

**Search for the Standard Model Higgs boson  
in the  $H \rightarrow \tau\tau \rightarrow \ell\nu_\ell\nu_\tau h\nu_\tau$  decay mode  
in proton-proton collisions at  $\sqrt{s} = 7$  TeV  
with the ATLAS experiment**

DIPLOMARBEIT

vorgelegt von

Helge Haß

10. April 2013



Prof. Dr. Karl Jakobs

Fakultät für Mathematik und Physik der  
*Albert-Ludwigs-Universität*  
Freiburg im Breisgau



**Search for the Standard Model Higgs boson  
in the  $H \rightarrow \tau\tau \rightarrow \ell\nu_\ell\nu_\tau h\nu_\tau$  decay mode  
in proton-proton collisions at  $\sqrt{s} = 7$  TeV  
with the ATLAS experiment**

DIPLOMARBEIT

vorgelegt von

Helge Haß

10. April 2013

Prof. Dr. Karl Jakobs

Fakultät für Mathematik und Physik der

*Albert-Ludwigs-Universität*

Freiburg im Breisgau



# Contents

<b>Introduction</b>	<b>1</b>
<b>1 Theoretical overview</b>	<b>3</b>
1.1 Standard Model of particle physics . . . . .	3
1.1.1 Particle content . . . . .	4
1.1.2 Interactions from local gauge symmetries . . . . .	5
1.2 The Higgs mechanism . . . . .	10
1.3 Phenomenology of hadron colliders . . . . .	13
1.3.1 Luminosity and cross section . . . . .	13
1.3.2 Parton distribution . . . . .	15
1.3.3 Hadronization, pile-up and underlying event . . . . .	16
1.4 Higgs boson physics at the LHC . . . . .	17
1.4.1 Higgs boson production processes . . . . .	17
1.4.2 The decay of the Higgs boson . . . . .	18
1.4.3 Constraints on the Higgs boson mass . . . . .	21
1.4.4 The discovery of a Higgs-like boson . . . . .	23
<b>2 The ATLAS experiment and the Large Hadron Collider</b>	<b>25</b>
2.1 The Large Hadron Collider LHC . . . . .	25
2.2 The ATLAS experiment . . . . .	26
2.2.1 Inner detector . . . . .	28
2.2.2 Calorimetry . . . . .	29
2.2.3 Muon system . . . . .	31
2.3 The trigger . . . . .	32
<b>3 Event reconstruction and identification</b>	<b>33</b>
3.1 Track and vertex reconstruction . . . . .	33
3.2 Muon reconstruction . . . . .	34
3.2.1 Reconstruction efficiency and momentum resolution of the muon . . . . .	35

3.3	Electron reconstruction and identification . . . . .	36
3.3.1	Electron energy scale and resolution . . . . .	37
3.4	Lepton isolation measures . . . . .	38
3.5	Jet reconstruction and energy scale . . . . .	38
3.5.1	Jet reconstruction . . . . .	38
3.5.2	Jet energy scale . . . . .	40
3.6	Identification of $b$ -jets . . . . .	41
3.7	Reconstruction and identification of hadronic $\tau$ leptons . . . . .	41
3.8	Transverse missing momentum . . . . .	45
<b>4</b>	<b>Statistical methods</b>	<b>47</b>
4.1	Methodology . . . . .	47
4.2	The Likelihood function . . . . .	48
4.3	Test statistics . . . . .	49
4.3.1	Test statistic for discovery . . . . .	51
4.3.2	Test statistic for upper limits . . . . .	51
4.3.3	Quoting median expected limits . . . . .	51
4.4	Approximate probability distributions . . . . .	52
4.4.1	The <i>Asimov</i> dataset . . . . .	53
4.5	Illustrations of the limit computation with approximative distributions . . . . .	54
4.5.1	Blinded upper limits . . . . .	55
4.5.2	The $CL_s$ procedure . . . . .	55
<b>5</b>	<b><math>H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}</math> analysis in data at <math>\sqrt{s} = 7</math> TeV</b>	<b>57</b>
5.1	Event samples . . . . .	57
5.1.1	Signal and background processes . . . . .	57
5.1.2	Data samples . . . . .	60
5.1.3	Simulated events . . . . .	63
5.1.4	Pile-up re-weighting . . . . .	66
5.2	Event and object preselection . . . . .	66
5.2.1	Object selection . . . . .	66
5.2.2	$\tau\tau$ mass reconstruction . . . . .	67
5.3	Background estimation . . . . .	72
5.3.1	OS-SS method . . . . .	73
5.3.2	$N_{\text{SS}}^{\text{Data}}$ and $r_{\text{QCD}}$ . . . . .	73
5.3.3	W+jets estimation and control samples . . . . .	75

5.3.4	$Z \rightarrow \ell\ell + \text{jets}$ estimation . . . . .	76
5.3.5	$t\bar{t}$ estimation and control samples . . . . .	77
5.3.6	Di-boson background . . . . .	78
5.3.7	$Z \rightarrow \tau\tau$ background . . . . .	78
5.3.8	Fake factor method . . . . .	81
<b>6</b>	<b>Re-analysis of the 2011 dataset</b>	<b>83</b>
6.1	Optimization of the VBF category . . . . .	83
6.2	The boosted category . . . . .	87
6.3	Suppression of fake tau backgrounds . . . . .	87
6.4	Analysis categories . . . . .	91
6.4.1	Kinematic distributions . . . . .	93
<b>7</b>	<b>A multivariate approach for the <math>H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}</math> search</b>	<b>95</b>
7.1	Boosted Decision Trees . . . . .	96
7.2	Event classification . . . . .	100
7.3	Training parameter optimization . . . . .	103
7.3.1	Input variable and sample optimization . . . . .	104
7.3.2	Cross evaluation . . . . .	107
7.4	Signal region distributions . . . . .	108
7.5	Further BDT studies . . . . .	112
<b>8</b>	<b>Exclusion limits on the Higgs boson production cross section</b>	<b>117</b>
8.1	Systematic uncertainties on signal and background estimates . . . . .	118
8.1.1	QCD scale uncertainties on the signal production cross section . . . . .	118
8.1.2	Systematic uncertainties on Monte Carlo methods and luminosity . . . . .	120
8.1.3	Trigger acceptance systematic uncertainties . . . . .	120
8.1.4	Reconstruction and identification uncertainties . . . . .	121
8.1.5	Uncertainties on the background estimation . . . . .	123
8.2	Exclusion limits of the cut-based analysis . . . . .	124
8.3	Exclusion limits of the BDT analyses . . . . .	125
8.4	Comparison between cut-based and BDT approaches . . . . .	127
<b>9</b>	<b>Summary and conclusions</b>	<b>131</b>
<b>A</b>	<b>Measures of sensitivity</b>	<b>137</b>
<b>B</b>	<b>Scheme of the <math>H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}</math> object selection</b>	<b>139</b>

---

<b>C</b>	<b>Input variables applicable in the multivariate approach</b>	<b>141</b>
<b>D</b>	<b>BDT control regions in the training scenario without <math>m_{\tau\tau}^{\text{MMC}}</math></b>	<b>151</b>
<b>E</b>	<b>Systematic variations within the multivariate approach</b>	<b>153</b>
<b>F</b>	<b>Cut-based systematic uncertainties</b>	<b>159</b>
<b>G</b>	<b>BDT-related systematic uncertainties</b>	<b>163</b>
G.1	Training with $m_{\tau\tau}^{\text{MMC}}$ . . . . .	163
G.2	Training without $m_{\tau\tau}^{\text{MMC}}$ . . . . .	165
	<b>Bibliography</b>	<b>167</b>
	<b>List of Acronyms</b>	<b>179</b>
	<b>List of Figures</b>	<b>181</b>
	<b>List of Tables</b>	<b>185</b>
	<b>Acknowledgment</b>	<b>187</b>

---

# Introduction

---

Throughout the history of mankind, it was desired to achieve knowledge about the unfragmentable constituents of matter. Our current knowledge of elementary particles and their interactions is condensed in the Standard Model of particle physics (SM).

The SM, developed over the past 50 years, laid the cornerstone of our current understanding of elementary particles, separated on the one hand into particles of half-integer spin, the so-called *fermions*, as integral part of matter. On the other hand, three of the four known particle interactions can be described within the SM, mediated by particles with integer spin known as *bosons*. A multitude of experiments were realized to probe the SM and confirmed its predictions at a high degree of precision. In addition, new elementary particles initially proposed by theory could be discovered<sup>1</sup>.

However, it was impossible to include fermion and boson masses into the SM framework until the Higgs mechanism "revealed" them in 1964 through spontaneous symmetry breaking [5–9]. After the  $\tau$  neutrino was discovered in 2000 [10], only one component was left to conclude the SM. The famous and long-sought Higgs boson, introduced by the Higgs mechanism, eluded its detection for decades.

Finally, in July 2012, the two multi-purpose experiments at the LHC, ATLAS and CMS, were able to independently claim discovery of a new Higgs-like boson within three bosonic decay channels [11,12]. However, the Higgs mechanism describes couplings to fermions as well, which are not yet observed in direct decays. Since the Higgs boson coupling is proportional to mass, the  $H \rightarrow \tau^+\tau^-$  decay with the  $\tau$  as heaviest lepton is the most promising to study the leptonic Yukawa coupling of the Higgs boson.

In this thesis, the search for the SM Higgs boson in the  $H \rightarrow \tau^+\tau^-$  decay with the ATLAS experiment at the LHC is presented. The  $\tau$  lepton with a mass of 1.78 GeV [13] can decay into either leptons or hadrons. Within this thesis, the particular decay of one  $\tau$  lepton into an electron or muon together with one hadronically decaying  $\tau$  lepton, both accompanied by neutrinos, is considered. Furthermore, the Higgs boson production via vector-boson fusion with its characteristic jet topology offers a sensitive signal region for the  $H \rightarrow \tau^+\tau^-$  analysis

---

<sup>1</sup>For example the W- [1, 2] and Z-bosons [3, 4].

[14] It is optimized with respect to the expected sensitivity of a Higgs boson signal in this thesis.

The analysis is performed in a mass range of  $110 \text{ GeV} \leq m_H \leq 145 \text{ GeV}$  and uses proton-proton collision data corresponding to an integrated luminosity of  $4.6 \text{ fb}^{-1}$ , recorded by the ATLAS experiment in 2011 at a center-of-mass energy of  $\sqrt{s} = 7 \text{ TeV}$ .

The work and results stated in this thesis were performed within the Higgs to  $\tau\tau$  subgroup of the ATLAS collaboration. The Chapters 5 and 6 are therefore closely related to the publication of Ref. [15].

In Chapter 1, the theoretical framework of the Standard Model together with the Higgs mechanism that introduces the masses of fermions and gauge bosons is described. In addition, the phenomenology of hadron colliders as well as the Higgs boson production and decay modes that are investigated at the ATLAS experiment are outlined. Chapter 2 describes the ATLAS detector at the LHC with its components designed to measure the tracks and energy depositions of particles emerging from proton-proton collisions at high resolution.

This is followed by Chapter 3, which associates the detector measurements to physical objects with the help of various reconstruction and identification algorithms. To conclude the foundation needed for the analysis, statistical methods to quantify the compatibility of observed data to events simulated by Monte Carlo methods are presented in Chapter 4.

An overall view of the  $H \rightarrow \tau^+\tau^- \rightarrow \ell \nu_\ell \nu_\tau h \nu_h$  analysis, consisting of its different signal and background processes together with the event selection and background estimation is given in Chapter 5. Successively, Chapter 6 outlines the physically motivated categorization of events together with optimization studies of the VBF topology as most promising decay mode to discover the leptonic Higgs boson decay.

As an alternative to an analysis based on background suppression cuts, Chapter 7 motivates a multivariate approach to improve the Higgs boson signal significance. Therefore, a full analysis based on Boosted Decision Trees (BDTs) [16] is established.

Based on results of the former chapters, Chapter 8 presents the upper exclusion limits on the Higgs boson production cross section, found with both cut-based and BDT analyses utilizing a profiled likelihood ratio. Finally, Chapter 9 summarizes the findings and concludes the thesis.

Within this chapter, a brief introduction of the Standard Model of particle physics (SM) together with the Higgs mechanism as generator of particle masses is given. To probe the predictions of the SM, proton-proton collisions are performed at the LHC. Therefore, important phenomenological aspects of hadron colliders along with the protons peculiarity as compound object are described.

### 1.1. Standard Model of particle physics

The Standard Model of particle physics exploits quantum field theory to build a framework that predicts the electromagnetic, weak and strong interactions between all known particles. The gravitation as fourth elementary force resists the implementation into the SM, yet, because there is no known way to implement general relativity into a quantum field theory. However, it plays a negligible role at today's collider energies.

The SM evolved during the 1960s and 70s based on experimental observations. Moreover, it predicted the existence of additional elementary particles like the top quark which was experimentally observed in 1995 with the Tevatron collider [17, 18].

The underlying fundamental principle is the *local gauge invariance* of certain symmetry groups with respect to space-time coordinates. The internal symmetry of the SM is given through  $SU(3)_C \times SU(2)_L \times U(1)_Y$ , which describes the strong, weak and electromagnetic interactions through the exchange of gauge fields<sup>1</sup> that are associated to the generators of the symmetry groups<sup>2</sup>. The exchange is carried out by eight gluon fields carrying *color charge*  $C$ , three gauge fields  $W_i$  carrying *weak isospin*  $T$  and the gauge field  $B_0$  which carries *weak hypercharge*  $Y$ . The latter are *mixed* and observed as  $W^\pm$  and  $Z$  bosons which mediate the weak interaction and as  $\gamma$  boson which mediates the electromagnetic interaction. Moreover, left-handed fields are grouped in doublets of  $SU(2)_L$ , while right-handed fields transform as  $SU(2)_L$  singlets.

---

<sup>1</sup>The corresponding particles of the gauge fields are referred to as bosons.

<sup>2</sup>The association of conserved quantities to symmetries is extensively used in physics and mathematically described by the Noether theorem [19].

However, it is impossible to introduce mass terms in the SM because they violate the gauge invariance in a way that cannot be compensated and thus render the theory unrenormalizable. Hence, the mass of elementary particles is obtained through the Higgs mechanism that reveals the mass by *spontaneous symmetry breaking* and is described in Section 1.2. Throughout the thesis, natural units, i.e.  $\hbar = c = 1$ , are used and thus energy, momentum and mass are expressed in units of GeV. The elementary principles of the SM are outlined below and follow closely the discussions in Refs. [20–22].

### 1.1.1. Particle content

The fundamental particles within the SM can be segmented in two groups. Those with half integer spin, referred to as *fermions*, are associated with matter, whereas bosons with integer spin serve as *force mediators*.

Fermions are further divided in leptons and quarks and classified in three generations with two leptons and quarks each. They are shown in Table 1.1, arranged according to their electric charge (Q) and ordered by ascending mass (m). On top, each lepton and quark has an anti-particle with opposite charge. Each lepton generation involves one neutrino, which is electrically neutral and considered to be massless in the SM<sup>3</sup>. Charged leptons and quarks interact via the electromagnetic (EM) force. Furthermore, all fermions interact through the *weak force*, while quarks differ from leptons by their color charge which makes them subject to the strong force.

Table 1.1.: Quarks and leptons of the Standard Model grouped in generations with their corresponding electrical charge [13].

	Electric charge [e]	1 <sup>st</sup> Generation		2 <sup>nd</sup> Generation		3 <sup>rd</sup> Generation	
Leptons	-1	Electron	$e^-$	Muon	$\mu^-$	Tau	$\tau^-$
	0	Electron neutrino	$\nu_e$	Muon neutrino	$\nu_\mu$	Tau neutrino	$\nu_\tau$
Quarks	$\frac{2}{3}$	Up	u	Charm	c	Top	t
	$-\frac{1}{3}$	Down	d	Strange	s	Bottom	b

Regarding the bosons, strong and electromagnetic interactions are transmitted through massless gluons and photons, whereupon weak interactions are mediated by massive  $W^\pm$  and Z bosons.

The weak force acts at a short range of  $10^{-16} - 10^{-17}$  m due to the high mass of its

<sup>3</sup>There are observations of neutrino-oscillations that would lead to a non-zero neutrino mass [23, 24].

corresponding gauge bosons.  $W$  and  $Z$  couple to all fermions as well as to each other and the  $W^\pm$  bosons carry electromagnetic charge in addition. Therefore, they are subject to the electromagnetic force, too.

Gluons mediate the strong force through their coupling to color charge. They carry a color together with an anti-color themselves and thus couple to other gluons as well as to quarks. The strong interaction has the prominent feature that its coupling constant gets infinitely small in the limit of small lengths (high energies) and infinitely large at large length scales.

This leads to characteristics referred to as *asymptotic freedom* and *confinement*. The former specifies the treatment of quarks and gluons as free particles. In contrast, the latter prevents quarks and gluons to be observable as free particles by excitation of new quark-antiquark pairs out of the strong field energy whenever their distance exceeds a specific threshold of about  $10^{-15}$  m [25].

Thus, quarks are confined to compound particles (hadrons) of either three (anti)quarks (baryons) or one quark together with an antiquark (mesons). While the quarks within a hadron are called *valence* quarks, there are additional gluons and quark-antiquark pairs (*sea quarks*) that inherit a part of the momentum of the hadron and have to be considered in the description of inelastic scattering (see Section 1.3).

Finally, the massless photon mediates the electromagnetic (EM) force and thereby couples to all particles that carry electric charge. Opposed to the other two forces, the electromagnetic force is not bound to a finite range and photons do not couple to themselves. The gauge bosons together with their intrinsic charge and mass are listed in Table 1.2.

Table 1.2.: Gauge bosons of the Standard Model with their corresponding charge and mass (specifications taken from Ref. [13]).

Interaction	Gauge boson	Q [e]	Mass
Electromagnetic	Photon ( $\gamma$ )	0	$< 1 \times 10^{-18}$ eV
Strong	8 gluons	0	0
Weak	$W^+$	+1	$80.385 \pm 0.015$ GeV
	$W^-$	-1	
	Z	0	$91.188 \pm 0.002$ GeV

### 1.1.2. Interactions from local gauge symmetries

To describe particle physics based on field theory and local gauge symmetries, the Lagrangian framework which is widely used in classical mechanics is adopted. To incorporate it in field

theory, the formalism is extended from a discrete system with coordinates  $q_i(t)$  to continuously varying fields described by  $\phi(\mathbf{x}, t)$ . This leads to a Lagrangian density  $\mathcal{L}$  (further simply denoted as Lagrangian).

Hence, the *Euler-Lagrange equation* for a Lagrangian  $\mathcal{L}$  is given by

$$\frac{\partial}{\partial x_\mu} \left( \frac{\partial \mathcal{L}}{\partial(\partial\phi/\partial x_\mu)} \right) - \frac{\partial \mathcal{L}}{\partial\phi} = 0. \quad (1.1)$$

### Quantum Electrodynamics

To describe Quantum Electrodynamics (QED), one starts with a relativistic equation that describes a freely propagating (*Dirac*) field  $\Psi$  of spin 1/2 and mass  $m$ , known as the *Dirac equation*

$$(i\gamma^\mu \partial_\mu - m)\Psi = 0. \quad (1.2)$$

Here, the  $\gamma^\mu$  denote the so-called *Dirac  $\gamma$ -matrices* described for instance in Ref. [22].

The Lagrangian equivalent to the Dirac equation is

$$\mathcal{L} = i\bar{\Psi}\gamma_\mu\partial^\mu\Psi - m\bar{\Psi}\Psi \quad (1.3)$$

with the adjoint spinor  $\bar{\Psi} = \Psi^\dagger\gamma^0$ . This Lagrangian is invariant under a *global* U(1) transformation

$$\Psi(x) \rightarrow \Psi'(x) = \exp^{i\alpha} \Psi(x), \quad \alpha \in \mathbb{R}. \quad (1.4)$$

On the other hand, local gauge invariance with a coordinate dependent  $\alpha(x)$  implies

$$\Psi(x) \rightarrow \Psi'(x) = \exp^{i\alpha(x)} \Psi(x), \quad (1.5)$$

with the consequence that the first term of the Lagrangian does not obey the local gauge invariance.

To compensate for the extra term arising when substituting Eq. 1.5 in Eq. 1.3, the derivative  $\partial_\mu$  has to be modified including a vector field  $A^\mu$ . This leads to

$$D_\mu \equiv \partial_\mu - ieA_\mu, \quad (1.6)$$

where the vector field  $A^\mu$  transforms as

$$A_\mu \rightarrow A_\mu + \frac{1}{e}\partial_\mu\alpha. \quad (1.7)$$

To get this new field in accordance with the physical photon field, a term corresponding to its kinematic energy has to be added ( $\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ ).

Finally, this leads to the Lagrangian of QED with a *massless* photon and a coupling between a Dirac particle with charge  $-e$  and the photon field  $A_\mu$  of

$$\mathcal{L} = \bar{\Psi}(i\gamma^\mu\partial_\mu - m)\Psi + e\bar{\Psi}\gamma^\mu A_\mu\Psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \quad (1.8)$$

A naive mass term of the form  $\frac{1}{2}m^2 A_\mu A^\mu$  would explicitly violate gauge invariance of the Lagrangian, since the additional terms in the Lagrangian which arise from Eq. 1.7 cannot be compensated. If the Lagrangian with broken symmetry would nevertheless be taken to compute propagators, diverging perturbative corrections would arise and render the theory unrenormalizable. More sophisticated techniques to generate particle masses are discussed in Section 1.2.

### Quantum Chromodynamics and non-Abelian symmetries

Quantum Chromodynamics (QCD) are described by the local gauge invariance under the special unitary symmetry group  $SU(3)_C$ . This introduces quark triplets based on three different intrinsic *color charges*. This additional quantum number was adopted because of the otherwise symmetrical wave function of the  $\Delta^{++}$  particle<sup>4</sup>. On the other hand, the branching ratio of quark and muon final states produced in inelastic  $e^+e^- \rightarrow Z$  collisions indicated three distinct quark generations [13]. The quark triplet  $q = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix}$  thereby stands for any of the six known quark flavors u, d, s, c, b and t.

The  $SU(3)_C$  phase space transformation is of the form

$$q(x) \rightarrow q'(x) \equiv \exp^{i\alpha_a(x)T_a} q(x), \quad (1.9)$$

where  $T_a$  is a set of 8 *non-Abelian*, linearly independent and traceless  $3 \times 3$  matrices that generate the  $SU(3)_C$  group.

Local gauge invariance can be restored by modifications similar to the QED, leading to

$$D_\mu = \partial_\mu - ig_s T_a G_\mu^a. \quad (1.10)$$

In this equation,  $g_s$  denotes the strong coupling constant and  $G_\mu^a$  constitutes eight gauge fields that correspond to eight different gluons as field mediators of the strong force. The  $G_\mu^a$

<sup>4</sup>The Pauli exclusion principle states that two identical fermions cannot occupy the same quantum state at the same time, which would be the case for three up quarks with spin  $\pm\frac{1}{2}$  as constituents of the  $\Delta^{++}$  particle.

transform as

$$G_\mu^a \rightarrow G_\mu^a - \frac{1}{g} \partial_\mu \alpha_a - f_{abc} \alpha_b G_\mu^c, \quad (1.11)$$

whereas the last term restores gauge invariance given the non-commuting gauge transformation. Hence, the final QCD Lagrangian is given by

$$\mathfrak{L} = \sum_q (\bar{q}(i\gamma^\mu \partial_\mu - m)q - g_s (\bar{q}\gamma^\mu T_a q) G_\mu^a) - \frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu}. \quad (1.12)$$

Equation 1.11 introduces the self-coupling of gluons, together with the coupling of gluons to quark fields, as remarkable feature of the non-Abelian  $SU(3)_C$  symmetry.

### Electroweak interactions

The unification of electromagnetic and weak interactions into a single framework, entitled *electroweak model* and described by a  $SU(2)_L \times U(1)_Y$  gauge symmetry, was a major development in particle physics and is also known as *Glashow-Salam-Weinberg* or GSW model [26–28].

Weak interactions are mediated by either  $W^\pm$  bosons, referred to as *charged weak current* or by the  $Z$  boson as *neutral weak current*. All fermions couple to the weak bosons and interactions including  $W^\pm$  bosons are the only known to change the quark flavor together with their electric charge.

Another peculiarity of weak interactions are their parity<sup>5</sup> violation processes. It was found that  $W$  bosons couple to left-handed fermions (right-handed anti-fermions) only, what results in a maximal parity violation [29].

To account for this behavior in the GSW model, left-handed fermions are grouped in doublets with *weak isospin*  $T_3 = \pm\frac{1}{2}$ , whereas right-handed fermions are grouped in isospin singlets with respect to the  $SU(2)_L$  symmetry with  $T_3 = 0$ . In addition, they carry a *weak hypercharge*  $Y$ . These intrinsic properties are pictured in Table 1.3 and connected to the measurable electric charge via the *Gell-Mann-Nishijima* equation

$$Q = T_3 + \frac{Y}{2}. \quad (1.13)$$

To summarize, the  $SU(2)_L \times U(1)_Y$  transformations of the left-handed doublets and the

---

<sup>5</sup>A parity operation corresponds to a point reflection.

Table 1.3.: The fermions with their corresponding quantum numbers. Left-handed fermions are denoted by  $L$ , right-handed by  $R$ .

	Generation			Quantum numbers			
	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>	I	$I_3$	Y	Q [e]
Quarks	$\begin{pmatrix} u \\ d \end{pmatrix}$	$\begin{pmatrix} c \\ s \end{pmatrix}$	$\begin{pmatrix} t \\ b \end{pmatrix}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{2}{3}$
	$u_R$	$c_R$	$t_R$	0	0	$\frac{4}{3}$	$\frac{2}{3}$
	$d_R$	$s_R$	$b_R$	0	0	$-\frac{2}{3}$	$-\frac{1}{3}$
Leptons	$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L$	$\frac{1}{2}$	$\frac{1}{2}$	-1	0
	$e_R^-$	$\mu_R^-$	$\tau_R^-$	0	0	-2	-1

right-handed singlets are

$$\begin{aligned}\chi_L &\rightarrow \chi'_L = \exp^{i\alpha(\mathbf{x})\cdot\mathbf{T}+i\beta(x)Y} \chi_L, \\ \psi_R &\rightarrow \psi'_R = \exp^{i\beta(x)Y} \psi_R,\end{aligned}\tag{1.14}$$

with local phases  $\alpha(x)$  and  $\beta(x)$  together with  $\mathbf{T}$  and  $Y$  as generators of the  $SU(2)_L$  and  $U(1)_Y$  groups of gauge transformations, respectively.

In analogy to the last two sections, gauge fields with corresponding coupling constants have to be introduced to render the Lagrangian invariant under local transformations of  $\chi$  and  $\psi$ . Thus, three  $SU(2)_L$  gauge fields  $W_\mu^a$  with coupling constant  $g$  together with one gauge field,  $B_\mu$  for the  $U(1)_Y$  symmetry with coupling constant  $g'$ , have to be established. Moreover, the derivative  $\partial_\mu$  has to be replaced by a covariant derivative  $D_\mu$  similar to Eq. 1.6 and Eq. 1.10, which yields

$$D_\mu = \partial_\mu + ig\frac{\tau_a}{2}W_\mu^a + ig'\frac{Y}{2}B_\mu.\tag{1.15}$$

To conclude, the Lagrangian for electroweak interactions is provided by

$$\mathcal{L}_{EW} = i\bar{\chi}_L^i\gamma^\mu D_\mu^L\chi_L^i + i\bar{\psi}_R^i\gamma^\mu D_\mu^R\psi_R^i - \frac{1}{4}W_{\mu\nu}^a W_a^{\mu\nu} - \frac{1}{4}B_{\mu\nu}^2.\tag{1.16}$$

As consequence of the non-Abelian  $SU(2)$  symmetry, the  $W_\mu^a$  gauge fields are self-interacting, whereas the  $B_\mu$  field does not couple to itself.

Together with the QCD Lagrangian introduced earlier, an entire picture of the elementary particle interactions within the SM is set. However, there remains one major distinction from the particles that were observed in experiments, as all gauge bosons and fermions introduced

in the former sections had to be massless to establish gauge invariance and render the theory renormalizable.

## 1.2. The Higgs mechanism

In order to generate masses for the three heavy gauge bosons ( $W^\pm$ ,  $Z$ ), the *Higgs mechanism* is introduced [5–9]. Therefore, the potential  $V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2$  is added to the electroweak Lagrangian in Eq. 1.16,

$$\mathfrak{L} = i\bar{\chi}_L^i\gamma^\mu D_\mu^L\chi_L^i + i\bar{\psi}_R^i\gamma^\mu D_\mu^R\psi_R^i - \frac{1}{4}W_{\mu\nu}^a W_a^{\mu\nu} - \frac{1}{4}B_{\mu\nu}^a B_a^{\mu\nu} - V(\phi). \quad (1.17)$$

$\phi$  is thereby chosen as an isospin doublet with  $Y=1$  and four real scalar fields  $\phi_i$ , given by

$$\phi = \begin{pmatrix} \phi_\alpha \\ \phi_\beta \end{pmatrix} = \sqrt{1/2} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \quad (1.18)$$

Concerning the potential  $V(\phi)$ , two scenarios have to be distinguished (pictured in Fig. 1.1). For  $\mu^2 > 0$  and  $\lambda > 0$ , the potential has a global minimum at  $\phi = 0$  and  $\mathfrak{L}$  results in 4 scalar particles  $\phi_i$  of mass  $\mu$  interacting with massless gauge bosons. On the other hand, the scenario with  $\mu^2 < 0$ ,  $\lambda > 0$  has infinite degenerated minima.

The minima in the latter scenario satisfy  $\phi^\dagger\phi = -\frac{\mu^2}{\lambda} \equiv v^2/2$ , where  $v$  denotes the vacuum expectation value. One of them has to be chosen as vacuum ground state, e.g.  $\phi_3^2 = v^2$ . By allocation of this arbitrary ground state, a *spontaneous symmetry breaking* is realized.

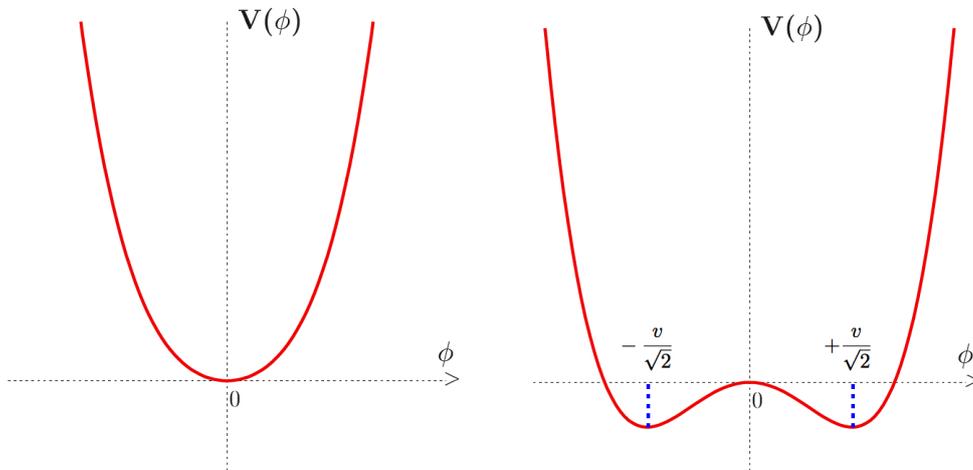


Figure 1.1.: The potential  $V(\phi)$  shown for  $\mu^2 > 0$ ,  $\lambda > 0$  (left plot) as well as  $\mu^2 < 0$ ,  $\lambda > 0$  (right plot) as a projection in one dimension.

To examine excitations of  $\phi(x)$ , it gets expanded around the ground state, leading to

$$\phi(x) = \sqrt{\frac{1}{2}} \cdot \exp(i\frac{\sigma_i}{2}\theta^i(x)) \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \quad (1.19)$$

with four real fields  $H(x)$  and  $\theta^i(x)$ , since we started with four scalar fields in the  $SU(2)$  doublet. However, due to the local gauge invariance, the  $\theta^i(x)$  can be gauged as  $\theta^i(x) = 0$  and correspond to three massless *Goldstone*<sup>6</sup> bosons.

The  $U(1)_{QED}$  symmetry of the vacuum associated to all  $H(x) \equiv 0$  states represents a massless degree of freedom identified with the photon. The chosen vacuum ground state corresponds to

$$\phi_0 \equiv \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (1.20)$$

and, if substituted in the primary Lagrangian (Eq. 1.17), yields self-coupling terms of

$$\frac{1}{2}(\frac{g}{2}v)^2(W_\mu^1W^{1,\mu} + W_\mu^2W^{2,\mu}) + \frac{1}{8}v^2(W_\mu^3, B_\mu) \begin{pmatrix} g^2 & -gg' \\ -gg' & g'^2 \end{pmatrix} \begin{pmatrix} W^{3,\mu} \\ B^\mu \end{pmatrix}. \quad (1.21)$$

To attain the actual observed physical fields,  $W_\mu^1$  and  $W_\mu^2$  are combined through

$$W_\mu^\pm = \frac{1}{2}(W_\mu^1 \pm W_\mu^2), \quad (1.22)$$

whereas the mass term results of a comparison to  $M_W^2 W^+ W^-$ .

The last term in Eq. 1.21 is off-diagonal in the  $W_\mu^3$  and  $B_\mu$  basis and the physical fields  $Z_\mu$  and  $A_\mu$  diagonalize the matrix by

$$Z_\mu = \cos \theta_W W_\mu^3 - \sin \theta_W B_\mu, \quad (1.23)$$

$$A_\mu = \sin \theta_W W_\mu^3 + \cos \theta_W B_\mu, \quad (1.24)$$

with a weak mixing angle  $\theta_W$ <sup>7</sup>.

By re-casting the electroweak Lagrangian (Eq. 1.16) in terms of the physical fields, one

<sup>6</sup>The *Goldstone theorem* states that there exist as many massless spin-0 particles as broken generators of the underlying symmetry.

<sup>7</sup>The *weak-mixing* angle can be expressed according to the last term of Eq. 1.21 by  $\cos \theta_W = \frac{g}{\sqrt{g^2+g'^2}}$  and  $\sin \theta_W = \frac{g'}{\sqrt{g^2+g'^2}}$ . The quantity of  $\theta_W$  has to be determined empirically, e.g. through W and Z mass measurements related by  $\cos \theta_W = \frac{m_W}{m_Z}$

can compare the components linked to  $A_\mu$  with the QED of Section 1.1.2, what reveals

$$e = g \sin \theta_W = g' \cos \theta_W, \quad Q = T_3 + \frac{Y}{2} \quad (1.25)$$

as discussed e.g. in Ref. [20].

To conclude, the three gauge bosons mediating the weak interaction obtained mass while the photon remains massless,

$$M_W = \frac{1}{2}gv, \quad M_Z = \frac{\sqrt{g^2 + g'^2}}{2}v = \frac{M_W}{\cos \theta_W}, \quad M_\gamma = 0. \quad (1.26)$$

Therefore, the three Goldstone bosons which correspond to three degrees of freedom became the longitudinal polarization of the now massive gauge bosons, which appear together with self-coupling terms in the Lagrangian. In addition, the remaining scalar particle H of the parametrization is called Higgs boson. It couples to the massive electroweak bosons plus itself and acquires a mass of

$$m_H = \sqrt{\lambda}v, \quad (1.27)$$

with a free parameter  $\mu$  that has to be determined experimentally.

Prior to the Higgs mechanism, the local gauge invariance of the Lagrangian was broken by insertion of fermion mass terms such as  $\mathfrak{L}_m = -m(\bar{\Psi}_L\Psi_R + \bar{\Psi}_R\Psi_L)$ . However, due to the additional scalar doublet introduced by the Higgs mechanism, fermion masses can be acquired by additional terms in the Lagrangian that correspond to *Yukawa interactions* between the Higgs boson field and the fermions. These terms are  $SU(2)_L \times U(1)_Y$  gauge invariant and result for instance for an electron in

$$\mathfrak{L}_{\text{Yukawa},e} = -\frac{G_e}{\sqrt{2}}v(\bar{e}_L e_R + \bar{e}_R e_L) - \frac{G_e}{\sqrt{2}}h(\bar{e}_L e_R + \bar{e}_R e_L) = -m_e \bar{e}e - m_e \bar{e}e \frac{h}{v} \quad (1.28)$$

with the electron mass  $m_e = \frac{G_e v}{\sqrt{2}}$  and an arbitrary coupling constant  $G_e$ .

Concerning the quarks, masses for both quarks in the doublet have to be generated. Thus, the doublet  $\phi_C = \sqrt{\frac{1}{2}} \begin{pmatrix} v + H(x) \\ 0 \end{pmatrix}$  is introduced.

To allow for changes of the quark family in weak interactions as observed in experiments, the quark eigenstates participating in weak interactions have to be superpositions of their mass eigenstates. This superposition is performed by the *Cabibbo-Kobayashi-Maskawa* (CKM)

matrix [29] through

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = M_{CKM} \times \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (1.29)$$

The final Lagrangian yields

$$\mathcal{L}_{\text{Yukawa},Q} = -G_d^{ij} \bar{Q}_L^i \phi d_R^j - G_u^{ij} \bar{Q}_L^i \phi_C u_R^j + \text{hermitian conjugates}, \quad (1.30)$$

where  $u_R^j, d_R^j$  are the corresponding up- and down-type quarks and the coupling constant and mixing between the generations is inherited by G.

The renormalization of the theory remains after applying the Higgs mechanism, which has been shown by 't Hooft in 1971 [30]. To conclude, the SM Lagrangian is derived from an  $SU(3)_C \times SU(2)_L \times U(1)_Y$  gauge symmetry and is given by

$$\mathcal{L}_{SM} = \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{EW}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}. \quad (1.31)$$

### 1.3. Phenomenology of hadron colliders

In contrast to electron-positron colliders which are used to achieve high precision measurements and operate at a specific center-of-mass energy, hadron-hadron colliders are mainly used as *discovery machines* due to the broad energy range and high maximal energy they can access. This section describes the underlying phenomenology of collider experiments followed by a theoretical model for hadron collisions. Since the colliding protons are composite objects, several effects based on the strong interactions of the partons within the proton need to be considered when analyzing proton-proton collisions. These are described below.

#### 1.3.1. Luminosity and cross section

The most important quantity when performing particle collisions at collider experiments is the expected number of events  $N_{\text{events}}$  per time, denoted as rate R, for each physical process. Two distinct parts contribute to the prediction of R, the theoretical cross section  $\sigma$  for the specified process as well as the *instantaneous Luminosity* L given by the collider, which concludes in

$$R = \frac{dN_{\text{events}}}{dt} = L \times \sigma. \quad (1.32)$$

By integrating Eq. 1.32 over time, the total number of events can be expressed using the integrated luminosity  $\mathcal{L} = \int L dt$ .

$L$  can be quoted for a storage ring with  $N_B$  particle bunches that collide with a bunch crossing frequency per second  $f$  and contain  $N_p$  protons in each bunch by

$$L \propto \frac{N_p^2 N_B f}{4\pi\sigma_x\sigma_y}, \quad (1.33)$$

where  $\sigma_x$  and  $\sigma_y$  correspond to the width of the particle bunches in x- and y-direction under assumption of Gaussian beam profiles.

The cross section on the other hand can be calculated independently within the SM. In its short form, it can be written as

$$\sigma = \int d\sigma = \int \frac{|\mathcal{M}|^2}{F} dQ, \quad (1.34)$$

where  $F$  and  $dQ$  are kinematic factors for the incident flux in the laboratory and the Lorentz invariant phase space factor, respectively. The whole physics is established by the invariant amplitude  $\mathcal{M}$  that describes the transition between the initial and final states. The expected cross sections of different physics processes at the Tevatron and the LHC are shown in Fig. 1.2 as a function of the  $pp$  center-of-mass energy  $\sqrt{s}$ .

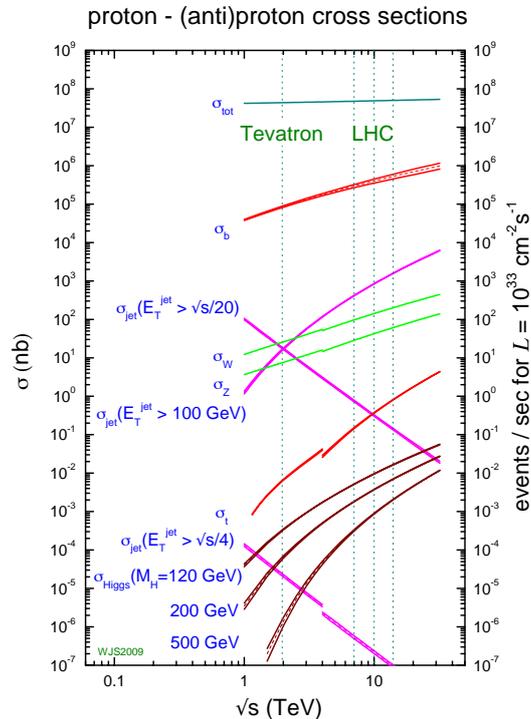


Figure 1.2.: Production cross sections for different center-of-mass energies at a logarithmic scale. In 2011, the LHC operated at  $\sqrt{s} = 7$  TeV, which is illustrated by the leftmost of the three dashed lines assigned to the LHC [31].

### 1.3.2. Parton distribution

When performing an experiment with hadrons in the initial state, Eq. 1.34 has to be extended since the proton is composed of valence quarks, sea-quarks and gluons, referred to as partons. From this variety of constituents, it is unknown which partons participated in a particular interaction of the  $pp$  collision. In addition, the momentum fraction of each parton with respect to its proton is uncertain.

The parton distribution functions (PDFs) parametrize the parton content and the probabilities of specific momentum fractions within a hadron. It is determined experimentally<sup>8</sup> and  $f_q(x)$  states the probability to have a parton  $q$  within the hadron that carries a relative momentum fraction  $x \in [0, 1]$  [32]. In Figure 1.3, the MSTW<sup>9</sup> set of PDFs, parametrized at NNLO, is illustrated.

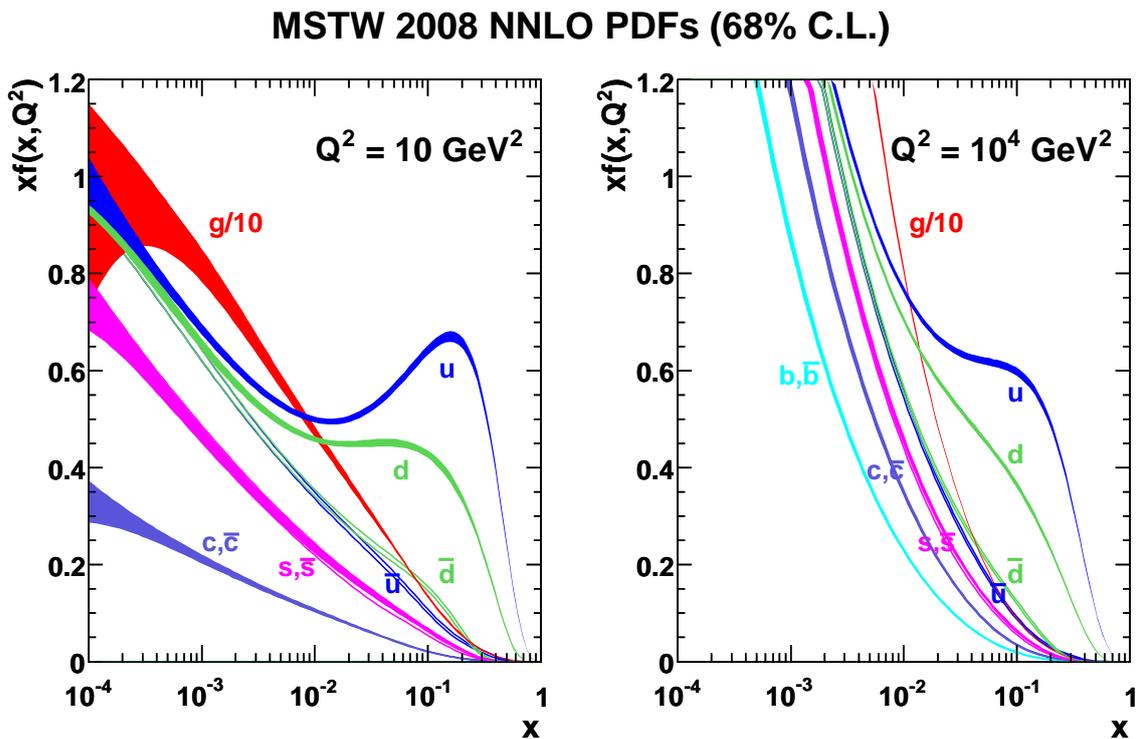


Figure 1.3.: MSTW parton distribution functions for the partons in a proton determined at next-to-next-to-leading-order for a momentum transfer of  $Q^2 = 10 \text{ GeV}^2$  (left) and  $Q^2 = 10^4 \text{ GeV}^2$  (right). The valence quarks (d and u) are depicted together with the gluons (scaled down by a factor of 10) and sea quarks. The 68 % confidence level (CL) is shown by the width of the distributions [33].

<sup>8</sup>PDFs are usually estimated from deep-inelastic scattering experiments which allow to explore the hadron structure.

<sup>9</sup>Martin-Stirling-Thorne-Watt.

Thus, the total cross section of an interaction of hadrons can be expressed as convolution of their hard scattering process (Eq. 1.34) and corresponding PDFs, given by

$$\sigma(P_1 P_2 \rightarrow X) = \sum_{p,q} \int_0^1 dx_p \int_0^1 dx_q f_p(x_p) f_q(x_q) \hat{\sigma}(p(x_p)q(x_q) \rightarrow X). \quad (1.35)$$

The cross section is summed over all partons  $p, q$  that can be the origin of the inelastic scattering of hadrons  $P_1$  and  $P_2$  into a final state  $X$ , whereas  $\hat{\sigma}(p(x_p)q(x_q) \rightarrow X)$  is taken from Eq. 1.34.

The cross section derived with PDFs similar to Eq. 1.35 neglects an important aspect of the proton structure. The radiation of gluons at initial or final state add contributions to the perturbative calculation of the cross section. This introduces an additional dependence of the PDF on the momentum transfer  $Q^2$  in the interaction and hence violates the *Bjorken Scaling*<sup>10</sup>.

The *DGLAP*<sup>11</sup> evolution equation describes the dependence of  $f_q(x, Q^2)$  on  $Q^2$  by

$$\frac{d}{d \log Q^2} f_q(x, Q^2) = \frac{1}{2\pi} \int_x^1 \frac{dy}{y} f_q(y, Q^2) \times \alpha_s P_{qq}\left(\frac{x}{y}\right) + \dots, \quad (1.36)$$

where the probability, that a parton of momentum fraction  $x$  came from a parent parton with momentum fraction  $y > x$  is given by  $P_{qq}\left(\frac{x}{y}\right)$ . The equation is integrated over all possible momenta between  $x$  and 1. Finally, it is extended by further additional processes in the scattering.

### 1.3.3. Hadronization, pile-up and underlying event

The colored quarks and gluons as result of the hard interaction outlined above are subject to the confinement caused by the strong interaction (Section 1.1.2). Hence, they undergo a complex *hadronization* process which leads to color neutral final states that are observed as collimated *hadron jets*.

Since the QCD coupling strength increases at large distances, the hadronization cannot be described by perturbation theory. Therefore, *fragmentation functions*  $D_p^h(z)$  are used to parameterize the probability to find a hadron  $h$  with momentum fraction  $z$  of the initial parton  $p$  [34]. As the PDFs, the fragmentation functions are determined experimentally.

Besides, pile-up events that are unrelated to the hard scattering are a side effect of  $pp$  collisions. They arise from simultaneous events initiated by the same or neighbor bunches

<sup>10</sup> *J. Bjorken* suggested in 1968, that the inelastic proton scattering is in fact an elastic scattering of free interacting partons for the limit of infinite momentum transfer.

<sup>11</sup> Dokshitzer-Gribov-Lipatov-Altarelli-Parisi

and have to be considered at the reconstruction of an event and its objects.

On the one hand, particles from additional proton interactions in the same bunch are scattered into the calorimeters, denoted as *in-time pile-up*, which deposit energy in the detectors and distort the energy determination and scaling of reconstructed particles. The contribution of in-time pile-up rises with the density inside a single bunch of the particle accelerator.

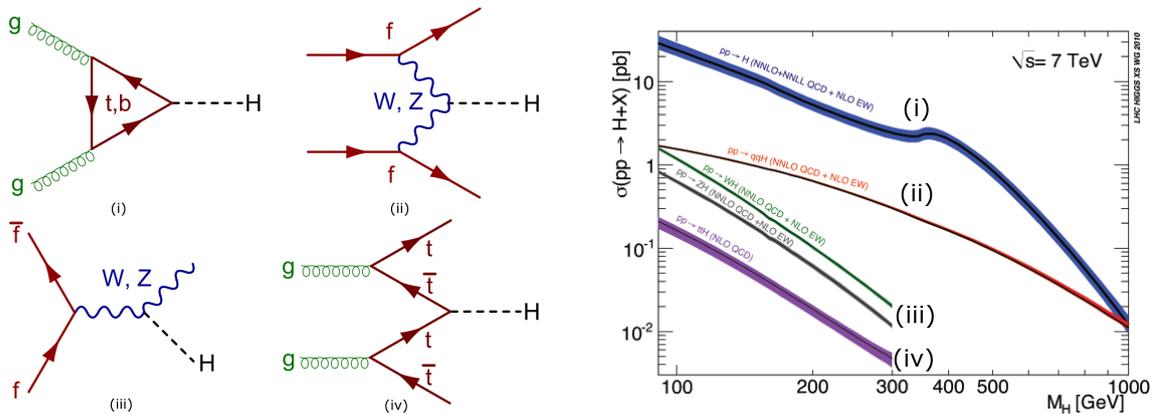
On the other hand, the bunch spacing of 50 ns at the LHC in 2011 introduces a sensitivity to the energy flow of past and future collisions, known as *out-of-time pile-up*. It occurs when the response time of the detector is larger than the bunch separation time.

Apart from pile-up, the remnants of the proton contribute to the event that is detected. The hadronization of the residual partons to color neutral states and its aggregate is called *underlying event*. Effects from pile-up and the underlying event have to be incorporated in the simulation of events by Monte Carlo methods as described in Section 5.1.3.

## 1.4. Higgs boson physics at the LHC

### 1.4.1. Higgs boson production processes

In proton-proton collisions at the LHC, several Higgs boson production processes are exploited. They are illustrated as leading order Feynman diagrams in Fig. 1.4a. Their corresponding cross section<sup>12</sup>  $\sigma$  at  $\sqrt{s} = 7$  TeV is displayed in Fig. 1.4b.



(a) Leading order Feynman diagrams of the Higgs boson production via gluon fusion (i), vector-boson fusion (ii), associated Higgs boson production (iii) and top-quark fusion (iv).

(b) Cross sections of the four Higgs boson production processes at  $\sqrt{s} = 7$  TeV as a function of the Higgs boson mass [35].

Commencing with the highest cross section, the gluon fusion (ggF) denotes the Higgs boson production where two gluons fuse via quark (mainly top) loops due to the fact that the Higgs boson coupling is proportional to the mass.

<sup>12</sup> $[\sigma] = \text{barn} = 10^{-24} \text{ cm}^2$ .

Furthermore, the vector-boson fusion (VBF) has roughly one tenth of the ggF cross section at  $m_H = 125$  GeV with  $\sqrt{s} = 7$  TeV and describes a production process where two vector bosons that were radiated off the incoming quarks fuse to a Higgs boson. An important feature of the VBF is, that the out-coming quarks form characteristic jets with high transverse momentum that usually end up in the forward detector region, referred to as *tagging jets*. This facet is used to get a clear Higgs boson signature, for instance in the here presented  $H \rightarrow \tau_{\text{lep}} \tau_{\text{had}}$  analysis.

Proceeding with the third-highest cross section, the Higgs boson is produced in association with a W or a Z boson. Thus, jets or leptons are added to the final state. The associated production with an additional lepton in the final state is the main production mode exploited within the  $H \rightarrow b\bar{b}$  decay channel that is described below, where the lepton is needed as trigger object.

Finally, the Higgs boson can be produced via the fusion of a top and anti-top quark pair. However, this production process has the lowest cross section and a complex final state with several jets.

#### 1.4.2. The decay of the Higgs boson

In order to get a SM that is consistent with theory and data of already performed precision measurements, the Higgs boson mass has to be above 100 GeV (see Section 1.4.3). It is very short-lived and has a predicted lifetime of around  $10^{-22}$  s [36] which solely allows for a discovery via its decay products. The decay of the SM Higgs boson is determined by its coupling to mass. A decay in two photons is possible through fermion or heavy-boson loops, dominated by the top quark loop due to its high mass.

Figure 1.4 shows the branching ratios for different Higgs boson decay modes. For Higgs boson masses above 180 GeV, the decays into two heavy vector-bosons (WW,ZZ) or  $t\bar{t}$  as heaviest quark are dominant. The latter starts to contribute at roughly 345 GeV since the top quark mass of 172.9 GeV [13] leads to at least one off-shell top quark below this barrier. The same argument holds for the ZZ and WW final states that fall steeply in BR at masses below  $\sim 180$  GeV and  $\sim 160$  GeV, respectively.

The other decay products are always produced on-shell for Higgs boson masses of interest, whereby the mass range up to 160 GeV is dominated by the  $b\bar{b}$  and  $\tau\tau$  final states. To conclude, a Higgs boson decay into two gluons or a  $c\bar{c}$  quark pair is hard to distinguish from background contaminations and will probably never be exploited at the LHC.

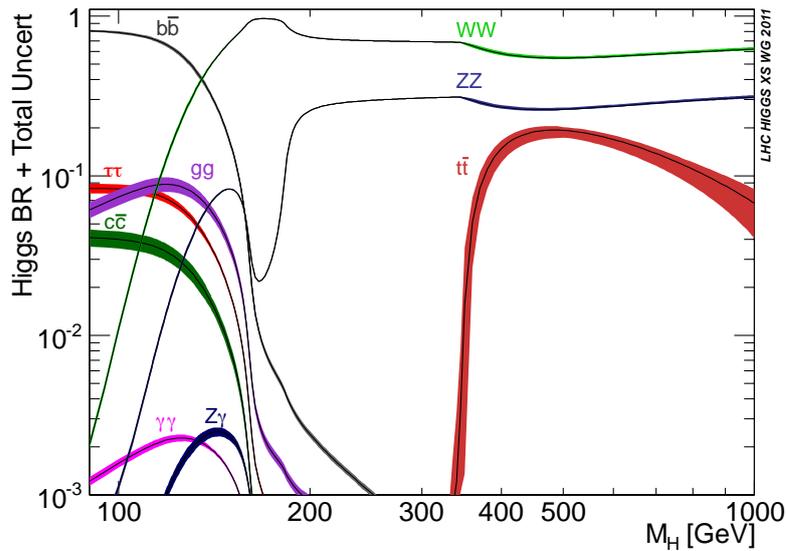


Figure 1.4.: Higgs boson branching-ratios for a potential Higgs boson mass between 80 and 1000 GeV in logarithmic scale. The total uncertainty is illustrated by the width of the lines corresponding to one  $\sigma$  [36].

Figure 1.5 shows the cross sections times branching ratios for the various decay channels. The decay modes usually demand at least one lepton in the final state, since the ATLAS trigger algorithm has to diminish the occurred events by a factor of roughly  $5 \times 10^6$  (see Section 2.3). Hence, without leptons, the jet thresholds have to be very high or a pre-scale<sup>13</sup> has to be applied in order to store events of a reasonable rate given the computing and storage resources.

Final states without jets or missing energy in the Higgs boson decay have the capability to find the Higgs boson with a clean signature and a narrow mass resolution. Therefore, the evidence for a new Higgs-like boson was first and foremost found in the  $H \rightarrow \gamma\gamma$  and  $H \rightarrow ZZ \rightarrow \ell^+\ell^-\ell^+\ell^-$  channels. In these analyses, all objects can be fully reconstructed and the full angular information is accessible. They also allow to measure the spin and coupling structure of the Higgs boson as soon as more data is available. The other ZZ decay possibilities are accompanied by either a Z decaying into two neutrinos or a quark-antiquark pair.

In the case of neutrinos which escape detection, only the total transverse missing momentum, reconstructed of all other objects in the event, is accessible. For gluon- or quark-initiated jets, a lot of background is expected from the parton remnants in the event. Thus, both suffer from a worsened resolution and provide additional experimental challenges.

<sup>13</sup>A pre-scale with a factor  $n$  indicates, that only every  $n$ -th event is recorded. This allows for a much lower trigger threshold.

Apart from the  $Z$  and  $\gamma$  bosons, the  $H \rightarrow W^+W^-$  decay has a high  $\sigma \times BR$  and thereby is a promising candidate to discover the Higgs boson. However, the  $H \rightarrow W^+W^-$  channel with leptons in the final state is accompanied by neutrinos which cause a broader mass resolution and do not allow a full reconstruction of all objects in the final state.

Amongst the possible fermionic Higgs boson decay modes, only the decay into  $b\bar{b}$  or  $\tau^+\tau^-$  have a sufficient<sup>14</sup>  $\sigma \times BR$  to allow for a search of fermionic Higgs boson decay modes within the ATLAS detector using the data recorded in 2011<sup>15</sup>.

The associated Higgs boson production with a leptonic decay of the  $W^\pm$  or  $Z$  boson is the most important in the  $b\bar{b}$  decay mode due to the presence of a lepton in the final state for triggering.

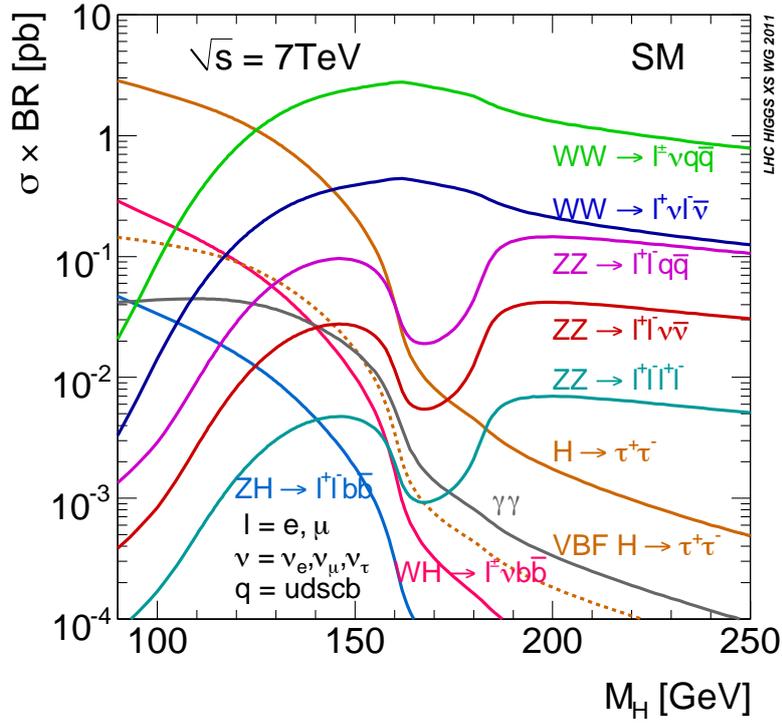


Figure 1.5.: Higgs boson production cross sections times branching-ratios of its various decay modes at  $\sqrt{s} = 7$  TeV. The  $H \rightarrow \tau\tau$  decay is illustrated by the solid orange line, whereas the dashed orange line illustrates its distinct VBF topology [37].

<sup>14</sup>Because the Higgs boson couples proportional to the mass.

<sup>15</sup>The top quark itself has the highest quark mass but decays within  $\sim 5 \times 10^{-25}$  s [13] into another quark and thus only its decay products can be reconstructed.

### The $\tau\tau$ final state

Last but not least, the  $\tau\tau$  final states offer a very high  $\sigma \times BR$  up to a Higgs boson mass of  $m_H = 150$  GeV. At a mass of approximately 125 GeV, where the Higgs-like boson was discovered in July 2012, the  $\sigma \times BR$  is  $\sim 1$  pb for the inclusive  $H \rightarrow \tau\tau$  decay as well as  $\sim 0.1$  pb for its VBF production mode. It provides a promising channel to probe the Yukawa coupling of the Higgs boson to leptons. Besides, the characteristic jet topology of the VBF channel can be exploited to achieve a high signal significance.

Since the  $\tau$  lepton has a mass of 1.777 GeV [13], it can decay into another lepton, accompanied by a  $\tau$  (anti)neutrino and the corresponding lepton (anti)neutrino as well as hadronically together with a  $\tau$  (anti)neutrino. The branching ratio of  $H \rightarrow \tau\tau$  into two hadronic ( $\tau_{\text{had}}\tau_{\text{had}}$ ), one hadronic and leptonic ( $\tau_{\text{had}}\tau_{\text{lep}}$ ) and two leptonic ( $\tau_{\text{lep}}\tau_{\text{lep}}$ ) final states is 42 %, 46 % and 12 %, respectively.

The  $\tau_{\text{lep}}\tau_{\text{lep}}$  final state has the advantage, that two leptons are present in the final state. They can be used for triggering with lower thresholds on the lepton  $p_T$ . Moreover, it does not suffer from inefficiencies in the hadronic  $\tau$  identification (see Section 3.7). In contrary, the  $ee$  and  $\mu\mu$  final states are affected by a high  $Z \rightarrow \ell\ell$  contamination and it has the lowest BR.

The  $\tau_{\text{had}}\tau_{\text{had}}$  decay on the other hand has a high BR but is subject to a high QCD background contamination. In addition, it has to deal with two  $\tau_{\text{had}}$  candidates in the final state that are difficult to identify. Also, no lepton is present in the final state which could be used for triggering.

Finally, the  $\tau_{\text{had}}\tau_{\text{lep}}$  channel is a comprise between high BR and leptons in the final state. They can be reconstructed and identified with high efficiency and used to trigger the event. The  $Z \rightarrow \tau\tau$  constitutes a large irreducible background of this decay mode since all other backgrounds can be diminished by background suppression cuts (more details are outlined in the Chapters 5 and 6).

#### 1.4.3. Constraints on the Higgs boson mass

Prior to 2012, the existence of the Higgs boson was not established despite tremendous efforts made in particle collisions. In addition, its mass is not predicted by the Higgs mechanism (see Eq. 1.27).

Theoretical constraints on the high and low Higgs boson mass bound can be computed at various energy scales  $\Lambda$  (illustrated in Fig. 1.6a).

Besides theoretical constraints on the Higgs boson mass for different cut-off scales, electroweak precision measurements can be used to probe the compatibility of different Higgs

boson masses  $m_H$  with the SM. This is possible because the  $W^\pm$  and  $Z$  boson mass measurements reach high accuracy and become sensitive to loop corrections including the top quark and the Higgs boson<sup>16</sup>.

If high- $Q^2$  measurements for the Z-pole, hadronic vacuum polarization, top and W mass are used as constraints to compare different Higgs mass hypotheses with the SM, an upper limit of  $m_H < 158$  GeV can be set at 95 % confidence level (CL). Moreover, a minimum in the  $\chi^2$  test is found at a Higgs boson mass of  $m_H = 94^{+29}_{-24}$  GeV [38]. The test results with comparison to the minimum are shown by the  $\Delta\chi^2$  curves in Fig. 1.6b. In addition, the yellow areas indicate where a Higgs boson could already be excluded at 95 % CL in the direct search by analyzing the compatibility of observed data with different Higgs boson masses. To conclude, these measurements gave amenities to a low SM Higgs boson mass, whereby masses below 114.4 GeV were already excluded at 95 % CL [38].

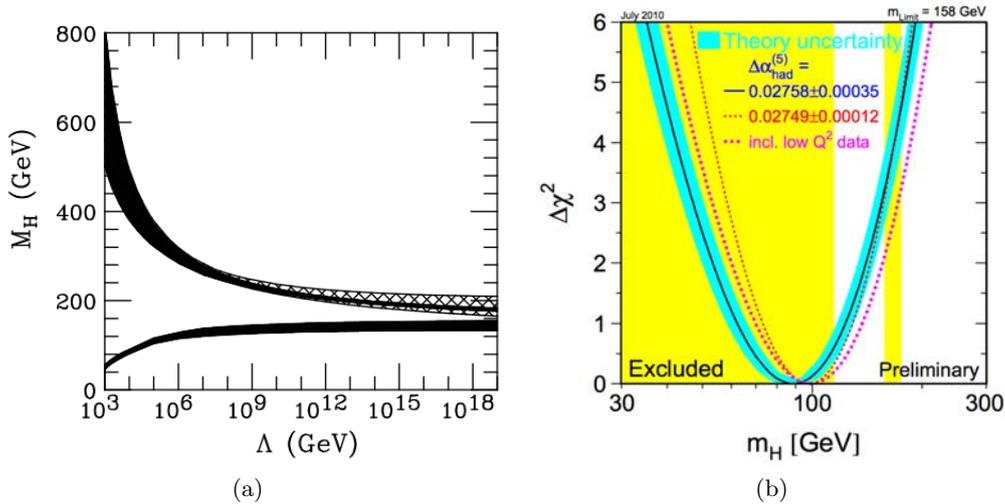


Figure 1.6.: (a) Theoretical upper and lower constraints on the Higgs boson mass [39]. The solid and cross-hatched areas indicate theory uncertainties on the Higgs boson mass. (b)  $\Delta\chi^2 = \chi^2 - \chi^2_{min}$  curves using an input from electroweak precision measurements carried out at the LEP, SLD, CDF and DØ as a function of the Higgs boson mass  $m_H$  with the assumption of the Standard Model. The blue band represents the estimated theoretical uncertainties due to missing higher-order corrections, whereas the pink dots illustrate the fit with a hadronic vacuum polarization term derived by theory. The yellow areas show the mass regions which could be excluded by the LEP and Tevatron experiments at 95 % CL [38].

