BERGISCHE
UNIVERSITÄT
WUPPERTAL

FACHBEREICH C
FACHGRUPPE PHYSIK

SEARCH FOR HIGGS BOSONS IN TWO-HIGGS-DOUBLET MODELS IN THE $h/H \to WW^{(*)} \to e\nu\mu\nu$ CHANNEL WITH THE ATLAS DETECTOR

Dissertation

von

Dipl.-Phys. Simon Köhlmann
aus Cuxhaven

Wuppertal, den 10. Juni 2014

Die Dissertation kann wie folgt zitiert werden:

urn:nbn:de:hbz:468-20140616-114252-7
[http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3Ahbz%3A468-20140616-114252-7]

Referent: Prof. Dr. Wolfgang Wagner, Fachbereich C - Experimentelle Teilchenphysik, Bergische Universität Wuppertal (DE)

Korreferent: Prof. Dr. Marumi Kado, Universite de Paris-Sud 11 (FR)

Tag der mündlichen Prüfung: 06. Juni 2014

Contents

Chapter 1

Introduction

Through scientific studies, physicists strive to understand the properties and dynamics of matter from the very small scale of subatomic particles up to the very large scale of the entire universe. At smallest scales, this continuing search for answers and for new phenomena leads to what is called particle physics. The Standard Model of particle physics (SM), which was developed in the 1960s and 1970s, describes the elementary particles of matter and the electromagnetic, weak and strong forces, all of them mediated by the exchange of gauge bosons. In order to test the SM and search for physics phenomena beyond the SM a particle accelerator, the Large Hadron Collider (LHC) [1], was built by the European Organisation for Nuclear Research (CERN). One of the most important goals of the LHC and the two multi-purpose particle detector experiments ATLAS [2] and CMS [3] is to prove or disprove the existence of the Higgs boson to gain insight into the Higgs field as the origin of electroweak symmetry breaking. The Higgs boson is the last missing entity of the SM.

On the 4th of July 2012 the ATLAS and CMS collaborations announced the observation of a Higgs-like boson at a mass of approximately 125 GeV [4, 5]. To date, all measurements concerning the production rates, the branching ratios and kinematic distributions are compatible with this particle being the Higgs boson predicted by the SM. However, the SM falls short of being a complete theory of fundamental interactions. Among other shortcomings the SM cannot explain the observed baryon asymmetry in the universe and does not provide any viable dark matter candidate as required by astronomical observations [6].

Alternative models are conceived to address and solve one or more of these problems. Since the discovery of the Higgs boson Two-Higgs-Doublet models (2HDMs) have attracted much attention in phenomenological studies [7–11], which provides a strong incentive for dedicated experimental investigations in this direction. While the SM Higgs sector is based on only one complex Higgs doublet, 2HDMs feature an additional second complex Higgs doublet [12] and are one of the simplest ways to extend the scalar sector of the SM. The two complex Higgs doublets give rise to five Higgs bosons: two CP-even scalars h and H, one pseudoscalar A (CP-odd), and two charged Higgs bosons $H^{\pm}$.

Two main categories of 2HDMs are distinguished depending on the different couplings of the two scalar fields h and H to fermions and weak gauge bosons. In the type-I model, all quarks couple to just one of the Higgs doublets. In the type-II model, the up-type quarks couple to one Higgs doublet and the down-type quarks couple to the other doublet. The observed rate of the Higgs boson at 125 GeV in various channels provides already constraints on 2HDMs [13], mainly reducing the parameter space of the type-II model. Searches for generic 2HDMs have been performed by the CDF collaboration at the Tevatron [14, 15].

The aim of the analysis presented in this thesis is to investigate the possibility that the boson observed by the ATLAS and CMS experiments at a mass of 125 GeV originates from a Higgs boson that is part of a 2HDM. In particular, it is assumed that the observed particle is the low mass CP-even Higgs boson h of a 2HDM. The analysis considers type-I and type-II models by searching for additional signal contributions arising from the higher mass CP-even boson H. The Higgs boson is reconstructed in the decay to a W boson pair. The W bosons cannot be observed directly with the detector because they decay immediately after production into leptons or hadrons. Both CP-even Higgs bosons are reconstructed in the dilepton decay channel, in which the W boson pair decays into two isolated opposite sign leptons and two neutrinos. The main backgrounds are non-resonant $W^{+}W^{-}$ production and top-quark-antiquark pair production.

The following two production modes of Higgs bosons are considered: the gluon-fusion process and the vector-boson-fusion (VBF) process. To be sensitive to both production mechanisms the analysis considers two different final states. In the first channel, two charged leptons and large missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$ are required (0-jet channel), and in the second channel, which is sensitive to the VBF process, two jets with high transverse momentum (2-jet channel) are reconstructed in addition. As a result, the selection of candidate events is based on the identification of electrons, muons, and jets and on the reconstruction of the $E_{\mathrm{T}}^{\mathrm{miss}}$.

In this thesis, data recorded with the ATLAS detector in 2012 are analysed. During the data taking period the LHC operated at a centre-of-mass energy of 8 TeV and the analysed data set corresponds to an integrated luminosity of 13 fb^{-1}. Artificial neural networks (NN) are used to enhance the sensitivity by combining the information contained in various kinematic variables. In the mass interval of 135 GeV $< m_H <$ 300 GeV no significant excess of events originating from the production of a heavy Higgs boson is observed. Expected and observed exclusion limits are calculated using a frequentist approach based on pseudo experiments, and contours of excluded parameter regions of the considered 2HDMs are determined.

Chapter 2

Theoretical Overview

There are four known fundamental forces in nature: electromagnetism, the weak and strong nuclear forces and gravity. The three former forces are described by the SM, while the latter is described by the theory of general relativity which describes large-scale physical phenomena like the distribution of matter in a galaxy. The SM, on the contrary, describes the small-scale structure of matter, which is governed by quantum mechanics, and incorporates the three non-gravitational forces.

The SM contains 61 elementary particles: 36 quarks, 12 leptons, eight gluons, two W bosons, one Z boson, one photon and one Higgs boson, which is the only scalar particle in the SM. The quarks and the leptons form the so-called matter particles, while the gluons, the W and Z bosons and the photon are force mediating particles, which carry the strong interaction, the weak interaction and the electromagnetic interaction, respectively.

The masses of the elementary particles are explained by the Brout-Englert-Higgs (BEH) mechanism [16–20]. According to this theory, particles gain mass by interacting with the Higgs field that exists throughout the universe. The Higgs field is a four-component scalar field that forms the simplest possible scalar structure to implement this mechanism: just one complex SU(2) doublet. The before mentioned Higgs boson is a quantum excitation of one of the four components of this Higgs field. For almost 50 years researchers have endeavoured to detect the Higgs boson, the last missing particle of the SM. In 2012, a new particle has been observed whose measured properties are in excellent agreement with those of a Higgs boson.

After the discovery of a Higgs boson, one of the most important open questions is now whether the observed Higgs boson is the one predicted by the SM or whether the new particle is part of a richer scalar sector. To answer this vital question two directions have been pursued. On the one hand, the properties of the observed boson are measured precisely. This comprises the spin and parity [21, 22] and the coupling strength [23–25]. On the other hand, experimentalists search for additional particles predicted in theories beyond the SM containing an extended Higgs sector. In this analysis, the latter approach is adopted. Two scenarios with natural flavour conservation are investigated, the so-called type-I and type-II models of the 2HDMs.

This chapter is organised as follows. The first part of this chapter describes the SM, its particle content and the fundamental forces which determine the interactions between these particles by mediating gauge bosons. Subsequently, the BEH mechanism is explained. The second part is about 2HDMs. A survey of the various types of 2HDMs is given, and the calculation of the production cross sections and decay rates is described.

2.1 The Standard Model of Particle Physics

The SM contains two different kinds of particles, the spin-1/2 fermions and the spin-1 vector bosons. Another elementary particle in the SM is the Higgs boson. It is massive and has no intrinsic spin. Each elementary particle has an antiparticle with complementary additive quantum numbers. All additive quantum numbers of the photon, the Z boson and the Higgs boson are zero and these bosons have therefore undistinguishable antiparticles. They are their own antiparticles. All experimentally verified elementary particles are shown in figure 2.1. The fermions are subdivided by colour charge into quarks and leptons. Quarks carry colour charge and participate in strong interactions, whereas leptons do not carry colour charge. Both, quarks and leptons, come in six flavours. The six flavours can be arranged into three generations with higher generation fermions in each row of figure 2.1 differing in mass but not in electric charge. For quarks these are the up- and down-, charm- and strange-, and top- and bottom-type quarks. For leptons these are the electron, muon, and tau with the corresponding neutrinos. All ordinary matter is made of particles of the first generation. The gauge bosons of the electromagnetic force, the weak force and the strong force – the photon γ, the $W^{\pm}$ and Z bosons, and the eight gluons g_i – are the force carriers of the fundamental interactions.

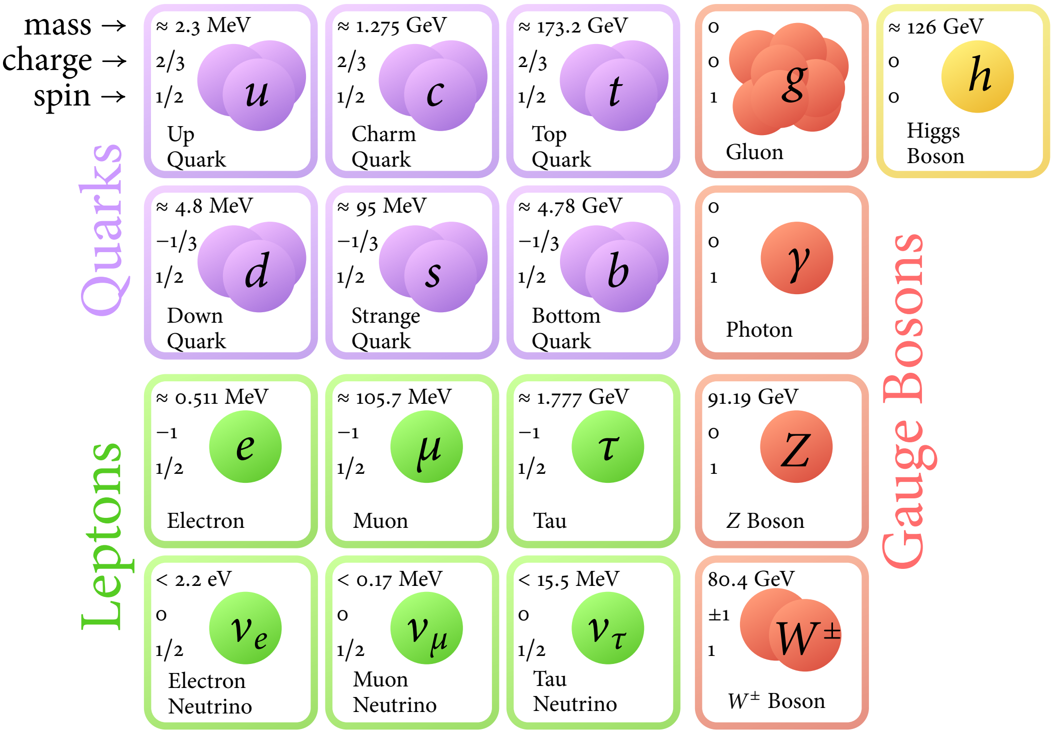


Figure 2.1: The elementary particles of the Standard Model [26]. The masses of the top and the bottom quark are updated according to [27] and [28], respectively.

While the photon and the eight gluons are massless, the $W^{\pm}$ and Z bosons acquire mass by coupling with the Higgs field. Figure 2.2 illustrates all possible interactions between fermions via gauge bosons, the self interactions of the weak bosons and the gluons, and interactions between massive particles and the Higgs boson.

While the electromagnetic current has the structure of a vector, i.e. the electromagnetic interaction is symmetric under parity transformations, the weak interaction violates parity maximally due to its $V-A$ structure, where V stands for vector and A stands for axial vector. To unify the electromagnetic and the weak forces into one theory the $V-A$ structure has to be absorbed in the parametrisation of the particles. This means that a fermion field ψ has to be decomposed into left-handed and right-handed projections

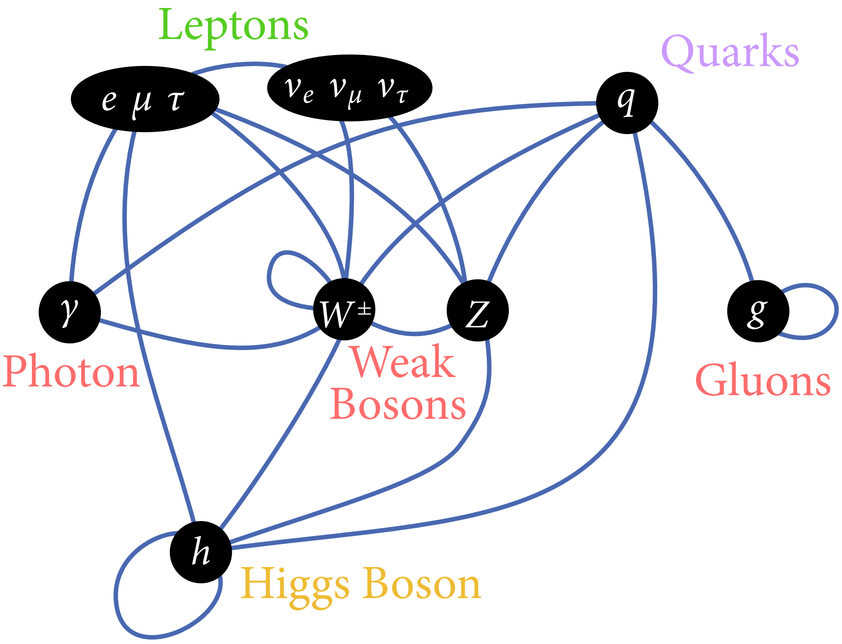


Figure 2.2: Interactions of the elementary particles of the Standard Model [29]. The neutrinos are considered to be massless in the SM, therefore interactions between the Higgs boson and the neutrinos are omitted.

$$\begin{aligned}\psi &= \frac{1}{2}\left(1-\gamma^5\right)\psi + \frac{1}{2}\left(1+\gamma^5\right)\psi \\ &= P_L\psi + P_R\psi \\ &= \psi_L + \psi_R\,,\end{aligned} \tag{2.1}$$

with

$$\gamma^5 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{2.2}$$

P_L and P_R are the left-handed and right-handed projection operators: $P_L = \left(1-\gamma^5\right)/2$ and $P_R = \left(1+\gamma^5\right)/2$. In the SM, ψ_R and ψ_L represent

$$e_R, \mu_R, \tau_R, u_R, d_R, c_R, s_R, t_R, b_R$$

and

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L, \begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L,$$

respectively.

Quantum field theory is the mathematical framework of the SM in which particles are described as excited states of underlying fields. The theory is consistent with quantum mechanics and special relativity and brings these two theories together to account for subatomic phenomena. The field equations are obtained from the Lagrangian density

$$\mathcal{L}_{\mathrm{SM}} = \mathcal{L}_{\mathrm{Gauge}} + \mathcal{L}_{\mathrm{Matter}} + \mathcal{L}_{\mathrm{Higgs}} + \mathcal{L}_{\mathrm{Yukawa}} \tag{2.3}$$

by minimising the action

$$S = \int \mathcal{L}\, \mathrm{d}^4 x\,. \tag{2.4}$$

The gauge Lagrangian $\mathcal{L}_{\mathrm{Gauge}}$ is given by a sum of independent kinetic terms and self-interaction terms of massless spin-1 gauge fields. The matter term $\mathcal{L}_{\mathrm{Matter}}$ consists of the kinetic terms of massless spin-1/2 particles, the leptons and the quarks, and their interaction among each other via spin-1 gauge bosons. The Higgs Lagrangian $\mathcal{L}_{\mathrm{Higgs}}$ and the last term, the Yukawa Lagrangian, $\mathcal{L}_{\mathrm{Yukawa}}$ describe the interaction between the Higgs field and the gauge bosons and between the Higgs field and the fermions, respectively.

The SM is constructed such that it possesses an additional special property: local gauge symmetry. A symmetry is a transformation of the fields Φ^a that leaves the action invariant

$$\delta\mathcal{L} = \sum_a \frac{\partial\mathcal{L}}{\partial\Phi^a}\, \delta\Phi^a = 0. \tag{2.5}$$

Here $\delta\Phi^a$ is an infinitesimal transformation

$$\delta\Phi^a = \varepsilon^r T_r{}^a{}_b\, \Phi^b \tag{2.6}$$

with generators T of the local group. To be a local gauge symmetry means that the continuous infinitesimal parameter ε^r is an arbitrary function of spacetime, i.e. $\varepsilon^r = \varepsilon^r(x_\mu)$. The real world is built from matter particles and local gauge symmetry requires the existence of gauge fields through which the matter particles interact with each other.

The SM is a non-Abelian gauge field theory with the symmetry group U(1) ⊗ SU(2) ⊗ SU(3) [30–32]. The corresponding charges, which are also the generators of these symmetry groups, are the weak hypercharge Y, the isospin I and the colour charge. The colour charge determines the strong interaction and can take one of three values: blue, green or red. The electric charge Q is related to the third component of the isospin I_3 and the weak hypercharge Y by the Gell-Mann-Nishijima formula [33–35]

$$Q = I_3 + \frac{1}{2} Y \,. \tag{2.7}$$

The Glashow-Weinberg-Salam model describes the electromagnetic force and the weak force in a unified electroweak theory based on the U(1) ⊗ SU(2) gauge group [30–32]. The electromagnetic and the weak gauge bosons originate from the vector-boson fields of the U(1) and SU(2) gauge groups, the B_0 and the three W_i, and arise due to the spontaneous symmetry breaking. The gauge bosons are linear combinations of B_0 and the W_i in terms of the weak mixing angle θ_W:

$$\begin{aligned} |\gamma\rangle &= \cos\theta_W |B_0\rangle + \sin\theta_W |W_1\rangle \\ |Z\rangle &= -\sin\theta_W |B_0\rangle + \cos\theta_W |W_1\rangle \\ |W^\pm\rangle &= \frac{1}{\sqrt{2}} \left(|W_2\rangle \pm i\,|W_3\rangle\right) \,. \end{aligned} \tag{2.8}$$

The strong interaction is described by quantum chromodynamics (QCD) which is a non-Abelian gauge theory based on the local symmetry group SU(3). The strong force has a unique property. Unlike all other forces the strong force does increase in strength the farther two colour-charged particles are moved away from each other. As a consequence of this feature two bound coloured particles cannot be isolated and therefore cannot be directly observed singly. This property is called confinement. Another consequence of the strength of the strong force being dependent on the distance is asymptotic freedom which means that two particles behave almost like free particles when they are close to each other [36, 37].

More information about the SM and gauge theories in general can be found in references [28, 38–40] and reference [41], respectively. Reference [42] discusses the general principles of quantum field theory and its most profound consequences.

2.1.1 The Brout-Englert-Higgs Mechanism and the Generation of Fermion Masses

The simplest implementation of the electroweak symmetry breaking introduces an extra Higgs field to the gauge theory, with the outcome of massive weak vector bosons, $W^{\pm}$ and Z, and a Higgs boson. Electroweak symmetry breaking is essential for the mass of fermions as well because in the unbroken phase, mass terms for fermions and gauge bosons are forbidden as they are not SU(2) ⊗ U(1) gauge invariant.

To explain the masses of elementary particles via the BEH mechanism, the SM introduces an SU(2) doublet

$$\phi = \begin{pmatrix} \phi^{+} \\ \phi^{0} \end{pmatrix} \tag{2.9}$$

with two complex scalar fields, the Higgs fields ϕ^{+} and ϕ^{0}. The Higgs potential is defined as follows:

$$V_{\text{Higgs}} = \frac{1}{2}\mu^{2}\phi^{\dagger}\phi + \frac{1}{4}\lambda\left(\phi^{\dagger}\phi\right)^{2} \tag{2.10}$$

with the free parameters μ and λ. As a result, the before mentioned Higgs Lagrangian becomes

$$\mathcal{L}_{\text{Higgs}} = \left(D_{\mu}\phi\right)^{\dagger}\left(D^{\mu}\phi\right) - \frac{1}{2}\mu^{2}\phi^{\dagger}\phi - \frac{1}{4}\lambda\left(\phi^{\dagger}\phi\right)^{2}, \tag{2.11}$$

where D_{μ} is the covariant derivative associated to SU(2) ⊗ U(1):

$$D_{\mu} = \partial_{\mu} + ig'\frac{1}{2}YB_{\mu} + ig\frac{1}{2}\vec{\sigma}\,\vec{W}_{\mu}\,. \tag{2.12}$$

g' and g are the gauge couplings of the U(1) and SU(2) groups, respectively. The weak hypercharge Y and the Pauli matrices $\vec{\sigma}$ are the generators of the U(1) and SU(2) groups, respectively.

