z \to \tau\tau$ background with the embedding technique has been applied to real collision data for the first time after its development on simulated $Z \to \mu\mu$ Monte Carlo samples. It could be established as standard background estimate for all $H \to \tau\tau$ final states and is a key ingredient especially for multi-dimensional analyses, which strongly rely on correlations of jet-related observables. Even for performance studies of hadronic τ decay objects in the detector, embedded events are widely used in order to evaluate the impact of pile-up in data on the τ_{h} identification and energy calibration.

With all background estimation techniques developed in the $Z \to \tau\tau$ analysis, an intermediate result for $H \to \tau\tau$ final states has been published on the dataset recorded by ATLAS in 2011 [153], which excluded an enhanced coupling of the Higgs boson to τ-leptons. Further invested effort led to a continuous development and refinement of the analysis and background estimation techniques [164], which eventually resulted in a significant improvement of the sensitivity as compared to expectations from earlier simulation studies (cf. [145]) with both a higher centre-of-mass energy ($\sqrt{s} = 14$ TeV) and a larger dataset (30fb^{-1})[1].

[1] For the ATLAS data taken at $\sqrt{s} = 8$ TeV in 2012, the instantaneous luminosity and thus the pile-up situation was comparable to the Monte Carlo studies based on design luminosity at $\sqrt{s} = 14$ TeV.

As the final result of this thesis, an expected sensitivity of 2.3σ could be achieved by employing a multi-variate analysis of $H \to \tau_{\text{lep}}\tau_{\text{had}}$ final states in 20.3 fb^{-1} of data recorded with ATLAS in 2012, which in combination with the corresponding searches in the $H \to \tau_{\text{lep}}\tau_{\text{lep}}$ and $H \to \tau_{\text{had}}\tau_{\text{had}}$ channels expected an evidence of 3.2σ for a SM Higgs boson with $m_H = 125$ GeV decaying to $\tau\tau$ final states. With the data observed in the corresponding BDT score distributions, this expected evidence was then confirmed by an observed excess over the background-only hypothesis of 3.2σ for the $H \to \tau_{\text{lep}}\tau_{\text{had}}$ channel alone and 4.1σ for the $H \to \tau\tau$ combination [171]. After completing this thesis, a combination of this result with the 4.6 fb^{-1} of data recorded in 2011 was submitted for final publication [182] with only small modifications to the analysis presented here.

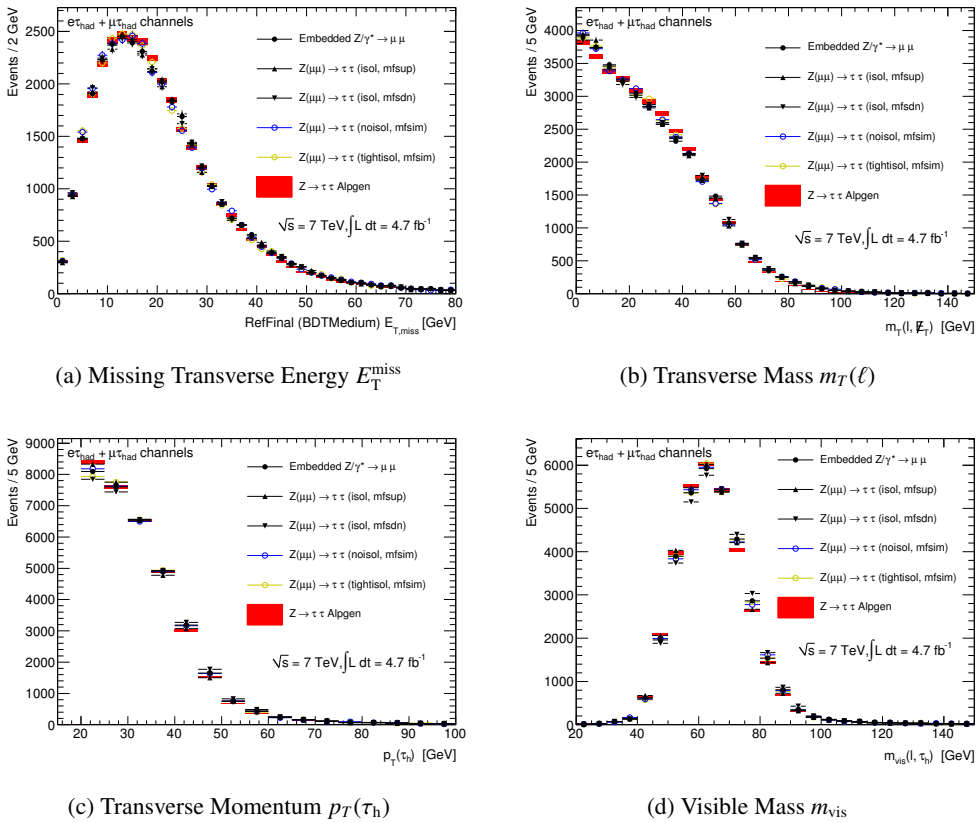
Together with the observed (expected) evidence of 3.2σ (3.7σ) in $H \to \tau\tau$ decays published by the CMS collaboration [183], this confirms the fermion coupling of the Higgs-like particle discovered in 2012 in vector boson final states. Interpreted together with an observed exclusion limit of $9.8 \times \sigma_{\text{SM}}$ on the $H \to \mu\mu$ process [184], this fermion coupling is consistent with the mass dependent coupling ($g_{Hff} \propto m_f$) as predicted in the Standard Model and is very different e.g. from the universal lepton coupling of the Z boson (cf. $\sigma(pp \to Z \to \ell\ell)$) independent of the lepton mass.

For a future mass measurement, alternative analysis strategies have been studied, which are more sensitive to the observed resonance mass $m(\tau, \tau)$ than the default BDT-based analysis. In the context of multi-variate analysis techniques, it could be shown that the separate but dedicated training of BDTs on very different background types can provide an improved sensitivity, while the separate BDTs can be less complex and thus more robust at the same time. The strategies presented in these studies can be applied in order to further improve the sensitivity of searches both for SM and MSSM $H \to \tau\tau$ decays or to design related analyses on spin- or CP-properties. These can eventually lead either to a confirmation of the SM properties of the observed Higgs boson or to the discovery of new physics beyond the Standard Model.

APPENDIX A

Search for SM $H \to \tau\tau$ Decays with 4.6 fb^{-1} of ATLAS Data

A.1 Embedding Validation



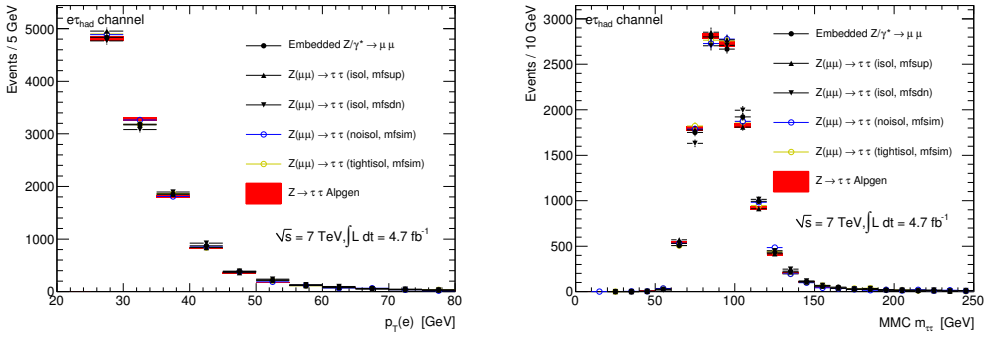
(a) Missing Transverse Energy $E_{\mathrm{T}}^{\mathrm{miss}}$

(b) Transverse Mass $m_T(\ell)$

(c) Transverse Momentum $p_T(\tau_{\mathrm{h}})$

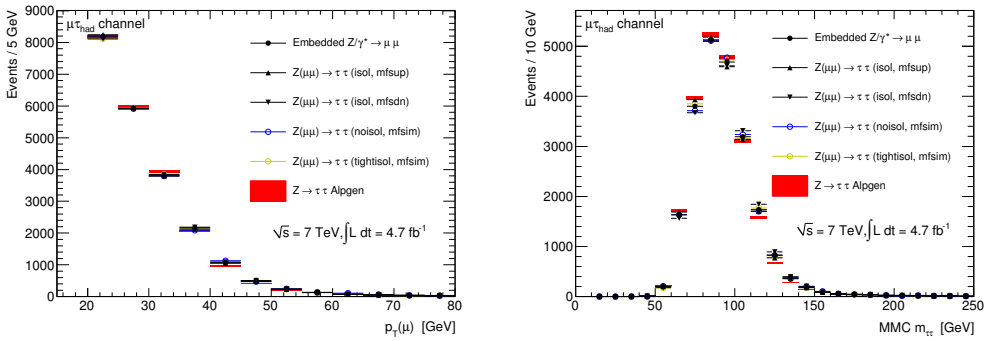
(d) Visible Mass m_{vis}

Figure A.1: Comparison of kinematic distributions between simulated $Z \to \tau\tau$ Monte Carlo events and the τ-embedded $Z \to \mu\mu$ data, shown for the inclusive selection defined in Sec. 6.1.1. Systematic uncertainties of the embedding method are included, which are estimated from variations of the muon isolation requirement (noisol, tightisol) and the cell energy subtraction (mfsup, mfsdn) as described in Sec. 5.2.8. All distributions are normalised to the event yield as predicted by the simulated Monte Carlo sample.



(a) Transverse Momentum $p_T(e)$

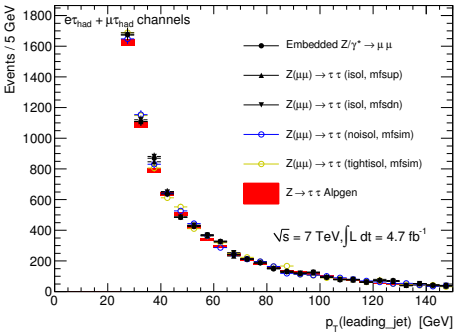
(b) MMC Mass $m_{\mathrm{MMC}}(\tau,\tau)$ $e\tau_{\mathrm{h}}$ Channel



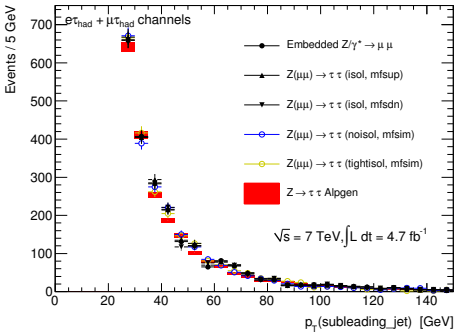
(c) Transverse Momentum $p_T(\mu)$

(d) MMC mass $m_{\mathrm{MMC}}(\tau,\tau)$ $\mu\tau_{\mathrm{h}}$ Channel

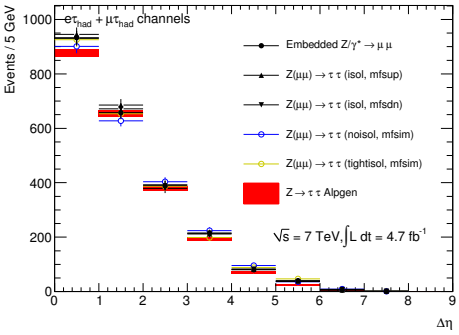
Figure A.2: Comparison of kinematic distributions between simulated $Z \to \tau\tau$ Monte Carlo events and the τ-embedded $Z \to \mu\mu$ data, shown for the inclusive selection defined in Sec. 6.1.1. Systematic uncertainties of the embedding method are included, which are estimated from variations of the muon isolation requirement (noisol, tightisol) and the cell energy subtraction (mfsup, mfsdn) as described in Sec. 5.2.8. All distributions are normalised to the event yield as predicted by the simulated Monte Carlo sample.



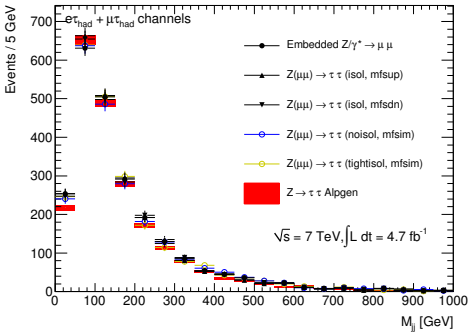
(a) Transverse Momentum $p_T(j_1)$



(b) Transverse Momentum $p_T(j_2)$

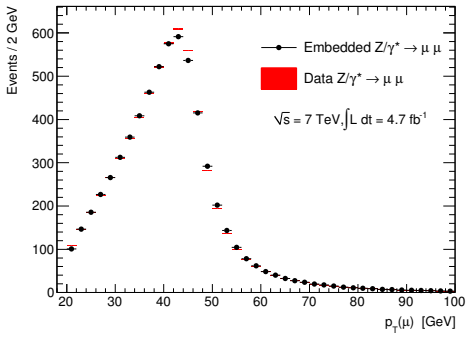


(c) Jet Pseudo-Rapidity Distance $\Delta\eta(j_1, j_2)$

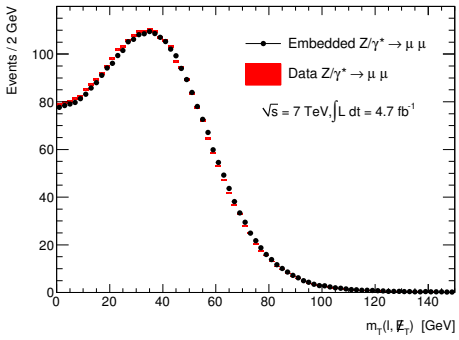


(d) Invariant Di-Jet Mass $m(j_1, j_2)$

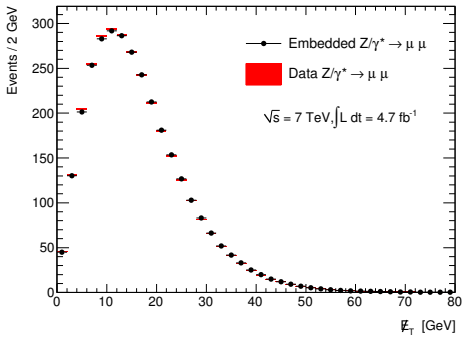
Figure A.3: Comparison of jet-related kinematic distributions between simulated $Z \to \tau\tau$ Monte Carlo events and the τ-embedded $Z \to \mu\mu$ data, shown for the inclusive selection defined in Sec. 6.1.1. Systematic uncertainties of the embedding method are included, which are estimated from variations of the muon isolation requirement (noisol, tightisol) and the cell energy subtraction (mfsup, mfsdn) as described in Sec. 5.2.8. All distributions are normalised to the event yield as predicted by the simulated Monte Carlo sample.



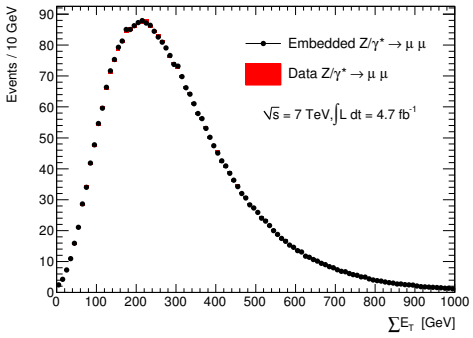
(a) Transverse Momentum $p_T(\mu)$



(b) Transverse Mass $p_T(\mu)$



(c) Missing Transverse Energy $E_{\mathrm{T}}^{\mathrm{miss}}$



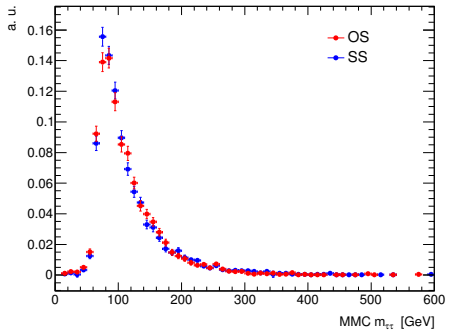
(d) Total Energy Sum $\sum E_T$

Figure A.4: Comparison of kinematic distributions obtained from the selected data and the embedded $Z \to \mu^+\mu^-$ sample. Small differences occur due to resolution effects from the simulation of the muons, which are folded with the resolution in the data. These effects are negligible for embedded τ decays.

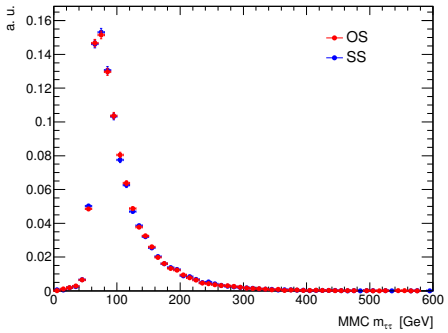
A.2 Validation of the Multijet Estimate

For the estimation of the multijet background in the analysis, the MMC mass shape obtained from the same-sign control region is extrapolated to the opposite-sign signal region via the ratio of OS and SS events in the anti-isolated control regions. This extrapolation, however, can only be performed if the MMC mass shape of multijet events is independent of their charge correlation. Therefore, this mass distribution is compared between OS and SS events in the anti-isolated control regions, which are dominated by the multijet background. Figures A.5-A.6 show these comparisons after subtracting the residual contributions from the electroweak backgrounds from the data. The agreement of the normalised distributions within their statistical uncertainties confirms that the MMC mass shapes are similar for OS and SS multijet events and can thus be extrapolated with a simple factor r_{QCD}.

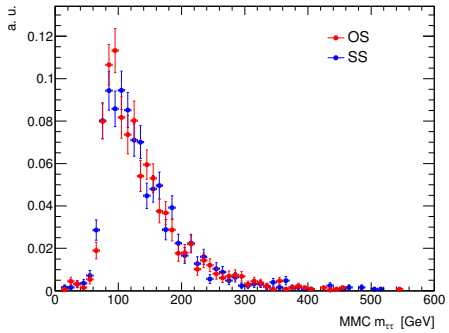
Also these extrapolation factors, which are obtained from the anti-isolated control regions, must be shown to be independent of the lepton isolation and not to be affected by inverting the selection. Figure A.7 therefore compares the distributions of the isolation observables for multijet events, which are again obtained by subtracting the electroweak contributions from the data. Their shapes are found to agree within uncertainties, so that the extrapolation factors r_{QCD} obtained from the anti-isolated control regions are confirmed to be valid for the extrapolation into the signal regions. Since also these comparisons are dominated by the multijet background, a potential signal contamination is expected to be negligible.



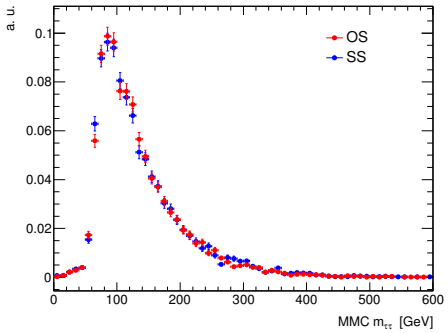
(a) MMC Mass $e\tau_{\mathrm{h}}$ Channel 0-jet (low-$E_{\mathrm{T}}^{\mathrm{miss}}$)



(b) MMC Mass $\mu\tau_{\mathrm{h}}$ Channel 0-jet (low-$E_{\mathrm{T}}^{\mathrm{miss}}$)

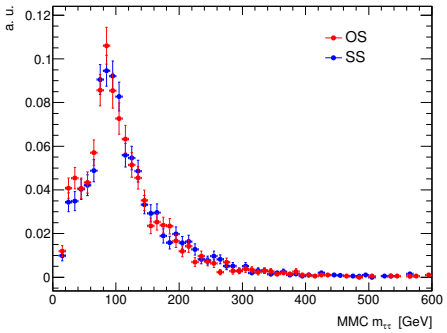


(c) MMC Mass $e\tau_{\mathrm{h}}$ Channel 0-jet (high-$E_{\mathrm{T}}^{\mathrm{miss}}$)

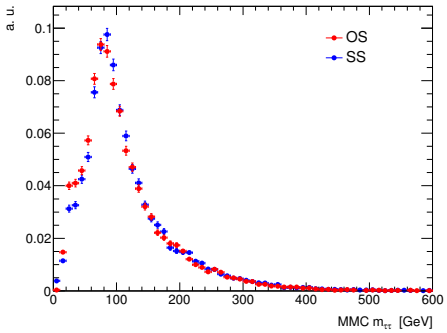


(d) MMC Mass $\mu\tau_{\mathrm{h}}$ Channel 0-jet (high-$E_{\mathrm{T}}^{\mathrm{miss}}$)

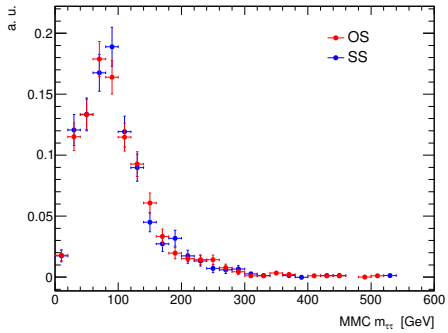
Figure A.5: Comparison of the MMC mass distributions for opposite-sign (red) and same-sign (blue) multijet events in the anti-isolated control regions for the 0-jet analysis categories.