<sup>16</sup>However, the top quark mass enters with quadratic dependence which infers much stronger constraints than the logarithmic dependence on the Higgs boson mass.

### 1.4.4. The discovery of a Higgs-like boson

As a result of its very low production cross section and high expected backgrounds, the Higgs boson eluded discovery for about 50 years. Finally, the CMS and ATLAS collaborations could claim discovery of a new, Higgs-like boson at roughly 125 GeV after scanning billions of events recorded in 2011 and in the first half of 2012. With a dataset corresponding to an integrated luminosity of  $\sim 11 \text{ fb}^{-1}$ , they were able to declare the discovery with 5 (CMS [12]) and 6 (ATLAS [11]) standard deviations, respectively. This translates to a probability (local  $p_0$ -value) of  $5 \times 10^{-6}$  and  $10^{-9}$ , that the observed data originated exclusively of SM processes without a Higgs boson.

Figure 1.7 shows the  $p_0$ -values and significances for the ATLAS and CMS experiment, respectively. The observed  $p_0$ -value describes the compatibility of the observed data with simulated events assuming an absence of a Higgs boson. The expected  $p_0$ -value on the other hand describes the compatibility of simulated events with a given SM Higgs boson signal to those without Higgs boson. At the time of the discovery, neither the ATLAS nor the CMS experiment saw a significant excess in their  $H \rightarrow \tau^+\tau^-$  analyses.

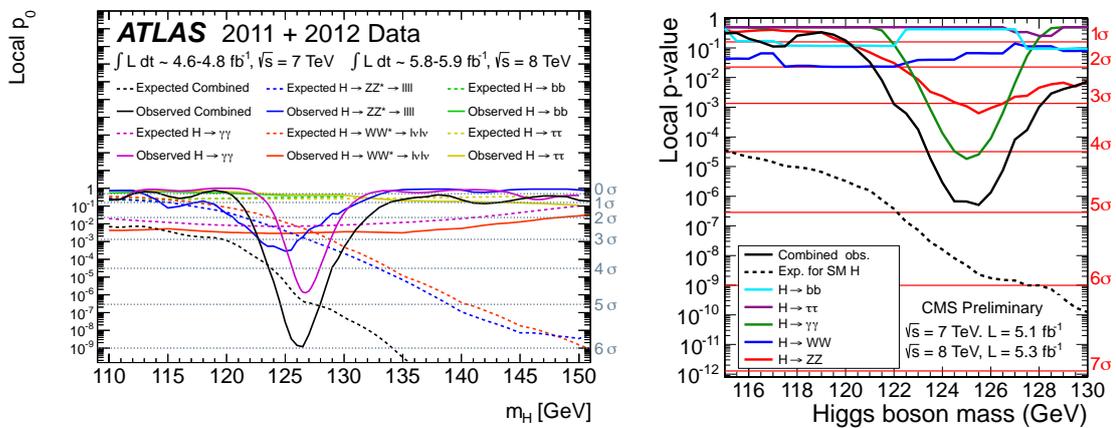


Figure 1.7.: The observed (solid lines) and expected (dashed lines) local  $p_0$ -values for the Higgs boson decay channels into  $\gamma\gamma$ ,  $WW$ ,  $ZZ$ ,  $b\bar{b}$  and  $\tau^+\tau^-$  separately as well as combined. The latter are depicted by the solid (dashed) black lines, whereas the  $H \rightarrow \tau^+\tau^-$  channel is shown in yellow and purple for the ATLAS (left) and CMS (right) experiments, respectively.

The most recent results for the Higgs boson mass including up to  $25 \text{ fb}^{-1}$  ( $12.2 \text{ fb}^{-1}$ ) of data recorded by the ATLAS [40] (CMS [41]) experiment in 2011 and 2012 suggest masses of

$$\begin{aligned}
 \text{ATLAS:} \quad & m_H = 125.5 \pm 0.2 \text{ (stat)} \quad {}^{+0.5}_{-0.6} \text{ (sys)} \text{ GeV}, \\
 \text{CMS:} \quad & m_H = 125.8 \pm 0.4 \text{ (stat)} \quad \pm 0.4 \text{ (sys)} \text{ GeV}.
 \end{aligned} \tag{1.37}$$

The signal strength of the Higgs boson for each decay channel as well as combined is shown in Fig. 1.8. For most channels, the signal strength for both the ATLAS and the CMS experiment is compatible with the one expected in the Standard Model at a mass of  $m_H = 125$  GeV, denoted by  $\mu = \frac{\sigma}{\sigma_{SM}} = 1$ . However, a small excess is observed in the  $H \rightarrow \gamma\gamma$  decay, where both the ATLAS and the CMS experiment measure values higher than the one predicted by theory.

For the combination of all analyzed Higgs boson decay channels, the ATLAS experiment states a signal strength of  $\mu = 1.30 \pm 0.20$  [42], whereas the CMS experiment obtains a best fit result of  $\mu = 0.88 \pm 0.17$  [12]. Thus, no significant deviations from the SM expectation are observed for the combined signal strength.

One of the upcoming challenges is to confirm whether the new particle has the properties postulated by the Standard Model. Therefore, the spin, CP eigenvalue and its couplings to fermions, bosons and its self-coupling have to be measured.

First studies of the  $H \rightarrow ZZ$  decay to determine the spin and parity of the Higgs-like boson favor the SM assignment of  $J^P = 0^+$  over the other examined states  $0^-$ ,  $1^+$ ,  $1^-$ ,  $2_m^+$  and  $2^-$  at confidence levels between 74 % and 99 % [43]. Moreover, separate fits to the different Higgs boson production modes indicate that the particle is indeed produced via ggF, VBF and associated production [44].

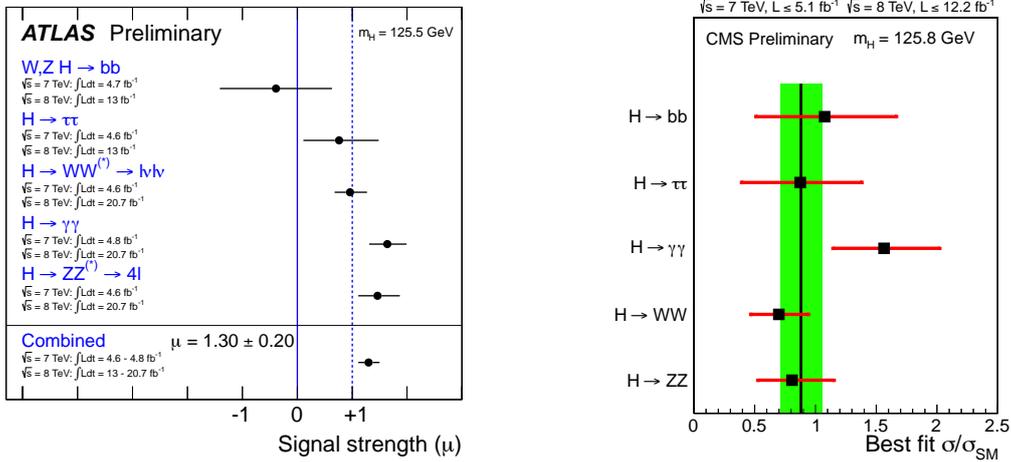


Figure 1.8.: The Higgs boson signal strength  $\mu = \frac{\sigma}{\sigma_{SM}}$  for different Higgs boson decay modes separately and combined. The ATLAS experiment (left) states the combined signal strength of  $\mu = 1.30 \pm 0.20$  below the others whereupon the dotted line illustrates the SM Higgs boson signal strength. The CMS experiment (right) depicts the combined signal strength of  $\mu = 0.88 \pm 0.17$  by a solid black line with its uncertainty in a green band.

---

# 2 The ATLAS experiment and the Large Hadron Collider

---

The  $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$  analysis described in this thesis is based on proton-proton ( $pp$ ) collision data recorded by the ATLAS experiment installed at the *Large Hadron Collider* (LHC). The LHC is briefly summarized in Section 2.1, followed by a description of the ATLAS experiment in Section 2.2. Finally, the trigger system to cope with the enormous amount of interactions occurring within the ATLAS detector is reviewed in Section 2.3.

## 2.1. The Large Hadron Collider LHC

The LHC is currently the world's most powerful particle accelerator and located at CERN<sup>1</sup> near Geneva, Switzerland. It is built in the former *Large Electron-Positron Collider* (LEP) tunnel with a circumference of 27 km and lies approximately 100 m below ground level. The LHC can store either two proton beams ( $pp$  physics) or two heavy ion beams (HI physics) that are brought to collision at four separate interaction points. The LHC design parameters for  $pp$  collisions imply a center-of-mass energy of  $\sqrt{s} = 14$  TeV along with an instantaneous luminosity of  $L = 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ .

The collision of electrons and positrons would yield a much cleaner final state due to interactions of the proton remnants in  $pp$  collisions. However, the synchrotron radiation of an accelerated particle with mass  $m$  is proportional to  $m^{-4}$  and results in a non-restorable energy loss.

The protons for the LHC originate from a hydrogen source and gain momentum in a sequence of accelerators. They are forced onto a circular orbit by a large number of superconducting dipole magnets. The LHC is capable of storing 2808 proton bunches, each composed of  $\sim 10^{11}$  protons. The temporal spacing between two bunches, the *bunch spacing*, can be as small as 25 ns. In 2011, the LHC operated at  $\sqrt{s} = 7$  TeV with a bunch spacing of 50 ns.

---

<sup>1</sup>Organisation européenne pour la recherche nucléaire, in English: European Organization for Nuclear Research.

Each of the four interaction points along the beam line is surrounded by an experiment to analyze the occurring particle collisions. These include two multi-purpose experiments, ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid) as well as two more specialized experiments, LHCb (Large Hadron Collider beauty) and ALICE (A Large Ion Collider Experiment). The Higgs boson search was one of the main incentives while designing the LHC and the ATLAS experiment. However, a diversity of other physical processes are explored, amongst others the search for supersymmetry and precision measurements of known physics.

## 2.2. The ATLAS experiment

To describe the ATLAS detector and the particles emerging from  $pp$  collisions, a particular coordinate system is defined. Its origin is the nominal interaction point with the beam pipe as the z-axis. In addition, the x-axis points to the center of the LHC ring and the y-axis points upwards. The azimuthal angle  $\phi$  is defined in the x-y plane and the polar angle  $\theta$  between z- and y-axis. Instead of the polar angle, the pseudorapidity is used, which is defined as  $\eta = -\ln \tan(\theta/2)$  and yields 0 for a particle emerging transverse to the beam axis. The geometrical distance within two detector points is measured by  $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ .

The requirements for the ATLAS detector systems were determined to harmonize with the challenge of finding the Higgs boson as well as new physics [45–47]. Leptonic decay modes of the Higgs boson<sup>2</sup> are most promising because of the high expected QCD background in hadron collisions. They have to be studied at a high momentum resolution (e.g. for  $E_T^{\text{miss}} = 0$ ) together with a good charge identification. To distinguish characteristic final states, the ATLAS detector has to provide good particle identification capabilities exploiting energy deposition, tracking and interaction vertices.

Thus, the active components of the detector serve different purposes:

- The *tracking system* in the inner detector (ID) provides good charged-particle momentum resolution and reconstruction abilities for secondary vertices that help with heavy-quark flavor tagging and the rejection of pile-up events.
- A *calorimeter system* separated into an electromagnetic (EM) calorimeter for electron and photon identification as well as a hadronic calorimeter to measure the energy of jets traversing the detector and to determine the transverse missing energy.
- To bend tracks of charged particles, a *solenoid* and a *toroid* magnet system provide the

---

<sup>2</sup>E.g.  $H \rightarrow Z^+Z^- \rightarrow l^+l^-l^+l^-$ .

required magnetic field.

- The *muon spectrometer* (MS) allows to detect muons over a wide momentum range and to measure their momentum together with their charge.

Additionally, a nearly full azimuthal coverage accompanied by a high acceptance in pseudo-rapidity is desirable. The ATLAS detector (Fig. 2.1) is built with an onion-like structure providing a forward-backward-symmetry and a coverage down to  $\simeq 1^\circ$  with respect to the beam axis.

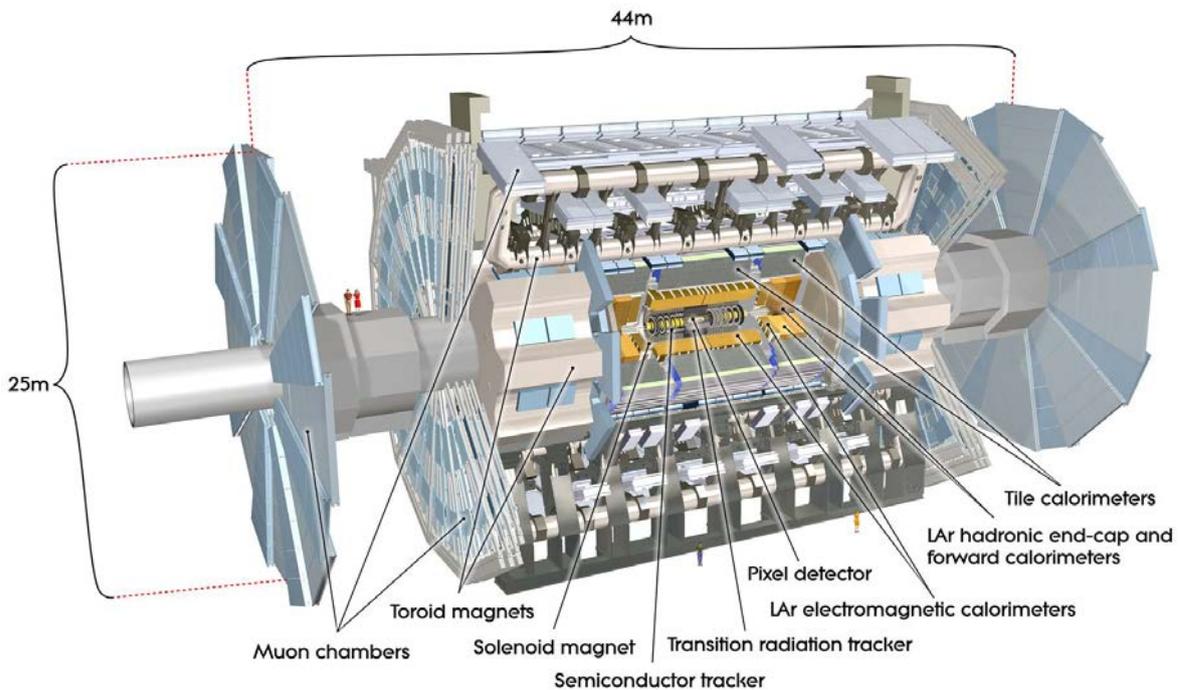


Figure 2.1.: Cut-away view of the ATLAS detector supplying the ID, the EM and hadronic calorimeter together with the MS [45].

Furthermore, the inelastic proton-proton cross section of 80 mb leads to  $10^9$  events/s at design luminosity. It presents a huge challenge for fast and radiation hard electronics together with the requirement for a good online trigger system to be able to store all important events.

The luminosity delivered to the ATLAS experiment by the LHC is measured by two small detector systems in the forward region at high  $\eta$ , namely the LUCID (LUMinosity measurement using Cerenkov Integrating Detector) and the ALFA (Absolute Luminosity For ATLAS) detectors [48]. In 2011, the ATLAS detector recorded data corresponding to an integrated luminosity of  $\int L dt = 5.25 \text{ fb}^{-1}$ .

### 2.2.1. Inner detector

The Inner Detector (ID) is built of three independent sub-systems, the *Pixel detector*, the *Semiconductor Tracker* (SCT) and the *Transition Radiation Tracker* (TRT). The ID, shown in Fig. 2.2, is immersed in a 2 T magnetic field generated by the central solenoid. It enables a transverse momentum measurement with an uncertainty of  $\frac{\sigma_{p_T}}{p_T} = 0.05 \% p_T \oplus 1 \%$ .

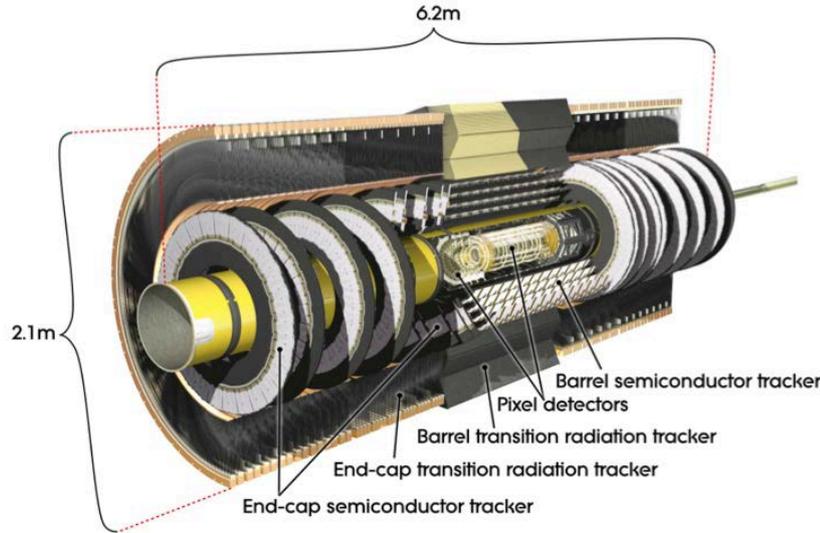


Figure 2.2.: Cut-away view of the ATLAS Inner detector (ID) [45].

#### Pixel Detector

The pixel detector is the innermost tracking device installed at around 50 mm distance to the beam pipe. It consists of more than 80 million readout channels and covers  $|\eta| < 2.5$ . It is constructed of three cylindrical layers in the barrel region along with  $2 \times 3$  disks in the end-cap regions. This assembly causes a track to traverse typically three pixel layers. The accuracy is  $10 \mu\text{m}$  for  $(R-\phi)$  in both regions and  $115 \mu\text{m}$  for  $z$  in the barrel and  $R$  in the end-cap region.

#### The Semiconductor Tracker

The SCT uses the same geometry as the pixel detector and consists of four cylindrical layers and  $2 \times 9$  disks in the barrel and end-cap regions, respectively. Together with the pixel detector, it forms the precision tracking detector system. It uses stereo strips arranged in a small angle of  $\sim 40 \text{ mrad}$  to measure two coordinates of each track, whereas in each layer one set of strips is parallel to the beam direction to determine  $R-\phi$ . The SCT provides a resolution

of  $17 \mu\text{m}$  in  $R\text{-}\phi$  for both barrel and end-cap module as well as  $580 \mu\text{m}$  for  $z$  (barrel) and  $R$  (end-cap disks).

### The Transition Radiation Tracker

The ID is completed by the TRT, which surrounds the pixel and SCT detectors and starts at a distance of approximately 55 cm to the beam pipe. About 50,000 straw tubes with a diameter of 4 mm and a length of 144 cm are installed parallel to the beam axis in the barrel region. In addition, around 120,000 straw tubes of 37 cm length are arranged radially in each end-cap region. Summed up, the TRT features approximately 350,000 readout channels.

The TRT geometry allows to follow tracks up to  $|\eta| = 2.0$ , however, it yields only  $R\text{-}\phi$  information with an accuracy of  $\sim 130 \mu\text{m}$  per straw tube provided through a drift time measurement. To recover precision, the TRT usually detects a large number of hits ( $\sim 36$  per track).

To enhance the electron identification capability of the ATLAS detector, transition-radiation photons that were created in a radiator between the straw tubes are detected. This is achieved by a  $Xe\text{-}CF_4$ -based gas mixture within the straw tubes, whereby  $Xe$  is selected due to its high absorption capability. The  $CF_4$  is added to achieve a faster drift-time of the electrons within the straw tubes.

#### 2.2.2. Calorimetry

Apart from the high-precision track measurement achieved by the ID, the calorimeter ensures a thorough and precise energy measurement of particles traversing the detector. To establish a good separation between photons and electrons against hadronically interacting particles, two different calorimeter technologies, the electromagnetic and the hadronic calorimeter, are used.

The calorimeters, installed between the ID and the MS, cover the region of  $|\eta| < 4.9$  and provide a granularity fine enough to provide lateral and longitudinal shower shapes needed for identification on top of the measurement of the full energy deposited by the particle of interest. The calorimeter depth is an important design parameter to ensure a good containment of electromagnetic and hadronic particle showers within their detector system. Also, it prevents a punch-through of hadronic showers into the MS.

An overview of both electromagnetic and hadronic calorimeters of the ATLAS detector are illustrated in Fig. 2.3.

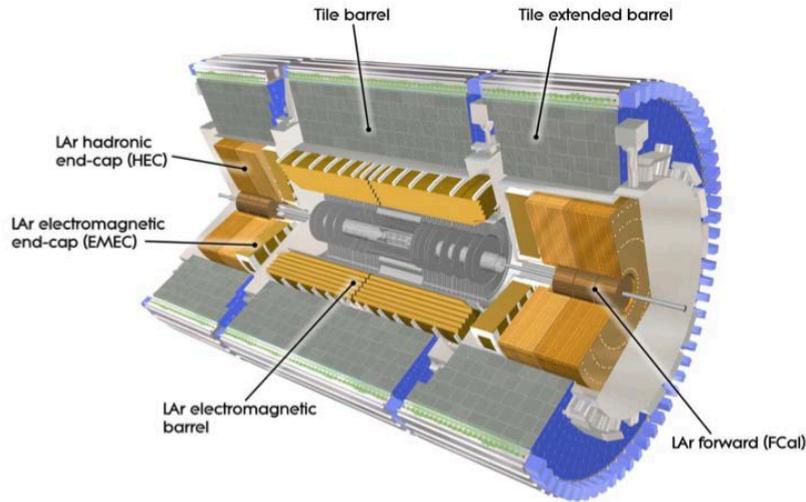


Figure 2.3.: The ATLAS calorimeter design, taken from Ref. [45].

### LAr electromagnetic calorimeter

The EM calorimeter is a liquid-argon (LAr) detector containing accordion-shaped kapton electrodes and lead absorber plates. The thickness of the lead plates is optimized in terms of energy resolution. It establishes a full azimuthal as well as an  $|\eta| < 3.2$  coverage. The latter is provided by a barrel calorimeter for  $|\eta| < 1.475$ , divided into two half-barrels with a small gap at  $z=0$  and completed by two coaxial end-cap wheels on each side, covering  $1.375 < |\eta| < 3.2$ .

To avoid two distinct vacuum walls, the barrel EM calorimeter shares a cryostat together with the central solenoid and the end-cap wheels share their cryostat with the hadronic end-cap calorimeter described below.

The region dedicated to high precision physics ( $|\eta| < 2.5$ ) is segmented into three layers with decreasing granularity. The first layer enables a precise position and shower shape determination. The second layer has the highest depth and adheres most of the energy deposited by electromagnetic particles. Finally, the third or *back layer* ensures that even high-energy electromagnetic particles deposit their entire energy in the EM calorimeter.

The EM calorimeter has an energy resolution of  $\frac{\sigma_E}{E} = 10\%/\sqrt{E} \oplus 3\%$ .

### Hadronic calorimeter

The hadronic calorimeter system utilizes three different technologies covering different  $\eta$  regions. Each region provides a total thickness in terms of interaction lengths of about  $10 \lambda$ . The innermost tile calorimeter (TileCal) covers  $|\eta| < 1.7$  and uses steel as absorber together

with scintillating tiles as active material. It is segmented into three layers and provides a resolution of  $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$  for the inner two layers, completed by  $\Delta\eta \times \Delta\phi = 0.2 \times 0.1$  for the outermost layer. The scintillating tiles are read out by wavelength shifting fibres followed by photomultiplier tubes.

The LAr hadronic end-cap (HEC) covers  $1.5 < |\eta| < 3.2$ , therefore overlapping on both sides with the other hadronic calorimeter systems. It is formed out of two wheels per side with 32 identical wedge-shaped modules each. In addition, each wheel is divided into two layers that are built of copper plates with LAr as active medium.

To conclude, the LAr forward calorimeter (FCal) envelopes the region  $3.1 < |\eta| < 4.9$ . It starts roughly 1.2 m behind the EM end-cap calorimeter to avoid neutron albedo into the ID. Hence, the truncated FCal required a high-density design. It is built of three layers, where the first one is optimized for electromagnetic particles and built of copper, followed by two layers containing tungsten to measure the energy deposited of the dominating hadronic interactions. LAr is used as sensitive medium.

The intrinsic energy resolution of the hadronic calorimeter with respect to jets is  $\frac{\sigma_E}{E} = 50 \text{ \%}/\sqrt{E} \oplus 3 \text{ \%}$  for the barrel and HEC as well as  $\frac{\sigma_E}{E} = 100 \text{ \%}/\sqrt{E} \oplus 10 \text{ \%}$  for the FCal.

### 2.2.3. Muon system

The muon spectrometer (MS) forms the outermost detector system of the ATLAS experiment. In the barrel region within  $|\eta| < 1.4$ , three cylindrical layers are arranged around the beam axis. For  $1.4 < |\eta| < 2.7$ ,  $2 \times 3$  planes perpendicular to the beam axis are installed.

The main functionality of the MS is based on the magnetic bending of muon tracks, provided by large superconducting *toroid* magnets. These are split in a barrel and two smaller end-cap magnets. In the region between barrel and end-cap magnets, referred to as *transition region* at  $1.4 < |\eta| < 1.6$ , a combination of both fields carries out the magnetic deflection. To optimize the bending power in the transition region, the end-cap magnets are rotated by  $22.5^\circ$  with respect to the barrel magnet.

To measure the tracks of traversing muons with high accuracy, the MS is instrumented with separate trigger and tracking chambers.

Monitored Drift Tubes (MDTs) represent the latter and provide a high-precision measurement of the track coordinates. To withstand the challenging background conditions in the most forward region at  $2.0 < |\eta| < 2.7$ , Cathode Strip Chambers (CSCs) with a higher granularity are used there instead of MDTs in the innermost layer.

In addition, the muons are triggered by Resistive Plate Chambers (RPCs) and Thin Gap Chambers (TGCs) in the barrel and end-cap region, respectively. They serve for a bunch-

crossing identification and provide well-defined  $p_T$  thresholds. Moreover, the RPCs and TGCs determine the muon track coordinates in the azimuthal coordinate, which is orthogonal to the ones measured by the MDTs and CSCs [45].

The MS aims at a resolution of  $\frac{\sigma_{p_T}}{p_T} = 10\%$  for muons with  $p_T = 1$  TeV.

### 2.3. The trigger

At its design parameters, the LHC will operate at an instantaneous luminosity of  $10^{34}$  cm<sup>-2</sup>s<sup>-1</sup> and thus generate  $pp$  interactions at a rate of  $\sim 1$  GHz. It is impossible and undesired to store all the data produced by the detector systems for each event, since most of them are associated to a low momentum transfer and of almost no interest.

For this purpose, a three stage trigger system is installed to reject non-significant events before data storage. They are referred to as *Level 1* (L1), *Level 2* (L2) and *Event Filter* (EF) [49, 50]. The final rate that is read-out and stored using the full detector information is about 200 Hz.

The L1 trigger system is installed at hardware level. It uses only coarse information of the calorimeter systems and the MS, therefore omitting the ID, to create a fast trigger decision. The rate is reduced to  $\sim 75$  kHz by the L1 trigger and several *Regions of Interest* (RoIs) are defined wherever a preferred particle signature is observed.

These contain  $\eta$ - and  $\phi$ -coordinates together with information on the criteria passed. The L2 selection is seeded by the RoI information and makes use of the full granularity of the calorimeters within the RoIs. The L2 trigger reduces the trigger rate to roughly 3.5 kHz and takes about 40 ms per event.

For the EF, offline algorithms that make use of informations of the detailed calorimeter structure, the MS and the ID are performed. Hence, the event rate is finally lowered to  $\sim 200$  Hz and recorded permanently using approximately 1.3 megabyte per event.

---

## 3 Event reconstruction and identification

---

To identify the underlying physics of a  $pp$  collision, it is important to reconstruct the particles traversing the ATLAS detector and reconstruct the event properties. The ATLAS detector consists of numerous detector systems that are themselves a compound of different modules. Each of these have a multitude of channels that interact with crossing particles. Based on this, the event information is stored whenever a positive trigger decision occurred.

The physical interpretation of the detector response is carried out by several algorithms collected in the ATLAS software framework ATHENA [51]. In the following, the identification and reconstruction algorithms for particles and event properties used in the  $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$  analysis are summarized. This includes particle tracks, primary interacting vertices, muons, electrons, jets, hadronic taus<sup>1</sup> and missing transverse energy.

### 3.1. Track and vertex reconstruction

The ability to reconstruct tracks of charged particles is a crucial ingredient for further identification of the particles that are used in physics analyses. The track reconstruction at the ATLAS experiment currently provides two distinct algorithms, namely the *inside-out* and the *outside-in* sequence. They are bound to  $|\eta| < 2.5$  due to the Inner Detector (ID) geometry [46, 52].

The *inside-out* technique is seeded by hits in the silicon tracker and succeeds to the outer border of the ID via a window search along the seed direction. Further hits are rejected or added to the initial track using a combinatorial Kalman filtering method [53]. To solve possible ambiguities arising from mis-identified tracks or overlapping segments, a score is assigned to each track that favors fully reconstructed tracks and hits in precise detector segments with respect to small track segments and less accurate detector parts. Tracks that passed these algorithm are extended into the Transition Radiation Tracker (TRT). The compatibility of hits is probed by either a helix fit to the track coordinates or by a more sophisticated algorithm, called deterministic annealing filter (DAF) [52]. The DAF clusters

---

<sup>1</sup>The notation electron, muon and tau is used for both leptons and their positively charged antiparticles.

a group of hits together and performs a line fit as described before. In addition, a weight is assigned to each hit in the group representing the likeliness to represent the true hit.

The very efficient *inside-out* technique is not extended to the TRT if not enough hits were assigned to a track or if the particle left an insufficient number of silicon hits in the ID. Thus, the *outside-in* sequence is used as supplement. TRT hits are thereby clustered and identified as tracks if they were not assigned to tracks reconstructed by the *inside-out* technique earlier on. They are traced back into the silicon detector which allows to find small track segments there.

In addition to the track reconstruction, it is desired to reconstruct interaction vertices across the beam line. As a first step, the reconstructed tracks fulfilling certain quality criteria are extrapolated to the beam axis  $z$  and the global maximum of this distribution along  $z$  serves as vertex seed. Subsequently, a  $\chi^2$  fit based on the  $z$  value is executed for each track with respect to the existing vertices. A new vertex seed is generated if the incompatibility of the track to each vertex is larger than  $7\sigma$ . This procedure is repeated until no track is unassigned, leaving a set of vertices along the  $z$ -axis. Finally, reconstructed vertices with only one allocated track are discarded [52].

Due to high pile-up conditions in the LHC data taking [54], the vertex originating from the hard interaction is of particular interest. Therefore, all vertices are ordered by their sum of squared momenta of all associated tracks ( $\sum p_T^2$ ). Then, the one with the highest  $\sum p_T^2$  is denoted as *primary vertex*.

For the use in physics analyses or particle reconstruction algorithms, track impact parameters quantifying the distance between the track and the primary vertex at closest approach are determined. This is done either along the beam direction, entitled *longitudinal impact parameter*  $z_0$ , or in its transverse plane, named *transverse impact parameter*  $d_0$ .

## 3.2. Muon reconstruction

To reconstruct muons in the ATLAS detector, several strategies are used to combine measurements performed by the muon spectrometer (MS) and the ID. Besides, different  $\eta$  and  $p_T$  regions demand specific approaches [55].

*Stand-alone* (SA) muons for example are solely reconstructed in the MS. By extrapolating the spectrometer track back to the beam line, the direction of flight and impact parameter of the muon is calculated. Energy losses in the calorimeters are hereby taken into account. This reconstruction algorithm relies entirely on the MS, which allows to reconstruct muons up to  $|\eta| < 2.7$ , but is in contrary vulnerable to contamination of particles that were not produced

in  $pp$  collisions.

*Segment tagged* (ST) muons on the other hand are seeded by an identified track in the ID which is extrapolated to the MS. In addition, the tracks are refitted after associating them to track segments in the MS to improve the resolution. ST muons help to enlarge the MS driven reconstruction to poorly covered regions and muons with low transverse momenta that reached the inner layer of the muon chambers, only. Difficulties in the reconstruction of MS tracks are most apparent in the transition region at  $|\eta| \sim 1.2$  (see Section 2.2.3), where only one chamber is traversed by the muons (which renders the SA reconstruction impracticable) and at  $|\eta| \sim 0$  where noticeable space is needed for services of the ID and calorimeters.

Finally, *Combined* (CB) muons are the muon candidates with highest purity. The reconstruction is carried out independently in the ID and MS. In the  $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$  analysis, the muon track combination is achieved using the so-called *STACO* algorithm [55]. Hereby, pairs of tracks are formed from reconstructed ID and MS tracks. The kinematic properties for each pair are assigned through a weighted combination of the ID and MS stand-alone measurements. A  $\chi^2$  value to quantify the likeliness of the ID and MS track to form one muon track, depending on several track quantities, is assigned. Then, all track pairs below a certain  $\chi^2$  threshold are evaluated. Subsequently, the muon candidate with the lowest  $\chi^2$  is stored and the corresponding tracks in the ID and MS are removed from the containers. Iteratively, all compatible tracks in the ID and MS are reconstructed as muon candidates until no match is left.

### 3.2.1. Reconstruction efficiency and momentum resolution of the muon

The muon reconstruction efficiency was measured in  $Z \rightarrow \mu^+\mu^-$  decays using a *tag-and-probe* method [56]. The tag muon is required to be reconstructed in both ID and MS, whereas the probe is either a SA muon, when the ID efficiency is to be measured, or a muon reconstructed by an ID track if the MS and matching efficiency is evaluated. The efficiencies and their corresponding scale factors between data and MC simulations were attained as a function of  $p_{\text{T}}$  and  $\eta$  of the probed muon [55].

The muon momentum resolution was extracted on the one hand from the width of the well known  $Z$  boson line shape<sup>2</sup> in  $Z \rightarrow \mu^+\mu^-$  events, utilizing CB muons that make use of two independent detector parts, the ID and the MS. Differences in the muon momentum resolution and a possible bias of the line shape between data and MC are thereby corrected.

On the other hand, the relative difference of two independent momentum measurements of the single muon  $W \rightarrow \mu\nu_{\mu}$  decay, carried out by either the MS or the ID, was exploited to

<sup>2</sup>The  $Z$  width is a convolution of the natural  $Z$  width and the muon momentum resolution.

correct the detector systems individually, because the difference is sensitive to the quadratic sum of both momentum resolutions. Finally, the muon momentum resolution was specified as a function of  $p_T$  and  $\eta$  for both ID and MS, resulting in a correction function to model simulated muons to match the performance in data (*smearing procedure*) [57].

### 3.3. Electron reconstruction and identification

For most Standard Model measurements and Higgs boson searches, it is essential to precisely reconstruct electrons with a high efficiency over a broad energy range<sup>3</sup>. Moreover, it is important to distinguish them from hadronic jets with a much larger production cross section.

Given the calorimetry of the ATLAS detector, an electron will usually deposit its whole energy in the electromagnetic calorimeter (EMC) (see Section 2.2.2), thus the reconstructed electrons are denoted as calibrated at the electromagnetic energy scale (EM scale). Besides, the electron as charged particle is supposed to leave a characteristic shower in the EMC that is mainly generated by photon emission via *bremsstrahlung* and photon conversion into an electron-positron pair. Hence, the challenge to precisely reconstruct and identify an electron is faced by a combination of the tracks reconstructed in the ID (for electrons within  $|\eta| < 2.47$ ) and energy depositions in the different EMC layers designed to enhance the identification efficiency (see Section 2.2.2). The acceptance of electrons utilized in this thesis is restricted to those reconstructed in the central detector region within  $|\eta| < 2.47$ , where additional information of the ID is taken into account [59].

In the central detector region, the reconstruction is initiated by energy clusters [60] in the EM that are afterwards matched to tracks from charged particles reconstructed in the ID. At first, seed clusters are searched by the *sliding-window* algorithm [60]. A rectangular window of fixed size is moved across the middle layer of the EM calorimeter, which is segmented into a grid of  $N_\eta \times N_\phi$  elements<sup>4</sup> of size  $\Delta\eta \times \Delta\phi$ . A precluster is formed whenever a local energy maximum exceeds 2.5 GeV. Finally, if the distance between two seed clusters is less than a predefined  $\Delta\eta_{\text{dupl}} \times \Delta\phi_{\text{dupl}}$ , the one possessing the lower energy is removed.

Electron candidates in the central detector are those seed clusters for which a track of the ID could be matched within specific  $\Delta\eta$  and  $\Delta\phi$  criteria. To eliminate ambiguities in case of several matched tracks, those with silicon hits are preferred and the track with the smallest  $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$  to the seed cluster is taken.

After a track was matched successfully, the electron cluster is rebuild using a different

<sup>3</sup>Inclusive electron searches showed that a background rejection of  $\sim 10^5$  is required to bring the background below the level of single isolated electrons [58].

<sup>4</sup>The energy of all cells within one element is summed over all longitudinal layers, forming a *tower*.

window. The cluster energy of this electron candidate is at last determined by summation of the measured cluster energy with estimations for the energy deposited in front of the EM calorimeter as well as estimations for longitudinal and lateral energy leakage. The  $\eta$  and  $\phi$  directions of the electron candidate are taken from the corresponding track at its vertex.

To reject jets that fake an electron candidate, three sets of reference named *loose++*<sup>5</sup>, *medium++* and *tight++* offer different working points of electron selection efficiency and background rejection. For the *loose++* identification, shower shape variables of the EM middle layer together with hadronic leakage variables are used. The latter describes energy depositions in the hadronic calorimeter which are used to distinguish between electrons and hadronically interacting particles. Adding variables from the EM strip layer together with track quantities provides the *medium++* selection. Finally, the *tight++* identification adds information to reject photon conversions along with information based on  $E^{calo}/p^{ID}$  measurements and the TRT [59].

To correct for differences in the reconstruction and identification between measured data and simulated events, the corresponding efficiencies were measured in  $Z \rightarrow e^+e^-$  events using a *tag-and-probe* method. The resulting scale factors are specified as function of  $\eta$  and  $p_T$ .

### 3.3.1. Electron energy scale and resolution

The EM scale is initially derived from test-beam measurements performed with detector components of the EM barrel calorimeter [46]. In addition, an in-situ calibration based on  $Z \rightarrow e^+e^-$  decays is applied. In the central region, the in-situ calibration is cross-checked with  $J/\Psi \rightarrow e^+e^-$  events that possess a softer  $p_T$  spectrum. Residual mis-calibrations from the test-beam calibration, which yields  $E^{\text{meas}}$ , are parametrized by

$$E^{\text{meas}} = E^{\text{true}}(1 + \alpha_i). \quad (3.1)$$

$E^{\text{true}}$  is thereby the true electron energy available in simulated events and  $\alpha_i$  measures the mis-calibration for different  $\eta$  regions  $i$ . The factor  $\alpha$  is obtained by an unbinned log-likelihood fit quantifying the compatibility of the reconstructed to the actual Z line shape.

The energy resolution was determined in  $Z \rightarrow e^+e^-$  events by fitting a Breit-Wigner function with fixed Z width convolved with a Crystal Ball function. On the one hand, the fit is used to adjust the Z peak position of MC events to match the peak obtained in data, which compensates remaining differences of the MC-based energy scale calibration above. Apart from that, smearing corrections to correct for discrepancies in the resolution between data

<sup>5</sup>The affix *++* is used to distinguish the identification criteria from those used in 2010 data.

and Monte Carlo simulations imply terms for electronic noise and the different calorimeter sampling layers [61].

### 3.4. Lepton isolation measures

To distinguish between fake leptons, leptons from semi-leptonic hadron decays and those from electroweak processes (prompt leptons), isolation criteria are an effective tool. Two kinds of isolation variables are computed for this purpose, the *track isolation* and the *calorimeter isolation*. The former defines the sum of transverse momenta of all tracks within a specified cone that satisfy  $p_T > 1$  GeV and excluding the lepton itself, normalized to the reconstructed lepton  $p_T$ . The latter determines a similar ratio based on the calorimeter energy deposits in a cone around the lepton direction, again excluding the lepton energy itself from the numerator.

Corrections to the isolation variables depend on the number of primary vertices to account for a bias from pile-up together with a correction for energy leakage of high- $p_T$  objects that is not incorporated in the object reconstruction.

### 3.5. Jet reconstruction and energy scale

In contradiction to electrons and muons, the interaction of quarks and gluons is determined by their color charge. As a characteristic of the strong interaction (QCD), they undergo a complex hadronization process until they are observed in the calorimeter as color singlet bound states of different type and charge (see Section 1.3.3). This leaves a signature of a collimated jet of particles that is detected mainly in the calorimeters and labeled as jet [62]. In addition, electrically charged constituents of the jet leave tracks in the ID. Several algorithms are available to reconstruct a jet out of its components by clustering the energy depositions and merging them accordingly.

#### 3.5.1. Jet reconstruction

The reassembling of hadronic energy clusters in the ATLAS experiment is typically carried out by the *topological clustering algorithm* [60]. Clusters are seeded whenever a certain signal-to-noise ratio  $t_{\text{seed}}$  is exceeded. Neighbour cells are added to this seed if they exceed a lower threshold  $t_{\text{low}}$ . To allow for further expansion, the neighbor cells are marked as seeds themselves if their signal-to-noise ratio is above  $t_{\text{medium}}$ .

To finalize the cluster, all direct neighbors on the outer perimeter which surpass  $t_{\text{cell}}$  are added. Direction and total energy of these combined clusters are given by their barycentre

and the sum of the individual cell energies, respectively.

At the ATLAS experiment, the common jet algorithm is the *anti-kt jet algorithm*<sup>6</sup> [63]. It clusters different objects through a distance measure. Two *entities* (particles, pseudojets)  $i$  and  $j$  are assigned a distance  $d_{i,j}$  by

$$d_{i,j} = \min(k_{t,i}^{-2}, k_{t,j}^{-2}) \frac{\Delta_{i,j}^2}{R^2}, \quad (3.2)$$

with  $\Delta_{i,j}^2 = (y_i - y_j)^2 + (\Phi_i - \Phi_j)^2$  and  $k_{t,i}$ ,  $y_i$  and  $\Phi_i$  as the transverse momentum, rapidity and azimuth of *entity*  $i$ , respectively. The parameter  $R$  specifies the implicit radius of the combined jet. This distance is compared to the distance of entity  $i$  with respect to the beam axis, given by

$$d_{i,B} = k_{k,i}^{-2}. \quad (3.3)$$

Whenever  $d_{i,j}$  is smaller than  $d_{i,B}$ , both  $i$  and  $j$  are combined, otherwise entity  $i$  is called a jet and removed from the list of objects. This sequence is repeated iteratively throughout all entities until none is left. The jet radius is an arbitrary choice, balancing between the inclusion of all genuine particles of a jet and rejecting extrinsic objects that did not originate of the actual quark or gluon.

The reconstructed jet collection still involves a number of undesired background processes. This involves calorimeter noise, cosmic ray muons or beam-gas events. As a result, certain quality criteria are adapted to reject fake jets while retaining true jet objects coming from  $pp$  collisions. These are partitioned in *Looser*, *Loose*, *Medium* and *Tight* working points and provide different background rejection and jet selection efficiencies [64]. In addition, pile-up events that are unrelated to the hard scattering are a side effect of the high instantaneous luminosity at the LHC. They introduce problems in particular for analyses relying on the jet multiplicity and jet energy determination (see Section 1.3.3). On the one hand, particles from additional proton interactions are scattered into the calorimeters, denoted *in-time pile-up* and adding energy to the transverse momentum of a reconstructed jet. On the other hand, the bunch spacing of 50 ns in 2011 introduces a sensitivity to the energy flow of past collisions, known as *out-of-time pile-up*. To account for these effects, the *jet vertex fraction* (JVF) is designed. It describes the ratio of the energy of tracks within a cone around the jet direction and originating from the primary vertex with respect to all tracks that are assigned

---

<sup>6</sup>The anti-kt algorithm has the advantage that it is infrared and collinear safe.

to one particular jet [65],

$$\text{JVF}(\text{jet } i, \text{vertex } j) = \frac{\sum_{\text{track} \in i \cap j} p_{\text{T}}^{\text{track}}}{\sum_{\text{track} \in i} p_{\text{T}}^{\text{track}}}. \quad (3.4)$$

The  $\text{JVF} \in [0, 1]$  can only be computed within  $|\eta| < 2.5$  because it relies on the track reconstruction carried out in the ID.

### 3.5.2. Jet energy scale

The *topo-clusters* that are constructed based on the description above are initially calibrated at the EM scale. Jets are a collection of topological-related energy depositions in both EM and hadronic calorimeters, which are further associated to tracks of charged particles that are measured in the ID. However, the calorimeter response to hadrons is usually lower with respect to electrons or photons due to inevitable energy losses, for example caused by neutrons escaping the detection (non-compensation) accompanied by energy depositions in inactive detector regions [66].

The jets used in the  $H \rightarrow \tau_{\text{lep}} \tau_{\text{had}}$  analysis are therefore calibrated based on the *local cluster weighting* scheme (LCW) [67]. It classifies each topo-cluster as being of either electromagnetic or hadronic nature. Based on this classification, corrections for dead material, losses due to noise threshold effects and the non-compensation of the hadronic calorimeter are applied. Moreover, energy corrections from single neutral and charged pion Monte Carlo simulations are taken into account.

Jets build of LCW clusters are afterwards corrected to match the energy scale measured from simulated jets on truth particle level. Hence, a detailed calibration scheme is exerted, set off by adjustments of the jet origin to point back to the primary vertex, followed by pile-up corrections. In addition, the inverse of the average energy response with respect to the true jet  $\mathfrak{R} = \frac{E_{\text{jet}}^{\text{LCW}}}{E_{\text{jet}}^{\text{truth}}}$  with dependence on  $\eta$  and  $p_{\text{T}}$  is applied.

On top of this energy calibration that relies on MC simulations only, an *in-situ* derived correction is used to compensate for residual differences between MC and data. It is achieved by comparing the balance between the  $p_{\text{T}}$  of a jet and a reference object,

$$\frac{\langle p_{\text{T}}^{\text{jet}} / p_{\text{T}}^{\text{ref}} \rangle_{\text{Data}}}{\langle p_{\text{T}}^{\text{jet}} / p_{\text{T}}^{\text{ref}} \rangle_{\text{MC}}}. \quad (3.5)$$

It is performed for  $\gamma$ +jet(s) events, Z+jet(s) events and events where a system of low- $p_{\text{T}}$  jets recoiled against a high- $p_{\text{T}}$  jet, in order to cover a broad momentum range.

### 3.6. Identification of $b$ -jets

In the search for a Higgs boson with one hadronically decaying  $\tau$ , it is important to reject jets originating from  $b$ -quarks, since they indicate a top quark in the final state.

A variety of distinct properties of the  $b$ -quark can be used to distinguish  $b$ -jets from jets originating from light quarks or gluons. The CKM matrix elements for the decay of the  $b$ -quark via flavor changing weak currents exhibit small transition probabilities for  $b \rightarrow c$  and  $b \rightarrow u$  on the order of 1 % [13]. Thus, the typical lifetime of a  $b$ -quark is about 1.5 ps, resulting in a flight path length on the order of several millimeters [46]. The most likely decay into a  $c$ -quark will yield a large number of tracks associated to this jet together with a high impact parameter.

To summarize, the characterization of  $b$ -jets, known as  $b$ -tagging, is performed by algorithms taking advantage of track impact parameters or reconstructed vertices of  $b$ - and  $c$ -hadron decay products. The  $IP3D$  algorithm [68] is an example of the former, utilizing a likelihood ratio technique. The signed transverse and longitudinal impact parameter significances,  $\frac{d_0}{\sigma_{d_0}}$  and  $\frac{z_0}{\sigma_{z_0}}$ , are compared to distributions obtained from Monte Carlo simulations [68].

In addition, vertex information is, for instance, implemented in the *JetFitter* algorithm. This algorithm tries to find a common line between primary vertex and the vertices of the  $b$ - and  $c$ -quark, which are reconstructed with a secondary vertex-based algorithm [68].

Moreover, these two algorithms are merged in a neural-network driven combination called *JetFitterCombNN*. On top, the output of  $IP3D$ , *JetFitterCombNN* and another secondary vertex-based algorithm (*SV1*) are combined in another neural network, the *MV1 b-tagger*.

The described  $b$ -tagging algorithms provide an output weight  $w$  that discriminates between  $b$ - and non- $b$ -jets. As the purity of  $b$ -jets increases with higher  $w$  [69], several working points resembling different nominal  $b$ -tagging efficiencies  $\epsilon_b^{sim}$  are stated for each tagging algorithm. These are derived from an inclusive sample of simulated  $t\bar{t}$  events and usually span a range of 60 % to 85 % signal efficiency [70].

Thereby, the  $b$ -tagging algorithm utilized in the  $H \rightarrow \tau_{lep}\tau_{had}$  analysis, the *JetFitterCombNN* at a working point of 70 % signal efficiency, yields a light jet rejection capability of approximately 99 % [68].

### 3.7. Reconstruction and identification of hadronic $\tau$ leptons

Tau leptons are the heaviest known leptons and the only ones decaying into hadrons. They possess a mass of 1.777 GeV and a lifetime of about  $2.9 \times 10^{-13}$  seconds [13], what results in a decay at very short distance to the interaction vertex. The  $\tau$  lepton can either decay

leptonically into electrons and muons, accompanied by two neutrinos and further denoted as  $\tau_\mu$  for  $\tau \rightarrow \mu\nu_\mu\nu_\tau$  with a branching ratio (BR) of  $BR(\tau_\mu) \approx 17.9\%$  [13], or  $\tau_e$  for  $\tau \rightarrow e\nu_e\nu_\tau$  with  $BR(\tau_e) \approx 17.4\%$  [13], respectively. It is very difficult to discriminate the leptonic  $\tau$  decay from a prompt electron or muon. Hence, no  $\tau_{\text{lep}}$  identification is realized.

The remaining branching ratio of  $BR(\tau_{\text{had}}) \approx 64.7\%$  [13] is related to its hadronic decay and further summarized as  $\tau_{\text{had}}$ . Almost all hadronic final states involve one or three charged pions, may involve neutral pions and always a  $\nu_\tau$  neutrino. Thus, the visible decay products of the hadronic  $\tau$  leave a collimated calorimeter shower. Performing a cut to maintain  $\tau_{\text{had}}$  candidates with one or three charged pion tracks only, as done in the  $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$  analysis, yields a signal efficiency of  $\sim 76\%$ , while rejecting roughly 63% of di-jet background events [71].

The  $\tau$  lepton identification is based on the visible decay product of the hadronically decaying  $\tau$ , denoted  $\tau_{\text{had-vis}}$ . Candidates are seeded by reconstructed jets with transverse momenta above 10 GeV and  $R = 0.4$ , using the algorithms described in Section 3.5.1. Due to the importance of track multiplicity and other track-based observables, the  $\tau$  acceptance is restricted to the coverage of the ATLAS detector tracking system and is therefore limited to  $|\eta| < 2.5$ . Their energy is reconstructed in a narrow cone of  $\Delta R < 0.2$  and calibrated to the true visible energy scale available in simulated events, denoted as  $\tau$  energy scale (TES) [72].

Relevant tracks are associated to a  $\tau_{\text{had-vis}}$  candidate if they lie within the core cone ( $\Delta R < 0.2$ ) and satisfy several other quality criteria. These tracks set up the *n-prongs* of the  $\tau$ . Tracks within the *isolation annulus* of  $0.2 < R < 0.4$  and satisfying the same merit as the core cone tracks are used to compute quantities used in the identification.

The task for the identification of a hadronic  $\tau$  is to distinguish it from jets initiated by quarks or gluons in QCD processes with a much higher cross section and rejecting the latter. The  $\tau$  objects used in this thesis are identified based on a Boosted Decision Tree (BDT). A BDT assigns a score to each event based on given input variables, whereby a higher BDT score refers to a more signal-like event topology (more details on the BDT methodology can be found in Chapter 7). The BDT classifier for the  $\tau$  ID is build on a set of input variables corresponding to a variety of properties based on the associated track and energy depositions [73].

- **Track and calorimeter radius ( $\mathbf{R}_{\text{track}}$  and  $\mathbf{R}_{\text{cal}}$ ):** The former describes the  $p_T$  weighted track width of all tracks associated to the  $\tau_{\text{had-vis}}$  candidate. The latter denotes the shower width in all cells of the EM and hadronic calorimeter, weighted by

the  $E_T$  of each calorimeter part:

$$R_{\text{track}} = \frac{\sum_i^{\Delta R_i < 0.4} p_{T,i} \Delta R_i}{\sum_i^{\Delta R_i < 0.4} p_{T,i}}, \quad R_{\text{cal}} = \frac{\sum_{i \in \text{cal}}^{\Delta R_i < 0.4} E_{T,i} \Delta R_i}{\sum_{i \in \text{cal}}^{\Delta R_i < 0.4} E_{T,i}}. \quad (3.6)$$

- **Leading track momentum and core energy fraction ( $\mathbf{f}_{\text{track}}$  and  $\mathbf{f}_{\text{core}}$ ):** This variables describe the transverse momentum fraction of the leading  $p_T$  core track ( $\Delta R < 0.1$ ) and the fraction of transverse energy in the core of the  $\tau_{\text{had-vis}}$  candidate, respectively. Both are normalized to the energy of all cells within  $\Delta R < 0.4$  of the  $\tau_{\text{had-vis}}$  candidate, calibrated at EM scale.
- **Number of isolation tracks ( $\mathbf{N}_{\text{track}}^{\text{iso}}$ ):** The number of tracks within  $0.2 < \Delta R < 0.4$  (isolation annulus) of the  $\tau_{\text{had-vis}}$  candidate.
- **Cluster and track mass ( $\mathbf{m}_{\text{clusters}}$  and  $\mathbf{m}_{\text{tracks}}$ ):** The invariant mass of all LCW scaled clusters that are associated to the seeding jet and of all tracks of the  $\tau_{\text{had-vis}}$  candidate within  $\Delta R < 0.4$ , respectively.
- **Maximum  $\Delta R$ :** The largest spatial distance between a core track and the associated  $\tau$  axis,  $\Delta R_{\text{max}}$ .
- **Leading cluster energy ratio ( $\mathbf{f}_{3, \text{lead clusters}}$ ):** The ratio of the three clusters with highest energy over the energy of all clusters associated to a  $\tau_{\text{had-vis}}$  candidate.
- **Significance of the flight path ( $\mathbf{S}_T^{\text{flight}}$ ) and the leading track impact parameter ( $\mathbf{S}_{\text{lead track}}$ ):** The significance is defined as ratio of the parameter value with respect to its uncertainty. These variables describe the decay length significance of the secondary vertex<sup>7</sup> and the transverse impact parameter significance, respectively.

In Figure 3.2, BDT distributions for 1-prong and 3-prong  $\tau_{\text{had-vis}}$  candidates, trained on these input variables, are displayed. A clear discrimination between signal events possessing a real  $\tau$  and a background from di-jet data events can be observed. To correct for discontinuities in the ID efficiency and provide approximately independent working points regarding the pile-up conditions, individual BDTs are constructed for various bins of the number of primary vertices.

Three distinct cut values on the BDT distributions yield a loose, medium and tight working point. These working points characterize a signal efficiency of 60 %, 50 % and 30 % for 1-prong candidates as well as 65 %, 55 % and 35 % for 3-prong candidates. According to the

<sup>7</sup>It is fitted for the core tracks of multi-prong  $\tau_{\text{had-vis}}$  candidates, since secondary vertices with only one assigned track are discarded during the reconstruction (see Section 3.1).

signal efficiencies of these working points, a background rejection of around 94 %, 96 % and 98.5 % is achieved for 1-prong together with approximately 97.5 %, 98.5 % and 99.6 % for 3-prong  $\tau_{\text{had-vis}}$  candidates [74].

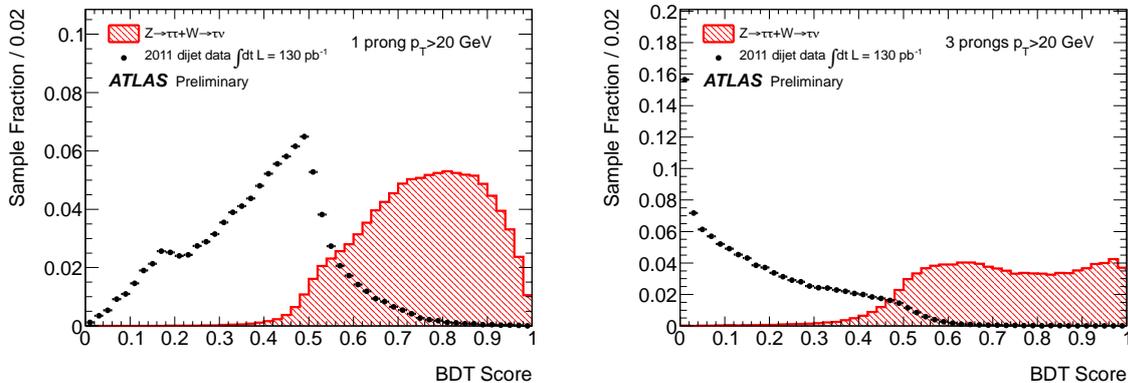


Figure 3.1.: The jet BDT score for 1-prong (left) and 3-prong (right)  $\tau_{\text{had-vis}}$  candidates. They display the distributions of  $Z \rightarrow \tau\tau + W \rightarrow \tau\nu$  events (red histograms) compared to background events composed of di-jet data events (black dots) for identified  $\tau$ s with  $p_T > 20$  GeV (plots taken from Ref. [73]).

Apart from a good signal efficiency, it is important to reject hadronic  $\tau$ -fake candidates arising from electrons or muons. For example, electrons can fake a 1-prong  $\tau_{\text{had-vis}}$  candidate. Several shower shape variables can be used to distinguish these from real hadronic taus. As for the identification, a BDT was performed specifying three working points which result in 95 %, 85 % and 75 %  $\tau_{\text{had-vis}}$  signal efficiency. It is accompanied by a background suppression of roughly 98.6 %, 99,5 % and 99.7 % [74].

Also, muons can mimic the  $\tau$  due to anomalous energy depositions or a coincidental overlap from other calorimeter depositions. In this case, cuts on different shower shape variables were optimized targeting a  $\tau_{\text{had-vis}}$  efficiency of 96 % along with a muon rejection of about 55 % [74].

The efficiency of the BDT based  $\tau_{\text{had-vis}}$  identification was determined based on a  $Z \rightarrow \tau_\mu \tau_{\text{had}}$  *tag-and-probe* measurement for each identification point, accompanied by different muon and electron veto settings [74]. An isolated muon is hereby taken as tag together with a  $\tau_{\text{had-vis}}$  candidate satisfying a probe definition that does not include any  $\tau$  identification. The identification efficiency  $\epsilon^{id}$  in the total probe sample is then defined as fraction of the probe candidates that pass different  $\tau_{\text{had-vis}}$  identifications. It was measured in data and Monte Carlo simulations. A data/MC correction factor given by

$$C_{\text{ID}}^{\text{Data/MC}} = \frac{\epsilon_{\text{ID}}^{\text{Data}}}{\epsilon_{\text{ID}}^{\text{MC}}} \quad (3.7)$$

was extracted. The results are pictured in Fig. 3.2 as a function of the  $\tau_{\text{had-vis}}$   $p_T$  for both 1-prong and 3-prong candidates.

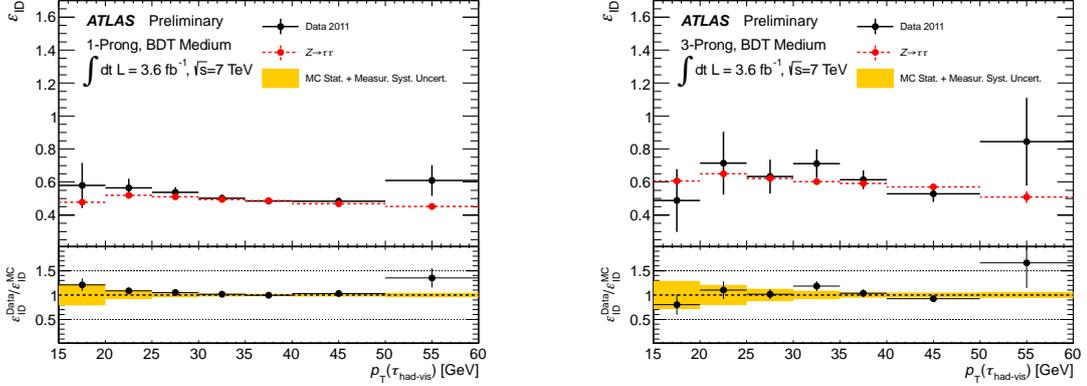


Figure 3.2.: Identification efficiencies for  $\tau_{\text{had-vis}}$  illustrated for 1-prong (left) and 3-prong (right) candidates. The plots are binned in  $p_T$  of the  $\tau_{\text{had-vis}}$  candidate and shown for the medium BDT working point. The errors on the measured data include systematic and statistical uncertainties while the error on the simulated  $Z \rightarrow \tau\tau$  efficiencies show only statistical uncertainties. In the lower part of the plots, the efficiency ratio with statistical uncertainties on the data points and MC statistical as well as systematic uncertainties of the measured data in the yellow bands is displayed (plots taken from Ref. [74]).

### 3.8. Transverse missing momentum

In  $pp$  collisions at the ATLAS detector, incoming partons inside the proton are assumed to have a negligible momentum transverse to the beam axis ( $p_T$ ). Due to conservation of the overall  $p_T$ , the missing transverse momentum  $\mathbf{E}_T^{\text{miss}}$  is defined as the event momentum imbalance in the transverse plane and obtained from the negative vector sum of the momenta of all detected particles [75]. Significant  $E_T^{\text{miss}}$  may originate from the presence of undetected particles, for example neutrinos. The reconstruction of  $E_{T,x(y)}^{\text{miss}}$  in the  $x(y)$  axis is performed by

$$E_{T,x(y)}^{\text{miss}} = E_{T,x(y)}^{\text{miss},e} + E_{T,x(y)}^{\text{miss},\gamma} + E_{T,x(y)}^{\text{miss},\tau} + E_{T,x(y)}^{\text{miss},jets} + E_{T,x(y)}^{\text{miss},soft jets} + E_{T,x(y)}^{\text{miss},cell-out} + (E_{T,x(y)}^{\text{miss},calo\mu}) + E_{T,x(y)}^{\text{miss},\mu}. \quad (3.8)$$

The equation includes energy deposits in calorimeters as well as muons reconstructed in the muon spectrometer. In regions covered by the ID ( $|\eta| < 2.5$ ), only muons with a matched track in the ID are considered (combined muons). The terms  $E_{T,x(y)}^{\text{miss},e}$ ,  $E_{T,x(y)}^{\text{miss},\gamma}$ ,  $E_{T,x(y)}^{\text{miss},\tau}$ ,  $E_{T,x(y)}^{\text{miss},jets}$  and  $E_{T,x(y)}^{\text{miss},soft jets}$  represent the contributions from the reconstructed high- $p_T$  objects based

on calorimeter cells. They correspond to electrons, photons, hadronically decaying  $\tau$ s, jets above 20 GeV as well as jets within  $7 \text{ GeV} < p_T < 20 \text{ GeV}$  (denoted as *soft jets*), respectively. Moreover, the track momentum is used to consider contributions from low- $p_T$  particles that did not reach the calorimeters or have topoclusters not assigned to any high- $p_T$  object, subsumed as  $E_{T,x(y)}^{\text{miss,cell-out}}$  [75]. To appropriately account for energy deposited by muons in the calorimeters, the term in parentheses,  $E_{T,x(y)}^{\text{miss,calo}\mu}$ , is computed separately for isolated and non-isolated muons and eventually added.

Limitations for the reconstruction of  $\mathbf{E}_T^{\text{miss}}$  are due to the finite resolution of energy measurements, the non complete coverage of the full solid angle by the detector and inactive detector material. On top, cosmic-ray muons traversing the detector and electronic noise can contribute to  $\mathbf{E}_T^{\text{miss}}$ .

In the search for new physics, the compatibility of observed data to theory predictions of the SM processes is examined to discover or exclude a possible signal at a certain probability. In addition, the sensitivity of an analysis can be expressed by an expected exclusion limit or an expected discovery significance, whereby pseudo-data is simulated with the help of MC methods, containing the expected signal and background rate.