The value of λ must be positive to have a well defined minimum of the energy. The value of μ^2 can be either positive or negative, but has to be chosen to be negative to obtain massive vector bosons. The Higgs potential obeys a global U(1) symmetry and its minimum is found at

$$v = \langle \phi^0 \rangle_{\text{min}} = \sqrt{\frac{-\mu^2}{\lambda}} \,. \tag{2.13}$$

v is the so-called vacuum expectation value.

After expanding the potential around the vacuum expectation value in the direction of $\mathfrak{R}\left(\phi^0\right)$, the Higgs field becomes

$$\phi = \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} , \tag{2.14}$$

where $h(x)$ represents the first order of the perturbation series. Due to the expansion, the partial derivatives in the Higgs field kinetic energy term in equation (2.11) break the local gauge invariance. By substitution of equation (2.14) for the Higgs field and using the relation $-\mu^2 = v^2\lambda$ from equation (2.13), one gets

$$\begin{aligned}\mathcal{L}_{\text{Higgs}} = &\frac{1}{2}\left(\partial_\mu h\right)\left(\partial^\mu h\right) - \frac{1}{4}\vec{W}_{\mu\nu}\vec{W}^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \\ &+ (v+h)^2\left(\frac{1}{8}g^2\vec{W}_\mu\vec{W}^\mu + \frac{1}{8}g'^2B_\mu B^\mu - \frac{1}{4}gg'B_\mu W_3^\mu\right) \\ &+ \frac{1}{2}\lambda v^2(v+h) - \frac{1}{4}\lambda(v+h)^2\end{aligned} \tag{2.15}$$

Each term of equation (2.15) has mass dimension four, otherwise the theoretical model would be non renormalisable [43, 44]. The mass terms must appear in the form $m_W^2 WW$ and $m_Z^2 ZZ$ because the vector fields themselves have mass dimension one. Therefore the $\vec{W}_\mu\vec{W}^\mu$ and $B_\mu B^\mu$ terms in the second line can be interpreted as the mass terms of the gauge bosons. But due to the extra mixing term that is the last term in the second line, the W_μ^i and B_μ fields do not represent the physical mass eigenstates. To obtain three physical massive bosons and one massless boson and remove the mixing term, the W_μ^i and B_μ fields have to be replaced as shown in equation (2.8). The masses of the $W^\pm$ and Z bosons are

$$m_W = \frac{1}{2} v g \tag{2.16}$$

$$m_Z = \frac{1}{2} v \sqrt{g^2 + g'^2} \tag{2.17}$$

Second order mass terms for the Higgs field h are presented in the third line. These are the manifestation of the Higgs field as a massive physical Higgs boson. Other terms containing mixing of the Higgs field h and vector-boson fields, B and $\vec{W}$, can be interpreted as interactions between the Higgs boson and the vector bosons.

A fermion mass term of the form

$$\begin{aligned} m_f \bar{\psi}\psi &= m_f \left(\bar{\psi}_R + \bar{\psi}_L\right)\left(\psi_L + \psi_R\right) \\ &= m_f \left(\bar{\psi}_R \psi_L + \bar{\psi}_L \psi_R\right) \end{aligned} \tag{2.18}$$

is not gauge invariant and therefore not allowed in the Lagrangian shown in equation (2.3), because the left-handed doublets $\bar{\psi}_L$ and right-handed singlets $\bar{\psi}_R$ transform differently under SU(2) ⊗ U(1):

$$\begin{aligned} \psi_L &\to e^{i\frac{1}{2}\vec{\alpha}(x)\vec{\sigma} + i\beta(x)Y} \psi_L \\ \psi_R &\to e^{i\beta(x)Y} \psi_R \,. \end{aligned}$$

$\vec{\alpha}(x)$ and $\beta(x)$ are continuous parameters.

In the SM, the mass of fermions are given through interactions between the fermion fields and the Higgs doublet: $\lambda_f \bar{\psi}_L \phi \psi_R$. These so-called Yukawa interactions are invariant under SU(2) ⊗ U(1). λ_f is the Yukawa coupling. The Lagrangian of the electron for example becomes

$$\begin{aligned} \mathcal{L}_e &= -\lambda_e \left(\bar{\psi}_L \phi \psi_R + \bar{\psi}_R \phi \psi_L\right) \\ &= -\frac{\lambda_e}{\sqrt{2}} v \bar{e} e - \frac{\lambda_e}{\sqrt{2}} h \bar{e} e \,. \end{aligned} \tag{2.19}$$

The mass term of the electron is represented by the first term in the last line with a mass of $m_e = \lambda_e v / \sqrt{2}$. The interaction between the electron and the

Higgs boson is represented by the last term with a coupling proportional to the electron mass. The masses of the up-type quarks are given in a similar way due to terms of the form

$$\mathcal{L}_{up} = \bar{\psi}_L \tilde{\phi}^c \psi_R \,, \tag{2.20}$$

where $\tilde{\phi}^c = -i\sigma_2\phi^*$.

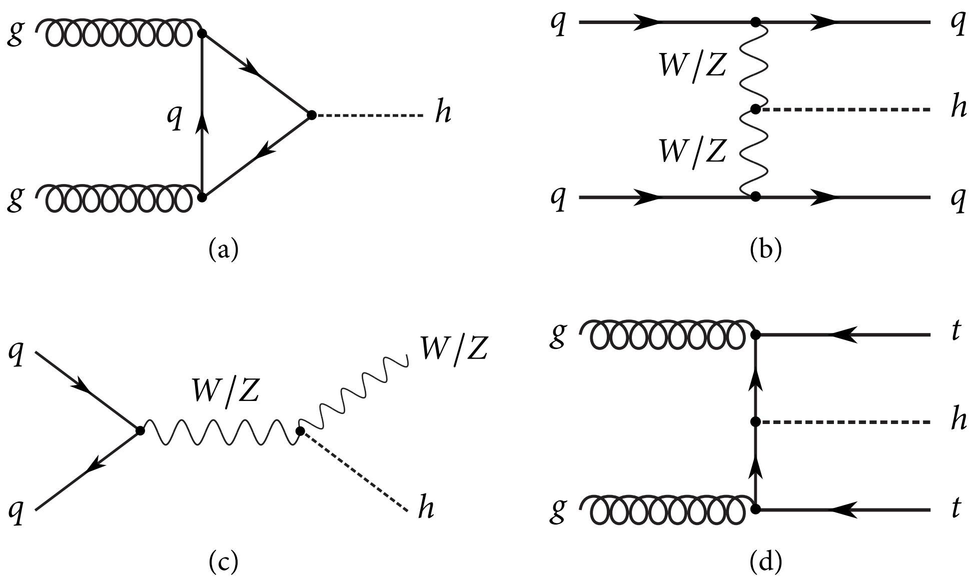
The spontaneous symmetry breaking via the Higgs fields is an elegant way to explain the masses of elementary particles and seems to be realised in nature [4]. The strong force, however, is not affected by the Brout-Englert-Higgs Mechanism and still confines the quarks into hadrons. As a result of the large binding energy, the masses of these bound states is almost unaffected of their constituents' masses and remain almost unchanged.

More information about the mechanism of electroweak symmetry breaking and in particular the Higgs particle of the SM can be found in the comprehensive review in reference [45].

2.1.2 Production and Decay of the Higgs Boson in the SM

At the LHC, the most important processes of the production of the Higgs boson in the SM are gluon fusion, VBF, Higgsstrahlung and associated production with a top-quark-antiquark ($t\bar{t}$) pair. Tree-level Feynman diagrams are shown in figure 2.3. The gluon fusion through a heavy-quark loop is the main production channel of the SM Higgs boson and is at least one order of magnitude larger than the cross section in the other channels as shown in figure 2.4. The production of the SM Higgs boson in association with two forward jets, denoted as VBF process, is the second main production channel, see figure 2.4. These two production modes as well as the production via Higgsstrahlung are considered in this analysis.

The cross sections of gluon-fusion production $\sigma(gg \to h_{\mathrm{SM}})$, VBF production $\sigma(qq \to qqh_{\mathrm{SM}})$, and Higgsstrahlung production $\sigma(pp \to Vh_{\mathrm{SM}})$, as well as the branching ratios $\mathcal{B}$ of the SM Higgs boson in various decay channels used in this analysis are listed in table 2.1 and table 2.2, respectively. Next-to-next-to-leading order (NNLO) QCD corrections and next-to-leading order (NLO) electroweak contributions to the SM cross sections and branching ratios are taken into account as described in [47] and the most recent values are listed in [48].



(a) (b) (c) (d)

Figure 2.3: Examples of tree-level Feynman diagrams for Higgs-boson production via gluon fusion (a), vector-boson fusion (b), Higgsstrahlung (c) and associated production with a top-quark-antiquark pair (d).

Up to 300 GeV the SM Higgs boson as well as the neutral Higgs bosons of the 2HDMs described in the next section can be assumed to be on-shell and therefore the narrow width approximation can be imposed, decoupling their production and decay. In addition, no colour and no spin information can be transmitted to the final outgoing particles at tree level and in all corrections that are calculated to date. As a result, there are no correlations between the production and decay process, which simplifies the calculation and simulation of Higgs signatures.

The coupling of the Higgs field to all particles is proportional to their masses and the decay rates are fixed by unitarity. The dominant decay mode for $m_{h_{\mathrm{SM}}} < 135$ GeV is $h_{\mathrm{SM}} \to b\bar{b}$ whereas for $m_{h_{\mathrm{SM}}} > 135$ GeV the off-shell decay into a W-boson pair, $h_{\mathrm{SM}} \to WW^*$, dominates over the other modes. In the mass region of the Higgs boson where the W bosons as well as the Z bosons become on-shell, the ratio of $h_{\mathrm{SM}} \to WW$ and $h_{\mathrm{SM}} \to ZZ$ decays is fixed by a relative factor of two because of the additional degree of freedom of the W-boson decay: W^+W^- and W^-W^+. The mass range of the Higgs boson

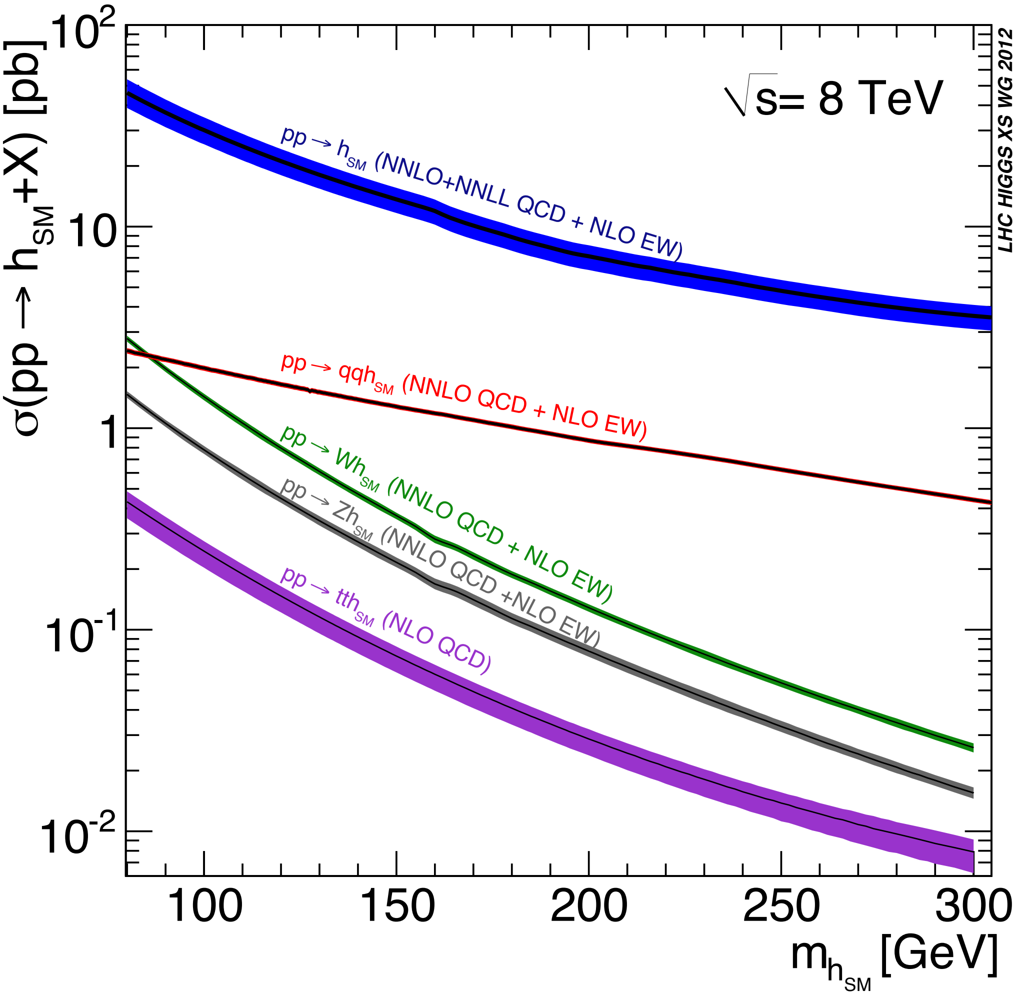


Figure 2.4: Standard-Model Higgs-boson production cross sections at $\sqrt{s} = 8$ TeV [46]. The line width indicates the uncertainty on the various cross sections due to theory and parametric uncertainties.

in table 2.1 and table 2.2 was chosen with regard to the sensitivity of the decay channel considered in this analysis.

Table 2.1: Cross sections of gluon-fusion, VBF and Higgsstrahlung processes for various mass points of the SM Higgs boson. V stands for the W boson and the Z boson, so that $\sigma(pp \rightarrow Vh_{\mathrm{SM}}) = \sigma(pp \rightarrow Wh_{\mathrm{SM}}) + \sigma(pp \rightarrow Zh_{\mathrm{SM}})$.

$m_{h_{\mathrm{SM}}}$[GeV]	$\sigma(pp \rightarrow h_{\mathrm{SM}})$[pb]	$\sigma(pp \rightarrow qqh_{\mathrm{SM}})$[pb]	$\sigma(pp \rightarrow Vh_{\mathrm{SM}})$[pb]
125	19.52	1.58	1.09
130	18.07	1.51	0.96
135	16.79	1.45	0.84
140	15.63	1.39	0.74
145	14.59	1.33	0.66
150	13.65	1.28	0.58
155	12.79	1.23	0.52
160	11.95	1.19	0.45
165	10.89	1.14	0.42
170	10.12	1.10	0.37
175	9.48	1.06	0.34
180	8.87	1.02	0.30
185	8.33	0.98	0.28
190	7.87	0.94	0.25
195	7.43	0.90	0.23
200	7.13	0.87	0.21
220	6.04	0.77	0.14
240	5.18	0.67	0.10
260	4.48	0.58	0.07
280	3.96	0.50	0.06
300	3.61	0.44	0.04

Table 2.2: Branching ratios for various mass points of the SM Higgs boson.

$m_{h_{\mathrm{SM}}}$[GeV]	$\mathcal{B}(h_{\mathrm{SM}} \to WW^{(*)})$	$\mathcal{B}(h_{\mathrm{SM}} \to ZZ^{(*)})$	$\mathcal{B}(h_{\mathrm{SM}} \to b\bar{b})$
125	22%	3%	58%
130	30%	4%	49%
135	40%	4%	40%
140	50%	7%	32%
145	60%	8%	23%
150	70%	8%	16%
155	79%	7%	9%
160	91%	4%	3%
165	96%	2%	1%
170	96%	2%	1%
175	96%	3%	1%
180	93%	6%	1%
185	84%	15%	$< 1\%$
190	79%	21%	$< 1\%$
195	76%	24%	$< 1\%$
200	74%	26%	$< 1\%$
220	71%	28%	$< 1\%$
240	70%	29%	$< 1\%$
260	70%	30%	$< 1\%$
280	70%	30%	$< 1\%$
300	69%	31%	$< 1\%$

2.2 Theory and Phenomenology of Two-Higgs-Doublet Models

In the SM, the Higgs sector has been chosen to be as simple as possible to create the masses of the weak bosons and the fermions. But what is seen to be the minimalistic description is not always supported by nature. As in the case of fermions, where instead of one generation three generations exist in reality, the Higgs sector could also have more fields and more particles than the minimum choice in the SM. The addition of a second complex Higgs doublet, leading to 2HDMs, is not only the most straightforward extension of the SM Higgs sector, but also capable of producing enough CP symmetry violation to explain the observed asymmetry between baryons and antibaryons in the universe. The CP symmetry is the product of two symmetries, the charge symmetry and the parity symmetry. Charge symmetry is the invariance of physical laws under the transformation of charge conjugation. Parity symmetry is the invariance of physical laws under the transformation of flipping the sign of all spatial coordinates of a particle simultaneously. The violation of the CP symmetry was discovered in 1964 [49] and is one of the Sakharov conditions [50], which must be satisfied to produce the huge imbalance between matter and antimatter in the universe [51]. The SM predicts CP symmetry breaking in weak interactions, which is experimentally verified, but is much too small to explain the observed matter-antimatter asymmetry of the universe.

2.2.1 The Models

An implementation of a non-minimal Higgs sector favoured by many theories are 2HDMs. These models introduce two complex scalar SU(2) doublets

$$\phi_1 = \begin{pmatrix} \phi_1^+ \\ \phi_1^0 \end{pmatrix} \text{ and } \phi_2 = \begin{pmatrix} \phi_2^+ \\ \phi_2^0 \end{pmatrix} \tag{2.21}$$

with Hypercharge +1.

One of the most serious problems of all 2HDMs is the possibility of tree level flavour-changing neutral currents (FCNC), which must be very small in order to satisfy the experimental bounds [52, 53]. To avoid tree level FCNC it is necessary and sufficient, that all fermions with the same quantum num-

Table 2.3: Type-I and type-II 2HDMs which lead to natural flavour conservation. In the type-I 2HDM, ϕ_u is the Yukawa coupling that determines the couplings to u_R^i, d_R^i and l_R^i. In the type-II 2HDM, ϕ_d is the Yukawa coupling that determines the coupling to d_R^i and l_R^i, while ϕ_d determines the coupling to u_R^i. The superscript i is the index of the various generations: $u_R^i = (u_R, c_R, t_R)$, $d_R^i = (d_R, s_R, b_R)$, $l_R^i = (e_R, \mu_R, \tau_R)$.

	Type I	Type II
u_R^i	ϕ_u	ϕ_u
d_R^i	ϕ_u	ϕ_d
l_R^i	ϕ_u	ϕ_d

bers couple to the same Higgs doublet [54, 55]. Following this condition, within the quark sector of the 2HDM only two possibilities exist. In the type-I 2HDM, all quarks couple to just one of the Higgs doublets with the coupling constant ϕ_u. In the type-II 2HDM, the up-type quarks couple to one Higgs doublet with the coupling constant ϕ_u and the down-type quarks to the other doublet with the coupling constant ϕ_d. The two models are presented in table 2.3. The type-I 2HDM can be enforced by imposing a discrete $\mathbb{Z}_2$ symmetry, where

$$\phi_1 \to -\phi_1 \text{ and } \phi_2 \to \phi_2 \,.$$

The type-II 2HDM can be enforced by a discrete $\mathbb{Z}_2$ symmetry, where

$$\phi_1 \to -\phi_1 \,, \phi_2 \to \phi_2 \text{ and } d_R^i \to -d_R^i \,.$$

These requirements lead to the following scalar Higgs potential

$$\begin{aligned} V =& m_{11}^2 \phi_1^\dagger \phi_1 + m_{22}^2 \phi_2^\dagger \phi_2 - \left(m_{12}^2 \phi_1^\dagger \phi_2 + \text{h.c.} \right) + \frac{1}{2}\lambda_1 \left(\phi_1^\dagger \phi_1 \right)^2 \\ &+ \frac{1}{2}\lambda_2 \left(\phi_2^\dagger \phi_2 \right)^2 + \lambda_3 \left(\phi_1^\dagger \phi_1 \right) \left(\phi_2^\dagger \phi_2 \right) + \lambda_4 \left(\phi_1^\dagger \phi_2 \right) \left(\phi_2^\dagger \phi_1 \right) \\ &+ \left\{ \frac{1}{2}\lambda_5 \left(\phi_1^\dagger \phi_2 \right)^2 + \text{h.c.} \right\} \end{aligned} \tag{2.22}$$

with h.c. as the abbreviation for Hermitian conjugate. All parameters in equation (2.22) are real [56].