(a) MMC Mass $e\tau_{\mathrm{h}}$ Channel 1-jet

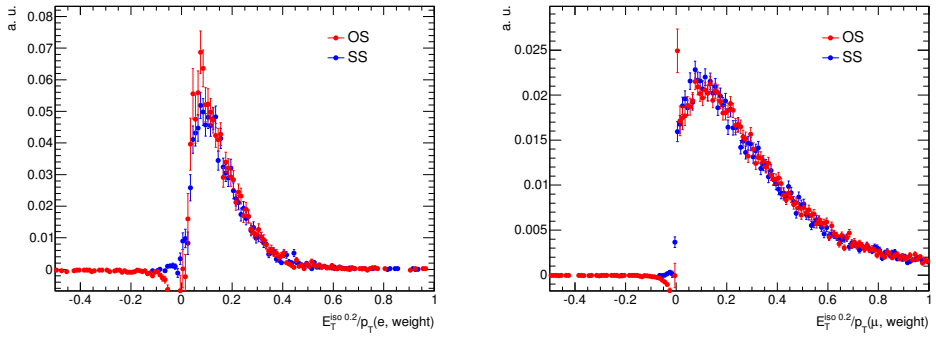


(b) MMC Mass $\mu\tau_{\mathrm{h}}$ Channel 1-jet



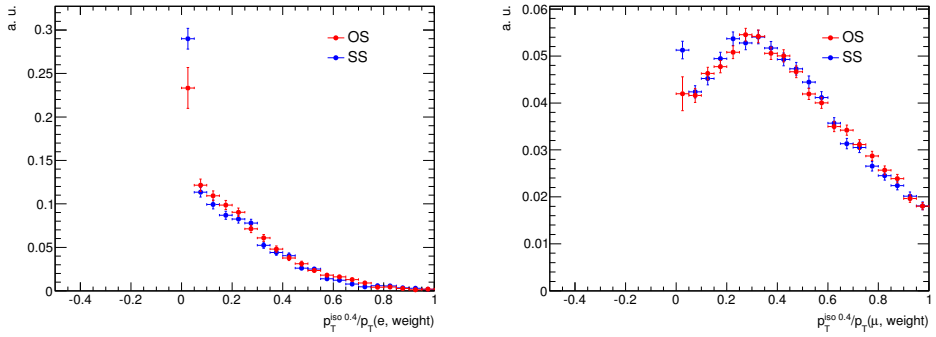
(c) MMC Mass $m_{\mathrm{MMC}}(\tau,\tau)$ VBF

Figure A.6: Comparison of the MMC mass distributions for opposite-sign (red) and same-sign (blue) multijet events in the anti-isolated control regions for the 1-jet and VBF analysis categories.



(a) $I_{ET}^{0.2}/p_T(e)$ Isolation ($e\tau_h$ Channel)

(b) $I_{ET}^{0.2}/p_T(\mu)$ Isolation ($\mu\tau_h$ Channel)

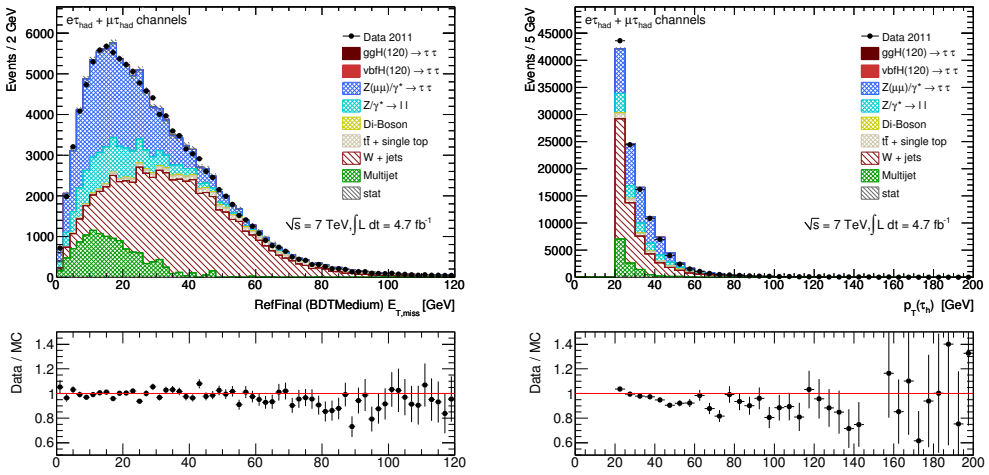


(c) $I_{PT}^{0.4}/p_T(e)$ Isolation ($e\tau_h$ Channel)

(d) $I_{PT}^{0.4}/p_T(\mu)$ Isolation ($\mu\tau_h$ Channel)

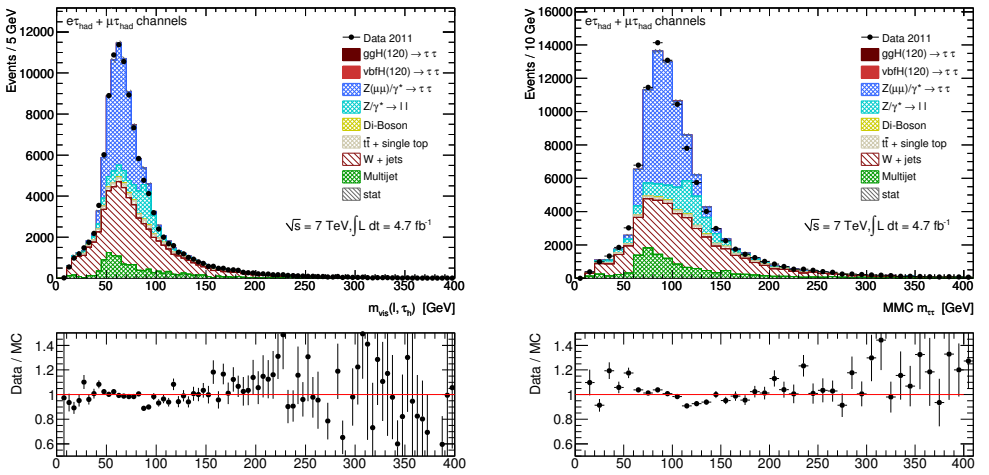
Figure A.7: Comparison of the isolation observables $I_{ET}^{0.2}/p_T$ and $I_{PT}^{0.4}/p_T$ between opposite-sign (red) and same-sign (blue) multijet events for the inclusive selection.

A.3 Kinematic Distributions



(a) Missing Transverse Energy $E_{\mathrm{T}}^{\mathrm{miss}}$

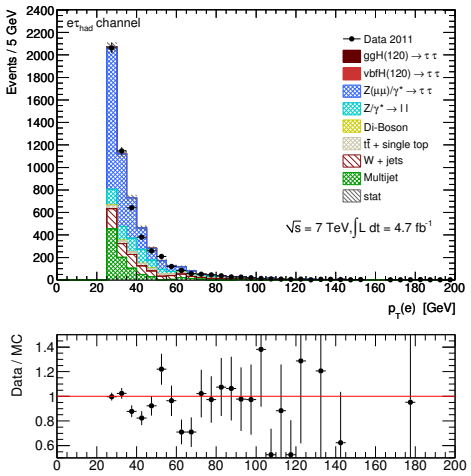
(b) Transverse Momentum $p_T(\tau_{\mathrm{h}})$



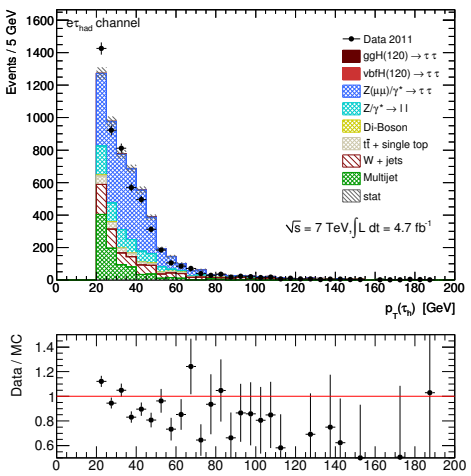
(c) Visible Mass $m_{\mathrm{vis}}(\ell, \tau_{\mathrm{h}})$

(d) MMC Mass $m_{\mathrm{MMC}}(\tau, \tau)$

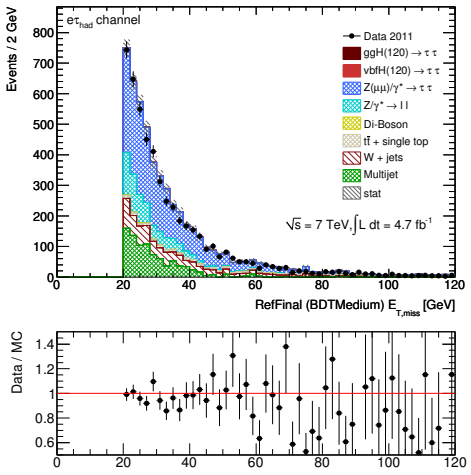
Figure A.8: Kinematic distributions of the data and the sum of the estimated background contributions at preselection level (cf. Sec. 6.1.1), shown combined for the $e\tau_{\mathrm{h}}$ and $\mu\tau_{\mathrm{h}}$ channels. The background components are estimated as described in Sec. 6.1.2.



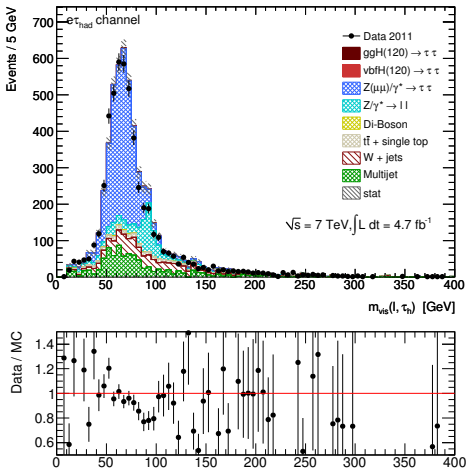
(a) Transverse Momentum $p_T(e)$



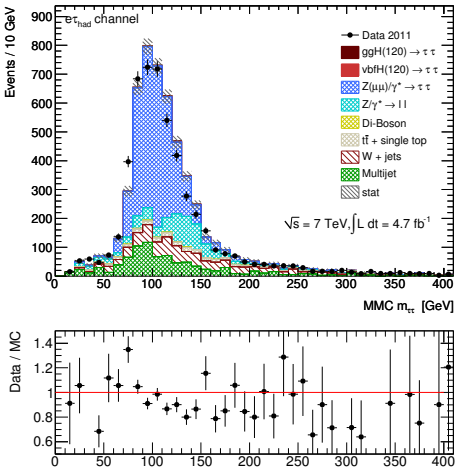
(b) Transverse Momentum $p_T(\tau_\mathrm{h})$



(c) Missing Transverse Energy $E_\mathrm{T}^\mathrm{miss}$

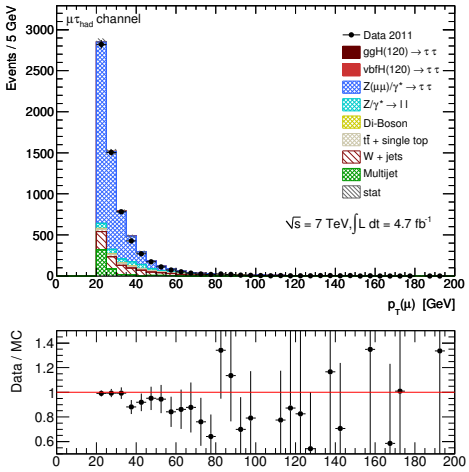


(d) Visible Mass $m_\mathrm{vis}(e, \tau_\mathrm{h})$

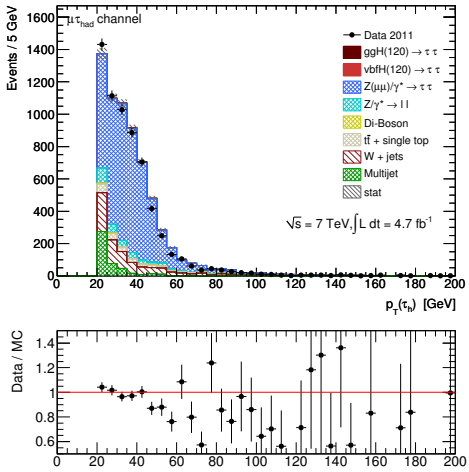


(e) MMC Mass $m_\mathrm{MMC}(\tau, \tau)$

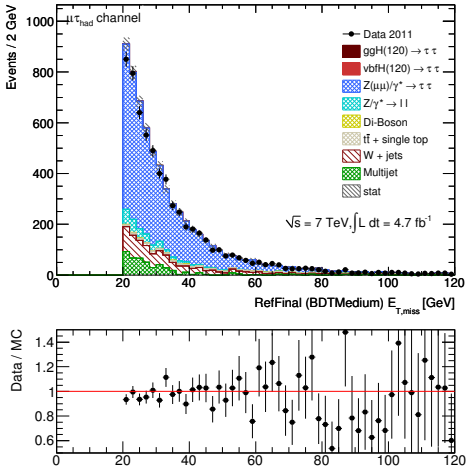
Figure A.9: Kinematic distributions of the data and the sum of the estimated background contributions at preselection level (cf. Sec. 6.1.1), shown for the $e\tau_\mathrm{h}$ channel. The different background components are estimated as described in Sec. 6.1.2.



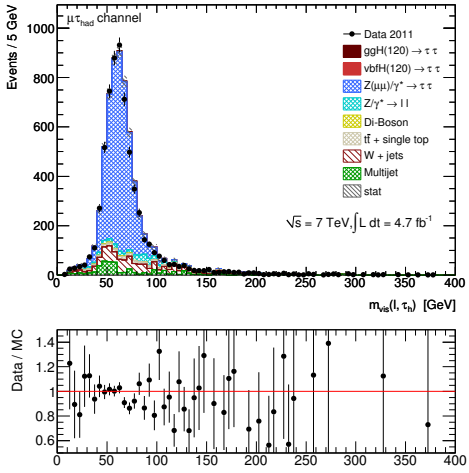
(a) Transverse Momentum $p_T(\mu)$



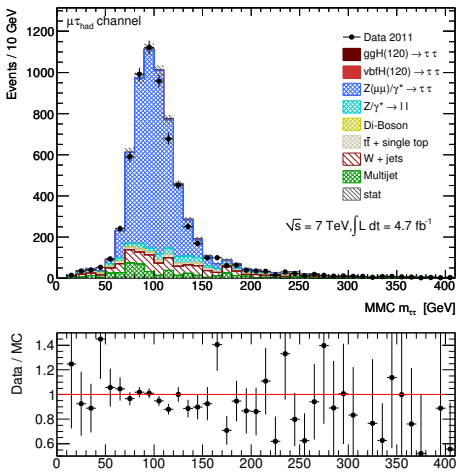
(b) Transverse Momentum $p_T(\tau_\mathrm{h})$



(c) Missing Transverse Energy $E_\mathrm{T}^\mathrm{miss}$

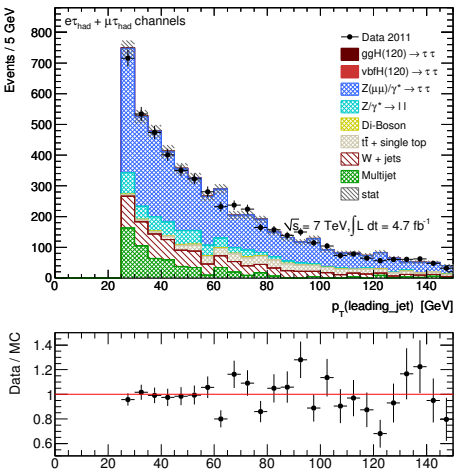


(d) Visible Mass $m_\mathrm{vis}(\mu, \tau_\mathrm{h})$

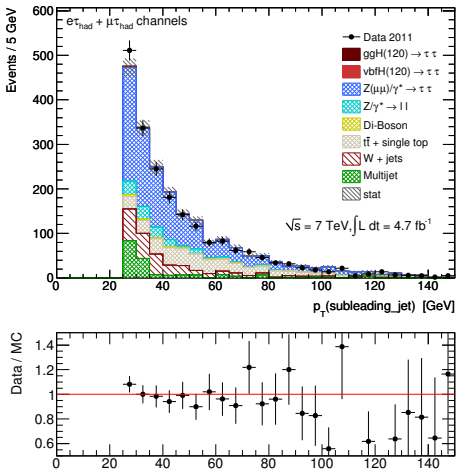


(e) MMC Mass $m_\mathrm{MMC}(\tau, \tau)$

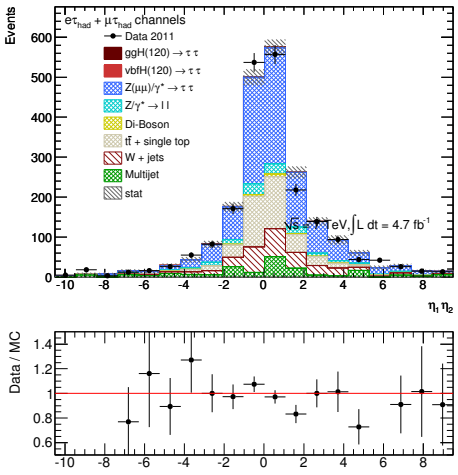
Figure A.10: Kinematic distributions of the data and the sum of the estimated background contributions at preselection level (cf. Sec. 6.1.1), shown for the $\mu\tau_\mathrm{h}$ channel. The different background components are estimated as described in Sec. 6.1.2.