In order to compute a statistical probability, bayesian and frequentist viewpoints have to be distinguished. Bayesian probabilities, also called evidential probabilities, express the subjective degree of belief given a prior probability which is updated in the light of new data. In contradiction, the frequentist probability interpretation describes the rate with which the measured result is true, denoted *relative frequency*. In the ATLAS collaboration, modified frequentist limits based on profile likelihood ratio tests are computed and the so-called  $CL_s$  technique [76] is used to set exclusion limits.

### 4.1. Methodology

In order to quote a discovery with a certain probability or exclude a signal down to a specific cross section, a dataset is analyzed by testing it against a set of hypotheses. The dataset is composed of either observed data or simulated events with an injected signal process of particular cross section. If the probability for the existence of a new, undiscovered process should be quantified, it is compared to background processes only which are generated by MC. Different from that, the significance of an underlying signal model is scanned if an upper limit on the signal strength within the dataset is quoted.

The probability, that the observed data can be explained by the fundamental hypothesis is called *p-value*. Sometimes, the equivalent significance  $Z$  of a standard Gaussian with cumulative distribution  $\Phi$  is stated, implying an upper-tail probability from  $Z$  to infinity that is equal to the p-value,

$$Z = \Phi^{-1}(1 - p). \quad (4.1)$$

To exclude a signal hypothesis, particle physicists tend to a required p-value of 0.05, corre-

sponding to a confidence level ( $CL$ ) of  $CL \equiv 1 - p = 95\%$  or  $Z=1.64$ .

In the following, the likelihood function and the different test statistics used for the purpose of discovery or exclusion are briefly summarized according to the discussion in Refs. [77] and [78]. Furthermore, approximation techniques to simplify the limit computation are outlined.

## 4.2. The Likelihood function

The likelihood function as premise to derive limits is constructed of Poisson distributions. The expectation value  $E[n_i]$  of the observed data  $n$  in bin  $i$  of a binned histogram can be written as

$$E[n_i] = \mu s_i + b_i. \quad (4.2)$$

In this equation,  $b_i$  characterizes the expected background events in bin  $i$  and  $s_i$  the signal estimate with  $\mu$  as scale factor for the signal strength<sup>1</sup>. The signal and background expectation in each bin is given by the total number of events ( $s_{\text{tot}}, b_{\text{tot}}$ ), weighted by the probability to end up in the range  $dx$  of bin  $i$ ,

$$s_i = s_{\text{tot}} \int_{\text{bin } i} f_s(x; \boldsymbol{\theta}_s) dx, \quad (4.3)$$

$$b_i = b_{\text{tot}} \int_{\text{bin } i} f_b(x; \boldsymbol{\theta}_b) dx. \quad (4.4)$$

The probability density functions (PDFs) for signal and background,  $f_s$  and  $f_b$ , are functions of an observable  $x$  and so-called *nuisance parameters* (NPs). The NPs provide a continuous parametrization of the effect of various systematic uncertainties on the PDF. All NPs are finally denoted by the vector  $\boldsymbol{\theta}$ .

The signal and background modelization is obtained using Monte Carlo simulations and measurements in different control regions ( see Section 5.3 ). To constrain the NPs, for instance the value of the jet energy scale, subsidiary measurements can be implemented in the likelihood function. These are carried out in distinct control regions. According to Eq. 4.2 and Eq. 4.4, the expectation value  $E[m_j]$  for the observed data  $m$  within an auxiliary measurement is given by

$$E[m_j] = u_{\text{tot}} \int_{\text{bin } j} f_u(y; \boldsymbol{\theta}) dy. \quad (4.5)$$

The observable  $y$  and the NPs of auxiliary measurements do not have to be equivalent to those used in the analysis whereupon the limit is computed.

However, normalization uncertainties with a corresponding prior are often implemented

---

<sup>1</sup> $\mu \equiv 1$  describes a cross section that is equal to the SM prediction.

in the likelihood function as simple replacement of subsidiary measurements. The priors  $\mathcal{P}$  are typically Gaussian or lognormal distributions around the evaluated mean ( $\tilde{L}$ ) and variance ( $\sigma_L$ ) of the supplementary measurement,  $\mathcal{P}(\tilde{L} \pm \sigma_L)$  [79]. Hence, Bayesian reasoning is introduced in this frequentist approach.

Finally, the likelihood function is composed of Poissonian terms for the analysis and the various subsidiary measurements, whereas the latter are often replaced by simple priors to parametrize the effect of a systematic uncertainty:

$$\mathcal{L}(\mu, \boldsymbol{\theta}) = \text{Pois}(n|\mu s(\boldsymbol{\theta}) + b(\boldsymbol{\theta})) \times \prod \text{Pois}(m|u(\boldsymbol{\theta})) \times \prod \mathcal{P}(L|\tilde{L}, \sigma_L). \quad (4.6)$$

For illustration, the likelihood function written out for an analysis with one subsidiary measurement yields

$$\mathcal{L}(\mu, \boldsymbol{\theta}) = \prod_{i=1}^N \frac{(\mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta}))^{n_i}}{n_i!} \exp^{-(\mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta}))} \prod_{j=1}^M \frac{(u_j(\boldsymbol{\theta}))^{m_j}}{m_j!} \exp^{-m_j}. \quad (4.7)$$

Chapter 8 describes how the different nuisance parameters are treated for the limit computation of the  $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$  analysis.

### 4.3. Test statistics

To interpret the likelihood function introduced above, the test statistics of the  $CL_s$  technique are based on the maximum likelihood<sup>2</sup> given by

$$\lambda(\mu, \text{obs}) = \frac{L(\mu, \hat{\boldsymbol{\theta}}(\mu, \text{obs}))}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})}. \quad (4.8)$$

$\lambda(\mu)$  describes the ratio of the likelihood maximized at a fixed  $\mu$ , thereof called *conditional* maximum-likelihood-estimator (MLE) of  $\boldsymbol{\theta}$ , denoted  $\hat{\boldsymbol{\theta}}(\mu, \text{obs})$ , normalized to the likelihood where all parameters are floating in the fit, what leads to the *unconditional* MLE with corresponding  $\hat{\mu}$  and  $\hat{\boldsymbol{\theta}}$  [81]. The former has a maximized value below the unconditional MLE for the reason that the NPs are optimized with respect to a fixed  $\mu$ .

Finally, a test statistic is chosen to evaluate the likelihood ratio, which is given by

$$q_\mu = -2 \ln \lambda(\mu) > 0. \quad (4.9)$$

Hence, a larger  $q_\mu$  corresponds to a higher incompatibility between the tested  $\mu$  and the

<sup>2</sup>The maximum likelihood is known as an unbiased estimator of the probed value ( $\mu$  in this case) that results in the best possible convergence because the variance of  $\mu$  lies on the Cràmer-Rao lower bound [80].

observed data.

To quantify this disagreement, the p-value of  $\mu$  is computed utilizing the PDF of  $q_\mu$  for a given hypothesis,  $f(q_\mu|\hat{\theta}(\mu, obs))$  (in the following denoted  $f(q_\mu|\mu)$ ). It measures the incompatibility of the hypothesis to the observed data,

$$p_\mu = \int_{q_{\mu,obs}}^{\infty} f(q_\mu|\mu) dq_\mu. \quad (4.10)$$

A thorough way of generating this PDF is accomplished by the simulation of pseudo-experiments. Therefore, every bin content for each process is varied due to a Poisson distribution reflecting the statistical uncertainty. In addition, each NP is randomized around its *conditional* fit value  $\hat{\theta}(\mu, obs)$ . The intention is to recalculate  $q_\mu$  for each pseudo-experiment and succeedingly approximate the PDF. This approach is named the *unconditional ensemble* [78] and reflects a scenario where the NPs used to generate the pseudo-experiments are taken after a fit to data that minimized the global likelihood, denoted by  $\hat{\theta}(\mu, obs)$ . The *conditional ensemble* on the other hand originates from the nominal<sup>3</sup> values for each NP.

According to Eq. 4.10 and Eq. 4.9, the test statistic obtained for the observed data can be classified within the PDF that reveals the probability to find just this incompatibility between the observed data and the underlying hypothesis. Figure 4.1 demonstrates how the p-value is obtained after computation of a probability function for  $q_\mu$ . In addition, the extraction of the significance in terms of standard deviations is pictured.

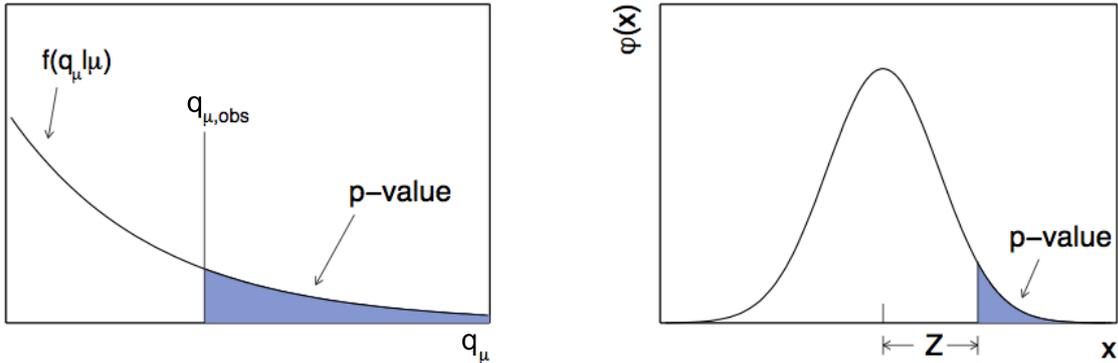


Figure 4.1.: Derivation of the p-value to set an upper limit for a given strength parameter  $\mu$  (left plot) and the corresponding significance  $Z$  in a normal Gaussian distribution  $\phi(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)$  related to this p-value (right plot) (taken from Ref. [77]).

The test statistic of Eq. 4.9 can be adjusted to serve different purposes. If for example a

<sup>3</sup>The nominal rate of each bin is attained by setting all NPs to 0 which corresponds to their mean value. Deviations of the NPs are given in units of  $\sigma$  within the limit computation.

positive signal process is taken as granted, the unconditional denominator may be restricted to positive  $\hat{\mu}$  and set to 0 for negative  $\hat{\mu}$ .

#### 4.3.1. Test statistic for discovery

If the observed data is tested against the background only hypothesis in order to probe a signal existence, the  $q_0$  test statistic is of interest. Furthermore, the hypothesis that the observed data arose from background processes only cannot be rejected if the unconditional signal strength parameter is below zero. This provides

$$q_0 = \begin{cases} -2 \ln \lambda(0) & \hat{\mu} \geq 0, \\ 0 & \hat{\mu} < 0 \end{cases} \quad (4.11)$$

which leads to a p-value of 100 % in the latter case. To finally claim discovery, the p-value

$$p_0 = \int_{q_{0,obs}}^{\infty} f(q_0|0) dq_0 \quad (4.12)$$

has to be less than  $4 \times 10^{-7}$ , corresponding to  $Z=5$ .

#### 4.3.2. Test statistic for upper limits

An upper limit quantifies the signal cross section that can be excluded at a given confidence level, usually 95 %. In order to do so, different signal strength parameters  $\mu$  are scanned to obtain  $\bar{\mu}$  with  $p_{\bar{\mu}} = 0.05$ . The corresponding test statistic is defined as

$$\tilde{q}_\mu = \begin{cases} -2 \ln \tilde{\lambda}(\mu) & \hat{\mu} \leq \mu, \\ 0 & \hat{\mu} > \mu. \end{cases} = \begin{cases} -2 \ln \frac{L(\mu, \hat{\theta}(\mu, obs))}{L(0, \hat{\theta}(0, obs))} & \hat{\mu} < 0, \\ -2 \ln \frac{L(\mu, \hat{\theta}(\mu, obs))}{L(\hat{\mu}, \hat{\theta})} & 0 \leq \hat{\mu} \leq \mu, \\ 0 & \hat{\mu} > \mu. \end{cases} \quad (4.13)$$

Data with  $\hat{\mu} > \mu$  is considered as more compatible with the signal strength tested in terms of exclusion, thus  $\tilde{q}_\mu = 0$  is taken in this case. In addition, the test statistic is limited to positive values for  $\hat{\mu}$ . In the following,  $q_\mu$  is used as substitution for  $\tilde{q}_\mu$  and  $q_0$  in general statements.

#### 4.3.3. Quoting median expected limits

To obtain an expectation on the sensitivity of an experiment, it is relevant to consider pseudo-data with a different signal strength  $\mu'$  as the one being tested ( $\mu$ ). Thus, a second sampling distribution is needed. For instance, if optimization studies are performed within an analysis

or varying theories are tested, the expected median significance for an upper limit on the signal cross section, considering no signal in the pseudo-data, is of substantial interest. Hence, pseudo-experiments with  $\mu' = 0$  have to be generated, providing  $f(q_\mu|\hat{\theta}(\mu' = 0, obs))$ . This distribution is set in contrast to the underlying hypothesis  $f(q_\mu|\mu)$  and embodies not only the median expected upper limit but their statistical variation, too, as shown in Fig. 4.2.

Besides, the expected discovery significance under assumption of a certain signal process is desired to measure the statistical impact of a presumed signal. Therefore, pseudo-experiments with  $\mu' = 1$  have to be generated. From these,  $f(q_0|\mu')$  is constructed and compared to the background-only hypothesis given by  $f(q_0|0)$ . However, this leaves the task to build two precise PDFs in order to obtain significances up to several  $\sigma$ . Thus, simplifications are needed and are presented in the following.

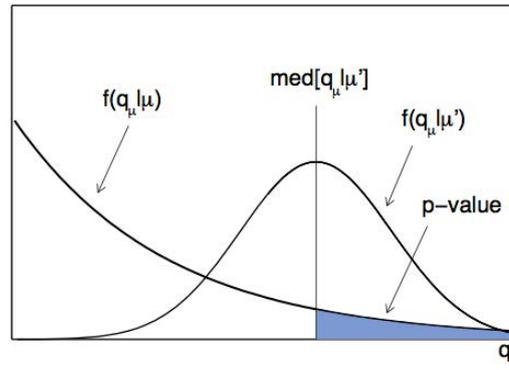


Figure 4.2.: The median of  $q_\mu$  and its resulting p-value under assumption of a different strength parameter  $\mu'$  [77].

#### 4.4. Approximate probability distributions

In the general case of the last section, a strength parameter  $\mu$  is tested against pseudo-data which is distributed corresponding to a different strength parameter  $\mu'$ . Thus, the PDF for this case,  $f(q_\mu|\theta(\mu', obs))$ , is required. Wald [82] demonstrated that  $q_\mu$  is distributed according to

$$q_\mu = -2 \ln \lambda(\mu) = \frac{(\mu - \hat{\mu})^2}{\sigma^2} + \mathcal{O}(1/\sqrt{N}), \quad (4.14)$$

whereby  $\hat{\mu}$  is following a Gaussian distribution with mean  $\mu'$  and standard deviation  $\sigma$ . The additional term corresponds to the size of the data sample,  $N$ . The mean  $\mu'$  is set by the conditional MLE,  $\hat{\theta}(\mu', obs)$ , which is used to simulate the pseudo-experiments.

Neglecting the  $\mathcal{O}(1/\sqrt{N})$  term, it can be shown [82] that  $q_\mu$  follows a *non central chi-square*

distribution<sup>4</sup> for one degree of freedom (d.o.f.),

$$f(q_\mu, \Lambda) = \frac{1}{2\sqrt{q_\mu}} \frac{1}{2\pi} \left[ \exp\left(-\frac{1}{2}(\sqrt{q_\mu} + \sqrt{\Lambda})^2\right) + \exp\left(-\frac{1}{2}(\sqrt{q_\mu} - \sqrt{\Lambda})^2\right) \right] \quad (4.15)$$

with a non centrality parameter  $\Lambda = \frac{(\mu - \mu')^2}{\sigma^2}$ , which is zero for  $\mu' = \mu$ , leading to

$$f(q_\mu, 0) = \frac{1}{\sqrt{q_\mu}} \frac{1}{2\pi} \exp^{-q_\mu/2}. \quad (4.16)$$

Hence, both PDFs of Fig. 4.2 can be approximated and all parameters, except of  $\sigma$ , are known.

#### 4.4.1. The Asimov dataset

Possessing the approximate distribution of Eq. 4.15, the undetermined Gaussian distribution of  $\hat{\mu}$  with unknown standard deviation  $\sigma$  remains to be specified. Hence, a single representative dataset with fixed  $\mu'$ , called *Asimov* dataset [46], is simulated. Besides, the *Asimov* dataset adheres NPs at their conditional MLE,  $\hat{\theta}(\mu', obs)$  [84]. Based on this, the *Asimov* profile likelihood ratio  $\lambda_A$  is derived:

$$\lambda_A(\mu) = \frac{L_A(\mu, \hat{\theta}(\mu, obs))}{L_A(\hat{\mu}, \hat{\theta})} = \frac{L_A(\mu, \hat{\theta}(\mu, obs))}{L_A(\mu', \hat{\theta}(\mu', obs))}. \quad (4.17)$$

A major advantage of the *Asimov* dataset is, that  $q_{\mu,A} = -2 \ln \lambda_A(\mu)$  is equivalent to the median limit one receives when testing  $\mu$  under assumption of a different signal strength  $\mu'$ . By means of Eq. 4.14 without higher order terms<sup>5</sup>, the variance characterizing the distribution of  $\hat{\mu}$  yields

$$\sigma_A^2 = \frac{(\mu - \mu')^2}{q_{\mu,A}}. \quad (4.18)$$

To summarize, the PDF approximation of Eq. 4.15 together with the *Asimov* dataset that allows to determine the variance needed in the case of  $\mu' \neq \mu$  are sufficient to compute an expected exclusion limit as well as an expected discovery significance.

<sup>4</sup> $q_\mu$  as test statistic was mainly chosen for the purpose to get  $f(q_\mu | \hat{\theta}(\mu, obs))$  independent of  $\theta$ . However, by introducing  $\mu'$ , the PDF is now related to a non-central chi-square distribution with dependence on  $\hat{\theta}$ . There are ongoing discussions to modify the test statistics in order to improve this behavior [83].

<sup>5</sup>Wald's approximation has the remarkable feature that higher-order terms do not alter the median as long as the monotonicity between  $q_\mu$  and  $\hat{\mu}$  holds.

## 4.5. Illustrations of the limit computation with approximative distributions

In order to quote either a discovery significance or an exclusion limit, the p-value as upper-tail integral of the PDF characterized by the underlying hypothesis needs to be computed. Therefore, the cumulative distribution related to Eq. 4.15,

$$F(q_\mu|\mu') = \Phi\left(\sqrt{q_\mu} + \frac{\mu - \mu'}{\sigma}\right) + \Phi\left(\sqrt{q_\mu} - \frac{\mu - \mu'}{\sigma}\right) - 1 \quad (4.19)$$

is required. Given this equation, the p-value and corresponding Z can be derived as

$$p_\mu = 1 - F(q_\mu|\mu') \quad \text{and} \quad Z_\mu = \Phi^{-1}(1 - p_\mu) = \Phi^{-1}(F(q_\mu|\mu')), \quad (4.20)$$

respectively.

In the case where real observed data is used to compute  $q_\mu$ , which is afterwards classified within the PDF  $f(q_\mu|\mu)$  as illustrated in Fig. 4.1,  $\mu'$  is equivalent to  $\mu$ . Thus, the cumulative distribution is reduced to

$$F(q_\mu|\mu) = 2\Phi(\sqrt{q_\mu}) - 1 \quad (4.21)$$

which leads to a p-value of

$$p_\mu = 2(1 - \Phi(\sqrt{q_\mu})). \quad (4.22)$$

As a subtlety, the special cases in Eq. 4.11 and Eq. 4.13 where  $q_\mu$  is set to zero lead to the desired behavior as they mark the hypothesis as always true, because  $\Phi(0) = 0.5$  which leads to  $F(q_\mu|\mu) = 0$  and  $p_\mu = 1$  in Eq. 4.21 and Eq. 4.22, respectively.

To quote an upper limit at a specific confidence level  $\alpha$ , the  $\bar{\mu}$  with  $p_{\bar{\mu}} = \alpha$  is necessary. Thus, Eq. 4.22 is solved for  $\mu$ , whereby  $q_\mu$  is taken from the approximation in Ref. [82],  $q_\mu = \frac{(\mu - \hat{\mu})^2}{\sigma^2}$ . The upper limit  $\mu_{up}$  is then given by

$$\mu_{up} = \hat{\mu} + \sigma\Phi^{-1}(1 - \alpha/2). \quad (4.23)$$

Note, that  $\sigma$  usually depends on the hypothesized  $\mu$ . Thus,  $\mu_{up}$  has in general to be found numerically.

In Section 4.4.1, the *Asimov* dataset was introduced with the prominent feature that  $q_{\mu,A}$  is the median of  $f(q_\mu|\mu')$ , what results in the median expected upper limit and median discovery significance, respectively. Hence, the median experimental sensitivity of  $\mu$  under assumption

of  $\mu'$ ,  $\text{med}[Z_\mu|\mu']$ , can be found by exchanging  $q_\mu$  with  $q_{\mu,A}$  in Eq. 4.20 and 4.21, leading to

$$\text{med}[Z_\mu|\mu'] = \Phi^{-1}(F(q_{\mu,A}|\mu')) = \Phi^{-1}(2\Phi(\sqrt{q_{\mu,A}}) - 1). \quad (4.24)$$

On top of a median upper limit, its variation (given by  $\hat{\mu} \pm N\sigma$ ,  $N \in \mathbb{N}$ ) is desired. The calculation follows 4.23 by replacing  $\hat{\mu}$  with  $\hat{\mu} \pm N\sigma$ , what implies

$$1 - \alpha/2 = \Phi\left(\frac{\mu_{up+N\sigma} - \mu' \pm N\sigma}{\sigma}\right) \implies \mu_{up\pm N\sigma} = \mu' + \sigma(\Phi^{-1}(1 - \alpha/2) \pm N). \quad (4.25)$$

#### 4.5.1. Blinded upper limits

Occasionally, a blinded upper limit is quoted. It is attained by pseudo-experiments that are generated using the *conditional ensemble* with  $\theta$  at its nominal values stated by the subsidiary measurements. Hence, a blinded limit utilizes no fit to data at all.

#### 4.5.2. The $CL_s$ procedure

The typical case when searching for new physics are overlapping PDFs for the background and the signal plus background hypotheses, accompanied by small signal rates. As a property of the profiled likelihood ratio, this can lead to an exclusion of  $\mu$  where the experiment is in fact not sensitive enough. To adjust for this manner, the  $CL_s$  method [85] determines the confidence level of exclusion by

$$CL_s \equiv 1 - p_s = 1 - \frac{p_{s+b}}{1 - p_b} \quad (4.26)$$

with the p-value  $p_{s+b}$  for the signal plus background hypothesis. Thereby, the p-value of the background only hypothesis ( $p_b$  with  $\mu = 0$ ) reduces the confidence level with rising compatibility of the dataset with the background only hypothesis.

Finally, the signal sensitivity  $p_s$  can be written in terms of the Asimov dataset as

$$p_s = \frac{1 - \Phi(\sqrt{q_\mu})}{\Phi(\sqrt{q_{\mu,A}} - \sqrt{q_\mu})}, \quad (4.27)$$

resulting in an upper limit of

$$\mu_{up+N\sigma} = \sigma(\Phi^{-1}(1 - \alpha\Phi(N)) + N). \quad (4.28)$$

The median upper limit,  $\mu_{up}^{med}$ , is hereby constantly taken for the calculation of  $\sigma$  [78].



---

# 5 $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$ analysis in data at $\sqrt{s} = 7$ TeV

---

Relying on the basic principles outlined in the previous chapters, the Higgs boson search in the  $\tau\tau$  final state with one leptonic and one hadronic  $\tau$  decay is presented in the following. This chapter starts by an explanation of the event selection and background estimation applied in the  $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$  analysis of the 2011 dataset recorded with the ATLAS detector. It is thereby closely related to the descriptions presented in Refs. [15, 86, 87] and lays the foundation for a topology-motivated event categorization and the multivariate approach introduced in the subsequent chapters.

The  $\tau\tau$  decay is an important channel for the search of a Higgs boson in an expected mass range up to 150 GeV (see Fig. 1.5). At a supposed Higgs boson mass of 125 GeV, the branching ratio of the  $H \rightarrow \tau\tau$  channel is  $\text{BR}[H \rightarrow \tau\tau] = 6.32 \pm 0.36 \%$  [36], which is by far the highest leptonic branching ratio<sup>1</sup>. In this thesis, the  $\tau\tau$  decay with one leptonic and one hadronic  $\tau$  in the final state is considered, contributing 46 % to the total branching ratio of the  $\tau\tau$  decay modes. Hence, the signal topology indicates exactly one electron or muon, one hadronic  $\tau$  candidate and missing transverse energy due to the escaping neutrinos from the  $\tau$  decays.

Section 5.1.1 outlines the different signal and background processes which are simulated with Monte Carlo methods and compared to the recorded data within the analysis. Thereafter, Section 5.2.1 describes the object selection based on the  $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$  topology followed by the mass reconstruction algorithm to obtain the mass of the di- $\tau$  system. Finally, in Section 5.3, data-driven background estimation techniques to model events arising from  $Z \rightarrow \tau\tau$  as irreducible background as well as jets and leptons faking the  $\tau_{\text{had}}$  candidate are presented.

## 5.1. Event samples

### 5.1.1. Signal and background processes

Several Higgs boson production and background processes are considered within this analysis and discussed in the following.

---

<sup>1</sup>The  $H \rightarrow \mu\mu$  decay for instance has a BR of  $\sim 0.022 \%$ .

### Higgs boson signature

The  $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$  process is characterized by exactly one lepton and one  $\tau_{\text{had}}$  candidate, whereby the  $\tau$  leptons are directed back-to-back in the Higgs boson rest frame. Since the SM Higgs boson has no electric charge, an opposite sign of the charges of the decay products is required. Besides, the neutrinos from the  $\tau$  decays lead to missing transverse energy which results in a not fully reconstructable final state. The number of additional jets associated with the Higgs boson decay affect the share of each Higgs boson production mode and offers distinct decay topologies to enhance the signal sensitivity. The various Higgs boson production modes are described in Section 1.4 and their distinct features are exploited within the analysis.

The gluon fusion (ggF) is the dominant production process, accompanied by zero jets at leading order. The decay products of the Higgs boson are roughly back-to-back in the laboratory frame for events without additional jets. This leads to a low missing transverse energy because both  $\tau$  decays are accompanied by neutrinos that escape the detector.

However, the ggF is subject to large QCD radiative corrections which involve quark or gluon-initiated jets. Moreover, the ggF Higgs boson production will have, on average, more associated jets than the main  $q\bar{q} \rightarrow Z \rightarrow \tau\tau$  background. A recoil of the Higgs boson against additional jets alters the back-to-back structure of the  $\tau$ s because it acquires a boost in the transverse plane. This results in a higher  $E_{\text{T}}^{\text{miss}}$  resolution and thus provides better background suppression prospects.

Another Higgs boson production mode accompanied by jets is the associated production, whereby the Higgs boson is boosted against the vector boson. Since only one lepton is desired in the final event, the selection is only sensitive to hadronic decays of the vector boson.

Finally, the VBF production mode offers a distinct jet topology. It is exploited by demanding two high- $p_{\text{T}}$  jets, the so-called *tagging jets* that possess a high di-jet mass and are well separated in  $\eta$ . Besides, little jet activity is expected near the Higgs boson decay products which are usually found between the two tagging jets in  $\eta$ .

The Higgs boson production via top anti-top fusion is set aside in the context of the  $H \rightarrow \tau\tau$  analysis because it has a production cross section of  $< 1$  % with respect to the gluon fusion at 125 GeV.

### $Z/\gamma^* \rightarrow \tau\tau$

The  $Z \rightarrow \tau\tau(+ \text{jets})$  background has the same final state as the Higgs boson since two real  $\tau$ s are present. Thus, the kinematic properties of the event are similar to the signal process which renders  $Z \rightarrow \tau\tau(+ \text{jets})$  as irreducible and most important background. Besides, it has a much higher cross section than the signal process (see Table 5.2) and in order to discriminate

it from the Higgs boson, its kinematic properties need to be determined with high precision. Therefore, efforts were made to determine the  $Z \rightarrow \tau\tau$  from data what results in the so-called *embedded* samples, described in Section 5.3.7. In phase space regions with a small amount of events in data-driven estimates, a MC sample with event filter on generator-level is used, whereby the MC jet kinematic properties are corrected based on data (see Section 5.3.7).

### W+jets

The associated production of W bosons with jets, referred to as W+jets in the following, is a major background for two reasons. It is produced with a very large cross section coupled with missing transverse energy due to the neutrinos from the W decay. Nearly all background events originate of a W boson decaying into a lepton and a neutrino, accompanied by a jet that fakes the hadronic  $\tau$ . This background can be suppressed by requirements on the angle between the hadronic  $\tau$  and  $E_T^{\text{miss}}$  because the W boson is typically boosted against the jet. Therefore, the main contribution of this process to the  $\tau_{\text{lep}}\tau_{\text{had}}$  final state arises from events where an additional jet is faking the hadronic  $\tau$ .

### QCD

The di- and multi-jet backgrounds are important in every Higgs boson analysis due to its very high production cross section. Events with one jet faking the hadronic  $\tau$  and another one faking an electron or muon can imitate the signal topology. In addition, real leptons can arise from semi-leptonic decays of B and D mesons through weak interactions. The phase space of di- and multi-jet backgrounds with a mis-identified lepton and hadronic  $\tau$  is very small and not necessarily well modeled in MC generators. Moreover, the quark-gluon fraction of jets and the  $\tau$  fake rate is known to be mis-modeled in simulated events. Thus data-driven techniques are implemented to predict the background from QCD events, described in detail in Sections 5.3 and 5.3.8.

### $t\bar{t}$

The decay of two top quarks into two bottom quarks and W bosons ( $t\bar{t} \rightarrow W^+bW^-\bar{b}$ ) is an important background. Because of the W bosons, real  $E_T^{\text{miss}}$  is present in the final state together with electrons, muons and  $\tau$  leptons. In addition, the b-jets and hadronically decaying W bosons can fake the hadronic  $\tau$ . Due to b quarks in the final state, this background can be suppressed with b-tagging (see Section 3.6).

### $Z/\gamma^* \rightarrow \ell\ell + \text{jets}$

The decay of  $Z/\gamma^*$  into an electron or muon pair with additional jets is present as background because both a lepton and a jet can be misidentified as hadronic  $\tau$ . However, the latter can be diminished by a di-lepton veto on lepton candidates after loose ID requirements. The  $E_{\text{T}}^{\text{miss}}$  of this background is caused by the finite  $E_{\text{T}}^{\text{miss}}$  resolution (see Section 3.8).

### Single-top

Single tops can be produced in the t- or s-channel or in association with a W boson, denoted tW-channel, as illustrated in Fig. 5.1. The top quark will further decay weakly through a W boson and may contribute as background if the W decays into a lepton. In addition, the  $\tau$  lepton is a misidentified jet (dominantly from a b quark) or a real  $\tau$  from the decay of the second W boson.

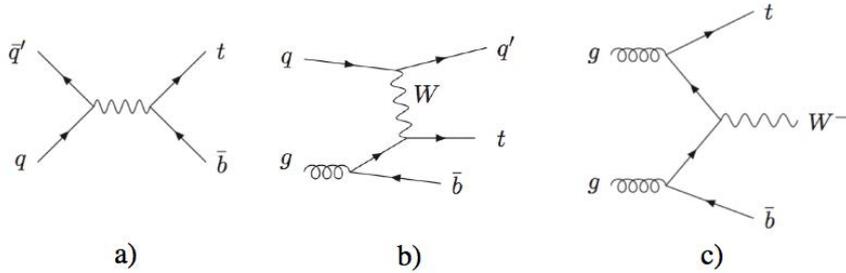


Figure 5.1.: Tree level Feynman diagrams of single top production in the s-channel (a), t-channel (b) and tW-channel (c) (taken from Ref. [88]).

### Di-boson WW/ZZ/WZ

Pairs of vector bosons can be produced via electroweak interactions. They either arise from quark radiation or quark-antiquark fusion. The signal topology can be faked by a WW, ZZ or WZ pair if one or both vector bosons decay into leptons. Also, a hadronic decay of one vector boson into  $q\bar{q}$  or  $qq'$  needs to be considered since a quark jet can fake the hadronic  $\tau$  candidate. However, this background represents only a small fraction of the total background in the  $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$  analysis.

#### 5.1.2. Data samples

The results of this thesis are based on data of proton-proton collisions at a center-of-mass energy of  $\sqrt{s} = 7$  TeV, recorded with the ATLAS detector in 2011. From the dataset with a corresponding integrated luminosity of  $5.61 \text{ fb}^{-1}$  delivered by the LHC in 2011, ATLAS

recorded  $5.25 \text{ fb}^{-1}$ . This corresponds to 93.6 % where all ATLAS sub-detectors were fully operational [89]. Some datasets are rejected if they fail quality criteria monitored by a data-quality sub-group which leads to an integrated luminosity of the relevant dataset of  $4.60 \pm 0.08 \text{ fb}^{-1}$ . The dataset exerted in the  $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$  analysis includes events that were triggered by either a single lepton trigger (SLT) or by a combined lepton+tau trigger (LTT). The latter can exploit a lower lepton  $p_{\text{T}}$  threshold due to the additional  $\tau$  object in the trigger requirement. It is considered if the event did not fulfill the SLT  $p_{\text{T}}$  threshold to avoid overlap, thus increasing the event yield of the analysis.

For the 2011 dataset, the data taking is split into eleven *periods* (B-M). They reflect different operation conditions at the ATLAS detector and the LHC. The trigger thresholds and isolation criteria applied in each data period used in the analysis are shown in Table 5.1. The value behind its corresponding object (e for electron, mu for  $\mu$  and tau for  $\tau$ ) indicates the  $p_{\text{T}}$  threshold (in GeV) used.

This trigger selection bears different efficiencies for both simulated and data events. Thus, a correction factor which is measured in a dedicated CR is applied to the simulated events. In the analysis, muons with a  $p_{\text{T}} < 25 \text{ GeV}$  are rejected due to a mis-modeling between data and simulated events. It was observed in the  $p_{\text{T}}$ -spectrum of all events containing a muon after preselection (Fig. 5.2). The origin is up-to-date not understood and under investigation.

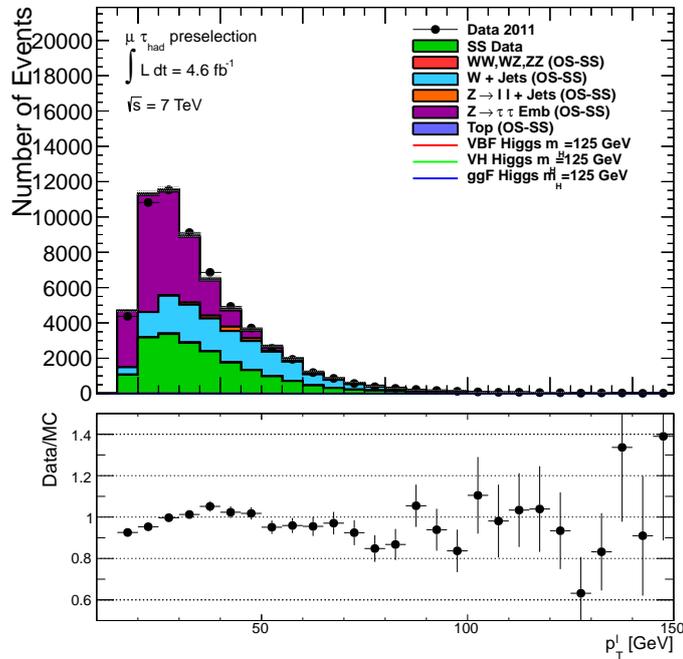


Figure 5.2.: Transverse momentum of the muon in MeV after preselection. The first two bins in the lower ratio plot display a mis-modeling between data and MC events below 25 GeV.

Table 5.1.: The specific triggers at event filter (EF) level used in particular run periods together with their integrated luminosity. The quantity succeeding the object (e for electron, mu for  $\mu$  and tau for  $\tau$ ) specifies the corresponding  $p_T$  threshold in GeV. The *medium* denotes certain ID criteria applied on trigger level. The *vh* following the electron threshold indicates the use of a hadronic core veto and variable thresholds at L1 trigger level (see Section 2.3). Finally, the *MG* in the muon trigger naming specifies that the *TrigMuonGirth* trigger algorithm was used [90, 91].

Period	electron trigger	muon trigger	electron+tau trigger	$\mathcal{L}$ [ $\text{pb}^{-1}$ ]
B-I	EF_e20_medium	EF_mu18_MG	EF_tau16_loose_e15_medium	1480.4
J	EF_e20_medium	EF_mu18_MG_medium	EF_tau16_loose_e15_medium	226.4
K	EF_e22_medium	EF_mu18_MG_medium	EF_tau20_medium_e15_medium	561.7
L-M	EF_e22vh_medium1	EF_mu18_MG_medium	EF_tau20_medium_e15vh_medium	2392.8
Total				4.6 $\text{fb}^{-1}$

### 5.1.3. Simulated events

To interpret the recorded data, relevant processes are simulated via Monte Carlo methods. A multitude of different generators is available to model the hard process [92–94], whereby the properties of the final state particles are given by probabilities derived in matrix element calculations (see Section 1.3).

The leptonic decays of the  $W$  and  $Z/\gamma^*$  bosons, accompanied by up to five jets, are generated via ALPGEN [95]. It is a tree-level matrix element (ME) event generator for up to five jets. As a subtlety, two distinct  $Z/\gamma^* + jets$  samples are generated for invariant di-lepton masses of  $10 \text{ GeV} < m_{\ell\ell} < 40 \text{ GeV}$  and  $m_{\ell\ell} > 40 \text{ GeV}$  in order to model the low-mass Drell-Yan tail separately. The single top  $s$ -,  $t$ - and  $tW$ -channels are generated by ACERMC [96].

$WW, WZ$  and  $ZZ$  di-boson samples and the decay of top quark pairs ( $t\bar{t}$ ) are generated with MC@NLO [97], which computes all diagrams at next-to-leading order, including virtual emissions. However, the gluon-induced  $gg \rightarrow WW$ , mediated through a box diagram, is modeled by GG2WW [98].

Concerning the signal samples, the gluon and vector-boson fusion processes are modeled with POWHEG [99]. Moreover, the Higgs boson radiation from a vector boson  $qq' \rightarrow WH/ZH$  is modeled by PYTHIA.

Various PDF sets are used within the MC event generators. The CTEQ6L1 PDFs [100] are implemented in the ALPGEN event generator. For the PYTHIA and ACERMC MC generators, the MRST2008 PDF set [101] is used. Finally, MC@NLO and POWHEG employ the CTEQ6.6 PDF set [100] for event generation.

Apart from the hard interaction, the initial and final state radiation (*parton showering*) together with the *hadronization* process need to be addressed. Hence, the simulated hard process is interfaced to a separate generator. The various background processes of the analysis are interfaced to HERWIG [102], while the signal processes are interfaced to PYTHIA [103].

In order to select the partons which are interfaced to the parton shower generator, the constituents of the generated event are split at an arbitrary cut-off scale to separate between partons of the hard process and those who evolved subsequently. However, this procedure bears the risk that some phase space regions are double-counted. If a hard, large-angle emission occurred during the evolution of a  $N$ -parton configuration, this can lead to a final state with  $(N+1)$  partons. Yet, these events may be covered by the appropriate  $(N+1)$ -parton configuration. To avoid this, the *MLM matching* scheme [104] is implemented in the ALPGEN event generator.

In addition, the underlying event of the various processes is modeled by JIMMY [105]. To model the decay of  $\tau$  leptons and QED radiations in the leading-logarithm approximation,

TAUOLA [106] and PHOTOS [107] are used. Finally, every simulated event is passed through a full simulation of the ATLAS detector incorporated in the GEANT4 [108, 109] framework which mimics the detector response and allows to reconstruct the event similar to data.

### Cross sections

The cross sections for the generated processes are computed in the perturbative expansion. For the Higgs boson production cross sections, up to next-to-next-to-leading-order (NNLO) QCD corrections are derived in all except the VBF production mode. In this case, the cross section is calculated at approximate NNLO. For each Higgs boson production mode, NLO electroweak corrections, assuming factorization between the QCD and EW calculations, are applied [37]. For the gluon fusion, a resummed computation of the Higgs boson transverse-momentum  $p_T$  is performed via HQT [36, 110]. It has full NNLL accuracy and is matched to NLO results at high  $p_T$ . The particular branching ratios of each Higgs boson production process are calculated through HDECAY [111].

The production cross sections of the Z and  $W^\pm$  gauge bosons are determined at NNLO with the DYNNLO [112] and FEWZ [113] generators, respectively. The single top quark production in the t-channel takes NNLO collinear and soft gluon corrections on the cross section into account [114]. For its s-channel production, a NNLL resummation of both corrections is calculated [115]. In addition, a soft-gluon resummation at NNLL accuracy is determined for the associated production of a single top quark with a W boson, which allows for an approximate NNLO accuracy of the production cross section [116]. Moreover, the diboson production cross sections are calculated at NLO using MCFM [117]. To conclude, the top quark pair production cross sections are calculated with HATHOR at approximative NNLO in QCD [118].

To obtain the expected number of events at a given integrated luminosity  $\int Ldt$ , an *event weight*  $w$  is applied on all simulated events  $N_{\text{events}}$  of a corresponding process:

$$w = \frac{\sigma \times BR \times k_{\text{factor}} \times \epsilon_{\text{filter}}}{N_{\text{events}}} \times \int Ldt, \quad (5.1)$$

where the  $k$ -factor estimates the corrections through higher orders in perturbation. The filter efficiency  $\epsilon_{\text{filter}}$  accounts for an event filter applied during MC event generation.

A summary of the production cross sections together with their corresponding generators is given in Table 5.2.

Table 5.2.: Cross sections  $\times$  BR of the  $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$  signal and background processes used in the analysis [119]. In case of a lepton in the final state, they are stated for one leptonic decay. The respective cross sections are derived including corrections up to the quoted perturbative order. The  $\sim$  NNLO denotes a computation at approximative NNLO. Furthermore, the employed event generator,  $k$ -factor and filter efficiencies are specified.

Process	Event generator	$\sigma$ [pb] · BR	$k$ -factor	$\epsilon_{\text{filter}}$
<b>Signal</b> ( $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$ )		$(m_H = 125 \text{ GeV})$		
ggF	POWHEG + PYTHIA	0.445 (NNLO)	1.0	1.0
VBF	POWHEG + PYTHIA	0.035 ( $\sim$ NNLO)	1.0	1.0
WH/ZH	PYTHIA	0.017 / 0.0092 (NNLO)	1.0	1.0
<b>Backgrounds</b>				
$W \rightarrow \ell + \text{Jets}$ ( $\ell = e, \mu, \tau$ )	ALPGEN + HERWIG	$8.70 \times 10^3$ (NNLO)	1.2	1.0
$Z/\gamma^* \rightarrow \ell\ell + \text{Jets}$ ( $M_{\ell\ell} > 40 \text{ GeV}$ )	ALPGEN + HERWIG	$0.856 \times 10^3$ (NNLO)	1.25	1.0
$Z/\gamma^* \rightarrow \ell\ell + \text{Jets}$ ( $10 \text{ GeV} < M_{\ell\ell} < 40 \text{ GeV}$ )	ALPGEN + HERWIG	$3.18 \times 10^3$ (NNLO)	1.22	1.0
$t\bar{t}$	MC@NLO + HERWIG	164.57 ( $\sim$ NNLO)	1.0	1.0
Single top				
$t$ -channel	ACERMC	6.93 (NNLO)	1.0	1.0
$s$ -channel	ACERMC	0.5 (NNLO)	1.0	1.0
$tW$ -channel	ACERMC	15.6 ( $\sim$ NNLO)	1.0	1.0
WW	MC@NLO (gg2WW)	4.84 (NLO)	1.0	0.327
WZ	MC@NLO	4.0 (NLO)	1.0	1.0
ZZ	MC@NLO	1.08 (NLO)	1.0	1.0

#### 5.1.4. Pile-up re-weighting

Following the simulation, the event is typically re-weighted to resemble an underlying reference. For instance, pile-up effects arising from simultaneous events initiated by the same or neighbor bunches (see Sections 3.5.1 and 1.3.3) are incorporated in the event simulation and modeled by predefined parameters. However, the real distribution of pile-up in recorded data is unknown a priori. As the contribution from pile-up events to the reconstruction of event objects and its energy measurements plays a significant role, differences between simulation and ATLAS data are corrected by re-weighting the distribution of the average number of interactions per bunch crossing  $\langle \mu \rangle$ .

## 5.2. Event and object preselection

Starting from events fulfilling the trigger conditions stated in Section 5.1.2, these have to fulfill certain quality criteria to be used in the  $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$  analysis. For example, an event has to possess at least four tracks associated to the primary vertex and events with failures in the calorimeter measurements are rejected. Moreover, events involving jets or electrons in inefficient LAr detector regions or accompanied by significant detector noise are dismissed [120]. The latter is determined via pulse shape differences of the measured pulse from the expectations obtained from simulations.

The event topology of the  $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$  analysis involves one isolated charged lepton together with one hadronic  $\tau$  and missing transverse energy. Therefore, different object requirements are imposed. The work of this thesis was performed within the Higgs to  $\tau\tau$  subgroup at the ATLAS collaboration, therefore the object preselection and background estimation described in the following are closely related to the results published in Refs. [15, 86, 119].

### 5.2.1. Object selection

Exactly one well isolated lepton in the central detector region with a  $p_{\text{T}} > 25$  GeV for the SLT and  $17 \text{ GeV} < p_{\text{T}}^e < 25$  GeV for the LTT trigger [121] is required. Electron candidates with an associated track in the transition region of the EM calorimeter ( $1.37 < |\eta| < 1.52$ ) are rejected due to the low reconstruction efficiency, high amount of dead material and complicated simulation conditions. A veto on additional leptons is applied to suppress events from  $Z \rightarrow \ell\ell$  ( $\ell = e, \mu$ ),  $t\bar{t}$  and single top processes. To enhance the rejection power of the di-lepton veto, a looser electron identification together with a discarded isolation requirement for the muon is chosen.

In addition, a hadronic  $\tau$  lepton identified at medium ID working point (see Section 3.7) with  $p_T^\tau > 20$  GeV and  $p_T^{\tau} > 25$  GeV is required for the SLT and LTT triggers, respectively. The lepton and the hadronic  $\tau$  are required to have an opposite charge.

Since tracks and calorimeter energy can be used to reconstruct and identify different objects at the same time, those that overlap in direction within  $\Delta R < 0.2$  are removed in a sequence with a priority given by  $\mu$ ,  $e$ ,  $\tau$ , jet. The jets participating in the analysis are reconstructed by the anti- $k_T$  jet algorithm based on LCW calibrated topo-clusters and a cone size of  $R = 0.4$ . In addition, the jets are required to have a  $p_T > 25$  GeV together with a JVF  $> 0.75$  in the barrel region (for details, see Section 3.5.1).

A detailed scheme of the object selection is pictured in Fig. B.1. Figures 5.3 and 5.4 compare data and simulated events for various kinematic distributions in the  $e + \tau_{\text{had}}$  and  $\mu + \tau_{\text{had}}$  channel, respectively. They demonstrate that the data are well understood. Table 5.3 displays the numbers of events passing the object selection described above for the predicted signal and background processes as well as the number of observed events in data. The utilized background estimation is described in Section 5.3.

The hereby defined *preselection* sets the foundation for the subsequent analyses described in the following chapters.

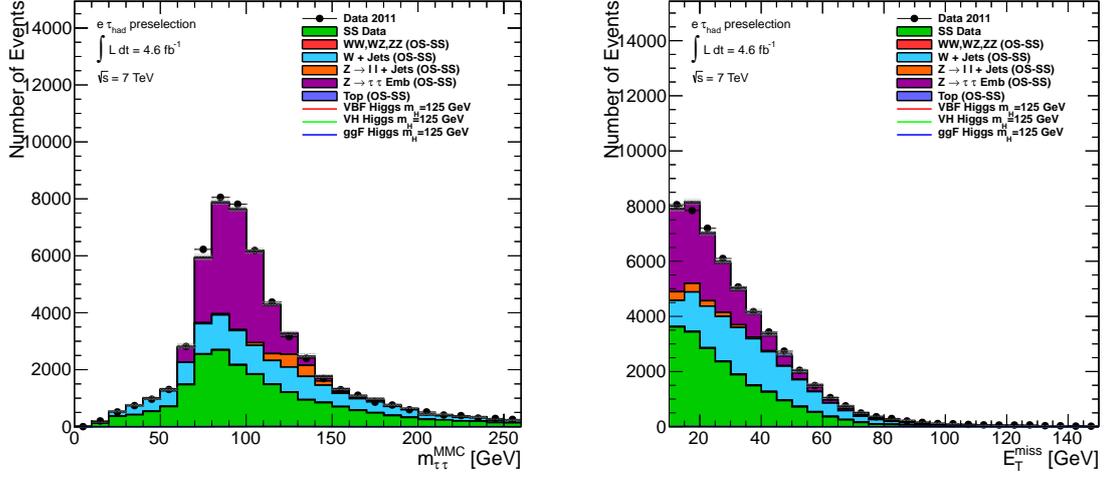
### 5.2.2. $\tau\tau$ mass reconstruction

Once a possible Higgs boson candidate is selected, it is desired to determine the invariant mass of the  $\tau\tau$  system. Due to the  $Z$  mass of 91.2 GeV [13], a possible Higgs boson at approximately 125 GeV with its small production cross section would be visible as a *bump* in the high mass tail of the most important  $Z \rightarrow \tau\tau$  background. If the  $\tau$  momenta could be fully reconstructed, the invariant mass would be derived as  $m_{\tau\tau} = |p_{\tau_1} + p_{\tau_2}|^2$ . However, in the  $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$  analysis, one neutrino is produced in the hadronic decay together with two neutrinos of the leptonic  $\tau$  decay. Since only the total transverse missing energy of the event is known, the four-vectors of the neutrinos cannot be determined.

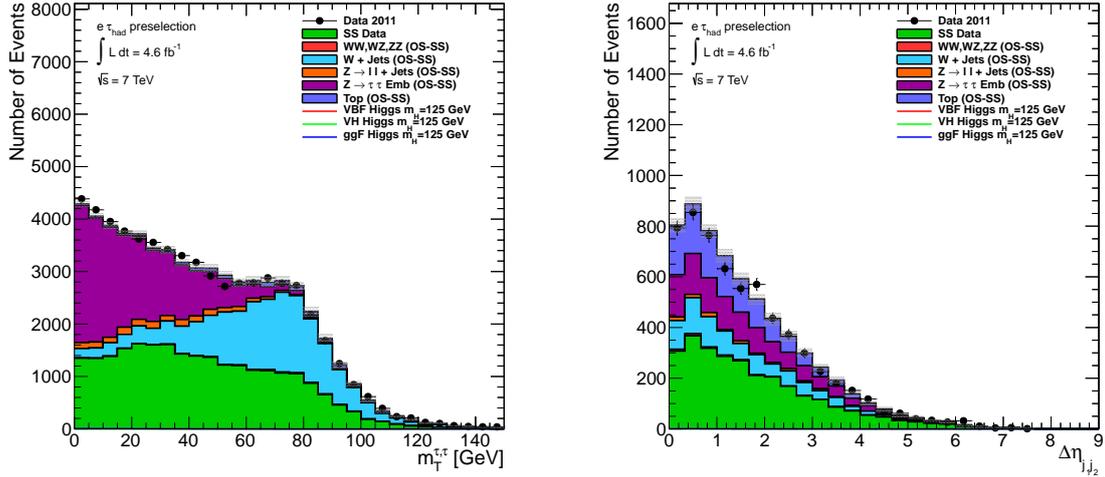
This characteristic demands for a mass reconstruction with the use of approximative algorithms. Thus, to reconstruct the invariant mass, assumptions have to be made for the different neutrinos that arise from the  $\tau$  decays, whereby the two neutrinos of the leptonic  $\tau$  decay are usually combined in a neutrino system.

Table 5.3.: Number of events passing the preselection enlisted for the single electron and muon trigger channels together with the electron +  $\tau_{\text{had}}$  trigger channel. The various background processes are estimated via the OS-SS method described in Section 5.3.1. OS refers to opposite sign charges and SS to same sign charges between the lepton and the  $\tau_{\text{had}}$ . The  $Z \rightarrow \tau\tau$  background is obtained through an *embedding* procedure, which is outlined in Section 5.3.7. The corresponding uncertainties are statistical only.

Process	Events			
	Total	Single electron trigger	Combined electron+tau trigger	Single muon trigger
Signal VBF, $m_H = 125$ GeV	10.19 ± 0.07	4.49 ± 0.05	1.10 ± 0.02	4.60 ± 0.05
Signal WH, $m_H = 125$ GeV	3.72 ± 0.09	1.63 ± 0.06	0.35 ± 0.03	1.74 ± 0.07
Signal ZH, $m_H = 125$ GeV	2.13 ± 0.05	0.94 ± 0.04	0.20 ± 0.02	0.97 ± 0.04
Signal ggF, $m_H = 125$ GeV	104.59 ± 0.80	44.7 ± 0.5	13.1 ± 0.3	46.79 ± 0.53
$Z \rightarrow \tau\tau$ embedded (OS-SS)	30961.74 ± 95.51	11777.68 ± 65.72	5531.94 ± 36.24	13651.12 ± 54.50
WW,WZ,ZZ (OS-SS)	578.59 ± 8.49	282.10 ± 5.95	22.52 ± 1.36	273.77 ± 5.90
$Z \rightarrow \ell\ell$ +Jets (OS-SS)	2526.37 ± 62.01	1469.95 ± 30.44	74.91 ± 7.12	981.44 ± 23.51
Top (OS-SS)	2865.67 ± 14.47	1470.23 ± 10.49	103.37 ± 2.54	1291.46 ± 9.64
W+Jets (OS-SS)	29147.85 ± 374.113	14497.66 ± 252.89	895.32 ± 54.87	13754.46 ± 273.09
SS Data	39339.00 ± 204.95	20782.00 ± 144.16	3516.00 ± 59.30	15042.06 ± 133.06
Total background	107126 ± 441.19	50279.63 ± 300.21	10144.06 ± 88.88	44615.68 ± 309.80
Data	106073	50584	10411	45078

(a) Reconstructed mass of the  $\tau\tau$  system.

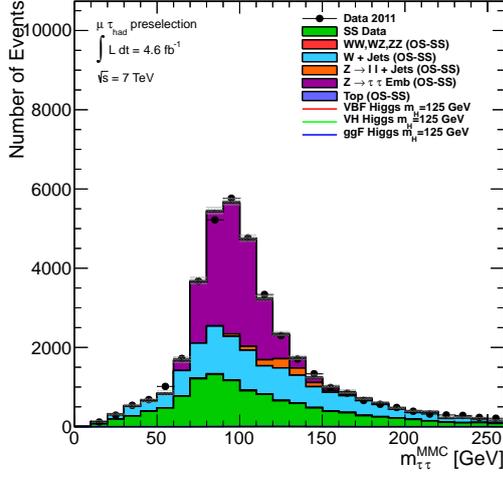
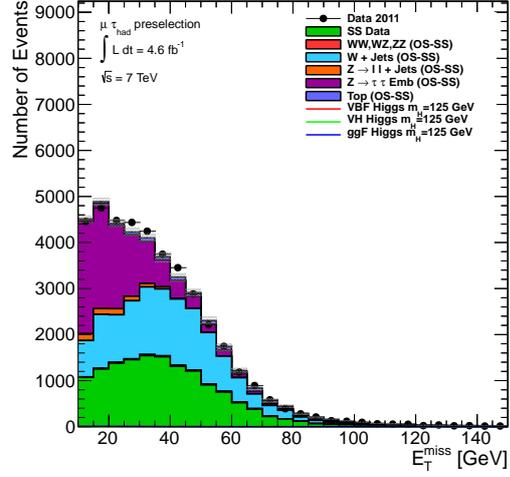
(b) Missing transverse energy.

(c) Transverse mass of electron and  $E_T^{\text{miss}}$ .(d)  $|\Delta\eta_{jj}|$  between the two jets with highest  $p_T$ .Figure 5.3.: Distributions of several kinematic variables in the  $e + \tau_{\text{had}}$  channel after the preselection described in Section 5.2.1.

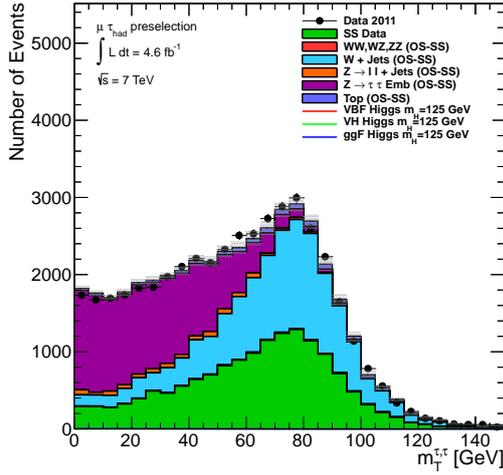
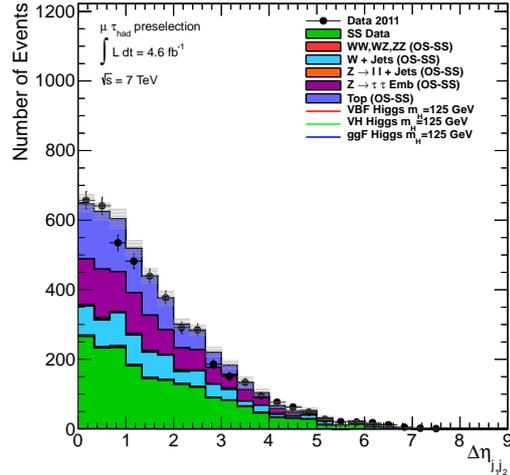
Hence, there are seven unknowns<sup>2</sup> for only four equations of the di-tau system:

$$\begin{aligned}
 E_x^{\text{miss}} &= \mathbf{p}_{\text{miss}_1} \sin \theta_{\text{miss}_1} \cos \phi_{\text{miss}_1} + \mathbf{p}_{\text{miss}_2} \sin \theta_{\text{miss}_2} \cos \phi_{\text{miss}_2}, \\
 E_y^{\text{miss}} &= \mathbf{p}_{\text{miss}_1} \sin \theta_{\text{miss}_1} \sin \phi_{\text{miss}_1} + \mathbf{p}_{\text{miss}_2} \sin \theta_{\text{miss}_2} \sin \phi_{\text{miss}_2}, \\
 m_{\tau_1}^2 &= m_{\text{miss}_1}^2 + m_{\text{vis}_1}^2 + 2\sqrt{\mathbf{p}_{\text{vis}_1}^2 + m_{\text{vis}_1}^2} \sqrt{\mathbf{p}_{\text{miss}_1}^2 + m_{\text{miss}_1}^2} - \\
 &\quad 2\mathbf{p}_{\text{vis}_1} \mathbf{p}_{\text{miss}_1} \cos \Delta\theta_{\nu m_1}, \\
 m_{\tau_2}^2 &= m_{\text{vis}_2}^2 + 2\sqrt{\mathbf{p}_{\text{vis}_2}^2 + m_{\text{vis}_2}^2} \sqrt{\mathbf{p}_{\text{miss}_2}^2 + m_{\text{miss}_2}^2} - \\
 &\quad 2\mathbf{p}_{\text{vis}_2} \mathbf{p}_{\text{miss}_2} \cos \Delta\theta_{\nu m_2}.
 \end{aligned} \tag{5.2}$$

<sup>2</sup>Due to the massless neutrinos, the x-, y- and z- components of the invisible momentum for both neutrinos as well as the invariant mass of the leptonic neutrino system are undefined.

(a) Reconstructed mass of the  $\tau\tau$  system.

(b) Missing transverse energy.

(c) Transverse mass of muon and  $E_T^{\text{miss}}$ .(d)  $|\Delta\eta_{jj}|$  between the two jets with highest  $p_T$ .Figure 5.4.: Distributions of several kinematic variables in the  $\mu + \tau_{\text{had}}$  channel after the preselection described in Section 5.2.1.

In this system of equations,  $E_{x,y}^{\text{miss}}$  are the x- and y-components of the  $E_T^{\text{miss}}$  vector and  $\mathbf{p}_{\text{vis},1,2}$ ,  $m_{\text{vis},1,2}$ ,  $\theta_{\text{vis},1,2}$  and  $\phi_{\text{vis},1,2}$  are the momenta, invariant masses, polar and azimuthal angles of the visible  $\tau$  products [36]. The unknown parameters are the momenta of the neutrino (system) for each  $\tau$  decay ( $\mathbf{p}_{\text{miss},1,2}$ ) together with the invariant mass of the neutrinos of the leptonic  $\tau$ ,  $m_{\text{miss},1}$ . Moreover,  $\Delta\theta_{\nu m,1,2}$  are the angles between  $\mathbf{p}_{\text{miss},1,2}$  and  $\mathbf{p}_{\text{vis},1,2}$  of each  $\tau$  lepton.

A straightforward approach to assign the leftover unknown parameters is to align all neutrinos with their corresponding visible  $\tau$  products. This is motivated by the large difference

between the assumed Higgs boson mass and the  $\tau$  mass that results in a high boost of the  $\tau$ s. The thereby reconstructed invariant mass is referred to as *collinear mass*.

In the  $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$  analysis, a more sophisticated approach, the missing mass calculator (MMC) [122] is applied. The MMC performs a scan in the  $(\phi_{\text{miss}_1}, \phi_{\text{miss}_2}, m_{\text{miss}_1})$  parameter space, leading to fully defined  $\mathbf{p}_{\text{miss}_{1,2}}$  and  $\mathbf{p}_{\text{vis}_{1,2}}$  vectors for every scan point. Hence, the system of equations is solved and allows to specify the  $\mathbf{p}_T$  of both  $\tau$ s together with a reconstructed  $m_{\tau\tau}^{\text{MMC}}$ .

For each grid point in the parameter space, the corresponding  $m_{\tau\tau}^{\text{MMC}}$  is weighted by the probability that this configuration occurs, which is derived in simulated  $Z/\gamma^* \rightarrow \tau\tau$  events. The probability is a compound of different terms, namely the  $\Delta\theta_{3D}$  between the directions of visible and invisible  $\tau$  decay products, their momenta and the  $E_T^{\text{miss}}$  resolution.

The  $\Delta\theta_{3D}$  are computed for bins of 5 GeV in the initial  $\tau$  momentum in a range between 10 GeV  $< p_\tau < 230$  GeV [122]. Besides, the  $\Delta\theta_{3D}$  distributions are provided for leptonic, 1- and 3-prong hadronic  $\tau$  decays, separately. To simplify the computation, these distributions are parametrized using a linear combination of Gaussian and Landau functions.

Examples of the probability distribution function  $\mathcal{P}(\Delta\theta, p_\tau)$  fitted to simulated decays are shown for a true  $\tau$  momentum of 45 GeV  $< p_\tau < 50$  GeV in Fig. 5.5.

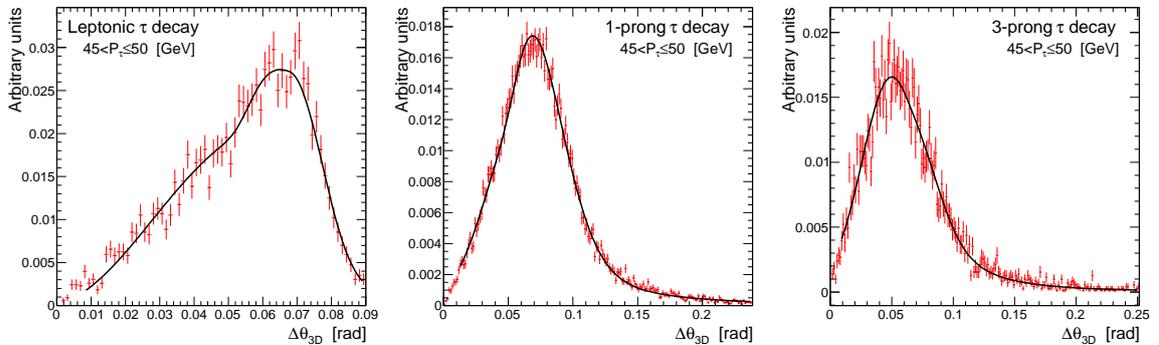


Figure 5.5.: Probability distribution functions  $\mathcal{P}(\Delta\theta, p_\tau)$  for a true  $\tau$  lepton momentum in the range 45 GeV  $< p_\tau \leq 50$  GeV for leptonic (left), 1-prong hadronic (middle) and 3-prong hadronic decays (right).