The neutral components ϕ_1^0 and ϕ_2^0 of the two doublets ϕ_1 and ϕ_2 acquire two different vacuum expectation values

$$\langle\phi_1\rangle_{\min} = \begin{pmatrix} 0 \\ v_1/\sqrt{2} \end{pmatrix} \text{ and } \langle\phi_2\rangle_{\min} = \begin{pmatrix} 0 \\ v_2/\sqrt{2} \end{pmatrix} \tag{2.23}$$

with $v^2 = v_1^2 + v_2^2$, where v is the SM vacuum expectation value defined in equation (2.13). After expanding the potential around the vacuum expectation values, the Higgs fields become

$$\phi_a = \begin{pmatrix} \phi_a^+ \\ (v_a + \rho_a + i\eta_a)/\sqrt{2} \end{pmatrix}, \quad a = 1, 2, \tag{2.24}$$

with $\rho_a = \Re(\phi_a^0) - v_a$ and $\eta_a = \Im(\phi_a^0)$.

The mass terms in the Lagrangian density for the neutral scalars, for the pseudoscalars and for the charged scalars are

$$\mathcal{L}_{\text{mass}}^{\phi^\pm} = \{m_{12}^2 - (\lambda_4 + \lambda_5) v_1 v_2\} \begin{pmatrix} \phi_1^- & \phi_2^- \end{pmatrix} \begin{pmatrix} v_2/v_1 & -1 \\ -1 & v_1/v_2 \end{pmatrix} \begin{pmatrix} \phi_1^+ \\ \phi_2^+ \end{pmatrix} \tag{2.25}$$

$$\mathcal{L}_{\text{mass}}^{\eta} = \frac{m_A^2}{v_1^2 + v_2^2} \begin{pmatrix} \eta_1 & \eta_2 \end{pmatrix} \begin{pmatrix} v_2^2 & -v_1 v_2 \\ -v_1 v_2 & v_1^2 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} \tag{2.26}$$

$$\mathcal{L}_{\text{mass}}^{\rho} = -\begin{pmatrix} \rho_1 & \rho_2 \end{pmatrix} \begin{pmatrix} m_{12}^2 v_2/v_1 + \lambda_1 v_1^2 & -m_{12}^2 + \lambda_{345} v_1 v_2 \\ -m_{12}^2 + \lambda_{345} v_1 v_2 & m_{12}^2 v_1/v_2 + \lambda_2 v_2^2 \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix}, \tag{2.27}$$

with $\lambda_{345} = \lambda_3 + \lambda_4 + \lambda_5$ [56].

The excitations of the different Higgs fields around their vacuum expectation values give rise to eight degrees of freedom, from which three can be absorbed to generate the longitudinal degrees of freedom of the SM weak bosons. This leaves a quintet of scalar particles: two CP-even neutral Higgs bosons h and H, a CP-odd neutral Higgs boson A, and two charged Higgs particles $H^\pm$.

The mixing of the different fields can be written as:

$$\begin{pmatrix} h \\ H \end{pmatrix} = \begin{pmatrix} \sin\alpha & -\cos\alpha \\ -\cos\alpha & -\sin\alpha \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix} \tag{2.28}$$

$$\begin{pmatrix} G \\ A \end{pmatrix} = \begin{pmatrix} \cos\beta & \sin\beta \\ \sin\beta & -\cos\beta \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} \tag{2.29}$$

$$\begin{pmatrix} G^\pm \\ H^\pm \end{pmatrix} = \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix} \begin{pmatrix} \phi_1^\pm \\ \phi_2^\pm \end{pmatrix}, \tag{2.30}$$

where G and $G^\pm$ are the Goldstone bosons, which can be transformed into the physical weak vector bosons Z and $W^\pm$ [56]. The rotation angle β diagonalises the mass-squared matrices of the pseudoscalars and of the charged scalars. The rotation angle α diagonalises the mass-squared matrix of the neutral scalars. An important parameter of the 2HDMs is the ratio of the vacuum expectation values of the two Higgs doublets: $\tan\beta = v_2/v_1$.

The main reason for extending the SM to 2HDMs is the possibility of generating new sources of CP violation. But with the aforementioned choice of the discrete $\mathbb{Z}_2$ symmetry CP violation arises only from the complex Yukawa couplings as in the case of the SM. Some ways to get new sources of CP symmetry violation, which are proposed in [57] are:

- The term proportional to m_{12}^2 breaks the discrete $\mathbb{Z}_2$ symmetry softly. This type of violation respects the discrete symmetry at small orders of perturbation theory and suppresses FCNC.

- The hard symmetry breaking violates the discrete $\mathbb{Z}_2$ symmetry through the complex parameter λ_6 and λ_7.

- The discrete symmetry can be dropped entirely and the FCNC is assumed to be suppressed by other mechanisms.

- The Yukawa couplings as well as the potential break the discrete symmetry, but these violations are small.

2.2.2 Production Cross Sections and Decay Rates

The two parameters α and β introduced in section 2.2.1 determine the interactions of the various Higgs fields with the vector bosons and with the fermions. They are thus crucial in discussing 2HDM phenomenology. The couplings in terms of α and β relative to the SM are given in table 2.4 [56]. In both models, the couplings of the neutral Higgs bosons to the W and Z bosons are the same: the coupling ξ_h^V of the light Higgs boson to either W or Z bosons is the same as the SM coupling times $\sin(\beta - \alpha)$ and the coupling ξ_H^V of the heavy Higgs boson to either W or Z bosons is the same as the SM coupling times $\cos(\beta-\alpha)$. The quadratic sum of the couplings of the Higgs bosons to one of the vector bosons is identical to one: $\left(\xi_h^V\right)^2+\left(\xi_H^V\right)^2 = \sin^2(\beta-\alpha)+\cos^2(\beta-\alpha) = 1$. While the couplings of the neutral Higgs bosons to up-type quarks, ξ_h^u and ξ_H^u, is the same for either the type-I or the type-II 2HDM, the couplings to down-type quarks, ξ_h^d and ξ_H^d, differ according to the particular model.

If $\alpha = 0$, the type-I 2HDM exhibits an interesting limit, the so-called fermiophobic limit. In this scenario, the heavy Higgs boson H has a zero coupling to fermions at tree level and becomes a fermiophobic Higgs boson.

In case $\alpha = \beta - \pi/2$, another relevant scenario appears in which the light Higgs boson becomes SM like and the heavy Higgs boson becomes a bosophobic Higgs boson. This scenario constrains the coupling of the heavy Higgs boson to the fermions too: $\xi_H^u = \xi_H^d = -\tan^{-1}\beta$ in the type-I 2HDM and $\xi_H^u = -\tan^{-1}\beta$, $\xi_H^d = \tan\beta$ in the type-II 2HDM.

The production of the light Higgs boson as well as the heavy Higgs boson can proceed through gluon fusion and VBF, with decays to SM channels. The implementation of the gluon-fusion coupling factors to the analysis is performed with `SusHi` [59], a program which calculates Higgs production cross sections in various models. As of version 1.0.3, it is possible to calculate the cross sections within several types of 2HDMs. The accuracy of the predictions calculated by `SusHi` is "almost" NNLO, which means that the NLO QCD contributions from the third family of quarks are fully taken into account. The NNLO QCD corrections are implemented by `ggh@nnlo` [60] which takes into account gluon fusion via a top-quark loop. The electroweak corrections up to NNLO are implemented as tabulated correction factors. The internal workflow of `SusHi` is given in figure 2.5, taken from [59].

Table 2.4: The couplings of the light (upper part) and the heavy (lower part) Higgs bosons, h and H, in the type-I and type-II 2HDM in terms of α and β relative to the couplings of the Standard Model [58].

Coupling	Type I	Type II
ξ_h^V	$\sin(\beta-\alpha)$	$\sin(\beta-\alpha)$
ξ_h^u	$\cos\alpha/\sin\beta$	$\cos\alpha/\sin\beta$
ξ_h^d	$\cos\alpha/\sin\beta$	$-\sin\alpha/\cos\beta$
ξ_h^l	$\cos\alpha/\sin\beta$	$-\sin\alpha/\cos\beta$
ξ_H^V	$\cos(\beta-\alpha)$	$\cos(\beta-\alpha)$
ξ_H^u	$\sin\alpha/\sin\beta$	$\sin\alpha/\sin\beta$
ξ_H^d	$\sin\alpha/\sin\beta$	$\cos\alpha/\cos\beta$
ξ_H^l	$\sin\alpha/\sin\beta$	$\cos\alpha/\cos\beta$

The VBF cross sections of the 2HDMs are calculated by multiplying the SM cross sections of table 2.1 by the coupling constants of table 2.4:

$$\sigma\left(pp \to qq\, h/H\right) = \sigma\left(pp \to qq\, h_{\mathrm{SM}}\right) \cdot \left(\xi_{h/H}^V\right)^2 . \tag{2.31}$$

The branching ratios of the neutral Higgs bosons of the 2HDMs into two W bosons, $\mathcal{B}\left(h/H \to W^+W^-\right)$, are calculated through

$$\begin{aligned}\mathcal{B}\left(h/H \to WW\right) &= \frac{\Gamma\left(h/H \to WW\right)}{\Gamma_{\mathrm{total}}} \\ &\simeq \frac{\Gamma\left(h_{\mathrm{SM}} \to WW\right)\left(\xi_{h/H}^V\right)^2}{\Gamma\left(h_{\mathrm{SM}} \to VV\right)\cdot\left(\xi_{h/H}^V\right)^2 + \Gamma\left(h_{\mathrm{SM}} \to bb\right)\cdot\left(\xi_{h/H}^d\right)^2} \\ &= \frac{\mathcal{B}\left(h_{\mathrm{SM}} \to WW\right)\left(\xi_{h/H}^V\right)^2}{\mathcal{B}\left(h_{\mathrm{SM}} \to VV\right)\cdot\left(\xi_{h/H}^V\right)^2 + \mathcal{B}\left(h_{\mathrm{SM}} \to bb\right)\cdot\left(\xi_{h/H}^d\right)^2} .\end{aligned} \tag{2.32}$$

The calculated values of cross section times branching ratio in terms of $\cos\alpha$ and m_H for the heavy Higgs boson are given in figure 2.6.

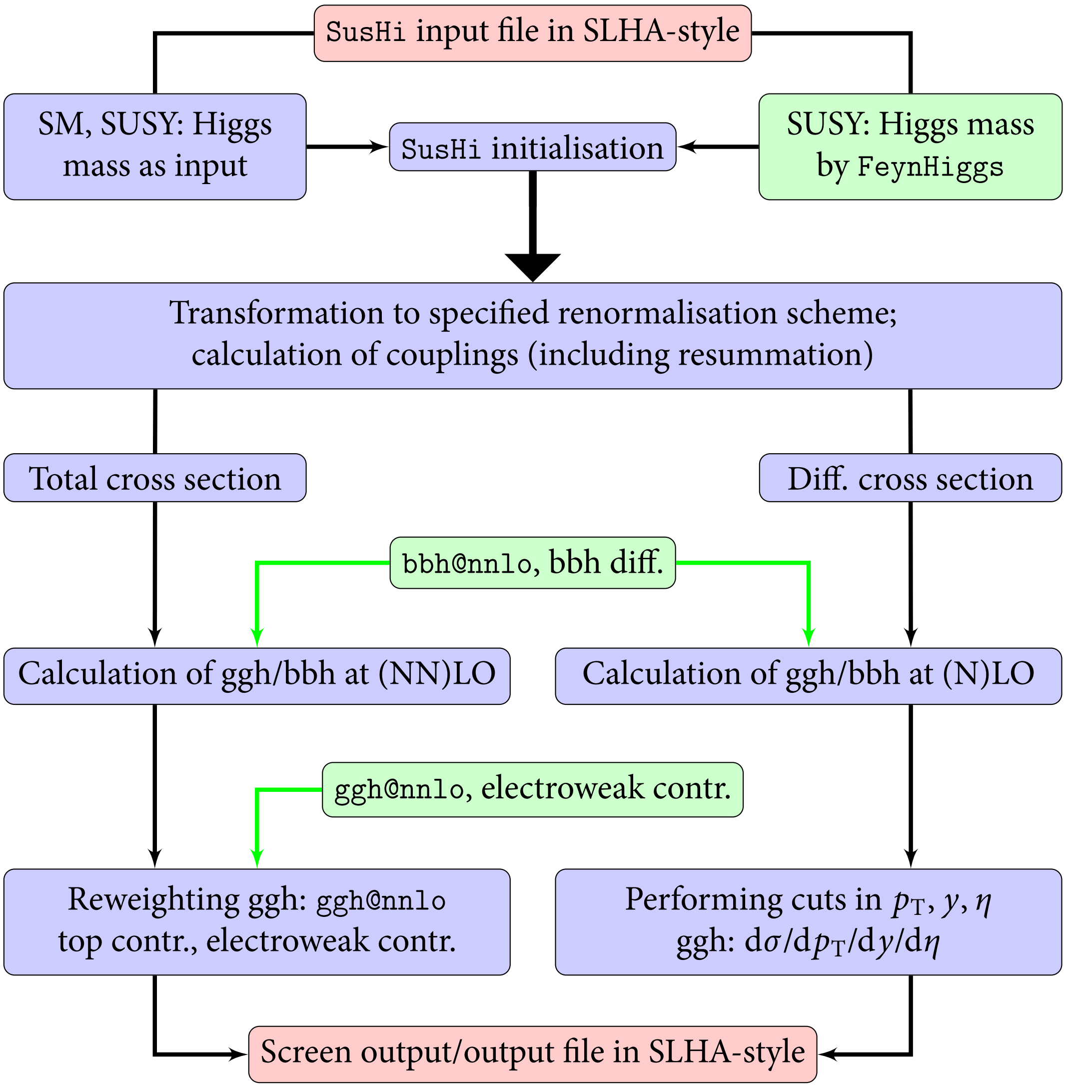
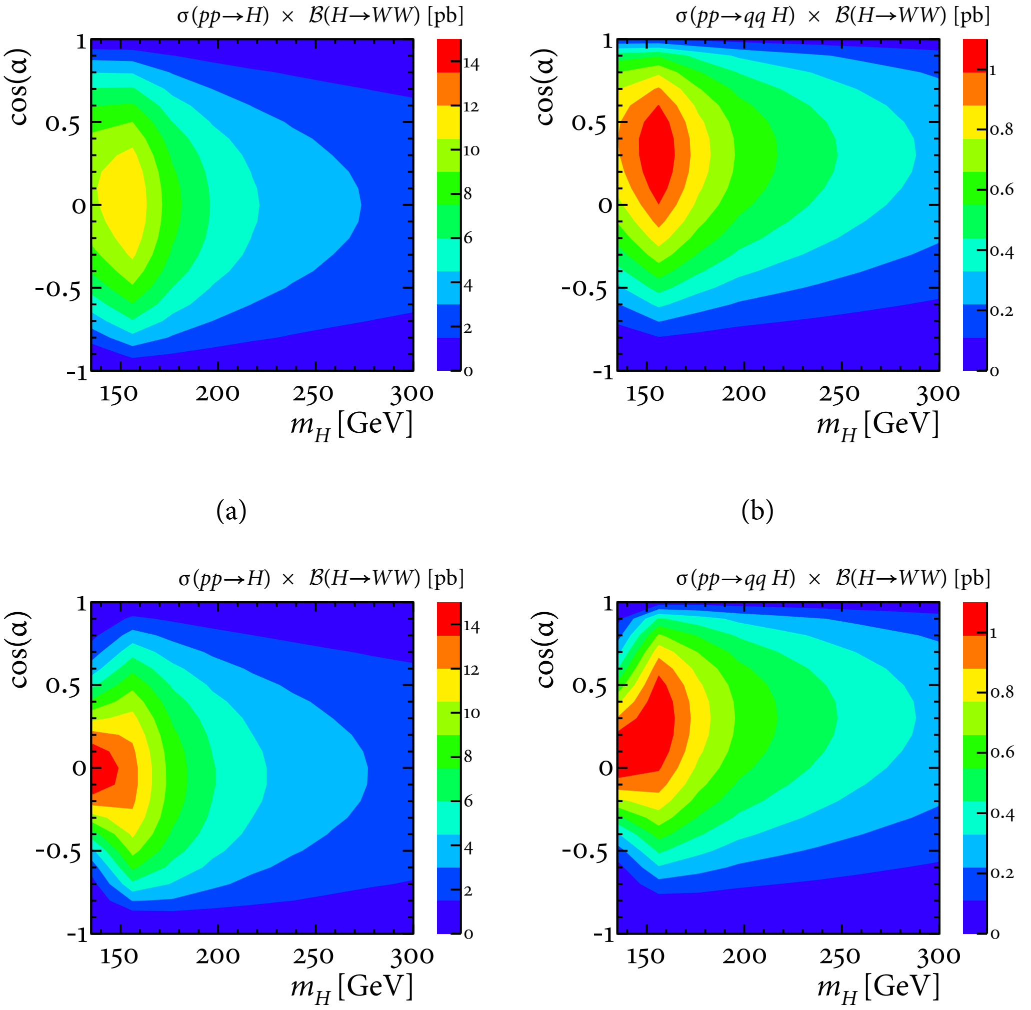


Figure 2.5: Internal workflow of `SusHi` [59]. Red boxes describe user interaction, green boxes refer to external code.



(a)

(b)

(c)

(d)

Figure 2.6: Cross section times branching ratio for the 2HDMs type-I and type-II in the $\cos\alpha$-m_{H} plane for $\tan\beta = 3$. Figure (a) and (b) show the gluon-fusion cross sections and VBF cross sections in the type-I 2HDM, respectively. Figure (c) and (d) show the gluon-fusion cross sections and VBF cross sections in the type-II 2HDM, respectively. The gluon fusion cross sections are calculated with `SusHi` [59].

Chapter 3

LHC and the ATLAS Detector

To test the predictions of the SM and search for physics phenomena beyond the SM, very high energies are required. The currently highest reachable energies are achieved by the accelerator complex at CERN, depicted in figure 3.1. The accelerator complex consists of a succession of machines, which accelerates, protons or heavy ions. For protons the energy of 4 TeV per beam is obtained in the LHC, which is the last element in this sequence.