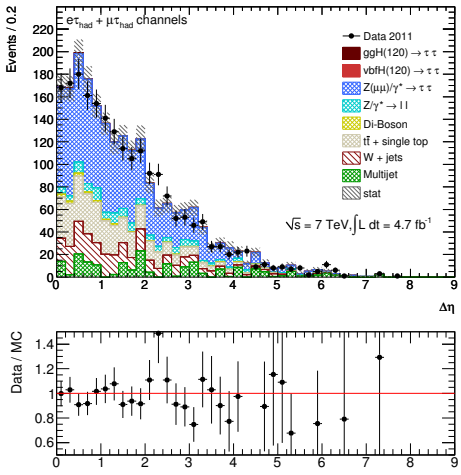
(a) Transverse Momentum $p_T(j_1)$



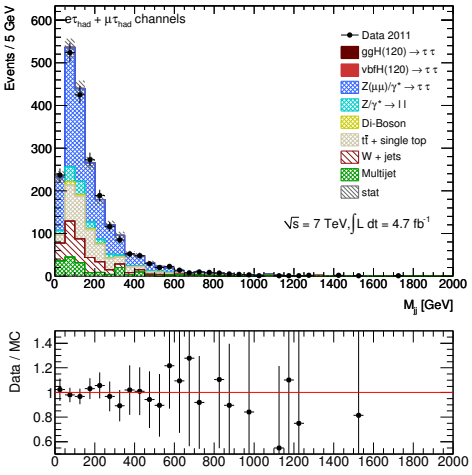
(b) Transverse Momentum $p_T(j_2)$



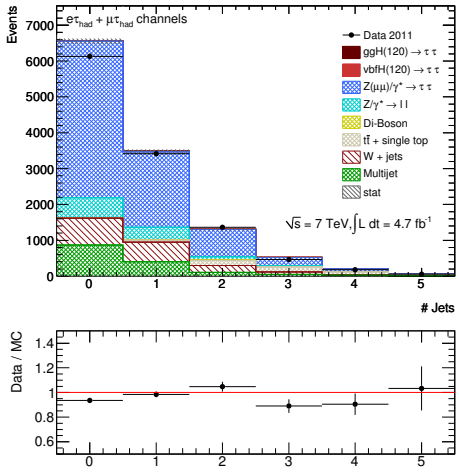
(c) Product of Jet Pseudo-Rapidities $\eta(j_1) \cdot \eta(j_2)$



(d) Jet Pseudo-Rapidity Distance $\Delta\eta(j_1, j_2)$



(e) Invariant Di-Jet Mass $m(j_1, j_2)$



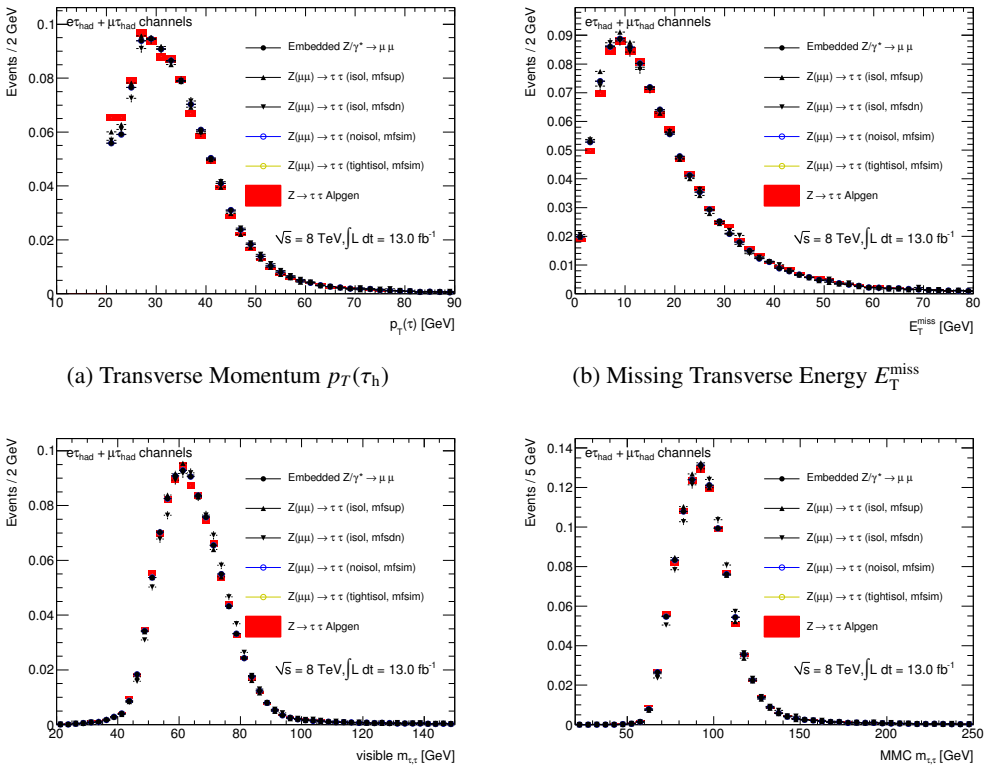
(f) Number of Jets per Event N_{jet}

Figure A.11: Jet-related kinematic distributions of the data and the sum of the estimated background contributions at preselection level (cf. Sec. 6.1.1), shown combined for the $e\tau_h$ and $\mu\tau_h$ channels. The different background components are estimated as described in Sec. 6.1.2.

APPENDIX B

Optimised Analysis with 4.6+13.0 fb^{-1} of ATLAS Data

B.1 Embedding Validation



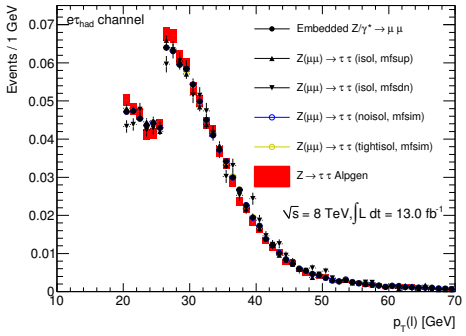
(a) Transverse Momentum $p_T(\tau_h)$

(b) Missing Transverse Energy E_T^{miss}

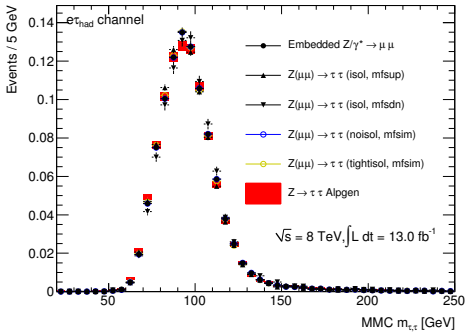
(c) Visible Mass m_{vis}

(d) MMC Mass $m_{MMC}(\tau,\tau)$

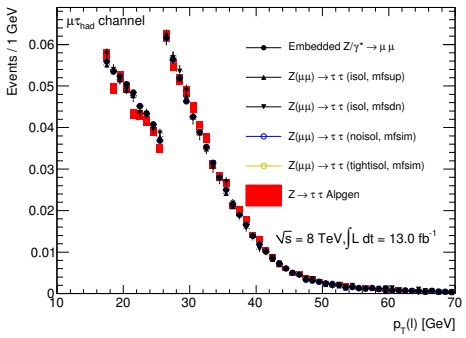
Figure B.1: Comparison of kinematic distributions between simulated $Z \to \tau\tau$ Monte Carlo events and the τ-embedded $Z \to \mu\mu$ data, shown for the inclusive selection defined in Sec. 6.2.1. Systematic uncertainties of the embedding method are included, which are estimated from variations of the muon isolation requirement (noisol, tightisol) and the cell energy subtraction (mfsup, mfsdn) as described in Sec. 5.2.8. All distributions are normalised to the event yield as predicted by the simulated Monte Carlo sample.



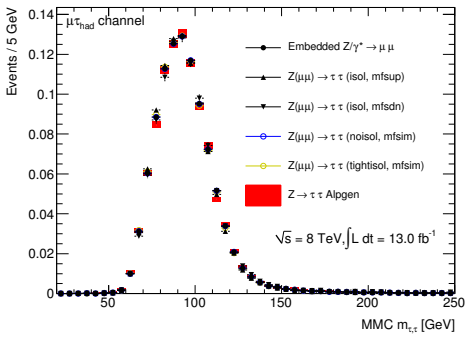
(a) Transverse Momentum $p_T(e)$



(b) MMC Mass $m_{\mathrm{MMC}}(\tau,\tau)$ $e\tau_{\mathrm{h}}$ Channel

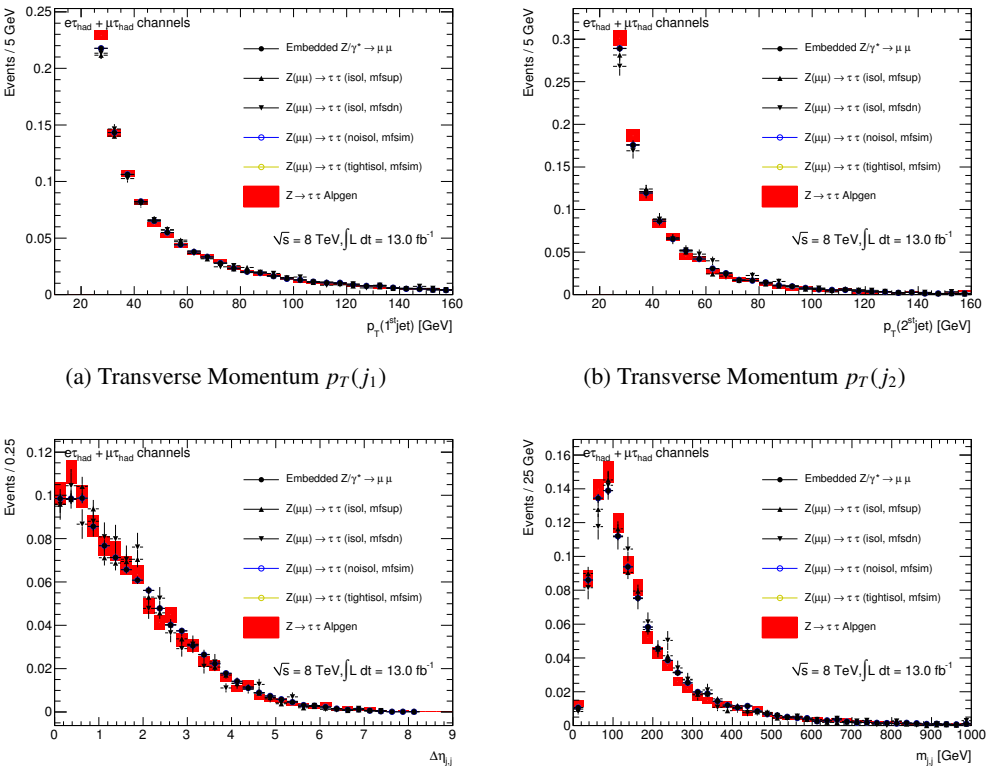


(c) Transverse Momentum $p_T(\mu)$



(d) MMC Mass $m_{\mathrm{MMC}}(\tau,\tau)$ $\mu\tau_{\mathrm{h}}$ Channel

Figure B.2: Comparison of kinematic distributions between simulated $Z \to \tau\tau$ Monte Carlo events and the τ-embedded $Z \to \mu\mu$ data, shown for the inclusive selection defined in Sec. 6.2.1. Systematic uncertainties of the embedding method are included, which are estimated from variations of the muon isolation requirement (noisol, tightisol) and the cell energy subtraction (mfsup, mfsdn) as described in Sec. 5.2.8. All distributions are normalised to the event yield as predicted by the simulated Monte Carlo sample.



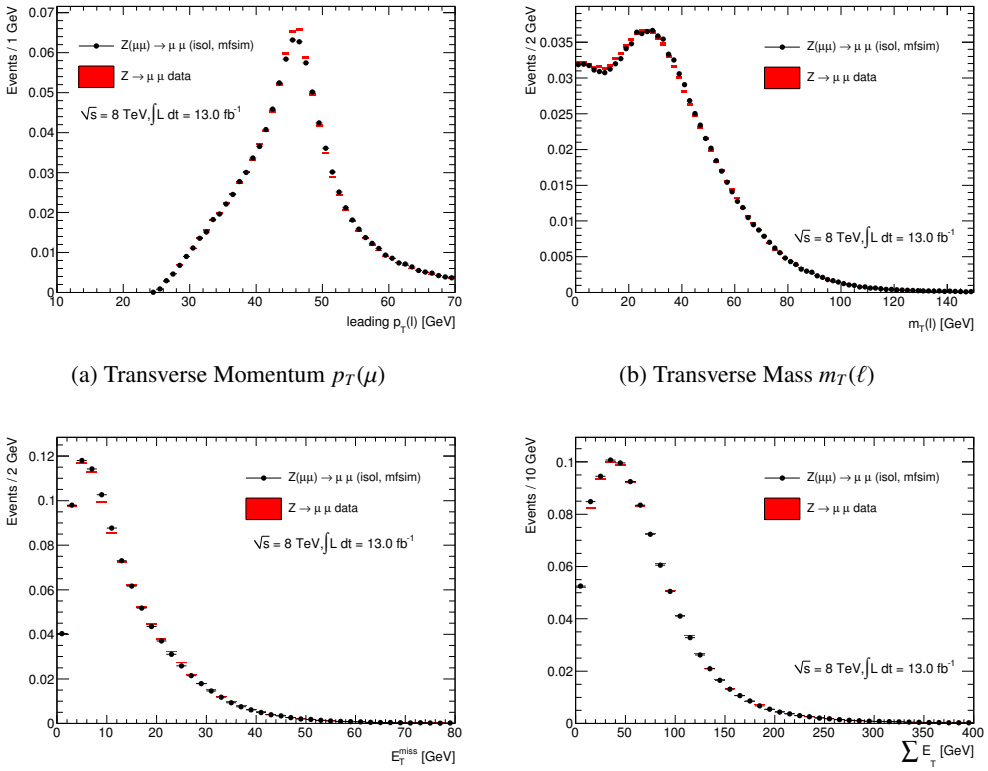
(a) Transverse Momentum $p_T(j_1)$

(b) Transverse Momentum $p_T(j_2)$

(c) Jet Pseudo-Rapidity Distance $\Delta\eta(j_1, j_2)$

(d) Invariant Di-Jet Mass $m(j_1, j_2)$

Figure B.3: Comparison of jet-related kinematic distributions between simulated $Z \to \tau\tau$ Monte Carlo events and the τ-embedded $Z \to \mu\mu$ data, shown for the inclusive selection defined in Sec. 6.2.1. Systematic uncertainties of the embedding method are included, which are estimated from variations of the muon isolation requirement (noisol,tightisol) and the cell energy subtraction (mfsup, mfsdn) as described in Sec. 5.2.8. All distributions are normalised to the event yield as predicted by the simulated Monte Carlo sample.



(a) Transverse Momentum $p_T(\mu)$

(b) Transverse Mass $m_T(\ell)$

(c) Missing Transverse Energy $E_{\mathrm{T}}^{\mathrm{miss}}$

(d) Total Energy Sum $\sum E_T$

Figure B.4: Comparison of kinematic distributions obtained from the selected data and the embedded $Z \to \mu^+\mu^-$ sample. Small differences occur due to resolution effects from the simulation of the muons, which are folded with the resolution in the data. These effects are negligible for embedded τ decays.

B.2 Event Selection

7 TeV		8 TeV	
VBF Category	Boosted Category	VBF Category	Boosted Category
▹ $p_{\mathrm{T}}^{\tau_{\mathrm{had}}}$ >30 GeV	–	▹ $p_{\mathrm{T}}^{\tau_{\mathrm{had}}}$ >30 GeV	▹ $p_{\mathrm{T}}^{\tau_{\mathrm{had}}}$ >30 GeV
▹ $E_{\mathrm{T}}^{\mathrm{miss}}$ >20 GeV	▹ $E_{\mathrm{T}}^{\mathrm{miss}}$ >20 GeV	▹ $E_{\mathrm{T}}^{\mathrm{miss}}$ >20 GeV	▹ $E_{\mathrm{T}}^{\mathrm{miss}}$ >20 GeV
▹ ≥ 2 jets	▹ $p_{\mathrm{T}}^{\mathrm{H}} > 100$ GeV	▹ ≥ 2 jets	▹ $p_{\mathrm{T}}^{\mathrm{H}} > 100$ GeV
▹ $p_{\mathrm{T}}^{j1}, p_{\mathrm{T}}^{j2} > 40$ GeV	▹ $0 < x_1 < 1$	▹ $p_{\mathrm{T}}^{j1} > 40, p_{\mathrm{T}}^{j2}$ >30 GeV	▹ $0 < x_1 < 1$
▹ $\Delta\eta_{jj} > 3.0$	▹ $0.2 < x_2 < 1.2$	▹ $\Delta\eta_{jj} > 3.0$	▹ $0.2 < x_2 < 1.2$
▹ $m_{jj} > 500$ GeV	▹ Fails VBF	▹ $m_{jj} > 500$ GeV	▹ Fails VBF
▹ centrality req.	–	▹ centrality req.	–
▹ $\eta_{j1} \times \eta_{j2} < 0$	–	▹ $\eta_{j1} \times \eta_{j2} < 0$	–
▹ $p_{\mathrm{T}}^{\mathrm{Total}} < 40$ GeV	–	▹ $p_{\mathrm{T}}^{\mathrm{Total}} < 30$ GeV	–
–	–	▹ p_{T}^{ℓ} >26 GeV	–
• m_{T} <50 GeV	• m_{T} <50 GeV	• m_{T} <50 GeV	• m_{T} <50 GeV
• $\Delta(\Delta R) < 0.8$	• $\Delta(\Delta R) < 0.8$	• $\Delta(\Delta R) < 0.8$	• $\Delta(\Delta R) < 0.8$
• $\sum \Delta\phi < 3.5$	• $\sum \Delta\phi < 1.6$	• $\sum \Delta\phi < 2.8$	–
–	–	• b-tagged jet veto	• b-tagged jet veto
1-jet Category	0-jet Category	1-jet Category	0-jet Category
▹ ≥ 1 jet, p_{T} >25 GeV	▹ 0 jets p_{T} >25 GeV	▹ ≥ 1 jet, p_{T} >30 GeV	▹ 0 jets p_{T} >30 GeV
▹ $E_{\mathrm{T}}^{\mathrm{miss}}$ >20 GeV	▹ $E_{\mathrm{T}}^{\mathrm{miss}}$ >20 GeV	▹ $E_{\mathrm{T}}^{\mathrm{miss}}$ >20 GeV	▹ $E_{\mathrm{T}}^{\mathrm{miss}}$ >20 GeV
▹ Fails VBF, Boosted	▹ Fails Boosted	▹ Fails VBF, Boosted	▹ Fails Boosted
• m_{T} <50 GeV	• m_{T} <30 GeV	• m_{T} <50 GeV	• m_{T} <30 GeV
• $\Delta(\Delta R) < 0.6$	• $\Delta(\Delta R) < 0.5$	• $\Delta(\Delta R) < 0.6$	• $\Delta(\Delta R) < 0.5$
• $\sum \Delta\phi < 3.5$	• $\sum \Delta\phi < 3.5$	• $\sum \Delta\phi < 3.5$	• $\sum \Delta\phi < 3.5$
–	• $p_{\mathrm{T}}^{\ell} - p_{\mathrm{T}}^{\tau} < 0$	–	• $p_{\mathrm{T}}^{\ell} - p_{\mathrm{T}}^{\tau} < 0$

Table B.1: Event requirements applied in the different categories of the optimised analysis [164]. Requirements marked with a triangle (▹) are categorisation requirements, meaning that if an event fails that requirement it is still considered for the remaining categories. Requirements marked with a bullet (•) are only applied to events passing all categorisation requirements in a category; events failing such requirements are discarded.

APPENDIX C

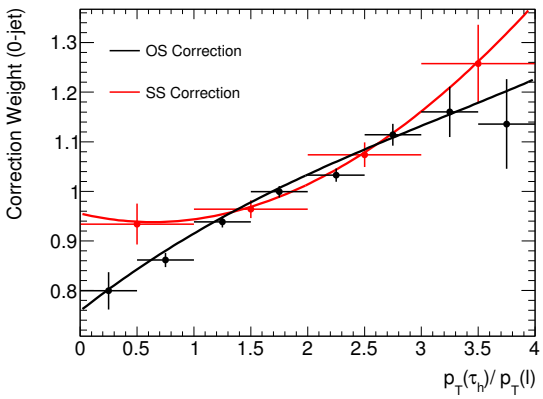
MVA-Based Search for SM $H \to \tau\tau$ Decays with 20.3 fb^{-1} of ATLAS Data

C.1 Background Estimation

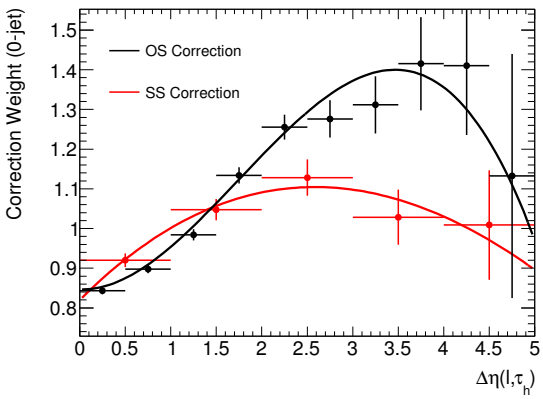
C.1.1 W+jets Re-weighting

Although the general jet$\to \tau_{\mathrm{h}}$ mis-identification probability is modelled better in the new ALPGEN+PYTHIA Monte Carlo simulation, a stronger dependence of the mis-identification on the kinematics of the jet$\to \tau_{\mathrm{h}}$ candidates is observed in the analysis categories and especially pronounced in the W+jets control regions. Therefore, correction factors are derived, which are applied as kinematic weights to all W+jets events. Since the mis-modelling is expected to depend both on the charge correlation of the (ℓ,jet$\to \tau_{\mathrm{h}}$) pair and the selection of additional jets, the corrections for OS and SS events are derived separately for the Rest category dominated by the W+1-jet process and for the 1-jet category with dominant W+2-jet contributions. The corresponding correction factors shown in Fig. C.1 are derived and applied in a two-step procedure: the first step derives $p_T(\ell)/p_T(\tau_{\mathrm{h}})$ dependent weights, which are obtained by scaling the mis-modelled $p_T(\ell)/p_T(\tau_{\mathrm{h}})$ distribution in the W+jets control region to the data[1]. The second step then repeats this procedure with $\Delta\eta(\ell, \tau_{\mathrm{h}})$ dependent weights. The binned correction factors are fitted with a third-order polynomial function, so that the two kinematic weights can be obtained from the two corresponding fit-functions for each simulated W+jets event. The application of these corrections results in good agreement of the modelled W+jets prediction and the data for all other observables. Systematic uncertainties are obtained by randomly varying the binned correction factors within their errors and re-fitting the polynomial. After repeating this procedure in 50,000 iterations, the two fitted functions pol(x) with the largest integrated difference $\int |\mathrm{pol}(x) - \tilde{\mathrm{pol}}(x)|dx$ to the default polynomial $\tilde{\mathrm{pol}}(x)$ in opposite direction are then selected to provide a systematic variation of the reweighting, which is expected to cover all related uncertainties arising both from the method and from the fluctuations of the data in the corresponding control distributions, which are used to derive the corrections.