As the performance of the MMC is highly correlated to the initial  $E_T^{\text{miss}}$ , its resolution is taken into account by inclusion of a scan on  $E_{x,y}^{\text{miss}}$  within three standard deviations, whereby  $\sigma$  is provided by the ATLAS  $E_T^{\text{miss}}$  group ([75]). Thus, the final event probability is given by

$$\mathcal{P}_{\text{event}} = \mathcal{P}(\Delta\theta_1, p_{\tau 1}) \times \mathcal{P}(\Delta\theta_2, p_{\tau 2}) \times \mathcal{P}(\Delta E_x^{\text{miss}}) \times \mathcal{P}(\Delta E_y^{\text{miss}}), \quad (5.3)$$

whereas the  $E_{x,y}^{\text{miss}}$  probabilities are defined as Gaussian probabilities according to the expected

resolution:

$$\mathcal{P}(E_{x,y}^{\text{miss}}) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\Delta E_{x,y}^{\text{miss}})^2}{2\sigma^2}\right), \quad (5.4)$$

where the difference of scan value and initial  $E_{\text{T}}^{\text{miss}}$  is denoted  $\Delta E_{x,y}^{\text{miss}}$ .

The efficiency and resolution of the *collinear approximation* and the MMC algorithm are compared in Fig. 5.6 and a significant improvement in terms of mass resolution of the MMC algorithm is clearly visible.

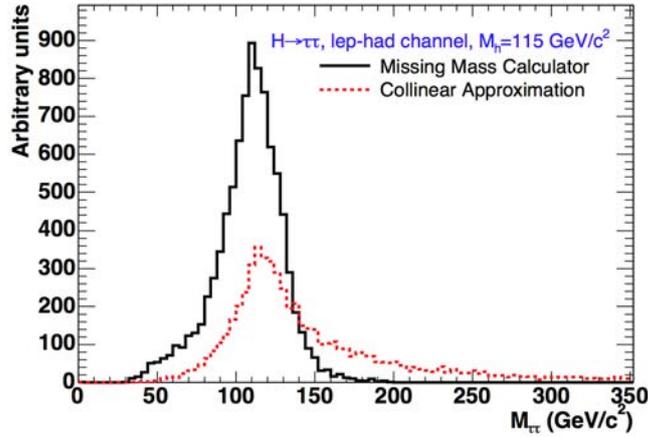


Figure 5.6.: Distributions of the reconstructed mass for the  $\tau\tau$  system. Therefore, Higgs boson events produced via gluon fusion at a mass of  $m_H = 115$  GeV are taken. The MMC (black solid line) is hereby compared to the collinear approximation (red dashed line). The normalization difference reflects a higher efficiency of the MMC method and a tail up to high masses in the collinear approximation is observed. It arises from roughly back-to-back  $\tau$  leptons (plot taken from Ref. [122]).

### 5.3. Background estimation

The background processes described in Section 5.1.1 can be classified into three different groups, whether the  $\tau_{\text{had}}$  is a true  $\tau$  or a mis-identified jet or lepton ( $e/\mu$ ).

The  $Z \rightarrow \tau\tau$  background provides a true lepton and  $\tau_{\text{had}}$  signature, which renders it as most important and irreducible background. These background events have a high charge correlation  $N_{OS} \gg N_{SS}$ <sup>3</sup> between both lepton and  $\tau_{\text{had}}$  since they arise from a neutral particle. Other background processes with a true lepton and a  $\tau_{\text{had}}$  signature are di-boson (VV with V=W,Z),  $t\bar{t}$  and single top decays.

Many background events contain a mis-identified  $\tau_{\text{had}}$  candidate. These  $\tau$  fakes can result e.g. from a jet that fulfills the  $\tau$  specifications. This occurs mainly in QCD, W+jets and top

<sup>3</sup>An opposite signed charge (OS) between the lepton and  $\tau_{\text{had}}$  is declared if  $q_l \cdot q_\tau < 0$ . Otherwise, the final state is identified as event with same sign charges (SS).

events. In addition, a lepton can fake the  $\tau_{\text{had}}$ , typically in the  $Z/\gamma^* \rightarrow ee/\mu\mu$  process.

The tau fake rate, both from leptons or jets, is badly modeled in simulations and it is therefore intended to extract these contributions from data. On top, a data-driven background estimation might reduce systematic uncertainties that arise from background predictions solely based on simulation.

Two background estimation techniques trying to model the  $\tau_{\text{had}}$  fakes by data are described in the following. Fake backgrounds are therefore taken from either same sign data events, denoted *OS-SS* method, or events using preselected  $\tau_{\text{had}}$  candidates that fail the  $\tau_{\text{had}}$  identification, the so-called *fake factor* method (see Section 5.3.8). The corresponding uncertainties are summarized in Section 8.1.

### 5.3.1. OS-SS method

The OS-SS background estimation method assumes that the shape of kinematic variables is identical for OS and SS events that pass the kinematic selection cuts of the analysis.

Hence, the expected number of background events  $N_{\text{OS}}^{\text{bkg}}$  is given by

$$N_{\text{OS}}^{\text{bkg}} = r_{\text{QCD}} \cdot N_{\text{SS}}^{\text{data}} + N_{\text{add-on}}^{Z \rightarrow \tau\tau} + N_{\text{add-on}}^{Z \rightarrow \ell\ell(\ell \rightarrow \tau)} + N_{\text{add-on}}^{Z \rightarrow \ell\ell + \text{jet}(\rightarrow \tau)} + N_{\text{add-on}}^{W + \text{jets}} + N_{\text{add-on}}^{\text{top}} + N_{\text{add-on}}^{\text{VV}}, \quad (5.5)$$

whereby the  $N_{\text{SS}}^{\text{data}}$  accounts for QCD events as well as contributions from other fake backgrounds. Thus, the residual backgrounds are added from MC simulations after subtraction of their estimated SS amount and called *add-on*. Whenever a sufficiently pure control region (CR) is accessible for a given background, a scaling factor  $k = N^{\text{data}}/N^{\text{MC}}$  is derived to account for differences between the simulation and the observed data for this background.

The  $r_{\text{QCD}}$  factor is obtained in a dedicated CR and compensates for differences in the flavor composition of the final state jets that are selected by the charge requirement and fake the  $\tau_{\text{had}}$  or lepton.

### 5.3.2. $N_{\text{SS}}^{\text{Data}}$ and $r_{\text{QCD}}$

The QCD background with a jet that fakes a prompt<sup>4</sup> lepton together with a jet that fakes the  $\tau_{\text{had}}$  is very hard to predict via simulations. Thus, same sign data events that fulfill the object selection stated in Section 5.2.1, except from the charge correlation requirement, are taken as estimate for QCD events.

<sup>4</sup>Non-prompt leptons from semi-leptonic decays of heavy flavor quarks are usually not isolated as described in Section 3.4 and therefore rejected in this analysis.

Therefore, the fraction between OS and SS di- and multi-jet events is estimated in a QCD-enriched control region and leads to a correction factor for  $N_{\text{SS}}^{\text{data}}$  represented by  $r_{\text{QCD}}$ . Neither missing transverse energy nor a high transverse mass of the lepton and  $\tau_{\text{had}}$  is expected for QCD events. This leads to a control region defined by  $E_{\text{T}}^{\text{miss}} < 15$  GeV and  $m_{\text{T}} < 30$  GeV<sup>5</sup> together with a BDT-loose working point for the  $\tau_{\text{had}}$  identification and a non-isolated lepton.

Contributions of electroweak and top background processes are subtracted based on simulations, whereby W+jets background events are normalized to data in the opposite and same sign region, respectively (see Section 5.3.3). To ensure a similar shape of di-jet events in the OS and SS region, several kinematic properties are compared. Figure 5.7 depicts an example of well compatible distributions for  $m_{\tau\tau}^{\text{MMC}}$ .

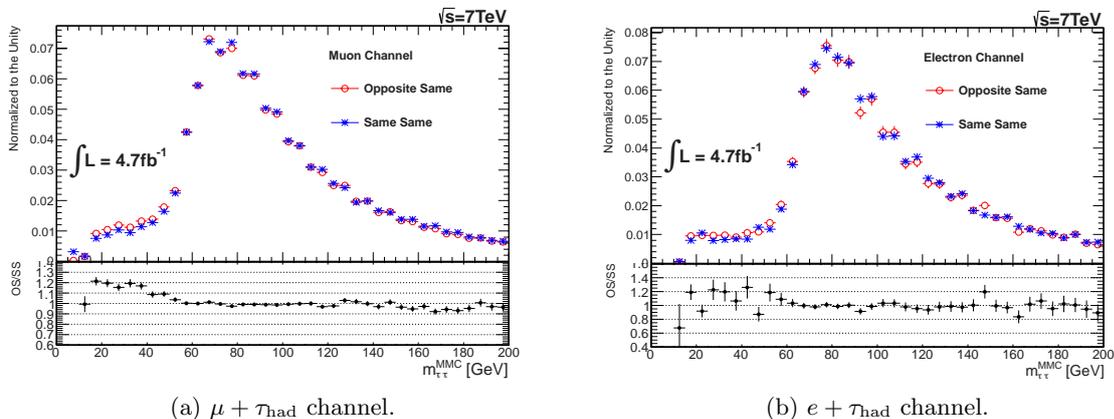


Figure 5.7.: Comparison of the opposite- and same-sign distributions of the reconstructed mass of the di- $\tau$  system (MMC) in the muon (left) and electron (right) channels shown for the QCD control region (plots taken from Ref. [87]).

The  $OS/SS$  ratio was determined within the CRs for different values of calorimeter and track isolation, denoted as  $etcone/p_{\text{T}}$  and  $ptcone/p_{\text{T}}$ , respectively. To estimate  $r_{\text{QCD}}$  for events passing the default lepton isolation requirements, a linear fit (see Fig. 5.8) was performed and extrapolated to the actual isolation used in the analysis.

Afterwards, the final  $r_{\text{QCD}}$  value was determined as mean value of both fits, which results in  $r_{\text{QCD}} = \frac{1}{2}(r_{\text{QCD}}^{etcone/p_{\text{T}}} + r_{\text{QCD}}^{ptcone/p_{\text{T}}})$  with an uncertainty defined to be  $\sigma_{\text{QCD}} = \frac{1}{2}|r_{\text{QCD}}^{etcone/p_{\text{T}}} - r_{\text{QCD}}^{ptcone/p_{\text{T}}}|$ . Additional uncertainties are derived by re-computation of  $r_{\text{QCD}}$  with a BDT-medium  $\tau$  ID requirement. Both uncertainties were added in quadrature as conservative approach. The dependence of  $r_{\text{QCD}}$  on the two isolation variables for a BDT-loose and BDT-medium  $\tau$  ID requirement in the  $\mu + \tau_{\text{had}}$  channel is shown in Fig. 5.8.

<sup>5</sup>The transverse mass is defined in Section 6.3.

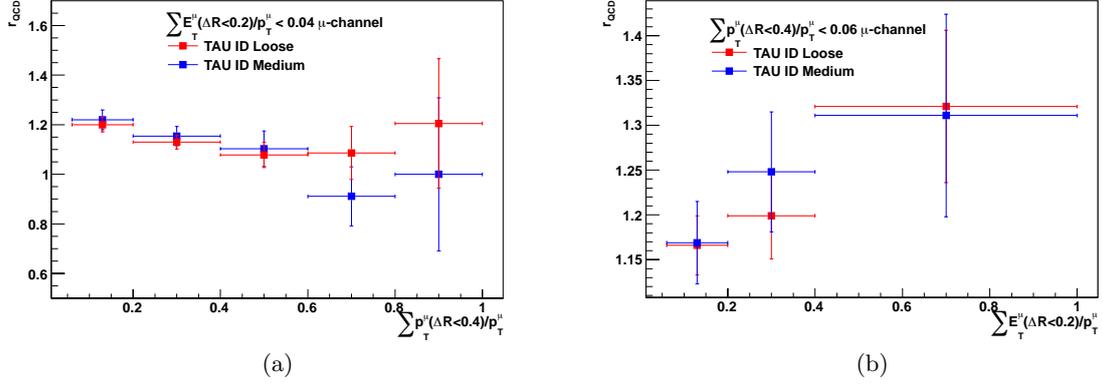


Figure 5.8.: Dependence of  $r_{\text{QCD}}$  on the track (left) and calorimeter (right) isolation variables for BDT-loose (red dots) and BDT-medium (blue dots) ID requirements, illustrated for the muon channel. A linear fit is performed to obtain the quantity of  $r_{\text{QCD}}$  for  $etcone/p_T = 0.04$  and  $ptcone/p_T = 0.06$ , the actual lepton isolation demanded in the analysis (figure taken from Ref. [87]).

Finally, the values of  $r_{\text{QCD}}$  used in the analysis are

$$r_{\text{QCD}}^{e+\tau_{\text{had}}} = 1.00 \pm 0.05, \quad r_{\text{QCD}}^{\mu+\tau_{\text{had}}} = 1.18 \pm 0.07. \quad (5.6)$$

### 5.3.3. W+jets estimation and control samples

The dominating fraction of W+jets background events contain a leptonic W decay in combination with a jet faking the  $\tau_{\text{had}}$  candidate. An excess in simulated events over those in collision data is observed after  $\tau$  identification cuts, caused by a mis-modeled quark-gluon fraction and  $\tau$  fake rate in simulated events. To account for this mis-modeling, a scale factor is derived in a dedicated control region defined by a high transverse-mass cut of  $m_T > 70$  GeV after preselection (Section 5.2.1).

The  $m_{\tau\tau}^{\text{MMC}}$  distributions for this control region are shown in Fig. 5.9. A clear excess of simulated events over observed data is visible. The scale factor  $k_W$  adjusts the number of simulated W+jets events in the control region ( $N_{W,\text{CR}}^{\text{MC}}$ ) to the expectation of W+jets events in the observed data. To obtain the latter, residual background contributions in the CR ( $N_{\text{non-W},\text{CR}}^{\text{MC}}$ ) are taken from MC simulation and subtracted from the observed data in the CR. Thus, the scale factor  $k_W$  is given by

$$k_W = \frac{N_{\text{CR}}^{\text{data}} - N_{\text{non-W},\text{CR}}^{\text{MC}}}{N_{W,\text{CR}}^{\text{MC}}}. \quad (5.7)$$

It is computed for each category defined in the following Chapters 6 and 7, separately. More-

over, the  $k_W$  are derived individually for opposite and same sign regions to account for  $\tau$  fake differences from quark- and gluon-initiated jets. The OS region has a higher amount of jets from quark hadronization, as the charge is coupled if the quark radiated the W boson. In contradiction, the same sign region has more gluon-initiated jets.

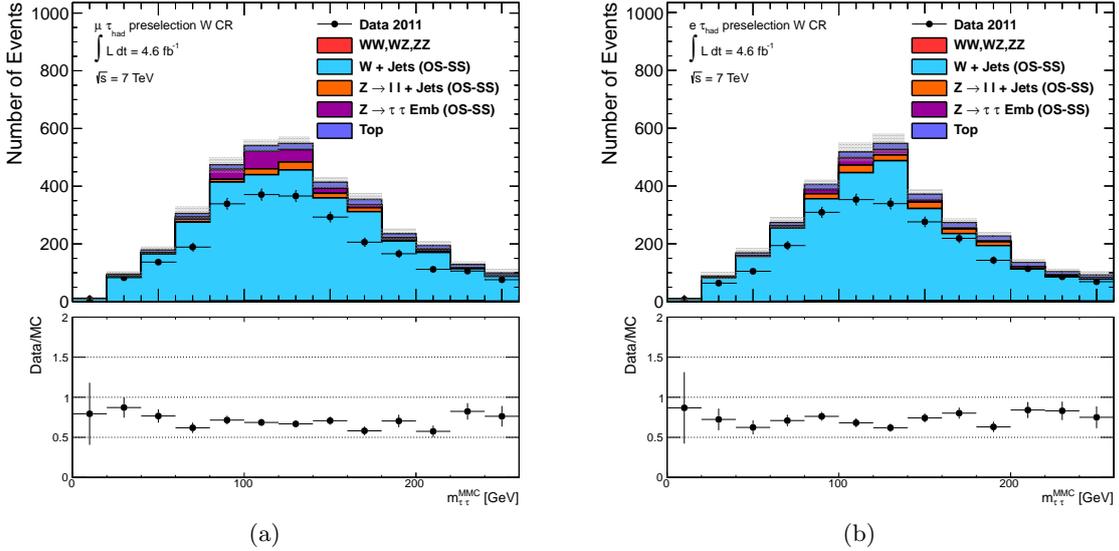


Figure 5.9.: Distributions of  $m_{\tau\tau}^{\text{MMC}}$  in the W control region for all events after preselection, shown for the OS  $\mu + \tau_{\text{had}}$  (left) and  $e + \tau_{\text{had}}$  (right) channel, separately. The lower plots display the data-to-MC ratio of the corresponding distributions, including their statistical uncertainties.

The uncertainties assigned to the  $k_W$  originate from statistical uncertainties (Eq. 5.7) propagated into the normalization uncertainty of the W+jets background estimation. Systematic variations are incorporated into the  $k_W$  by supplementary computations for every systematic variation.

#### 5.3.4. $Z \rightarrow \ell\ell + \text{jets}$ estimation

The  $Z \rightarrow \ell\ell + \text{jets}$  background has to be split into events where the lepton faked the  $\tau_{\text{had}}$  and those with a jet faking the  $\tau_{\text{had}}$ , in the following denoted by the object in parentheses. For the  $Z \rightarrow \mu\mu(\mu \rightarrow \tau)$  background, no data-driven estimation is applied, since the  $\mu$ -veto used in the  $\tau_{\text{had}}$  ID during preselection has a very high efficiency which results in a negligible background fraction ( $< 1\%$ ) of this specific background.

### $Z \rightarrow ee(e \rightarrow \tau)$

Concerning the  $Z \rightarrow ee(e \rightarrow \tau)$  background which accounts for approximately 34 % of the  $Z \rightarrow \ell\ell + \text{jets}$  background after preselection, a control region is defined after object selection as described in Section 5.2.1. The invariant mass of the electron and visible  $\tau_{\text{had}}$  candidate is required to be within  $80 \text{ GeV} < m_{e,\tau} < 100 \text{ GeV}$ . Events with any additional jets are rejected and only  $\tau$ s with one identified track are taken into consideration, because an electron mainly mimics a 1-prong  $\tau$  decay. To further enhance the  $Z \rightarrow ee(e \rightarrow \tau)$  background and diminish signal contributions, the transverse mass has to be between  $30 \text{ GeV} < m_T < 40 \text{ GeV}$ .

A scale factor  $k_{Z \rightarrow ee(e \rightarrow \tau)}$  is derived similar to Eq. 5.7. For the computation of the k-factor as ratio between data and simulated events, other background contributions estimated by MC are subtracted from data. This includes events where the reconstructed  $\tau$  is not matched to a true electron in simulation. Finally, a scale factor of  $k_{Z \rightarrow ee(e \rightarrow \tau)} = 1.0 \pm 0.24$  is used within the analysis. The uncertainty results from the statistical uncertainties of the data and MC events in the control region.

### $Z \rightarrow \ell\ell + \text{jet}(\rightarrow \tau)$

To compensate for the mis-modeled  $\text{jet} \rightarrow \tau_{\text{had}}$  fake rate in simulated events, this particular background is adjusted in a CR of two oppositely charged leptons ( $ee$  or  $\mu\mu$ ) that satisfy the same identification criteria as the analysis plus a  $\tau_{\text{had}}$  satisfying the BDT-medium ID criteria.

To account for differences in the quark-gluon fraction of the jets faking the  $\tau_{\text{had}}$  candidate, the CR is split into events with 0, 1 or 2 additional jets of  $p_T > 25 \text{ GeV}$ . After subtraction of non  $Z \rightarrow \ell\ell + \text{jet}(\rightarrow \tau)$  background events, including those where a true lepton is matched to the identified  $\tau$ , the k-factor is derived as ratio of data to simulated events. Assigned uncertainties are based on the statistical uncertainties of data events and background predictions.

Thus, a scale factor of  $k_{Z \rightarrow \ell\ell + \text{jet}(\rightarrow \tau)} = 0.56 \pm 0.03$  (stat.) is used for the 0- and 1-jet categories, whereas  $k_{Z \rightarrow \ell\ell + \text{jet}(\rightarrow \tau)} = 1.00 \pm 0.34$  (stat.) is assigned to VBF signal regions.

#### 5.3.5. $t\bar{t}$ estimation and control samples

The  $t\bar{t}$  background plays an important role in the VBF Higgs boson production mode, as it is characterized by multiple high- $p_T$  jets in association with the lepton and  $\tau_{\text{had}}$ . Thus, a distinct CR is used to correct the MC@NLO prediction.

It is defined based on the preselection with an additional transverse mass requirement ( $m_T > 30 \text{ GeV}$ ), at least two jets with  $p_T > 25 \text{ GeV}$  and at least one b-jet. The *b-tagging* efficiency is chosen to be 70 % (for *b-tagging*, see Section 3.6). The k-factors are measured

for  $e + \tau_{\text{had}}$ ,  $\mu + \tau_{\text{had}}$  and in both OS and SS, separately.

Concerning the overall uncertainty of the top background estimation, statistical uncertainties arising from the data and MC events are added in quadrature to the systematic errors originating from the b-tagging efficiency uncertainty, the  $\tau$  ID uncertainty and the  $\tau$  and jet energy scale uncertainties (see Chapter 3 and Section 8.1).

The  $t\bar{t}$  scale factors used within the analysis are

$$\underbrace{k_{\text{top}}^{\text{OS}} = 0.99 \pm 0.11, k_{\text{top}}^{\text{SS}} = 1.13 \pm 0.22}_{e+\tau_{\text{had}} \text{ channel}} \text{ and } \underbrace{k_{\text{top}}^{\text{OS}} = 0.99 \pm 0.10, k_{\text{top}}^{\text{SS}} = 1.13 \pm 0.22}_{\mu+\tau_{\text{had}} \text{ channel}}. \quad (5.8)$$

### 5.3.6. Di-boson background

The background with two heavy vector bosons in the final state is the smallest background in this analysis with an event yield of  $< 1\%$  (see Table 5.3) after preselection. Therefore, its predictions are taken from Monte Carlo simulation.

### 5.3.7. $Z \rightarrow \tau\tau$ background

As described above, the  $Z \rightarrow \tau\tau$  process has the same final state and similar event properties as the signal, but a much higher cross section. Due to the  $Z$  boson mass of 91.2 GeV, a potential Higgs boson at  $\sim 125$  GeV would be observed as a small *bump* in the high  $Z$  mass tail. Hence, it is crucial to model the  $m_{\tau\tau}$  shape well.

The  $Z \rightarrow \tau\tau$  background is modeled from data whenever a sufficient number of events is available. Otherwise, simulated events with a truth-object filter on generator level are taken after application of a data-based correction according to jet-related quantities.

### Embedding procedure

The naive approach to take  $Z \rightarrow \tau\tau$  events directly from data cannot be realized because the identification of hadronic  $\tau$ s is not efficient enough and the distinction between the  $Z$  and a possible Higgs boson is not entirely achievable. Thus,  $Z \rightarrow \mu\mu$  decays are chosen as a nearly signal free sample to serve for a data-driven estimation of  $Z \rightarrow \tau\tau$  decays. Thereby, the  $Z$  candidate can be selected with a high efficiency and its kinematic properties do not depend on the particular decay.

In the *embedding* procedure, the reconstructed tracks of the muon and its energy depositions in the calorimeters are subtracted from the  $Z$  decay, whereby the energy depositions are obtained from MC simulation. The tracks and calorimeter depositions of the muons are replaced with those of  $\tau$ s, which are taken from MC simulation and possess the kinematic

properties of the initial muons ( $p_\tau = \sqrt{E_\mu^2 - m_\tau^2}$ ). The  $\tau$  decay is simulated by TAUOLA [106], whereas final state QED radiations are simulated by PHOTOS [107].

Finally, the  $Z \rightarrow \tau\tau$  decay with replaced muons is submitted to the full event reconstruction chain to account for the modified track and cell quantities. The embedding procedure is illustrated in Fig. 5.10.

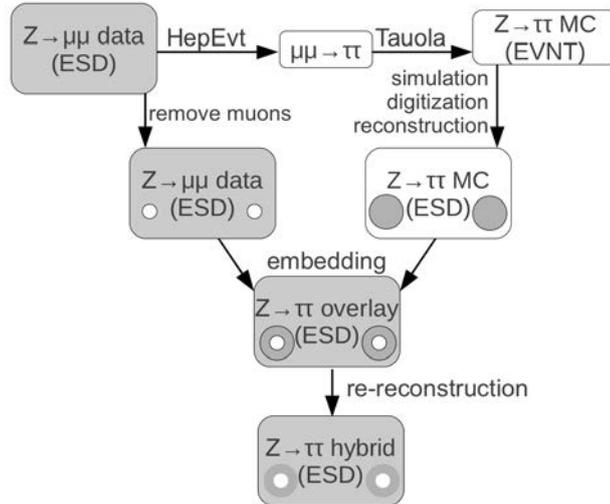


Figure 5.10.: Scheme of the embedding procedure seeded by a data event with two isolated muons possessing a high momentum. ESD (event summary data) is a dataset format containing a detailed output of the detector reconstruction for selected events. It allows to perform particle reconstruction and identification algorithms. EVNT is a type of dataset for MC events. It is generated from HepEvt, an object oriented event record for high energy physics MC generators. The EVNT dataset is passed through a full simulation of the ATLAS detector to obtain the ESD format for simulated events.

Since the trigger decision cannot be implemented in the procedure, the embedded sample has to be normalized to MC in the end. Therefore, the amount of embedded events is adjusted to those in simulated  $Z \rightarrow \tau\tau$  events after preselection.

Systematic effects on top of those applied on the simulated objects are included by variation of the muon isolation criteria and the simulated cell energy prior to their subtraction (see Section 8.1).

### VBF-filtered samples

The amount of events within the embedded sample is limited to the integrated luminosity recorded by the ATLAS detector and the embedding efficiency, what inhibits a precise modeling in the characteristic VBF topology. Thus, an alternative estimation based on MC is used, referred to as *VBF-filtered* sample. To populate the phase space region of interest, a loose filter on jets is applied on truth particle level of ALPGEN MC events.

The filter demands two jets with  $p_T > 15$  GeV that possess an invariant mass of  $m_{jj} > 200$  GeV and a  $|\Delta\eta_{jj}| > 2.0$ .

Since the VBF filter procedure is based on simulated events, a comparison of jet-related kinematic distributions from  $Z \rightarrow \mu\mu$  events<sup>6</sup> between data and non-filtered ALPGEN MC was performed and revealed a mis-modeling as expected, since especially soft jet radiation is not well described within simulations. To compensate for this, a re-weighting of the  $|\Delta\eta_{jj}|$  distributions is carried out in order to get a similar shape between MC and data events. Figure 5.11 illustrates the  $|\Delta\eta_{jj}|$  distributions in the  $Z \rightarrow \mu\mu$  CR before and after re-weighting.

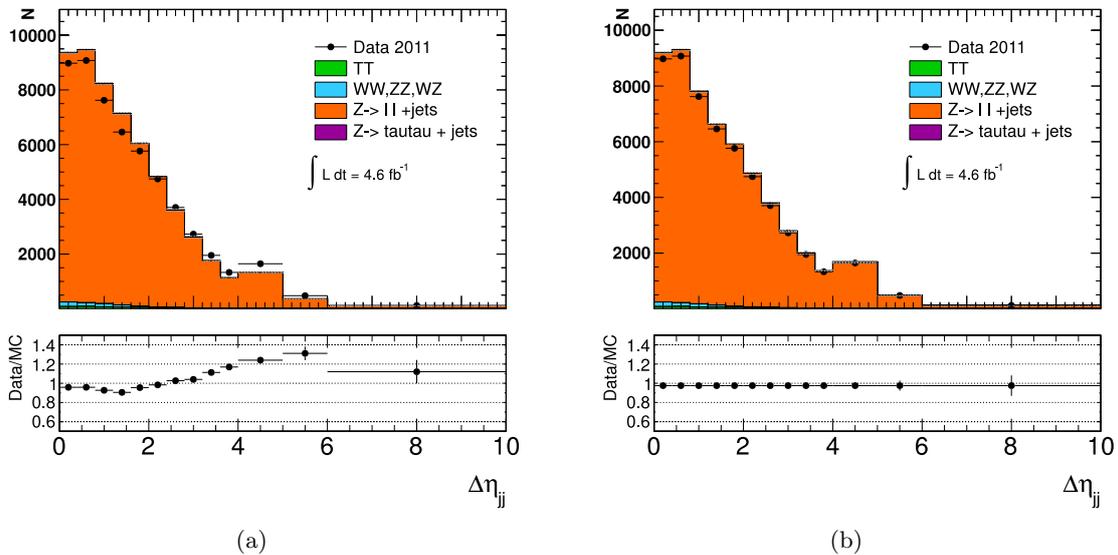


Figure 5.11.: Comparison of  $|\Delta\eta_{jj}|$  between the two jets with highest transverse momentum in  $Z \rightarrow \mu\mu$  events from data corresponding to an integrated luminosity of  $4.6 \text{ fb}^{-1}$  and ALPGEN MC before (a) and after (b) re-weighting. The lower plots display the data-to-MC ratio of the corresponding distributions, including their statistical uncertainties.

A conservative systematic uncertainty of  $\pm 5\%$  is assigned to this correction to compensate for the residual mis-modeling observed after the full VBF selection, shown in Fig. 5.12 (see Section 6.4 for more details).

<sup>6</sup>The  $Z \rightarrow \mu\mu$  events were obtained by requiring two isolated leptons with  $|m_{\ell\ell} - 91.2 \text{ GeV}| < 15 \text{ GeV}$ .

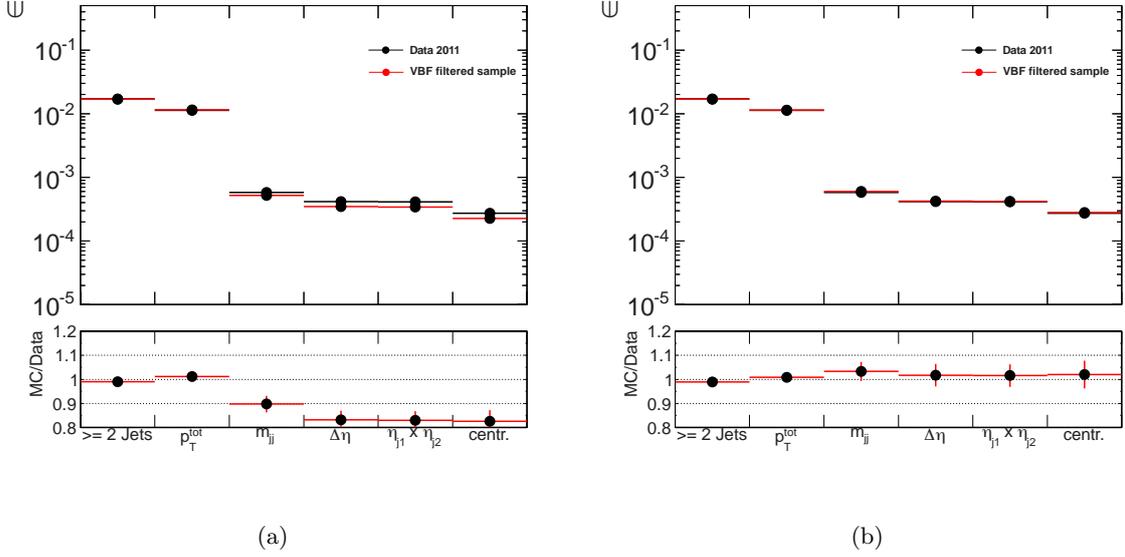


Figure 5.12.: Distributions of  $|\Delta\eta_{jj}|$  for the VBF-filtered sample before (a) and after (b) re-weighting. The lower plots display the MC-to-data ratio of the corresponding distributions, including their statistical uncertainties (plot taken from Ref. [87]).

### 5.3.8. Fake factor method

The fake estimation based on SS data can suffer from a low amount of events, especially in the VBF topology. Therefore, the *fake-factor* (FF) method outlined in the following is implemented and allows for a data-driven fake- $\tau$  estimation in phase space regions with few SS data events. It exploits the fact that a  $\tau$  ID is required in the analysis and obtains an orthogonal control sample by inverting the  $\tau$  ID cut.

A scale factor, referred to as fake factor (FF), is applied to re-weight the  $\tau$  fake events ( $N_{\text{anti-}\tau}^{\text{data}}$ ) to the amount anticipated in the signal region ( $N_{\text{Bkg.}}^{\text{Est.}}$ ). The latter is given by

$$N_{\text{Bkg.}}^{\text{Est.}} = N_{\text{anti-}\tau}^{\text{data}} \times FF. \quad (5.9)$$

The fake factor is computed as ratio of total events fulfilling the  $\tau$ -ID to those failing it (anti- $\tau$ ),

$$FF = \frac{N_{\tau\text{-ID}}}{N_{\text{anti-}\tau}}. \quad (5.10)$$

The  $\tau$ -fake background is dominated by W+jets and QCD events. The  $\tau_{\text{had}}$  candidate is thereby mainly faked by quark-initiated jets in the former and gluon-initiated jets in latter case. Hence, corresponding fake factors are computed in two distinct control regions and combined afterwards.

Similar to the background estimation via SS data, a control region of high transverse mass

is taken to attain a nearly uncontaminated W+jets sample. Non-isolated leptons are taken to obtain a QCD control region. Both fake factors are finally combined via

$$FF_{\text{MIX}} = R_{\text{W+jets}} \cdot FF_{\text{W+jets}} + (1 - R_{\text{W+jets}}) \cdot FF_{\text{QCD}}, \quad (5.11)$$

where  $R_{\text{W+jets}}$  is the fraction of W+jets events in all anti- $\tau$  events in the signal region (SR).

The fake factors are derived for SLT and LTT as well as for 1-prong and 3-prong  $\tau$  candidates, separately. To estimate an uncertainty on the fake-factor method, the FF was studied in a quark- and gluon-rich sample represented by W+jet events without additional jets<sup>7</sup> and QCD events, respectively.

Figure 5.13 depicts the nominal FF together with the FF for quark- and gluon-enriched samples of 1 prong (left) and 3 prong (right)  $\tau$  candidates in the single lepton trigger.

Finally, a conservative uncertainty of  $\pm 50\%$  is assigned to  $FF_{\text{MIX}}$  for the 1 prong and 3 prong  $\tau$  candidates and the single and lepton+tau trigger, respectively. It accounts for the imperfect knowledge of the quark and gluon jet flavor composition in simulated events.

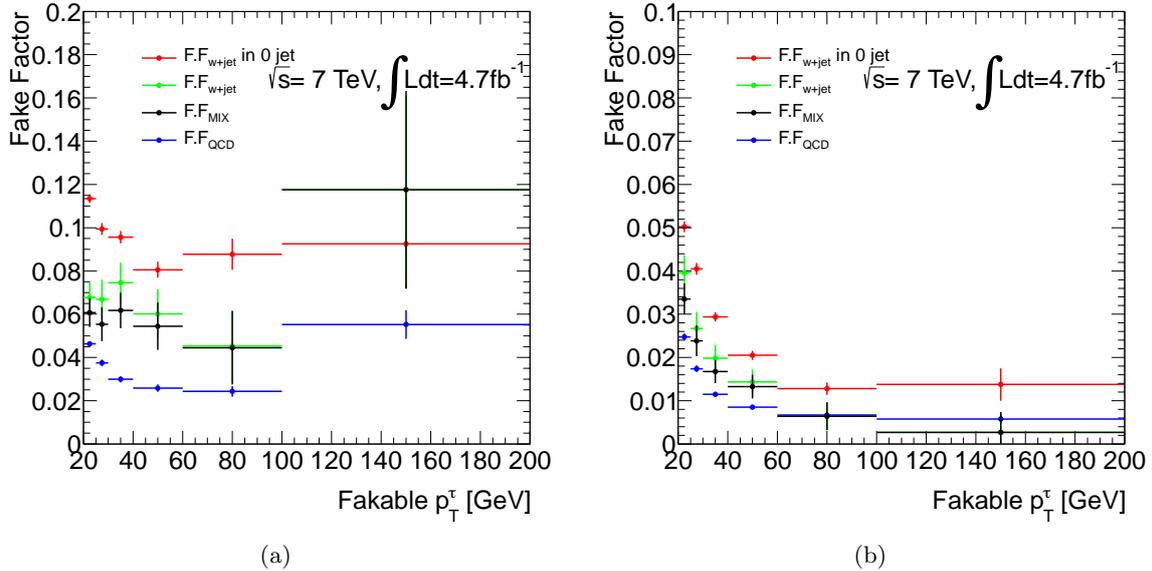


Figure 5.13.: The fake-factor dependence on  $p_T^\tau$  for 1 prong (a) and 3 prong (b)  $\tau$  candidates in the SLT. The uncertainty on  $FF_{\text{MIX}}$  is derived via a quark- and a gluon-enriched sample represented by  $FF_{\text{W+jets}}$  and  $FF_{\text{QCD}}$ , respectively (plot taken from Ref. [87]).

<sup>7</sup>The W background with exactly one jet is dominated by a quark that radiated a W boson.

---

# 6

## Re-analysis of the 2011 dataset

---

Subsequent to their preselection, the events are categorized to benefit from different Higgs boson production modes and enhance the signal-to-background ratio. For one, the Higgs boson production at the LHC is dominated by gluon fusion, whereby the VBF production mode can be accentuated by its characteristic topology. Besides, the transverse boost of the  $\tau\tau$  system leads to a better  $E_T^{\text{miss}}$  resolution and can be exploited to enhance the separation between signal-like and background-like events.

The  $H \rightarrow \tau\tau$  subgroup of the ATLAS collaboration has the objective to establish the SM Higgs boson existence in a leptonic decay mode. Apart from its sole existence, the analysis has the intention to distinguish between different Higgs boson production modes to probe the predicted ratios of the coupling strength parameters given by the SM.

Triggered by optimization prospects, a re-analysis [15] of the  $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$  channel following an already published analysis on the 2011 dataset [86] was carried out. Thus, various quantities were established to cope with background events arising from mis-identified  $\tau_{\text{had}}$  candidates, which are presented within this chapter. Furthermore, the *Boosted* category which was implemented to benefit from the better mass resolution of a collimated  $\tau\tau$  system is described. In addition, an optimization of the VBF classification cuts was performed as part of this thesis.

### 6.1. Optimization of the VBF category

To take advantage of the VBF production topology and separate the  $Z \rightarrow \tau\tau$  background from signal events, quantities based on the two *tagging jets* (Section 1.4) are exploited. The initial analysis of the 2011 dataset [86] was taken as premise for the optimization. Thereby, the background estimation was accomplished by the OS-SS method (Section 5.3.1) together with the *embedding* technique (Section 5.3.7) to model the  $Z \rightarrow \tau\tau$  background.

The study was performed by continuous variation of the cut thresholds on jet-related variables used to define the VBF category, namely the transverse momenta of the *tagging jets*, their spatial separation and invariant mass. Prior to the re-analysis, events accompanied

by jets were allocated to the VBF category if they fulfilled specific VBF classification cuts that are outlined below. In addition, a 1J category was established for remaining events which were accompanied by at least one jet. To optimize the VBF categorization, both categories had to be taken into account since altered jet cuts in the VBF category change the 1J event composition and vice versa.

Initially, all events in the  $e + \tau_{\text{had}}$  and  $\mu + \tau_{\text{had}}$  channels with a missing transverse energy of  $E_{\text{T}}^{\text{miss}} > 20$  GeV and a transverse mass (Section 6.3) of  $m_{\text{T}} < 30$  GeV were assigned to the signal region. Then, events with additional jets were classified in the following way:

1. For the VBF category, at least two jets with  $p_{\text{T}} > 25$  GeV were demanded. In addition, the two jets with highest transverse momenta had to reside in opposite hemispheres of the detector ( $\eta_{j1} \times \eta_{j2} < 0$ ) and had to be well separated in pseudo-rapidity ( $\Delta\eta_{j1j2} > 3.0$ ). Finally, their invariant mass was required to fulfill  $m_{j1j2} > 300$  GeV and the lepton and  $\tau_{\text{had}}$  candidates had to reside between these two jets in pseudo-rapidity.
2. Events that failed the categorization above were subsumed in the 1J category.

Figure 6.1 shows the  $m_{\tau\tau}^{\text{MMC}}$  distributions of both categories, which were used in the first analysis of the 2011 dataset corresponding to an integrated luminosity of  $\int \mathcal{L} dt = 4.6 \text{ fb}^{-1}$ . They display a reasonable agreement between recorded data and simulated events.

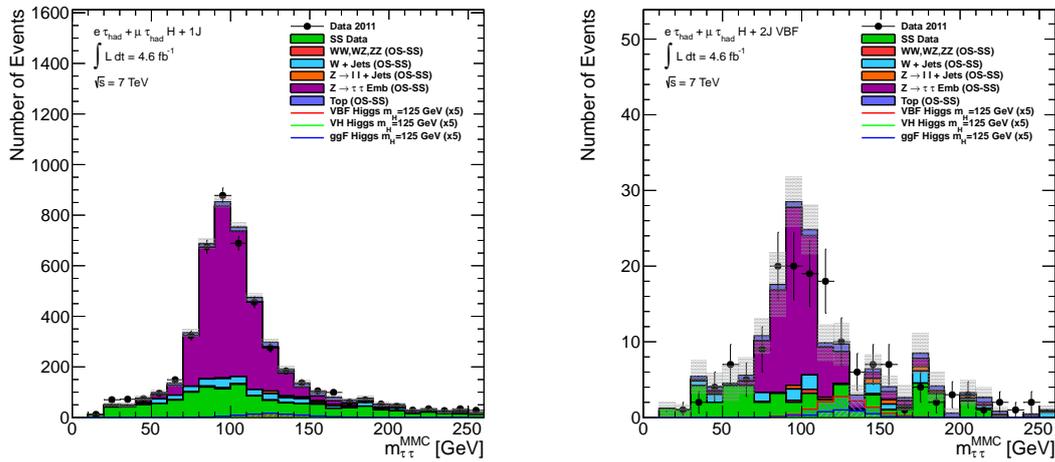


Figure 6.1.: Distributions of the reconstructed mass of the di- $\tau$  system for the 1J (left) and VBF (right) categories used in the first analysis of the 2011 dataset.

Originating from these, a four-dimensional scan in the  $(p_{\text{T}}^{j1}, p_{\text{T}}^{j2}, m_{j1j2}, \Delta\eta_{j1j2})$  parameter space was performed with the purpose to attain an optimal VBF classification given the underlying background estimation techniques. The considered phase space was constrained

by  $25 \text{ GeV} < (p_T^{j_1}, p_T^{j_2}) < 50 \text{ GeV}$ ,  $250 \text{ GeV} < m_{j_1 j_2} < 550 \text{ GeV}$  and  $2.5 < \Delta\eta_{j_1 j_2} < 5.5$ . Thereby, the parameter space was partitioned in a grid consisting of 5 GeV steps in  $p_T^{j_1}$  and  $p_T^{j_2}$ , 0.5 steps in  $\Delta\eta_{j_1 j_2}$  and 50 GeV steps in  $m_{j_1 j_2}$ .

For each point of this grid, the expected 95 % CL exclusion limit on the Higgs boson production cross section, normalized to the SM prediction  $\sigma_{\text{SM}}$  (in the following called exclusion limit<sup>1</sup>) and with  $m_{\tau\tau}^{\text{MMC}}$  as discriminating variable was calculated. The statistical uncertainties on the data and MC events as well as the systematic uncertainties on the jet and  $\tau$  energy scale were incorporated in the limit computation. Detailed descriptions of the limit computation and the various systematic uncertainties are given in Chapter 4 and Section 8.1.

The nominal categorization yielded an expected upper limit on the signal production cross section of  $6.29 \cdot \sigma_{\text{SM}}$ , which was defined as reference to quantify possible improvements. The four-dimensional scan resulted in an optimal set of cuts given by

$$p_T^{j_1} = 50 \text{ GeV}, \quad p_T^{j_2} = 25 \text{ GeV}, \quad m_{j_1 j_2} = 450 \text{ GeV}, \quad \Delta\eta_{j_1 j_2} = 3.0, \quad (6.1)$$

which concluded in an expected limit of  $5.45 \cdot \sigma_{\text{SM}}$  and corresponds to an improvement of 13.35 %. This result indicated a more rigorous set of cuts to take advantage of the characteristic VBF topology. To illustrate the deviation in each direction, Figs. 6.2a and 6.2b show the expected exclusion limits as functions of  $\Delta\eta_{j_1 j_2}$  and  $p_T^{j_2}$ . The remaining parameters are fixed to the values of the above defined optimum.

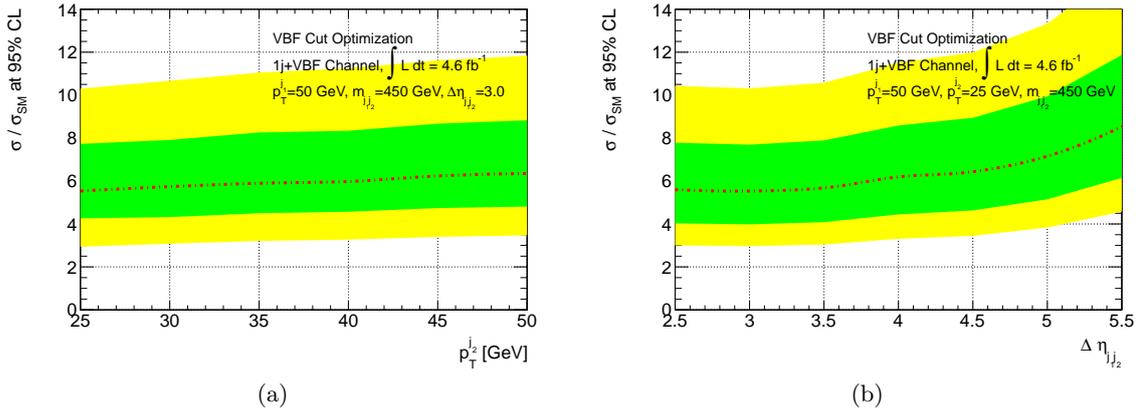


Figure 6.2.: Dependence of the expected exclusion limit as a function of (a)  $p_T^{j_2}$  and (b)  $\Delta\eta_{j_1 j_2}$ . The remaining variables are set to their optimal value given by Eq. 6.1.

<sup>1</sup>More precise, the expected limit is set on the production cross section multiplied with the  $H \rightarrow \tau_{\text{lep}} \tau_{\text{had}}$  BR, the acceptance and the efficiency of the ATLAS detector.

Figure 6.3 displays the expected exclusion limits in the  $(p_T^{j_1}, m_{j_1 j_2})$  parameter plane for fixed  $\Delta\eta_{j_1 j_2} = 3.0$  and  $p_T^{j_2} = 25$  GeV.

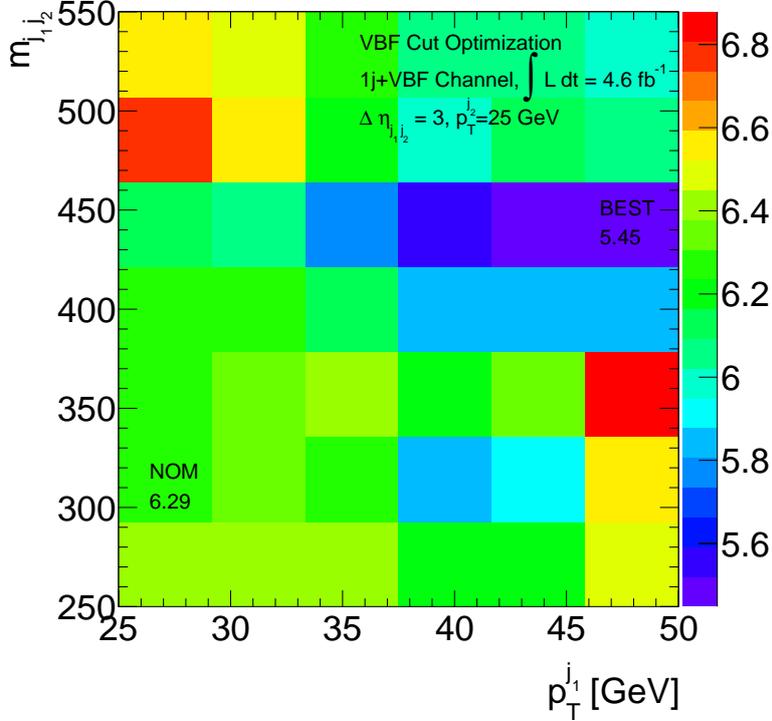


Figure 6.3.: The expected exclusion limit  $\sigma/\sigma_{\text{SM}}$  as a function of  $p_T^{j_1}$  and  $m_{j_1 j_2}$  for fixed  $\Delta\eta_{j_1 j_2} = 3.0$  and  $p_T^{j_2} = 25$  GeV. NOM identifies the initial VBF categorization cuts of  $p_T^{j_1} = 25$  GeV and  $m_{j_1 j_2} = 300$  GeV with a corresponding limit of  $6.29 \cdot \sigma_{\text{SM}}$ . The optimal point found in the 4D scan is at the same  $\Delta\eta_{j_1 j_2}$  and  $p_T^{j_2}$  value together with  $p_T^{j_1} = 50$  GeV and  $m_{j_1 j_2} = 450$  GeV. It is labeled BEST in the 2D plot and corresponds to a limit of  $5.45 \cdot \sigma_{\text{SM}}$ .

The latter illustrates significant differences of close-by parameters, for example between  $p_T^{j_1} = 25$  GeV with a  $m_{j_1 j_2}$  of 450 GeV and 500 GeV or between  $p_T^{j_1} = 50$  GeV and a  $m_{j_1 j_2}$  of 350 GeV and 400 GeV. This revealed a possible dependence on statistical fluctuations of the underlying discriminant, which drive the profiling of nuisance parameters during the limit setting procedure. In addition, it demonstrated problems in the background estimation if the OS-SS method with an embedded  $Z \rightarrow \tau\tau$  sample was used, especially at tight cut parameters that substantially restricted the phase space. Therefore, different methods were developed to describe the  $Z \rightarrow \tau\tau$  and fake- $\tau$  backgrounds. This led to the VBF-filtered  $Z \rightarrow \tau\tau$  samples together with the fake factor method (see Sections 5.3.7 and 5.3.8).

To illustrate the lower number of events and the higher signal sensitivity in the modified VBF category, the corresponding distributions for the reconstructed mass of the  $\tau\tau$  system are shown before and after optimization in Fig. 6.4.

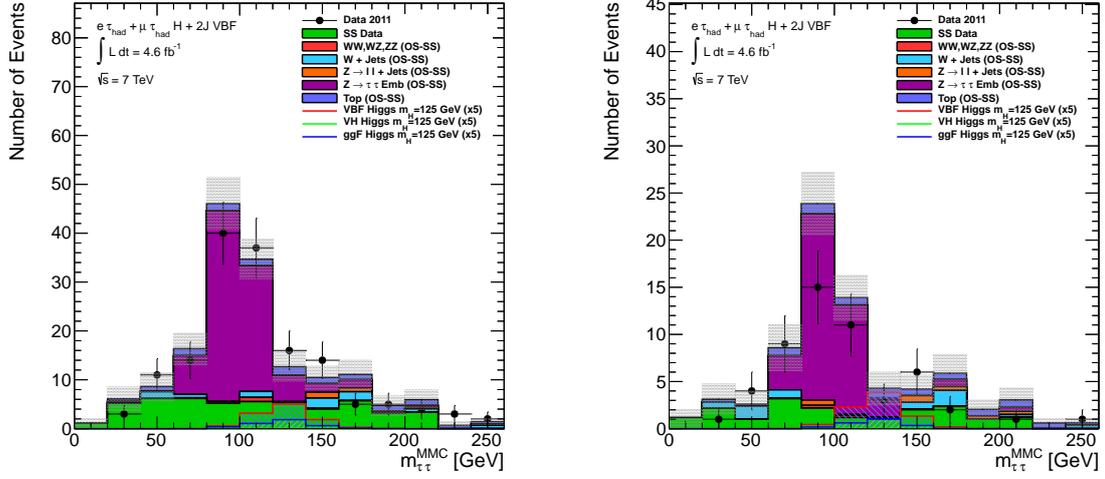


Figure 6.4.: The  $m_{\tau\tau}^{MMC}$  distributions for the old (left) and the optimized (right) VBF classification.

## 6.2. The boosted category

A possible boost of the Higgs boson is acquired when it recoils against jets or other vector bosons, for example in the associated production. It can be exploited by the transverse Higgs boson momentum  $p_T^H$ , which is reconstructed from its visible and invisible decay products via

$$p_T^H = |\mathbf{p}_T^\ell + \mathbf{p}_T^{\tau_{had}} + \mathbf{E}_T^{miss}|. \quad (6.2)$$

Dividing the events of the former 1J category into those above a certain  $p_T^H$  threshold (in the following called *Boosted* category) and failing it resulted in a higher sensitivity compared to the previous 1J category alone. This is caused by an improved  $E_T^{miss}$  resolution and more collinear  $\tau$  decays that render the di- $\tau$  mass reconstruction more performant.

Moreover, the QCD and fake- $\tau$  background processes typically lead to softer objects with a different origin as the lepton and are therefore separated by the transverse momentum cut. The improved resolution of  $m_{\tau\tau}^{MMC}$  as well as the better separation between signal and background is illustrated in Fig. 6.5.

## 6.3. Suppression of fake tau backgrounds

To diminish the contamination from fake- $\tau$  background events, several quantities based on the momenta and angular distributions of the Higgs boson decay products were implemented in the re-analysis of the 2011 dataset and are outlined in the following.

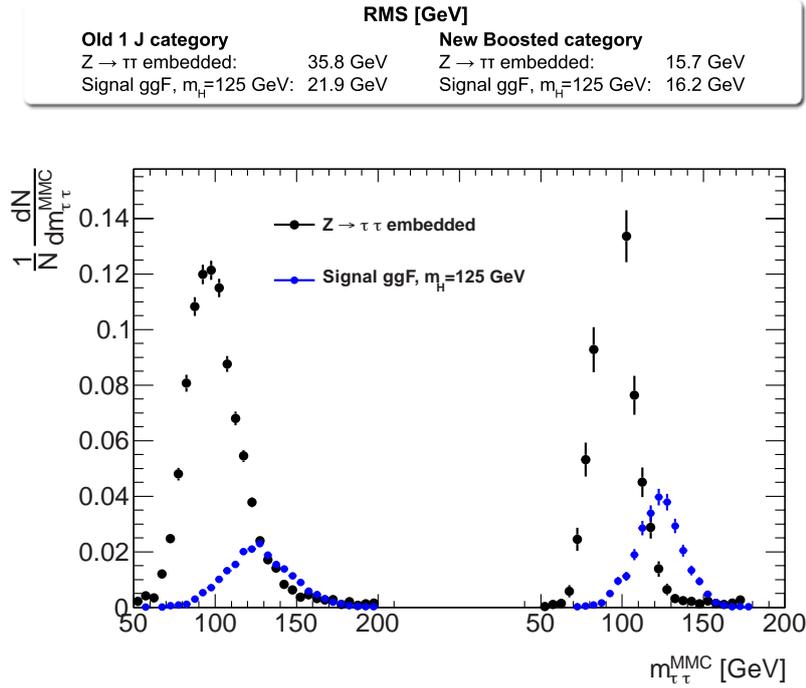


Figure 6.5.: Distributions of  $m_{\tau\tau}^{\text{MMC}}$  for the old 1J category as well as the newly defined *Boosted* category. The embedded sample (black dots) is normalized to unity and compared to the ggF signal sample at  $m_H = 125$  GeV (blue dots). The signal is normalized to the signal-to-background ratio and multiplied by a factor of 50 for better visualization. In addition, the root mean squared (RMS) is specified.

- $m_T$ : The transverse mass ( $m_T$ ) is used to discriminate between signal events that typically possess a low  $m_T$  and background processes with a leptonic  $W$  boson decay ( $W \rightarrow e/\mu + \nu$ ) leading to a high  $m_T \leq m_W$ . It is defined as

$$m_T = \sqrt{2p_T^{\text{lep}} E_T^{\text{miss}} (1 - \cos\Delta\phi)}, \quad (6.3)$$

where  $\Delta\phi$  is the angular difference in the transverse plane (see Section 2.2) between the lepton with transverse momentum  $p_T^{\text{lep}}$  and the missing transverse energy  $E_T^{\text{miss}}$ . In this analysis,  $m_T$  discriminates between the signal region and different control regions (see Sections 5.3.3 to 5.3.5).

- $\sum \Delta\phi$ : To exploit the angular distribution of the final state, a quantity based on the  $\Delta\phi$  between  $E_T^{\text{miss}}$  and both lepton and hadronic  $\tau$  is used. It is given by

$$\sum \Delta\phi = |\phi_{\text{lep}} - \phi_{E_T^{\text{miss}}}| + |\phi_{\tau} - \phi_{E_T^{\text{miss}}}| \quad (6.4)$$

and takes advantage of the fact, that the  $E_T^{\text{miss}}$  of a Higgs boson decay is usually residing

in between the visible  $\tau$  decay products in the plane perpendicular to the beam axis, which yields  $\sum \Delta\phi < \pi$ . In contrast, fake- $\tau$  backgrounds can have a significantly larger  $\sum \Delta\phi$ . Hence, various cuts are applied to enhance the signal significance in each category. Thereby, a large signal efficiency of  $\sim 99\%$  is retained to prevent a bias that may result in additional systematic uncertainties.

- **$\Delta\Delta R$** : If the Higgs boson recoils against additional jets or vector bosons, it acquires a boost in the transverse plane. Thus, the Higgs boson decay products are assumed to be more collinear, measured by  $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ . The correlation between  $\Delta R^{\ell,\tau}$  and the transverse momentum of the lepton- $\tau$  system ( $p_T^{\ell,\tau}$ ) is shown in Fig. 6.6a. To suppress non-resonant processes, the agreement between the  $\Delta R^{\ell,\tau}$  of an event with the expectation for a signal event ( $\Delta R_{\text{pred}}^{\ell,\tau}$ ) is tested. Therefore, the  $\Delta R^{\ell,\tau}(p_T^{\ell,\tau})$  dependence is parametrized by a Landau function for the simulated signal events in each category. An exemplary fit of the VBF category is illustrated in Fig. 6.6b.

The final variable is obtained through the difference between  $\Delta R_{\text{pred}}^{\ell,\tau}$  and the measured  $\Delta R^{\ell,\tau}$  of an event,

$$\Delta\Delta R^{\ell,\tau} = |\Delta R^{\ell,\tau} - \Delta R_{\text{pred}}^{\ell,\tau}|. \quad (6.5)$$

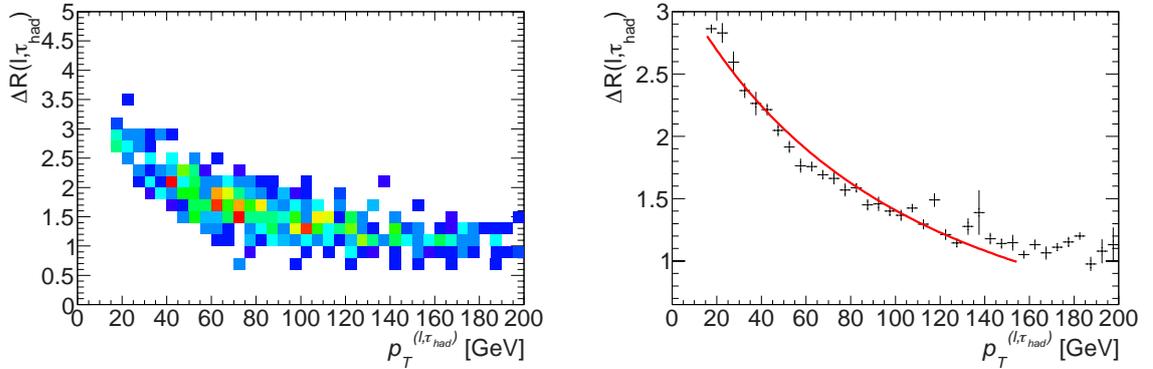


Figure 6.6.: Correlation between  $\Delta R^{\ell,\tau}$  and the transverse momentum of the lepton- $\tau$  system,  $p_T^{\ell,\tau}$ , illustrated for the VBF production process. On the left plot, different colors indicate the amount of events in a distinct 2D-tile (dark blue stands for a small, red for a high event count). The right plot displays the mean of each  $p_T^{\ell,\tau}$  bin with its corresponding statistical error and the fit of a Landau function to determine  $\Delta R_{\text{pred}}^{\ell,\tau}(p_T^{\ell,\tau}, \text{category})$  (plot taken from Ref. [87]).

A cut on this quantity which varies with the event category is applied. It results in a signal efficiency of roughly 99 %, while rejecting up to 50 % of the non-resonant background.

- **$\mathbf{x}_\ell$  and  $\mathbf{x}_h$ :**  $x_\ell$  and  $x_h$  express the momentum fraction of the two visible  $\tau$  decay products with respect to their parent  $\tau$  lepton. They are calculated under the assumption that the respective neutrino(s) of each  $\tau$  decay are collinear to their visible counterpart (collinear mass approximation).

Since the visible energy fraction is between zero and one for real  $\tau$ s, values beyond these bounds indicate that the collinear assumption is wrong. This helps to suppress background events while keeping a high fraction of signal events.

- **$\mathbf{p}_T^{\text{total}}$ :** Following the VBF signal topology description given in Section 5.1.1, little jet activity is expected besides the two *tagging jets* from the scattered partons. This gives rise to a new quantity imposed in the VBF category, the transverse boost of the entire VBF system, which is defined as

$$p_T^{\text{total}} = |\mathbf{p}_T^\ell + \mathbf{p}_T^{\tau_{\text{had}}} + \mathbf{p}_T^{j1} + \mathbf{p}_T^{j2} + \mathbf{E}_T^{\text{miss}}|. \quad (6.6)$$

- **$\mathbf{p}_T^{\text{asym}}$ :** Since the leptonic  $\tau$  decay is accompanied by two neutrinos in contrary to one neutrino in the  $\tau_{\text{had}}$  decay, the quantity

$$p_T^{\text{asym}} = p_T^\ell - p_T^\tau \quad (6.7)$$

is assumed to peak at negative values for the signal and  $Z \rightarrow \tau\tau$  processes.

In contrary, a broad  $p_T^{\text{asym}}$  distribution with positive mean is detected for fake- $\tau$  backgrounds. The identified lepton is required to pass specific isolation criteria and is mainly produced in electroweak processes. Thus, it usually acquires a higher momentum with respect to the jet that is mis-identified as  $\tau_{\text{had}}$  candidate.

- **$Z \rightarrow \ell\ell$  rejection cuts:** The  $Z \rightarrow \ell\ell$ +jets background can contribute if either a lepton or a jet is mis-identified as  $\tau_{\text{had}}$  (see Section 5.3.4). In the former occasion, both final state objects have the same origin. This leads to a reconstructed mass close to the mass of a typical signal event together with a significant amount of events in categories with zero jets. Hence, an additional lepton fake rejection, on top of the electron and muon veto already applied in the  $\tau_{\text{had}}$  identification algorithm, is introduced.

For the  $Z \rightarrow ee$  ( $e \rightarrow \tau_{\text{had}}$ ) background, a narrow window in the central detector region ( $|\eta| < 0.05$ ) is excluded because limited TRT and calorimeter information is available (see Section 2.2.1). This cut is carried out on one-prong  $\tau_{\text{had}}$  candidates in all categories except of the Boosted and VBF category, which do not suffer from a significant

$Z \rightarrow ee(e \rightarrow \tau_{\text{had}})$  contamination.

On the other hand,  $Z \rightarrow \mu\mu(\mu \rightarrow \tau_{\text{had}})$  events appear when a large amount of bremsstrahlung is induced by the muon. This background can be diminished by the requirement of a minimum fraction of transverse energy from the  $\tau_{\text{had}}$  candidate within the electromagnetic calorimeter ( $f_{\text{EM}} > 0.1$ ). This cut is solely applied on  $\tau_{\text{had}}$  candidates with one reconstructed track. Moreover, the invariant mass of the muon and  $\tau_{\text{had}}$ -track is required to be within 10 GeV of the  $Z$  mass.

## 6.4. Analysis categories

Subsuming the recent sections, the event classification was revisited to improve the signal sensitivity achieved in Ref. [86]. Therefore, more stringent classification cuts were determined within the existing VBF category and an increased mass reconstruction ability was exploited for boosted Higgs boson candidates. In addition, characteristic kinematic and angular distributions were used to separate the signal from fake- $\tau$  events in each category.

Thus, a set of optimal fake- $\tau$  suppression cuts as well as the best-performant  $p_{\text{T}}^{\text{H}}$  threshold were determined via different figures of merit<sup>2</sup>. Moreover, a VBF optimization study similar to the approach presented in Section 6.1 was performed to attain an optimal set of cuts. Thereby, VBF-filtered  $Z$  samples and the fake factor method to estimate the fake- $\tau$  background contribution were used.

The final VBF classification cuts used in the re-analysis differ from the optimal cuts determined in Section 6.1. The VBF category of the re-analysis requires a higher threshold of the sub-leading jet and a higher invariant di-jet mass. These modifications can be motivated by the increased number of events available for fake backgrounds and in VBF-filtered samples.