There are seven experiments at the LHC recording events from collisions at the four interaction points where the particle beams collide: ALICE [62], ATLAS [2], CMS [3], LHCb [63], LHCf [64], MoEDAL [65] and TOTEM [66]. A Toroidal LHC ApparatuS (ATLAS) and the Compact Muon Solenoid (CMS) are multipurpose detectors exploring a wide range of physics processes, ranging from confirmation and improved measurements of the SM, including the Higgs boson, to investigations of new physics phenomena, which are predicted by distinct theories. A Large Ion Collider Experiment (ALICE) is dedicated to heavy ion collisions to explore the generated quark-gluon plasma in which the temperature is large enough to abrogate the confinement of quarks and gluons. Such a state of matter is supposed to have existed at the beginning of the universe. The Large Hadron Collider beauty (LHCb) experiment is specialised to bottom-quark physics to find and measure CP symmetry violating processes. The Large Hadron Collider forward (LHCf) experiment is designed to study π^0-mesons produced in the forward region of collisions close to the beam. The Monopole and Exotics Detector At the LHC (MoEDAL) is dedicated to direct searches for magnetic monopoles or dyons and other highly ionising stable (or pseudo-stable) massive particles. The LHC's seventh detector, TOTal Elastic and diffractive cross section Measurement (TOTEM), measures the

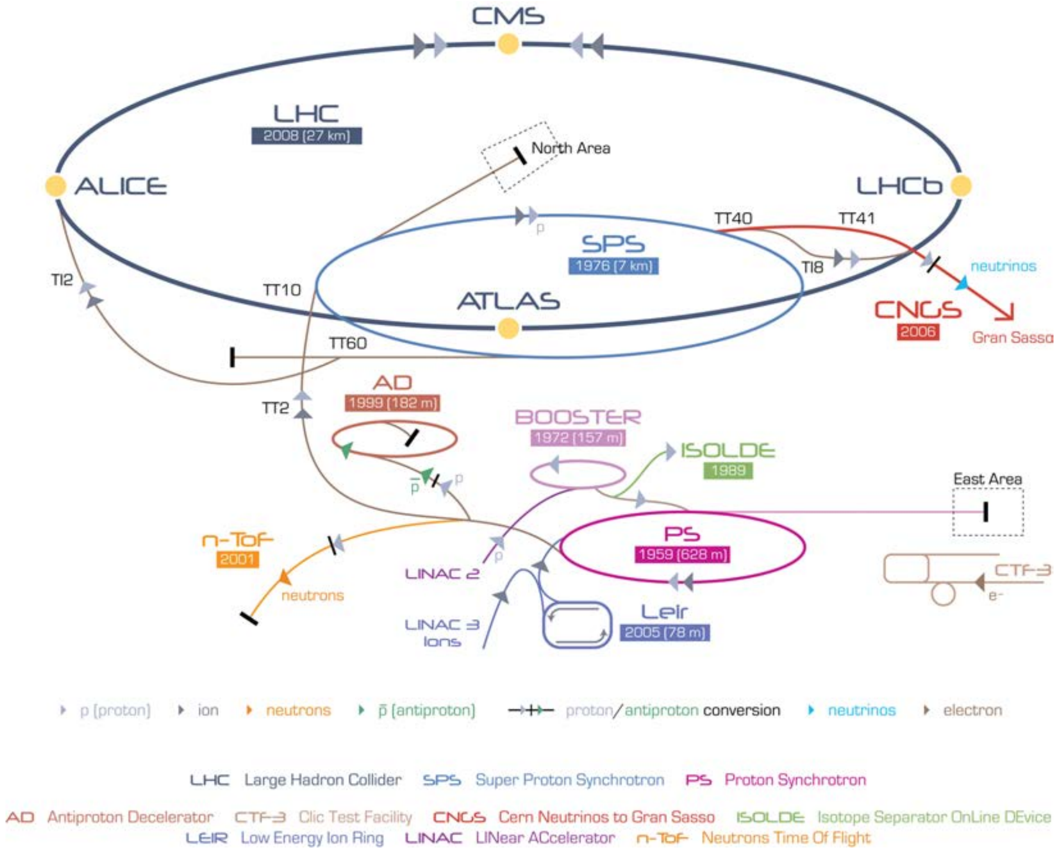


Figure 3.1: The accelerator complex at CERN showing the LHC and the smaller particle accelerators, which are used to boost the particles to their final energies [61].

total proton-proton cross section, studies elastic and diffractive scattering and accurately monitors the luminosity of the LHC.

This chapter contains two sections, section 3.1 giving a short overview of the accelerator complex at CERN and section 3.2 explaining in more detail the ATLAS detector, whose recorded data are used in this thesis.

3.1 The Large Hadron Collider

The protons used by the LHC originate from a bottle of hydrogen gas. The hydrogen atoms are passed through an electric field to strip off the electrons and yield protons [67]. The starting point for these protons, see figure 3.1, is the Linear Accelerator 2 (Linac 2) [68] in which they are accelerated to 50 MeV for injection into the second accelerator in the chain, the Proton Synchrotron Booster (PSB) [69]. The PSB accelerates the protons to 1.4 GeV followed by the Proton Synchrotron (PS) [70]. The PS is responsible for providing 81 proton bunches, which are discrete packets of protons, and pushes these to an energy of 25 GeV. The protons are then injected into the second-largest machine in the accelerator complex, the Super Proton Synchrotron (SPS) [71], where the beam energy is increased to 450 GeV. Finally, the protons are transferred to two separate beam pipes of the LHC in such a way that one beam of protons runs clockwise and the other one runs counterclockwise. In 2012, the LHC operated with an energy of 4 TeV per beam, but it is designed to reach eventually an energy of 7 TeV per beam.

The beams are kept on their circular path around the accelerator ring of the LHC by a strong magnetic field maintained by 1232 superconducting electromagnets. These electromagnets are made of copper-clad niobium-titanium and their operating temperature of 1.9 K is provided by super-fluid helium. To accelerate the bunches the LHC uses for each beam eight radio-frequency cavities, which are metallic chambers that contain an oscillating electromagnetic field [72]. Each cavity delivers 2 MV at an oscillation frequency of 400 MHz.

One of the most important parameters of particle colliders is the instantaneous luminosity L, which is a measure of how efficiently an accelerator produces collision events. The instantaneous luminosity is the number of collisions that can be produced in a detector per square centimetre per second at the interaction point:

$$L = \frac{f \cdot N^2 \cdot k_b \cdot F}{4 \cdot \pi \cdot \varepsilon^* \cdot \beta^*} \,. \tag{3.1}$$

f is the crossing frequency of the bunches, N is the bunch intensity (particles per bunch), k_b is the number of bunches and F is the geometric luminosity reduction factor due to the crossing angle at the interaction point. ε^* and β^*

Table 3.1: Peak performance for the year 2010 to 2012 in comparison to the nominal values [73, 74].

	2010	2011	2012	Nominal
Bunch Spacing [ns]	150	50	50	25
k_b	368	1380	1380	2808
β^* [m] (ATLAS and CMS)	3.5	1.0	0.6	0.55
N_{max}	1.2×10^{11}	1.5×10^{11}	1.7×10^{11}	1.2×10^{11}
ε^* [μm]	~ 2.0	~ 2.4	~ 2.5	3.75
L_{peak} [$\mathrm{cm}^{-2}\mathrm{s}^{-1}$]	2.1×10^{32}	3.7×10^{33}	7.7×10^{33}	1.0×10^{34}

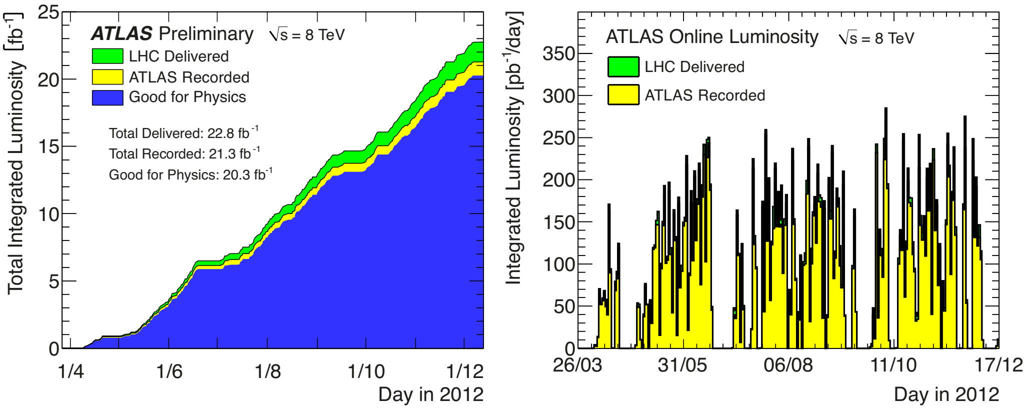
are the normalised transverse beam emittance and the amplitude function at the interaction point, respectively. Different values of peak performance through the years are shown in table 3.1.

The integral of the delivered luminosity over time is called integrated or cumulative luminosity:

$$L_{\mathrm{int}} = \int L \,\mathrm{d}t \,. \tag{3.2}$$

It is a measure of the amount of collected data in a given time period. The delivered and recorded total integrated luminosity and integrated luminosity per day for proton-proton collisions in 2012 are shown in figure 3.2. The first 13 fb^{-1} of the 2012 data set are analysed in this thesis. The total number of events N is proportional to the integrated Luminosity as well as to the cross section σ of a specific process:

$$N = \sigma \cdot L_{\mathrm{int}} \,. \tag{3.3}$$



(a)

(b)

Figure 3.2: Integrated luminosity for proton-proton collisions at a centre-of-mass energy of 8 TeV. Figure (a) shows the total integrated luminosity versus the day in 2012 and figure (b) shows the integrated luminosity per day versus the day in 2012. The graphs show the luminosity delivered (green) to ATLAS as well as the luminosity recorded (yellow) by ATLAS. | ATLAS Experiment © 2013 CERN

3.2 The ATLAS Experiment

The ATLAS detector, which is shown as an overview in figure 3.3, consists of four major components around the interaction point: the inner detector (ID) [75, 76] to track charged particles, the calorimeters to measure the energy carried by the particles, the muon spectrometer (MS) to identify and measure the momenta of muons and the magnet system to bend charged particles for momentum measurement. The expected energy and momentum resolutions and geometric acceptances of the various subdetectors are summarised in table 3.2.

ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point in the centre of the detector and the z-axis along the beam direction. The z-axis is parallel to the anti-clockwise beam viewed from above. Using a Cartesian coordinate system, the x-axis points towards the centre of the LHC tunnel and the y-axis points vertical upwards [1]. Cylindrical

---

[1] Actually the y-axis is slightly different from vertical due to the general tilt of the LHC tunnel.

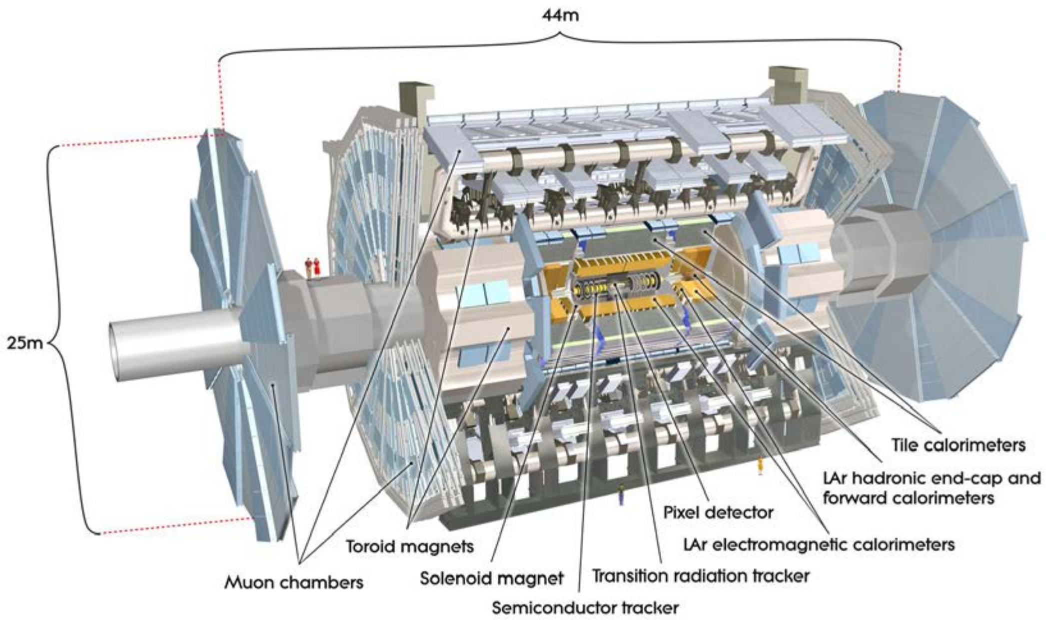


Figure 3.3: Schematic view of the ATLAS detector including all subdetectors.
| ATLAS Experiment © 2013 CERN

Table 3.2: Performance goals of the various subdetectors of the ATLAS detector. The units of E and p_T are GeV.

Detector Component	Required Resolution	η Coverage Measurement	Trigger
Tracking	$\sigma_{p_T}/p_T = 0.05\% \cdot p_T \oplus 1\%$	±2.5	-
EM Calorimetry	$\sigma_E/E = 10\% / \sqrt{E} \oplus 0.7\%$	±3.2	±2.5
Hadronic Calorimetry			
Barrel and End-Cap	$\sigma_E/E = 50\% / \sqrt{E} \oplus 3\%$	±3.2	±3.2
Forward	$\sigma_E/E = 100\% / \sqrt{E} \oplus 10\%$	$3.1 < \|\eta\| < 4.9$	$3.1 < \|\eta\| < 4.9$
Muon Spectrometer	$\sigma_{p_T}/p_T = 10\%$ at $p_T = 1$ TeV	±2.7	±2.4

coordinates are often used for convenience. In this case, the azimuthal angle ϕ is measured from the x-axis in the x-y-plane, so that $\phi = \pi/2$ is pointing up. The polar angle Θ is measured with respect to the positive z-axis. Instead of Θ usually the rapidity

$$y := \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \tag{3.4}$$

or the pseudorapidity

$$\eta := -\ln \left(\tan \frac{\Theta}{2} \right) \tag{3.5}$$

are used. The difference between the rapidities of two particles $\Delta y = y_1 - y_2$ is invariant under Lorentz transformations along the z-axis. η is a zero-mass approximation of y that can be applied for highly relativistic particles. While the region with high values of $|\eta|$ are regarded as "forward", the region with small values of $|\eta|$ are called "central". To measure distances between directions of two objects the angular distance ΔR defined as

$$\Delta R := \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \tag{3.6}$$

is used. The transverse momentum p_T, defined as

$$p_\mathrm{T} := \sqrt{p_x^2 + p_y^2}\,, \tag{3.7}$$

is the projection of the momentum p onto the x-y-plane and the energy E_T, defined as

$$E_\mathrm{T} := E \sin\Theta = \sqrt{m^2 + p_\mathrm{T}^2}\,, \tag{3.8}$$

which is used for calorimetric measurements. For massless particles E_T is identical to the magnitude of the transverse momentum.

3.2.1 The Inner Detector

The ATLAS inner detector (ID) is the innermost part of a series of ever-larger concentric detectors around the interaction point. It is designed to precisely measure trajectories of charged particles in a high occupancy environment of numerous tracks within the region $|\eta| < 2.5$. This includes momentum measurement, pattern recognition and the identification of primary vertices that are the space points where inelastic proton-proton collisions have occurred. For the momentum measurement a solenoid magnet [77] that surrounds the ID produces a magnetic field of 2 T to curve even very energetic particles. The ID consists of three subdetectors, from inside out:

- The pixel detector [78], see figure 3.4, allows an accurate reconstruction of vertices. It consists of three barrel layers and three end-cap discs per side with a spatial resolution of 10 μm in r-ϕ-direction, i.e. the concentric plane around the beam axis, and of 115 μm along the beam axis. The pixel detector provides three measurement space points per track.

- The semiconductor tracker (SCT) [79] contributes to the measurements of momentum, impact parameter, and vertex position. The detecting medium is formed by silicon microstrip sensors. The SCT consists of four barrel layers and nine end-cap discs per side with a spatial resolution of 17 μm in r-ϕ-direction and of 580 μm along the beam axis. To measure three dimensional space points, single-sided sensors are glued back-to-back, such that a small stereo angle exists between them, to form double-sided modules. The SCT provides at least four measurement space points per track if all modules work with maximum efficiency.

- The transition radiation tracker (TRT) [80] is designed as a combined tracking and electron identification detector within the region $|\eta| < 2.0$. Its detecting elements are thin-walled straw drift tubes with a resolution of 130 μm. When a charged particle traverses the straw, it ionises the gas. The accrued free electrons drift towards the wire, cascade in the strong electric field close to the wire and produce a detectable signal. The space between the straws is filled with material of widely varying indices of refraction. At the boundary between two of these media relativistic charged particles emit photons in strongly forward direction. The emitted photons interact with the straw gas resulting in a higher signal strength. The probability of transition radiation increases with the relativistic Lorentz factor $\gamma = 1/\sqrt{1 - v^2/c^2}$, where v and c are the speed of the charged particle and the speed of light, respectively. Electrons produce significantly more transition radiation than pions because of their higher relativistic Lorentz factor. The higher amplitude of the signal

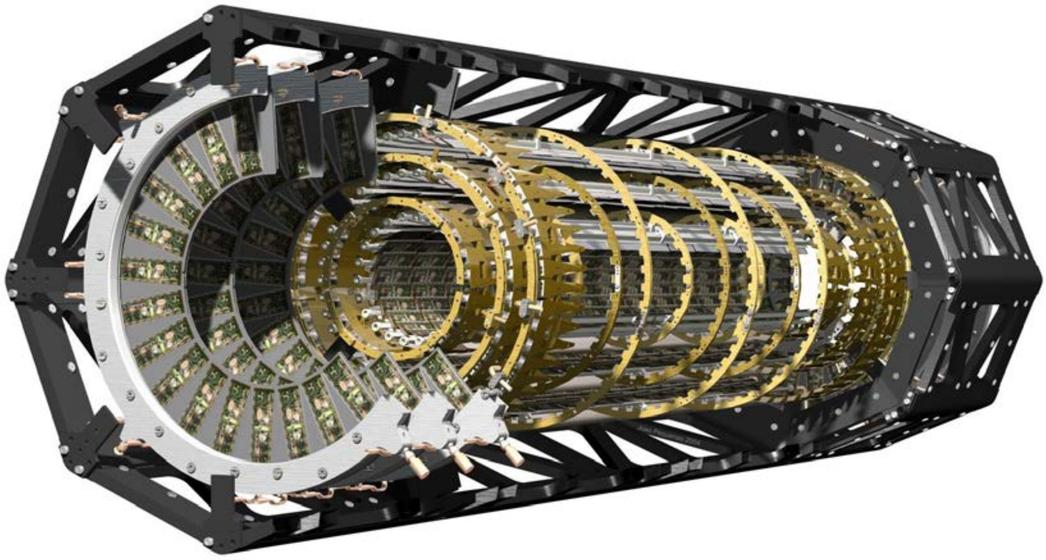


Figure 3.4: The ATLAS pixel detector. | ATLAS Experiment © 2013 CERN

is used to distinguish electrons from pions. The TRT provides an average of 36 measurement space points per track for charged particles with transverse momenta of $p_\mathrm{T} > 0.5$ GeV.

3.2.2 The Calorimeter System

The ATLAS calorimeter system [81–83], which is situated outside the solenoid magnet, measures the energies of charged and neutral particles. It is built up of two different calorimeters: the inner electromagnetic calorimeter [84] to measure the energy from electromagnetically interacting particles, which produce electromagnetic showers of electrons and photons, and the outer hadronic calorimeter [85] to measure the energy from all particles that pass through the electromagnetic calorimeter.

The electromagnetic calorimeter is a sampling detector, which covers a pseudorapidity range of $|\eta| < 3.2$ and the whole range in the ϕ-direction. It consists of a barrel part within the range $|\eta| < 1.5$, the endcap within the range $1.5 < |\eta| < 2.5$, and the forward calorimeter within the range $2.5 < |\eta| < 3.2$. Lead and copper are used as the energy absorbing passive materials and liquid argon (LAr) is used as the active sampling material.

The hadronic calorimeter is mainly used to measure the energy of jets by stopping them. It is divided into three components: the tile calorimeter (TileCal), which is subdivided into a barrel and one extended barrel per side, operating in the range $|\eta| < 1.7$, the LAr hadronic endcap calorimeter (HEC) within the range $1.5 < |\eta| < 3.2$, and the LAr forward calorimeter (FCal) within the range $3.1 < |\eta| < 4.9$. The TileCal uses steel as absorber material and scintillating tiles as active material. The HEC is a sampling LAr calorimeter with parallel plates of copper as absorber material.

3.2.3 The Muon Spectrometer

The muon spectrometer (MS) [86] surrounds the calorimeter system and is the outermost part of the ATLAS detector. It provides tracking information of muons to determine their momenta in the pseudorapidity region of $|\eta| < 2.7$. The MS consists of three parts:

- The barrel toroid and the endcap toroid provide a magnetic field of 0.5 T and 1 T, respectively, to bend the muon tracks and thus enable the momentum measurement.
- The cathode strip chambers (CSCs) operating in the range $2.0 < |\eta| < 2.7$ and the monitored drift tubes (MDTs) operating in the range $|\eta| < 2.7$, both provide tracking information with high spatial resolution of the outgoing muons. The CSCs are multi-wire proportional chambers installed in the endcaps. They have a high counting rate capability as well as an excellent spatial resolution of 50 μm per chamber. The MDTs consist of three layers of drift tubes with an aluminium casing providing a resolution of 35 μm per chamber.
- A set of trigger chambers with an accurate time resolution within the region of $|\eta| < 2.4$.

3.2.4 Trigger and Data Acquisition System

Due to large amounts of raw data per second produced by the ATLAS detector an effective trigger system, which selects the interesting events and rejects the other ones as much as possible, is a necessity. The ATLAS trigger and data acquisition system [87–89] reduces the output rate of events in real time. It consists of three trigger levels of which each refines the decision of the previous level:

- The hardware-implemented level-1 trigger is based on a calorimeter trigger processor for electrons, photons, jets, taus, $E_{\mathrm{T}}^{\mathrm{miss}}$, and total transverse energy, and a muon trigger processor for high transverse momentum muons in barrel and endcaps. It reduces the nominal bunch crossing rate of 20 MHz to 75 kHz within a total latency of 2 μs maximum, including the propagation delays on cables.