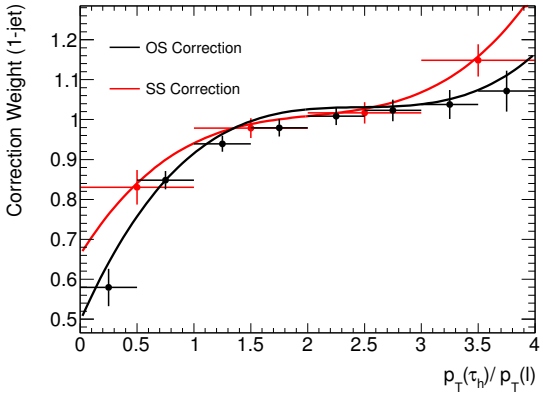
[1] For this scaling of W+jets Monte Carlo to the data, residual contributions predicted for other backgrounds are subtracted beforehand.



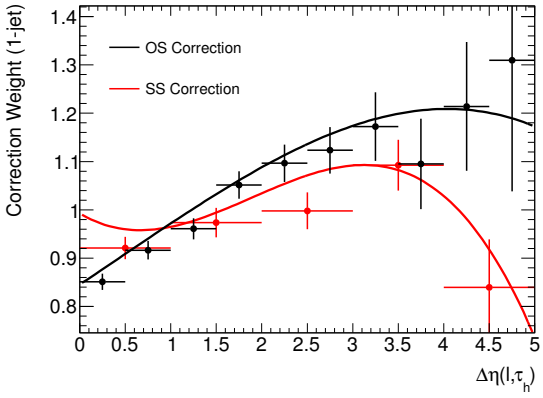
(a) Rest Category: $p_T(\ell)/p_T(\tau_\mathrm{h})$



(b) Rest Category: $\Delta\eta(\ell, \tau_\mathrm{h})$



(c) 1-jet Category: $p_T(\ell)/p_T(\tau_\mathrm{h})$



(d) 1-jet Category: $\Delta\eta(\ell, \tau_\mathrm{h})$

Figure C.1: Correction weights for the W+jets Monte Carlo in the Rest and 1-jet category, binned in the observables $p_T(\ell)/p_T(\tau_\mathrm{h})$ and $\Delta\eta(\ell, \tau_\mathrm{h})$. The distributions of these weights are fitted with a third-order polynomial function.

C.1.2 Background Summary

Process/Category	0-jet OS	0-jet SS
$ZH(H \to \tau\tau)$ (m_H = 125 GeV)	$7.3 \pm 0.10^{+8.6}_{-8.7}$	$0.1 \pm 0.01^{+0.01}_{-0.01}$
$WH(H \to \tau\tau)$ (m_H = 125 GeV)	$11.8 \pm 0.17^{+8.7}_{-8.7}$	$1.23 \pm 0.05^{+0.13}_{-0.12}$
$VBF(H \to \tau\tau)$ (m_H = 125 GeV)	$34.5 \pm 0.2^{+8.6}_{-8.8}$	$0.09 \pm 0.01^{+0.02}_{-0.02}$
$gg \to H \to \tau\tau$ (m_H = 125 GeV)	$232 \pm 2^{+28}_{-24}$	$2.9 \pm 0.2^{+0.7}_{-0.6}$
$Z \to \tau\tau$	$172466 \pm 184.57^{+8538}_{-8606}$	$1571 \pm 46^{+122}_{-167}$
$Z \to \ell\ell(\ell \to \tau_{\mathrm{h}})$	$11773 \pm 217^{+1803}_{-1826}$	$244 \pm 32^{+65}_{-69}$
$Z \to \ell\ell(j \to \tau_{\mathrm{h}})$	$10605 \pm 264^{+370}_{-411}$	$9361 \pm 235^{+277}_{-237}$
Di-boson	$1543 \pm 23^{+149}_{-151}$	$226 \pm 8^{+26}_{-29}$
Top	$8760 \pm 60^{+428}_{-347}$	$2017 \pm 34^{+207}_{-155}$
W+jets	$121477 \pm 746^{+10548}_{-9370}$	$47194 \pm 679^{+5630}_{-5368}$
QCD	$62455 \pm 878^{+9368}_{-9368}$	–
Total Background	$377620 \pm 1187^{+16345}_{-15652}$	–
Data	378168	115347

Table C.1: Predicted event yields for the 0-jet OS and SS categories. The quoted uncertainties correspond to the statistical and systematic uncertainties before the global fit. The background components are estimated as described in Sec. 7.4.

Process/Category	1-jet OS	1-jet SS
$ZH(H \to \tau\tau)$ (m_H = 125 GeV)	$7.3 \pm 0.1^{+8.6}_{-8.7}$	$0.31 \pm 0.02^{+0.03}_{-0.03}$
$WH(H \to \tau\tau)$ (m_H = 125 GeV)	$11.8 \pm 0.2^{+8.7}_{-8.7}$	$0.97 \pm 0.05^{+0.10}_{-0.11}$
$VBF(H \to \tau\tau)$ (m_H = 125 GeV)	$34.5 \pm 0.2^{+8.6}_{-8.8}$	$0.54 \pm 0.03^{+0.05}_{-0.06}$
$gg \to H \to \tau\tau$ (m_H = 125 GeV)	$232 \pm 2^{+28}_{-24}$	$2.7 \pm 0.2^{+0.8}_{-0.7}$
$Z \to \tau\tau$	$40791 \pm 110^{+2116}_{-2059}$	$725 \pm 21^{+67}_{-102}$
$Z \to \ell\ell(\ell \to \tau_{\text{h}})$	$2501 \pm 55^{+531}_{-457}$	$86 \pm 22^{+31}_{-24}$
$Z \to \ell\ell(j \to \tau_{\text{h}})$	$6970 \pm 110^{+316}_{-570}$	$6998 \pm 108^{+323}_{-644}$
Di-boson	$1918 \pm 25^{+167}_{-169}$	$583 \pm 13^{+53}_{-51}$
Top	$8760 \pm 60^{+428}_{-347}$	$2017 \pm 34^{+207}_{-155}$
W+jets	$42409 \pm 330^{+6171}_{-4219}$	$22956 \pm 330^{+3132}_{-2510}$
QCD	$21601 \pm 4^{+3086}_{-3240}$	–
Total Background	$123921 \pm 578^{+7258}_{-5678}$	–
Data	123195	53006

Table C.2: Predicted event yields for the 1-jet OS and SS categories. The quoted uncertainties correspond to the statistical and systematic uncertainties before the global fit. The background components are estimated as described in Sec. 7.4.

C.2 Control Distributions

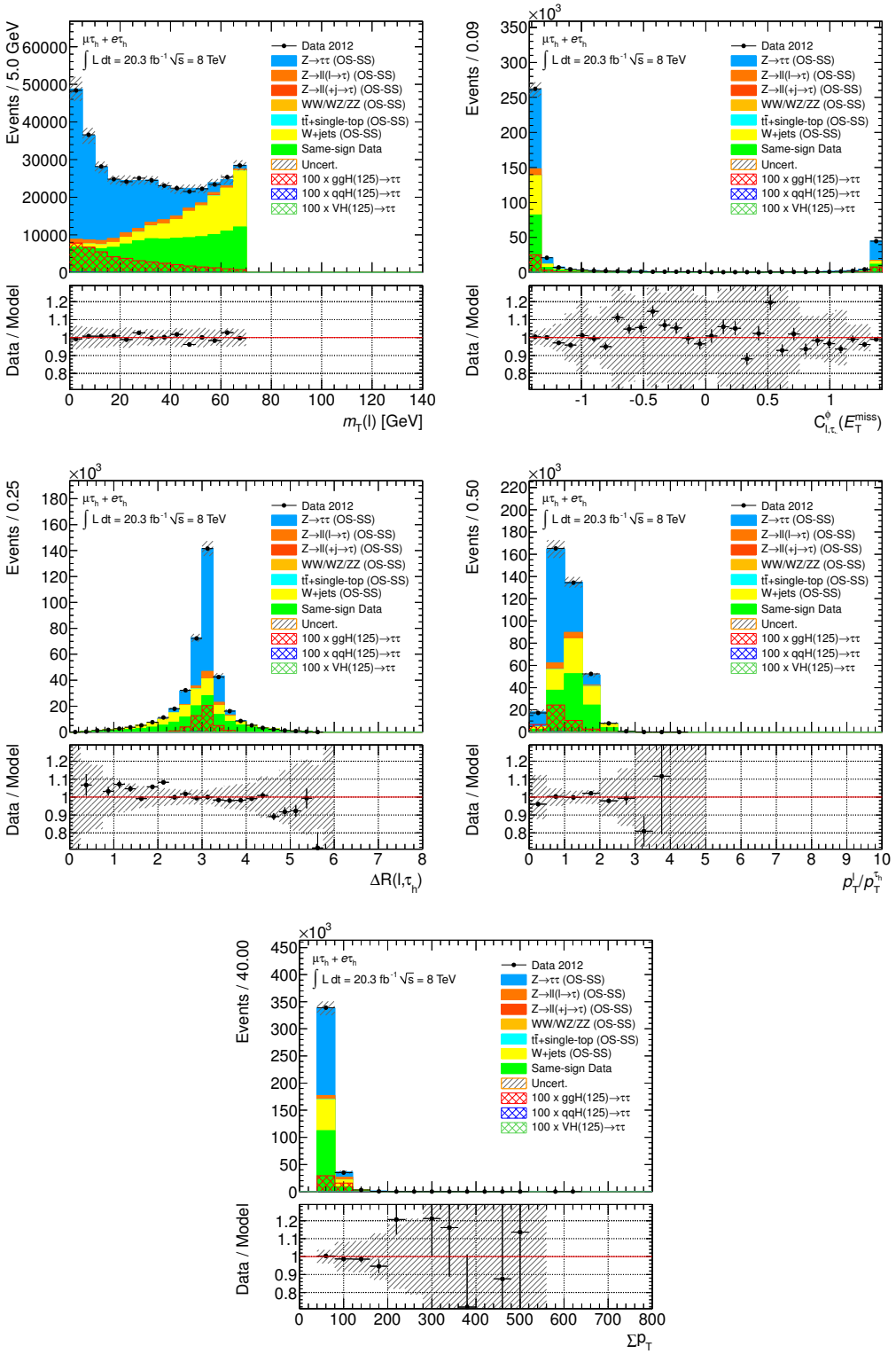


Figure C.2: Distributions of the input observables for the BDT training in the Rest category. The background components are estimated as described in Sec. 7.4.

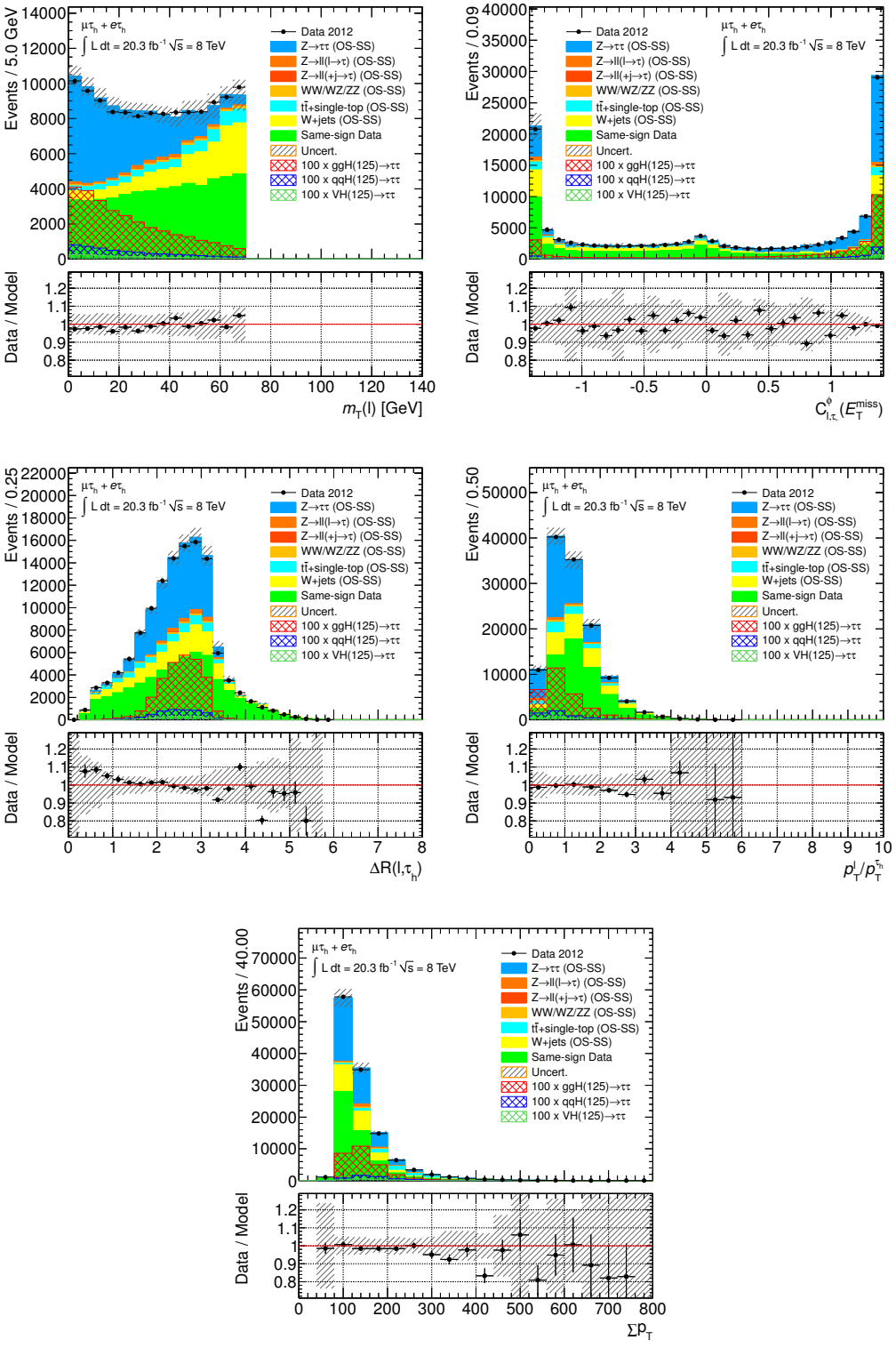


Figure C.3: Distributions of the input observables for the BDT training in the 1-jet category. The background components are estimated as described in Sec. 7.4.

C.3 Systematic Uncertainties

C.3.1 Normalisation Uncertainties

Systematic	qqH	ggH	WH	ZH	$Z \to \ell\ell(\ell \to \tau_{\text{h}})$	$WW/WZ/ZZ$
BR_tautau	$^{+5.7}_{-5.7}$	$^{+5.7}_{-5.7}$	$^{+5.7}_{-5.7}$	$^{+5.7}_{-5.7}$	< 0.5	< 0.5
Gen_Qmass_ggH	< 0.5	$^{+18.0}_{-18.0}$	< 0.5	< 0.5	< 0.5	< 0.5
QCDscale_V	< 0.5	< 0.5	< 0.5	< 0.5	$^{+1.0}_{-1.0}$	< 0.5
QCDscale_VH	< 0.5	< 0.5	$^{+1.0}_{-1.0}$	$^{+1.0}_{-1.0}$	< 0.5	< 0.5
QCDscale_VV	< 0.5	< 0.5	< 0.5	< 0.5	< 0.5	$^{+5.0}_{-5.0}$
QCDscale_ggH2in	< 0.5	$^{+25.0}_{-20.0}$	< 0.5	< 0.5	< 0.5	< 0.5
QCDscale_qqH	$^{+2.1}_{-2.1}$	< 0.5	< 0.5	< 0.5	< 0.5	< 0.5
pdf_Higgs_gg	< 0.5	$^{+8.0}_{-7.0}$	< 0.5	< 0.5	< 0.5	< 0.5
pdf_Higgs_qq	$^{+3.0}_{-3.0}$	< 0.5	$^{+3.0}_{-3.0}$	$^{+3.0}_{-3.0}$	< 0.5	< 0.5
pdf_qq	< 0.5	< 0.5	< 0.5	< 0.5	$^{+4.0}_{-4.0}$	$^{+4.0}_{-4.0}$
UE_gg	< 0.5	$^{+30.0}_{-30.0}$	< 0.5	< 0.5	< 0.5	< 0.5
UE_qq	$^{+6.0}_{-6.0}$	< 0.5	< 0.5	< 0.5	< 0.5	< 0.5

Table C.3: Summary of theoretical systematic uncertainties in % for the signal and background components in the VBF category. Numbers marked with < 0.5 are not considered in the fit model. For uncertainties marked with S, additional bin-by-bin uncertainties are taken into account as BDT score shape systematics.

Systematic	qqH	ggH	WH	ZH	$Z \to \tau\tau$	$Z \to \ell\ell(\ell \to \tau_h)$	$Z \to \ell\ell(j \to \tau_h)$	Top	$WW/WZ/ZZ$
BTag_BEFF	< 0.5	< 0.5	< 0.5	$^{-0.5}_{+0.5}$ S	< 0.5	< 0.5	< 0.5	$^{-7.8}_{+8.0}$ S	< 0.5
BTag_CEFF	< 0.5 S	< 0.5	$^{-1.0}_{+1.0}$ S	$^{-0.7}_{+0.7}$ S	< 0.5	< 0.5	< 0.5	< 0.5	< 0.5
EL_EFF	$^{+0.9}_{-0.9}$	$^{+0.8}_{-0.8}$	$^{+0.9}_{-0.9}$	$^{+0.9}_{-0.9}$	< 0.5	$^{+1.0}_{-1.0}$	$^{+1.5}_{-1.5}$	$^{+0.8}_{-0.8}$	$^{+0.8}_{-0.8}$
EL_EFF_Emb	< 0.5	< 0.5	< 0.5	< 0.5	$^{+0.6}_{-0.6}$	< 0.5	< 0.5	< 0.5	< 0.5
EL_RES	< 0.5	< 0.5	$^{-0.9}_{+2.7}$	$^{+0.0}_{+1.8}$	< 0.5	$^{+16.2}_{-1.5}$ S	$^{+0.5}_{-1.7}$	< 0.5	$^{+1.4}_{-0.0}$
EL_SCALE	< 0.5	$^{+1.1}_{-0.8}$	$^{+1.5}_{-2.2}$	$^{+0.2}_{-1.0}$	$^{+1.0}_{-0.9}$	$^{+18.2}_{-1.1}$	$^{+1.8}_{-1.9}$	$^{+0.8}_{-0.6}$	$^{+0.6}_{-1.1}$
LUMI	$^{-2.8}_{+2.8}$	$^{-2.8}_{+2.8}$	$^{-2.8}_{+2.8}$	$^{-2.8}_{+2.8}$	< 0.5	$^{-2.8}_{+2.8}$	< 0.5	< 0.5	$^{-2.8}_{+2.8}$
MU_EFF	$^{+1.1}_{-1.1}$	$^{+1.2}_{-1.2}$	$^{+1.0}_{-1.0}$	$^{+1.2}_{-1.2}$	< 0.5	$^{+0.9}_{-0.9}$	< 0.5	$^{+1.2}_{-1.2}$	$^{+1.1}_{-1.1}$
MU_EFF_Emb	< 0.5	< 0.5	< 0.5	< 0.5	$^{+0.7}_{-0.7}$	< 0.5	< 0.5	< 0.5	< 0.5
MU_SCALE	< 0.5	< 0.5	$^{+0.4}_{-1.0}$	$^{+1.6}_{-1.6}$	< 0.5	< 0.5	< 0.5	< 0.5	< 0.5
PU_RESCALE	< 0.5 S	< 0.5 S	< 0.5 S	< 0.5 S	< 0.5	$^{-1.4}_{+1.0}$	< 0.5	$^{+0.5}_{-0.5}$	< 0.5
TAU_EFAKE	< 0.5	< 0.5	< 0.5	< 0.5	< 0.5	$^{+13.8}_{-13.8}$ S	< 0.5	< 0.5	< 0.5
TAU_ID	$^{+3.3}_{-3.3}$	$^{+3.3}_{-3.3}$	$^{+3.4}_{-3.4}$	$^{+3.5}_{-3.5}$	$^{+3.3}_{-3.3}$	< 0.5	< 0.5	$^{+1.6}_{-1.6}$	$^{+1.4}_{-1.4}$
TAU_MFAKE	< 0.5	< 0.5	< 0.5	< 0.5	< 0.5	$^{+6.2}_{-6.2}$	< 0.5	< 0.5	< 0.5
TES_FAKE	< 0.5	< 0.5	< 0.5	< 0.5	< 0.5	$^{+18.1}_{-2.7}$ S	$^{+6.8}_{-10.2}$	$^{+3.3}_{-4.2}$	$^{+2.9}_{-2.9}$
TES_TRUE	$^{+1.4}_{-1.3}$	$^{+3.1}_{-3.2}$	$^{+0.0}_{-0.9}$	$^{+0.1}_{-1.6}$	$^{+2.8}_{-2.7}$	< 0.5	< 0.5	< 0.5	< 0.5

Table C.4: Summary of experimental systematic uncertainties (except jet-related) in % for the signal and background components in the VBF category. Numbers marked with < 0.5 are not considered in the fit model. For uncertainties marked with S, additional bin-by-bin uncertainties are taken into account as BDT score shape systematics.