Table 6.1 outlines the final categorization scheme of the re-analysis [15], starting by the preselection described in Section 5.2.1. The preselection is followed by the event categorization on the basis of kinematic properties of jets in the event or the di- $\tau$  system. To obtain the signal region, various cuts are applied on the variables defined in Section 6.3.

To illustrate the sensitivity of each category, a signal significance (Eq. A.6) is stated in Table 6.1. It has the highest value within the VBF topology. Moreover, a clear advantage can be discovered for the Boosted category with respect to the 1J category. The 0J category has the worst signal significance since it suffers from a bad signal-to-background separation and a high amount of irreducible  $Z \rightarrow \tau\tau$  events.

<sup>2</sup>These are derived in Appendix A and based on the limit procedure of Chapter 4.

Table 6.1.: Event categorization applied in the cut-based re-analysis of the 2011 dataset. An event and object preselection according to Section 5.2.1 is performed at the beginning. Afterwards, the events are classified to enhance the signal significance and emphasize different Higgs boson production modes. Subsequently, additional signal region cuts are applied in order to suppress fake- $\tau$  background processes. The various background rejection cuts are introduced in Sections 6.1 to 6.3. The signal significance is derived on the basis of Eq. A.6.

Preselection as described in Section 5.2.1 plus $E_{\tau}^{\text{miss}} > 20$ GeV			
Combination of $e + \tau_{\text{had}}$ and $\mu + \tau_{\text{had}}$ events			
VBF	Boosted	1+ Jet	0J
- At least two jets with $p_{\text{T}} > 40$ GeV	- not VBF	- not VBF/Boosted	- exactly zero jets
- $\Delta\eta_{j_1 j_2} > 3.0$ , $m_{j_1 j_2} > 500$ GeV	- $p_{\text{T}}^{\text{H}} > 100$ GeV	- Jets with $p_{\text{T}} > 25$ GeV	
- $\eta_{j_1} \times \eta_{j_2} < 0$ and $p_{\text{T}}^{\text{total}} < 40$ GeV	- $0 < x_{\ell} < 1.0$ and $0.2 < x_h < 1.2$		
- $\min(\eta_{j_1}, \eta_{j_2}) < \eta_{\text{T}}^{\text{had}} < \max(\eta_{j_1}, \eta_{j_2})$			
Signal region cuts applied after event classification			
- $m_{\text{T}} < 50$ GeV	- $m_{\text{T}} < 50$ GeV	- $m_{\text{T}} < 50$ GeV	- $m_{\text{T}} < 30$ GeV
- $\Delta\Delta R < 1.0$	- $\Delta\Delta R < 0.6$	- $\Delta\Delta R < 0.6$	- $\Delta\Delta R < 0.5$
- $\sum \Delta\phi < 2.9$	- $\sum \Delta\phi < 1.6$	- $\sum \Delta\phi < 3.5$	- $\sum \Delta\phi < 3.5$
		- $Z \rightarrow \ell\ell$ rejection cuts	- $Z \rightarrow \ell\ell$ rejection cuts
			- $p_{\text{T}}^{\ell} - p_{\text{T}}^{\tau_{\text{had}}} < 0$
Signal significance			
0.668	0.423	0.317	0.267

### 6.4.1. Kinematic distributions

Figure 6.7 depicts the final  $m_{\tau\tau}^{\text{MMC}}$  distributions for the categorization outlined in Table 6.1. These are used to compute an exclusion limit on the Higgs boson production cross section. Thereby, the compatibility of measured data to the background processes and a possible Higgs boson predicted by the SM is probed (see Section 8.2).

The Higgs boson processes are multiplied by a factor of 50 except for the sensitive VBF category (Fig. 6.7f), where the signal is multiplied by a factor of 5. All distributions display a reasonable agreement between data and Monte Carlo within their statistical uncertainties. As described above, the more rigid VBF classification as well as the new Boosted category enhance the signal sensitivity with respect to the analysis performed previously on the 2011 dataset.

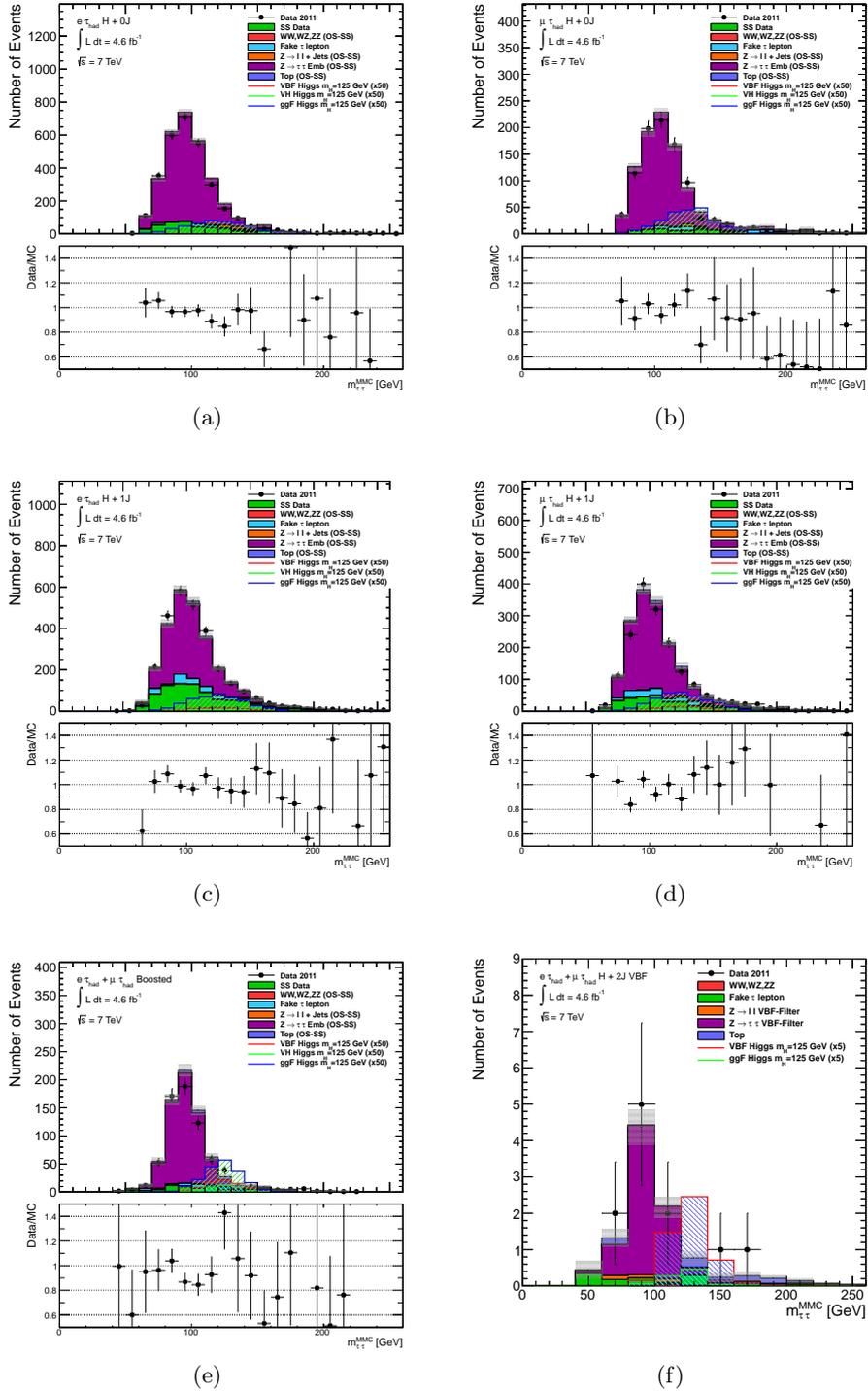


Figure 6.7.: Distributions of the reconstructed mass for the di- $\tau$  system for events without (a,b) or with at least one reconstructed jet (c,d) that fail the VBF and Boosted classification. They are split into the  $e + \tau_{\text{had}}$  (left column) and  $\mu + \tau_{\text{had}}$  (right column) channel. In addition, the Boosted (e) and VBF (f) categories are depicted, where both  $e + \tau_{\text{had}}$  and  $\mu + \tau_{\text{had}}$  events are combined to compensate for the low event count. The background is modeled by embedded  $Z \rightarrow \tau\tau$  events and the OS-SS method (see Section 5.3) in all categories except VBF. In this case, the fake factor method in combination with VBF-filtered  $Z$ +jets samples (see Sections 5.3.7 and 5.3.8) are used. Each distribution is accompanied by its corresponding data-to-MC ratio, including their statistical uncertainties.

---

# 7

## A multivariate approach for the

### $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$ search

---

The cut-based analysis presented in Chapter 6 utilizes one- or two-sided cuts on distinct variables in order to enhance the signal-to-background ratio. Thereby, signal events are usually discarded as well because of the imperfect discrimination of the variables.

A multivariate analysis (MVA) is able to outperform a cut-based approach due to the use of multiple variables with their entire shape and correlation information [123]. The MVA employs machine learning algorithms to combine the given input variables into a single powerful discriminant.

The analysis described in this chapter is based on Boosted Decision Trees (BDTs) [16, 124], implemented in the *Toolkit for Multivariate Data Analysis* (TMVA) [125]. BDTs are commonly applied in physics analyses<sup>1</sup> and possess well understood advantages and drawbacks. Moreover, they provide an excellent performance when compared to other multivariate techniques like artificial neural-networks (ANN) [126], linear methods like the Fisher classifier [127] or a one-dimensional likelihood<sup>2</sup> classification [125]. All these MVA techniques were tested on the Boosted category, defined in Table 7.2, with the default configurations provided in TMVA.

Their performance is illustrated by the ROC<sup>3</sup>-curves in Fig. 7.1. At a signal efficiency of 70 %, the BDT achieves a background rejection of approximately 90.8 %. In contrast, a background rejection of  $\sim 64.0$  %,  $\sim 87.9$  % and  $\sim 89.7$  % is realized in case of the Fisher, one-dimensional Likelihood and ANN classifier, respectively.

This chapter provides a brief overview of the BDT technique in Section 7.1, followed by an optimization study to improve the sensitivity of the analysis based on the BDT configuration, input variables and sample selection. Finally, the agreement between recorded data and simulated events of the BDT classifier is examined in the SR as well as in several CRs.

---

<sup>1</sup>For example in the hadronic  $\tau$  ID requirements described in Section 3.7.

<sup>2</sup>The one-dimensional likelihood estimator ignores correlations between the variables by simple multiplication of the signal PDFs.

<sup>3</sup>The receiver operating characteristic (ROC) curve depicts the background rejection for various signal efficiency thresholds.

The studies performed within this thesis pioneered the BDT approach which is currently established for the 2012 dataset within the  $H$  to  $\tau\tau$  subgroup at the ATLAS collaboration. The achieved results regarding the cross validation (Section 7.3.2), pre-categorization and background estimation were taken as foundation for more elaborate studies on the data corresponding to an integrated luminosity of  $20.3 \text{ fb}^{-1}$  at center-of-mass energy of  $\sqrt{s} = 8 \text{ TeV}$ .

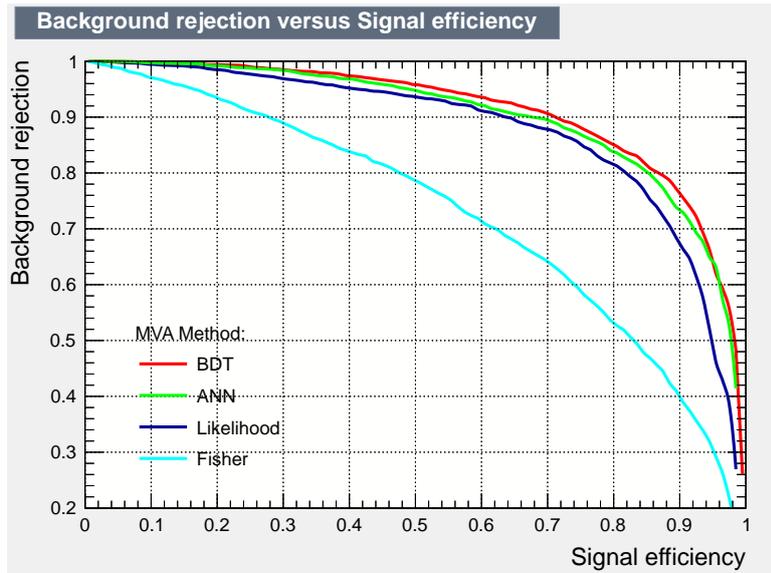


Figure 7.1.: ROC-curves for various multivariate and linear classifiers, trained on the Boosted category specified in Table 7.2. Thereby, Boosted Decision Trees (BDTs) are compared to an artificial neural-network (ANN), a one-dimensional Likelihood estimator and the Fisher classifier.

## 7.1. Boosted Decision Trees

The BDT algorithm maps the multi-dimensional phase space spanned by the input variables to assign a *score* on an event depending on the signal-likeness of the region it falls into. The classifier is thereby constructed on a known composition of signal and background events that are typically used in the analysis.

As illustrated in Fig. 7.2, a single tree of the BDT algorithm consists of multiple nodes on different layers. The training of a single tree starts with a root node that is split in two subsets. This procedure is successively repeated on each node of the tree. To determine the splitting criterion, all variables are scanned with a predefined number of cuts ( $n\text{Cuts}$ ) on each step. Then, the particular variable and cut value is chosen based on a specified metric.

Within this analysis, the *Gini index* is used. It is defined as

$$G(x, c) = p(x, c) \cdot (1 - p(x, c)) \quad , \quad \text{with} \quad p(x, c) = \frac{N_{\text{signal}}(x, c)}{N_{\text{bkg.}}(x, c)} \quad (7.1)$$

at tested variable  $x$  and cut value  $c$ . The highest *Gini index* is chosen and the splitting continues until a stopping condition is reached. In this tree, the maximum amount of consecutive cuts defines its depth and every event is eventually assigned to a specific node, denoted leaf. According to the majority of events, a final leaf is denoted *signal-like* (S) or *background-like* (B) (Fig. 7.2). In the configuration of this analysis, an event that ends up in a S(B) leaf triggers a numerical response of  $+(-) 1$  from that tree.

Stopping criteria are for instance a fixed minimum number of events ( $n\text{Events}_{\text{min}}$ ) in each final leaf or a maximal depth of the concluding tree ( $\text{Depth}_{\text{max}}$ ).

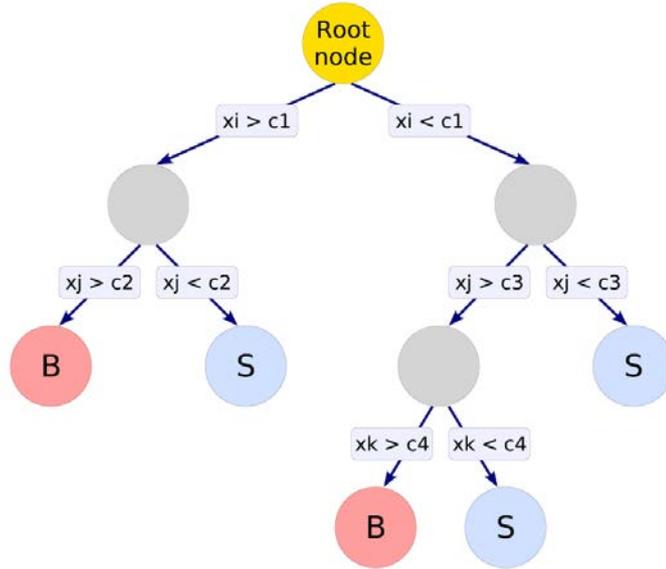


Figure 7.2.: A single tree with four cuts, concluding in five leaves assigned S for signal-like or B for background-like (plot taken from Ref. [125]).

A shortcoming of the tree structure is its instability with regard to statistical fluctuations of the training sample. A BDT constructed with a specific training sample can in principle distinguish perfectly between the signal and background events in this sample. Thus, an independent sample is employed to *test* the classifier. If the BDT *score* on the testing sample is significantly lower than the one on the training sample, a so-called *overtraining* may have occurred.

This case indicates that the tree exploits statistical fluctuations that lead to an artificially high performance on the training sample and cause a sub-optimal performance on an independent sample in return.

A way to reduce *overtraining* is to limit the tree size as described above. In this analysis, the simulated events for the various background and signal processes are split in two halves according to their event number, whereby one half is used to train the classifier and the other one to test it.

However, a tree that consists of only a few leaves with a high amount of events in each is a weak classifier in terms of performance. Thus, a collection of trees ( $N_{\text{trees}}$ ) is successively trained on all events, but with a higher weight on mis-classified events of the previous tree. This strategy, which causes each tree to pay more attention to events that were mis-classified in the previous tree is referred to as *boosting* [123, 128].

In this thesis, the adaptive boosting technique (*AdaBoost*) is chosen, which applies a weight  $\alpha \geq 1$  to all mis-classified events in tree  $i$ , given by

$$\alpha_i = \frac{1 - \text{err}_i}{\text{err}_i}, \text{ with } \text{err}_i = \frac{N_{\text{mis-classified}}}{N_{\text{all}}} \leq 0.5. \quad (7.2)$$

Subsequently, all events are re-weighted to regain the initial normalization.

Finally, all trees are combined in a single classifier according to

$$y(\mathbf{x}) = \frac{1}{N_{\text{trees}}} \sum_{i=1}^{N_{\text{trees}}} \ln(\alpha_i) \cdot h_i(\mathbf{x}), \quad (7.3)$$

whereby  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  represents the vector of input variables. Moreover,  $h_i(\mathbf{x}) = +1(-1)$  depends on whether the evaluated event was assigned to a signal-(background-)like leaf in tree  $i$ . Due to  $\ln(\alpha_i)$ , a larger weight is applied on trees with a lower error rate. Since *overtraining* cannot be diminished entirely, the training events are discarded afterwards. The remaining events are used to evaluate the tree and obtain the final BDT score distribution (*testing*).

A supplemental technique to reduce *overtraining* and reverse insignificant splitting procedures is denoted as *pruning* [129]. Every branch of the tree is examined in a bottom-up approach and a particular splitting is withdrawn if it did not yield a predefined gain in terms of mis-classification. In this case, the previous node, which contains a higher amount of events, is taken as final leaf.

To determine the optimal configuration, specified in Table 7.1, several BDT configurations were compared based on their ROC-curves. Exemplary configurations are shown in Fig. 7.3. In every tested configuration, specific parameters were altered with respect to the optimal configuration. To study the performance of another boosting technique, the so-called *gradient boost* [130], a distinct BDT was trained with the default gradient boost configuration provided in TMVA [125].

Table 7.1.: BDT configurations applied in this analysis.

Option	Chosen value
$N_{\text{trees}}$	1000
$\text{Depth}_{\text{max}}$	3
Boosting type	AdaBoost
Separation type	Gini Index
Pruning method	No pruning
nCuts	20
$n\text{Events}_{\text{min}}$	150

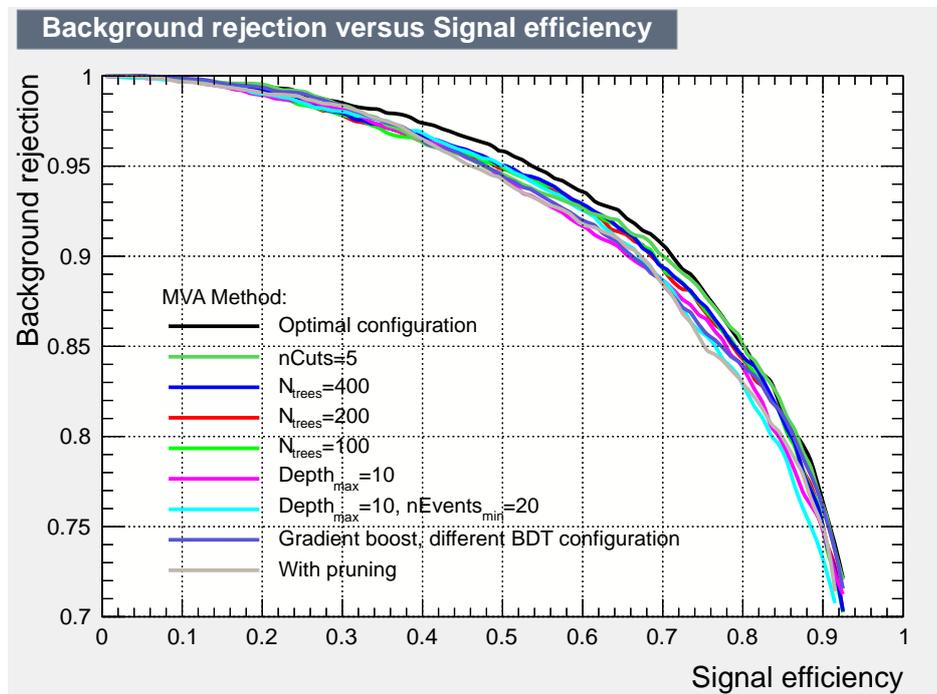


Figure 7.3.: ROC-curves for various BDT configurations, trained on the Boosted category which is defined in Table 7.2. The optimal configuration is outlined in Table 7.1. In the succeeding configurations, the specified training parameters were modified with respect to the optimal configuration. Thus, the dependence of the classifier on the different training parameters was examined. For the gradient boosting technique, which has a different foundation within TMVA, a BDT was trained with the default configuration provided in TMVA. The legend is ordered based on the classifiers background rejection capability at a signal efficiency of 70 %.

If the maximal depth of each tree was increased to ten layers, a decrease in the BDTs background rejection capability over a wide range of the signal efficiency could be observed. This could be caused by statistical fluctuations due to the lower number of events in the final leaves. In addition, a lower number of single trees, which are combined to the final classifier

in the end, led to a lower background rejection. Moreover, if a coarse grid was chosen to scan the cut-values on each variable (here,  $n\text{Cuts}=5$  was assigned), a decrease in signal significance occurred, which was enhanced in the region with low signal efficiency. This region corresponds to high values of the BDT score, where a finer scan improves the separation between signal-like and background-like phase space areas. The *pruning* algorithm might have caused an overall lower background rejection capability because of the particular tree depth and high minimal event yield specified in the residual BDT parameters. The determined optimum of the BDT configuration parameters might have reduced the *overtraining* already to a great extent.

## 7.2. Event classification

The BDT analysis is based on all events fulfilling the preselection of Section 5.2.1. Prior to the training, a loose categorization is applied to separate specific event topologies. However, it preserves a large number of events and disposes the actual signal extraction and background separation to the BDT itself.

The final signal region and classification cuts are listed in Table 7.2. Thereby, the  $E_{\text{T}}^{\text{miss}}$  and  $m_{\text{T}}$  cuts are similar to those applied in the cut-based analysis in order to suppress fake backgrounds and define an orthogonal W+jets control region. In addition, a converging<sup>4</sup>  $m_{\tau\tau}^{\text{MMC}}$  and a minimum  $\tau_{\text{had}}$  momentum are implemented to enhance the rejection of fake- $\tau$ .

The VBF category is constructed in a way that ensures a high efficiency of the VBF-filtered Z samples (see Section 5.3.7). Thus, the VBF-filtered samples with a large number of events are used for the BDT training and evaluation in this case, whereas the embedded sample is taken in the other categories. A distinct study considering modified  $Z \rightarrow \tau\tau$  templates in the training is given in the next section. Concerning the data-driven fake- $\tau$  estimation, the OS-SS method outlined in Section 5.3.1 is employed in all categories. This is justified by the sufficient SS data event count when applying only loose VBF cuts. Possible improvements are discussed in Section 8.4.

The OS-SS method considers additional backgrounds by their *add-on* shape (Eq. 5.5) obtained by subtraction of the simulated SS events from simulated OS events. In this analysis, only those with opposite signed charge are taken for the residual background contribution. To restore their initial relative weight, background  $i$  is scaled according to

$$N_{\text{est.}}^i = \frac{N_{\text{add-on}}^i}{N_{\text{OS}}^i}. \quad (7.4)$$

<sup>4</sup>The MMC algorithm (Section 5.2.2) assigns a negative value if the parameter fit did not converge.

Table 7.2.: Signal region cuts applied on top of the preselection (Section 5.2.1). These are followed by a loose categorization of events in order to enhance certain Higgs boson production modes and exploit a transverse momentum of the  $\tau\tau$  system. The relative signal composition of each category is specified in the last row.

<b>Additional signal region cuts</b>	
$m_{\tau\tau}^{\text{MMC}} > 50 \text{ GeV}$ ,	$p_{\text{T}}^{\tau} > 30 \text{ GeV}$ ,
$p_{\text{T}}^{\tau} > 30 \text{ GeV}$ ,	$E_{\text{T}}^{\text{miss}} > 20 \text{ GeV}$ ,
	$m_{\text{T}} < 70 \text{ GeV}$
<b>BDT event categorization</b>	
Category	0J
<b>VBF</b> - At least two jets with $p_{\text{T}}^{\text{H}} > 40 \text{ GeV}$ , $p_{\text{T}}^{\text{J}} > 30 \text{ GeV}$ $\Delta\eta_{j_1, j_2} > 2.0$ and $m_{j_1, j_2} > 200 \text{ GeV}$ - All jets fail the JetFitterCombNN <i>b-tagger</i> at 70 % signal efficiency	<b>Boosted</b> - failing VBF categorization $p_{\text{T}}^{\text{H}} > 80 \text{ GeV}$
<b>Signal composition</b> $\sim 64 \%$ VBF, $\sim 34 \%$ ggF, $\sim 2 \%$ VH	$> 95 \%$ ggF, $\sim 3 \%$ VBF, $\sim 2 \%$ VH

Figure 7.4 illustrates, that the treatment of *add-on* backgrounds via OS-SS or with a re-weighted OS sample does provide significant shape differences but yield a similar maximal signal significance in the end. The figure of merit is based on the Asimov approximation derived in Appendix A:

$$\text{Significance} = \left[ 2 \left( (s+b) \ln \left[ \frac{(s+b)(b+\sigma_b^2)}{b^2 + (s+b)\sigma_b^2} \right] - \frac{b^2}{\sigma_b^2} \ln \left[ 1 + \frac{\sigma_b^2 s}{b(b+\sigma_b^2)} \right] \right) \right]^{\frac{1}{2}}. \quad (7.5)$$

This equation includes the simulated signal ( $s$ ) and background ( $b$ ) events as well as the statistical uncertainty of the background ( $\sigma_b^2$ ). It is a simplification of the limit setting procedure of Chapter 4 that neglects systematic uncertainties and does not perform a profiling of nuisance parameters with respect to a specified dataset (comparable to a blinded limit, Section 4.5.1).

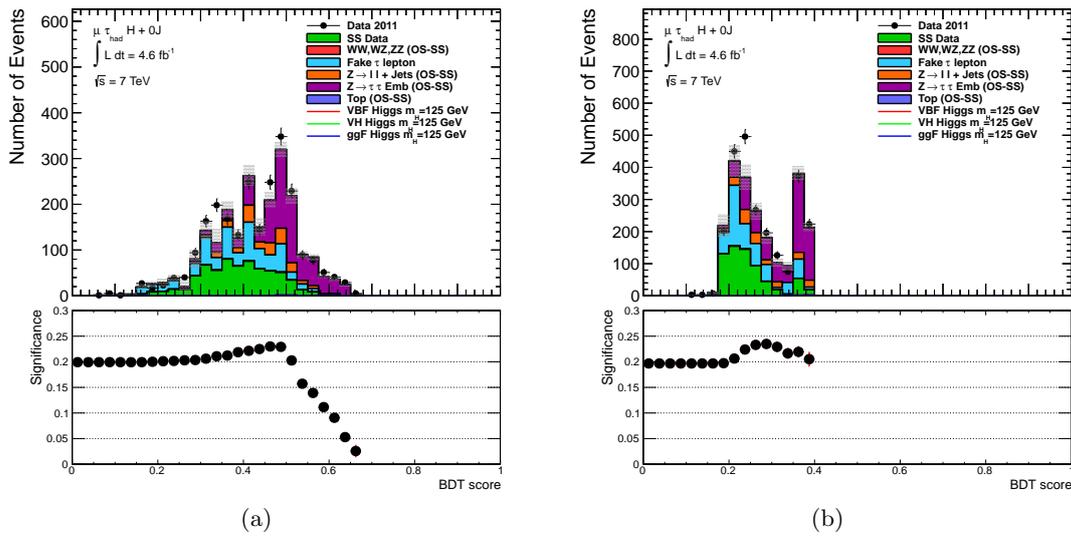


Figure 7.4.: (a) BDT score distribution for a training with OS events only, which are weighted to their relative amount obtained in the OS-SS background estimation. (b) BDT score distribution obtained via training with OS and SS events, where the SS events attained a negative weight. The lower plots display the signal significance for different BDT cut thresholds. Therefore, every point states the signal significance of all events above the particular BDT value.

Each point in the lower region of Fig. 7.4 refers to the significance of the events above the specific BDT score. Hence, the maximum significance indicates the highest signal-to-background ratio that can be achieved and is a measure for the separation between signal and background events in the distribution.

### 7.3. Training parameter optimization

A multitude of input variables can be used to train the BDT classifier, whereas each of them is required to display an agreement between data and MC. An elementary set of variables is defined by the variables used in the cut-based analysis plus additional variables which are introduced in the following:

- **$E_T^{\text{miss}} - \phi$  centrality:** This variable quantifies the  $\phi$  direction of  $E_T^{\text{miss}}$  with respect to the selected  $\tau_{\text{had-vis}}$  and lepton. All vectors are normalized to unity and  $E_T^{\text{miss}}$  is projected in the  $\tau_{\text{had-vis}}$ -lepton plane:

$$\mathbf{E}_T^{\text{miss}} - \phi \text{ centrality} = \frac{A + B}{\sqrt{A^2 + B^2}} \quad \text{with} \quad (7.6)$$

$$A = \frac{\sin \phi_{E_T^{\text{miss}}} - \phi_{\text{lep}}}{\sin \phi_\tau - \phi_{\text{lep}}}, \quad B = \frac{\sin \phi_\tau - \phi_{E_T^{\text{miss}}}}{\sin \phi_\tau - \phi_{\text{lep}}}.$$

The  $E_T^{\text{miss}} - \phi$  centrality yields 1 if  $E_T^{\text{miss}}$  is perfectly aligned with either of them and rises to  $\sqrt{2}$  in the center. Furthermore, a value below 1 is assigned if  $E_T^{\text{miss}}$  lies outside of both. This occurs frequently in events with mis-identified  $\tau_{\text{had}}$ .

- **$\eta$  centrality:** The  $\eta$  centrality is defined as

$$C_{\eta_{j1}, \eta_{j2}}(\eta_{\text{lep}}) = \exp \left[ \frac{-1}{\left(\frac{\eta_{j1} - \eta_{j2}}{2}\right)^2} \cdot \left( \eta_{\text{lep}} - \frac{(\eta_{j1} + \eta_{j2})}{2} \right)^2 \right]. \quad (7.7)$$

It quantifies the position of the lepton (or  $\tau$ ) in between the two *tagging jets*. It is 1 if the lepton is aligned with one of the *tagging jets* and reaches the value  $1/e$  in the center of the  $\eta$ -range spanned by the two jets. If the lepton (or  $\tau$ ) is outside this range, the value of  $C_{\eta_{j1}, \eta_{j2}}(\eta_{\text{lep}})$  is below  $1/e$ .

- **$\sum \mathbf{p}_T$ :** This quantity describes the scalar momentum sum of the selected lepton and  $\tau$  including the  $E_T^{\text{miss}}$  and possible jets.
- **Sphericity:** The sphericity measures the isotropy of the final state. It takes the momentum vectors of all event objects into account. A normed sphericity tensor is built via

$$S^{\alpha, \beta} = \frac{\sum_i \mathbf{p}_i^\alpha \mathbf{p}_i^\beta}{\sum_i |\mathbf{p}_i|^2}, \quad \alpha, \beta = 1, 2, 3. \quad (7.8)$$

Thereof, the eigenvalues  $\lambda_1 \geq \lambda_2 \geq \lambda_3$ , featuring  $\lambda_1 + \lambda_2 + \lambda_3 = 1$ , are computed and the sphericity is defined as  $S = \frac{3}{2}(\lambda_2 + \lambda_3)$ . It attains values between 0 and 1, whereas

1 specifies an isotropic event and 0 stands for an event shape where all vectors lie on one axis.

All variables exploited within the analysis are outlined in Table 7.3. Some of them are solely explored for the VBF topology because they rely on specific quantities of the *tagging jets*. The corresponding distributions are given in Appendix C. No significant mis-modeling between data and simulated events is observed.

Table 7.3.: Set of input variables used in all categories (left column) and those relying on di-jet quantities, exploited for the VBF category only (right column).

Variables for all categories	Additional VBF variables
$m_T$	$m_{j_1, j_2}$
$\Delta R$	$\eta_{j_1} \times \eta_{j_2}$
$\sum \Delta\phi$	$\Delta\eta_{j_1, j_2}$
$p_T^H$	$\eta$ centrality
$p_T^{\text{asym}}$	$p_T^{\text{total}}$
$E_T^{\text{miss}} - \phi$ centrality	
$m_{\tau\tau}^{\text{MMC}}$	
$\sum p_T$	
$E_T^{\text{miss}}$	
$x_l, x_h$	
sphericity	

### 7.3.1. Input variable and sample optimization

After a compendium of possible input variables is defined, an optimal combination has to be determined according to a specific figure of merit, e.g. the ROC curves (see Section 7.1). Another possibility to select significant input variables is the variable ranking provided by TMVA, described below. Thereby, the number of cuts on the specific variable within the collection of trees is counted, whereas each is weighted by the separation gain and the number of events in the corresponding node [16]. This classification was used in first place to discard input variables with negligible impact.

For the final specification, the limit setting procedure (Chapter 4) was utilized and the optimization was based on the expected upper limit on the signal production cross section at a 95 % confidence level, normalized to the SM cross section  $\sigma_{\text{SM}}$  (in the following referred to as expected limit). Thereby, only statistical uncertainties of the data and MC events were taken into account in order to analyze multiple configurations on a short timescale.

To model the  $Z \rightarrow \tau\tau$  background, various samples could be selected. According to Section 5.3.7 and the following, an ALPGEN MC sample, one obtained through *embedding* or the VBF-filtered  $Z$  sample, featuring a high event count, could be used. Furthermore, differences when applying additional signal region cuts (Table 7.2) on the training sample were studied.

Figure 7.5 depicts several expected exclusion limits for varying training configurations in the 0J, Boosted and VBF category, respectively. The configurations are described in detail in Table 7.4. In each case, it was found that the limit is improved if the signal region cuts were carried out on both training and testing samples, since the background composition of the training sample and its kinematic properties are thereby closer to the signal region. Moreover, an increased number of events in training due to additional ALPGEN MC events for the  $Z \rightarrow \tau\tau$  background helped in the 0J and Boosted category. However, ALPGEN events were not taken for testing due to the known jet-topology mis-modeling, which was the initial motivation to implement the *embedding* procedure.

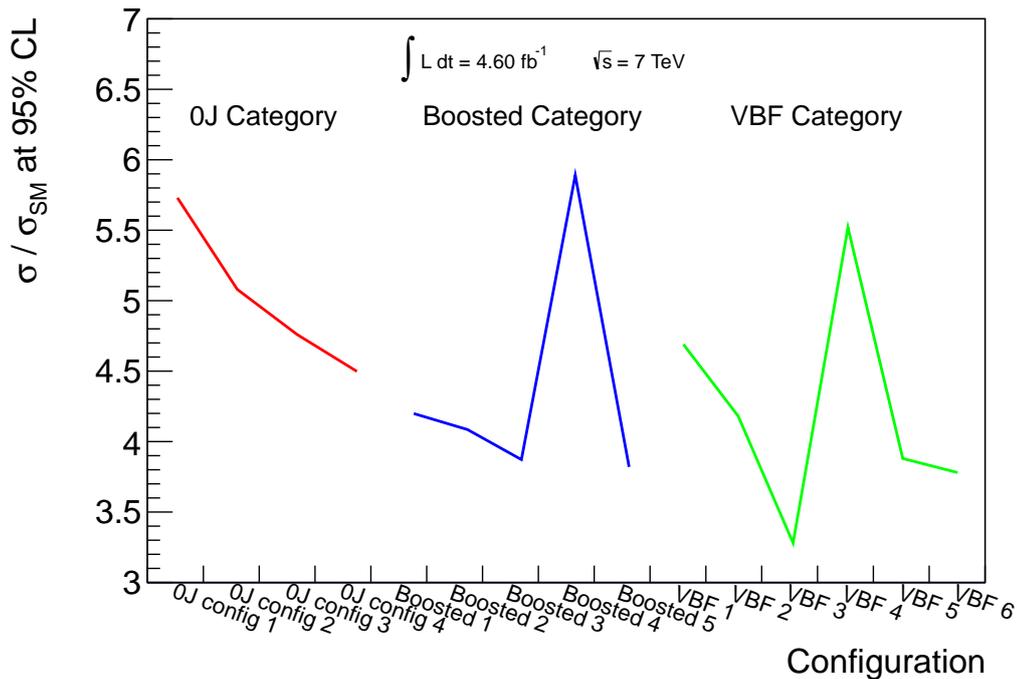


Figure 7.5.: Expected exclusion limits obtained during optimization of the signal region cuts, input variables and event samples. The corresponding configurations are specified in Table 7.4.

In contrary, additional training events did not help in case of the VBF category. This may result from their small statistical impact<sup>5</sup> or from disadvantages due to statistical fluctuations of the combined sample.

<sup>5</sup>The *embedding* sample accounts for roughly 10 % in the VBF topology after combination with the VBF-filtered sample.

Table 7.4.: Detailed listing of input variables, event samples and eventual signal region cuts used in the training. The rightmost column states the corresponding expected exclusion limit, whereas the colored configurations indicate the final ones chosen for the 0J, Boosted and VBF category, respectively. The initial configuration in each category utilizes no signal region cuts (see Table 7.2) for the training events. In addition, the training started with *embedded*  $Z \rightarrow \tau\tau$  samples in the 0J and Boosted category, whereas VBF-filtered samples were taken for the VBF category.

Configuration	Description	Limit [ $-\frac{\sigma}{\sigma_{\text{SM}}}$ ]
0J config 1	Variables: $m_{\text{T}}, \sum \Delta\phi, \Delta R, \sum p_{\text{T}}^{\text{H}}, p_{\text{T}}^{\text{H, asym}}, E_{\text{T}}^{\text{miss}} - \phi$ centrality, $m_{\tau\tau}^{\text{MMC}}$	5.73
0J config 2	config 1 plus sphericity, $\sum p_{\text{T}}$	5.08
0J config 3	config 2 plus signal region cuts in training	4.76
0J config 4	<b>config 3 plus ALPGEN <math>Z \rightarrow \tau\tau</math> events in training</b>	<b>4.50</b>
Boosted 1	Variables: $m_{\text{T}}, p_{\text{T}}^{\text{H}}, \Delta R, \sum \Delta\phi, m_{\tau\tau}^{\text{MMC}}$	4.20
Boosted 2	Boosted 1 with ALPGEN $Z \rightarrow \tau\tau$ events in training	4.09
Boosted 3	Boosted 2 plus $E_{\text{T}}^{\text{miss}} - \phi$ centrality, $\sum p_{\text{T}}$ , sphericity	3.87
Boosted 4	Boosted 3, but only ALPGEN $Z \rightarrow \tau\tau$ events in training; entire <i>embedded</i> sample in testing	5.89
Boosted 5	<b>Boosted 2 with additional signal region cuts according to Table 7.2</b>	<b>3.83</b>
VBF 1	Variables: $m_{\text{T}}, \sum \Delta\phi, m_{j_1, j_2}, \sum p_{\text{T}}, \eta_{j_1} \cdot \eta_{j_2}, \Delta R, p_{\text{T}}^{\text{H}}, p_{\text{T}}^{\text{H, asym}}, m_{\tau\tau}^{\text{MMC}}$	4.69
VBF 2	$E_{\text{T}}^{\text{miss}} - \phi$ centrality, sphericity, $p_{\text{T}}^{\text{total}}, \Delta\eta_{j_1, j_2}$ and $\eta$ centrality	4.18
VBF 3	VBF1 minus $E_{\text{T}}^{\text{miss}} - \phi$ centrality, sphericity, $p_{\text{T}}^{\text{total}}$ and $\Delta\eta_{j_1, j_2}$ <b>VBF2 with signal region cuts in training</b>	<b>3.28</b>
VBF 4	VBF3 with <i>embedded</i> $Z \rightarrow \tau\tau$ instead of VBF-filtered events	5.52
VBF 5	VBF3 with the entire <i>embedded</i> plus half of the VBF-filtered events in training	3.88
VBF 6	VBF3 trained against the VBF Higgs boson production mode only	3.78

To justify the use of VBF-filtered samples in the BDT, their capability to model the input variables and their correlation had to be confirmed. According to Section 5.3.7, they were re-weighted according to differences between data and MC events in the one-dimensional  $\Delta\eta_{j_1, j_2}$  distribution. The impact of this procedure can be easily examined in a cut-based analysis. However, in the case of a multivariate analysis, complex phase space cuts involving the correlations of the input variables are used to carve out regions with high signal significance. Hence, the VBF-filtered and embedded samples had to be compared in the context of their final BDT score distribution for the VBF category. In Figure 7.6, the BDT score distributions for both are displayed together with their statistical uncertainties. No significant deviation was observed and the VBF-filtered samples were eventually used for the VBF category.

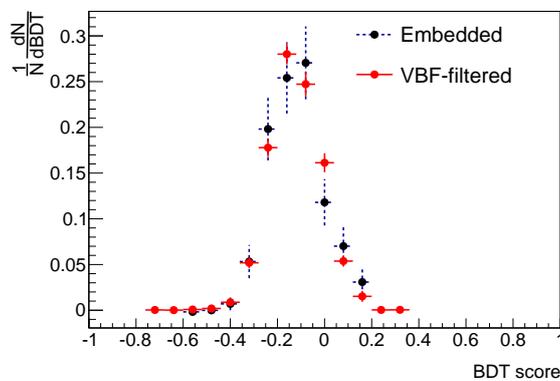


Figure 7.6.: BDT score distribution of the VBF category defined in Table 7.2. The results obtained by using the *embedded* (black dots) or the VBF-filtered (red dots)  $Z \rightarrow \tau\tau$  samples are shown together with their statistical uncertainties.

### 7.3.2. Cross evaluation

A major drawback of multivariate analyses is that the training sample has to be discarded in the final analysis to avoid a bias due to *overtraining*. In the event sample optimization, an improvement through additional training events was observed. In contrary, more events in the evaluation lead to smaller statistical uncertainties.

To benefit from both, two independent BDT classifiers are trained. In this analysis, events with an odd event number are used to train a BDT at first, while those with even event number are used to test it and vice versa. This technique, referred to as *cross evaluation*, leads to two distinct BDT distributions. After careful validation of those, all tested events are combined to establish one BDT distribution in the end. Figure 7.7 illustrates, that the statistical uncertainties of this technique are smaller due to the higher quantity of simulated events. Moreover, the signal significance (see Section 7.2) does not alter significantly.

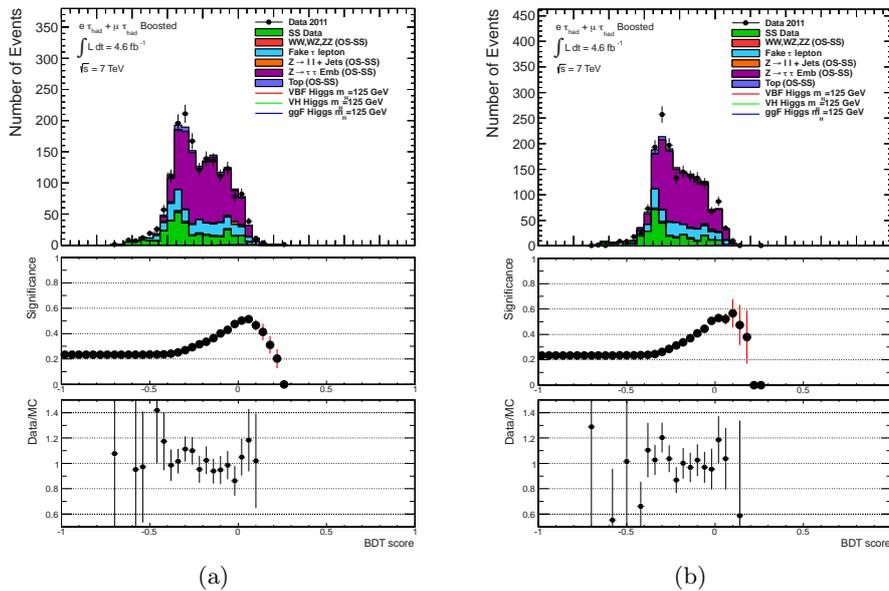


Figure 7.7.: BDT score distribution with their corresponding data/MC ratio and signal significance. Thereby, the *cross evaluation* (a) is compared to an evaluation with discarded training events (b) in the Boosted category. The signal significance for all events above a specific BDT score threshold (Eq. 7.5) is displayed in the middle plots. The data-to-MC ratio, incorporating the statistical uncertainties of both, is shown in the lower plots.

## 7.4. Signal region distributions

Finally, Figure 7.8 displays the BDT score distributions obtained from *cross evaluation* for all categories defined in Table 7.2. The corresponding training configuration is outlined by the colored rows in Table 7.4 for the 0J, Boosted and VBF category, respectively. The signal significance and its separation from background is much higher in the Boosted and VBF topology compared to the 0J category, because either the transverse boost of the di- $\tau$  system or the two characteristic *tagging jets* are exploited to distinguish between signal and background. The data shows a reasonable agreement with the subsumed background predictions in low BDT score regions, where no signal contribution is expected.

Apart from the signal region, the agreement between data and simulated events in the BDT score distributions is examined in various control regions. In each CR, a specific background process is enhanced and a corresponding data-driven scale factor is usually obtained. Due to the correlations between input variables that are taken into account in a BDT classifier, its outcome has to be inspected in these control regions. A mis-modeling would reveal a potential dependence of the input variables and correlations exploited in the BDT training.

In this analysis, the control regions outlined in Sections 5.3.3 and the following are em-

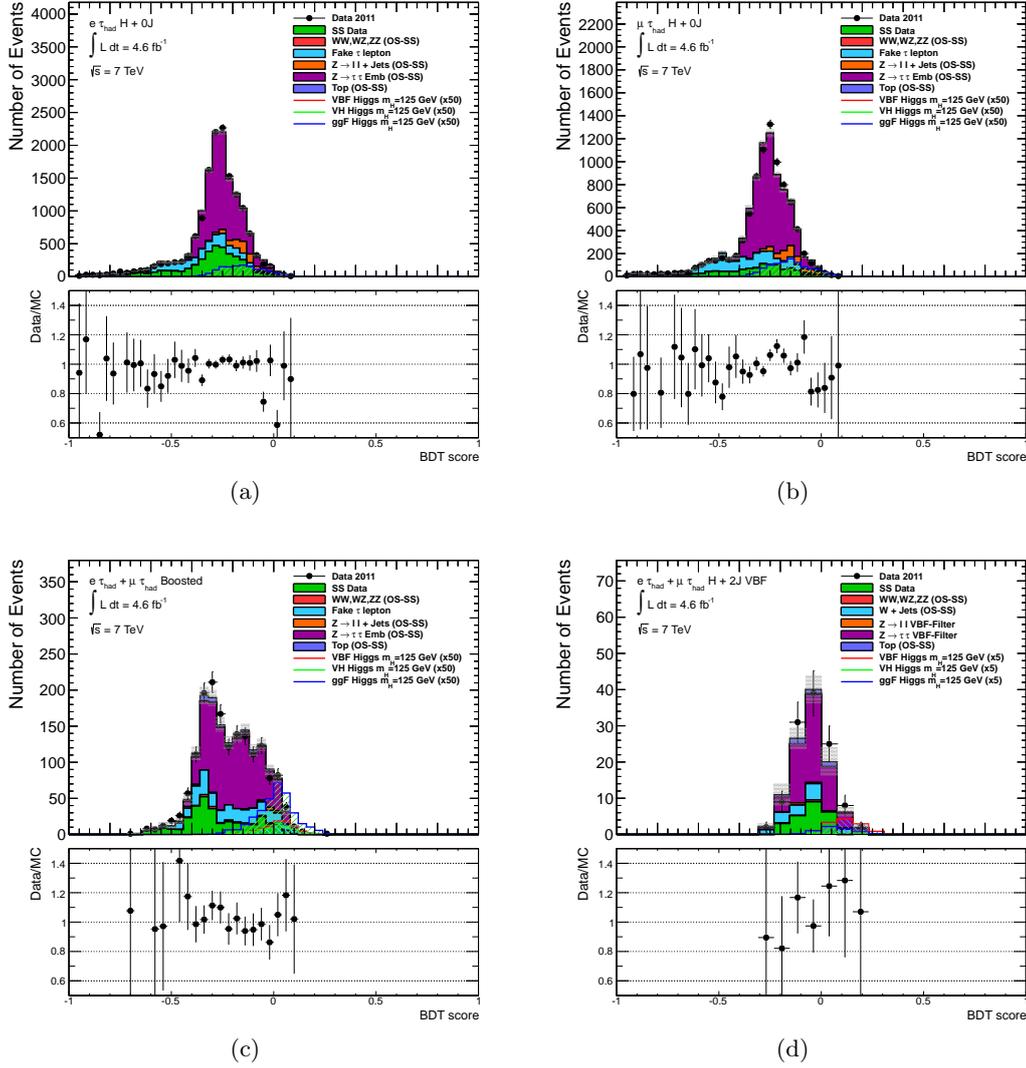


Figure 7.8.: Final BDT score distributions of the signal region for the 0J (a, b) together with the Boosted (c) and VBF (d) category. The former is divided into the  $e + \tau_{\text{had}}$  (top left) and  $\mu + \tau_{\text{had}}$  (top right) channels. The signal is scaled by 50 in all categories except of VBF, where it is multiplied by a factor of 5. Each distribution is accompanied by its corresponding data-to-MC ratio, which includes the statistical uncertainties of both.

employed. The W CR contains all events with a  $m_T > 70 \text{ GeV}$  on top of the preselection described in Section 5.2.1. Likewise, the  $t\bar{t}$  control region is based on the preselection with the additional requirement of  $m_T > 55 \text{ GeV}$  and at least one b-jet passing the 70 % *b-tagging* working point. To conclude, the  $Z \rightarrow \ell\ell + \text{jets}$  CR applies several cuts on top of the 0J category. The invariant mass of the electron and visible  $\tau_{\text{had}}$  candidate is required to be within  $80 \text{ GeV} < m_{e,\tau} < 100 \text{ GeV}$ , the transverse mass has to be between  $30 \text{ GeV} < m_T < 40 \text{ GeV}$  and only  $\tau$ s with one associated track are taken into consideration.

In Figure 7.9, the BDT score distributions are displayed for the W and top CRs for the Boosted and VBF category. Furthermore, Fig. 7.10 illustrates the BDT score distributions for the corresponding CRs for the 0J category. Neither a slope in the ratio between data and MC events nor a constant bias between data and simulation is observed.

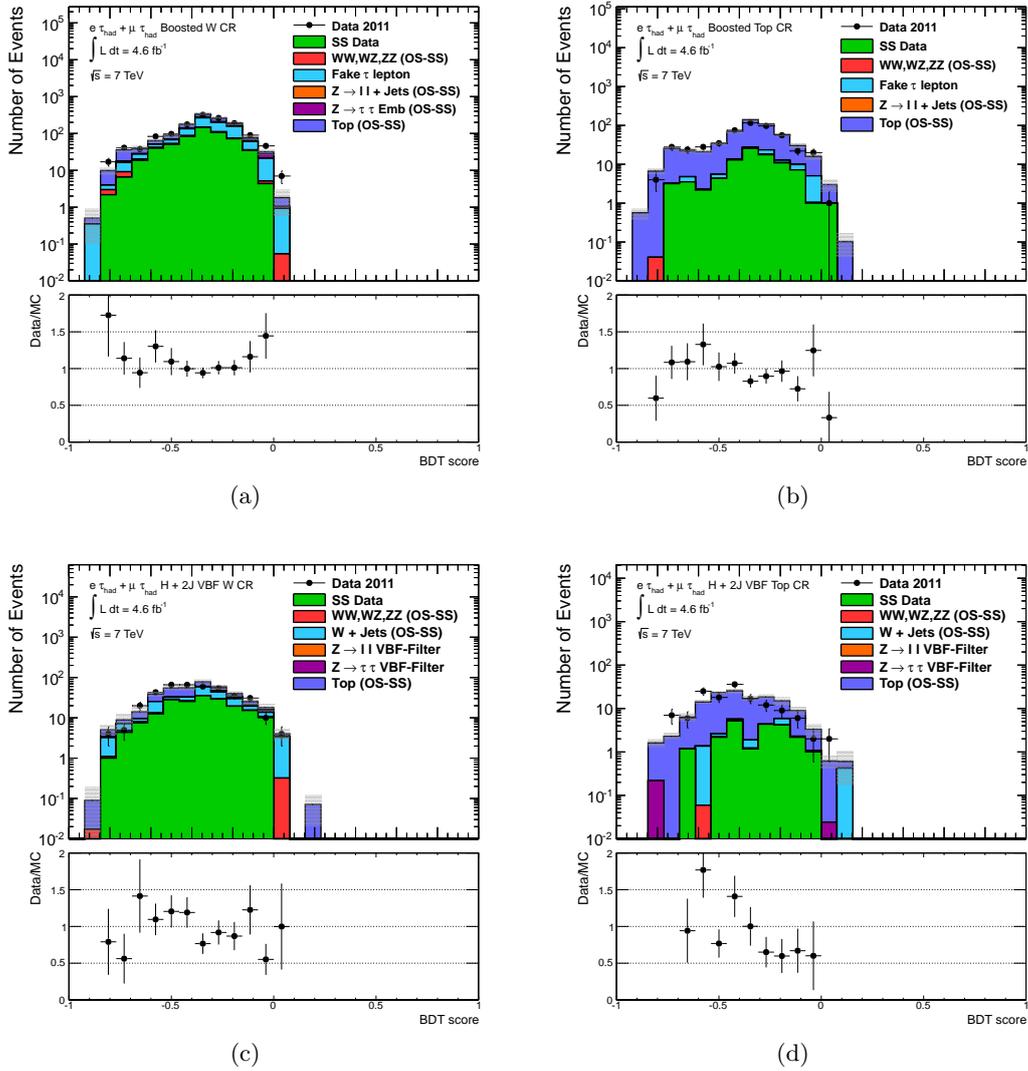


Figure 7.9.: BDT score distributions for the W (left column) and top (right column) control regions, shown for the Boosted (a, b) and VBF (c, d) categories. Both  $e + \tau_{\text{had}}$  and  $\mu + \tau_{\text{had}}$  channels are combined. Each distribution is accompanied by its corresponding data-to-MC ratio which incorporates the statistical uncertainties of both.

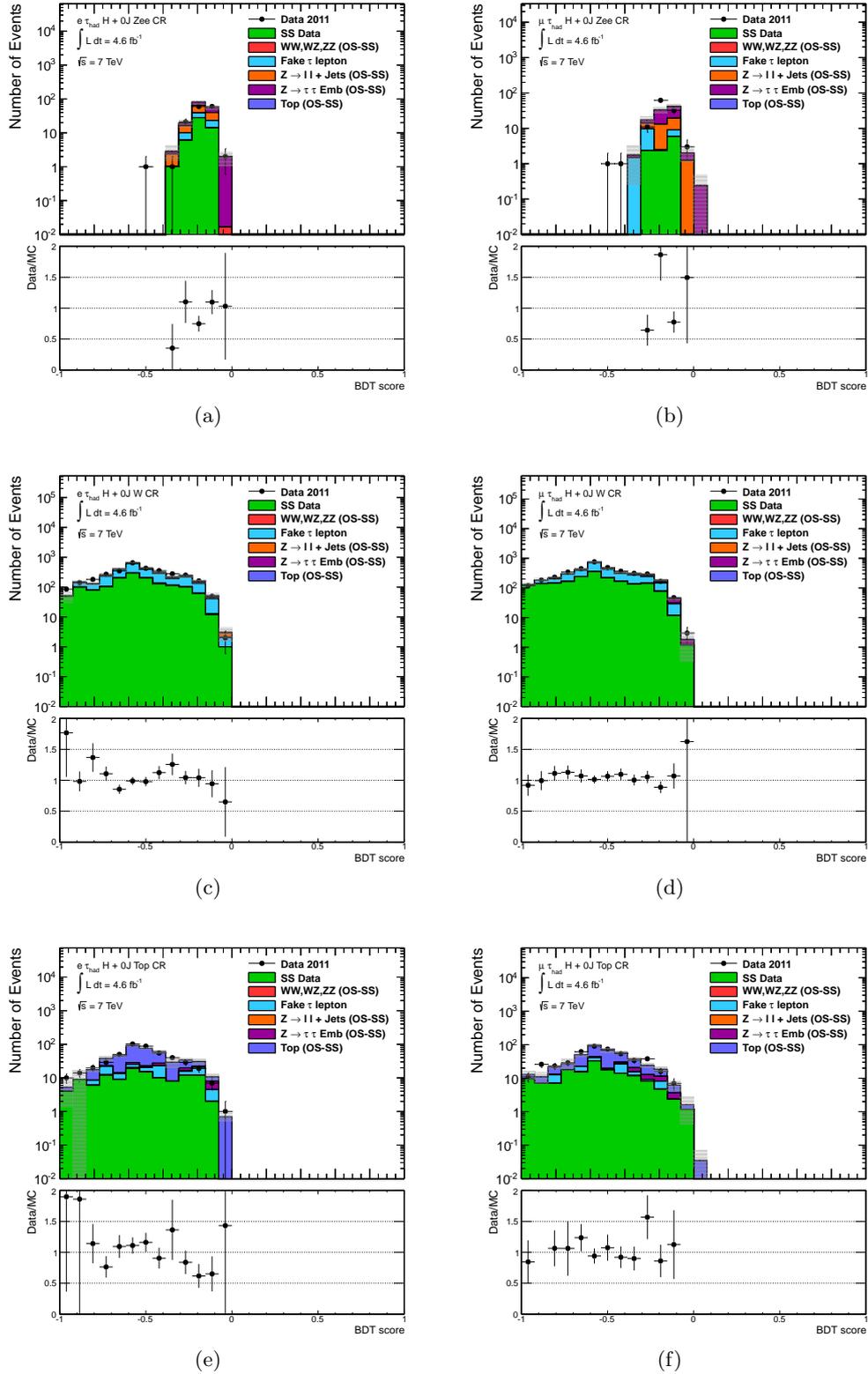


Figure 7.10.: BDT score distributions for the control regions of the 0J category, split into the  $e + \tau_{\text{had}}$  channel (left column) and  $\mu + \tau_{\text{had}}$  channel (right column). The  $Z \rightarrow ee$  (a) and  $Z \rightarrow \mu\mu$  (b) CR is depicted in the top row, the W CRs (c, d) in the middle row and the top CRs (e, f) in the lower row. Each distribution is accompanied by its corresponding data-to-MC ratio, including the statistical uncertainties of both.

## 7.5. Further BDT studies

According to Table 7.4,  $m_{\tau\tau}^{\text{MMC}}$  is involved in the training of each category. The implementation of the reconstructed mass of the di- $\tau$  system in the BDT training was debated within the Higgs working group of the ATLAS collaboration.

For one, it is desired to provide a mass distribution in the end, where a possible Higgs boson signal can be observed as a bump in the high mass tail of the distribution. Moreover, the BDT distribution as final observable within the limit computation bears the risk that additional systematic uncertainties are not properly estimated. For example, the scale factors which are obtained in specific CRs in order to normalize the expected background contributions to data can possess different values in different regions of the BDT score distributions. In the background estimation of Section 5.3, these scale factors are derived independently from the final discriminant. In addition, the theory uncertainties on the Higgs boson production modes (see Section 8.1.1) are sensitive to jet-related kinematic properties, which are again highly correlated to the BDT score. To illustrate the dependence on the latter, Fig. 7.11 displays a two-dimensional scatter plot of  $\Delta\eta_{j_1, j_2}$  against the BDT score.

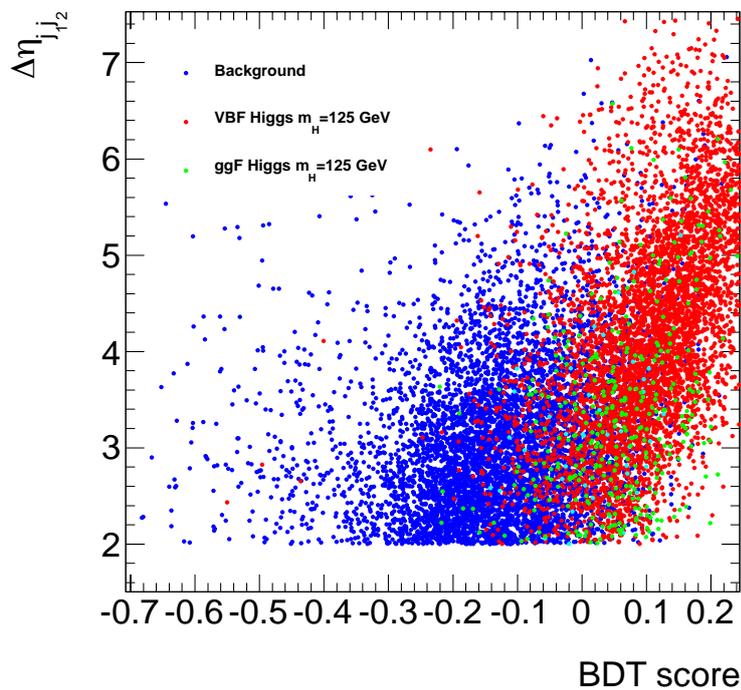


Figure 7.11.: Two-dimensional scatter plot of  $\Delta\eta_{j_1, j_2}$  against the BDT score, displayed for the VBF category. The blue dots show the distribution of background events in the two-dimensional plane. In addition, the red and green dots depict the signal events for the VBF and gluon fusion production modes, respectively

A clear correlation is observed, especially in the region of high BDT scores, where signal events with a large  $\Delta\eta_{j_1, j_2}$  are allocated. This region has the highest signal sensitivity and is therefore crucial for the computation of upper limits on the Higgs production cross section.

To account for the observed behavior, the input variables and their entire correlations have to be considered on reconstruction level to assess the impact of theoretical scale uncertainties on the BDT, which is not available for the MC generators used to derive these uncertainties. As a result, an alternative training without  $m_{\tau\tau}^{\text{MMC}}$  was tested. It should quantify the importance of the reconstructed mass of the  $\tau\tau$  system within the BDT classifier and provide an alternative approach for the signal extraction. Therefore, the BDT score is cut at a specific value. The  $m_{\tau\tau}^{\text{MMC}}$  distributions for all events above this threshold are used to compute exclusion limits similar to the cut-based approach (see Chapters 6 and 8).

BDT distributions in a training similar to the baseline defined above, but without  $m_{\tau\tau}^{\text{MMC}}$ , are shown in Fig. 7.13 and yield a decent agreement between data events and MC prediction. The corresponding BDT score distributions for the CRs are pictured in Appendix D. All distributions for the 0J, Boosted and VBF category show a reasonable agreement between data and simulated events. Hence, no indication for a mis-modeling is observed.

To optimize the signal significance, several cut thresholds on the BDT distributions were scanned for each category. For each cut value, an expected exclusion limit with the full systematic (Section 8.1) and statistical uncertainties was computed for the resulting  $m_{\tau\tau}^{\text{MMC}}$  distribution. The expected exclusion limits are illustrated for both the Boosted and the VBF category in Fig. 7.12. The lowest limit was chosen for the final BDT-cut threshold.

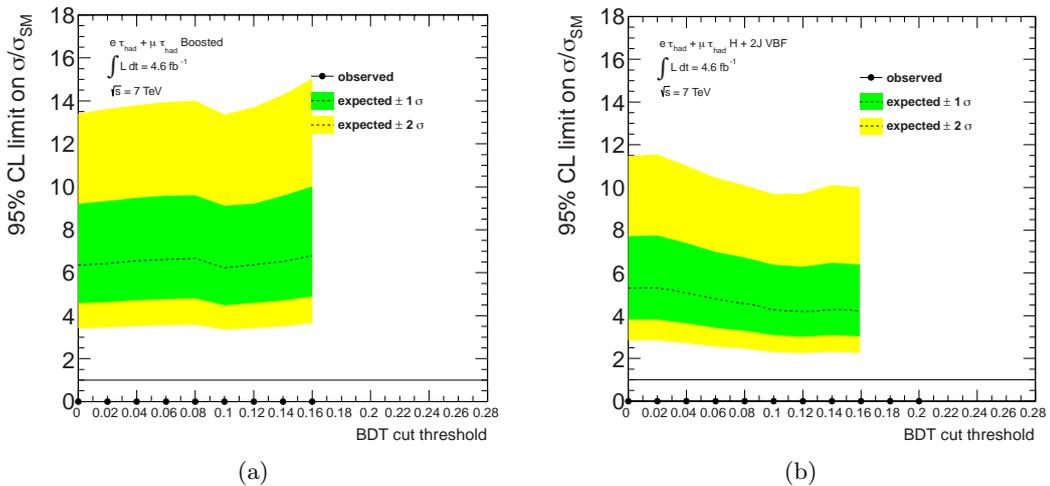


Figure 7.12.: Expected exclusion limits on the Higgs boson cross section on 95 % CL, normalized to the SM cross section  $\sigma_{\text{SM}}$  for the Boosted (a) and VBF (b) category. The x-axis value denotes the cut value on the BDT score.