- The software-based level-2 trigger receives candidate events that are selected before by the level-1 trigger and stored in readout buffers. This high-level trigger refines the selection to an output rate of ~ 5 kHz within an average latency of up to ~ 10 ms.

- The software-based event filter (EF), the second high-level trigger, is the last step of the trigger system and implies a full offline event reconstruction. The event builder collects all fragments from one event, which are stored in readout buffers, into a single memory of an EF processor. As a result, the EF has access to the complete properties of an event and can apply time-consuming pattern recognition algorithms to reduce the final output rate of candidate events to ~ 600 Hz. The average latency in the EF is about 4 s. Since the EF provides an offline-like environment, the selection software of the EF is based on the ATLAS offline reconstruction and analysis framework.

Chapter 4

Data Samples and Samples of Simulated Events

The interpretation of almost all physics results at the LHC requires the modeling of hard interactions, i.e. the collisions of one proton with another proton at high energies. For this purpose a Monte Carlo (MC) event generator is utilised, which simulates high-energy particle-physics events in full detail down to the level of individual stable particles. An MC simulation samples configurations of the phase space, which is a representation of the coordinates and momenta of all particles, such that each phase-space dimension corresponds to a particle degree of freedom. The aim is to generate a large number of simulated events consistent with the kinematic properties of final-state particles produced in the real world and consistent with their probability to be produced. The final-state particles generated by MC event generators are used as input for a detector simulation program, which simulates interactions with the detector material to allow a precise prediction for the entire system of experimental setup. The ATLAS detector simulation includes the digitisation in which the electronic response is modelled, taking into account resolution and detector response effects. After the detector simulation, the event reconstruction is performed with the same algorithms as used for collision data.

4.1 Parton-Shower Monte-Carlo Event Generators

A typical MC event generator simulates a hard subprocess as shown as a big red blob in figure 4.1. The hard scattering process can be described using perturbation theory in which some partons, i.e. quarks and gluons, of the colliding protons interact at a high momentum scale. The momentum of these interacting constituents of the colliding particles are determined by sampling the parton distribution functions (PDFs), shown in figure 4.2. The parton distribution functions represent the probability of finding a specific parton with a given longitudinal momentum fraction x in a proton at the energy scale Q^2. The outgoing particles of the hard subprocess are fundamental objects of the SM or hypothetical particles of new theories.

All particles with colour charge produced during the hard subprocess can themselves emit further gluons. Gluons can split into quark-antiquark pairs leading to a parton shower (red lines in figure 4.1). Electroweak interactions, which are shown as small red blobs in figure 4.1, are also part of the parton shower. Even though the hard subprocess can be considered as the starting point of the main process, the order of the parton shower following the hard subprocess is not entirely correct. Radiation corrections like the initial-state radiations, which are also part of the showering model, occur before the hard scattering process. Initial-state radiations arise by a parton of the incoming protons emitting coloured or charged objects, which may give large corrections to the overall topology of the events. Initial-state radiations are shown as pink lines in figure 4.1 going off the lines, which produce the hard subprocess.

Since the higher-order corrections developed in the parton shower cannot be calculated exactly, an approximation scheme is used. In each order, the dominant contribution are calculated by the collinear parton splitting of a parton of the type i into two partons $i \to j + k$:

$$\mathrm{d}\sigma_{n+1} \approx \mathrm{d}\sigma_n \frac{\alpha_s}{2\pi} \frac{\mathrm{d}\theta^2}{\theta^2} \mathrm{d}z \, \mathrm{d}\phi \, P_{ij}(z, \phi) \, . \tag{4.1}$$

α_s is the strong coupling constant, θ and ϕ are the opening angle and the azimuthal angle of the outgoing particles, respectively, and $P_{ij}(z, \phi)$ is the $i \to j$ Altarelli-Parisi splitting function [92], which describes the distribution of the energy fraction z of i carried by j. The sequential application of equation (4.1)

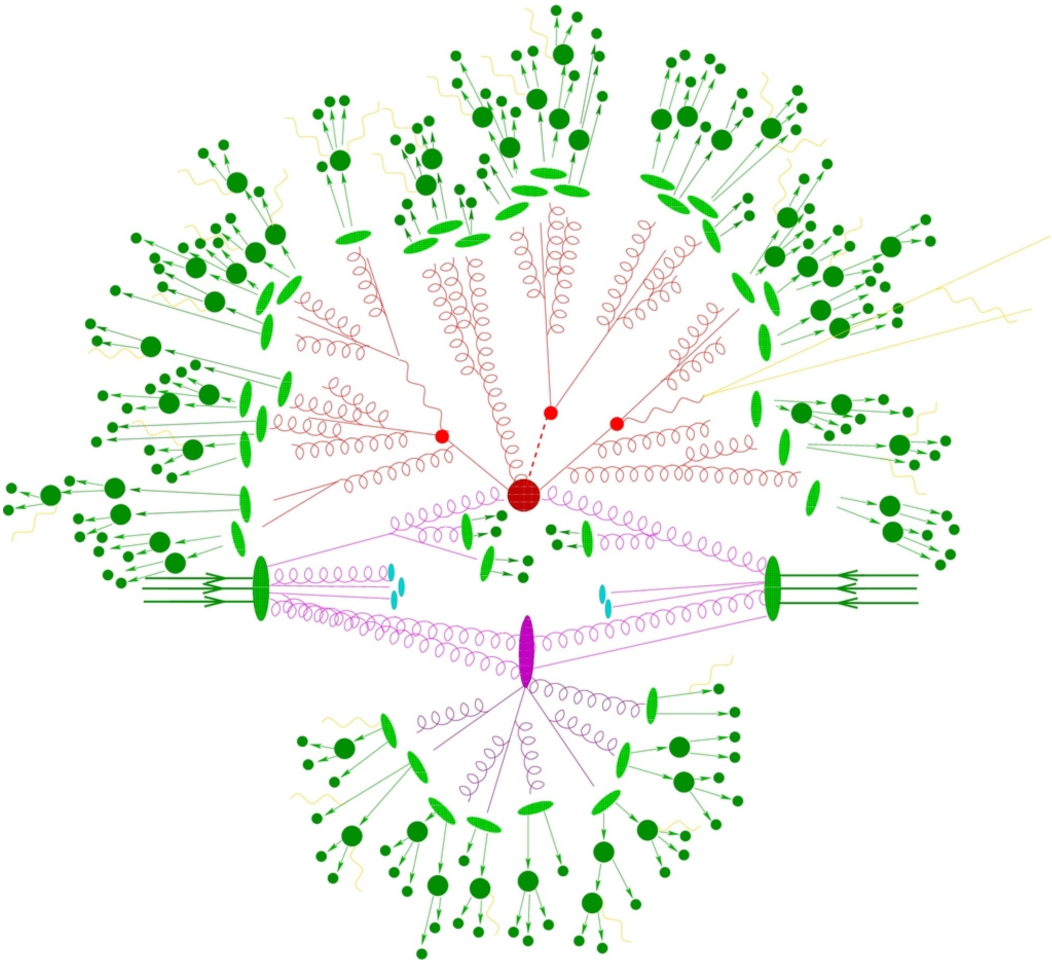


Figure 4.1: Simulation of a high-energy collision by a parton shower Monte Carlo event generator [90].

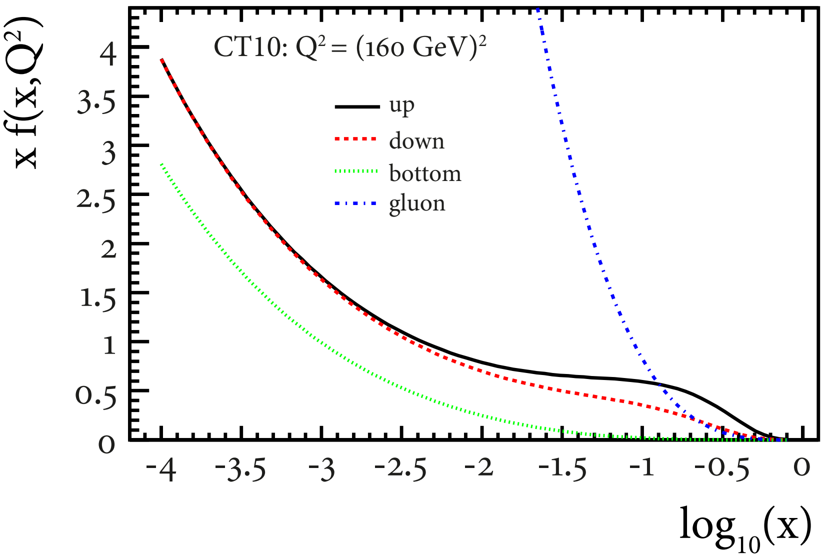


Figure 4.2: The `CT10` parton distribution functions [91] for up quarks (black), down quarks (red dashed), bottom quarks (green dotted) and gluons (blue dashed-dotted).

produces a parton shower for each coloured parton of the hard subprocess. For each iteration the splitting values of z, θ and ϕ are generated by MC generators. For more details the reader is referred to the publication [93] explaining the QCD factorisation theorem and the comprehensive handbook [94] describing perturbative QCD.

During the parton shower the interaction scale drops and the partons coalesce into composite states, the hadrons. This process is called hadronisation and cannot be calculated with perturbative techniques. The hadronisation process is shown in figure 4.1 as light green elliptic blobs. In contrast to the non-observable coloured particles occurring during the parton shower, the hadrons are colourless particles, which after subsequent decays become stable as shown as dark green blobs in figure 4.1.

As an approximation of the hadronisation process the Lund string model is used, which takes advantage of the fact that at large distances the potential energy of coloured particles increases linearly with the distance between them. Besides the string model, the cluster-fragmentation model based on a property called preconfinement [95], in which the partons are clustered in colourless quark-antiquark pairs, is applied in generators used at ATLAS.

The hard scattering leaves the coloured proton remnant behind, which takes the rest of the collision energy. The proton remnant is colour connected to the hard interaction and is therefore part of the same fragmentation model. The underlying event is the sum of all processes in a single particle collision not directly associated with the hard scattered subprocess. It is depicted as pink lines in figure 4.1 resulting in the pink elliptic blob, which also undergoes the parton shower and the hadronisation as described before.

At high-energy colliders like the LHC, it is important to consider the possibility that several proton-proton interactions can occur in the same bunch crossing. Typically, only one of these collisions occurs at a high momentum scale, resulting in an interesting event for high-energy physics research. The additional overlaid interactions are called pile-up events and are disturbing the reconstruction of the hard-scattering event, but need to be simulated properly to give a realistic representation of the LHC collision environment.

There are several matrix-element generators for the simulation of the hard subprocess available. The generated events of these matrix-element generators are fed to one of the three main general-purpose event generators, which provide showering and hadronisation. The following matrix-element generators are relevant for this analysis:

- `ALPGEN` [96] is dedicated to generate multiparton hard subprocesses. A computation of exact matrix elements for a large set of parton-level processes is performed at the leading order (LO) in QCD and electroweak interactions. The different emissions of `ALPGEN` and the parton-shower generator may produce the same final-state kinematics as shown in figure 4.3. To avoid this double counting a matrix-element and parton-shower matching scheme [97–100] has to be applied. For this purpose `ALPGEN` implemented the so-called MLM matching [101] in which the event is kept if each parton with large transverse momentum can be matched to a jet. Otherwise it is rejected. ALPGEN utilises helicity correlation in top-quark and gauge-boson decays, and includes masses for the bottom quark as well as for the top quark. The generated events feature full information on their colour structure and flavour structure.

- `GG2WW` [102] is an event generator for loop-induced WW production, e.g.

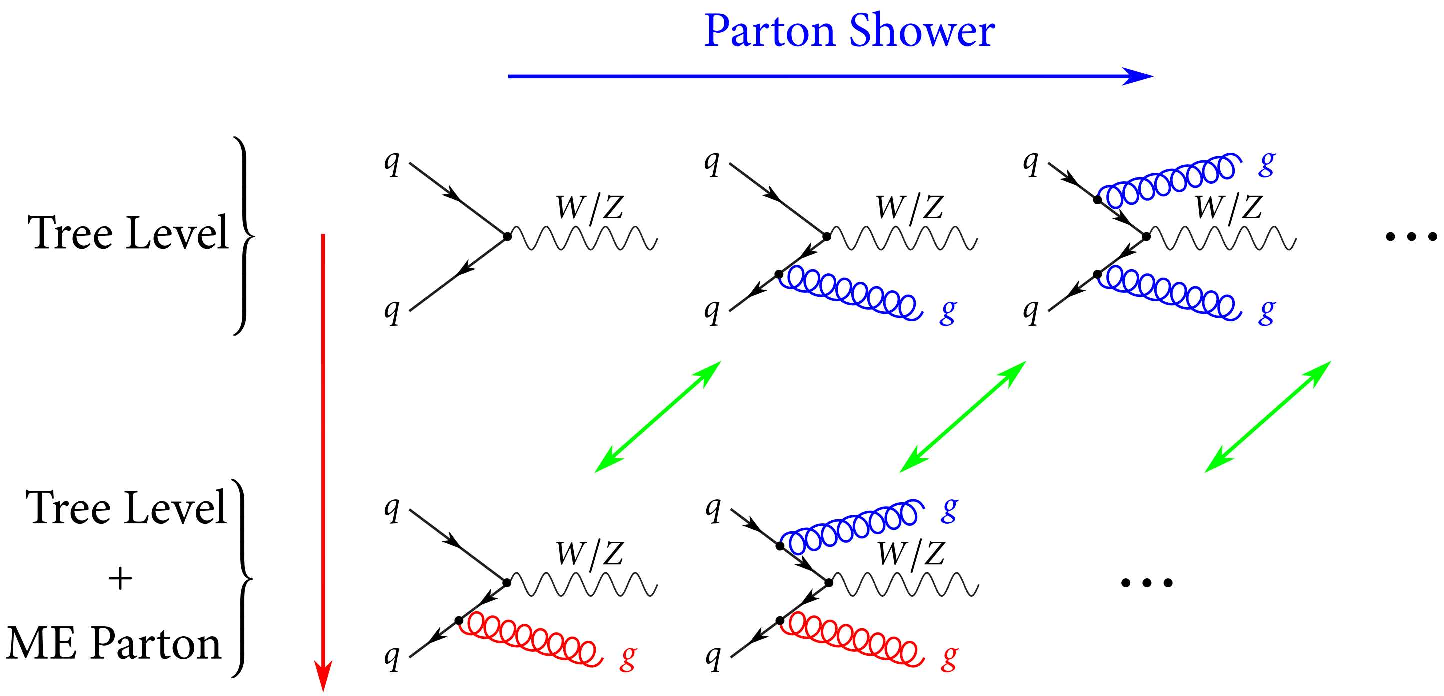


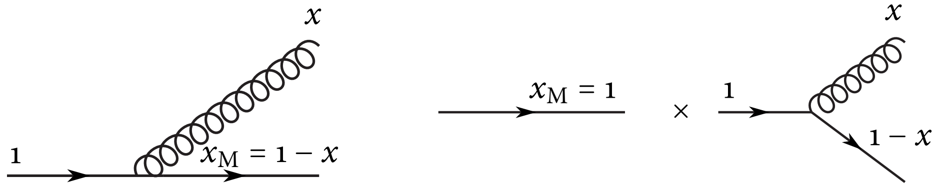
Figure 4.3: Pictorial view of double counting between the real-emission matrix element (bottom) and the parton shower (top).

$gg \to H \to W^+W^- \to$ leptons processes. Contributions from intermediate top-bottom massive-quark loops and the Higgs signal process are included. `GG2WW` computes cross sections and differential distributions, whereby arbitrary invariant masses of the W bosons are allowed.

- `MadGraph` [103–105] is a multipurpose matrix-element generator designed for any model that can be written in the form of a Lagrangian. For all relevant subprocesses that correspond to a specific final state the Feynman diagrams and the amplitudes are generated. `MadGraph` provides full automation and optimisation of NLO computations in the SM and in renormalisable theories beyond the SM.

- `MC@NLO` [106] is capable of computing partonic hard subprocess, including full NLO QCD corrections. Generated processes include the hadroproduction of single-vector bosons and Higgs bosons, vector-boson pairs, heavy-quark pairs, single top-quark production with and without an associated W boson or charged Higgs boson in type-I and type-II 2HDMs, lepton pairs, and Higgs bosons in association with a W boson or a Z boson.

  The absence of a well defined initial-parton configuration after the LO matrix-element calculation results in the before mentioned double counting. An additional problem occurs as part of the double counting by an NLO matrix-element generator, like `MC@NLO`, during the matching with the parton shower: There is no well-defined starting energy scale $x_{\mathrm{M}}(x)$ of the parton-shower generator. The NLO matrix element may generate a two-parton configuration with $x_{\mathrm{max}} < 1$ as shown in figure 4.4(a), or the Born and virtual parts of the NLO matrix element produce a one-parton configuration with $x_{\mathrm{max}} = 1$ as shown in figure 4.4(b). The Kinoshita-Lee-Nauenberg theorem [107, 108] only ensures the cancellation of divergences if one sums over all infrared and real-emission singularities of the same order. But the infrared divergences may occur in the real emission of the matrix-element generator as well as in the parton-shower generator. `MC@NLO` is specifically designed to work with the parton-shower MC simulation implemented in `HERWIG` [109] facing the matching of NLO+PS without double counting by a so-called modified subtraction.



(a): Real emission of matrix element.

(b): Born contribution supplemented by gluon emission in the parton shower.

Figure 4.4: An illustration of two different kinds of emission resulting in different starting energy scales in the parton shower.

- `AcerMC` [110] is an MC event generator simulating SM processes. The hard subprocess generated by `AcerMC` is completed by the initial-state and final-state radiation, hadronisation and decays using either `PYTHIA` or `HERWIG`. While the modified Kajantie-Byckling formalism [111] is used for phase-space construction, further smoothing of the phase space is obtained by using a modified `ac-VEGAS` algorithm [112]. This leads to an efficient phase-space sampling as an distinctive feature of `AcerMC`. In addition, the `AcerMC` generator incorporates the ACOT prescription [113–115] to consistently describe heavy quarks as massive particles in the initial state of specific processes, such as associated Zb production and t-channel single top-quark production.

- `POWHEG` [116, 117] is a shower-independent method to interfacing NLO QCD computations with parton-shower generators like `HERWIG` and `PYTHIA`. It is possible to generate only positive weighted events by using the exact NLO matrix elements.

The following multipurpose generators are used, mainly for parton-shower simulation and hadronisation:

- `HERWIG` [118] is a general-purpose MC event generator for simulation of hadronic final states. It includes the simulation of hard lepton-lepton, lepton-hadron and hadron-hadron scattering and soft hadron-hadron collisions. `HERWIG` provides an implementation of perturbative QCD as complete as possible, and includes a simple model of non-perturbative effects. The hadronisation process uses the before mentioned cluster-fragmentation model and a cluster-based simulation of the underlying event. Multiple parton interactions are not available within `HERWIG`. Therefore, the program `JIMMY` [119] is used to simulate the underlying event.

- `PYTHIA6` [120] is a prevalent general-purpose MC event generator with many implementations different from `HERWIG`. For the hadronisation mechanism the string model is used as opposed to the cluster-fragmentation model in `HERWIG`. The initial-state and final-state QCD-radiation algorithm is based on p_{T}-ordered evolution, while `HERWIG` uses angular ordering in the parton shower. The successor to `PYTHIA6` is `PYTHIA8`, which accommodates some new physics features. In `PYTHIA8`, the final-state evolution is interleaved with the initial-state evolution and the simulation of multiple interactions into one common sequence of p_{T}-decreasing evolution. The production of photons, charmonia and bottomonia, low-mass Drell-Yan pairs, and the t-channel $\gamma^*/Z/W^\pm$ exchange is also included during the calculation of multiple interactions.

4.2 Monte Carlo Samples

The signal of this analysis consists of two neutral Higgs bosons produced mainly via gluon-fusion and vector-boson-fusion processes with the subsequent decay $h/H \rightarrow WW^{(*)} \rightarrow e\nu\mu\nu$. The dominant backgrounds are non-resonant W^+W^- production and top-quark-antiquark pair production. Additional background processes are Z/γ^* + jets and the diboson processes $W(Z/\gamma^*)$ and ZZ.

In the presented analysis, the production of the two CP-even Higgs bosons h and H of the 2HDMs is modelled with samples of simulated events generated for SM Higgs-boson studies, which feature the same kinematic properties. The production cross sections are scaled according to the parameters of the 2HDMs as described in section 2.2.2. The mass of the light Higgs boson h is set to $m_h = 125$ GeV. The mass of the heavy Higgs boson H is varied between 135 and 300 GeV, using steps of 5 GeV between 135 and 200 GeV and steps of 20 GeV between 200 and 300 GeV. In this mass region, the natural width of the Higgs bosons, see figure 9.2, is negligible with respect to the experimental resolution.