Systematic	qqH	ggH	WH	ZH	$Z \to \ell\ell(\ell \to \tau_{\mathrm{h}})$	$Z \to \ell\ell(j \to \tau_{\mathrm{h}})$	Top	$WW/WZ/ZZ$
JER	< 0.5 S	$^{-2.4}_{+2.4}$ S	$^{-13.4}_{+13.4}$	$^{+7.9}_{-7.9}$	$^{-2.1}_{+2.1}$ S	$^{-12.1}_{+12.1}$ S	< 0.5 S	$^{-8.2}_{+8.2}$ S
JES_Detector	$^{+0.6}_{-0.6}$	$^{+0.8}_{-0.8}$	$^{+0.9}_{-1.5}$	$^{+6.6}_{-6.6}$	$^{+0.0}_{+18.6}$ S	$^{+4.4}_{-2.2}$	$^{+1.8}_{-1.7}$	$^{+1.7}_{-3.3}$
JES_Eta_StatMethod	$^{+0.8}_{-0.8}$	$^{+0.9}_{-1.9}$	$^{+2.5}_{-3.0}$	$^{+3.5}_{-1.9}$	$^{+0.0}_{+16.0}$ S	$^{+3.7}_{-2.1}$	$^{+0.0}_{+2.0}$	$^{+0.0}_{-0.8}$
JES_Modelling	$^{+2.2}_{-2.5}$ S	$^{+5.4}_{-5.9}$	$^{+7.3}_{-9.2}$	$^{+6.4}_{-12.5}$	$^{+17.7}_{-0.0}$ S	$^{+8.0}_{-5.5}$	$^{+1.2}_{-1.0}$	$^{+6.0}_{-9.0}$
JES_PileRho_TAU_GG	< 0.5	$^{+2.8}_{-3.3}$	< 0.5	< 0.5	< 0.5	< 0.5	$^{+1.1}_{-0.2}$	< 0.5
JES_PileRho_TAU_QG	< 0.5	< 0.5	< 0.5	< 0.5	$^{+0.0}_{+13.9}$ S	$^{+4.9}_{-2.5}$	< 0.5	< 0.5
JES_PileRho_TAU_QQ	$^{+1.3}_{-1.3}$	< 0.5	$^{+6.3}_{-5.7}$	$^{+2.7}_{-5.1}$	< 0.5	< 0.5	< 0.5	$^{+2.2}_{-2.7}$
JES_Statistical	$^{+0.5}_{-0.5}$	$^{+1.0}_{-0.8}$	$^{+0.0}_{-1.8}$	$^{+2.3}_{-1.1}$	$^{+1.1}_{-2.0}$	$^{+1.1}_{-1.2}$	$^{-0.5}_{+0.7}$	$^{+0.8}_{-0.0}$
JES_Eta_Modelling	$^{+4.1}_{-4.2}$ S	$^{+8.8}_{-9.7}$	$^{+8.6}_{-13.1}$	$^{+14.0}_{-16.4}$	$^{+25.7}_{-0.0}$ S	$^{+12.1}_{-12.5}$	$^{+5.2}_{-5.9}$	$^{+5.7}_{-9.0}$
JES_FlavComp_TAU_G	< 0.5	$^{+4.0}_{-5.4}$	< 0.5	< 0.5	$^{+14.9}_{-0.0}$ S	$^{+8.0}_{-5.5}$	< 0.5	< 0.5
JES_FlavComp_TAU_Q	$^{+2.1}_{-2.2}$ S	< 0.5	$^{+6.6}_{-7.0}$	$^{+8.9}_{-13.2}$	< 0.5	< 0.5	$^{+2.0}_{-1.4}$	$^{+1.2}_{-6.4}$
JES_FlavResp	$^{+1.4}_{-1.3}$ S	$^{+2.5}_{-2.5}$	$^{+4.6}_{-7.8}$	$^{+1.9}_{-6.2}$	$^{+0.0}_{+14.2}$ S	$^{+7.4}_{-4.2}$	$^{+1.4}_{-0.7}$	$^{+3.6}_{-5.6}$
JES_Flavb	< 0.5	< 0.5	< 0.5	< 0.5	< 0.5	< 0.5	$^{-1.9}_{+0.9}$	< 0.5
JES_Mu	< 0.5	< 0.5	$^{-1.6}_{-0.0}$	$^{+1.7}_{-1.2}$	$^{+0.0}_{+1.3}$	$^{+2.7}_{-1.7}$	$^{+1.5}_{-0.5}$	$^{+1.2}_{-0.0}$
JES_NPV	$^{+0.6}_{-0.5}$	$^{+1.4}_{-2.0}$	$^{+1.4}_{-3.9}$	< 0.5	$^{+10.2}_{-3.9}$	$^{+3.4}_{-1.1}$	$^{+2.4}_{-1.9}$	$^{+1.4}_{-0.9}$
JVF	< 0.5	$^{+2.5}_{-0.0}$	$^{+10.4}_{-0.0}$	$^{+4.1}_{-0.4}$	$^{+8.4}_{-0.0}$	$^{+2.1}_{-0.4}$	$^{+1.0}_{-0.0}$	$^{+0.0}_{+0.6}$
MET_RESOSOFT	< 0.5	$^{-0.9}_{+1.1}$	$^{-1.2}_{+0.5}$	< 0.5	$^{+2.1}_{-0.0}$ S	$^{+0.5}_{-0.1}$	< 0.5	$^{-2.8}_{+1.2}$
MET_SCALESOFT	< 0.5	$^{-0.9}_{+0.3}$	$^{+0.7}_{-3.2}$	$^{+0.0}_{+1.6}$	$^{-1.6}_{+1.9}$	$^{+0.6}_{-0.1}$	$^{+0.1}_{-1.5}$	$^{+1.4}_{-1.6}$

Table C.5: Summary of jet-related experimental systematic uncertainties in % for the signal and background components in the VBF category. Numbers marked with < 0.5 are not considered in the fit model. For uncertainties marked with S, additional bin-by-bin uncertainties are taken into account as BDT score shape systematics.

Systematic	qqH	ggH	WH	ZH	$Z \to \ell\ell(\ell \to \tau_h)$	$WW/WZ/ZZ$
BR_tautau	$^{+5.7}_{-5.7}$	$^{+5.7}_{-5.7}$	$^{+5.7}_{-5.7}$	$^{+5.7}_{-5.7}$	< 0.5	< 0.5
Gen_Qmass_ggH	< 0.5	$^{+29.0}_{-29.0}$	< 0.5	< 0.5	< 0.5	< 0.5
QCDscale_V	< 0.5	< 0.5	< 0.5	< 0.5	$^{+1.0}_{-1.0}$	< 0.5
QCDscale_VH	< 0.5	< 0.5	$^{+4.1}_{-4.0}$	$^{+4.1}_{-4.0}$	< 0.5	< 0.5
QCDscale_VV	< 0.5	< 0.5	< 0.5	< 0.5	< 0.5	$^{+5.0}_{-5.0}$
QCDscale_ggH1in	< 0.5	$^{+29.0}_{-22.0}$	< 0.5	< 0.5	< 0.5	< 0.5
QCDscale_ggH2in	< 0.5	$^{-5.0}_{+5.0}$	< 0.5	< 0.5	< 0.5	< 0.5
QCDscale_qqH	$^{+1.4}_{-1.4}$	< 0.5	< 0.5	< 0.5	< 0.5	< 0.5
pdf_Higgs_gg	< 0.5	$^{+8.0}_{-7.0}$	< 0.5	< 0.5	< 0.5	< 0.5
pdf_Higgs_qq	$^{+3.0}_{-3.0}$	< 0.5	$^{+3.0}_{-3.0}$	$^{+3.0}_{-3.0}$	< 0.5	< 0.5
pdf_qq	< 0.5	< 0.5	< 0.5	< 0.5	$^{+4.0}_{-4.0}$	$^{+4.0}_{-4.0}$

Table C.6: Summary of theoretical systematic uncertainties (except jet-related) in % for the signal and background components in the Boosted category. Numbers marked with < 0.5 are not considered in the fit model. For uncertainties marked with S, additional bin-by-bin uncertainties are taken into account as BDT score shape systematics.

Systematic	qqH	ggH	WH	ZH	$Z \to \tau\tau$	$Z \to \ell\ell(\ell \to \tau_h)$	$Z \to \ell\ell(j \to \tau_h)$	Top	$WW/WZ/ZZ$
BTag_BEFF	< 0.5	< 0.5 S	< 0.5	$^{-0.5}_{+0.6}$ S	< 0.5	< 0.5	< 0.5	$^{-9.6}_{+10.1}$ S	< 0.5
BTag_CEFF	< 0.5 S	< 0.5 S	$^{-1.2}_{+1.2}$ S	$^{-0.9}_{+0.9}$ S	< 0.5	< 0.5	< 0.5	$^{-0.5}_{+0.5}$	$^{-0.6}_{+0.6}$
EL_EFF	$^{+0.9}_{-0.9}$	$^{+0.9}_{-0.9}$	$^{+0.9}_{-0.9}$	$^{+0.9}_{-0.9}$	< 0.5	$^{+1.0}_{-1.0}$	$^{+1.1}_{-1.1}$	$^{+0.9}_{-0.9}$	$^{+0.8}_{-0.8}$
EL_EFF_Emb	< 0.5	< 0.5	< 0.5	< 0.5	$^{+0.6}_{-0.6}$	< 0.5	< 0.5	< 0.5	< 0.5
EL_RES	< 0.5	< 0.5	< 0.5	< 0.5	< 0.5	$^{+0.3}_{-0.8}$	$^{+0.0}_{+2.8}$	< 0.5	< 0.5
EL_SCALE	< 0.5	< 0.5	$^{+0.5}_{-0.7}$	< 0.5	$^{+2.3}_{-1.0}$	$^{+2.9}_{-1.8}$	$^{+2.1}_{-2.2}$	$^{+0.5}_{-0.5}$	$^{+0.3}_{-0.5}$
LUMI	$^{-2.8}_{+2.8}$	$^{-2.8}_{+2.8}$	$^{-2.8}_{+2.8}$	$^{-2.8}_{+2.8}$	< 0.5	$^{-2.8}_{+2.8}$	< 0.5	< 0.5	$^{-2.8}_{+2.8}$
MU_EFF	$^{+1.1}_{-1.1}$	$^{+1.0}_{-1.0}$	$^{+1.1}_{-1.1}$	$^{+1.1}_{-1.1}$	< 0.5	$^{+1.0}_{-1.0}$	$^{+0.8}_{-0.8}$	$^{+1.1}_{-1.1}$	$^{+1.1}_{-1.1}$
MU_EFF_Emb	< 0.5	< 0.5	< 0.5	< 0.5	$^{+0.6}_{-0.6}$	< 0.5	< 0.5	< 0.5	< 0.5
MU_SCALE	< 0.5	< 0.5	$^{+0.4}_{-0.6}$	$^{-0.6}_{+0.9}$	< 0.5	< 0.5	< 0.5	< 0.5	< 0.5
PU_RESCALE	< 0.5 S	$^{+0.6}_{-0.6}$ S	< 0.5 S	$^{+0.7}_{-0.7}$ S	< 0.5 S	$^{-0.3}_{+0.8}$ S	$^{-0.4}_{+0.7}$ S	< 0.5 S	< 0.5
TAU_EFAKE	< 0.5	< 0.5	< 0.5	< 0.5	< 0.5	$^{+14.2}_{-14.5}$ S	< 0.5	< 0.5	< 0.5
TAU_ID	$^{+3.3}_{-3.3}$	$^{+3.3}_{-3.3}$	$^{+3.3}_{-3.3}$	$^{+3.3}_{-3.3}$	$^{+3.2}_{-3.2}$	< 0.5	< 0.5	$^{+1.1}_{-1.1}$	$^{+1.6}_{-1.6}$
TAU_MFAKE	< 0.5	< 0.5	< 0.5	< 0.5	< 0.5	$^{+6.4}_{-6.4}$ S	< 0.5	< 0.5	< 0.5
TES_FAKE	< 0.5	< 0.5	< 0.5	< 0.5	< 0.5	$^{+6.3}_{-1.2}$	$^{+4.5}_{-2.7}$	$^{+3.5}_{-3.5}$	$^{+3.3}_{-3.8}$
TES_TRUE	$^{+1.1}_{-0.3}$	$^{+2.2}_{-2.1}$	$^{+0.8}_{-1.1}$	$^{+1.4}_{-0.8}$ S	$^{+1.2}_{-1.4}$ S	< 0.5	< 0.5	< 0.5	< 0.5

Table C.7: Summary of experimental systematic uncertainties (except jet-related) in % for the signal and background components in the Boosted category. Numbers marked with < 0.5 are not considered in the fit model. For uncertainties marked with S, additional bin-by-bin uncertainties are taken into account as BDT score shape systematics.

Systematic	qqH	ggH	WH	ZH	$Z \to \ell\ell(\ell \to \tau_h)$	$Z \to \ell\ell(j \to \tau_h)$	Top	$WW/WZ/ZZ$
JER	< 0.5 S	$^{+0.6}_{-0.6}$ S	$^{+1.0}_{-1.0}$ S	$^{-1.0}_{+1.0}$	$^{-0.5}_{+0.5}$	$^{-3.1}_{+3.1}$	$^{-1.9}_{+1.9}$	< 0.5
JES_Detector	< 0.5 S	$^{+0.8}_{-0.7}$ S	$^{+1.6}_{-1.1}$	$^{+1.0}_{-0.5}$	$^{+2.3}_{-1.5}$	$^{+2.6}_{-2.5}$	$^{+1.0}_{-0.7}$	$^{+0.5}_{-0.9}$
JES_Eta_StatMethod	< 0.5	$^{+0.9}_{-0.8}$	$^{+1.9}_{-1.0}$	< 0.5	< 0.5	$^{+3.9}_{-4.0}$	< 0.5	< 0.5
JES_Modelling	< 0.5 S	$^{+1.8}_{-2.0}$	$^{+1.2}_{-1.7}$	$^{+2.3}_{-1.7}$ S	$^{+5.9}_{-4.1}$	$^{+6.2}_{-6.8}$	$^{+2.0}_{-1.7}$	$^{+2.1}_{-2.0}$
JES_PileRho_TAU_GG	< 0.5	$^{+1.4}_{-0.7}$	< 0.5	< 0.5	< 0.5	< 0.5	< 0.5	< 0.5
JES_PileRho_TAU_QG	< 0.5	< 0.5	< 0.5	< 0.5	$^{+2.1}_{-2.9}$	$^{+3.1}_{-2.2}$	< 0.5	< 0.5
JES_PileRho_TAU_QQ	< 0.5	< 0.5	$^{+0.5}_{-0.8}$	$^{+0.8}_{-0.5}$ S	< 0.5	< 0.5	< 0.5	$^{+0.9}_{-0.9}$
JES_Statistical	< 0.5	$^{+0.2}_{-0.7}$	< 0.5	$^{+0.0}_{-0.9}$	$^{-0.9}_{+0.4}$	$^{+2.5}_{-2.3}$	< 0.5	< 0.5
JES_Eta_Modelling	< 0.5 S	$^{+1.4}_{-1.9}$ S	$^{+1.3}_{-1.9}$	$^{+1.6}_{-2.1}$	$^{+4.2}_{-4.0}$	$^{+2.8}_{-3.4}$	$^{+1.5}_{-1.5}$	$^{+2.2}_{-1.9}$
JES_FlavComp_TAU_G	< 0.5	$^{+2.3}_{-3.0}$ S	< 0.5	< 0.5	$^{+7.8}_{-7.0}$	$^{+4.9}_{-7.3}$	< 0.5	< 0.5
JES_FlavComp_TAU_Q	$^{+0.7}_{-0.8}$ S	< 0.5	$^{+2.2}_{-2.8}$ S	$^{+2.2}_{-1.7}$ S	< 0.5	< 0.5	$^{+2.5}_{-2.3}$	$^{+2.5}_{-2.9}$
JES_FlavResp	$^{+0.0}_{-0.6}$ S	$^{+1.7}_{-1.5}$	$^{+1.1}_{-1.8}$	$^{+1.6}_{-2.1}$	$^{+5.3}_{-5.0}$	$^{+4.8}_{-4.1}$	$^{+1.3}_{-1.5}$	$^{+1.6}_{-1.4}$
JES_Flavb	< 0.5	< 0.5	< 0.5	< 0.5	< 0.5	< 0.5	$^{+0.9}_{-0.1}$	< 0.5
JES_Mu	< 0.5	$^{-0.4}_{+0.5}$	< 0.5	$^{+0.6}_{-0.1}$	< 0.5	$^{+1.0}_{-0.7}$	$^{+0.7}_{-0.3}$	$^{+0.5}_{-0.5}$
JES_NPV	< 0.5	$^{+0.6}_{-0.7}$	$^{+0.0}_{-0.8}$	$^{+0.0}_{-0.7}$	$^{+1.4}_{-1.3}$	$^{+1.1}_{-2.1}$	< 0.5	$^{+0.7}_{-0.6}$
MET_RESOSOFT	< 0.5	$^{+0.0}_{+0.7}$	$^{+0.6}_{-0.3}$	$^{-0.2}_{+0.7}$	$^{+1.5}_{-0.2}$	< 0.5	$^{-0.7}_{+0.2}$	$^{-0.3}_{+0.7}$
MET_SCALESOFT	< 0.5	< 0.5	< 0.5	< 0.5	< 0.5	$^{-2.4}_{+1.4}$	< 0.5	< 0.5

Table C.8: Summary of jet-related experimental systematic uncertainties in % for the signal and background components in the Boosted category. Numbers marked with < 0.5 are not considered in the fit model. For uncertainties marked with S, additional bin-by-bin uncertainties are taken into account as BDT score shape systematics.