In case of the Boosted and VBF categories, cut values of 0.1 and 0.12, respectively, were chosen. For both 0J categories, a cut value of 0.04 was chosen. Above a certain cut threshold, the number of events was too low to accurately model the background contributions and no limit was computed.

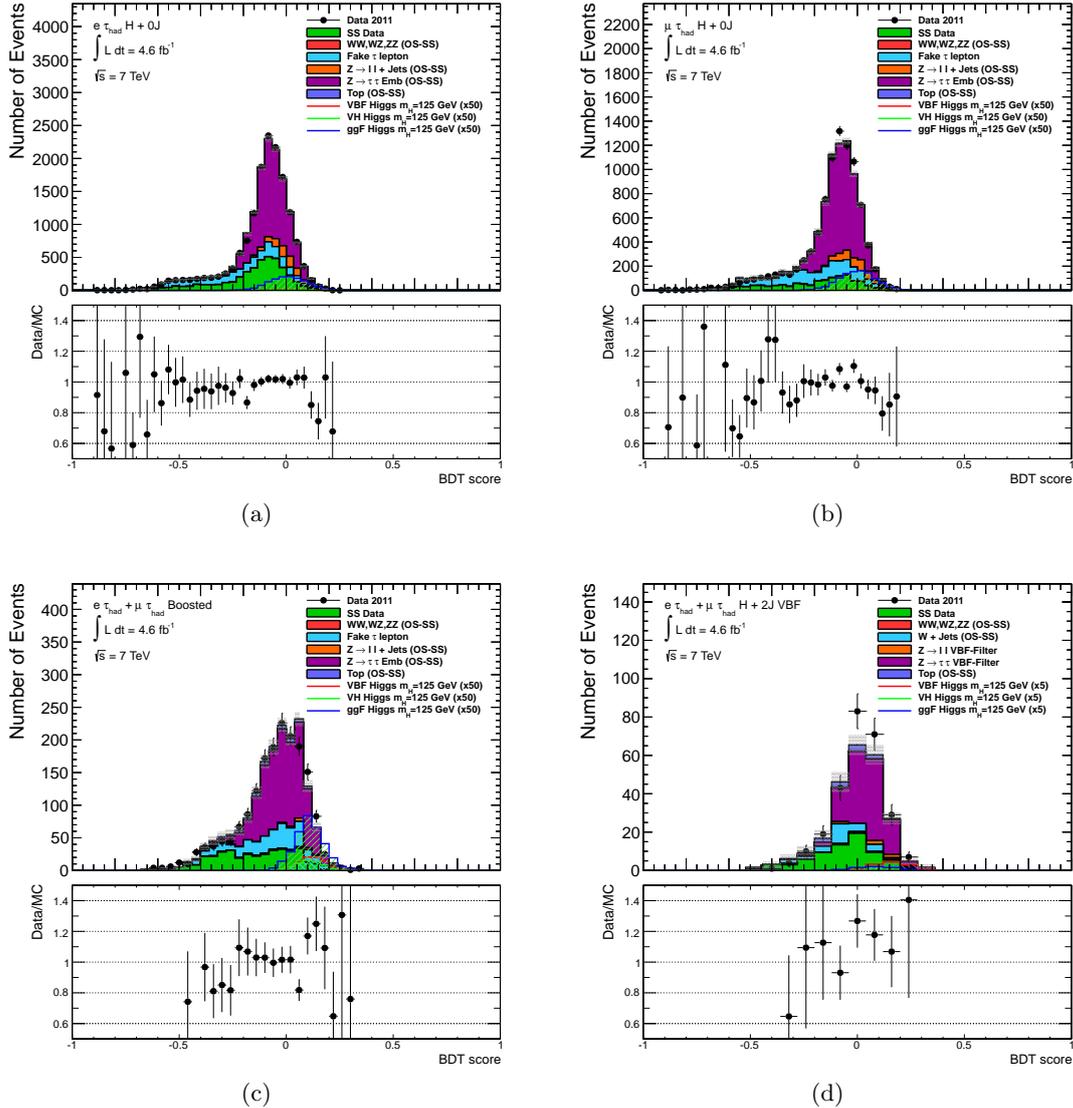


Figure 7.13.: Final distributions of the BDT score in the signal region for the 0J (a,b) together with the Boosted (c) and the VBF (d) category, similar to Fig. 7.8. The  $m_{\tau\tau}^{\text{MMC}}$  was discarded in the BDT training. Each distribution is accompanied by its corresponding data-to-MC ratio, incorporating the statistical uncertainties of both.

The final distributions of the reconstructed mass of the  $\tau\tau$  system, after a cut on the BDT score to achieve the best possible performance, are displayed in Fig. 7.14. The event yield of all categories is drastically diminished with respect to Fig. 7.8, whereby the signal sensitivity

in all categories improved compared to the  $m_{\tau\tau}^{\text{MMC}}$  distributions of the cut-based analysis (Fig. 6.7). The data events show a reasonable agreement with respect to the simulated background processes.

A detailed statistical evaluation and comparison of the outcome for both BDT approaches together with the cut-based analysis of Chapter 6 is carried out in Chapter 8. Thereby, distinct advantages and drawbacks are discussed.

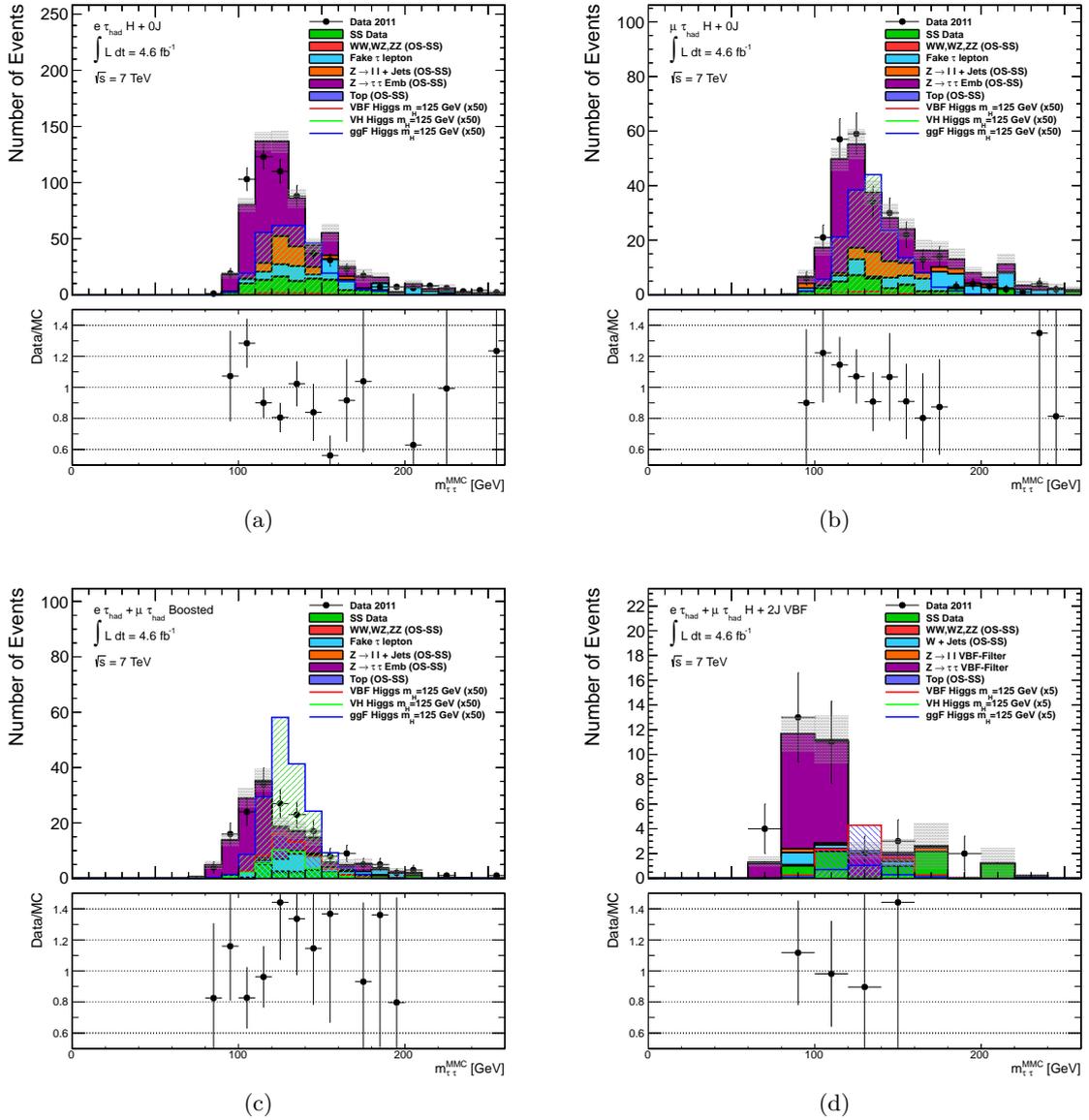


Figure 7.14.: Final distributions of  $m_{\tau\tau}^{\text{MMC}}$  in the signal region. The  $m_{\tau\tau}^{\text{MMC}}$  distributions are shown for the 0J category divided in the  $e + \tau_{\text{had}}$  (a) and  $\mu + \tau_{\text{had}}$  (b) channels, together with the Boosted (c) and the VBF (d) category. The signal is scaled by a factor of 50 in all categories except of VBF, where it is multiplied by a factor of 5. Each distribution is accompanied by its corresponding data-to-MC ratio, including the statistical uncertainties of both.



---

# 8 Exclusion limits on the Higgs boson production cross section

---

For the final distributions of the analyses established in Chapters 6 and 7, no significant excess of data events above SM background expectations is observed (see Figs. 6.7 and 7.8). However, a dataset with a corresponding integrated luminosity of  $5 \text{ fb}^{-1}$  at a center-of-mass energy of  $\sqrt{s} = 7 \text{ TeV}$  does not allow for a discovery in the  $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$  analysis. This judgement is based on the experimental challenge of identifying a hadronically decaying  $\tau$ , the irreducible  $Z \rightarrow \tau\tau$  background with a much larger production cross section than the SM Higgs boson and the presence of neutrinos that prohibit a straight forward mass reconstruction with high resolution.

Thus, upper exclusion limits on the production cross section of the SM Higgs boson are determined in this chapter. A given cross section is excluded if the underlying hypothesis is rejected with a confidence level of at least 95 %. To quantify the compatibility of the observed data with background and signal predictions, a profiled likelihood technique, outlined in Chapter 4, is used.

Systematic uncertainties are incorporated as nuisance parameters and considered as either channel-specific or common for all regarded categories. The latter are hereby taken as fully correlated and involve the uncertainty on the luminosity, energy scales and resolutions as well as uncertainties on the acceptance (e.g. trigger and ID uncertainties). For the limit setting procedure, each systematic uncertainty is treated as *normalization uncertainty* (Section 4.2), modeled by a Gaussian distribution as prior [131].

To estimate the impact of a specific uncertainty, the full analysis is reproduced for a variation of  $\pm 1$  standard deviation of the corresponding systematic uncertainty. For each category, the resulting shape deviations of the final observable are compared to its nominal shape. If they are not covered by the statistical uncertainty of the nominal distribution, the particular systematic uncertainty is considered in the limit setting procedure by additional degrees of freedom. These contain the full shape information for the final discriminant based on the  $\pm 1$  standard deviations of the systematic uncertainty [131]. In the other cases, the

impact on the acceptance within each defined category is determined and implemented in the limit setting procedure.

Exemplary systematic variations of the BDT analysis are shown in Appendix G. As outcome of this study, the shape information for the jet and  $\tau$  energy scale uncertainties as well as the uncertainty on the subtracted muon calorimeter depositions in the *embedding* procedure are taken into account. These are described in more detail together with all other sources of systematic uncertainties in the following.

The statistical uncertainties on the other hand are considered by Poissonian distributions, expressed for the initial amount of simulated events.

## 8.1. Systematic uncertainties on signal and background estimates

Every correction factor of the analyses is subject to systematic uncertainties. The correction factors are derived in auxiliary measurements and compensate for small differences between data and simulated events caused by the ATLAS detector response and various modeling parameters of the Monte Carlo generators. Systematic uncertainties are therefore assigned on simulation parameters, pile-up corrections (Section 5.1.4) and trigger efficiencies (Section 5.1.2). Also, the object identification and reconstruction efficiencies, energy resolutions and energy scales (outlined in Chapter 3) are subject to systematic uncertainties. Moreover, systematic uncertainties are evaluated to express the lack of precision on the different data-driven background estimation techniques within this analysis (Section 5.3).

### 8.1.1. QCD scale uncertainties on the signal production cross section

The cross sections of the inclusive Higgs boson production,  $\sigma_{TOT}$ , are known up to NNLO (Section 5.1.3) and listed in Ref. [37]. For the  $q\bar{q} \rightarrow H$  production modes<sup>1</sup>, an uncertainty due to individual variations of the renormalization and factorization scales in a range of  $\frac{m_H}{2} < \mu_R(\mu_F) < 2 \cdot m_H$  is estimated. The nominal values are chosen to be  $\mu_R = \mu_F = m_H$ . An uncertainty of  $\pm 2.4\%$  ( $\pm 1.2\%$ ) is assigned for events that end up in the VBF (non-VBF) categories [37].

Regarding the Higgs boson production via gluon fusion ( $gg \rightarrow H$ ), the differential cross sections accompanied by specific requirements on jet-related kinematics are sensitive to higher order QCD corrections and suffer from significant uncertainties. The LHC Higgs boson cross section working group presented in Ref. [132] that the scale uncertainty on an exclusive  $N$ -jet cross section,  $\sigma_N$ , is underestimated with respect to the theoretical resummed calculation. To

<sup>1</sup>These include the VBF, WH and ZH Higgs boson production modes.

assign the uncertainties correctly, they recommend to calculate inclusive  $N$ -jet cross sections,  $\sigma_{\geq N}$ . These are afterwards treated as uncorrelated and exclusive cross section uncertainties are calculated from the difference

$$\Delta_N^2 = \Delta_{\geq N}^2 - \Delta_{\geq N+1}^2. \quad (8.1)$$

To estimate the inclusive cross section uncertainties on the Higgs boson production via gluon fusion, the HNNLO [133] MC generator is employed for all  $gg \rightarrow H$  configurations except for the one that fulfills the VBF classification cuts,  $\sigma(gg \rightarrow H)_{VBF}$ . Since HNNLO is able to simulate the gluon fusion production mode at next-to-next-to-leading order, the underlying H+2 jets configuration of the VBF category, because of the two *tagging jets* (Section 1.4), can be calculated at leading-order only [15]. Due to the large higher order corrections on the gluon fusion production mode, it was intended to model this configuration at NLO. Thus, the  $\sigma(gg \rightarrow H)_{VBF}$  configuration is generated via the MCFM [117] generator in order to reduce the systematic uncertainties.

The agreement between the MCFM samples and the nominal ggF signal samples is discussed in Ref. [134]. The nominal ggF signal samples are generated by POWHEG plus PYTHIA (see Section 5.1.3) and re-weighted to the  $p_T$  spectrum of the Higgs boson obtained from NNLL+NNLO calculations of HQT. An agreement of both samples within  $\pm 10\%$  is observed and justifies the use of the errors estimated with MCFM (Table 8.1).

Table 8.1 outlines the theory uncertainties for the gluon fusion production mode used within the cut-based and multivariate approaches of this thesis. They are determined through the maximum spread in the cross sections when varying the renormalization and factorization scales independently in a range of  $\frac{m_H}{2} < \mu_R(\mu_F) < 2 \cdot m_H$ . The nominal values are chosen to be  $\mu_R = \mu_F = m_H$ .

Table 8.1.: Uncertainties of the Higgs boson production via gluon fusion resulting from individual variations of the renormalization and factorization scales (taken from Ref. [15]).

Generator $m_H$ [GeV]	HNNLO			MCFM
	$\Delta(\sigma_{TOT})$	$\Delta(\sigma_{\geq 1})$	$\Delta(\sigma_{Boosted})$	$\Delta(\sigma_{VBF})$
125	$\pm 10\%$	$\pm 20\%$	$\pm 20\%$	$\pm 23.5\%$

Since the VBF category relies significantly on forward jets, its acceptance is sensitive to variations of the underlying event (UE) model. An uncertainty of  $\pm 10\%$  ( $\pm 30\%$ ) due to the UE is estimated for the VBF (ggF) Higgs boson production modes.

All scale uncertainties on the signal production cross section are derived for the jet-related cuts of the cut-based analysis (see Chapter 6). This approach is applied to the BDT approach as well for two reasons. To obtain the impact of theoretical uncertainties on the BDT input variables and their correlation, they are required on reconstruction level which is not available for the scale variations within MCFM or HNNLO. Furthermore, the BDT score is in general correlated to jet-related kinematic properties (Fig. 7.11). Thus, a cut on a particular BDT value as applied in Section 7.5 would have to be considered.

A more rigid classification leads to larger estimated QCD scale uncertainties. This can be observed in a comparison between the primary cut-based analysis [86] of the 2011 dataset and the re-analysis presented in Chapter 6 [15]. Thus, a conservative estimate for the nominal uncertainties in the BDT approaches is chosen. However, any shape distortions that may arise from a BDT classifier are not taken into account in this approximation. The impact of shape distortions of the signal samples on the expected exclusion limit is discussed in Section 8.3.

### 8.1.2. Systematic uncertainties on Monte Carlo methods and luminosity

The processes simulated by various Monte Carlo generators are normalized to their specific cross section. To estimate the corresponding systematic uncertainties, the impact when varying different MC tuning parameters (see Section 5.1.3) is studied.

The uncertainties on the cross sections caused by uncertainties on the underlying PDFs and strong coupling constant  $\alpha_s$  are determined via the prescription given in Ref. [135]. Thus, an uncertainty of  $\pm 8\%$  is assigned to gluon-initiated processes, e.g. the Higgs boson production via gluon fusion. An uncertainty of  $\pm 4\%$  is estimated for quark-initiated processes. These include the  $q\bar{q} \rightarrow H$  production modes as well as the di-boson and  $Z \rightarrow \tau\tau$  background processes, which are not normalized to data in specific CRs.

The uncertainty on the luminosity measurement of the 2011 dataset recorded with the ATLAS detector is estimated as  $\pm 1.8\%$  [136]. As an example, Table 8.2 summarizes the different systematic uncertainties resulting from scale and PDF variations for the ALPGEN Z+Jets samples.

### 8.1.3. Trigger acceptance systematic uncertainties

Events passing the single lepton or the combined electron+tau trigger are used within the analyses (Table 5.1). The corresponding uncertainties on the lepton and  $\tau$  triggers are assumed to be uncorrelated. Efficiency scale factors as a function of  $\eta$  and  $\phi$  are determined to compensate for a different acceptance of data and simulated events. The scale factors of the

Table 8.2.: Systematic uncertainties for ALPGEN Z+jets MC samples for the 2011 dataset (taken from Ref. [15]). Thereby, the minimum requirements for the  $p_T$  and  $\Delta R$  of the final state partons are modified. Their default is set to  $p_T > 15$  GeV and  $\Delta R > 0.7$  within the ATLAS ALPGEN configuration [137]. The default scale definition is computed via the final state partons from the Z boson, only. In the case of *iqopt2*, the remaining final state partons are included as well. The MLM matching scheme is described in Section 5.1.3 and the kt factor modifies the scale at which a showering occurs within ALPGEN.

Source	Systematic Uncertainty
ALPGEN Z+jets MC	
Minimum final state parton $p_T > 12$ GeV	-0.21 %
Minimum final state parton $\Delta R > 0.4$	1.92 %
Different scale definition ( <i>iqopt2</i> )	0.83 %
Different PDF set (MRST2001J [138])	5.33 %
$Q^2$ Scale (double)	-0.12 %
$Q^2$ Scale (half)	0.13 %
MLM matching condition ( $p_T$ 15 GeV)	-10.6 %
MLM matching condition ( $\Delta R$ 0.4 )	3.36 %
kt factor (double)	0.30 %
kt factor (half)	-0.37 %

electron triggers used in the analysis are found to be within 0.958 and 1.002 with an uncertainty of roughly  $\pm 0.5$  % [121]. Moreover, muon trigger efficiency scale factors are measured in the  $Z \rightarrow \mu\mu$  process with an uncertainty of approximately  $\pm 1$  %, as reported in Ref. [139]. Finally, the scale factor uncertainties on the tau triggers are found to be within  $\pm 2$  % and  $\pm 4$  %, depending on the momentum of the  $\tau$  [119,140].

#### 8.1.4. Reconstruction and identification uncertainties

##### Electrons

For the electrons, systematic uncertainties on the calorimeter isolation correction, the ID and reconstruction efficiencies as well as the energy scale and resolution are considered. A systematic uncertainty of approximately  $\pm 3$  % is estimated for the ID and reconstruction efficiency, depending on the momentum of the electron [61]. Moreover, an uncertainty of  $\pm 2$  % ( $\pm 4$  %) is given for the corrections of the EM calorimeter isolation quantities. For the electron energy scale, an uncertainty of approximately  $\pm 1$  % is assigned on electrons, depending on their momentum and detector region [61]. Further, the energy resolution in simulated events has an uncertainty in a range of  $\pm 1$  % to  $\pm 1.8$  % [59].

## Muons

Regarding the muons, a shift of the energy scale is applied as described in Section 3.2. Systematic variations of both scales for the inner detector and muon spectrometer are considered and derived in different  $\eta$  regions. They have a relative uncertainty of roughly  $\pm 1\%$  [57]. In addition, systematic uncertainties for both identification and isolation of the muon are implemented, whereas the latter has an amount of  $\pm 1\%$  for muons with a  $p_T > 20$  GeV. The identification uncertainty on the other hand has values below  $\pm 1\%$  [55].

## Taus

As outlined in Section 3.7, a dedicated  $\tau$  energy scale (TES) is applied to all  $\tau_{\text{had}}$  candidates. The systematic uncertainties on the TES have an amount of  $\sim \pm 3\%$  for all measured bins which depend on the number of tracks in the inner cone of the  $\tau_{\text{had}}$  candidate, its energy and pseudorapidity [72]. Moreover, a  $\tau$  identification efficiency uncertainty of  $\pm 4\%$  and  $\pm 8\%$  is assigned to  $\tau_{\text{had}}$  candidates with  $p_T > 22$  GeV and  $20 \text{ GeV} < p_T < 22$  GeV, respectively [74].

## Jet properties

Jets have a significant impact on the event acceptance of the two most sensitive categories (Boosted and VBF) defined in the cut-based and multivariate approaches. Systematic uncertainties on all scale calibrations of jets based on LCW topo-clusters are derived in Section 3.5.1. This includes MC-based scale factors as well as the residual *in-situ* calibration of the jet energy scale. In addition, several effects such as pile-up [65], differences in the jet flavor composition, different simulation conditions and close-by jets are addressed [67]. The combination of all effects adds up to a systematic uncertainty between  $\pm 1\%$  and  $\pm 7\%$  that is estimated in several bins of the  $p_T$  and  $\eta$  of the jet [67]. Variations of the jet energy scale are applied to all jets that did not overlap with the identified electrons, muons or hadronic  $\tau$ s.

A supplemental measurement is performed to derive the *b-tagging* uncertainties for the different algorithms and working points described in Section 3.6. These are required whenever a veto on b-jets is implemented in the analysis categorization (see Table 7.2). The uncertainties are based on the 2011 dataset and are determined in  $t\bar{t}$  events. For the *JetFitterCombNN* b-tagger used in this thesis, an overall uncertainty within  $\pm 5\%$  to  $\pm 15\%$ , depending on the jet momenta [69, 70], is applied on top of the jet scale uncertainties.

### Missing transverse energy

To estimate uncertainties on  $E_T^{\text{miss}}$ , which is calculated from all reconstructed objects in the event, the corresponding systematic uncertainties of each object has to be taken into account. Thereby, scale uncertainties on the clusters assigned to  $\tau$ s or jets are treated as fully correlated. Besides, uncertainties on the remaining terms of the  $E_T^{\text{miss}}$  calculation,  $E_T^{\text{miss},\text{soft jets}}$  and  $E_T^{\text{miss},\text{cell-out}}$ , have to be considered (see Eq. 3.8). This is achieved by dedicated MC simulations including JES variations, different shower and UE models. A combination of these results in a total uncertainty of  $\sim \pm 10\%$  on the  $E_T^{\text{miss},\text{soft jets}}$  and  $\sim \pm 13\%$  on the  $E_T^{\text{miss},\text{cell-out}}$  term [75].

For the overall systematic uncertainty on  $E_T^{\text{miss}}$ , the terms for each object and the residual  $E_T^{\text{miss},\text{soft jets}}$  and  $E_T^{\text{miss},\text{cell-out}}$  terms are treated individually. Their impact is measured in  $W \rightarrow e\nu$  and  $W \rightarrow \mu\nu$  events. This results in an average uncertainty of approximately  $\pm 2.6\%$ , rising with larger  $\sum E_T^{\text{miss}}$  [75].

#### 8.1.5. Uncertainties on the background estimation

Apart from systematic uncertainties on theory and MC predictions, event triggers and the various objects, systematic uncertainties are assigned to the background estimation techniques described in Section 5.3. The corresponding uncertainties are summarized in Table 8.3.

Table 8.3.: Summary of systematic uncertainties on the background estimation methods described in Section 5.3. The systematic uncertainties for the W+jets background are derived individually for each category and for each systematic uncertainty. As an example, the quoted numbers refer to the BDT approach and the 0J category in case of *non-VBF*. They are displayed as  $\Delta k_W^{\text{OS}}/\Delta k_W^{\text{SS}}$ .

Source	Assigned uncertainty		
	non-VBF		VBF
	$e\mathcal{T}_{\text{had}}$	$\mu\mathcal{T}_{\text{had}}$	$e\mathcal{T}_{\text{had}} + \mu\mathcal{T}_{\text{had}}$
$r_{\text{QCD}}$	$\pm 5.1\%$	$\pm 5.6\%$	$\pm 5.1\% / \pm 5.6\%$
$k_{Z \rightarrow \ell\ell(\rightarrow\tau)}$	$\pm 24\%$	-	$\pm 24\% / -\%$
$k_{Z \rightarrow \ell\ell + \text{jet}(\rightarrow\tau)}$	$\pm 5.4\%$	$\pm 5.4\%$	$\pm 34\%$
$k_{\text{top}}^{\text{OS}}$	$\pm 10\%$		
$k_{\text{top}}^{\text{SS}}$	$\pm 22\%$		
VBF-filtered Z samples	-		$\pm 5\%$
Fake-factor method	-		$\pm 50\%$
$k_W^{\text{OS,SS}}$	$\pm 1.3\% / \pm 2.1\%$	$\pm 1.4\% / \pm 2.8\%$	$\pm 14.3\% / \pm 13.7\%$

Apart from these uncertainties, a production cross section uncertainty of  $\pm 5\%$  and  $\pm 7\%$

[119] is considered for the di-boson and  $Z \rightarrow \tau\tau$  samples, respectively, which are not normalized to data in specific CRs.

### $Z \rightarrow \tau\tau + \text{jets}$ embedding uncertainties

In the embedding technique outlined in Section 5.3.7, muon tracks and their calorimeter energy depositions are removed from data events. They are replaced by simulated  $\tau$ s. To estimate uncertainties on the embedding procedure itself, both track identification requirements and the subtracted calorimeter energies are varied.

Hence, the track isolation criteria are dropped or tightened and the simulated calorimeter depositions are scaled by  $\pm 30\%$  as a conservative estimate in order to take non-considered effects into account [119].

## 8.2. Exclusion limits of the cut-based analysis

Based on the  $m_{\tau\tau}^{\text{MMC}}$  distributions of the cut-based analysis (Fig. 6.7), upper limits are derived on the signal production cross section at 95 % CL (further referred to as exclusion limit) as a function of  $m_H$ . Figure 8.1 shows the final limit for all categories.

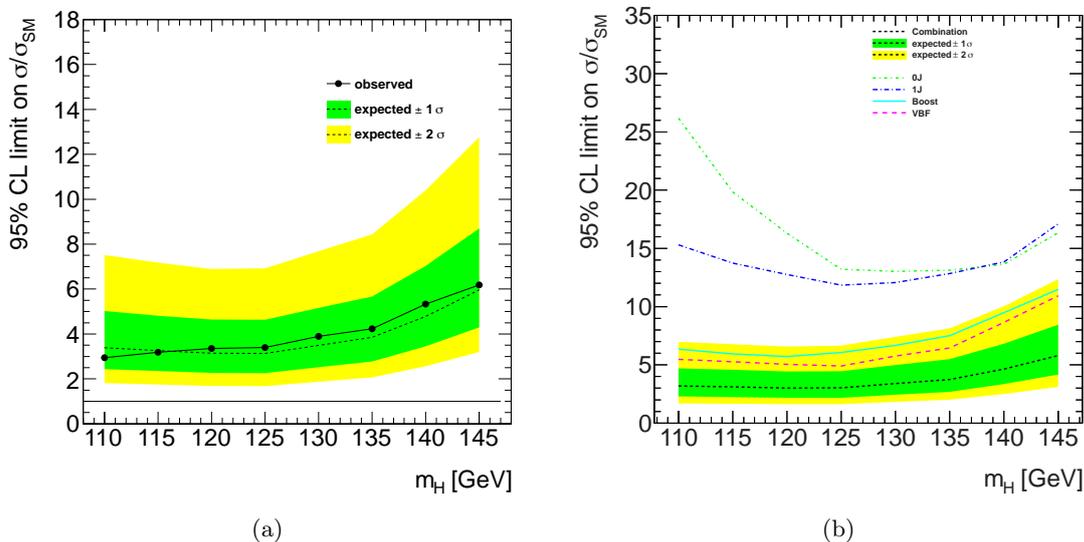


Figure 8.1.: Exclusion limits as a function of the Higgs boson mass in the range between 110 GeV and 145 GeV. They illustrate the excluded signal cross section at 95 % confidence level, normalized to the SM cross section  $\sigma_{\text{SM}}$ . The left plot shows the combined expected (dashed line) and observed (solid line) limits for all categories of the cut-based analysis. The right plot displays the expected exclusion limits for the different categories separately as well as their combination. In both plots, the green and yellow bands demonstrate the one and two standard deviations of the expected limit (see Section 4.5).

At  $m_H = 125$  GeV, the expected exclusion limit for all categories combined has a median value of  $3.08 \cdot \sigma_{\text{SM}}$ , whereas the observed limit excludes 3.45 times the SM Higgs boson cross section.

As described in Sections 6.1 and 6.2, the Boosted and VBF categories are the most sensitive ones and contribute significantly to the final limit (see Fig. 8.1(b)). Moreover, the 0J channel yields the lowest sensitivity due to its bad signal-to-background ratio.

Tables F.1 to F.6 present the corresponding systematic uncertainties. They are classified according to Section 8.1 and split into the different categories.

### 8.3. Exclusion limits of the BDT analyses

For the BDT analysis presented in Section 7.4, the BDT score distributions of Fig. 7.8 are exploited to set an upper exclusion limit on the Higgs boson production cross section. Alternatively, a BDT was trained without the reconstructed mass of the di- $\tau$  system (denoted *BDT-cut*). Thereby, a cut was applied on the BDT score to obtain the  $m_{\tau\tau}^{\text{MMC}}$  distributions for all events above a specific threshold. In this case, the exclusion limit is computed on the corresponding  $m_{\tau\tau}^{\text{MMC}}$  distributions given in Fig. 7.13. Figure 8.2 depicts the limits for both BDT approaches as a function of  $m_H$ .

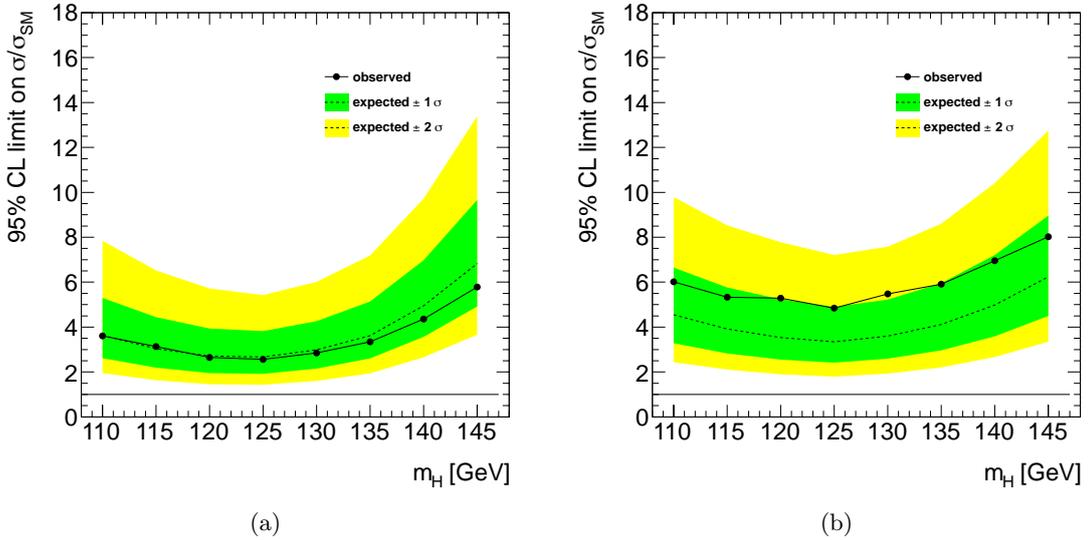


Figure 8.2.: Expected (dashed line) and observed (solid line) exclusion limits for the BDT approaches and a Higgs boson mass in the range between 110 GeV and 145 GeV. (a) The limits for a BDT classifier with  $m_{\tau\tau}^{\text{MMC}}$  in the training, where the BDT score is taken as observable. (b) The BDT approach without  $m_{\tau\tau}^{\text{MMC}}$  in the training, followed by a limit setting based on the distributions of the reconstructed mass of the di- $\tau$  system after a cut on the BDT score.

An expected (observed) limit of  $2.66 \cdot \sigma_{\text{SM}}$  ( $2.56 \cdot \sigma_{\text{SM}}$ ) is obtained for the fit on the BDT score distributions at  $m_H = 125$  GeV, whereby the *BDT-cut* approach yields an expected (observed) limit of  $3.33 \cdot \sigma_{\text{SM}}$  ( $4.80 \cdot \sigma_{\text{SM}}$ ) at  $m_H = 125$  GeV. The associated systematic uncertainties are summarized in Appendix G.

In Section 8.1.1, the treatment of theoretical uncertainties on the signal production cross section based on the cut-based categorization was mentioned. To analyze the sensitivity of the exclusion limit on these, an exemplary limit of the former BDT approach with doubled theoretical uncertainties was computed. At  $m_H = 125$  GeV, this led to an expected exclusion limit of  $2.93 \cdot \sigma_{\text{SM}}$ , which corresponds to a degradation of roughly 10 % regarding the nominal result of the BDT approach.

As a second approach, a systematic shape uncertainty was introduced to the signal and implemented as additional nuisance parameter within the limit computation. In this case, the remaining theory uncertainties were taken at their nominal values. To modify the signal shape, its distribution was multiplied with a linear function of arbitrary slope and normalized to the nominal signal yield at the end. To determine the actual shape uncertainty, the variation of the rightmost signal bin with respect to its nominal yield was examined for every category. The variation pictured in Fig. 8.3(a) corresponds to a shape deviation of  $\pm 60$  %.

Finally, the upper expected exclusion limit was computed for distinct steps between  $\pm 0$  % and  $\pm 60$  %, taking all systematic and statistical uncertainties into account. The impact on the upper expected Higgs boson production cross section at 95 % CL and  $m_H = 125$  GeV is illustrated in Fig. 8.3.

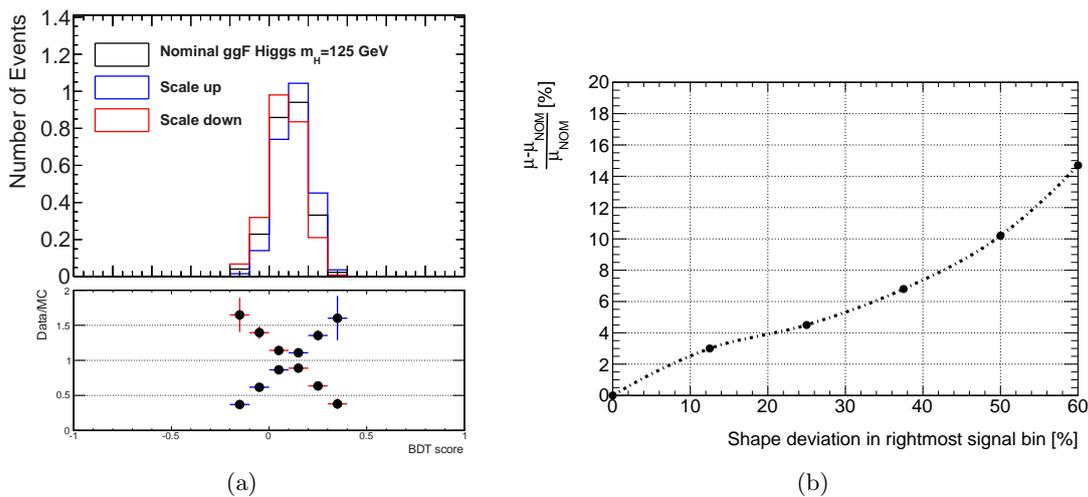


Figure 8.3.: Shape variation of the BDT score distribution after applying a linear shape distortion (a) and the resulting effect as a function of the shape variation of the rightmost signal bin on  $\mu$  (b), which is the fitted upper limit on the production cross section, at  $m_H = 125$  GeV.

The outermost point of  $\pm 60\%$  uncertainty on the signal shape corresponds to an expected exclusion limit of  $3.05 \cdot \sigma_{\text{SM}}$ . Additional studies are necessary to estimate the theory uncertainties and their dependence on the BDT score distributions. For example, the impact of QCD scale variations on various jet-related kinematic quantities can be estimated within the HNNLO and MCFM generators and applied on the input variables of the BDT. Moreover, samples with modified renormalization and factorization scales can be passed through a full simulation of the ATLAS detector in order to study their impact on the final BDT distributions. However, the expected exclusion limit at  $m_H = 125$  GeV has an increased significance with respect to the cut-based analysis, even after assumption of artificially high uncertainties.

## 8.4. Comparison between cut-based and BDT approaches

To summarize the results of the previous sections, no significant excess of data events over the background predictions is observed. The observed exclusion limits reside within one standard deviation of the expected exclusion limit.

Figure 8.4 compares the median expected exclusion limits of the two BDT approaches (see Chapter 7) with respect to the cut-based approach (Chapter 6) for distinct categories. An overall improvement of the BDT technique can be observed within the VBF category. It should be noted that the OS-SS method is used to model the fake- $\tau$  background contribution within the BDT approaches. However, the fake factor method would allow to model the background coming from mis-identified  $\tau$ s more precisely and with a lower statistical uncertainty (see Section 5.3.8). This could lead to substantial differences in the significant region of high BDT scores, where only few background events are expected. In order to achieve a reliable fake factor estimation, the dependence of the fake factor on the final BDT score distributions would have to be measured. This study was not performed within this thesis and could be considered in further analyses.

In the non-VBF categories, the cut-based exclusion limit resides in between the two BDT approaches. The *BDT-cut* technique displays a higher exclusion limit than the cut-based analysis at all considered Higgs boson masses. However, the modified event categorization (Table 7.2) should be emphasized in this case. Especially, the VBF and Boosted category are defined on a weaker set of cuts, followed by a rejection of all events with additional jets that fail both classification cuts.

Thus, the cut-based analysis has an extra 1J category with a large number of events that did not get a high transverse boost nor possess two characteristic *tagging jets*. Moreover, no information about  $m_{\tau\tau}^{\text{MMC}}$  is given in the BDT training, which may lead to an imperfect

signal significance in terms of the  $m_{\tau\tau}^{\text{MMC}}$  distributions.

The *BDT-cut* approach disposes a large number of events, including signal, to optimize the exclusion limit. A recycling of all events that fail the final BDT cut threshold would restore the otherwise lost signal events. In addition, they might be combined with all events that fail the VBF and Boosted categorization in the beginning, setting up a new category to enhance the signal extraction. This could be part of further studies.

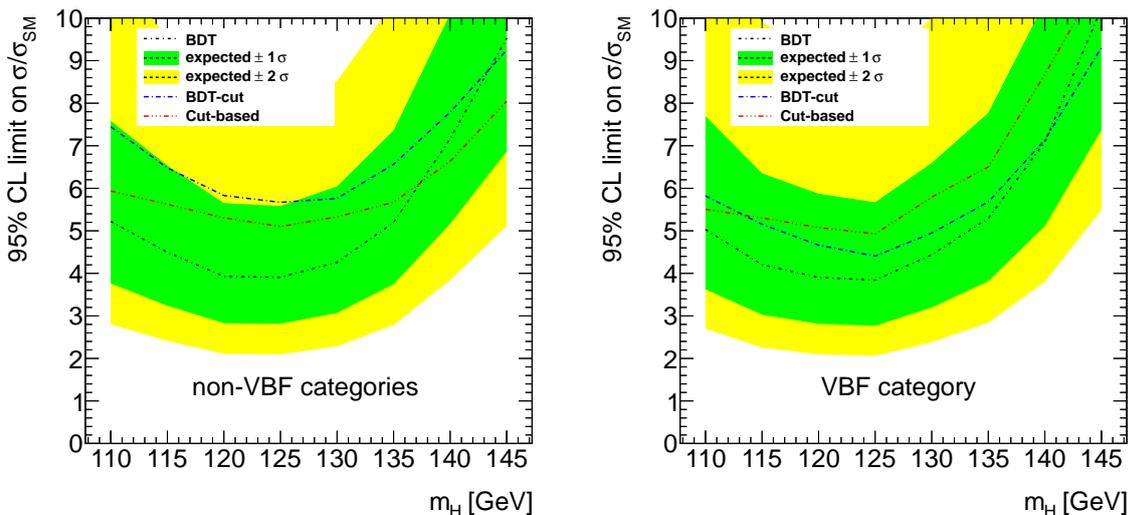


Figure 8.4.: Expected exclusion limits at 95 % CL and normalized to the SM cross section  $\sigma_{SM}$  for the BDT trained with variables including  $m_{\tau\tau}^{\text{MMC}}$  (black dashed line), the *BDT-cut* approach (blue dashed line) and the cut-based approach (red dashed line). They are computed for all corresponding categories combined and shown as a function of the Higgs boson mass in the range between 110 GeV and 145 GeV. The left plot shows all categories except VBF, whereas the right plot shows a distinct comparison of the expected exclusion limits obtained for the VBF category. In both cases, the green and yellow bands illustrate the one and two  $\sigma$  variation on the BDT exclusion limit.

As a subtlety, the BDT approaches are more sensitive to the probed Higgs boson mass. This reflects that the underlying BDT classifier was trained with a specific Higgs boson mass of  $m_H = 125$  GeV. Since a Higgs-like boson was observed in this mass range, the choice of input variables and event samples was optimized with respect to this particular mass. A straight-forward expansion of the BDT training on all Higgs boson masses can cause significant deviations in the classifier. This may alter the agreement between data and simulated events and lead to fluctuations in the exclusion limit. Hence, the performance of one technique with regard to another is quantified at  $m_H = 125$  GeV.

Finally, Figure 8.5 depicts the combined expected exclusion limits for the analyzed approaches. The *BDT-cut* approach suffers from the drawbacks mentioned above. A degrada-

tion of roughly 8 % on the upper exclusion limit at  $m_H = 125$  GeV, with respect to the cut-based analysis published in Ref. [15], is observed. However, the *BDT-cut* limit distribution exhibits the same trend as the cut-based analysis at Higgs boson masses of  $m_H \geq 125$  GeV. This reflects the separation between signal and background events reached in the reconstructed mass of the di- $\tau$  system. For a large assumed Higgs boson mass, the irreducible  $Z \rightarrow \tau\tau$  background can be clearly separated within the MMC algorithm.

In contrary, the BDT approach that was optimized within this thesis and incorporated  $m_{\tau\tau}^{\text{MMC}}$  in its training procedure results in an upper exclusion limit which is approximately 14 % better than the cut-based benchmark. This illustrates the effectiveness of a Boosted Decision Tree, which takes a compound of input variables and their correlations into consideration to maximize the signal sensitivity. Through *cross evaluation* (Section 7.3.2), the major drawback of a BDT, namely the rejection of the training sample, can be bypassed. However, a sufficient number of simulated and observed events is necessary in the signal region, since statistical fluctuations can diminish the advantages of a MVA analysis.

To summarize, the BDT algorithm is able to significantly improve the upper exclusion limits on the SM Higgs boson production cross section in the  $\tau\tau$  final state at the ATLAS experiment. Nevertheless, more elaborate studies have to be carried out in order to obtain conclusive results.

Based on the results of this thesis, a BDT is currently established as baseline analysis within the Higgs physics working group of the ATLAS experiment. Thereby, several approaches for the training and signal extraction are pursued. In addition to the analyses presented in this thesis, a two-dimensional fit between the BDT output and  $m_{\tau\tau}^{\text{MMC}}$  is tested. Thereby, the BDT is trained without  $m_{\tau\tau}^{\text{MMC}}$ . Moreover, optimization studies based on the expected exclusion limits are carried out to obtain an optimal BDT configuration and binning of the final discriminant. Also, more sophisticated quantities, exploiting the angular structure and kinematic properties of the Higgs boson decay products, are analyzed. To provide a more precise background estimation in the regions with a high signal significance, a filter technique similar to the VBF-filtered samples (Section 5.3.7) is applied to the W+jets sample. For both W+jets and  $Z \rightarrow \tau\tau$  backgrounds, several filtered samples for different restrictions of the phase space are produced. In addition, the fake factor method (Section 5.3.8) with a BDT-dependent fake factor is applied for the Boosted and VBF categories.

To conclude, a preliminary answer on the existence of a leptonic coupling may be established for the recently observed Higgs-like boson with the increased dataset of the 2012 data taking period. Official results of the ATLAS collaboration based on the full 2011 and 2012 datasets are anticipated for the Summer 2013 conferences.

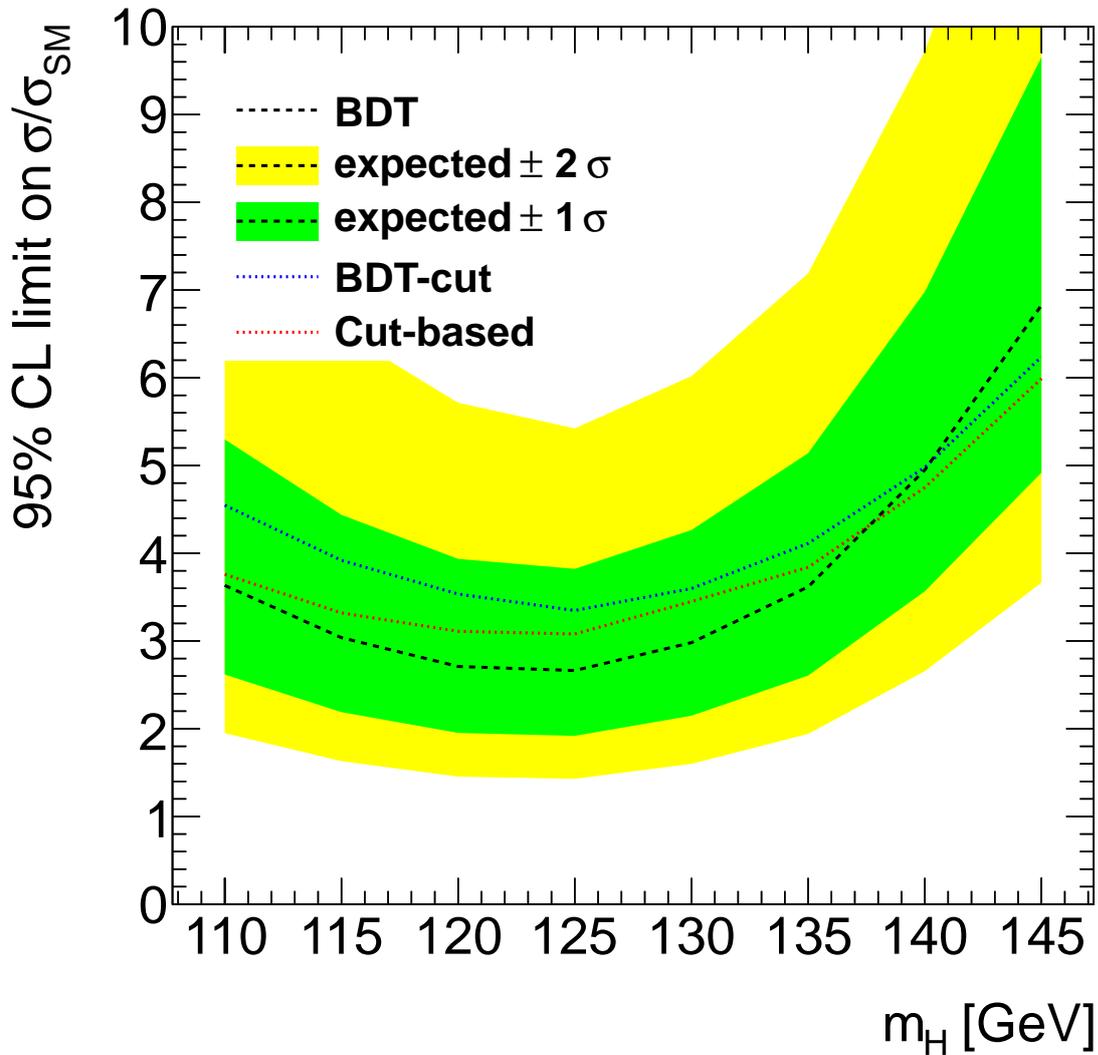


Figure 8.5.: Expected exclusion limits as a function of the Higgs boson mass in the range between 110 GeV and 145 GeV. They display the excluded Higgs boson production cross section at 95 % confidence level, normalized to the SM cross section  $\sigma_{SM}$ , for the BDT trained with variables including  $m_{\tau\tau}^{MMC}$  (black dashed line), the *BDT-cut* approach (blue dashed line) and the cut-based approach (red dashed line). They are computed for all corresponding categories combined. The green and yellow bands demonstrate the one and two  $\sigma$  variation of the BDT approach.

---

A search for the Standard Model Higgs boson in the  $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$  decay channel has been presented. It is based on data corresponding to an integrated luminosity of  $\mathcal{L} = 4.6 \text{ fb}^{-1}$ . The data was recorded with the ATLAS detector in 2011 at a center-of-mass energy of  $\sqrt{s} = 7 \text{ TeV}$ . The Standard Model of particle physics adheres our current knowledge of elementary particles together with their interactions. However, a spontaneous symmetry breaking is required in order to include fermion and boson masses into the SM framework.

It is achieved by the *Higgs mechanism*, which introduces a massive scalar boson, the *Higgs boson*. This couples to the mass of other particles and can be solely detected via its decay products. The ATLAS experiment, installed at the LHC, exploits several decay modes to establish the existence and measure the intrinsic properties of the Higgs boson. In July 2012, a yet unknown particle was observed in three bosonic decays,  $\gamma\gamma$ ,  $WW$  and  $ZZ$ , with a mass of approximately 125 GeV. To date, the measured signal strengths of this resonance within the three channels is compatible with the expectations from a SM Higgs boson. Also, the spin and parity states examined for this Higgs-like resonance favor the SM assignment of  $J^P = 0^+$ .

The  $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$  analysis concentrates on a neutral final state with one  $\tau$  lepton decaying into an electron or muon while the other  $\tau$  lepton decays hadronically. Both are accompanied by neutrinos that escape detection and lead to missing transverse energy. Thus, the exact  $\tau$  momenta and directions cannot be reconstructed, which introduces a challenge in the mass reconstruction of  $m_{\tau\tau}$  faced by the missing mass calculator (MMC) algorithm.

The  $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$  analysis has to cope with the irreducible  $Z \rightarrow \tau\tau$  process featuring a much larger production cross section. Moreover, events with a mis-identified  $\tau_{\text{had}}$  candidate constitute an important background. To obtain a precise estimation of those backgrounds and diminish uncertainties arising from jet and  $\tau$ -fake mis-modelings in simulation, data-driven methods are employed. For  $Z \rightarrow \tau\tau$ , the *embedding* algorithm was developed. It inserts simulated  $\tau$  decays into Z candidates selected from  $Z \rightarrow \mu\mu$  events in data. To model contributions involving a mis-identified  $\tau_{\text{had}}$ , a control region with  $\tau$  candidates with the same electric charge (same-sign  $\tau$ s) is defined. Residual backgrounds from  $W$ +jets,  $Z \rightarrow \ell\ell$ +jets

and  $t\bar{t}$  events are normalized to data in specific control regions.

The separation in characteristic event topologies is achieved by a classification based on the jet-multiplicity of the final state and kinematic properties of the reconstructed  $\tau\tau$  system. In order to enhance the Higgs boson production via vector-boson fusion (VBF), two *tagging jets* with a large transverse momentum and high spatial separation in  $\eta$  are required. Furthermore, a category is defined for events with a sufficient transverse boost of the  $\tau\tau$  system in order to benefit from a higher  $E_T^{\text{miss}}$  resolution and better mass reconstruction ability.

The work related to this thesis had a twofold purpose. For one, an optimization study was carried out on the VBF category defined in the primary cut-based analysis of the 2011 dataset. A cut-based approach utilizes subsequent cuts in order to suppress background events and improve the separation between signal and background predictions. To enhance the signal significance within the VBF topology, a four-dimensional grid scan was performed. Therefore, the thresholds on the transverse momenta of the *tagging jets*, their opening angle in  $\eta$  and invariant mass were varied. To express the sensitivity at each grid point, an upper limit on the Higgs boson production cross section at a CL of 95 % was derived. Thereby, the most important systematic uncertainties for the jet and  $\tau$  energy scale as well as the statistical uncertainties for the events obtained from data and simulation were taken into account. The outcome of this study indicated a more rigid set of cuts for the VBF topology.

Finally, the compatibility between observed data and the SM expectations was evaluated within each category defined in the re-analysis of the 2011 dataset. It utilized a profiled likelihood approach based on the corresponding  $m_{\tau\tau}^{\text{MMC}}$  distributions. A combination of all categories yielded an upper expected (observed) exclusion limit on the Higgs boson production cross section of  $3.08 \cdot \sigma_{\text{SM}}$  ( $3.45 \cdot \sigma_{\text{SM}}$ ) at 95 % CL and a Higgs boson mass of  $m_H = 125$  GeV.

The main focus of this thesis rested on the formation of a multivariate analysis. Therefore, boosted decision trees (BDTs) were used to separate signal-like and background-like events according to predetermined categories. The classification was carried out through a looser set of cuts with respect to the cut-based approach. This preserved a larger number of events and ensured a high efficiency of the BDT algorithm.

A BDT maps the multi-dimensional phase space spanned by given input variables to assign a score on an event depending on its topological properties. A tree is built by successive division of a given training sample that represents the signal region. Eventually, every event ends up in a leaf and is labeled as either signal- or background-like. To enhance the performance of the BDT, a successive collection of single trees is trained. Thereby, each tree assigns a higher weight to mis-classified events of the previous tree. The so-called *adaptive boosting* was employed and leads to a single and powerful classifier.

To obtain a high signal significance, several training setups were examined and their performance was evaluated based on the expected exclusion limit. This involved the optimization of the BDT configuration, the composition of input variables and different estimations of the irreducible  $Z \rightarrow \tau\tau$  background. Thereby, the BDT score distributions served as final discriminant. An expected (observed) upper limit on the Higgs boson production cross section of  $2.66 \cdot \sigma_{\text{SM}}$  ( $2.56 \cdot \sigma_{\text{SM}}$ ) was derived at 95 % CL and  $m_H = 125$  GeV.

In this approach, a possible Higgs boson cannot be observed by an increase of observed data events in the proximity of a specific mass of the di- $\tau$  system. To resolve this drawback, the  $m_{\tau\tau}^{\text{MMC}}$  distributions were restored as final observable. Hence, another BDT was trained, this time omitting the former considered  $m_{\tau\tau}^{\text{MMC}}$  as input variable in order to diminish possible correlations between both. The reconstructed mass of the di- $\tau$  system was used as final discriminant, however, in order to achieve a better signal-to-background ratio, a cut on the BDT score was applied.

This independent approach led to an expected (observed) upper limit of  $3.33 \cdot \sigma_{\text{SM}}$  ( $4.80 \cdot \sigma_{\text{SM}}$ ) at  $m_H = 125$  GeV. Like in the cut-based analysis, no significant excess of data above the SM background predictions was observed for both examined BDTs.

Finally, the expected exclusion limits at  $m_H = 125$  GeV were compared as measure of sensitivity of the analyses. By this criterion, the former BDT approach resulted in a significance gain of approximately 13 % with respect to the cut-based analysis. However, the BDT classifier constructed without  $m_{\tau\tau}^{\text{MMC}}$  shows a sensitivity that is roughly 8 % lower than the one obtained in the cut-based analysis on the same measure. This can be explained by the fact that the cut thresholds on the BDT score rejected a lot of signal events that were not absorbed in separate categories.

To find a concluding answer on the existence of the  $\tau\tau$  decay mode of the long-sought Higgs boson, an increased dataset compared to the one studied in this thesis is necessary. Based on the studies carried out in this thesis, a BDT analysis is currently established in the Higgs physics working group of the ATLAS experiment. The employed categories and background estimation techniques are comparable to those used in this thesis. It is anticipated that official ATLAS results based on the full 2011 and 2012 datasets are presented at conferences in Summer 2013. The dataset corresponds to an integrated luminosity of  $4.6 \text{ fb}^{-1}$  and  $20.3 \text{ fb}^{-1}$  at center-of-mass energies of  $\sqrt{s} = 7$  TeV and  $\sqrt{s} = 8$  TeV, respectively.

With the improved sensitivity of multivariate techniques, the important question whether the observed new boson couples to leptons, as predicted in the Standard Model, can be addressed. The BDT offers the possibility to maximize the signal sensitivity for a given dataset. Hence, it provides the baseline analysis method of this substantial investigation.



# Appendices



---

# A

# Measures of sensitivity

---

In order to optimize the sensitivity of an analysis, the figure of merit plays an important role. Since the final exclusion limits are computed via a profiled likelihood ratio as derived in Section 4.2, an optimization should be based on simplifications of it.

Assuming a counting experiment (e.g. one bin) with known background and an absence of systematics, the likelihood  $\mathcal{L}(\mu)$  corresponding to a certain signal strength  $\mu$  is reduced to

$$\mathcal{L}(\mu) = \frac{(\mu s + b)^n}{n!} \exp^{-(\mu s + b)}. \quad (\text{A.1})$$

Here,  $s$  and  $b$  represent the signal and background expectations given by Monte Carlo simulations and  $n$  the observed data. Taking the test statistics for discovery ( $q_0$ , see Eq. 4.11) and assuming that our signal process is responsible for any additional observed events ( $\hat{\mu}s = n - b$ ),  $q_0$  is given by

$$q_0 = -2 \ln \frac{L(0)}{L(\hat{\mu})} = -2 \ln \frac{(b)^n n! \exp^{-b}}{n! (\hat{\mu}s + b)^n \exp^{-(\hat{\mu}s + b)}} = 2(n \times \ln \frac{n}{b} + b - n). \quad (\text{A.2})$$

Using Wilks' theorem [81] (see 4.24) and the Asimov data set for the expected significance of a nominal signal process ( $\mu = 1$ ), thus replacing  $n$  by  $s + b$ , the significance yields

$$\text{med}[Z_0|1] = \sqrt{q_{0,A}} = \sqrt{2((s + b) \ln(1 + \frac{s}{b}) - s)}. \quad (\text{A.3})$$

$\text{med}[Z_0|1]$  represents the median Z-value for the background only hypothesis ( $\mu = 0$ ) with a nominal signal ( $\mu = 1$ ) in pseudo-data. An often used figure of merit is derived by the Taylor expansion of  $s \ll b$ , leading to

$$\text{med}[Z_0|1] = \frac{s}{\sqrt{b}}(1 + \mathcal{O}(s/b)). \quad (\text{A.4})$$

An intuitive explanation of this expression is that the signal quantity is divided by the statistical uncertainty of the observed data. An expansion of this simple model can be constructed by an auxiliary measurement in a background-only environment. This has m

observed events and a mean of  $\tau \times b$  expected events. Taking the scale factor  $\tau$  as known with negligible uncertainty, this yields

$$\frac{(\tau b)^m}{m!} \exp^{-\tau b}. \quad (\text{A.5})$$

The term is multiplied to Eq. A.1 to achieve the full likelihood function. Recapitulating the steps from above with the Asimov data set for the nominal signal process ( $s = n + b$ ) and the mean background prediction  $m = \tau b$  for the supplementary measurement provides

$$\begin{aligned} \text{med}[Z_0|1] = & \left[ 2 \left( (s + b) \ln \left[ \frac{(s + b)(b + \sigma_b^2)}{b^2 + (s + b)\sigma_b^2} \right] - \frac{b^2}{\sigma_b^2} \ln \left[ 1 + \frac{\sigma_b^2 s}{b(b + \sigma_b^2)} \right] \right) \right]^{\frac{1}{2}} = \\ & \frac{s}{\sqrt{b + \sigma_b^2}} (1 + \mathcal{O}(s/b) + \mathcal{O}(\sigma_b^2/b)). \end{aligned} \quad (\text{A.6})$$

The Taylor expansion was hereby performed in terms of  $s/b$  and  $\sigma_b^2/b$ .

The uncertainty of the background, using the control measurement  $m$ , is represented by  $\sigma_b^2$ . The estimator for  $b$  is  $\hat{b} = \frac{m}{\tau}$  and the variance for  $m$  is equal to its median  $\tau b$  in a Poissonian distribution, therefore

$$V[\hat{b}] \equiv \sigma_b^2 = \frac{b}{\tau}. \quad (\text{A.7})$$

The separation power between background and signal in a discriminating variable with  $n$  bins can be quantified by considering the bins as uncorrelated, thus computing the squared sum of one of the formulas stated above, e.g.

$$\text{med}[Z_0|1] = \sqrt{\sum_{i=0}^n \text{med}[Z_0|1]_i^2}. \quad (\text{A.8})$$

---

## B Scheme of the $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$ object selection

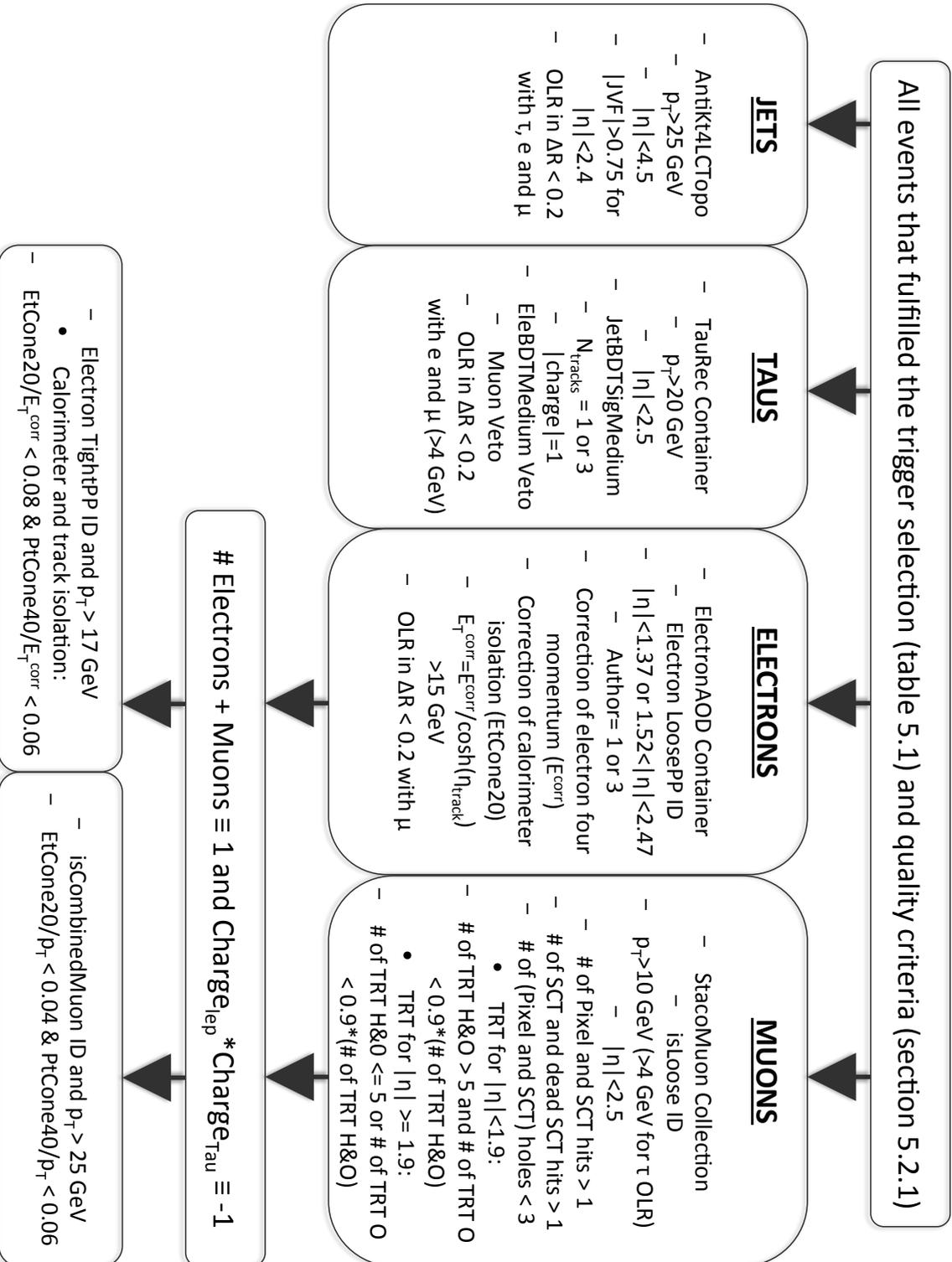
---

Figure B.1 depicts the object selection strategy used as foundation for both cut-based and multivariate approach. It shows the particular cuts applied to jets, taus, electrons and muons in order to obtain the final objects. In addition, the characteristic di-lepton veto to obtain  $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$  events only and reject  $Z \rightarrow \ell\ell + jets$  background events, performed with a loose lepton preselection, is illustrated. The ID working points and the definitions of the various selection quantities are given in Chapter 3.