The samples of simulated events for the Higgs-boson gluon-fusion and the VBF processes are generated with the `POWHEG` package, interfaced to `PYTHIA8` for showering and hadronisation as listed in table 4.1. The associated WH and ZH production processes are modelled using `PYTHIA8`, with the Higgs boson decaying to W^+W^-, while the associated W boson or Z boson decays to all possible modes. The W bosons that stem from the Higgs boson are forced to decay leptonically. The `CT10` set of PDFs is used for the gluon-fusion, the VBF, and the WH/ZH samples.

The cross section of the gluon-fusion process has been computed with NNLO QCD corrections [60, 121, 122], NLO electroweak corrections [123, 124], and corrections arising from the resummation of soft-gluon terms [125]. When rescaling the SM gluon-fusion cross section to a specific 2HDM, the different scaling of the top-quark-loop and bottom-quark-loop contributions, as well as their interference, is properly accounted for by using calculations that provide the relevant split of these three contributions [59, 126]. The expectation value of VBF Higgs-boson production is computed by using theoretical cross-section predictions that include full NLO QCD and electroweak corrections [127–129] and approximate NNLO QCD corrections [130]. The theoretical cross sections for the Higgs-strahlung processes WH and ZH are calculated with NNLO QCD corrections [131] and NLO electroweak corrections [132]. The relative uncertainty on the signal cross sections (gluon fusion, VBF and WH/ZH) is determined following Ref. [46, 133]. The relative cross-section uncertainty of these processes is assumed to be 25% for gluon fusion in the 0-jet channel and 30% in the 2-jet channel, while the uncertainty on the combined VBF/WH/ZH category is 10%. These uncertainties also account for uncertainties in the modelling of the underlying event. The cross-section uncertainties in the gluon-fusion process are treated uncorrelated between the 0-jet channel and the 2-jet channel, since they arise from different sources.

Inclusive W and Z/γ^* vector-boson production is simulated using the LO generator `ALPGEN`. The parton-level events are showered with `HERWIG` in connection with the `JIMMY` [119] underlying-event model. W+jets and Z+jets events with up to five additional partons are generated. The joint cross-section uncertainty for Drell-Yan production and Z+jets is 34%.

The irreducible diboson processes WW, $W(Z/\gamma^*)$ and ZZ are generated using `POWHEG`. The `POWHEG` events are showered with `PYTHIA6`. An additional contribution to the continuum WW background from gluon-initiated diagrams is modelled using gg2`WW` [102], also interfaced to `HERWIG` and `JIMMY`. The matrix-element generators `MADGRAPH` and `ALPGEN` (interfaced to `HERWIG`) are used to model $W(Z/\gamma^*)$ with $m_{(Z/\gamma^*)} < 7$ GeV and $W\gamma$, respectively. The treatment of the diboson backgrounds is strongly affected by theoretical uncertainties. A cross-section uncertainty of 25% is used for these backgrounds. The uncertainty corresponds to the difference between the LO and the NLO prediction.

Samples of the t-channel single top-quark process are produced with the `AcerMC` program linked to `PYTHIA6` for showering and hadronisation. The s-channel single top-quark process and Wt production are generated using `MC@NLO` and the resulting events are showered with `HERWIG+JIMMY`. Samples modelling $t\bar{t}$-pair production are generated with `MC@NLO+HERWIG+JIMMY`. All top-quark processes are produced with a top-quark mass of 172.5 GeV. The mere cross-section uncertainty of the top-quark processes is 6% [134–138].

Table 4.1 summarises the cross sections used to normalise each of the signal and background processes. More details on the event generators applied in this analysis and on the theoretical cross sections used to normalise these samples can be found in Ref. [139], describing the SM Higgs-boson analysis in the $h_{\mathrm{SM}} \to WW^{(*)}$ channel.

After the event-generation step, all samples are passed through the full simulation of the ATLAS detector [140] based on `GEANT4` [141] and are then reconstructed using the same procedure as for collision data. The simulation includes the effect of pile-up at a variable rate, and the events are weighted to match the conditions of the data sample using the average number of collisions per bunch crossing.

Table 4.1: Cross-sections at the centre-of-mass energy of $\sqrt{s}$ = 8 TeV for signal and background processes (given for $m_{h_{\text{SM}}}$ = 125 GeV in the case of the signal processes). The cross sections are combined with the branching fractions for processes for which a decay mode is specified. Processes are inclusive if no decay mode is specified. The generators used for the simulation of the various processes are also indicated.

Process	Generator (Tunes)	$\sigma \cdot \mathcal{B}$ [pb]
ggF	`POWHEG+PYTHIA8` (AU2)	0.441
VBF	`POWHEG+PYTHIA8` (AU2)	0.035
WH/ZH	`PYTHIA8` (AU2)	0.025
$q\bar{q}/g \to WW$	`POWHEG+PYTHIA8` (AU2)	5.68
$gg \to WW$	`GG2WW+HERWIG` (AUET2)	0.16
$t\bar{t}$	`MC@NLO+HERWIG` (AUET2)	238.1
tW/tb	`MC@NLO+HERWIG` (AUET2)	28
tqb	`AcerMC+PYTHIA6` (AUET2B)	88
Inclusive W	`ALPGEN+HERWIG` (AUET2)	$37 \cdot 10^3$
Inclusive $Z/\gamma^\star$	`ALPGEN+HERWIG` (AUET2)	$16 \cdot 10^3$
$Z^{(*)}Z^{(*)} \to 4l$	`POWHEG+PYTHIA8` (AU2)	0.73
$W(Z/\gamma^*)(m_{(Z/\gamma^*)} > 7)$ GeV	`POWHEG+PYTHIA8` (AU2)	3.63
$W(Z/\gamma^*)(m_{(Z/\gamma^*)} < 7)$ GeV	`MADGRAPH+PYTHIA6` (AUET2B)	14.3
$W\gamma$	`ALPGEN+HERWIG` (AUET2)	369

4.3 Data Samples

The analysis described in this thesis uses LHC proton-proton collision data at a centre-of-mass energy of 8 TeV collected with the ATLAS detector [142] between April and September 2012. Stringent detector and data quality requirements are applied, resulting in a dataset corresponding to an integrated luminosity of 13.0 fb^{-1}.

The detector control system is used to oversee the conditions of the individual sub detectors. The information about the quality of the data taking is stored in the ATLAS conditions database and is summarised to provide data quality status flags. In physics analyses, the data-quality information is used through dedicated lists of runs and luminosity blocks, the so-called good

Table 4.2: Data-taking periods and corresponding integrated luminosity.

Data Period	Corresponding Luminosity
A	0.740 fb^{-1}
B1-B3	0.299 fb^{-1}
B4-B8	1.964 fb^{-1}
B8-B11	1.194 fb^{-1}
B12-B14	1.640 fb^{-1}
C1-C9	1.321 fb^{-1}
D1-D8	3.301 fb^{-1}
E1-E5	2.281 fb^{-1}

run lists. A luminosity block is a small data block of roughly two minutes of data taking, in which the luminosity is supposed to remain approximately constant.

Events used in this analysis are required to be recorded with optimal data-taking conditions for the detector system. For all events the good run list denoted as *data12_8TeV.periodAllYear_DetStatus-v53-pro13-04_CoolRunQuery-00-04-08_All_Good.xml* is used. The *data12* and *8TeV* means that the proton-proton collision data was collected in 2012 at a centre-of-mass energy of 8 TeV. The number in *v53* and in *pro13-04* represent the data-quality virtual-flag version and the ATLAS-metadata-interface tag corresponding to the processing campaign, respectively. The term *CoolRunQuery-00-04-08* indicates the version of the defect logic used to calculate if data are good, and the suffix *All_Good* gives the name of the good run list template. The total number of events included in this good run list is 3,149,698,147 divided in 308 runs.

The data run is divided in data periods, such that they represent data with a coherent configuration of the detector and the trigger and data acquisition system. These data periods are listed in table 4.2.

Candidate events of the $h/H \to WW^{(*)} \to e\nu\mu\nu$ process are recorded using unprescaled single-electron [143, 144] and single-muon [145, 146] triggers with a p_{T} threshold of 24 GeV.

- **Electron trigger**
  For electrons a logical "OR" of *EF_e24vhi_medium* and *EF_e60_medium* triggers is used. The numbers after *EF_e* in the trigger names stand for the nominal p_{T}-threshold values in GeV. The tightness in the electron identification criteria is indicated by the suffix *medium*. The *vh* means that the trigger has η-dependent p_{T} thresholds and a hadronic leakage cut between one and two percent at the level 1 trigger. The *i* means that an isolation criterion is applied during the online event selection. In case of the *EF_e24vhi_medium* trigger a relative track isolation cut is used at the EF selection stage, see section 3.2.4:

  $$\frac{\sum p_{\mathrm{T}}}{p_{\mathrm{T}}(e)} < 0.1$$

  $\sum p_{\mathrm{T}}$ is the scalar sum of the transverse momenta of tracks having $p_{\mathrm{T}} >$ 1 GeV found in the inner detector in a cone of $\Delta R = 0.2$ around the electron candidate, after subtracting the p_{T} of the electron.

- **Muon trigger**
  For muons a logical "OR" of *EF_mu24i_tight* and *EF_mu36_tight* triggers is used. The number after *EF_mu*, the suffix *tight* and the *i* after the threshold value in the trigger names have the meaning as explained for the electron trigger. For the *EF_mu24i_tight* trigger a relative track isolation cut is used at the EF selection stage:

  $$\frac{\sum p_{\mathrm{T}}}{p_{\mathrm{T}}(\mu)} < 0.12$$

  $\sum p_{\mathrm{T}}$ is the scalar sum of the transverse momenta of tracks having $p_{\mathrm{T}} >$ 1 GeV found in the inner detector in a cone of $\Delta R = 0.2$ around the muon candidate, after subtracting the p_{T} of the muon.

Chapter 5

Physics Objects and Event Reconstruction

The large number of around 100 million individual readout channels in a high-occupancy environment with on average 20 inelastic proton-proton interactions per bunch crossing (see figure 5.1) requires a sophisticated processing of the information of the various detector components to form objects

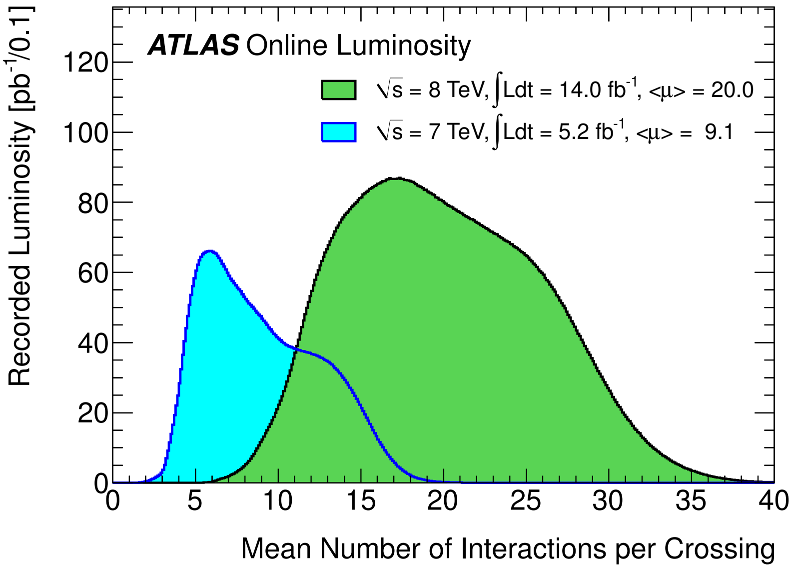


Figure 5.1: The luminosity-weighted distribution of the mean number of interactions per crossing for the 2011 and 2012 data [147, 148].

like charged particle tracks and energy clusters. These reconstructed objects are used to select, identify and build high-level physics objects like electrons, muons and jets to reconstruct the event. After particle identification and event reconstruction an optimised event selection is necessary to enhance the fraction of signal candidate events.

5.1 Object Reconstruction

A calorimeter cell (CaloCell), providing only one energy measurement, is the smallest data object used by the electromagnetic calorimeter as well as the hadronic calorimeter. Each CaloCell is uniquely identified by its geometrical position in an η-ϕ grid and includes the information of time and quality of the measured signal. In order to reconstruct the full energy of prototype physics objects, the second step of calorimeter reconstruction is to build calorimeter clusters (CaloCluster) by combining several CaloCells. To produce a CaloCluster, ATLAS uses the topological clustering algorithm [149]. This algorithm is designed to group all neighbouring CaloCells with signals sufficiently above the expected noise and to sum the total deposited energy within each cluster. The topological clustering algorithm also includes a mechanism to search for local maxima in terms of energy and split these topologically connected areas originating from different close-by particles [150].

Charge particle tracks are reconstructed in three subsequent steps: seed finding, track finding and track fitting. Pattern recognition algorithms are used as seed finder by combing triples of hits in the pixel layer with hits in the first SCT layer and checking if they constitute a valid track candidate. The parameters of each track like position, direction and bending are then transported through the detector layers by including real hits one-by-one via the Kalman algorithm [151]. In the last step, a final track fit based on the Kalman-filter results is performed and applies corrections to optimally determine the found track candidates and resolve ambiguities where track candidates share hits. At the end of the track fit procedure the quality of the track candidate is evaluated to decide whether it is accepted as a real particle track.

A typical collision event features several primary vertices (PVs) as well as secondary vertices (SVs) arising from decays of long-lived particles like bottom quarks or photon conversions. PVs are space points where inelastic proton-proton interactions have occurred.

It is important for all physics analyses to identify and find the location of the PV that stems from the hard inelastic subprocess of interest and to distinguish this PV from the several additional vertices originating from low-p_T proton-proton collision events.

The reconstruction of PVs can be subdivided in two tasks: PV finding and PV fitting. The before mentioned reconstructed tracks are used as input to the vertex finder. While the finder associates the tracks to a particular vertex candidate, the PV fitter is responsible for the reconstruction of the vertex position, the corresponding covariance matrix and the quantities related to the quality of the fit. Since these two processes are not easily distinguishable from each other, two approaches are implemented in ATLAS: the fitting-after-finding approach and the finding-through-fitting approach.

The *InDetPriVxFinder* algorithm [152] implements the fitting-after-finding approach. Here the reconstructed tracks compatible with the beam spot, i.e. the luminous interaction region, are preselected and then clustered. The centre of the cluster is defined as the z value at the point of the closest approach. The obtained vertex candidates are then iteratively fitted whereby the outliers, i.e. tracks for which the probability of the χ^2 between the vertex estimate and the trajectory in question is less than 8%, are rejected in each iteration. The fitting procedure ends as soon as no incompatible tracks are left or the cluster size becomes too small to continue.

The *AdaptiveMultiVertexFinder* algorithm [153, 154] is used as the finding-through-fitting approach, which is the default strategy for PV finding in ATLAS, because it provides a better treatment of outliers. In the first step of the *AdaptiveMultiVertexFinder* algorithm, all reconstructed tracks that are compatible with the beam spot are pre-selected. These pre-selected tracks are then fitted to an independent PV candidate, which is formed by the obtained track cluster in z projection, i.e. the z impact parameter. In contrast to the fitting-after-finding approach, reconstructed tracks that are classified as outliers during the first iteration are used to create a new vertex seed as a PV candidate. In the next iteration, the two vertices undergo a simultaneous adaptive fit in which both candidates compete with each other in order to gain more tracks. At the end of the fit all reconstructed tracks that do not enter any vertex with a probability of $\chi^2 > 1\%$ have no influence on the fit result. The vertex with the highest $\sum p_{\mathrm{T}}^2$ of constituent tracks is selected as the hard-scatter PV.

5.2 Particle Identification

Starting from the basic building blocks described in the last section high-level objects are identified and reconstructed. The analysis presented in this thesis makes use of electron and muon candidates, jets, and $E_{\mathrm{T}}^{\mathrm{miss}}$ (missing transverse momentum) signalling the presence of neutrinos which do not interact with the detector. A number of requirements are applied in order to ensure a high quality of reconstructed particles and to remove mis-reconstructed events. Events are selected if they contain at least one good PV candidate with at least three associated tracks with $p_{\mathrm{T}} > 0.4$ GeV. Events that are compatible with non-collision backgrounds are excluded, this includes cosmic rays passing through the detector, beam backgrounds, and events with special instrumental effects like noise bursts in certain parts of the detector.

5.2.1 Electron Candidates

A good electron identification is crucial inasmuch as the majority of the W+jets background originates from fake electrons. The electron candidates are required to fulfil stringent criteria regarding calorimeter shower shape, track quality, track-cluster matching, and transition radiation energy to ensure high identification quality.

The electron isolation based on the topological cluster algorithm [149] is used to reduce the pile-up dependence. The energy of an electron candidate is taken from the LAr electromagnetic calorimeter by summing the energy of uncalibrated topological clusters contained in a cone of size $\Delta R = 0.3$, which is represented as a yellow circle in figure 5.2. A rectangle of 5×7 electromagnetic calorimeter cells centred on the electron candidate is removed to subtract the energy of the electron itself. The energy that leaks outside of this rectangle is called the isolation energy of the electron candidate. The measured isolation energy is modified by a correction derived from a single electron MC sample. The correction is applied separately in at most 10 bins in pseudo-rapidity. In addition, the scalar sum of the transverse momenta of tracks having $p_{\mathrm{T}} > 1$ GeV

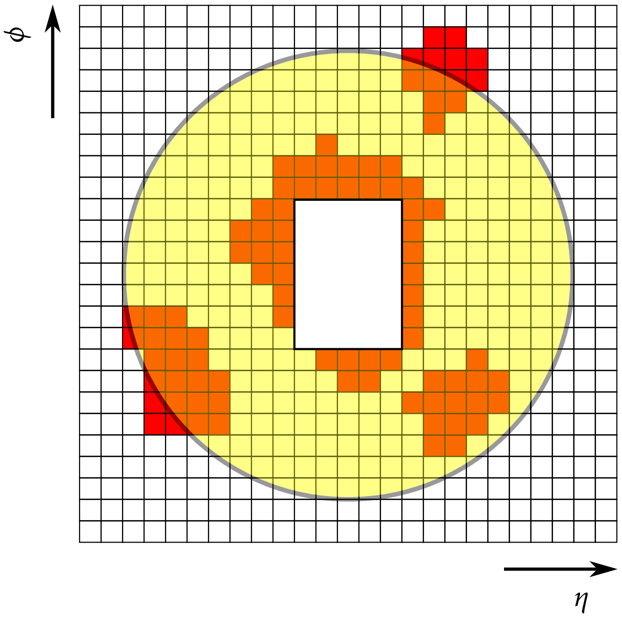


Figure 5.2: A pictorial view of the computation of topological calorimeter isolation $\sum(p_{\mathrm{T}}^{\mathrm{calo}})$. The grid represents the electromagnetic calorimeter granularity.

in a cone of $\Delta R = 0.3$, after subtracting the p_{T} of the electron, is used as a tracking based discriminator. The advantage of the track isolation criteria is the ability to reject tracks from the SVs. Electron candidates are required to be isolated by placing cuts on the isolation energy of the calorimeter $\sum(p_{\mathrm{T}}^{\mathrm{calo}})$ and the isolation energy of the tracking system $\sum(p_{\mathrm{T}}^{\mathrm{track}})$. For an electron candidate the fractional track based isolation $\sum(p_{\mathrm{T}}^{\mathrm{track}})/p_{\mathrm{T}}$ is required to be below 0.16, and this is tightened to 0.12 for candidates with p_{T} below 25 GeV which have more background, while $\sum(p_{\mathrm{T}}^{\mathrm{calo}})/p_{\mathrm{T}}$ is required to be below 0.16.

If two electron candidates are reconstructed within a cone of $\Delta R = 0.1$, the candidate with the lower p_{T} is discarded. Electron candidates with transverse energy above ~ 3 GeV are selected from topological clusters and matched to tracks. η and ϕ of the electron candidate stem from the track of the inner detector with at least one hit in the pixel layer and at least six hits in the SCT.

It is further required that the electron track points back to the PV with

Table 5.1: Electron-isolation and impact-parameter cuts.