C.3.2 Systematic Uncertainties after the Final Fit

Nuisance Parameter	$\hat{\theta} \pm \Delta\hat{\theta}$	pre-fit $\Delta\hat{\mu}(\Delta\theta)$	post-fit $\Delta\hat{\mu}(\Delta\hat{\theta})$
MET_RESOSOFT	$0.14^{+0.99}_{-0.96}$	$^{+0.02}_{+0.03}$	$^{+0.02}_{+0.02}$
MET_SCALESOFT	$0.01^{+1.01}_{-1.01}$	< 0.01	< 0.01
JER	$-0.27^{+0.87}_{-0.82}$	$^{-0.02}_{+0.00}$	$^{-0.02}_{+0.00}$
JES_Detector	$0.00^{+0.92}_{-0.90}$	$^{-0.08}_{+0.10}$	$^{-0.07}_{+0.09}$
JES_Eta_StatMethod	$0.11^{+0.95}_{-0.88}$	$^{-0.02}_{+0.07}$	$^{-0.02}_{+0.05}$
JES_Modelling	$-0.13^{+0.97}_{-0.96}$	$^{-0.17}_{+0.15}$	$^{-0.14}_{+0.16}$
JES_PileRho_TAU_GG	$-0.02^{+0.99}_{-0.99}$	$^{-0.02}_{+0.01}$	$^{-0.02}_{+0.01}$
JES_PileRho_TAU_QG	$0.06^{+0.93}_{-0.88}$	$^{-0.00}_{+0.06}$	$^{-0.00}_{+0.05}$
JES_PileRho_TAU_QQ	$0.00^{+1.00}_{-1.00}$	$^{-0.02}_{+0.02}$	$^{-0.02}_{+0.02}$
JES_Statistical	$0.01^{+1.00}_{-1.00}$	$^{-0.01}_{+0.01}$	$^{-0.01}_{+0.01}$
JES_Eta_Modelling	$-0.02^{+0.99}_{-0.97}$	$^{-0.26}_{+0.28}$	$^{-0.25}_{+0.28}$
JES_Flavb	$0.02^{+1.00}_{-1.00}$	$^{+0.01}_{-0.01}$	$^{+0.01}_{-0.01}$
JES_FlavComp_TAU_G	$-0.13^{+0.98}_{-0.98}$	$^{-0.10}_{+0.11}$	$^{-0.09}_{+0.12}$
JES_FlavComp_TAU_Q	$0.03^{+1.00}_{-1.00}$	$^{-0.09}_{+0.09}$	$^{-0.09}_{+0.08}$
JES_FlavResp	$0.03^{+0.92}_{-0.89}$	$^{-0.10}_{+0.11}$	$^{-0.09}_{+0.09}$
JES_Mu	$-0.02^{+0.99}_{-0.99}$	< 0.01	< 0.01
JES_NPV	$-0.09^{+0.98}_{-0.99}$	$^{-0.03}_{+0.03}$	$^{-0.03}_{+0.02}$
JVF	$-0.01^{+1.00}_{-1.00}$	$^{-0.03}_{+0.00}$	$^{-0.03}_{+0.00}$

Table C.9: List of all jet- and $E_{\mathrm{T}}^{\mathrm{miss}}$-related nuisance parameters, their fitted value and uncertainty from the global fit and the pre- and post-fit impact on the uncertainty $\Delta\mu$ of the fitted signal strength $\hat{\mu}$.

Nuisance Parameter	$\hat{\theta} \pm \Delta\hat{\theta}$	pre-fit $\Delta\mu(\Delta\theta)$	post-fit $\Delta\hat{\mu}(\Delta\hat{\theta})$
EMB_ISOL	$0.57^{+0.82}_{-0.81}$	$^{+0.01}_{-0.05}$	$^{+0.02}_{-0.03}$
EMB_MFS	$0.83^{+0.86}_{-0.81}$	$^{-0.00}_{-0.18}$	$^{-0.01}_{+0.01}$
Fake_boost	$0.24^{+0.13}_{-0.15}$	$^{-0.38}_{+0.03}$	$^{-0.04}_{+0.05}$
Fake_boost_RW	$0.86^{+0.80}_{-0.81}$	$^{-0.03}_{+0.35}$	$^{-0.19}_{+0.17}$
Fake_btop	$-0.05^{+0.99}_{-0.99}$	< 0.01	< 0.01
Fake_vbf	$0.18^{+0.18}_{-0.17}$	$^{-0.47}_{+0.11}$	$^{-0.09}_{+0.04}$
Fake_vbf_RW	$0.02^{+0.91}_{-0.92}$	$^{-0.09}_{+0.04}$	$^{-0.08}_{+0.04}$
Fake_vtop	$0.06^{+1.00}_{-1.00}$	$^{+0.01}_{-0.00}$	$^{+0.01}_{-0.00}$
Zll_HERWIG	$0.08^{+0.99}_{-0.98}$	$^{-0.01}_{+0.01}$	$^{-0.01}_{+0.01}$
Zlt_vbf_DETAJJ	$-0.07^{+1.00}_{-1.00}$	$^{-0.02}_{+0.02}$	$^{-0.02}_{+0.02}$
BTag_BEFF	$-0.06^{+0.99}_{-0.99}$	$^{+0.04}_{-0.05}$	$^{+0.04}_{-0.05}$
BTag_CEFF	$-0.01^{+1.00}_{-1.00}$	< 0.01	< 0.01
BTag_LEFF	$0.00^{+1.00}_{-1.00}$	< 0.01	< 0.01
EL_EFF_Emb	$0.00^{+1.00}_{-1.00}$	< 0.01	< 0.01
EL_EFF	$-0.02^{+1.00}_{-1.00}$	$^{-0.03}_{+0.03}$	$^{-0.03}_{+0.03}$
EL_EFF_Emb	$0.00^{+1.00}_{-1.00}$	< 0.01	< 0.01
EL_RES	$0.10^{+0.87}_{-0.93}$	$^{+0.00}_{+0.04}$	$^{+0.00}_{+0.04}$
EL_SCALE	$-0.05^{+1.01}_{-1.01}$	$^{-0.04}_{+0.01}$	$^{-0.03}_{+0.01}$
LUMI	$-0.04^{+1.00}_{-1.00}$	$^{-0.10}_{+0.08}$	$^{-0.10}_{+0.08}$
MU_EFF_Emb	$-0.01^{+1.00}_{-1.00}$	< 0.01	< 0.01
MU_EFF	$0.00^{+1.00}_{-1.00}$	$^{-0.04}_{+0.03}$	$^{-0.04}_{+0.03}$
MU_EFF_Emb	$-0.01^{+1.00}_{-1.00}$	< 0.01	< 0.01
MU_SCALE	$0.00^{+1.00}_{-1.00}$	< 0.01	< 0.01
PU_RESCALE	$-0.23^{+0.98}_{-0.99}$	$^{-0.01}_{+0.01}$	$^{-0.01}_{+0.01}$
TAU_EFAKE	$-0.70^{+0.96}_{-0.95}$	$^{-0.91}_{+0.03}$	$^{-0.12}_{+0.09}$
TAU_ID	$-0.02^{+1.00}_{-1.00}$	$^{-0.09}_{+0.08}$	$^{-0.09}_{+0.08}$
TAU_MFAKE	$0.01^{+1.00}_{-1.00}$	$^{-0.02}_{+0.02}$	$^{-0.02}_{+0.02}$
TES_FAKE	$-0.22^{+0.91}_{-0.97}$	$^{-0.02}_{+0.03}$	$^{-0.02}_{+0.04}$
TES_TRUE	$-0.58^{+0.91}_{-0.72}$	$^{-0.25}_{+0.04}$	$^{-0.12}_{+0.07}$

Table C.10: List of all experimental nuisance parameters (except jet- and $E_{\mathrm{T}}^{\mathrm{miss}}$-related), their fitted value and uncertainty from the global fit and the pre- and post-fit impact on the uncertainty $\Delta\mu$ of the fitted signal strength $\hat{\mu}$.

Nuisance Parameter	$\hat{\theta} \pm \Delta\hat{\theta}$	pre-fit $\Delta\hat{\mu}(\Delta\theta)$	post-fit $\Delta\hat{\mu}(\Delta\hat{\theta})$
norm_Top_boost	$0.99^{+0.07}_{-0.06}$	$^{+0.24}_{-0.16}$	$^{+0.20}_{-0.20}$
norm_Top_vbf	$0.98^{+0.09}_{-0.09}$	$^{+0.30}_{-0.16}$	$^{+0.24}_{-0.20}$
norm_Zll_boost	$1.02^{+0.19}_{-0.15}$	$^{+0.28}_{-0.32}$	$^{+0.31}_{-0.29}$
norm_Zll_vbf	$1.01^{+0.25}_{-0.20}$	$^{+0.25}_{-0.29}$	$^{+0.27}_{-0.28}$
norm_Ztt	$0.96^{+0.06}_{-0.05}$	$^{+0.11}_{-0.02}$	$^{+0.06}_{-0.07}$
stat_boost_bin_11	$1.04^{+0.05}_{-0.05}$	$^{-0.01}_{+0.12}$	$^{-0.06}_{+0.06}$
stat_boost_bin_12	$0.95^{+0.05}_{-0.05}$	$^{-0.13}_{+0.01}$	$^{-0.07}_{+0.07}$
stat_boost_bin_13	$1.03^{+0.08}_{-0.07}$	$^{-0.08}_{+0.17}$	$^{-0.13}_{+0.12}$
stat_boost_bin_14	$0.90^{+0.10}_{-0.09}$	$^{-0.19}_{-0.01}$	$^{-0.10}_{+0.10}$
stat_vbf_bin_10	$1.01^{+0.05}_{-0.05}$	$^{-0.13}_{+0.17}$	$^{-0.16}_{+0.14}$
stat_vbf_bin_11	$1.01^{+0.15}_{-0.14}$	$^{-0.19}_{+0.17}$	$^{-0.20}_{+0.16}$
BR_tautau	$0.00^{+1.00}_{-1.00}$	$^{-0.15}_{+0.14}$	$^{-0.15}_{+0.14}$
Gen_Qmass_ggH	$-0.12^{+0.99}_{-0.97}$	$^{-0.91}_{+0.19}$	$^{-0.24}_{+0.21}$
pdf_Higgs_gg	$-0.03^{+1.00}_{-1.00}$	$^{-0.08}_{+0.06}$	$^{-0.07}_{+0.06}$
pdf_Higgs_qq	$0.01^{+1.00}_{-1.00}$	$^{-0.05}_{+0.05}$	$^{-0.05}_{+0.05}$
pdf_qq	$-0.06^{+1.00}_{-1.00}$	$^{-0.03}_{+0.02}$	$^{-0.03}_{+0.02}$
QCDscale_ggH1in	$-0.11^{+0.97}_{-0.97}$	$^{-0.19}_{+0.10}$	$^{-0.16}_{+0.10}$
QCDscale_ggH2in	$0.04^{+0.99}_{-1.00}$	$^{-0.07}_{+0.04}$	$^{-0.07}_{+0.04}$
QCDscale_ggH3in	$-0.01^{+1.00}_{-1.00}$	$^{+0.05}_{-0.06}$	$^{+0.05}_{-0.06}$
QCDscale_qqH	$0.01^{+1.00}_{-1.00}$	$^{-0.03}_{+0.03}$	$^{-0.03}_{+0.03}$
QCDscale_VH	$-0.01^{+1.00}_{-1.00}$	$^{-0.00}_{+0.01}$	$^{-0.00}_{+0.01}$
QCDscale_V	$-0.02^{+1.00}_{-1.00}$	$^{-0.00}_{+0.00}$	$^{-0.00}_{+0.00}$
QCDscale_VH	$-0.01^{+1.00}_{-1.00}$	$^{-0.00}_{+0.01}$	$^{-0.00}_{+0.01}$
QCDscale_VV	$0.02^{+1.00}_{-1.00}$	$^{-0.01}_{+0.01}$	$^{-0.01}_{+0.01}$
UE_gg	$0.02^{+0.99}_{-0.99}$	$^{-0.11}_{+0.10}$	$^{-0.11}_{+0.10}$
UE_qq	$0.04^{+1.00}_{-1.00}$	$^{-0.09}_{+0.08}$	$^{-0.09}_{+0.08}$

Table C.11: List of all normalisation factors, statistics- and theory-related nuisance parameters, their fitted value and uncertainty from the global fit and the pre- and post-fit impact on the uncertainty $\Delta\mu$ of the fitted signal strength $\hat{\mu}$.

Source of Uncertainty	Relative Uncertainty $\Delta\mu(\Delta\theta)/\Delta\hat{\mu}$ (post-fit)
norm_Zll_boost	0.31
JES_Eta_Modelling	0.28
norm_Zll_vbf	0.27
norm_Top_vbf	0.24
Gen_Qmass_ggH	0.21
norm_Top_boost	0.20
Fake_boost_RW	0.17
JES_Modelling	0.16
stat_vbf_bin_11	0.16
BR_tautau	0.14
stat_vbf_bin_10	0.14
JES_FlavComp_G	0.12
stat_boost_bin_13	0.12
QCDscale_ggH1in	0.10
stat_boost_bin_14	0.10
UE_gg	0.10
JES_FlavResp	0.09
TAU_EFAKE	0.09
JES_Detector	0.09
JES_FlavComp_Q	0.08
LUMI	0.08
TAU_ID	0.08
UE_qq	0.08
TES_TRUE	0.07
stat_boost_bin_12	0.07
norm_Ztt	0.06
stat_boost_bin_11	0.06
pdf_Higgs_gg	0.06
JES_Eta_StatMethod	0.05
QCDscale_ggH3in	0.05
JES_PileRho_QG	0.05
Fake_boost	0.05
pdf_Higgs_qq	0.05

Table C.12: The important sources ($\geq 5\%$) of uncertainty on the measured signal strength parameter $\hat{\mu}$, given as relative uncertainties in units of $\Delta\hat{\mu}$.

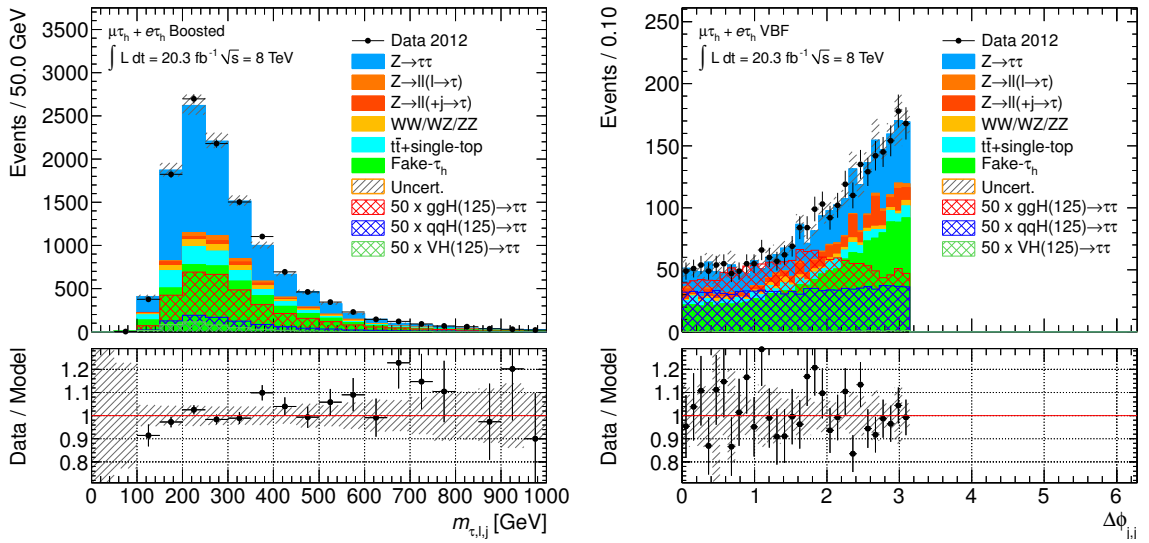
Source of Uncertainty	Relative Uncertainty $\Delta\mu(\Delta\theta)/\Delta\hat{\mu}$ (post-fit)
TES_FAKE	0.04
EL_RES	0.04
QCDscale_ggH2in	0.04
Fake_vbf_RW	0.04
Fake_vbf	0.04
BTag_BEFF	0.04
MU_EFF	0.03
QCDscale_qqH	0.03
EL_EFF	0.03
JES_PileRho_QQ	0.02
pdf_qq	0.02
MET_RESOSOFT	0.02
EMB_ISOL	0.02
JES_NPV	0.02
Zlt_vbf_DETAJJ	0.02
TAU_MFAKE	0.02
JES_Statistical	0.01
EL_SCALE	0.01
Fake_vtop	0.01
QCDscale_VV	0.01
JES_PileRho_GG	0.01
PU_RESCALE	0.01
Zll_HERWIG	0.01
JES_Flavb	0.01
EMB_MFS	0.01
QCDscale_VH	0.01
QCDscale_V	<0.01
MET_SCALESOFT	<0.01
JER	<0.00
Fake_btop	<0.01
BTag_CEFF	<0.01
MU_EFF_Emb	<0.01
MU_SCALE	<0.01
BTag_LEFF	<0.01
EL_EFF_Emb	<0.01
JVF	<0.01
JES_Mu	<0.01

Table C.13: Other sources of uncertainty (< 5%) on the measured signal strength parameter $\hat{\mu}$, given as relative uncertainties in units of $\Delta\hat{\mu}$.

APPENDIX D

Studies Towards a Mass Measurement

D.1 Control Distributions



(a) Boosted (b) VBF

Figure D.1: Distributions of the jet-related input observables for the training of the de-correlated BDTs in the (a) Boosted and (b) VBF category. The background components are estimated with the techniques described in section 7.4.

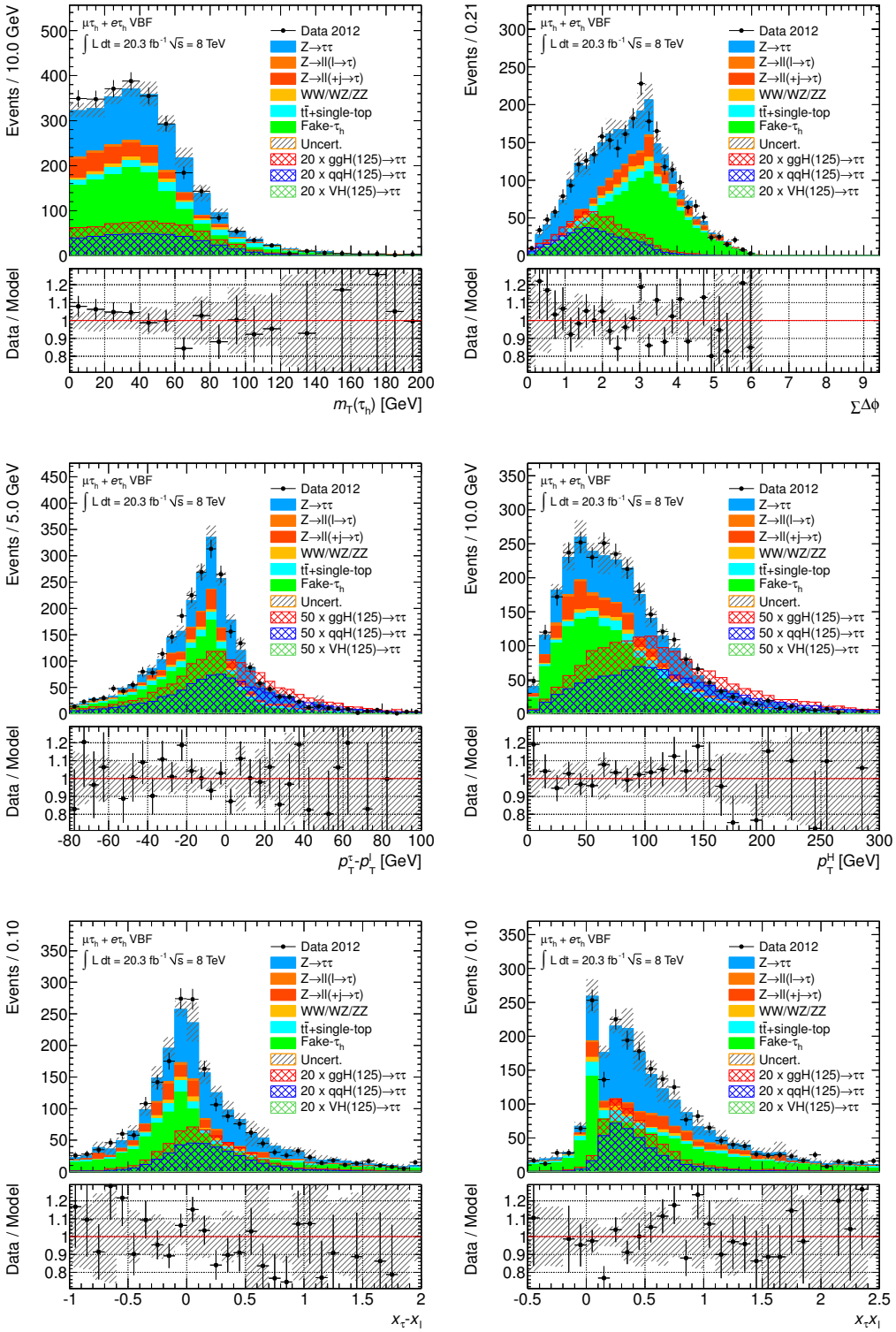


Figure D.2: Distributions of the input observables for the training of the de-correlated BDTs in VBF category. The background components are estimated with the techniques described in section 7.4.