Moreover, the smearing processes applied to the leptons is outlined in Sections 3.2 and 3.3. The jet and muon containers used for the  $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$  analysis are chosen of a set of different reconstruction algorithms described in Sections 3.2 and 3.5.1, respectively.

Finally, several muon selection cuts are based on hits in the inner detector, which is described in detail in Section 2.2.1.

Figure B.1.: The object preselection after trigger decisions (see Table 5.1) and event quality criteria cuts (see Section 5.2.1), followed by the di-lepton veto of electron and muon. If exactly one lepton is present after preselection, the final lepton selection is applied. OLR stands for overlap removal, and the H and O in the TRT selection of muons denote hits and outliers, respectively.



---

## **C** **Input variables applicable in the multivariate approach**

---

Prior to the optimization of the selection of input variables, their agreement between data and MC simulated events has to be confirmed. Therefore, all variables enlisted in Table 7.1 are plotted according to the categories defined in Table 7.2. None of the variables display a major mis-modeling caused by either an overall bias or a slope within the distribution.

The input variables of the 0J and Boosted category are depicted in Figs. C.1 to C.4 and Figs. C.5 to C.6, respectively. Finally, the variables of the VBF category, involving the same variables as used in the 0J and Boosted category, but with additional variables to exploit the characteristic VBF topology are shown in Figs. C.7 to C.9.

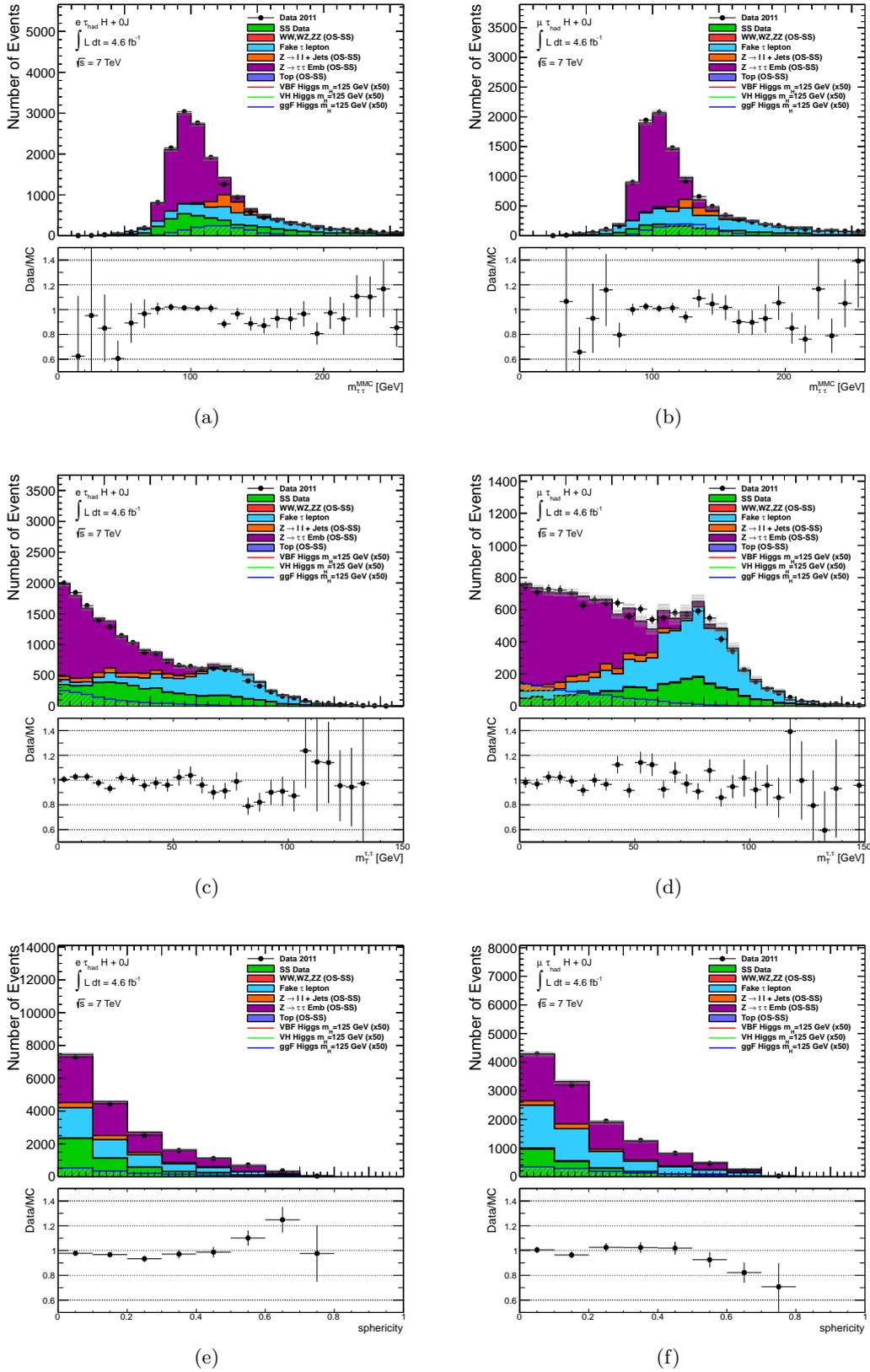


Figure C.1.: The distributions of  $m_{\tau\tau}^{\text{MMC}}$  (a,b),  $m_T$  (c,d) and sphericity (e,f) for the 0J category in the  $e + \tau_{\text{had}}$  (left column) and  $\mu + \tau_{\text{had}}$  (right column) channels. Each distribution is accompanied by its corresponding data-to-MC ratio, including the statistical uncertainties of both.

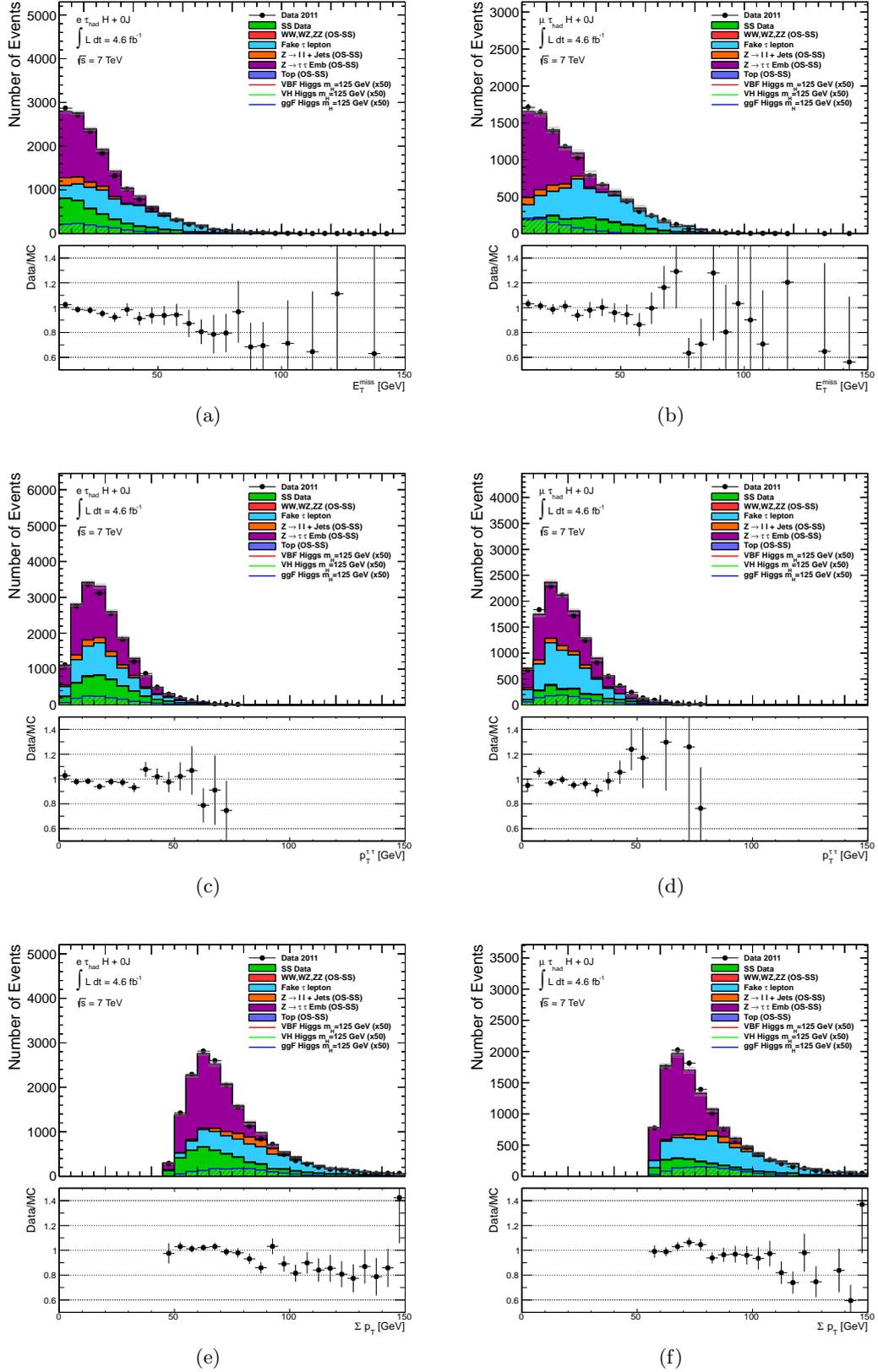


Figure C.2.: The distributions of  $E_T^{\text{miss}}$  (a,b),  $p_T^H$  (c,d) and  $\sum p_T$  (e,f) for the 0J category in the  $e + \tau_{\text{had}}$  (left column) and  $\mu + \tau_{\text{had}}$  (right column) channels. Each distribution is accompanied by its corresponding data-to-MC ratio, including the statistical uncertainties of both.

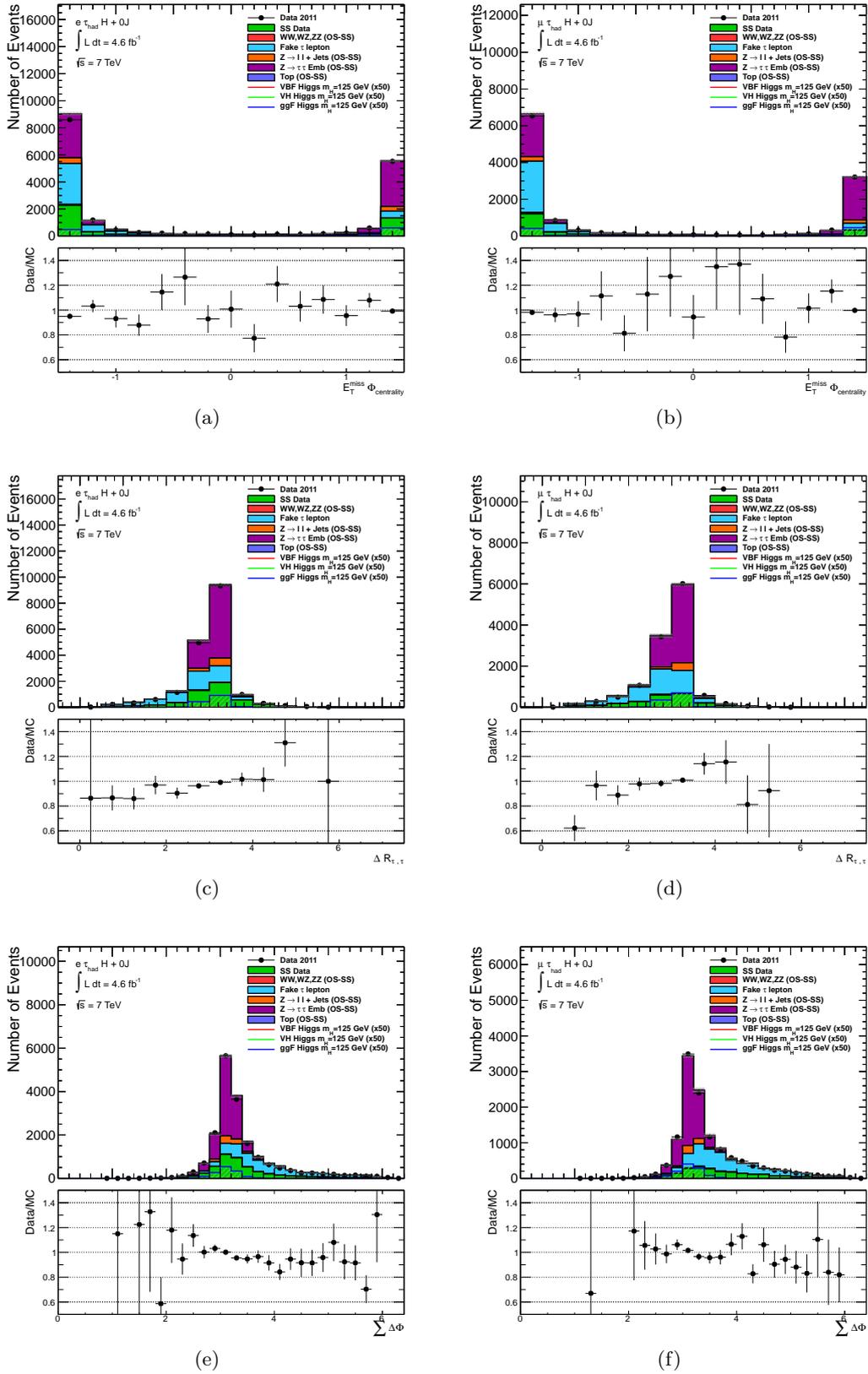


Figure C.3.: The distributions of the  $E_T^{\text{miss}} - \phi$  centrality (a,b),  $\Delta R$  (c,d) and  $\sum \Delta\phi$  (e,f) for the 0J category in the  $e + \tau_{\text{had}}$  (left column) and  $\mu + \tau_{\text{had}}$  (right column) channels. Each distribution is accompanied by its corresponding data-to-MC ratio, including the statistical uncertainties of both.

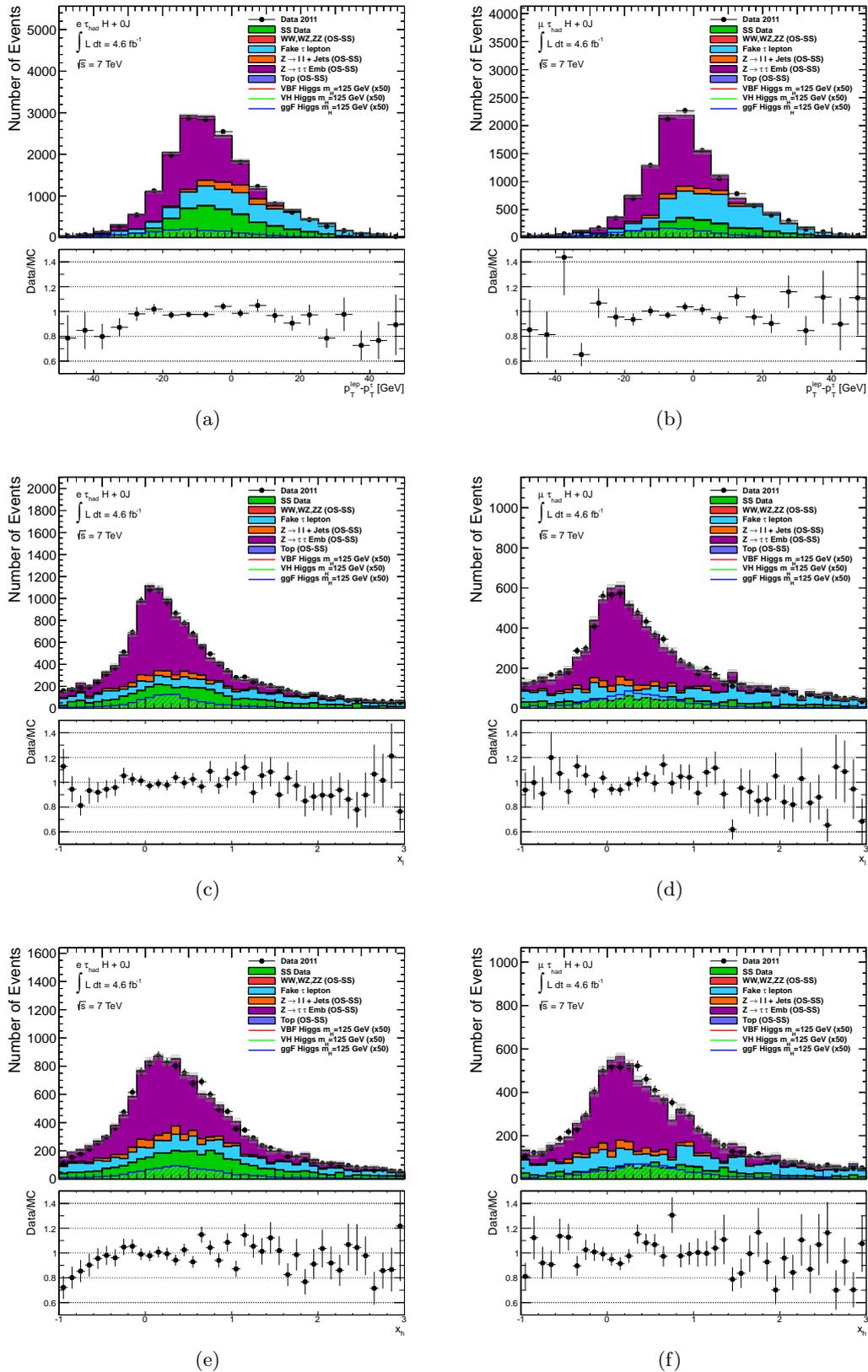


Figure C.4.: The distributions of the  $p_T$  asymmetry (a,b),  $x_l$  (c,d) and  $x_h$  (e,f) for the 0J category in the  $e + \tau_{\text{had}}$  (left column) and  $\mu + \tau_{\text{had}}$  (right column) channels. Each distribution is accompanied by its corresponding data-to-MC ratio, including the statistical uncertainties of both.

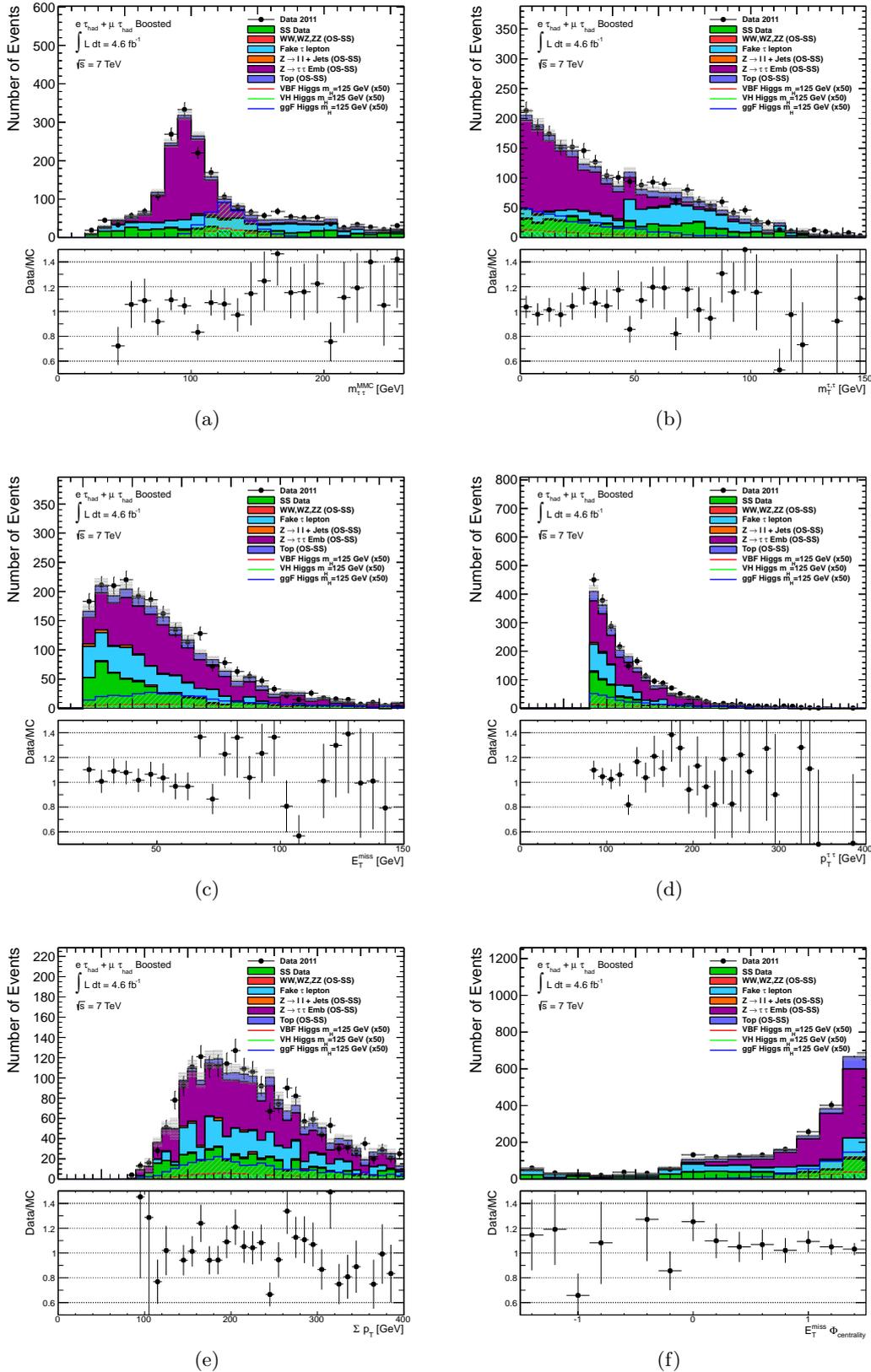


Figure C.5.: From top left to bottom right, these plots show the distributions of  $m_{\tau\tau}^{MMC}$ ,  $m_{\tau\tau}^H$ ,  $E_T^{miss}$ ,  $p_T^H$ ,  $\sum p_T$  and  $E_T^{miss} - \phi$  centrality, respectively, for the Boosted category with the  $e + \tau_{had}$  and  $\mu + \tau_{had}$  events combined. Each distribution is accompanied by its corresponding data-to-MC ratio, including the statistical uncertainties of both.

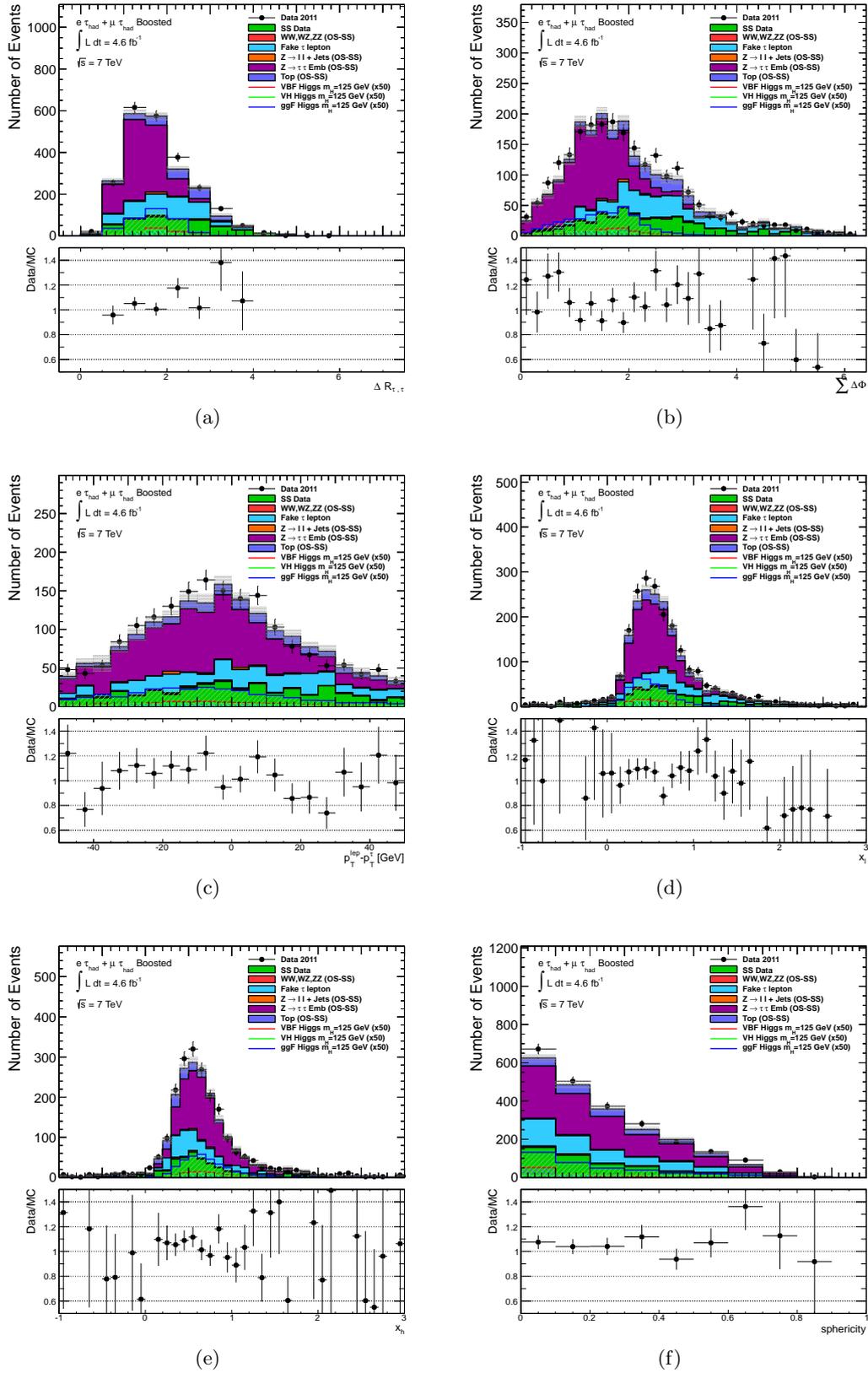


Figure C.6.: From top left to bottom right, these plots show the distributions of  $\Delta R$ ,  $\sum \Delta\phi$ ,  $p_T$  asymmetry,  $x_l$ ,  $x_h$  and the sphericity, respectively, for the Boosted category with the  $e + \tau_{\text{had}}$  and  $\mu + \tau_{\text{had}}$  events combined. Each distribution is accompanied by its corresponding data-to-MC ratio, including the statistical uncertainties of both.

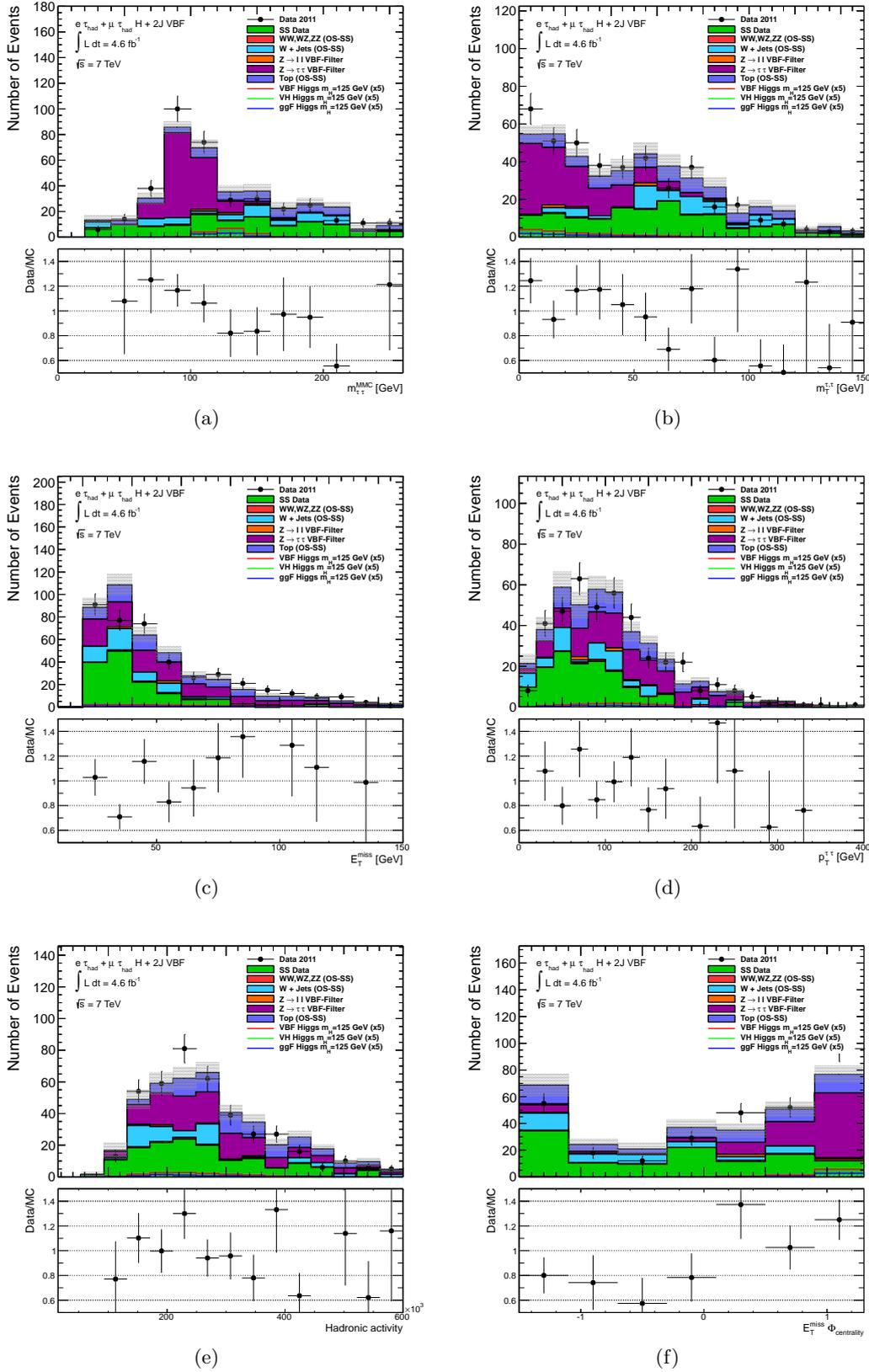


Figure C.7.: From top left to bottom right, these plots show the distributions of  $m_{\tau\tau}^{MMC}$ ,  $m_{\tau\tau}^H$ ,  $E_T^{miss}$ ,  $p_T^H$ ,  $\sum p_T$  and  $E_T^{miss} - \phi_{centrality}$ , respectively, for the VBF category with the  $e + \tau_{had}$  and  $\mu + \tau_{had}$  events combined. Each distribution is accompanied by its corresponding data-to-MC ratio, including the statistical uncertainties of both.

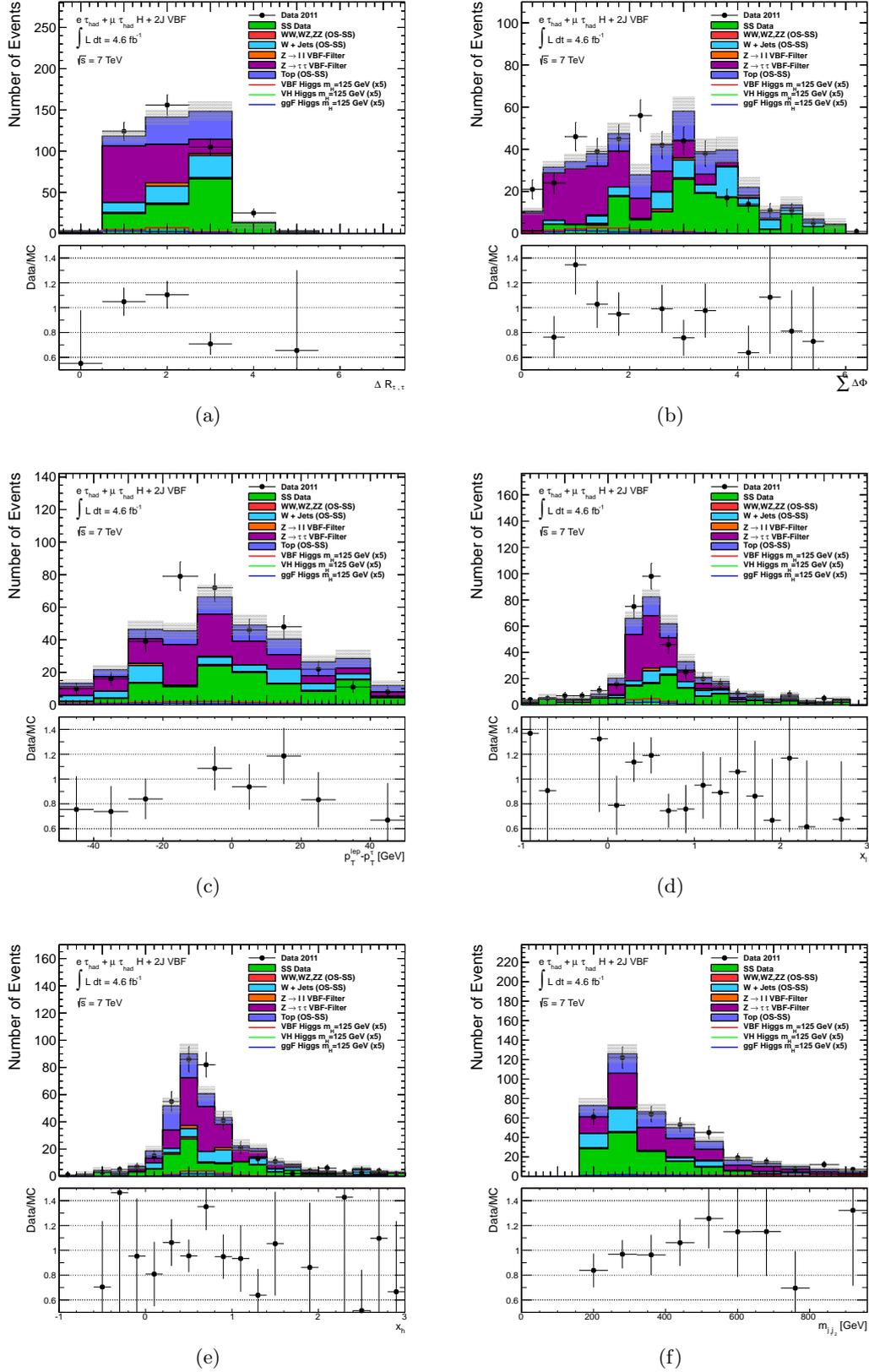


Figure C.8.: From top left to bottom right, these plots show the distributions of  $\Delta R$ ,  $\sum \Delta\phi$ ,  $p_T$  asymmetry,  $x_l$ ,  $x_h$  and  $m_{j_1,j_2}$ , respectively, for the VBF category with the  $e + \tau_{\text{had}}$  and  $\mu + \tau_{\text{had}}$  events combined. Each distribution is accompanied by its corresponding data-to-MC ratio, including the statistical uncertainties of both.

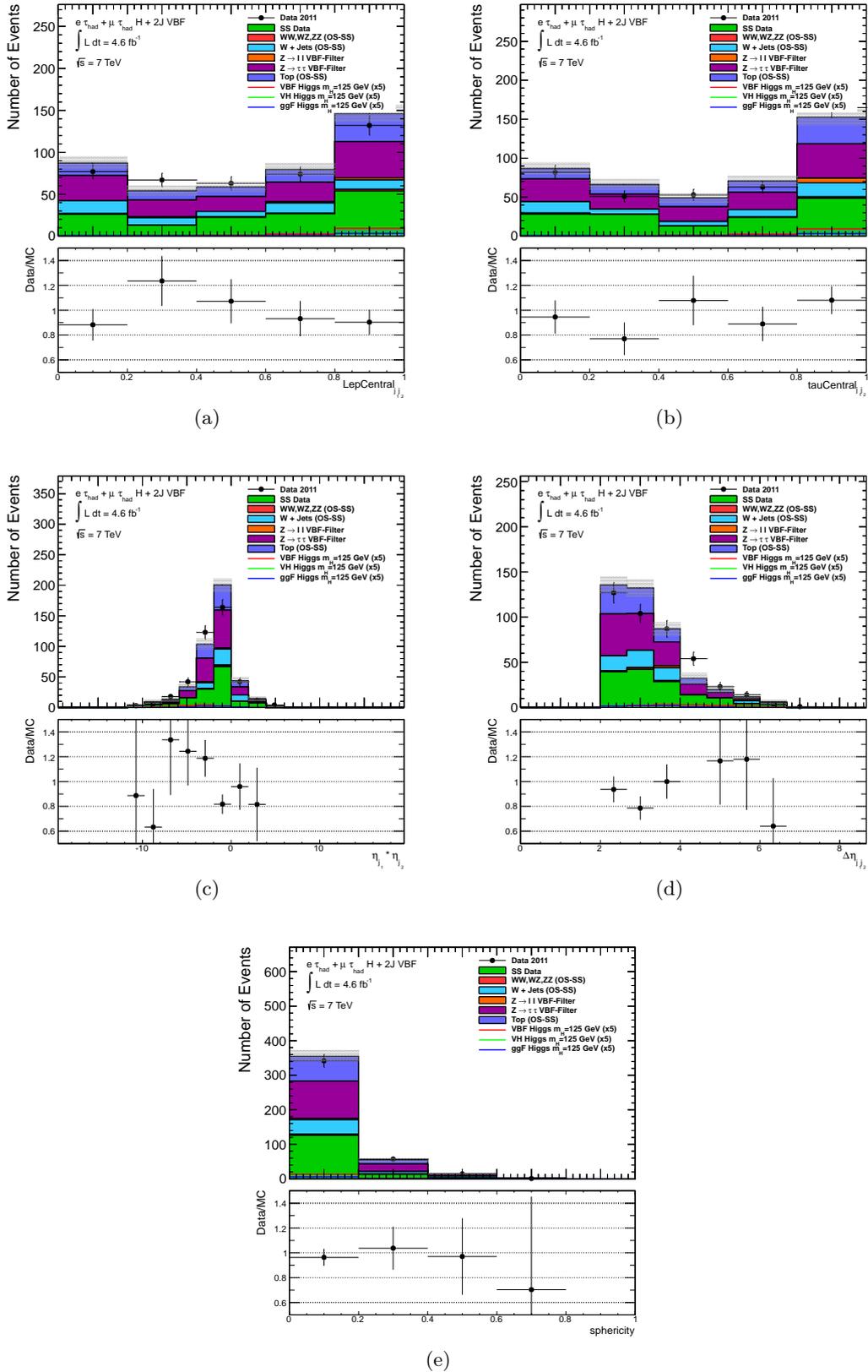


Figure C.9.: From top left to bottom, these plots show the lepton centrality,  $\tau$  centrality,  $\eta_{j_1} \cdot \eta_{j_2}$ ,  $\Delta\eta_{j_1, j_2}$  and sphericity, respectively, for the VBF category with the  $e + \tau_{\text{had}}$  and  $\mu + \tau_{\text{had}}$  events combined. Each distribution is accompanied by its corresponding data-to-MC ratio, including the statistical uncertainties of both.

# D

## BDT control regions in the training

### scenario without $m_{\tau\tau}^{\text{MMC}}$

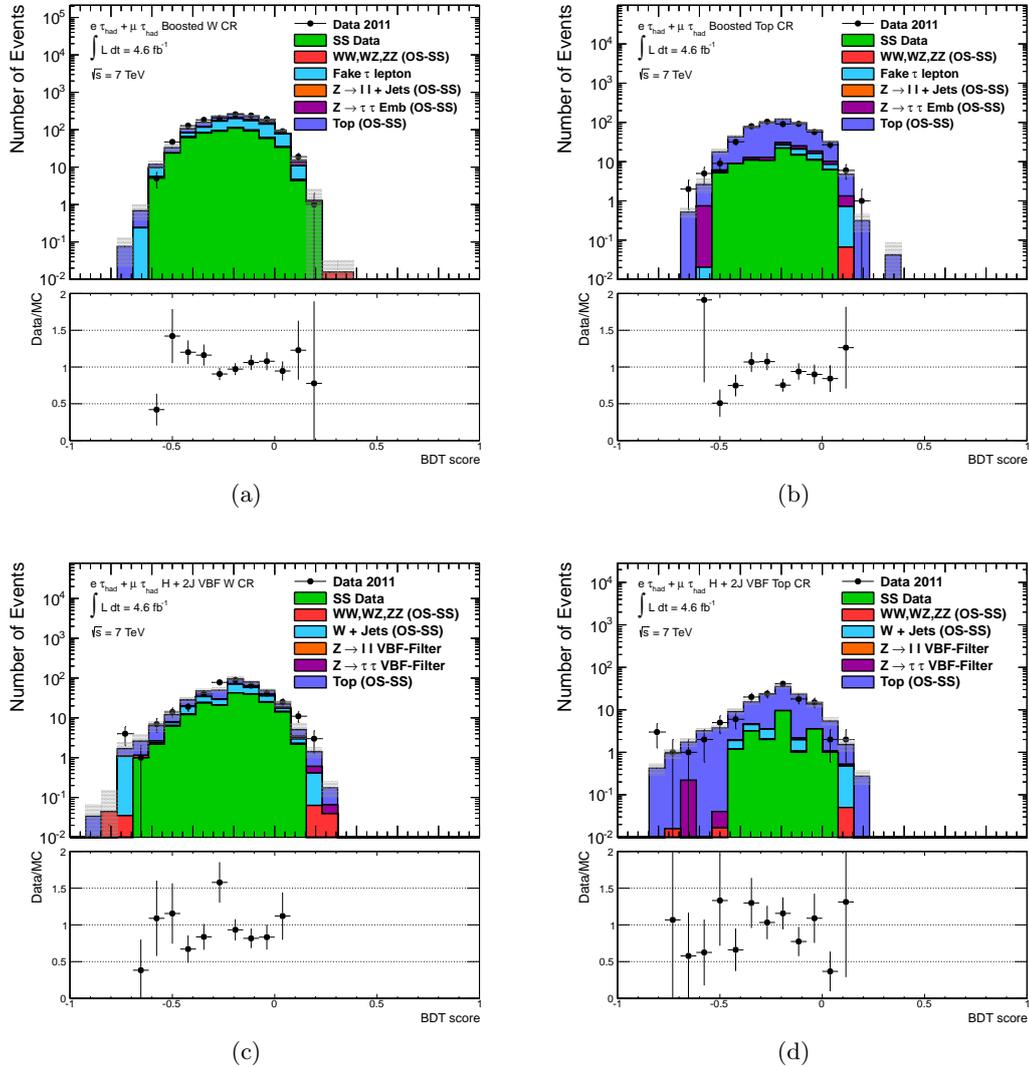


Figure D.1.: BDT score distributions for the W (left column) and top (right column) control regions for the Boosted (a,b) and VBF (c,d) category. Both  $e + \tau_{\text{had}}$  and  $\mu + \tau_{\text{had}}$  channels are combined. Each distribution is accompanied by its corresponding data-to-MC ratio, including the statistical uncertainties of both.

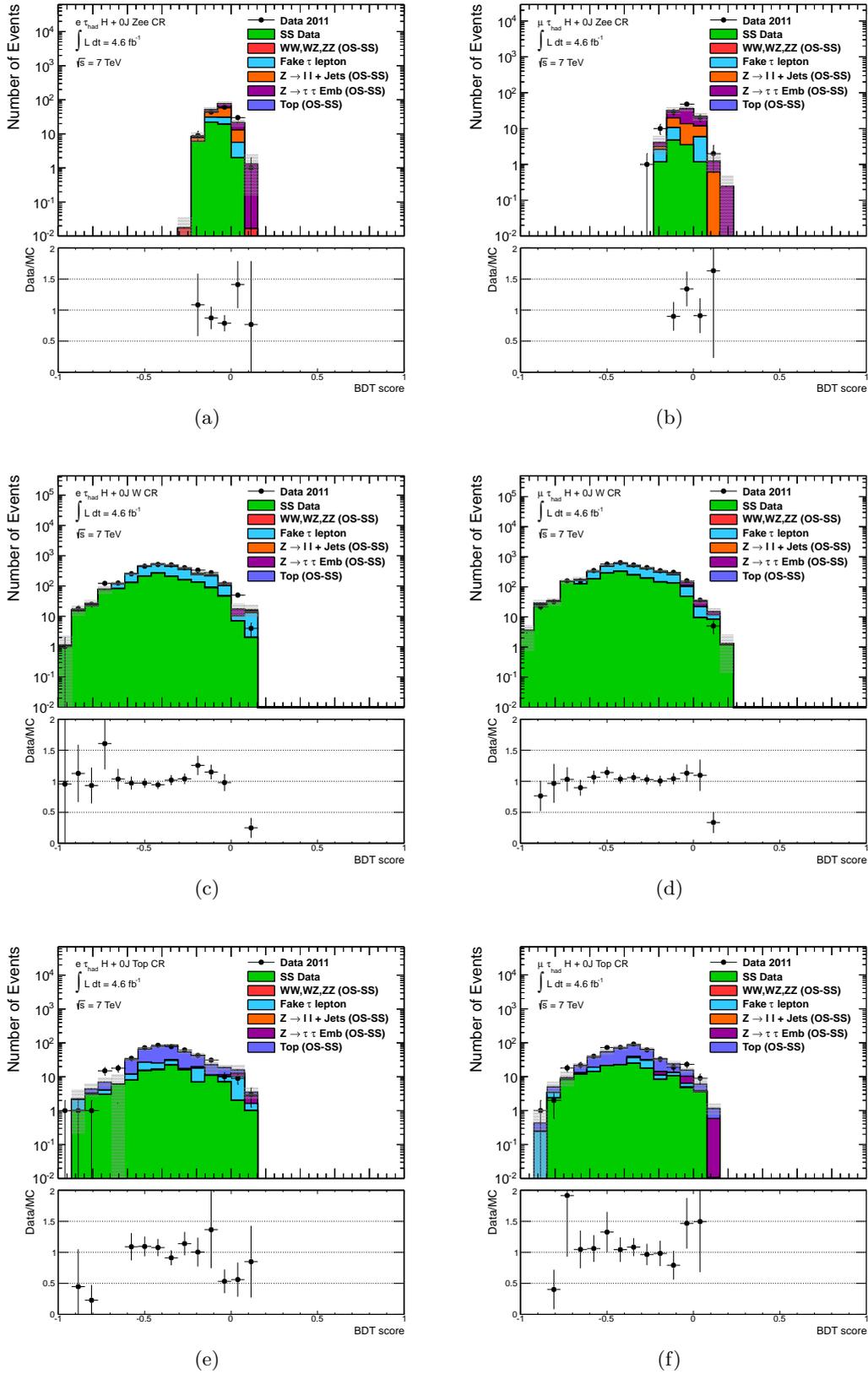


Figure D.2.: BDT score distributions for several control regions for the 0J category, split into the  $e + \tau_{\text{had}}$  (left column) and  $\mu + \tau_{\text{had}}$  (right column) channels. The  $Z \rightarrow ee$  and  $Z \rightarrow \mu\mu$  CR distributions are depicted in the top row, the W CR distributions in the middle row and the top CR distributions in the lower row. Each distribution is accompanied by its corresponding data-to-MC ratio, including the statistical uncertainties of both.

---

# E

## Systematic variations within the multivariate approach

---

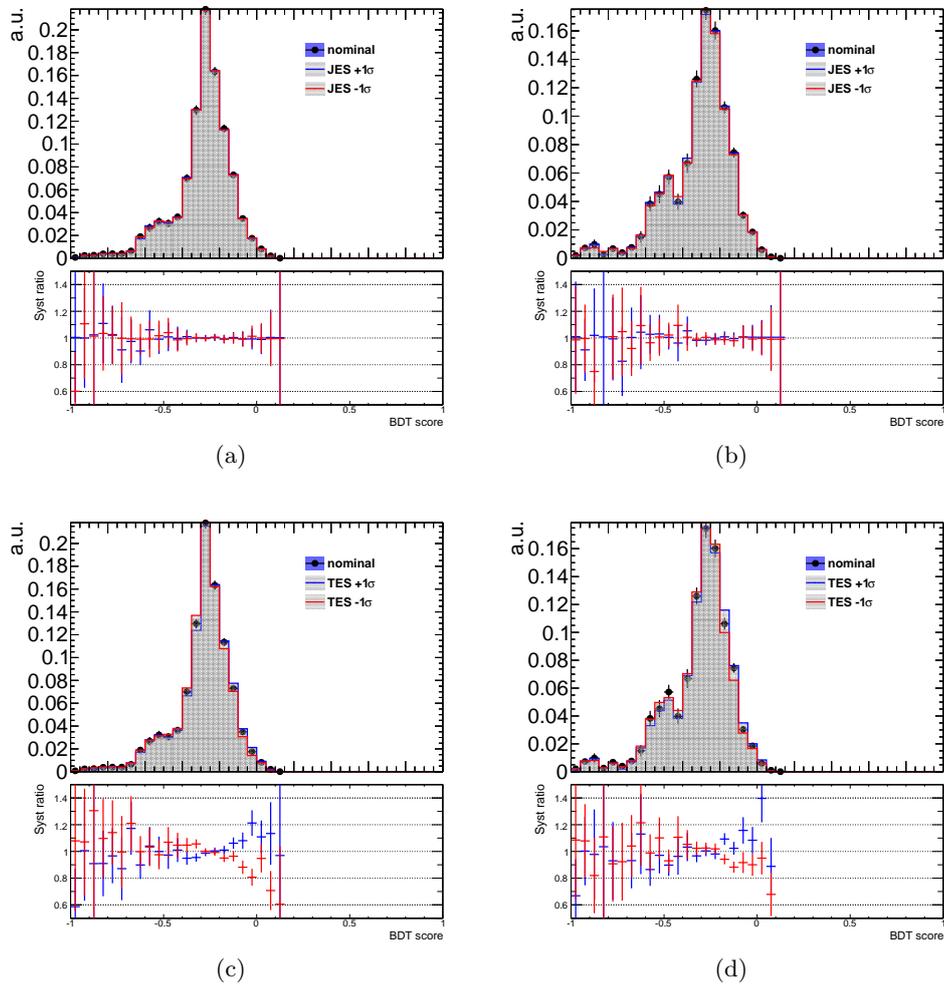


Figure E.1.: BDT score distributions for systematic variations of the jet and  $\tau$  energy scale for the 0J category. Thereby, the  $e + \tau_{\text{had}}$  channel (left column) and the  $\mu + \tau_{\text{had}}$  channel (right column) are plotted separately. The lower plot of each distribution compares the  $\pm 1\sigma$  variation to the nominal BDT score distribution within the statistical uncertainty of the latter.

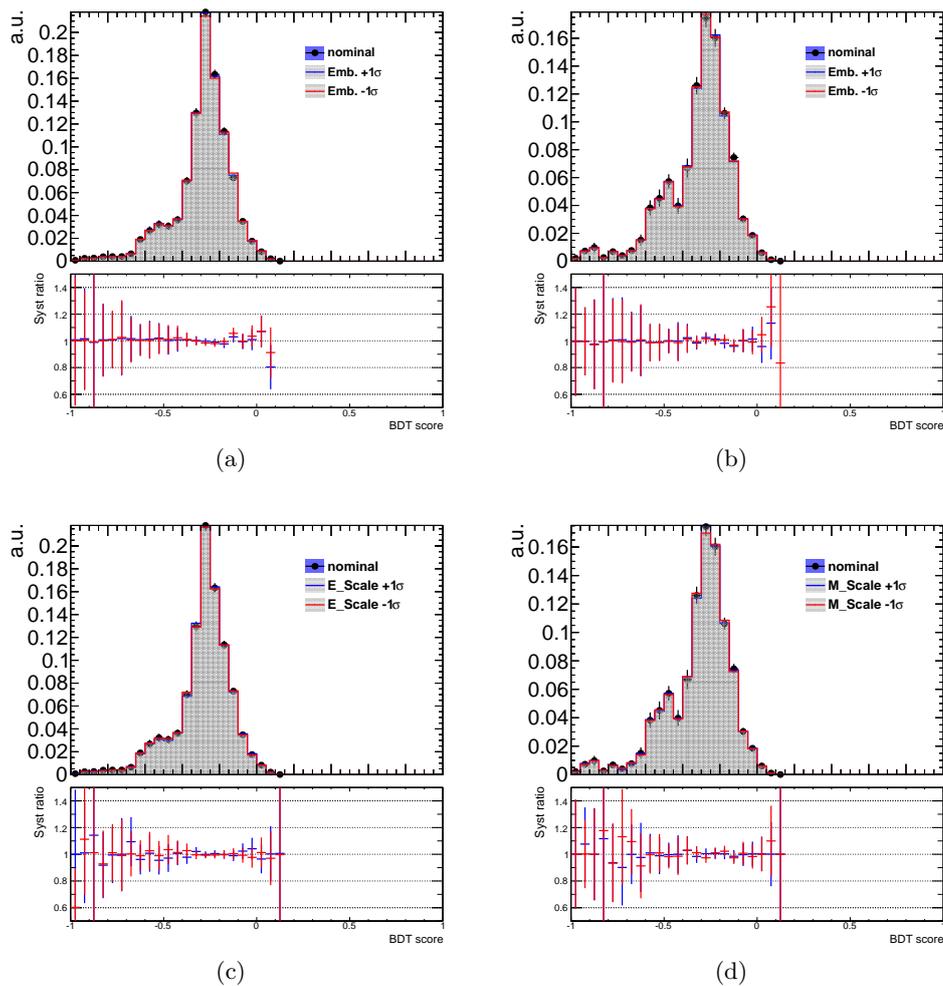


Figure E.2.: BDT score distributions for systematic variations of the embedding procedure (a,b), electron (c) and muon (d) energy scale for the 0J category. The lower plot of each distribution compares the  $\pm 1\sigma$  variation to the nominal BDT score distribution within the statistical uncertainty of the latter.

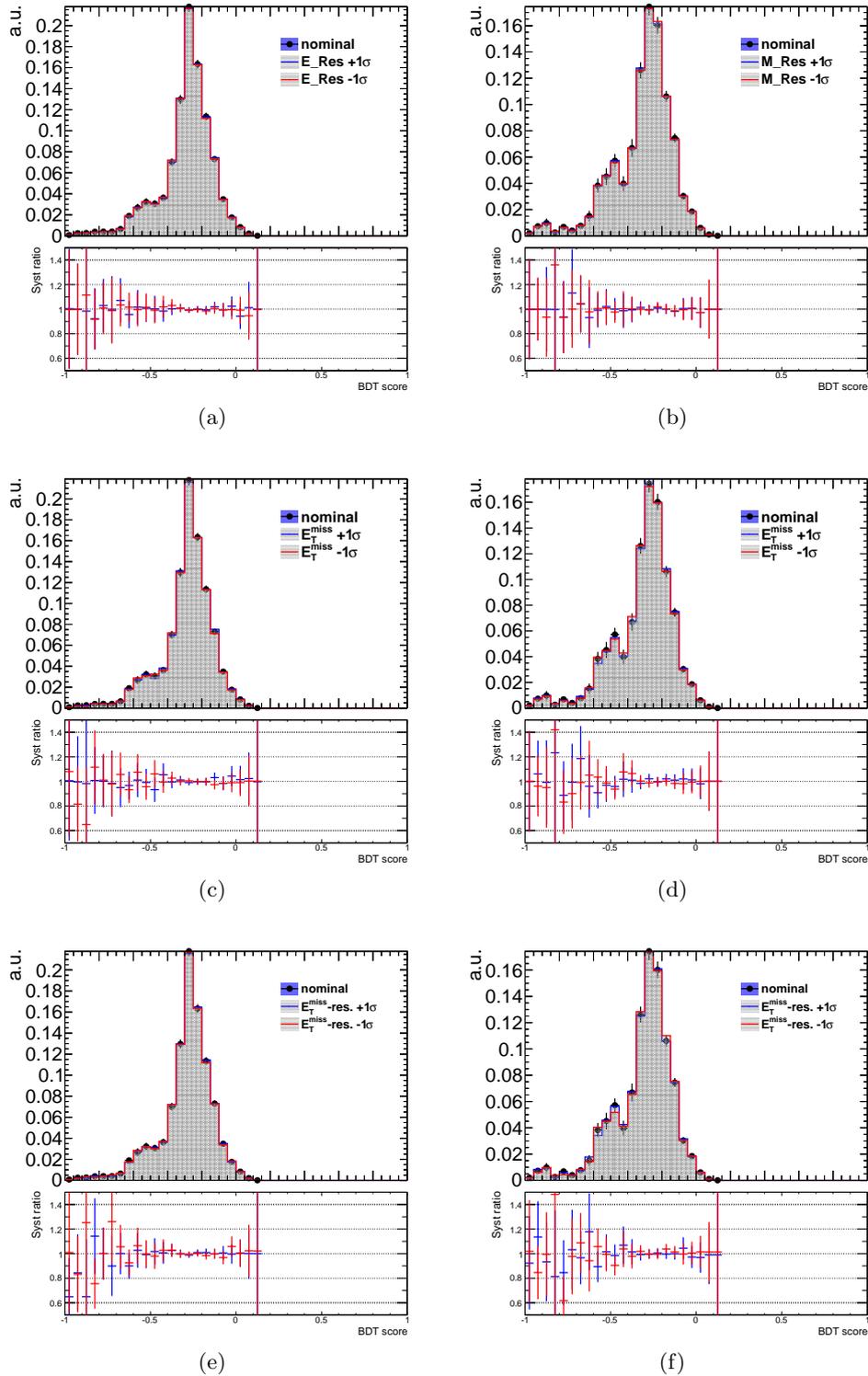


Figure E.3.: From top left to bottom right, these plots show the BDT score distributions for systematic variations of the the electron resolution, muon ID efficiency,  $E_{\text{T}}^{\text{miss}}$  energy scale (middle row) and resolution (lower row) for the 0J category. Thereby, the  $e + \tau_{\text{had}}$  channel (left column) and the  $\mu + \tau_{\text{had}}$  channel (right column) are plotted separately. The lower plot of each distribution compares the  $\pm 1\sigma$  variation to the nominal BDT score distribution within the statistical uncertainty of the latter.

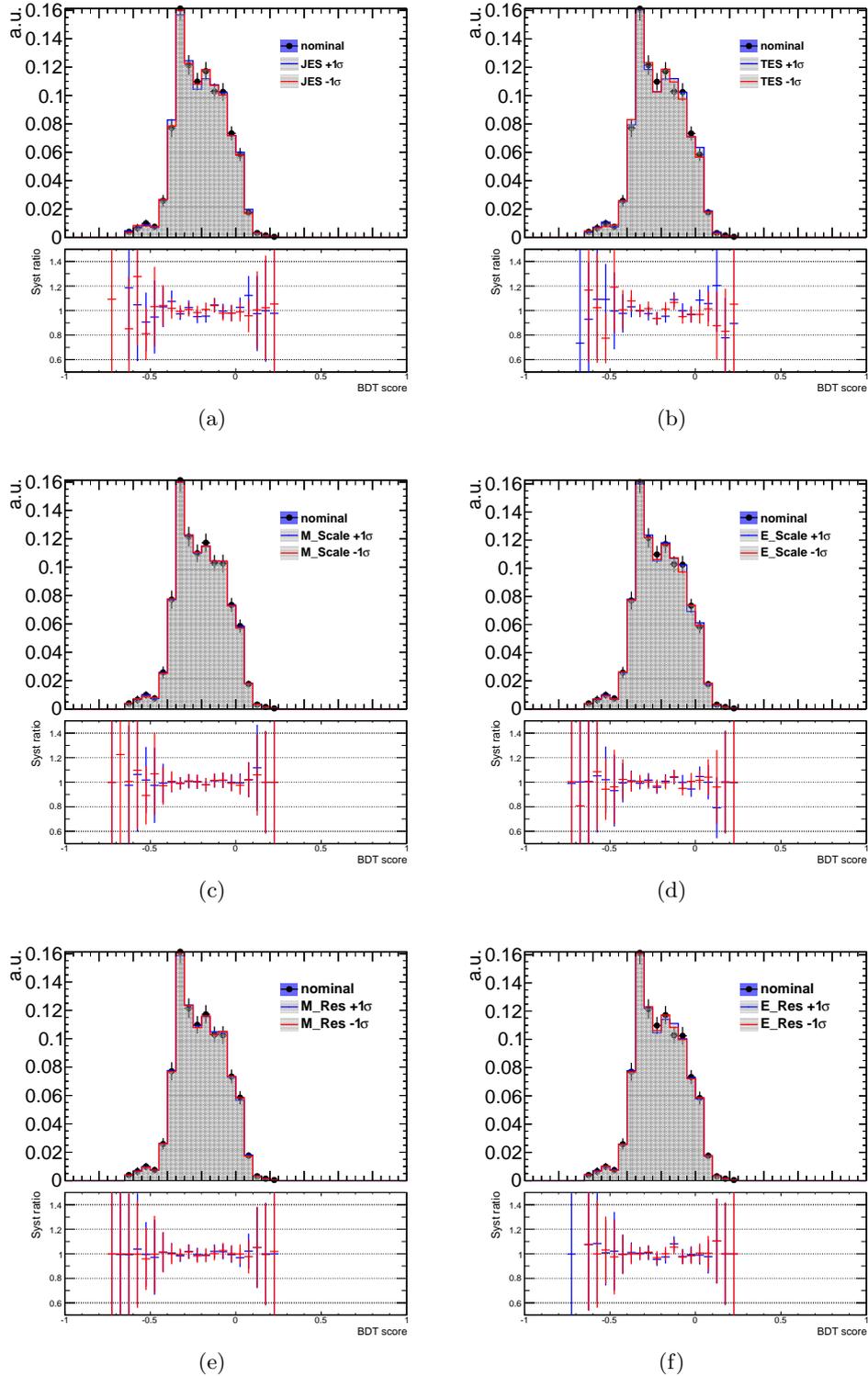


Figure E.4.: From top left to bottom right, these plots show the BDT score distributions for systematic variations of the jet and  $\tau$  energy scale, the muon and electron energy scale as well as the muon ID efficiency and electron resolution for the Boosted category. Thereby, the  $e + \tau_{\text{had}}$  and the  $\mu + \tau_{\text{had}}$  channels are combined. The lower plot of each distribution compares the  $\pm 1\sigma$  variation to the nominal BDT score distribution within the statistical uncertainty of the latter.