Cut	Value
$\sum(p_{\mathrm{T}}^{\mathrm{calo}})/p_{\mathrm{T}}$	< 0.16
$\sum(p_{\mathrm{T}}^{\mathrm{track}})/p_{\mathrm{T}}$	$\begin{cases} < 0.12 & \text{if } 15\ \mathrm{GeV} < p_{\mathrm{T}} < 25\ \mathrm{GeV} \\ < 0.16 & \text{if } p_{\mathrm{T}} > 25\ \mathrm{GeV} \end{cases}$
d_0 significance	< 3.0
$z_0 \cdot \sin\theta$	< 0.4 mm

transverse impact-parameter significance, i.e. the ratio of the transverse impact parameter to its uncertainty, less than 3. The longitudinal impact parameter z_0 of the track has to be compatible with the PV such that $z_0 \cdot \sin\theta < 0.4$ mm. Electron candidates are excluded if they lie in the pseudorapidity ranges $1.37 < |\eta| < 1.52$ or $|\eta| > 2.47$. The transverse energy is required to be larger than 15 GeV. Table 5.1 summarises the electron-isolation and impact-parameter cuts.

5.2.2 Muon Candidates

Muons are reconstructed by taking a statistical combination of the matched tracks in the Inner Detector (ID) and in the Muon Spectrometer (MS). The muon candidates are required to have a sufficiently large number of hits in the silicon detector and the Transition Radiation Tracker (TRT) to fulfil the highest purity of all muon types in ATLAS. Muon tracks are required to have at least one hit in the pixel detector, and four or more hits in the Semiconductor Tracker (SCT). Tracks are vetoed if they have more than two missing hits in the SCT and pixel detectors, as well as tracks with an excessive number of outlier hits, i.e. hits far away from the fitted tracks, in the TRT. Within the pseudorapidity region $0.1 < |\eta| < 1.9$ more than five TRT hits are required, $n_{\mathrm{TRT}}^{\mathrm{hits}} > 5$, and the number of TRT outliers on the muon track must satisfy $n_{\mathrm{TRT}}^{\mathrm{outliers}} < 0.9n$, where $n = n_{\mathrm{TRT}}^{\mathrm{hits}} + n_{\mathrm{TRT}}^{\mathrm{outliers}}$. For $|\eta| \geq 1.9$ it is required that $n_{\mathrm{TRT}}^{\mathrm{outliers}} < 0.9n$, if $n > 5$.

The momentum as measured using the ID is required to agree with the momentum measured using the MS after correcting for the predicted muon

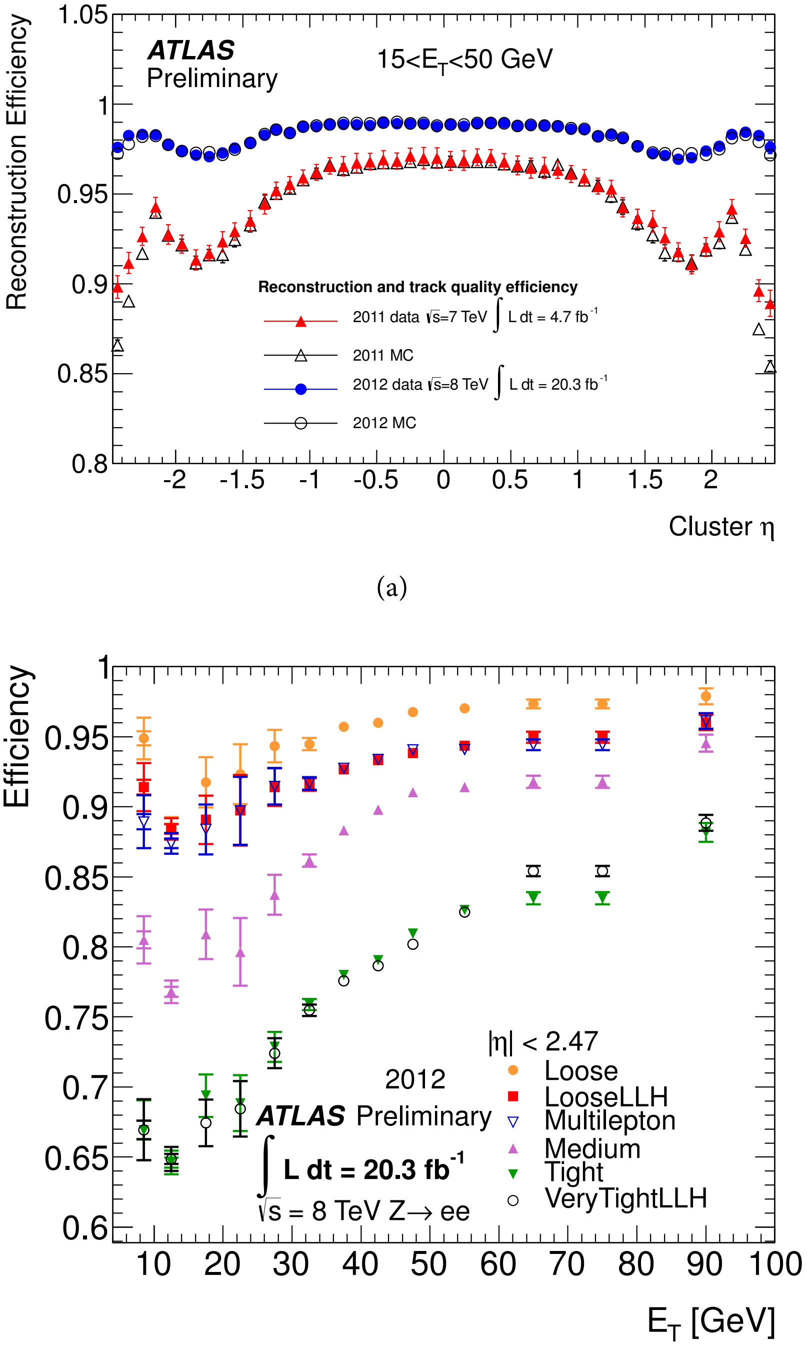
Table 5.2: Muon-isolation and impact-parameter cuts.

Cut	Value
$\sum(p_{\mathrm{T}}^{\mathrm{calo}})/p_{\mathrm{T}}$	$\begin{cases} < (0.014 \cdot p_{\mathrm{T}}/\mathrm{GeV}) - 0.15 \text{ and} \\ < 0.20 \end{cases}$
$\sum(p_{\mathrm{T}}^{\mathrm{track}})/p_{\mathrm{T}}$	$\begin{cases} < (0.01 \cdot p_{\mathrm{T}}/\mathrm{GeV}) - 0.105 \text{ and} \\ < 0.15 \end{cases}$
d_0 significance	< 3.0
$z_0 \cdot \sin\theta$	< 1.0 mm

energy loss in the calorimeter. Muon candidates are required to have $p_{\mathrm{T}} > 15$ GeV and $|\eta| < 2.4$, and must satisfy a similar isolation requirement. z_0 of the track has to be compatible with the PV such that $z_0 \cdot \sin\theta < 1.0$ mm. The required calorimeter isolation is $\sum(p_{\mathrm{T}}^{\mathrm{calo}})/p_{\mathrm{T}} < (0.014 \cdot p_{\mathrm{T}}/\mathrm{GeV}) - 0.15$ and $\sum(p_{\mathrm{T}}^{\mathrm{calo}})/p_{\mathrm{T}} < 0.20$, while for the track based isolation $\sum(p_{\mathrm{T}}^{\mathrm{track}})/p_{\mathrm{T}} < (0.01 \cdot p_{\mathrm{T}}/\mathrm{GeV}) - 0.105$ and $\sum(p_{\mathrm{T}}^{\mathrm{track}})/p_{\mathrm{T}} < 0.15$ is required. If an electron lies in a cone of $\Delta R = 0.1$ around a reconstructed muon, the electron candidate is ignored. Table 5.2 shows the muon-isolation and impact-parameter cuts.

5.2.3 Efficiency Calibration

The reconstruction, identification and trigger efficiencies of electron and muon candidates are measured using tag-and-probe methods on samples enriched with $Z \rightarrow \ell\ell$, $J/\psi \rightarrow \ell\ell$, or $W^{\pm} \rightarrow \ell\nu$ ($\ell = e, \mu$) events [155–157] in order to correct for differences between MC and data. Figure 5.3(a) shows the improvement of the electron reconstruction efficiency in 2012 as a function of the electron cluster η with respect to the 2011 data taking. The electron reconstruction efficiency is shown for data (filled markers) and MC (open markers) with E_{T} in between 15–50 GeV. In the barrel region, the efficiency is increased in 2012 by ~ 2% compared to 2011, while the efficiency increase amounts to ~ 8% in the endcap region. Except some smaller discrepancies in the forward region, figure 5.3(a) shows a good agreement between data



(a)

(b)

Figure 5.3: Reconstruction efficiency (a) and identification efficiency (b) of electrons are shown as a function of the pseudorapidity η and transverse energy E_{T}, respectively [158].

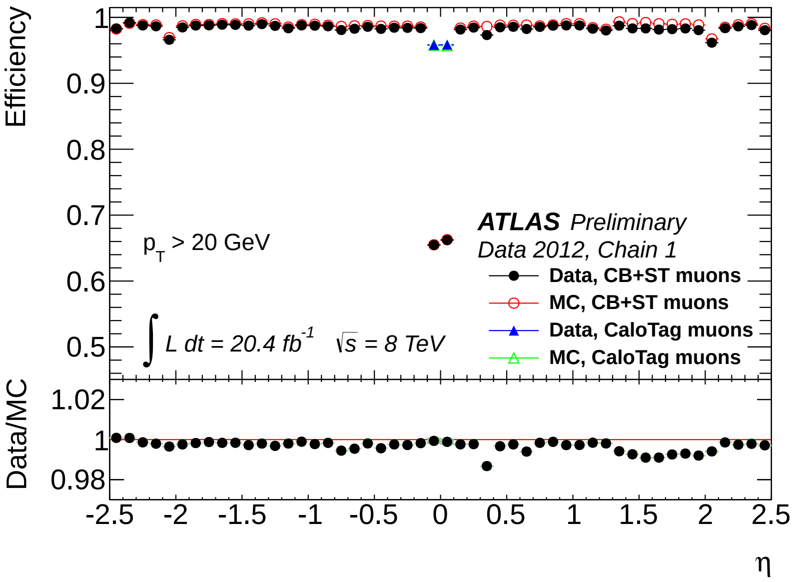


Figure 5.4: The reconstruction efficiency of muons is shown as a function of the pseudorapidity η [157]. The CaloTag muons are used only in the central region $|\eta| < 0.1$ while elsewhere the combined (CB) muon is used. The lower histogram shows the ratio between the measured and predicted efficiencies.

and MC. The electron identification efficiency in 2012 as a function of the electron E_T is shown in figure 5.3(b) with different sets of quality cuts and a pseudorapidity of $|\eta| < 2.47$.

The reconstruction efficiency of muons with $p_\mathrm{T} > 20$ GeV is shown as a function of muon η in figure 5.4. This efficiency is strongly affected by two main gaps of acceptance in the MS. These acceptance losses occur in the region at $2.0 < \eta < 2.1$ and in the region at $|\eta| < 0.1$, where the MS is only partially equipped. In the region at $1.5 < \eta < 2.2$, the modelling of the ID reconstruction efficiency is not correct due to non-operating pixel modules.

5.2.4 Jet Candidates

Jets are reconstructed from topological CaloClusters using the anti-k_t algorithm [159]. The anti-k_t jet-finding algorithm belongs to the class of sequentialrecombination algorithms. These clustering algorithms are all based on a specific way to measure the distance between all pairs of entities (particles, pseudojets) i and j

$$d_{ij} = \min\left(p_{\mathrm{T},i}^{2m}, p_{\mathrm{T},j}^{2m}\right) \frac{\left(\Delta R_{ij}\right)^2}{R^2} \tag{5.1}$$

and the beam distance between all entities i and the beam of the LHC

$$d_{iB} = p_{\mathrm{T},i}^{2m} . \tag{5.2}$$

The parameter m defines the specific algorithm: $m = -1$ corresponds to the anti-k_t algorithm, while $m = 0$ and $m = +1$ corresponds to the Cambridge/Aachen algorithm and the (inclusive-)k_t algorithm, respectively. The radius parameter $R = 0.4$ is used. For each entity i the smallest distance of all d_{ij} and d_{iB} has to be identified. If a d_{ij} is the smallest one, then i and j are combined to a new pseudo particle. If d_{iB} is the smallest distance, then i is removed from the list of candidates and called a jet.

The anti-k_t algorithm as well as the Cambridge/Aachen algorithm and the (inclusive-)k_t algorithm are infrared and collinear safe to all orders of perturbation theory. Collinear unsafe means that two hard particles 1 and 2 with transverse momenta $p_{\mathrm{T},1} > p_{\mathrm{T},2}$ may change their p_{T} order, when the hardest particle is split into a nearly-collinear pair $(p_{\mathrm{T,1a}}, p_{\mathrm{T,1b}})$. The consequence would be a drastically different clustering of particles to jets as shown in figure 5.5. This effect originates from the squared matrix element, which is proportional to $1/\left(p_1 + p_2\right)^2 = 1/2E_1E_2(1 - \cos^2\theta)$ and is collinear divergent for $\theta \to 0$ and $\theta \to \pi$. E_i and p_i are the energies and momenta of the pseudo particles, respectively, and θ is the solid angle between both particles. Infrared-unsafe algorithms may also produce a different final set of jets. Adding a new soft seed particle may lead to new stable cones, see figure 5.5. This corresponds to the infrared divergence of the squared matrix element for $E_i \to 0$.

After the application of the jet-finding algorithm, the average additional energy due to pile-up events is subtracted from the energy of the prototype jets measured in the calorimeter. After that, the prototype jets need to be calibrated to correct the reconstructed energy that may be impaired due to a non-uniform detector response, instrumental inefficiencies, e.g. dead detector material and

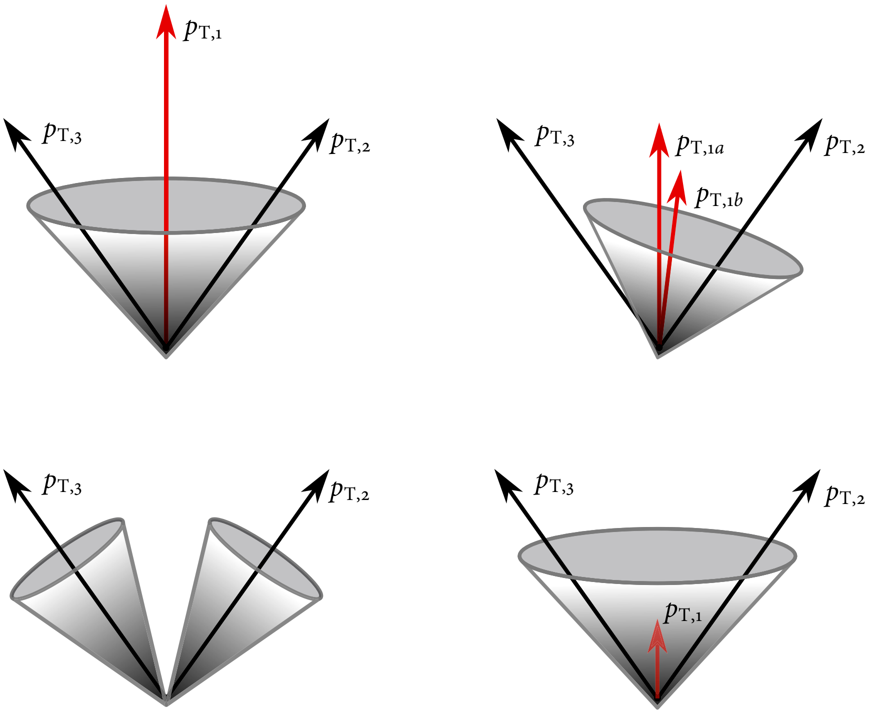


Figure 5.5: Depiction of a collinear-unsafe (top) and an infrared-unsafe (bottom) jet-finder algorithm.

gaps between several subdetectors, and detector independent effects, e.g. out-of-cluster energy and the underlying event [150, 160].

In a first step, the jet energy of the reconstructed jets is calibrated. This calibration is based on truth jets, which are formed by using the same jet finding algorithm on stable MC truth particles in simulated events. The corrected jet energy is calculated by

$$E_{\text{jet}}^{\text{reco}} = \sum_{i\in\text{jet}} w_i(\rho_i)E_i\,, \tag{5.3}$$

where w_i is a weighting function, which depends on the energy density ρ_i in cell i of the calorimeter, and E_i is the energy deposited in the ith cell. After that first step of calibration the jet energy scale does not depend any more on the characteristics of the calorimeter.

In the second step, the detector independent effects are corrected through an in-situ validation using suitable physics processes. The validation procedure of the jet energy scale checks the uniformity of the calibration as a function of ϕ and η by looking at QCD dijet events. The absolute hadronic energy scale is determined using Z/γ^*+jets events. The momentum is balanced between the jet and the Z/γ^*-boson. The Z/γ^*-boson is reconstructed via the decay to two electrons or two muons. An improved calibration can be derived by relating the hadronic energy scale of the jets to the better understood energy of the electromagnetic-interacting objects.

The selected jets are required to have $p_\mathrm{T} > 25$ GeV at the calibrated hadronic energy scale. This threshold is increased to 30 GeV in the forward region ($|\eta| > 2.5$) which is more sensitive to misreconstruction of jets arising from pile-up events. Only jets within $|\eta| < 4.5$ are used. If a jet candidate lies within a cone of $\Delta R = 0.3$ around a reconstructed electron candidate, the jet candidate is ignored.

To reject jets from pile-up events, a quantity called jet-vertex fraction (ε_JVF) [161] is defined as the ratio of $\sum p_{\mathrm{T},i\in\mathrm{PV}}$ for all tracks within the jet that originate from the PV associated to the hard-scattering collision to the $\sum p_{\mathrm{T},i}$ of all tracks matched to the jet:

$$\varepsilon_\mathrm{JVF}(j, \mathrm{PV}) = \frac{\sum p_{\mathrm{T},i\in\mathrm{PV}}}{\sum p_{\mathrm{T},i}} . \tag{5.4}$$

See figure 5.6 for an illustration of the jet-vertex-fraction discriminant. It is required that $\varepsilon_\mathrm{JVF} > 0.5$. For jets having no matched tracks, this criteria is omitted.

Selected jets are identified as b-quark jets by reconstructing SVs (see figure 5.7) from the tracks associated with each jet and combining lifetime related information with an NN [162]. The jets passing the identification requirements are called b-tagged jets. The chosen identification requirement has an efficiency of 85% for b-quark jets in $t\bar{t}$ events.

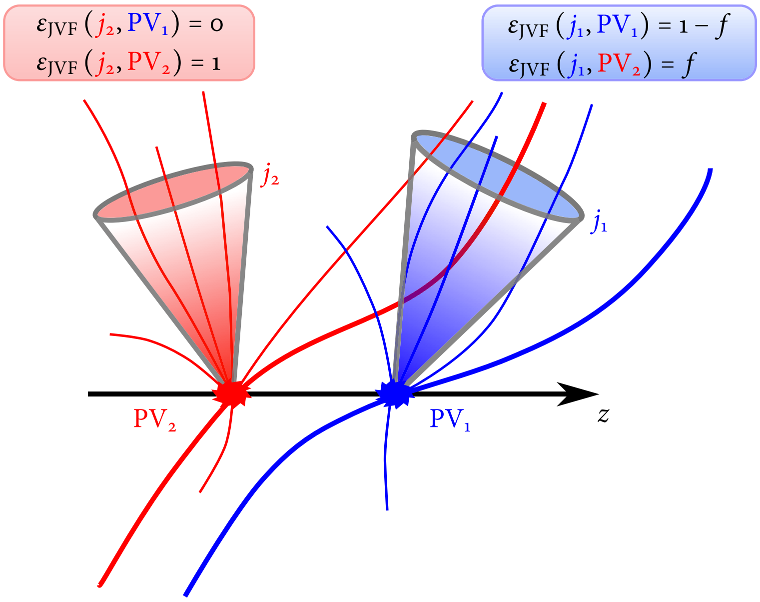


Figure 5.6: A pictorial view of the jet-vertex-fraction discriminant.