Bibliography

[1] P. W. Higgs, "Broken symmetries, massless particles and gauge fields", *Phys.Lett.* 12 (1964) 132–133, DOI: 10.1016/0031-9163(64)91136-9.

[2] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons", *Phys.Rev.Lett.* 13 (1964) 321–323, DOI: 10.1103/PhysRevLett.13.321.

[3] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons", *Phys.Rev.Lett.* 13 (1964) 508–509, DOI: 10.1103/PhysRevLett.13.508.

[4] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, "Global conservation laws and massless particles", *Phys. Rev. Lett.* 13 (1964) 585–587, DOI: 10.1103/PhysRevLett.13.585.

[5] The LHC Study Group, "The Large Hadron Collider: Conceptual design" (1995).

[6] G. Aad et al., "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys.Lett.* B716 (2012) 1–29, DOI: 10.1016/j.physletb.2012.08.020, eprint: 1207.7214.

[7] S. Chatrchyan et al., "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys.Lett.* B716 (2012) 30–61, DOI: 10.1016/j.physletb.2012.08.021, eprint: 1207.7235.

[8] G. Aad et al., "Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC", *Phys.Lett.* B726 (2013) 88–119, DOI: 10.1016/j.physletb.2013.08.010, eprint: 1307.1427.

[9] L. Landau, "On the angular momentum of a two-photon system", *Dokl.Akad.Nauk Ser.Fiz.* 60 (1948) 207–209.

[10] C.-N. Yang, "Selection Rules for the Dematerialization of a Particle Into Two Photons", *Phys.Rev.* 77 (1950) 242–245, DOI: 10.1103/PhysRev.77.242.

[11] G. Aad et al., "Evidence for the spin-0 nature of the Higgs boson using ATLAS data", *Phys.Lett.* B726 (2013) 120–144, DOI: 10.1016/j.physletb.2013.08.026, eprint: 1307.1432.

[12] The Nobel Foundation, "The Nobel Prize in Physics 2013", *Nobelprize.org. Nobel Media AB 2013. Web. 22 Jun 2014.* (2013), URL: http://www.nobelprize.org/nobel_prizes/physics/laureates/2013/.

[13] M. Baak et al.,
"The Electroweak Fit of the Standard Model after the Discovery of a New Boson at the LHC",
Eur.Phys.J. C72 (2012) 2205, DOI: 10.1140/epjc/s10052-012-2205-9,
eprint: 1209.2716.

[14] J. Beringer et al., "Review of Particle Physics (RPP)", *Phys.Rev.* D86 (2012) 010001,
DOI: 10.1103/PhysRevD.86.010001.

[15] J. Thomson, "Cathode rays", *Phil.Mag.* 44 (1897) 293–316.

[16] H. Geiger and E Marsden, "On a Diffuse Reflection of the α-Particles",
Proc. R. Soc. Lond. A 82 (1909) 495–500.

[17] E. Rutherford,
"The scattering of alpha and beta particles by matter and the structure of the atom",
Phil.Mag. 21 (1911) 669–688.

[18] S. P. Martin, "A Supersymmetry Primer" (1997), eprint: hep-ph/9709356.

[19] D. I. Kazakov, "Beyond the standard model (in search of supersymmetry)" (2000),
eprint: hep-ph/0012288.

[20] N. Cabibbo, "Unitary Symmetry and Leptonic Decays", *Phys.Rev.Lett.* 10 (1963) 531–533,
DOI: 10.1103/PhysRevLett.10.531.

[21] M. Kobayashi and T. Maskawa,
"CP Violation in the Renormalizable Theory of Weak Interaction", *Prog.Theor.Phys.* 49 (1973)
652–657, DOI: 10.1143/PTP.49.652.

[22] H. P. Nilles, "Supersymmetry, Supergravity and Particle Physics", *Phys.Rept.* 110 (1984)
1–162, DOI: 10.1016/0370-1573(84)90008-5.

[23] T. W. B. Kibble, "Symmetry breaking in non-Abelian gauge theories", *Phys. Rev.* 155 (1967)
1554–1561, DOI: 10.1103/PhysRev.155.1554.

[24] J. Goldstone, "Field Theories with Superconductor Solutions", *Nuovo Cim.* 19 (1961)
154–164, DOI: 10.1007/BF02812722.

[25] L. Alvarez-Gaume, and J. Ellis, "Eyes on a prize particle", *Nature Physics* 7 (2011) 2–3,
DOI: 10.1038/nphys1874.

[26] F. Halzen and A. D. Martin,
Quarks and Leptons: An Introductory Course in Modern Particle Physics, 1984,
ISBN: 9780471887416.

[27] S. Bethke, "World Summary of α_s (2012)", *Nucl.Phys.Proc.Suppl.* 234 (2013) 229–234,
DOI: 10.1016/j.nuclphysbps.2012.12.020, eprint: 1210.0325.

[28] G. Hinshaw et al., "Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP)
Observations: Cosmological Parameter Results", *Astrophys.J.Suppl.* 208 (2013) 19,
DOI: 10.1088/0067-0049/208/2/19, eprint: 1212.5226.

[29] V. Hess, "Über Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten",
Physikalische Zeitschrift 13 (1912) 1084–1091.

[30] E. O. Lawrence and M. S. Livingston,
"The production of high speed protons without the use of high voltages", *Phys.Rev.* 38 (1931)
834, DOI: 10.1103/PhysRev.38.834.

[31] O. S. Bruning et al., "LHC Design Report. 1. The LHC Main Ring", *CERN-2004-003-V-1, CERN-2004-003* (2004), ed. by O. S. Bruning.

[32] O. Buning et al., "LHC Design Report. 2. The LHC infrastructure and general services", *CERN-2004-003-V-2, CERN-2004-003* (2004), ed. by O. Buning.

[33] M. Benedikt et al., "LHC Design Report. 3. The LHC injector chain", *CERN-2004-003-V-3, CERN-2004-003* (2004), ed. by M. Benedikt.

[34] R. K. Ellis, W. J. Stirling and B. R. Webber, "QCD and Collider Physics", *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.* 8 (1996).

[35] A. D. Martin et al., "MRST2001: Partons and α_s from precise deep inelastic scattering and Tevatron jet data", *Eur.Phys.J.* C23 (2002) 73–87, DOI: 10.1007/s100520100842, eprint: hep-ph/0110215.

[36] A. D. Martin et al., "Parton distributions for the LHC", *Eur. Phys. J. C* 63 (2009) 189–285, DOI: 10.1140/epjc/s10052-009-1072-5, eprint: 0901.0002.

[37] J. W. Stirling, *private communication* (May 2014), URL: http://www.hep.ph.ic.ac.uk/~wstirlin/plots/plots.html.

[38] A. Djouadi, M. Spira and P. M. Zerwas, "Production of Higgs bosons in proton colliders: QCD corrections", *Phys. Lett. B* 264 (1991) 440–446, DOI: 10.1016/0370-2693(91)90375-Z.

[39] S. Dawson, "Radiative corrections to Higgs boson production", *Nucl. Phys. B* 359 (1991) 283–300, DOI: 10.1016/0550-3213(91)90061-2.

[40] M. Spira et al., "Higgs boson production at the LHC", *Nucl. Phys. B* 453 (1995) 17–82, DOI: 10.1016/0550-3213(95)00379-7, eprint: 9504378.

[41] R. Harlander and W. B. Kilgore, "Next-to-next-to-leading order Higgs production at hadron colliders", *Phys. Rev. Lett.* 88 (2002) 201801, DOI: 10.1103/PhysRevLett.88.201801, eprint: 0201206.

[42] C. Anastasiou and K. Melnikov, "Higgs boson production at hadron colliders in NNLO QCD", *Nucl. Phys. B* 646 (2002) 220, DOI: 10.1016/S0550-3213(02)00837-4, eprint: 0207004.

[43] V. Ravindran, J. Smith and W. L. van Neerven, "NNLO corrections to the total cross section for Higgs boson production in hadron hadron collisions", *Nucl. Phys. B* 665 (2003) 325, DOI: 10.1016/S0550-3213(03)00457-7, eprint: 0302135.

[44] S. Catani et al., "Soft-gluon re-summation for Higgs boson production at hadron colliders", *JHEP* 0307 (2003) 028, eprint: 0306211.

[45] U. Aglietti, R. Bonciani, G. Degrassi, and A. Vicini, "Two loop light fermion contribution to Higgs production and decays", *Phys. Lett. B* 595 (2004) 432, DOI: 10.1016/j.physletb.2004.06.063, eprint: 0404071.

[46] S. Actis et al., "NLO Electroweak Corrections to Higgs Boson Production at Hadron Colliders", *Phys. Lett. B* 670 (2008) 12–17, DOI: 10.1016/j.physletb.2008.10.018, eprint: 0809.1301.

[47] C. Anastasiou, R. Boughezal, and F. Petriello,
"Mixed QCD-electroweak corrections to Higgs boson production in gluon fusion",
JHEP 0904 (2009) 003, DOI: 10.1088/1126-6708/2009/04/003, eprint: 0811.3458.

[48] D. de Florian and M. Grazzini,
"Higgs production through gluon fusion: updated cross sections at the Tevatron and the LHC",
Phys. Lett. B 674 (2009) 291, DOI: 10.1016/j.physletb.2009.03.033, eprint: 0901.2427.

[49] J. Baglio and A. Djouadi, "Higgs production at the lHC", *JHEP* 1103 (2011) 055,
DOI: 10.1007/JHEP03(2011)055, eprint: 1012.0530.

[50] T. Han and S. Willenbrock, "QCD correction to the p p → W H and Z H total cross- sections",
Phys. Lett. B 273 (1991) 167–172, DOI: 10.1016/0370-2693(91)90572-8.

[51] O. Brein, A. Djouadi, and R. Harlander,
"NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders",
Phys. Lett. B 579 (2004) 149, DOI: 10.1016/j.physletb.2003.10.112, eprint: 0307206.

[52] M. L. Ciccolini, S. Dittmaier, and M. Kramer,
"Electroweak radiative corrections to associated W H and Z H production at hadron colliders",
Phys. Rev. D 68 (2003) 073003, DOI: 10.1103/PhysRevD.68.073003, eprint: 0306234.

[53] M. Ciccolini, A. Denner and S. Dittmaier, "Strong and electroweak corrections to the production of Higgs+2jets via weak interactions at the LHC", *Phys. Rev. Lett.* 99 (2007) 161803, DOI: 10.1103/PhysRevLett.99.161803, eprint: 0707.0381.

[54] M. Ciccolini, A. Denner and S. Dittmaier,
"Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC",
Phys. Rev. D 77 (2008) 013002, DOI: 10.1103/PhysRevD.77.013002, eprint: 0710.4749.

[55] K. Arnold et al.,
"VBFNLO: A Parton level Monte Carlo for processes with electroweak bosons",
Comput. Phys. Commun. 180 (2009) 1661–1670, DOI: 10.1016/j.cpc.2009.03.006,
eprint: 0811.4559.

[56] P. Bolzoni, F. Maltoni, S.-O. Moch, and M. Zaro,
"Higgs production via vector-boson fusion at NNLO in QCD", *Phys. Rev. Lett.* 105 (2010) 011801, DOI: 10.1103/PhysRevLett.105.011801, eprint: 1003.4451.

[57] LHC Higgs Cross Section Working Group et al.,
"Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables",
CERN-2011-002 (CERN, Geneva, 2011), eprint: 1101.0593.

[58] LHC Higgs Cross Section Working Group et al.,
"Handbook of LHC Higgs Cross Sections: 2. Differential Distributions",
CERN-2012-002 (CERN, Geneva, 2012), eprint: 1201.3084.

[59] LHC Higgs Cross Section Working Group et al.,
"Handbook of LHC Higgs Cross Sections: 3. Higgs Properties",
CERN-2013-004 (CERN, Geneva, 2013), eprint: 1307.1347.

[60] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language", *Nucl.Phys.* B126 (1977) 298, DOI: 10.1016/0550-3213(77)90384-4.

[61] V. Sudakov, "Vertex parts at very high-energies in quantum electrodynamics",
Sov.Phys.JETP 3 (1956) 65–71.

[62] B. Andersson et al., "Parton Fragmentation and String Dynamics", *Phys.Rept.* 97 (1983) 31–145, DOI: 10.1016/0370-1573(83)90080-7.

[63] B. Webber, "A QCD Model for Jet Fragmentation Including Soft Gluon Interference", *Nucl.Phys.* B238 (1984) 492, DOI: 10.1016/0550-3213(84)90333-X.

[64] T. Sjostrand, "A Model for Initial State Parton Showers", *Phys.Lett.* B157 (1985) 321, DOI: 10.1016/0370-2693(85)90674-4.

[65] Jean-Luc Caron, "Overall view of LHC experiments", *LHC-PHO-1998-349* (1998), URL: https://cds.cern.ch/record/841555.

[66] M. Lamont, "Status of the LHC", *J.Phys.Conf.Ser.* 455 (2013) 012001, DOI: 10.1088/1742-6596/455/1/012001.

[67] G. Aad et al., "Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC", *Eur.Phys.J.* C73 (2013) 2518, DOI: 10.1140/epjc/s10052-013-2518-3, eprint: 1302.4393.

[68] The ATLAS Collaboration, "Luminosity Public Results" (May 2014), URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults.

[69] G. Aad et al., "The ATLAS Experiment at the CERN Large Hadron Collider", *JINST* 3 (2008) S08003, DOI: 10.1088/1748-0221/3/08/S08003.

[70] G. Aad et al., "Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics" (2009), eprint: 0901.0512.

[71] S. Agostinelli et al., "GEANT4: A Simulation toolkit", *Nucl.Instrum.Meth.* A506 (2003) 250–303, DOI: 10.1016/S0168-9002(03)01368-8.

[72] G. Aad et al., "The ATLAS Simulation Infrastructure", *Eur.Phys.J.* C70 (2010) 823–874, DOI: 10.1140/epjc/s10052-010-1429-9, eprint: 1005.4568.

[73] G. Aad et al., "Performance of the ATLAS Trigger System in 2010", *Eur.Phys.J.* C72 (2012) 1849, DOI: 10.1140/epjc/s10052-011-1849-1, eprint: 1110.1530.

[74] The ATLAS Collaboration, "Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment", *ATLAS-CONF-2012-042* (2012).

[75] W. Lampl et al., "Calorimeter clustering algorithms: Description and performance", *ATL-LARG-PUB-2008-002, ATL-COM-LARG-2008-003* (2008).

[76] G. Aad et al., "Measurement of the $W \to \ell\nu$ and $Z/\gamma^* \to \ell\ell$ production cross sections in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector", *JHEP* 1012 (2010) 060, DOI: 10.1007/JHEP12(2010)060, eprint: 1010.2130.

[77] The ATLAS Collaboration, "Preliminary results on the muon reconstruction efficiency, momentum resolution, and momentum scale in ATLAS 2012 pp collision data", *ATLAS-CONF-2013-088* (2013), URL: http://cds.cern.ch/record/1580207.

[78] G. Aad et al., "Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data", *Eur.Phys.J.* C72 (2012) 1909, DOI: 10.1140/epjc/s10052-012-1909-1, eprint: 1110.3174.

[79] M. Cacciari, G. P. Salam and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* 04 (2008) 063, DOI: 10.1088/1126-6708/2008/04/063, eprint: 0802.1189.

[80] M. Cacciari and G. P. Salam, "Dispelling the N^3 myth for the k_t jet-finder", *Phys. Lett. B* 641 (2006) 57, DOI: 10.1016/j.physletb.2006.08.037, eprint: hep-ph/0512210.

[81] T. Barillari et al., "Local hadronic calibration", *ATL-LARG-PUB-2009-001-2* (2009).

[82] G. Aad et al., "Single hadron response measurement and calorimeter jet energy scale uncertainty with the ATLAS detector at the LHC", *Eur.Phys.J.* C73 (2013) 2305, DOI: 10.1140/epjc/s10052-013-2305-1, eprint: 1203.1302.

[83] G. Aad et al., "Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV", *Eur.Phys.J.* C73 (2013) 2304, DOI: 10.1140/epjc/s10052-013-2304-2, eprint: 1112.6426.

[84] S. Psoroulas, "Measurement of jets production in association with a Z boson and in the search for the SM Higgs boson via $H \to \tau\tau \to \ell\ell + 4\nu$ with ATLAS", PhD thesis: Rheinische Friedrich-Wilhelms Universität Bonn, 2013, URL: http://hss.ulb.uni-bonn.de/2013/3130/3130.pdf.

[85] The ATLAS Collaboration, "Performance of the ATLAS Secondary Vertex b-tagging Algorithm in 7 TeV Collision Data", *ATLAS-CONF-2010-042* (2010), URL: https://cds.cern.ch/record/1277682.

[86] The ATLAS Collaboration, "Performance of Impact Parameter-Based b-tagging Algorithms with the ATLAS Detector using Proton-Proton Collisions at $\sqrt{s} = 7$ TeV", *ATLAS-CONF-2010-091* (2010), URL: https://cds.cern.ch/record/1299106.

[87] The ATLAS Collaboration, "Performance of the Reconstruction and Identification of Hadronic Tau Decays in ATLAS with 2011 Data", *ATLAS-CONF-2012-142* (2012).

[88] The ATLAS Collaboration, "Determination of the tau energy scale and the associated systematic uncertainty in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector at the LHC in 2012", *ATLAS-CONF-2013-044* (2013), URL: https://cds.cern.ch/record/1544036.

[89] The ATLAS Collaboration, "Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons", *ATLAS-CONF-2011-077* (2011), URL: https://cds.cern.ch/record/1353226.

[90] G. Aad et al., "Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at $\sqrt{s}$=8 TeV" (2014), eprint: 1412.7086.

[91] The ATLAS Collaboration, "Performance of Missing Transverse Momentum Reconstruction in ATLAS with 2011 Proton-Proton Collisions at $\sqrt{s} = 7$ TeV", *ATLAS-CONF-2012-101* (2012), URL: http://cds.cern.ch/record/1463915.

[92] G. Aad et al., "Performance of Missing Transverse Momentum Reconstruction in Proton-Proton Collisions at 7 TeV with ATLAS", *Eur.Phys.J.* C72 (2012) 1844, DOI: 10.1140/epjc/s10052-011-1844-6, eprint: 1108.5602.

[93] The ATLAS Collaboration, "Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at 8 TeV", *ATLAS-CONF-2013-082* (2013), URL: https://cds.cern.ch/record/1570993.

[94] G. Cowan et al., "Asymptotic formulae for likelihood-based tests of new physics", *Eur.Phys.J.* C71 (2011) 1554, DOI: 10.1140/epjc/s10052-011-1554-0, eprint: 1007.1727.

[95] D. L. Rainwater, D. Zeppenfeld and K. Hagiwara, "Searching for $H \to \tau^+\tau^-$ in weak boson fusion at the CERN LHC", *Phys.Rev.* D59 (1998) 014037, DOI: 10.1103/PhysRevD.59.014037, eprint: hep-ph/9808468.

[96] V. Khoze et al., "A Z monitor to calibrate Higgs production via vector boson fusion with rapidity gaps at the LHC", *Eur.Phys.J.* C26 (2003) 429–440, DOI: 10.1140/epjc/s2002-01069-2, eprint: hep-ph/0207365.