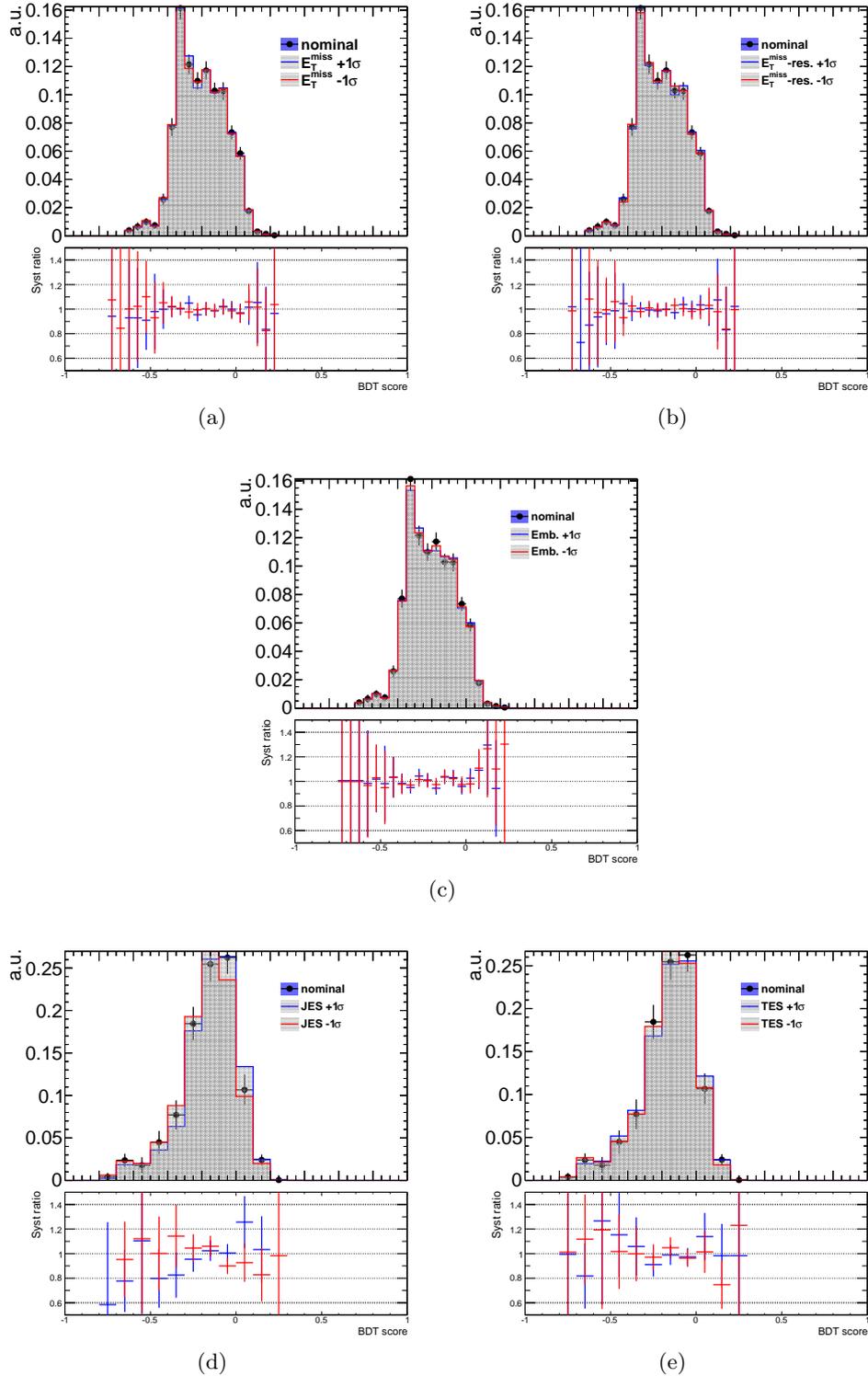


Figure E.5.: The first three plots show the BDT score distributions for systematic variations of the  $E_T^{\text{miss}}$  energy scale (a) and resolution (b) as well as the embedding (c) for the Boosted category. In addition, the jet (lower left) and  $\tau$  (lower right) energy scale systematic variation for the VBF category is depicted. Thereby, the  $e + \tau_{\text{had}}$  and the  $\mu + \tau_{\text{had}}$  channels are combined. The lower plot of each distribution compares the  $\pm 1\sigma$  variation to the nominal BDT score distribution within the statistical uncertainty of the latter.

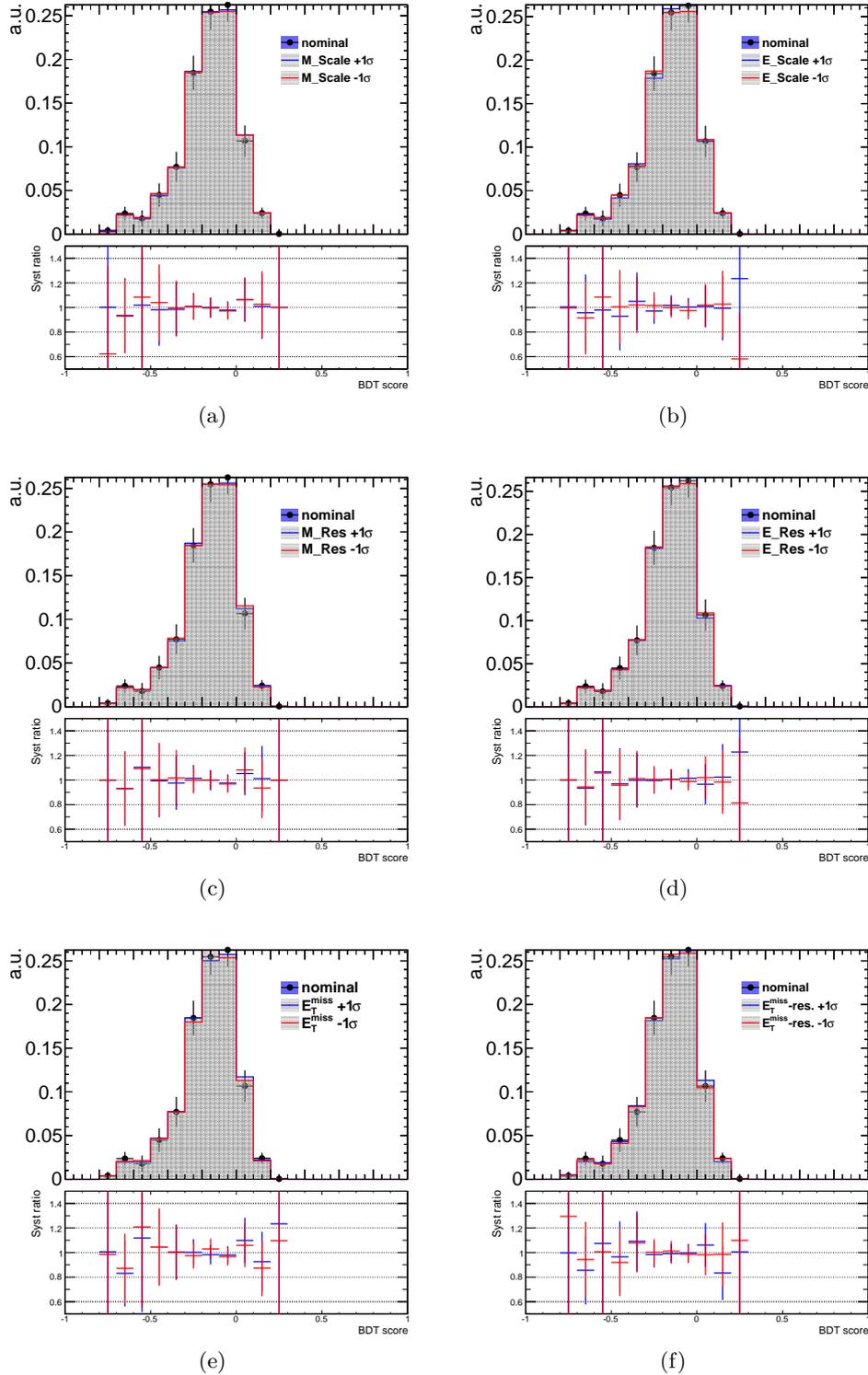


Figure E.6.: From top left to bottom right, these plots show the BDT score distributions for systematic variations of the muon and electron energy scale, the muon ID efficiency, electron resolution,  $E_T^{\text{miss}}$  energy scale and resolution for the VBF category. Thereby, the  $e + \tau_{\text{had}}$  and the  $\mu + \tau_{\text{had}}$  channels are combined. The lower plot of each distribution compares the  $\pm 1\sigma$  variation to the nominal BDT score distribution within the statistical uncertainty of the latter.

# F Cut-based systematic uncertainties

Table F.1.: Combined systematic uncertainties for the 0 jet category in the  $e + \tau_{\text{had}}$  channel.

Event sample	$\tau$ ID	Bkd est.	Ele	Emb.	JER	JES	$E_{\text{T}}^{\text{miss}}$	TES	Theory
WW,WZ,ZZ	4.7%	1.3%	3.5%		0.55%	6.3%	4.1%	10%	6.4%
VBF (125 GeV)	4.1%		3.5%		0.52%	11%	0.77%	6.5%	4.1%
WH (125 GeV)	4.0%		6.1%		3.7%	3.8%		8.7%	4.1%
ZH (125 GeV)	4.1%		5.7%		3.6%	8.6%	1.5%	7.6%	4.1%
ggF (125 GeV)	4.1%		3.4%		0.32%	3.1%	1.2%	7.4%	22%
W+Jets	0.42%	2.4%	5.0%		0.86%	1.5%	4.6%	4.4%	
$Z \rightarrow \ell\ell$	17%	2.8%	4.1%		0.76%	0.58%	1.1%	14%	13%
$Z \rightarrow \tau\tau$	1.6%	0.04%	3.7%	0.86%				8.9%	13%
Top	2.7%	14%	5.1%		10%	37%	2.5%	4.8%	

Table F.2.: Combined systematic uncertainties for the 0 jet category in the  $\mu + \tau_{\text{had}}$  channel.

Event sample	$\tau$ ID	Bkd est.	Emb.	JER	JES	$E_{\text{T}}^{\text{miss}}$	Muon	TES	Theory
WW,WZ,ZZ	3.7%	0.29%		0.48%	0.82%	0.39%	3.5%	7.7%	6.4%
VBF (125 GeV)	4.0%			1.4%	12%	1.4%	5.2%	8.2%	4.1%
WH (125 GeV)	4.0%			2.9%	8.9%	5.4%	4.6%	3.2%	4.1%
ZH (125 GeV)	4.0%			4.8%	4%	4.0%	4.9%	4.0%	4.1%
ggF (125 GeV)	4.0%			1.1%	3.1%	0.85%	5.0%	7.8%	22%
W+Jets	0.39%	3.2%		1.5%	2.3%	1.2%	4.4%	3.4%	
$Z \rightarrow \ell\ell$	0.26%	12%		0.71%	5.5%	13%	6.9%	23%	13%
$Z \rightarrow \tau\tau$	1.2%	0.04%	2.0%				1.1%	9.5%	13%
Top	2.5%	13%		3.7%	11%	4.4%	3.8%	15%	

Table F.3.: Combined systematic uncertainties for the 1 jet category in the  $e + \tau_{\text{had}}$  channel.

Event sample	$\tau$ ID	Bkd est.	Ele	Emb.	JER	JES	$E_{\text{T}}^{\text{miss}}$	TES	Theory
WW,WZ,ZZ	3.3%	1.2%	2.9%		2.6%	2.1%	1.7%	1.6%	6.4%
VBF (125 GeV)	4.0%		2.9%		0.36%	1.1%	0.09%	2.7%	11%
WH (125 GeV)	4.0%		3.1%		2.7%	2.3%	0.28%	2.8%	4.7%
ZH (125 GeV)	4.0%		2.9%		1.4%	3.1%	0.31%	1.3%	4.7%
ggF (125 GeV)	4.0%		2.9%		1.1%	6.6%	0.31%	2.5%	35%
W+Jets	0.18%	7.2%	5.0%		1.6%	6.0%	0.38%	4.9%	
$Z \rightarrow \ell\ell$	4.3%	35%	7.5%		6.5%	26%	4.5%	9.8%	13%
$Z \rightarrow \tau\tau$	0.95%	0.1%	2.9%	3.2%				4.7%	13%
Top	2.0%	15%	2.6%		1.1%	2.1%	0.73%	1.2%	

Table F.4.: Combined systematic uncertainties for the 1 jet category in the  $\mu + \tau_{\text{had}}$  channel.

Event sample	$\tau$ ID	Bkd est.	Emb.	JER	JES	$E_{\text{T}}^{\text{miss}}$	Muon	TES	Theory
WW,WZ,ZZ	3.4%	1.5%		0.9%	2.9%	0.77%	3.5%	5.1%	6.4%
VBF (125 GeV)	4.0%			0.05%	0.46%	0.09%	3.5%	1.9%	11%
WH (125 GeV)	4.0%			1.4%	1.4%	1.2%	4.2%	0.42%	4.7%
ZH (125 GeV)	4.0%			2.6%	2.7%	0.85%	3.7%	2.0%	4.7%
ggF (125 GeV)	4.0%			1.6%	6%	0.41%	3.6%	2.8%	35%
W+Jets	0.13%	11%		2.9%	7.2%	2.8%	2.8%	5.9%	
$Z \rightarrow \ell\ell$	0.12%	17%		26%	25%	8.8%	2.9%	5.7%	13%
$Z \rightarrow \tau\tau$	0.66%	0.07%	1.7%				1.1%	4.0%	13%
Top	2.0%	15%		0.63%	1.4%	0.32%	2.6%	2.6%	

Table F.5.: Combined systematic uncertainties for the Boosted category for the  $e + \tau_{\text{had}}$  plus  $\mu + \tau_{\text{had}}$  channels, combined.

Event sample	$\tau$ ID	Bkd	Ele	Emb.	JER	JES	$E_{\text{T}}^{\text{miss}}$	Muon	TES	Theory
WW,WZ,ZZ	3.9%	0.18%	1.3%		0.7%	6.4%	0.99%	1.5%	2.3%	6.4%
VBF (125 GeV)	4.0%		1.3%		0.98%	2.3%	0.10%	1.5%	0.28%	4.1%
WH (125 GeV)	4.0%		1.5%		0.26%	2.8%	0.76%	1.6%	0.89%	4.1%
ZH (125 GeV)	4.0%		1.5%		0.02%	4.0%	1.3%	1.2%	0.99%	4.1%
ggF (125 GeV)	4.0%		1.4%		1.1%	3.7%	0.24%	1.4%	0.52%	24%
W+Jets	0.07%	9.2%	2.6%		2.6%	9.2%	0.56%	1.4%	8.2%	
$Z \rightarrow \ell\ell$	6.5%	9.1%	7.1%		17%	27%	11%	3.0%	5.6%	13%
$Z \rightarrow \tau\tau$	0.53%	0.04%	1.1%	5.3%				0.6%	0.78%	13%
Top	2.1%	16%	1.5%		0.45%	7.0%	0.78%	1.2%	0.38%	

Table F.6.: Combined systematic uncertainties for the VBF category for the  $e + \tau_{\text{had}}$  plus  $\mu + \tau_{\text{had}}$  channels, combined.

Event sample	$\tau$ ID	Bkd est.	Ele	F.F.	JER	JES	$E_{\text{T}}^{\text{miss}}$	Muon	TES	Theory
WW,WZ,ZZ	3.1%	0.18%	10%		17%	10%	2.7%	1.5%	20%	6.4%
VBF (125 GeV)	4.0%		1.4%		1.2%	10%	1.7%	1.7%	2.3%	11%
ggF (125 GeV)	4.0%		1.7%		4.8%	10%	5.7%	1.0%	3.4%	35%
Fake-Tau				50%						
$Z \rightarrow \ell\ell$	8.5%	36%	16%		3.2%	21%	15%			13%
$Z \rightarrow \tau\tau$	4.0%	0.13%	1.6%		1.7%	13%	1.6%	2.0%	0.85%	13%
Top	1.9%	11%	1.4%		3.6%	6.2%	1.7%	1.1%	4.3%	



# G BDT-related systematic uncertainties

## G.1. Training with $m_{\tau\tau}^{\text{MMC}}$

Table G.1.: Combined systematic uncertainties for the 0 jet category in the  $e + \tau_{\text{had}}$  channel.

Event sample	$\tau$ ID	Bkd est.	Ele	Emb.	JER	JES	$E_{\text{T}}^{\text{miss}}$	TES	Theory
WW,WZ,ZZ	3.9%	0.44%	2.9%		0.35%	2.6%	0.51%	5.3%	6.4%
VBF (125 GeV)	4.0%		3.0%		0.41%	12%	0.65%	3.4%	4.1%
WH (125 GeV)	4.0%		3.1%		4.5%	9.3%	1.5%	0.85%	4.1%
ZH (125 GeV)	4.0%		3.4%		5.7%	9.3%	$\sim 0\%$	3.1%	4.1%
ggF (125 GeV)	4.0%		2.9%		0.74%	2.9%	0.22%	3.0%	22%
W+Jets	0.15%	2.5%	4.3%		0.87%	1.2%	0.39%	2.1%	
$Z \rightarrow \ell\ell$	17%	3%	2.5%		1.1%	2.2%	0.57%	4.6%	13%
$Z \rightarrow \tau\tau$	0.97%	0.05%	3.2%	1.6%				6.6%	13%
Top	2.7%	14%	3.4%		0.74%	15%	1.1%	6.3%	

Table G.2.: Combined systematic uncertainties for the 0 jet category in the  $\mu + \tau_{\text{had}}$  channel.

Event sample	$\tau$ ID	Bkd est.	Emb.	JER	JES	$E_{\text{T}}^{\text{miss}}$	Muon	TES	Theory
WW,WZ,ZZ	3.8%	0.64%		0.5%	2.3%	1.1%	2.0%	4.4%	6.4%
VBF (125 GeV)	4.0%			0.02%	15%	0.24%	2.0%	2.3%	4.1%
WH (125 GeV)	4.0%			2.4%	5.0%	0.87%	2.0%	3.6%	4.1%
ZH (125 GeV)	4.0%			2.2%	9.2%	$\sim 0\%$	2.0%	1.0%	4.1%
ggF (125 GeV)	4.0%			0.83%	3.3%	0.31%	2.0%	3.2%	22%
W+Jets	0.02%	3.6%		0.06%	2.4%	0.76%	2.0%	2.2%	
$Z \rightarrow \ell\ell$	0.01%	4.7%		0.86%	2.1%	0.94%	2.0%	3.8%	13%
$Z \rightarrow \tau\tau$		0.03%	2.5%				1.0%	7.0%	13%
Top	2.7%	13%		3.0%	15%	1.1%	2.2%	4.5%	

Table G.3.: Combined systematic uncertainties for the Boosted category for the  $e + \tau_{\text{had}}$  plus  $\mu + \tau_{\text{had}}$  channels, combined.

Event sample	$\tau$ ID	Bkd	Ele	Emb.	JER	JES	$E_{\text{T}}^{\text{miss}}$	Muon	TES	Theory
WW,WZ,ZZ	3.4%	1.6%	1.9%		0.88%	4.0%	2.5%	2.1%	4.0%	6.4%
VBF (125 GeV)	4.0%		1.6%		0.46%	1.2%	0.80%	0.87%	1.8%	4.1%
WH (125 GeV)	4.0%		1.6%		0.73%	1.2%	0.50%	0.90%	2.0%	4.1%
ZH (125 GeV)	4.0%		1.6%		0.3%	3.1%	0.28%	0.88%	2.5%	4.1%
ggF (125 GeV)	4.0%		1.6%		0.64%	3.6%	1.3%	0.80%	2.3%	24%
W+Jets	0.1%	6.8%	3.5%		1.4%	3.4%	3.2%	0.91%	2.5%	
$Z \rightarrow \ell\ell$	3.6%	19%	10%		3.1%	20%	9.4%	0.9%	2.6%	13%
$Z \rightarrow \tau\tau$	0.5%	0.07%	1.4%	3.3%				0.54%	3.1%	13%
Top	2.0%	16%	1.5%		0.25%	3.1%	1.1%	0.91%	3.6%	

Table G.4.: Combined systematic uncertainties for the VBF category for the  $e + \tau_{\text{had}}$  plus  $\mu + \tau_{\text{had}}$  channels, combined.

Event sample	$\tau$ ID	Bkd est.	Ele	JER	JES	$E_{\text{T}}^{\text{miss}}$	Muon	TES	Theory
WW,WZ,ZZ	3.0%	1.7%	2.5%	9.4%	5.3%	4.2%	1.4%	6.5%	6.4%
VBF (125 GeV)	4.0%		1.6%	0.06%	7.2%	0.01%	0.90%	2.5%	11%
WH (125 GeV)	3.9%		1.0%	5.1%	15%	$\sim 0\%$	1.2%	3.4%	4.7%
ZH (125 GeV)	4.0%		0.82%	5.3%	7.4%	$\sim 0\%$	1.4%	4.9%	4.7%
ggF (125 GeV)	4.0%		1.6%	2.0%	11%	0.26%	0.86%	1.8%	35%
W+Jets	0.04%	11%	4.9%	3.2%	9.6%	7.8%	6.8%	11%	
ZVBFFLL	2.3%	40%	5.9%	0.46%	27%	10%	2.9%	3.4%	13%
ZVBFFTT	4.0%	0.13%	1.7%	0.98%	13%	0.78%	0.87%	3.6%	13%
Top	1.9%	16%	1.8%	4.1%	3.8%	1.7%	1.0%	4.7%	

**G.2. Training without  $m_{\tau\tau}^{\text{MMC}}$** Table G.5.: Combined systematic uncertainties for the 0 jet category in the  $e + \tau_{\text{had}}$  channel.

Event sample	$\tau$ ID	Bkd est.	Ele	Emb.	JER	JES	$E_{\text{T}}^{\text{miss}}$	TES	Theory
WW,WZ,ZZ	6.4%	1.5%	7.1%		1.1%	8.2%	13%	24%	6.4%
VBF (125 GeV)	4.0%		3.1%		4.6%	13%	3.0%	18%	4.1%
WH (125 GeV)	4.0%		2.5%		13%	13%	8.6%	6.4%	4.1%
ZH (125 GeV)	4.0%		2.8%		63%	51%		30%	4.1%
ggF (125 GeV)	4.1%		3.2%		0.1%	3.5%	0.8%	18%	22%
W+Jets	0.2%	4.8%	9.1%		5.2%	3.7%	3.4%	22%	
$Z \rightarrow \ell\ell$	19%	1.8%	9.0%		0.6%	7.8%	20%	10%	13%
$Z \rightarrow \tau\tau$	1.5%	0.03%	3.6%	0.92 %				25%	13%
Top	2.4%	13%	2.5%		13%	27%	10%	4.6%	

Table G.6.: Combined systematic uncertainties for the 0 jet category in the  $\mu + \tau_{\text{had}}$  channel.

Event sample	$\tau$ ID	Bkd est.	Emb.	JER	JES	$E_{\text{T}}^{\text{miss}}$	Muon	TES	Theory
WW,WZ,ZZ	3.8%	0.96%		17%	28%	18%	1.9%	8.7%	6.4%
VBF (125 GeV)	3.9%			1.2%	6.4%	4.3%	3.5%	13%	4.1%
WH (125 GeV)	4.0%					34%	1.9%		4.1%
ZH (125 GeV)	4.0%			17%		27%	2.1%		4.1%
ggF (125 GeV)	4.0%			1.5%	4.4%	2.5%	2.0%	23%	22%
W+Jets		2.4%		3.0%	6.0%	8.3%	3.4%	5.1%	
$Z \rightarrow \ell\ell$		1.7%		0.2%	4.8%	5.4%	2.1%	16%	13%
$Z \rightarrow \tau\tau$		0.05%	1.85%				1.0%	22%	13%
Top	3.3%	21%		7.0%	25%	15%	2.0%	17%	

Table G.7.: Combined systematic uncertainties for the Boosted category for the  $e + \tau_{\text{had}}$  plus  $\mu + \tau_{\text{had}}$  channels, combined.

Event sample	$\tau$ ID	Bkd	Ele	Emb.	JER	JES	$E_{\text{T}}^{\text{miss}}$	Muon	TES	Theory
WW,WZ,ZZ	3.9%	1.1%	2.7%		0.64%	20%	3.5%	0.87%	9.0%	6.4%
VBF (125 GeV)	4.0%		1.6%		1.5%	5.2%	0.81%	0.83%	1.6%	4.1%
WH (125 GeV)	4.0%		1.7%		2.9%	11%	1.1%	0.91%	0.32%	4.1%
ZH (125 GeV)	4.0%		1.5%		1.7%	9.1%	1.1%	0.87%	1.6%	4.1%
ggF (125 GeV)	4.0%		1.7%		2.0%	8.3%	0.86%	0.80%	3.0%	24%
W+Jets	0.09%	4.6%	2.3%		1.6%	6.7%	10%	0.89%	1.6%	
$Z \rightarrow \ell\ell$	7.1%	12%	8.5%		25%	21%	10%	1.7%	52%	13%
$Z \rightarrow \tau\tau$	0.79%	0.06%	1.7%	5.5%				0.47%	2.9%	13%
Top	2.2%	20%	2.7%		2.6%	9.4%	6.9%	1.4%	4.6%	

Table G.8.: Combined systematic uncertainties for the VBF category for the  $e + \tau_{\text{had}}$  plus  $\mu + \tau_{\text{had}}$  channels, combined.

Event sample	$\tau$ ID	Bkd est.	Ele	JER	JES	$E_{\text{T}}^{\text{miss}}$	Muon	TES	Theory
WW,WZ,ZZ	5.2%	0.98%	3.1%	19%	3.8%	7.7%	3.1%	5.8%	6.4%
VBF (125 GeV)	4.0%		1.6%	0.06%	8.4%	0.07%	0.91%	3.0%	11%
WH (125 GeV)	4.0%		0.4%	18%	11%	$\sim 0\%$	1.6%	8.1%	4.7%
ZH (125 GeV)	4.0%		0.6%		1.9%	$\sim 0\%$	1.5%	7.9%	4.7%
ggF (125 GeV)	4.0%		1.8%	5.6%	14%	0.14%	1.1%	3.8%	35%
W+Jets		11%	28%	8.3%	12%	24%	1.4%	14%	
ZVBFFLL	1.0%	49%	5.4%	24%	54%	12%	0.52%	25%	13%
ZVBFFTT	3.9%	0.12%	2.0%	1.0%	17%	1.3%	0.85%	3.5%	13%
Top	2.9%	17%	1.1%	26%	8.6%	5.3%	1.3%	15%	

---

## Bibliography

---

- [1] M. Banner et al., *Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN pp collider*, Physics Letters B 122 (1983), no. 56 pp. 476 – 485.
- [2] G. Arnison et al., *Further evidence for charged intermediate vector bosons at the SPS collider*, Physics Letters B 129 (1983), no. 34 pp. 273 – 282.
- [3] P. Bagnaia et al., *Evidence for  $Z0e+e$  at the CERN pp collider*, Physics Letters B 129 (1983), no. 12 pp. 130 – 140.
- [4] G. Arnison et al., *Experimental observation of lepton pairs of invariant mass around  $95 \text{ GeV}/c^2$  at the CERN SPS collider*, Physics Letters B 126 (1983), no. 5 pp. 398 – 410.
- [5] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. 13 (1964) pp. 321–323.
- [6] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. 13 (1964) pp. 508–509.
- [7] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. 145 (1966) pp. 1156–1163.
- [8] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. 13 (1964) pp. 585–587.
- [9] T. W. B. Kibble, *Symmetry Breaking in Non-Abelian Gauge Theories*, Phys. Rev. 155 (1967) pp. 1554–1561.
- [10] The DONUT Collaboration Collaboration, K. Kodama et al., *Observation of tau neutrino interactions*, Phys.Lett. B504 (2001) pp. 218–224, hep-ex/0012035.

- [11] The ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Physics Letters B 716 (2012), no. 1 pp. 1 – 29.
- [12] The CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B 716 (2012) pp. 30–61. 59 p.
- [13] The Particle Data Group Collaboration, J. Beringer et al., *Review of Particle Physics*, Phys. Rev. D 86 (2012) p. 010001.
- [14] S. Asai et al., *Prospects for the search for a standard model Higgs boson in ATLAS using vector boson fusion*, The European Physical Journal C - Particles and Fields 32 (2004) pp. s19–s54.
- [15] The ATLAS Collaboration, *Search for the Standard Model Higgs boson in  $H \rightarrow \tau\tau$  decays in proton-proton collisions with the ATLAS detector*, ATLAS-CONF-2012-160, CERN, Geneva, Nov, 2012.
- [16] L. Breiman, J. Friedman, C. J. Stone and R. A. Olshen, *Classification and regression trees*. Chapman & Hall/CRC, 1984.
- [17] F. Abe et al., *Evidence for top quark production in  $p p$  collisions at  $\sqrt{s} = 1.8$  TeV*, Phys. Rev. D 50 (1994) pp. 2966–3026.
- [18] The DØ Collaboration, *Observation of the top quark*, Il Nuovo Cimento A Series 11 109 (1996) pp. 755–769.
- [19] E. Noether, *Invariant Variation Problems*, Gott.Nachr. 1918 (1918) pp. 235–257, physics/0503066.
- [20] F. Halzen and A. D. Martin, *Quarks and Leptones: An Introductory Course in Modern Particle Physics*. Jan., 1984.
- [21] A. Pich, *The Standard model of electroweak interactions*. 2007.
- [22] D. Griffiths, *Introduction to elementary particles*. 2008.
- [23] M. Maltoni, T. Schwetz, M. Trtola and J. W. F. Valle, *Status of global fits to neutrino oscillations*, New Journal of Physics 6 (2004), no. 1 p. 122.
- [24] The Super-Kamiokande Collaboration, Y. Fukuda et al., *Evidence for Oscillation of Atmospheric Neutrinos*, Phys. Rev. Lett. 81 (1998) pp. 1562–1567.

- [25] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions*, Phys. Rev. Lett. 30 (1973) pp. 1346–1349.
- [26] S. L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics 22 (1961), no. 4 pp. 579 – 588.
- [27] A. Salam and J. Ward, *Weak and electromagnetic interactions*, Il Nuovo Cimento Series 10 11 (1959) pp. 568–577.
- [28] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. 19 (1967) pp. 1264–1266.
- [29] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Progress of Theoretical Physics 49 (1973), no. 2 pp. 652–657.
- [30] G. 't Hooft and M. Veltman, *Regularization and renormalization of gauge fields*, Nuclear Physics B 44 (1972), no. 1 pp. 189 – 213.
- [31] J. M. Campbell, J. W. Huston and W. J. Stirling, *Hard interactions of quarks and gluons: a primer for LHC physics*, Reports on Progress in Physics 70 (2007), no. 1 p. 89.
- [32] A. Martin, W. Stirling, R. Thorne and G. Watt, *Parton distributions for the LHC*, The European Physical Journal C 63 (2009) pp. 189–285.
- [33] G. Watt and R. Thorne, *Study of Monte Carlo approach to experimental uncertainty propagation with MSTW 2008 PDFs*, Journal of High Energy Physics 2012 (2012) pp. 1–38.
- [34] J. C. Collins, L. Frankfurt and M. Strikman, *Factorization for hard exclusive electroproduction of mesons in QCD*, Phys. Rev. D 56 (1997) pp. 2982–3006.
- [35] The ATLAS Collaboration, *ATLAS Sensitivity Prospects for 1 Higgs Boson Production at the LHC Running at 7, 8 or 9 TeV*, ATL-PHYS-PUB-2010-015, CERN, Geneva, Nov, 2010.
- [36] LHC Higgs Cross Section Working Group et al., *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*, ArXiv e-prints (2012) arXiv:1201.3084 [hep-ph].
- [37] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino and R. Tanaka (Eds.), *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*, CERN-2011-002 (CERN, Geneva, 2011) arXiv:1101.0593 [hep-ph].

- [38] ALEPH Collaboration and CDF Collaboration and D0 Collaboration and DELPHI Collaboration and L3 Collaboration and OPAL Collaboration and SLD Collaboration and LEP Electroweak Working Group and Tevatron Electroweak Working Group and SLD electroweak heavy flavour group, *Precision Electroweak Measurements and Constraints on the Standard Model*, 2010.
- [39] T. Hambye and K. Riesselmann, *Matching conditions and Higgs boson mass upper bounds reexamined*, Phys. Rev. D 55 (1997) pp. 7255–7262.
- [40] The ATLAS Collaboration, *Combined measurements of the mass and signal strength of the Higgs-like boson with the ATLAS detector using up to 25 fb<sup>-1</sup> of proton-proton collision data*, ATLAS-CONF-2013-014, CERN, Geneva, Mar, 2013.
- [41] The CMS Collaboration, *Combination of standard model Higgs boson searches and measurements of the properties of the new boson with a mass near 125 GeV*, CMS-PAS-HIG-12-045, CERN, Geneva, 2012.
- [42] The ATLAS Collaboration, *Combined coupling measurements of the Higgs-like boson with the ATLAS detector using up to 25 fb<sup>-1</sup> of proton-proton collision data*, ATLAS-CONF-2013-034, CERN, Geneva, Mar, 2013.
- [43] The ATLAS Collaboration, *Measurements of the properties of the Higgs-like boson in the four lepton decay channel with the ATLAS detector using 25 fb<sup>-1</sup> of proton-proton collision data*, ATLAS-CONF-2013-013, CERN, Geneva, Mar, 2013.
- [44] The ATLAS Collaboration, *Measurements of the properties of the Higgs-like boson in the two photon decay channel with the ATLAS detector using 25 fb<sup>-1</sup> of proton-proton collision data*, ATLAS-CONF-2013-012, CERN, Geneva, Mar, 2013.
- [45] The ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation 3 (2008), no. 08 p. S08003.
- [46] The ATLAS Collaboration, *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, arXiv:0901.0512 (2008).
- [47] D. Froidevaux and P. Sphicas, *General-Purpose Detectors for the Large Hadron Collider*, Annual Review of Nuclear and Particle Science 56 (2006) pp. 375–440.
- [48] The ATLAS Collaboration, *Luminosity Determination in pp Collisions at sqrt(s) = 7 TeV using the ATLAS Detector in 2011*, ATLAS-CONF-2011-116, CERN, Geneva, Aug, 2011.

- [49] W. Buttinger, *The ATLAS Level-1 Trigger System*, ATL-DAQ-PROC-2012-024, CERN, Geneva, Jun, 2012.
- [50] I. Hristova, *The evolution and performance of the ATLAS calorimeter-based triggers in 2011 and 2012*, ATL-DAQ-PROC-2012-051, CERN, Geneva, Oct, 2012.
- [51] Calafiura, P. and Lavrijsen, W. and Leggett, C. and Marino, M. and Quarrie, D., *The athena control framework in production, new developments and lessons learned*, 2005.
- [52] T. Cornelissen et al., *The new ATLAS track reconstruction (NEWT)*, Journal of Physics: Conference Series 119 (2008), no. 3 p. 032014.
- [53] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl.Instrum.Meth. A262 (1987) pp. 444–450.
- [54] The ATLAS Collaboration, *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*, ATLAS-CONF-2012-042, CERN, Geneva, Mar, 2012.
- [55] The ATLAS Collaboration, *Muon reconstruction efficiency in reprocessed 2010 LHC proton-proton collision data recorded with the ATLAS detector*, ATLAS-CONF-2011-063, CERN, Geneva, Apr, 2011.
- [56] The ATLAS Collaboration, *Determination of the muon reconstruction efficiency in ATLAS at the Z resonance in proton-proton collisions at  $\sqrt{s}=7$  TeV*, ATLAS-CONF-2011-008, CERN, Geneva, Feb, 2011.
- [57] The ATLAS Collaboration, *Muon Momentum Resolution in First Pass Reconstruction of pp Collision Data Recorded by ATLAS in 2010*, ATLAS-CONF-2011-046, CERN, Geneva, Mar, 2011.
- [58] The ATLAS Collaboration, *Expected electron performance in the ATLAS experiment*, ATL-PHYS-PUB-2011-006, CERN, Geneva, Apr, 2011.
- [59] The ATLAS Collaboration, *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, The European Physical Journal C 72 (2012) pp. 1–46.
- [60] W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr, 2008.

- [61] The ATLAS Collaboration, D. Tsionou, *Electron and Photon Performance Measurements with the ATLAS Detector*, ATL-PHYS-PROC-2012-234, CERN, Geneva, Oct, 2012.
- [62] S. Ellis, J. Huston, K. Hatakeyama, P. Loch and M. Tonnesmann, *Jets in hadron-hadron collisions*, Prog.Part.Nucl.Phys. 60 (2008) pp. 484–551, arXiv:0712.2447 [hep-ph].
- [63] M. Cacciari, G. P. Salam and G. Soyez, *The anti- $k$   $t$  jet clustering algorithm*, Journal of High Energy Physics 2008 (2008), no. 04 p. 063.
- [64] The ATLAS Collaboration, *Selection of jets produced in proton-proton collisions with the ATLAS detector using 2011 data*, ATLAS-CONF-2012-020, CERN, Geneva, Mar, 2012.
- [65] P Loch et al., *Pile-up corrections for jets from proton-proton collisions at  $\sqrt{s} = 7$  TeV in ATLAS in 2011*, ATLAS-COM-CONF-2012-078, CERN, Geneva, May, 2012.
- [66] P. Giovannini and the ATLAS Liquid Argon Calorimeter Group, *Local hadron calibration with ATLAS*, Journal of Physics: Conference Series 293 (2011), no. 1 p. 012057.
- [67] S Adomeit et al., *Jet energy scale and its systematic uncertainty in proton-proton collisions at  $\sqrt{s}=7$  TeV with ATLAS 2011 data*, ATLAS-COM-CONF-2012-171, CERN, Geneva, Aug, 2012.
- [68] The ATLAS Collaboration, *Commissioning of the ATLAS high-performance  $b$ -tagging algorithms in the 7 TeV collision data*, ATLAS-CONF-2011-102, CERN, Geneva, Jul, 2011.
- [69] The ATLAS Collaboration, *Measurement of the Mistag Rate with 5  $\text{fb}^1$  of Data Collected by the ATLAS Detector*, ATLAS-CONF-2012-040, CERN, Geneva, Mar, 2012.
- [70] The ATLAS Collaboration, *Measuring the  $b$ -tag efficiency in a top-pair sample with 4.7  $\text{fb}^{-1}$  of data from the ATLAS detector*, ATLAS-CONF-2012-097, CERN, Geneva, Jul, 2012.
- [71] The ATLAS Collaboration, *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons*, ATLAS-CONF-2011-077, CERN, Geneva, May, 2011.

- [72] The ATLAS Collaboration, *Determination of the tau energy scale and the associated systematic uncertainty in proton-proton collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector at the LHC in 2011*, ATLAS-CONF-2012-054, CERN, Geneva, Jun, 2012.
- [73] The ATLAS Collaboration, *Performance of the Reconstruction and Identification of Hadronic Tau Decays with ATLAS*, ATLAS-CONF-2011-152, CERN, Geneva, Nov, 2011.
- [74] The ATLAS Collaboration, *Performance of the Reconstruction and Identification of Hadronic Tau Decays in ATLAS with 2011 Data*, ATLAS-CONF-2012-142, CERN, Geneva, Oct, 2012.
- [75] The ATLAS Collaboration, *Performance of missing transverse momentum reconstruction in proton-proton collisions at  $\sqrt{s} = 7$  TeV with ATLAS*, The European Physical Journal C 72 (2012) pp. 1–35.
- [76] A. L. Read, *Presentation of search results: the  $CL_s$  technique*, Journal of Physics G: Nuclear and Particle Physics 28 (2002), no. 10 p. 2693.
- [77] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, The European Physical Journal C 71 (2011) pp. 1–19.
- [78] The ATLAS Collaboration, *Frequentist Limit Recommendation*, Draft 1.6 (2011) pp. 1–9.
- [79] J. Conway, *Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra*, in *PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments*, 2011.
- [80] P. Stoica and N. Arye, *MUSIC, maximum likelihood, and Cramer-Rao bound*, Acoustics, Speech and Signal Processing, IEEE Transactions on 37 (1989) pp. 720–741.
- [81] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, The Annals of Mathematical Statistics 9 (1938), no. 1 pp. pp. 60–62.
- [82] A. Wald, *Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large*, Transactions of the American Mathematical Society 54 (1943), no. 3 pp. pp. 426–482.

- [83] K. Cranmer, *Combined searches for the Higgs boson with ATLAS and CMS*, in *PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments*, 2011.
- [84] G. Cowan, *Use of the profile likelihood function in searches for new physics*, in *PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments*, 2011.
- [85] A L Read, *Modified frequentist analysis of search results (the  $CL_s$  method)*, 2000.
- [86] The ATLAS Collaboration, *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) p. 070, [arXiv:1206.5971](https://arxiv.org/abs/1206.5971) [hep-ex].
- [87] D Alvarez et al., *Search for the Standard Model Higgs boson in  $H \rightarrow \tau\tau$  decays in proton-proton collisions with the ATLAS detector*, ATLAS-COM-CONF-2012-196, CERN, Geneva, Nov, 2012.
- [88] K. Lannon, F. Margaroli and C. Neu, *Measurements of the Production, Decay and Properties of the Top Quark: A Review*, Eur.Phys.J. C72 (2012) p. 2120, [arXiv:1201.5873](https://arxiv.org/abs/1201.5873) [hep-ex].
- [89] The ATLAS Collaboration, *Luminosity plots of the ATLAS detector*, [https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults#2011\\_pp\\_Collisions](https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults#2011_pp_Collisions), 2012.
- [90] The ATLAS Collaboration, *Trigger naming convention for the ATLAS detector*, <https://twiki.cern.ch/twiki/bin/viewauth/Atlas/TriggerMenuConvention>.
- [91] The ATLAS Collaboration, *TrigMuGirl as Muon Event Filter algorithm*, <https://twiki.cern.ch/twiki/bin/viewauth/Atlas/TrigMuonEFMuGirlMerge>.
- [92] Andy Buckley et al., *General-purpose event generators for LHC physics*, Physics Reports 504 (2011), no. 5 pp. 145 – 233.
- [93] The ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, The European Physical Journal C 70 (2010) pp. 823–874.
- [94] M. A. Dobbs et. al, *Les Houches Guidebook to Monte Carlo Generators for Hadron Collider Physics*, 2004.

- [95] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP 0307 (2003) p. 001, [hep-ph/0206293](#).
- [96] 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*, [hep-ph/0405247](#).
- [97] S. Frixione, F. Stoeckli, P. Torrielli, B. R. Webber and C. D. White, *The MCaNLO 4.0 Event Generator*, [arXiv:1010.0819](#) [hep-ph].
- [98] T. Binoth, M. Ciccolini, N. Kauer and M. Kramer, *Gluon-induced W-boson pair production at the LHC*, JHEP 0612 (2006) p. 046, [hep-ph/0611170](#).
- [99] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, Journal of High Energy Physics 2007 (2007), no. 11 p. 070.
- [100] Pavel M. Nadolsky et al., *Implications of CTEQ global analysis for collider observables*, Phys.Rev. D78 (2008) p. 013004, [arXiv:0802.0007](#) [hep-ph].
- [101] A. Martin, W. Stirling, R. Thorne and G. Watt, *Parton distributions for the LHC*, Eur.Phys.J. C63 (2009) pp. 189–285, [arXiv:0901.0002](#) [hep-ph].
- [102] G. Corcella et al., *HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP 0101 (2001) p. 010, [hep-ph/0011363](#).
- [103] T. Sjöstrand, S. Mrenna and P. Skands, *PYTHIA 6.4 physics and manual*, Journal of High Energy Physics 2006 (2006), no. 05 p. 026.
- [104] Johan 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) pp. 473–500, [arXiv:0706.2569](#) [hep-ph].
- [105] J. Butterworth, J. R. Forshaw and M. Seymour, *Multiparton interactions in photoproduction at HERA*, Z.Phys. C72 (1996) pp. 637–646, [hep-ph/9601371](#).
- [106] S. Jadach, J. H. Kühn and Z. Was, *TAUOLA - a library of Monte Carlo programs to simulate decays of polarized leptons*, Computer Physics Communications 64 (1991), no. 2 pp. 275 – 299.

- [107] E. Barberio and Z. Was, *PHOTOS - a universal Monte Carlo for QED radiative corrections: version 2.0*, Computer Physics Communications 79 (1994), no. 2 pp. 291 – 308.
- [108] S. Agostinelli et al., *Geant4 - a simulation toolkit*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 506 (2003), no. 3 pp. 250 – 303.
- [109] D. Costanzo et al., *The Geant4-Based Simulation Software of the ATLAS Detector*, in *Nuclear Science Symposium Conference Record, 2006. IEEE*, vol. 1, pp. 5 –11, 29 2006-nov. 1, 2006.
- [110] G. Bozzi, S. Catani, D. de Florian and M. Grazzini, *Transverse-momentum resummation and the spectrum of the Higgs boson at the LHC*, Nucl.Phys. B737 (2006) pp. 73–120, [hep-ph/0508068](#).
- [111] A. Djouadi, J. Kalinowski and M. Spira, *HDECAY: A Program for Higgs boson decays in the standard model and its supersymmetric extension*, Comput.Phys.Commun. 108 (1998) pp. 56–74, [hep-ph/9704448](#).
- [112] S. Catani, L. Cieri, G. Ferrera, D. de Florian and M. Grazzini, *Vector boson production at hadron colliders: A Fully exclusive QCD calculation at NNLO*, Phys.Rev.Lett. 103 (2009) p. 082001, [arXiv:0903.2120](#) [hep-ph].
- [113] R. Gavin, Y. Li, F. Petriello and S. Quackenbush, *FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order*, Comput.Phys.Commun. 182 (2011) pp. 2388–2403, [arXiv:1011.3540](#) [hep-ph].
- [114] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, Phys.Rev. D83 (2011) p. 091503, [arXiv:1103.2792](#) [hep-ph].
- [115] N. Kidonakis, *NNLL resummation for s-channel single top quark production*, Phys.Rev. D81 (2010) p. 054028, [arXiv:1001.5034](#) [hep-ph].
- [116] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-*, Phys.Rev. D82 (2010) p. 054018, [arXiv:1005.4451](#) [hep-ph].
- [117] J. M. Campbell and R. Ellis, *MCFM for the Tevatron and the LHC*, Nucl.Phys.Proc.Suppl. 205-206 (2010) pp. 10–15, [arXiv:1007.3492](#) [hep-ph].

- [118] M. Aliev et al., *HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR*, Comput.Phys.Commun. 182 (2011) pp. 1034–1046, arXiv:1007.1327 [hep-ph].
- [119] S Banerjee et al., *Re-optimized Search for Standard Model  $H \rightarrow \tau_{lep} + \tau_{had}$  with the ATLAS Detector in 7 TeV Proton-Proton Collisions*, ATL-COM-PHYS-2012-1087, CERN, Geneva, Jul, 2012.
- [120] The ATLAS Collaboration, *Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of  $\sqrt{s} = 7$  TeV*, 2010.
- [121] The ATLAS Collaboration, *Performance of the ATLAS Electron and Photon Trigger in  $p$ - $p$  Collisions at  $\sqrt{s} = 7$  TeV in 2011*, ATLAS-CONF-2012-048, CERN, Geneva, May, 2012.
- [122] A. Elagin, P. Murat, A. Pranko and A. Safonov, *A new mass reconstruction technique for resonances decaying to  $\tau$ s*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 654 (2011), no. 1 pp. 481 – 489.
- [123] J. Friedman, T. Hastie and R. Tibshirani, *The elements of statistical learning*, vol. 1. Springer Series in Statistics, 2001.
- [124] J. R. Quinlan, *Induction of decision trees*, Machine learning 1 (1986), no. 1 pp. 81–106.
- [125] A. Hoecker et al., *TMVA - Toolkit for Multivariate Data Analysis*, ArXiv Physics e-prints (2007) physics/0.
- [126] Byron P. Roe et al., *Boosted decision trees as an alternative to artificial neural networks for particle identification*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 543 (2005), no. 23 pp. 577 – 584.
- [127] R. A. Fisher, *The use of multiple measurements in taxonomic problems*, Annals of Human Genetics 7 (1936), no. 2 pp. 179–188.
- [128] Y. Freund and R. E. Schapire, *A decision-theoretic generalization of on-line learning and an application to boosting*, Journal of computer and system sciences 55 (1997), no. 1 pp. 119–139.

- [129] J. R. Quinlan, *Simplifying decision trees*, International journal of man-machine studies 27 (1987), no. 3 pp. 221–234.
- [130] J. H. Friedman, *Stochastic gradient boosting*, Computational Statistics & Data Analysis 38 (2002), no. 4 pp. 367–378.
- [131] L. Moneta, K. Cranmer, G. Schott and W. Verkerke, *The RooStats project*, in *Proceedings of the 13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research. February 22-27, 2010, Jaipur, p.57*, 2010. [arXiv:1009.1003](https://arxiv.org/abs/1009.1003) [physics.data-an].
- [132] I. W. Stewart and F. J. Tackmann, *Theory Uncertainties for Higgs and Other Searches Using Jet Bins*, Phys.Rev. D85 (2012) p. 034011, [arXiv:1107.2117](https://arxiv.org/abs/1107.2117) [hep-ph].
- [133] S. Catani and M. Grazzini, *HNNLO: A Monte Carlo program to compute Higgs boson production at hadron colliders*, PoS RADCOR2007 (2007) p. 046, [arXiv:0802.1410](https://arxiv.org/abs/0802.1410) [hep-ph].
- [134] R. Aben et al., *Monte Carlo generators and theoretical aspects for the search of the Higgs boson in the  $H \rightarrow WW \rightarrow l\nu l\nu$  decay mode using  $4.7 \text{ fb}^{-1}$  of data collected with the ATLAS detector at  $\sqrt{s} = 7 \text{ TeV}$* , ATL-COM-PHYS-2011-1726, CERN, Geneva, Dec, 2011.
- [135] Michiel Botje et al., *The PDF4LHC Working Group Interim Recommendations*, [arXiv:1101.0538](https://arxiv.org/abs/1101.0538) [hep-ph].
- [136] The ATLAS Collaboration, *Improved Luminosity Determination in  $pp$  Collisions at  $\sqrt{s} = 7 \text{ TeV}$  using the ATLAS Detector at the LHC*, ATLAS-CONF-2012-080, CERN, Geneva, Jul, 2012.
- [137] The ATLAS Collaboration, *ATLAS MC generator production scripts*, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/AlpgenProductionScripts>.
- [138] A. D. Martin, R. Roberts, W. Stirling and R. Thorne, *MRST2001: Partons and  $\alpha_s$  from precise deep inelastic scattering and Tevatron jet data*, Eur.Phys.J. C23 (2002) pp. 73–87, [hep-ph/0110215](https://arxiv.org/abs/hep-ph/0110215).
- [139] The ATLAS Collaboration, *Performance of the ATLAS muon trigger in 2011*, ATLAS-CONF-2012-099, CERN, Geneva, Jul, 2012.
- [140] The ATLAS Collaboration, *Performance of the ATLAS tau trigger in 2011*, ATLAS-CONF-2013-006, CERN, Geneva, Jan, 2013.

---

## List of Acronyms

---

SM	Standard Model
QCD	Quantum Chromodynamics
QED	Quantum Electrodynamics
GSW	Glashow-Salam-Weinberg
CKM	Cabibbo-Kobayashi-Maskawa
BR	Branching Ratio
CL	Confidence Level
ggF	gluon-gluon Fusion
VBF	Vector-boson Fusion
VH	Associated Higgs boson production
ttH	top pair Fusion
LHC	Large Hadron Collider
LEP	Large Electron-Positron Collider
HI	Heavy Ion
ATLAS	A Toroidal LHC ApparatuS
CMS	Compact Muon Solenoid
LHCb	Large Hadron Collider beauty
ALICE	A Large Ion Collider Experiment
ID	Inner Detector
EM	Electromagnetic
MS	Muon Spectrometer
LUCID	LUMinosity measurement using a Cerenkov Integrating Detector
ALFA	Absolute Luminosity For ATLAS
TRT	Transition Radiation Tracker
SCT	Semiconductor Tracker
EMC	Electromagnetic Calorimeter
LAr	Liquid Argon
TileCal	Tile Calorimeter
HEC	Hadronic End-cap Calorimeter
FCal	Forward Calorimeter
MDT	Monitored Drift Tube
CSC	Cathode Strip Chamber

---

RPC	Resistive Plate Chamber
TGC	Thin Gap Chamber
L1	Level 1 Trigger
L2	Level 2 Trigger
EF	Event Filter
RoF	Region of Interest
BDT	Boosted Decision Tree
JVF	Jet Vertex Fraction
SR	Signal Region
SA muon	Stand-Alone muon
ST muon	Segment Tagged muon
CB muon	Combined muon
EM Scale	Electromagnetic Energy Scale
DAF	Deterministic Annealing Filter
JES	Jet Energy Scale
JVF	Jet Vertex Fraction
LCW	Local Cluster Weighting
TES	Tau Energy Scale
MLE	Maximum-Likelihood Estimator
pdf	Probability Density Function
LR	Likelihood-Ratio
MVUE	Minimum-Variance Unbiased Estimator
LLR	Log-Likelihood Ratio
NP	Nuisance Parameter
d.o.f	Degree of Freedom
MMC	Missing Mass Calculator
SLT	Single Lepton Trigger
LTT	Lepton + tau Trigger
FF	Fake Factor
MLE	Maximum-Likelihood Estimator
PDF	Parton Distribution Function
BR	Branching Ratio
OS	Opposite Sign Charges
SS	Same Sign Charges
CR	Control Region
SR	Signal Region
MET	Missing Transverse Energy
mT	Transverse Mass
MVA	Multivariate Analysis
ANN	Artificial Neural Network
TMVA	Toolkit for a Multivariate Data Analysis

# List of Figures

1.1	The potential $V(\phi)$ as projection in one dimension. . . . .	10
1.2	Production cross sections for different center-of-mass energies. . . . .	14
1.3	MSTW parton distribution functions. . . . .	15
1.4	Higgs boson branching ratios for masses between 80 and 1000 GeV. . . . .	19
1.5	Higgs boson production cross sections times branching-ratios of its various decay modes at $\sqrt{s} = 7$ TeV. . . . .	20
1.6	Theoretical and experimental constraints on the Higgs boson mass. . . . .	22
1.7	Observed and expected $p_0$ values for ATLAS and CMS. . . . .	23
1.8	Higgs boson signal strength derived by the ATLAS and CMS experiments. . .	24
2.1	Cut-away view of the entire ATLAS detector. . . . .	27
2.2	Cut-away view of the ATLAS Inner detector (ID) [45]. . . . .	28
2.3	The ATLAS calorimeter design, taken from Ref. [45]. . . . .	30
3.1	The jet BDT score for 1- and 3-prong $\tau_{\text{had-vis}}$ candidates. . . . .	44
3.2	Identification efficiencies for 1- and 3-prong $\tau_{\text{had-vis}}$ candidates. . . . .	45
4.1	Illustration of the p-value derivation and its corresponding Z. . . . .	50
4.2	Median of $q_\mu$ and its p-value with strength parameter $\mu'$ . . . . .	52
5.1	Tree level Feynman diagrams for single top production. . . . .	60
5.2	Transverse momentum of the muon after preselection. . . . .	61
5.3	Kinematic distributions for the $e + \tau_{\text{had}}$ channel after preselection. . . . .	69
5.4	Kinematic distributions for the $\mu + \tau_{\text{had}}$ channel after preselection. . . . .	70
5.5	Probability distribution functions used in the MMC algorithm. . . . .	71
5.6	Comparison of distributions for the MMC and collinear approach. . . . .	72
5.7	Comparison of the OS and SS distributions in the QCD CR. . . . .	74
5.8	$r_{\text{QCD}}$ for several $etcone/p_T$ and $ptcone/p_T$ values. . . . .	75
5.9	W CR in the $\mu + \tau_{\text{had}}$ and $e + \tau_{\text{had}}$ channel, prior to scaling. . . . .	76
5.10	Scheme of the embedding procedure. . . . .	79

5.11	$ \Delta\eta_{jj} $ distributions before and after data-driven re-weighting. . . . .	80
5.12	Distributions of the $ \Delta\eta_{jj} $ re-weighting for the VBF-filtered sample. . . . .	81
5.13	FF dependence on $p_T^\tau$ for 1 prong and 3 prong $\tau_{\text{had}}$ candidates . . . . .	82
6.1	$m_{\tau\tau}^{\text{MMC}}$ distributions for the past 1J and VBF category. . . . .	84
6.2	Expected exclusion limits for the VBF-optimization as function of $p_T^{j_2}$ and $\Delta\eta_{j_1j_2}$ . . . . .	85
6.3	Expected exclusion limits for the VBF-optimization as function of $p_T^{j_1}$ and $m_{j_1j_2}$ . . . . .	86
6.4	$m_{\tau\tau}^{\text{MMC}}$ distributions for old (left) and optimized (right) VBF classification. . . . .	87
6.5	$m_{\tau\tau}^{\text{MMC}}$ distributions and RMS for the primer 1J and new Boosted category. . . . .	88
6.6	Correlation between $\Delta R^{\ell,\tau}$ and the $p_T$ of the lepton- $\tau$ system. . . . .	89
6.7	$m_{\tau\tau}^{\text{MMC}}$ distributions for the 2011 re-analysis categories. . . . .	94
7.1	ROC-curves for various multivariate and linear classifiers. . . . .	96
7.2	Illustration of a single tree classifier. . . . .	97
7.3	ROC-curves for various BDT configurations. . . . .	99
7.4	Significance for BDT distributions with OS events only or additional SS events. . . . .	102
7.5	Expected exclusion limits for several BDT configurations. . . . .	105
7.6	BDT score distributions for <i>embedded</i> and VBF-filtered samples. . . . .	107
7.7	BDT distributions with and without cross evaluation. . . . .	108
7.8	BDT score distributions of the signal region in all categories. . . . .	109
7.9	W and top CR distributions for the Boosted and VBF category. . . . .	110
7.10	$Z \rightarrow \ell\ell$ , W and top CR distributions for the 0J category. . . . .	111
7.11	Scatter plot of $\Delta\eta_{j_1,j_2}$ against the BDT score in the VBF category. . . . .	112
7.12	Various exclusion limits to optimize the BDT-cut threshold. . . . .	113
7.13	BDT score distributions of the signal region without $m_{\tau\tau}^{\text{MMC}}$ in training. . . . .	114
7.14	$m_{\tau\tau}^{\text{MMC}}$ mass distributions obtained after a cut on the BDT score. . . . .	115
8.1	Cut based exclusion limits for $110 \text{ GeV} \leq m_H \leq 145 \text{ GeV}$ . . . . .	124
8.2	Exclusion limits for the BDT approaches and $110 \text{ GeV} \leq m_H \leq 145 \text{ GeV}$ . . . . .	125
8.3	Limit deviation after an introduced shape dependence. . . . .	126
8.4	Limit comparison between cut-based and BDT approaches for the VBF and non-VBF categories. . . . .	128
8.5	Limits on the cut-based and BDT approaches for $110 \text{ GeV} \leq m_H \leq 145 \text{ GeV}$ . . . . .	130
B.1	Detailed illustration of the object preselection. . . . .	140
C.1	The distributions of $m_{\tau\tau}^{\text{MMC}}$ , $m_T$ and sphericity for the 0J category. . . . .	142

C.2	The distributions of $E_T^{\text{miss}}$ , $p_T^H$ and $\sum p_T$ for the 0J category. . . . .	143
C.3	The distributions of the $E_T^{\text{miss}} - \phi$ centrality, $\Delta R$ and $\sum \Delta\phi$ for the 0J category.	144
C.4	The distributions of the $p_T$ asymmetry, $x_l$ and $x_h$ for the 0J category. . . . .	145
C.5	The distributions of $m_{\tau\tau}^{\text{MMC}}$ , $m_T$ , $E_T^{\text{miss}}$ , $p_T^H$ , $\sum p_T$ and $E_T^{\text{miss}} - \phi$ centrality for the Boosted category. . . . .	146
C.6	The distributions of $\Delta R$ , $\sum \Delta\phi$ , $p_T$ asymmetry, $x_l$ , $x_h$ and the sphericity for the Boosted category. . . . .	147
C.7	The distributions of $m_{\tau\tau}^{\text{MMC}}$ , $m_T$ , $E_T^{\text{miss}}$ , $p_T^H$ , $\sum p_T$ and $E_T^{\text{miss}} - \phi$ centrality for the VBF category. . . . .	148
C.8	The distributions of $\Delta R$ , $\sum \Delta\phi$ , $p_T$ asymmetry, $x_l$ , $x_h$ and $m_{j_1, j_2}$ for the VBF category. . . . .	149
C.9	lepton centrality, $\tau$ centrality, $\eta_{j_1} \cdot \eta_{j_2}$ , $\Delta\eta_{j_1, j_2}$ and sphericity in the VBF category	150
D.1	W and top CRs for the Boosted and VBF category in the <i>BDT-cut</i> approach.	151
D.2	$Z \rightarrow \ell\ell$ , W and top CRs for the 0J category in the <i>BDT-cut</i> approach. . . . .	152
E.1	Systematic variation of the jet and $\tau$ energy scale for the 0J category. . . . .	153
E.2	Embedding, electron and muon energy scale variation for the 0J category. . . . .	154
E.3	Systematic variation of the electron resolution, muon ID efficiency, $E_T^{\text{miss}}$ energy scale and resolution for the 0J category. . . . .	155
E.4	Systematic variation of the jet and $\tau$ energy scale, muon and electron energy scale, muon ID efficiency and electron resolution for the Boosted category. . . . .	156
E.5	Systematic variation of the $E_T^{\text{miss}}$ energy scale and resolution and the embedding for the Boosted category. Moreover, systematic variations of the jet and $\tau$ energy scale for the VBF category. . . . .	157
E.6	Systematic variation of the muon and electron energy scale, muon ID efficiency, electron resolution, $E_T^{\text{miss}}$ energy scale and resolution for the VBF category. . . . .	158



# List of Tables

1.1	Quarks and leptons of the Standard Model. . . . .	4
1.2	Gauge bosons of the Standard Model with their corresponding charge and mass. . . . .	5
1.3	Fermions with their corresponding quantum numbers. . . . .	9
5.1	The specific triggers at event filter (EF) level used in this thesis. . . . .	62
5.2	$\sigma \times \text{BR}$ of the $H \rightarrow \tau_{\text{lep}} \tau_{\text{had}}$ signal and background processes. . . . .	65
5.3	Numbers of events passing the preselection. . . . .	68
6.1	Event categorization applied in the cut-based re-analysis of the 2011 dataset. . . . .	92
7.1	BDT configurations applied in this analysis. . . . .	99
7.2	Classification of events in the BDT approach. . . . .	101
7.3	Possible input variables for the BDT approach. . . . .	104
7.4	Details on various BDT configurations used to optimize it. . . . .	106
8.1	Uncertainties of the Higgs boson production via gluon fusion . . . . .	119
8.2	Systematic uncertainties for ALPGEN $Z$ +jets MC samples. . . . .	121
8.3	Summary of systematic uncertainties on the background estimation methods. . . . .	123
F.1	Combined systematic uncertainties for the 0 jet category in the $e + \tau_{\text{had}}$ channel. . . . .	159
F.2	Combined systematic uncertainties for the 0 jet category in the $\mu + \tau_{\text{had}}$ channel. . . . .	159
F.3	Combined systematic uncertainties for the 1 jet category in the $e + \tau_{\text{had}}$ channel. . . . .	160
F.4	Combined systematic uncertainties for the 1 jet category in the $\mu + \tau_{\text{had}}$ channel. . . . .	160
F.5	Combined systematic uncertainties for the Boosted category. . . . .	161
F.6	Combined systematic uncertainties for the VBF category. . . . .	161
G.1	Combined systematic uncertainties for the 0 jet category in the $e + \tau_{\text{had}}$ channel. . . . .	163
G.2	Combined systematic uncertainties for the 0 jet category in the $\mu + \tau_{\text{had}}$ channel. . . . .	163
G.3	Combined systematic uncertainties for the Boosted category. . . . .	164
G.4	Combined systematic uncertainties for the VBF category. . . . .	164
G.5	Combined systematic uncertainties for the 0 jet category in the $e + \tau_{\text{had}}$ channel. . . . .	165

G.6	Combined systematic uncertainties for the 0 jet category in the $\mu + \tau_{\text{had}}$ channel.	165
G.7	Combined systematic uncertainties for the Boosted category. . . . .	166
G.8	Combined systematic uncertainties for the VBF category. . . . .	166

---

# Acknowledgement

---

First of all I want to thank Professor Karl Jakobs for his motivation and support during my thesis. He gave me the opportunity to work in his excellent group and contribute to the exciting search for the Higgs boson at the ATLAS experiment.

Many thanks go to Romain Madar and Nils Ruthmann, with whom i participated in the analysis and who assisted and guided me wherever i had questions or problems. Moreover, it was a great pleasure to work with them and receive some of their knowledge. Aside from working, I had a great time in the office with Nils, Sascha and Johanna with a lot of entertaining discussions that shortened the days. However, they took their time to explain everything I sought to know.

The conversations with Romain were very helpful and illuminating, yet he was never hesitant to have some fun in between. We sat together for many hours breeding over my thesis despite of other important tasks he had to fulfill. However, he had the amusement to read my first *golden* limit computation chapter which remains to be interpreted by forthcoming generations.

Nils contributed to my thesis with invaluable suggestions and improvements. Due to him, I was able to perform the presented studies with all their technical and theoretical facets.

Also I want to express my gratefulness to all the talented and enjoyable people in the group who were always there for discussions and diversion.

Above all, I want to thank my family for their unconditional love and all they rendered possible. They encouraged me by all means and enabled me to focus on my studies and travel around the world without restrictions. Last but not least, I want to thank Katharina for her patience and love. There were many weekends and evenings where she had to share me with my work and did not complain at all but strengthened me without exception.



## EIDESSTATTLICHE ERKLÄRUNG

Hiermit versichere ich, die vorliegende Arbeit selbständig verfasst zu haben. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen entnommen sind, habe ich als solche kenntlich gemacht. Die Arbeit wurde noch nicht anderweitig als Diplomarbeit eingereicht

Freiburg, den 10. April 2013

Helge Haß