[97] S. Chatrchyan et al., "Measurement of the hadronic activity in events with a Z and two jets and extraction of the cross section for the electroweak production of a Z with two jets in pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* 1310 (2013) 062, DOI: 10.1007/JHEP10(2013)062, eprint: 1305.7389.

[98] T. Sjostrand, S. Mrenna and P. Z. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* 0605 (2006) 026, DOI: 10.1088/1126-6708/2006/05/026, eprint: hep-ph/0603175.

[99] J. Butterworth, J. R. Forshaw and M. Seymour, "Multiparton interactions in photoproduction at HERA", *Z.Phys.* C72 (1996) 637–646, DOI: 10.1007/s002880050286, eprint: hep-ph/9601371.

[100] G. Corcella et al., "HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)", *JHEP* 0101 (2001) 010, DOI: 10.1088/1126-6708/2001/01/010, eprint: hep-ph/0011363.

[101] G. Corcella et al., "HERWIG 6.5 release note" (2002), eprint: hep-ph/0210213.

[102] K. Melnikov and F. Petriello, "Electroweak gauge boson production at hadron colliders through $O(\alpha_s^2)$", *Phys. Rev. D* 74 (2006) 114017, DOI: 10.1103/PhysRevD.74.114017, eprint: 0609070.

[103] S. Catani et al., "Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO", *Phys.Rev.Lett.* 103 (2009) 082001, DOI: 10.1103/PhysRevLett.103.082001, eprint: 0903.2120.

[104] R. Gavin et al., "FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order", *Comput.Phys.Commun.* 182 (2011) 2388–2403, DOI: 10.1016/j.cpc.2011.06.008, eprint: 1011.3540.

[105] M. L. Mangano et al., "ALPGEN, a generator for hard multiparton processes in hadronic collisions", *JHEP* 0307 (2003) 001, DOI: 10.1088/1126-6708/2003/07/001, eprint: hep-ph/0206293.

[106] S. Hoeche et al., "Matching parton showers and matrix elements" (2006), eprint: hep-ph/0602031.

[107] J. Alwall et al., "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions", *Eur.Phys.J.* C53 (2008) 473–500, DOI: 10.1140/epjc/s10052-007-0490-5, eprint: 0706.2569.

[108] T. Gleisberg et al., "Event generation with SHERPA 1.1", *JHEP* 0902 (2009) 007, DOI: 10.1088/1126-6708/2009/02/007, eprint: 0811.4622.

[109] R. Hamberg, W. van Neerven and T. Matsuura,
"A Complete calculation of the order α_s^2 correction to the Drell-Yan K factor",
Nucl.Phys. B359 (1991) 343–405, DOI: 10.1016/0550-3213(91)90064-5.

[110] S. Frixione and B. R. Webber,
"Matching NLO QCD computations and parton shower simulations", *JHEP* 06 (2002) 029,
eprint: hep-ph/0204244.

[111] S. Alioli, P. Nason, and C. Oleari, and E. Re,
"NLO Higgs boson production via gluon fusion matched with shower in POWHEG",
JHEP 0904 (2009) 002, eprint: 0812.0578.

[112] P. Nason and C. Oleari,
"NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG",
JHEP 1002 (2010) 037, eprint: 0911.5299.

[113] M. Cacciari et al., "Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation", *Phys.Lett.* B710 (2012) 612–622,
DOI: 10.1016/j.physletb.2012.03.013, eprint: 1111.5869.

[114] M. Czakon and A. Mitov,
"Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders" (2011), eprint: 1112.5675.

[115] M. Czakon, P. Fiedler and A. Mitov,
"Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$",
Phys.Rev.Lett. 110.25 (2013) 252004, DOI: 10.1103/PhysRevLett.110.252004,
eprint: 1303.6254.

[116] P. Bärnreuther, M. Czakon and P. Fiedler, "Virtual amplitudes and threshold behaviour of hadronic top-quark pair-production cross sections", *JHEP* 1402 (2014) 078,
DOI: 10.1007/JHEP02(2014)078, eprint: 1312.6279.

[117] J. M. Campbell, R. K. Ellis and C. Williams, "Vector boson pair production at the LHC", *JHEP* 1107 (2011) 018, DOI: 10.1007/JHEP07(2011)018, eprint: 1105.0020.

[118] B. P. Kersevan and E. Richter-Was, "The Monte Carlo event generator AcerMC version 2.0 with interfaces to PYTHIA 6.2 and HERWIG 6.5" (2004), eprint: 0405247.

[119] N. Kidonakis, "NNLL resummation for s-channel single top quark production",
Phys. Rev. D 81 (2010) 054028, DOI: 10.1103/PhysRevD.81.054028, eprint: 1001.5034.

[120] N. Kidonakis, "Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-", *Phys.Rev.* D82 (2010) 054018,
DOI: 10.1103/PhysRevD.82.054018, eprint: 1005.4451.

[121] N. Kidonakis, "Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production", *Phys. Rev. D* 83 (2011) 091503,
DOI: 10.1103/PhysRevD.83.091503, eprint: 1103.2792.

[122] T. Binoth et al., "Gluon-induced W-boson pair production at the LHC", *JHEP* 0612 (2006) 046, DOI: 10.1088/1126-6708/2006/12/046, eprint: hep-ph/0611170.

[123] D. de Florian et al.,
"Transverse-momentum resummation: Higgs boson production at the Tevatron and the LHC",
JHEP 1111 (2011) 064, DOI: 10.1007/JHEP11(2011)064, eprint: 1109.2109.

[124] S. Jadach et al., "The tau decay library TAUOLA: Version 2.4", *Comput. Phys. Commun.* 76 (1993) 361–380, DOI: 10.1016/0010-4655(93)90061-G.

[125] P. Golonka and Z. Was, "PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays", *Eur. Phys. J. C* 45 (2006) 97–107, DOI: 10.1140/epjc/s2005-02396-4, eprint: 0506026.

[126] H.-L. Lai et al., "New parton distributions for collider physics", *Phys.Rev.* D82 (2010) 074024, DOI: 10.1103/PhysRevD.82.074024, eprint: 1007.2241.

[127] J. Pumplin et al., "New generation of parton distributions with uncertainties from global QCD analysis", *JHEP* 0207 (2002) 012, DOI: 10.1088/1126-6708/2002/07/012, eprint: hep-ph/0201195.

[128] A. Sherstnev and R. Thorne, "Parton Distributions for LO Generators", *Eur.Phys.J.* C55 (2008) 553–575, DOI: 10.1140/epjc/s10052-008-0610-x, eprint: 0711.2473.

[129] P. M. Nadolsky et al., "Implications of CTEQ global analysis for collider observables", *Phys.Rev.* D78 (2008) 013004, DOI: 10.1103/PhysRevD.78.013004, eprint: 0802.0007.

[130] R. K. Ellis et al., "Higgs Decay to tau+ tau-: A Possible Signature of Intermediate Mass Higgs Bosons at the SSC", *Nucl.Phys.* B297 (1988) 221, DOI: 10.1016/0550-3213(88)90019-3.

[131] A. Elagin et al., "A New Mass Reconstruction Technique for Resonances Decaying to di-tau", *Nucl.Instrum.Meth.* A654 (2011) 481–489, DOI: 10.1016/j.nima.2011.07.009, eprint: 1012.4686.

[132] The ATLAS Collaboration, "Search for neutral MSSM Higgs bosons decaying to $\tau^+\tau^-$ pairs in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector", *ATLAS-CONF-2011-132* (2011), URL: http://cds.cern.ch/record/1383835.

[133] C. Boddy et al., "Measurement of $Z \to \tau\tau$ production cross-section in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector - Support Note for lep-had channels", *ATL-COM-PHYS-2011-416* (2011), URL: https://cds.cern.ch/record/1346327.

[134] A. Wald, "Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large", *Transactions of the American Mathematical Society* 54.3 (1943) 426–482.

[135] S. Wilks, "The large-sample distribution of the likelihood ratio for testing composite hypotheses", *Ann. Math. Statist.* 9 (1938) 60–62.

[136] R. Brun and F. Rademakers, "ROOT: An object oriented data analysis framework", *Nucl.Instrum.Meth.* A389 (1997) 81–86, DOI: 10.1016/S0168-9002(97)00048-X.

[137] W. Verkerke and D. P. Kirkby, "The RooFit toolkit for data modeling", *eConf* C0303241 (2003) MOLT007, eprint: physics/0306116.

[138] L. Moneta et al., "The RooStats Project", *PoS* ACAT2010 (2010) 057, eprint: 1009.1003.

[139] F. James and M. Roos, "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations", *Comput.Phys.Commun.* 10 (1975) 343–367, DOI: 10.1016/0010-4655(75)90039-9.

[140] G. Aad et al., "Measurement of the Z to tau tau Cross Section with the ATLAS Detector", *Phys.Rev.* D84 (2011) 112006, DOI: 10.1103/PhysRevD.84.112006, eprint: 1108.2016.

[141] The ATLAS Collaboration, "Muon reconstruction efficiency in reprocessed 2010 LHC proton-proton collision data recorded with the ATLAS detector", *ATLAS-CONF-2011-063* (2011), URL: http://cds.cern.ch/record/1345743.

[142] The ATLAS Collaboration, "Measurement of the Mis-identification Probability of τ Leptons from Hadronic Jets and from Electrons" (2011), URL: http://cds.cern.ch/record/1375550.

[143] K. Nakamura et al., "Review of particle physics", *J.Phys.* G37 (2010) 075021, DOI: 10.1088/0954-3899/37/7A/075021.

[144] M. Schmitz, "Sensitivity of the ATLAS Experiment to discover the Decay $H \to \tau\tau \to \ell\ell + 4\nu$ of the Standard Model Higgs Boson produced in Vector Boson Fusion", PhD thesis: Rheinische Friedrich-Wilhelms Universität Bonn, 2011, URL: http://hss.ulb.uni-bonn.de/2011/2554/2554.pdf.

[145] N. Möser, "A Sensitivity Study for Higgs Boson Production in Vector Boson Fusion in the $H \to \tau\tau \to \ell h + 3\nu$ Final State with ATLAS", PhD thesis: Rheinische Friedrich-Wilhelms Universität Bonn, 2012, URL: http://hss.ulb.uni-bonn.de/2012/2776/2776.pdf.

[146] G. Altarelli, R. Kleiss and C. Verzegnassi, "Z Physics at LEP-1. Proceedings, Workshop, Geneva, Switzerland, September 4-5, 1989. Vol. 3: Event Generators and Software" (1989), ed. by G. Altarelli.

[147] Z. Czyczula, T. Przedzinski and Z. Was, "TauSpinner Program for Studies on Spin Effect in tau Production at the LHC", *Eur.Phys.J.* C72 (2012) 1988, DOI: 10.1140/epjc/s10052-012-1988-z, eprint: 1201.0117.

[148] D. Alvarez et al., "Search for the Standard Model Higgs boson $H \to \tau\tau$ decays with the ATLAS detector", *ATL-COM-PHYS-2013-722* (2013), URL: https://cds.cern.ch/record/1553809.

[149] G. Aad et al., "Search for neutral MSSM Higgs bosons decaying to $\tau^+\tau^-$ pairs in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector", *Phys.Lett.* B705 (2011) 174–192, DOI: 10.1016/j.physletb.2011.10.001, eprint: 1107.5003.

[150] G. Aad et al., "Search for charged Higgs bosons decaying via $H^+ \to \tau\nu$ in top quark pair events using pp collision data at $\sqrt{s} = 7$ TeV with the ATLAS detector", *JHEP* 1206 (2012) 039, DOI: 10.1007/JHEP06(2012)039, eprint: 1204.2760.

[151] J. Liebal, "PhD thesis in preparation", PhD thesis: Rheinische Friedrich-Wilhelms Universität Bonn, 2015.

[152] G. Aad et al., "Search for the neutral Higgs bosons of the Minimal Supersymmetric Standard Model in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector", *JHEP* 1302 (2013) 095, DOI: 10.1007/JHEP02(2013)095, eprint: 1211.6956.

[153] G. Aad et al., "Search for the Standard Model Higgs boson in the H to $\tau^+\tau^-$ decay mode in $\sqrt{s} = 7$ TeV pp collisions with ATLAS", *JHEP* 1209 (2012) 070, DOI: 10.1007/JHEP09(2012)070, eprint: 1206.5971, URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/HIGG-2012-07/.

[154] The ATLAS Collaboration,
"Performance of the Reconstruction and Identification of Hadronic Tau Decays with ATLAS", *ATLAS-CONF-2011-152* (2011), URL: https://cds.cern.ch/record/1398195.

[155] I. W. Stewart and F. J. Tackmann,
"Theory Uncertainties for Higgs and Other Searches Using Jet Bins", *Phys. Rev. D* 85 (2012) 034011, DOI: 10.1103/PhysRevD.85.034011, eprint: 1107.2117.

[156] The ATLAS Collaboration,
"Procedure for the LHC Higgs boson search combination in summer 2011", *ATL-PHYS-PUB-2011-011, CMS-NOTE-2011-005* (2011),
URL: http://cds.cern.ch/record/1375842.

[157] J. M. Campbell and R. K. Ellis,
"An update on vector boson pair production at hadron colliders", *Phys. Rev. D* 60 (1999) 113006, DOI: 10.1103/PhysRevD.60.113006, eprint: 9905386.

[158] S. Moch and P. Uwer,
"Theoretical status and prospects for top-quark pair production at hadron colliders", *Phys. Rev. D* 78 (2008) 034003, DOI: 10.1103/PhysRevD.78.034003, eprint: 0804.1476.

[159] M. Beneke et al., "Threshold expansion of the $gg(q\bar{q}) \to Q\bar{Q} + X$ cross section at $O(\alpha_S^4)$", *Phys.Lett.* B690 (2010) 483–490, DOI: 10.1016/j.physletb.2010.05.038,
eprint: 0911.5166.

[160] M. Botje et al., "The PDF4LHC Working Group Interim Recommendations" (2011),
eprint: 1101.0538.

[161] R. D. Ball et al.,
"Impact of Heavy Quark Masses on Parton Distributions and LHC Phenomenology", *Nucl. Phys. B* 849 (2011), eprint: 1101.1300.

[162] P. Z. Skands, "Tuning Monte Carlo Generators: The Perugia Tunes", *Phys.Rev.* D82 (2010) 074018, DOI: 10.1103/PhysRevD.82.074018, eprint: 1005.3457.

[163] G. Aad et al., "Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector at the LHC", *Eur.Phys.J.* C71 (2011) 1630,
DOI: 10.1140/epjc/s10052-011-1630-5, eprint: 1101.2185.

[164] The ATLAS Collaboration, "Search for the Standard Model Higgs boson in $H \to \tau\tau$ decays in proton-proton collisions with the ATLAS detector", *ATLAS-CONF-2012-160* (2012),
URL: http://cds.cern.ch/record/1493624.

[165] The ATLAS Collaboration, "Calibrating the b-Tag Efficiency and Mistag Rate in 35 pb^{-1} of Data with the ATLAS Detector", *ATLAS-CONF-2011-089* (2011),
URL: http://cds.cern.ch/record/1356198.

[166] The ATLAS Collaboration, "Measurement of the b-tag Efficiency in a Sample of Jets Containing Muons with 5 fb^{-1} of Data from the ATLAS Detector",
ATLAS-CONF-2012-043 (2012), URL: http://cds.cern.ch/record/1435197.

[167] The ATLAS Collaboration,
"Measurement of the Mistag Rate with 5 fb^{-1} of Data Collected by the ATLAS Detector", *ATLAS-CONF-2012-040* (2012), URL: http://cds.cern.ch/record/1435194.

[168] J. Kraus, "W -> tau nu Cross Section Measurement and Search for the Standard Model Higgs Boson in the tau(lep) tau(had) Final State with the ATLAS Detector", PhD thesis: Rheinische Friedrich-Wilhelms Universität Bonn, under review, 2015.

[169] L. Breiman et al., *Classification and Regression Trees*, New York: Chapman & Hall, 1984 358.

[170] A. Hocker et al., "TMVA - Toolkit for Multivariate Data Analysis", *PoS* ACAT (2007) 040, eprint: physics/0703039.

[171] The ATLAS Collaboration, "Evidence for Higgs Boson Decays to the $\tau^+\tau^-$ Final State with the ATLAS Detector", *ATLAS-CONF-2013-108* (2013), URL: http://cds.cern.ch/record/1632191.

[172] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting", *J. Comput. Syst. Sci.* 55 (1997) 119, ISSN: 0022-0000, DOI: 10.1006/jcss.1997.1504.

[173] J. Friedman, "Stochastic gradient boosting", *Comput. Stat. Data Anal.* 38 (2002) 367.

[174] The ATLAS Collaboration, "Jet energy scale and its systematic uncertainty in proton-proton collisions at $\sqrt{(s)} = 7$ TeV with ATLAS 2011 data", *ATLAS-CONF-2013-004* (2013), URL: http://cds.cern.ch/record/1509552.

[175] G. Aad et al., "Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector", *Eur.Phys.J.* C73 (2013) 2306, DOI: 10.1140/epjc/s10052-013-2306-0, eprint: 1210.6210.

[176] J. Rojo et al., "The impact of heavy quark mass effects in the NNPDF global analysis", *PoS* DIS2010 (2010) 173, eprint: 1007.0354.

[177] The ATLAS Collaboration, "ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11", *ATL-PHYS-PUB-2011-009* (2011), URL: http://cds.cern.ch/record/1363300.

[178] A. Kolmogorov, "Sulla determinazione empirica di una legge di distribuzione", *G. Ist. Ital. Attuari* 4 (1933) 83–91.

[179] N. Smirnov, "Table for Estimating the Goodness of Fit of Empirical Distributions", *The Annals of Mathematical Statistics* 19.2 (June 1948) 279–281, DOI: 10.1214/aoms/1177730256, URL: http://dx.doi.org/10.1214/aoms/1177730256.

[180] J. H. Friedman, "Data Analysis Techniques for High-Energy Particle Physics" (1974).

[181] A. David et al., "LHC HXSWG interim recommendations to explore the coupling structure of a Higgs-like particle" (2012), eprint: 1209.0040.

[182] G. Aad et al., "Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector", *submitted to JHEP* (2015), eprint: 1501.04943.

[183] S. Chatrchyan et al., "Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons", *JHEP* 1405 (2014) 104, DOI: 10.1007/JHEP05(2014)104, eprint: 1401.5041.

[184] The ATLAS Collaboration, "Search for a Standard Model Higgs boson in $H \to \mu\mu$ decays with the ATLAS detector.", *ATLAS-CONF-2013-010* (2013), URL: http://cds.cern.ch/record/1523695.

Danksagung

Danken möchte ich zuerst Prof. Dr. Norbert Wermes, der mir die Gelegenheit gegeben hat, in seiner Arbeitsgruppe zu promovieren, und mich während dieser Zeit nicht nur betreut, sondern in allen Angelegenheiten ständig unterstützt hat. Ein besonderer Dank gilt Dr. Jürgen Kroseberg, der mir sowohl bei der Koordination und Präsensation meiner Studien in der ATLAS Kollaboration als auch bei der Erstellung meiner Dissertation zu jeder Tages- und Nachtzeit mit hilfreichen Ratschlägen zur Seite stand.

Vielen Dank an Dr. Nicolas Möser und Dr. Martin Schmitz für die technische Einarbeitung, an Jana Kraus für die Zusammenarbeit am Analyse-Framework und an Jessica Liebal für die Übernahme der Verantwortung für das Embedding in ATLAS. Ebenfalls möchte ich mich bei allen Mitgliedern der ATLAS $H \rightarrow \tau\tau$ Analyse-Gruppe für die gemeinsame Arbeit bedanken, ohne die eine Veröffentlichung der ATLAS-Resultate nicht möglich gewesen wäre - besonders bei Koji Nakamura für die erste $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$ Analyse und bei Nils Ruthmann für die BDT-Analyse.

Nicht zuletzt möchte ich mich bei meinen Eltern bedanken, die mich während des Studiums und der Promotion ständig unterstützt haben, so dass ich neben der Arbeit noch Zeit für einen Ausgleich durch die Musik finden konnte.

Ganz besonders aber danke ich meiner Verlobten, Constanze Lademann, ohne die ich meine Promotion weder begonnen noch durchgehalten noch abgeschlossen hätte.