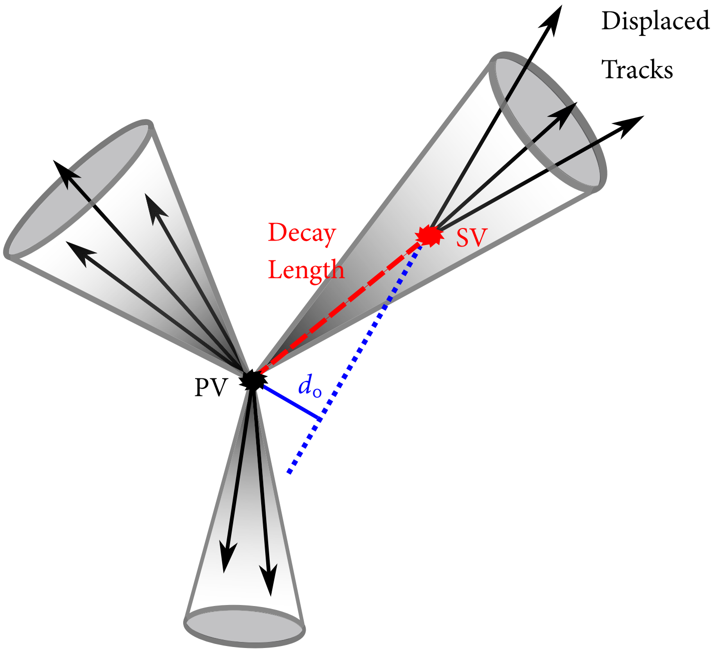


Figure 5.7: An illustration of a secondary-vertex reconstruction. d_0 stands for the impact parameter.

5.2.5 Missing Transverse Energy

The $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ is reconstructed starting from topological energy clusters in the calorimeters [163]. Corrections for measured muons are applied. To reduce the effect of mismeasurements of jets and leptons leading to artificial $E_{\mathrm{T}}^{\mathrm{miss}}$, the event selection uses the $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ variable which is defined as

$$E_{\mathrm{T,rel}}^{\mathrm{miss}} = \begin{cases} E_{\mathrm{T}}^{\mathrm{miss}} \sin(\Delta\phi_{\mathrm{min}}) & \text{if } \Delta\phi_{\mathrm{min}} < \pi/2 \\ E_{\mathrm{T}}^{\mathrm{miss}} & \text{if } \Delta\phi_{\mathrm{min}} \geq \pi/2 \end{cases} \tag{5.5}$$

where $\Delta\phi_{\mathrm{min}}$ is the minimum separation in the azimuth angle ϕ between any lepton or any jet with $p_{\mathrm{T}} > 25$ GeV and $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$. Figure 5.8 illustrates the effect of mis-reconstructing the energy of the nearest object $\vec{\ell}\,/\,\vec{j}$ on $E_{\mathrm{T}}^{\mathrm{miss}}$ as well as on $\vec{E}_{\mathrm{T,rel}}^{\mathrm{miss}}$. $\vec{E}_{\mathrm{T,rel}}^{\mathrm{miss}}$ is less prone to variation of the energy of the nearest objects due to misreconstruction.

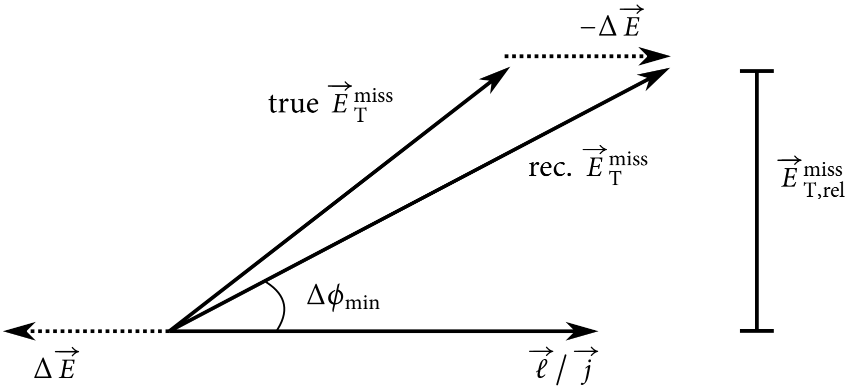


Figure 5.8: Illustration of $E_{\mathrm{T,rel}}^{\mathrm{miss}}$, which is the magnitude of the projection of $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ onto the direction perpendicular of the nearest object, $\vec{\ell}\,/\,\vec{j}$, in the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$-$\left(\vec{\ell}\,/\,\vec{j}\right)$ plane. The misreconstruction of the energy of $\vec{\ell}\,/\,\vec{j}$ leads to $-\Delta\vec{E}$, which changes the direction and the magnitude of the true $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$. However, $\vec{E}_{\mathrm{T,rel}}^{\mathrm{miss}}$ is invariant under misreconstruction of the energy of $\vec{\ell}\,/\,\vec{j}$.

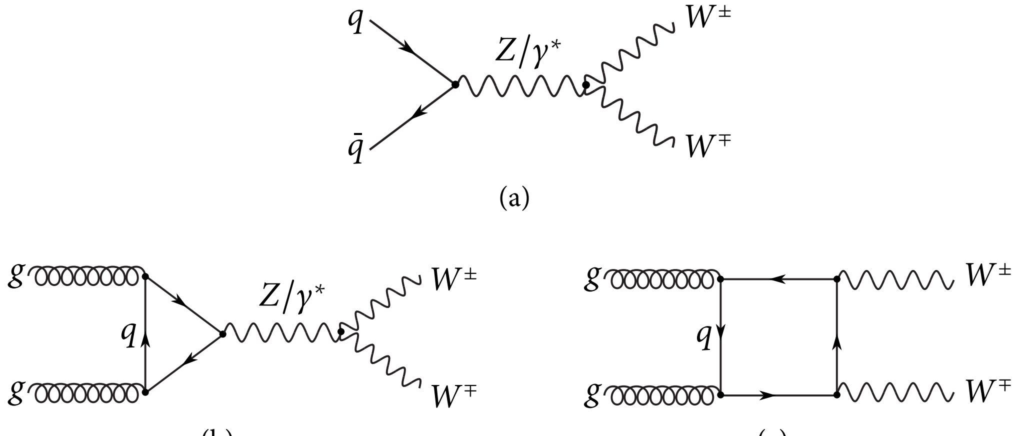
5.3 Event Selection

The presence of the high pile-up environment has a significant impact on the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$, leading to an intense increase of the DY background in the same-flavour final state. This analysis includes therefore only the different-flavour final state: $h/H \to WW^{(*)} \to e\nu\,\mu\nu$.

Many background processes produce final states that look similar to the Higgs boson in the $h/H \to WW^{(*)} \to e\nu\,\mu\nu$ decay channel. The main background contributions come from the irreducible non-resonant WW background (see figure 5.9), from $t\bar{t}$ production (see figure 5.10(a)) and single top-quark production (see figure 5.10(b) – 5.10(d)). The WW events exhibit exactly the same final state as the signal if both W bosons decay leptonically. Another important background is W+jets production (see figure 5.11(a)) in which one of the jets is misidentified as a lepton. Other backgrounds are due to the production of Z+jets (see figure 5.11(b)) and non-WW diboson production (see figure 5.12), i.e. $WZ/ZZ/W\gamma^{(*)}$. The number of expected events of the diboson background, the $Z/\gamma^* +$ jets background and the top-quark backgrounds are estimated by using the theoretical cross sections, detector acceptances from simulated events, and data-driven correction factors for identification and reconstruction efficiencies.

In order to enhance signal events relative to background events, the following preselection requirements for $h/H \to WW^{(*)} \to e\nu\,\mu\nu$ candidates are imposed. These requirements are the same as for the SM $h_{\mathrm{SM}} \to WW^{(*)}$ analysis [139]. Exactly two charged leptons of different flavour (electron or muon) and opposite charge are required. The lepton with the highest p_{T} is called the leading lepton ℓ_1 and must have $p_{\mathrm{T}} > 25$ GeV, while the second lepton ℓ_2 fulfils $p_{\mathrm{T}} > 15$ GeV. The invariant mass of the two leptons $m(\ell_1\ell_2)$ is required to be larger than 10 GeV. Events with $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ smaller than 25 GeV are removed. This provides a strong suppression against multijet production via QCD processes and against the Z+jets background which originates predominantly from the $Z \to \tau^+\tau^-$ decay channel. Figure 5.13 shows the number of events for each background process and the SM Higgs-boson distribution as a function of jet multiplicity after the preselection cuts are applied.

After the preceding cuts the events are divided into two analysis channels: the 0-jet channel, in which exactly zero jets are required, and the 2-jet channel, which features exactly two jets. In the 0-jet channel, additional requirements

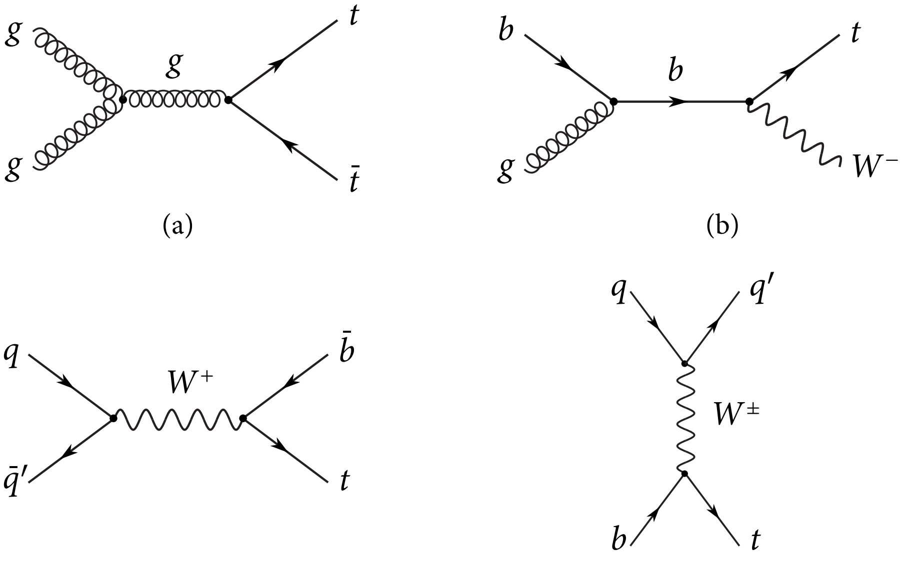


(a)

(b)

(c)

Figure 5.9: Examples of Feynman diagrams for WW production: tree-level Feynman diagram in $q\bar{q}$ annihilation (a) and loop-induced gluon-fusion Feynman diagrams (b)–(c).



(a)

(b)

(c)

(d)

Figure 5.10: Examples of tree-level Feynman diagrams for $t\bar{t}$ production (a) and single top-quark production in association with a W boson (b), single top-quark production in the s-channel (c), and single top-quark production in the t-channel (d).

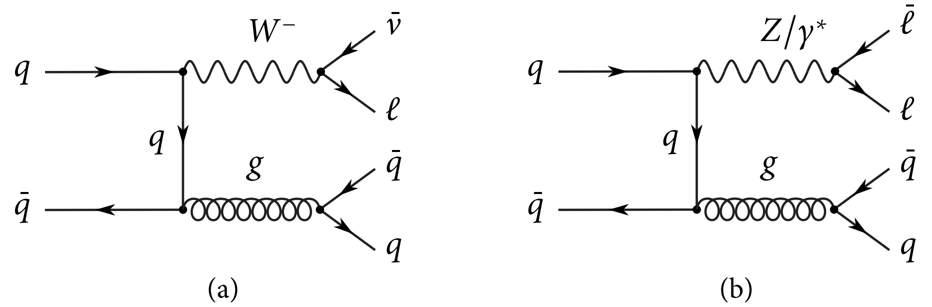


Figure 5.11: Examples of tree-level Feynman diagrams for W-boson production (a) and Z-boson production (b) in association with quarks.

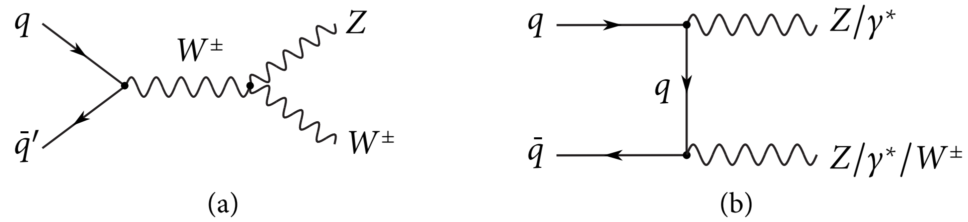


Figure 5.12: Examples of tree-level Feynman diagrams for non-WW diboson production.

are applied on the absolute value of the difference in the azimuth angles ϕ of the two leptons, $|\Delta\phi(\ell_1, \ell_2)| < 2.4$, and on the invariant mass of the dilepton system, $m(\ell_1\ell_2) < 75$ GeV. Both cuts are illustrated in figure 5.14. These two cuts remove background in regions where only a small signal contribution is expected.

Processes involving top-quarks form a substantial background and have the same kinematic signature concerning the leptons and $E_{\mathrm{T,rel}}^{\mathrm{miss}}$. Therefore, events containing a b-tagged jet are rejected in the 2-jet channel. The two selected jets are required to lie in opposite pseudorapidity hemispheres, $\eta(j_1) \times \eta(j_2) < 0$, and the two selected leptons have to have $m(\ell_1\ell_2) < 80$ GeV. An additional cut is placed on the transverse mass of the lepton-$E_{\mathrm{T}}^{\mathrm{miss}}$ system m_{T} which is defined as [164]

$$m_{\mathrm{T}} = \sqrt{(E_{\mathrm{T}}(\ell_1\ell_2) + E_{\mathrm{T}}^{\mathrm{miss}})^2 - (\vec{p}_{\mathrm{T}}(\ell_1\ell_2) + \vec{E}_{\mathrm{T}}^{\mathrm{miss}})^2} \tag{5.6}$$

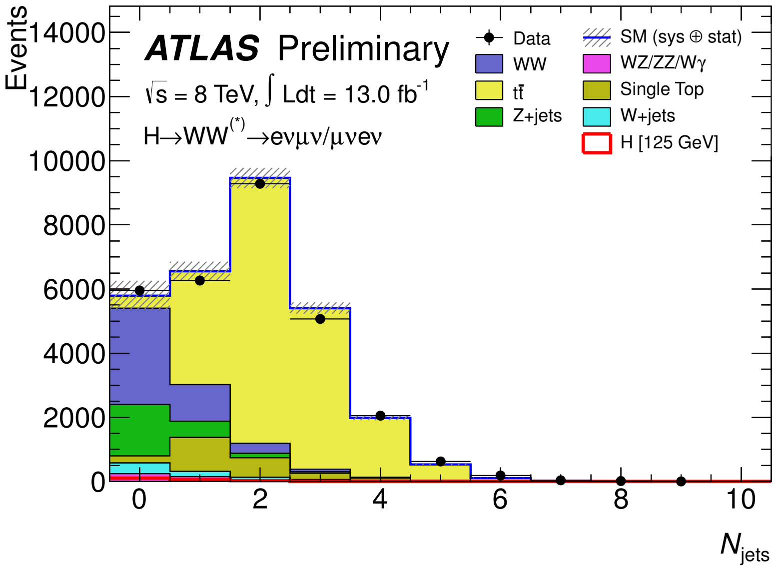


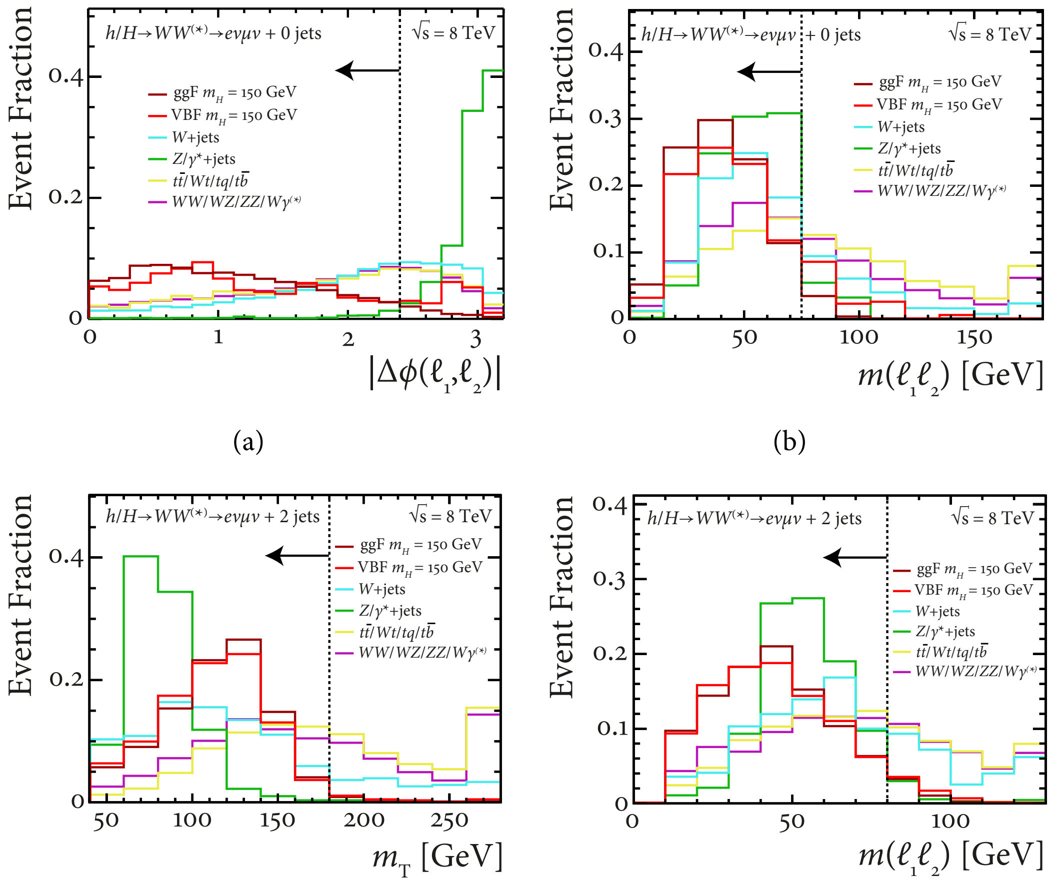
Figure 5.13: Jet multiplicity distribution after the preselection cuts are applied [139].

with

$$E_{\mathrm{T}}(\ell_1\ell_2) = \sqrt{\vec{p}_{\mathrm{T}}^{\,2}(\ell_1\ell_2) + m^2(\ell_1\ell_2)}\,. \tag{5.7}$$

It is required that $m_{\mathrm{T}} < 180$ GeV. The cuts on $m(\ell_1\ell_2)$ and m_{T} are illustrated in figure 5.14.

The W+jets background consist of events in which W bosons are produced in association with jets, and one jet is misidentified as a lepton. The W+jets background is estimated using a data-driven technique [139, 165] because the rate at which jets are misidentified as leptons is not accurately described by the detector simulation. A control sample is defined by one lepton that satisfies all identification requirements as defined in the main analysis and one so-called anti-identified lepton that fails these criteria, but passes a looser selection. The anti-identified electrons are required to have at least one hit in the pixel layer and at least three hits in the SCT. The calorimeter based isolation must satisfy $\sum(p_{\mathrm{T}}^{\mathrm{calo}})/p_{\mathrm{T}} < 0.3$, while for the track based isolation $\sum(p_{\mathrm{T}}^{\mathrm{track}})/p_{\mathrm{T}} < 0.16$ is required. No d_0-significance cut is required for anti-identified muons. Particles passing this alternative lepton definition are referred to as the denominator objects and are treated as identified leptons. Except for the lepton identifica-



(a) (b)

(c) (d)

Figure 5.14: Illustration of cuts applied in addition to the event preselection. Normalised distributions of $|\Delta\phi(\ell_1, \ell_2)|$ (a) without cuts on $|\Delta\phi(\ell_1, \ell_2)|$ and $m(\ell_1\ell_2)$, and $m(\ell_1\ell_2)$ (b) without a cut on $m(\ell_1\ell_2)$, both in the 0-jet channel. Normalised distributions of m_{T} (c) without cuts on m_{T} and $m(\ell_1\ell_2)$, and $m(\ell_1\ell_2)$ (d) without cut on $m(\ell_1\ell_2)$, both in the 2-jet channel. The vertical dashed line illustrates the additional cuts which have been applied in this analysis. The red shapes represent the signal for m_H = 150 GeV. The last bin of the histogram represents an overflow bin.

tion the events in the W+jets-enriched control sample pass all other selections as required for the signal samples. The number of these events is denoted as $N_{\text{Lepton+Denominator}}$. To obtain an estimate of the number of W+jets in the signal sample, the efficiency $\varepsilon_{\text{fake}}$ to misidentify an anti-identified lepton as a lepton is given by

$$\varepsilon_{\text{fake}} = \frac{N_{\text{Lepton}}}{N_{\text{Denominator}}} , \tag{5.8}$$

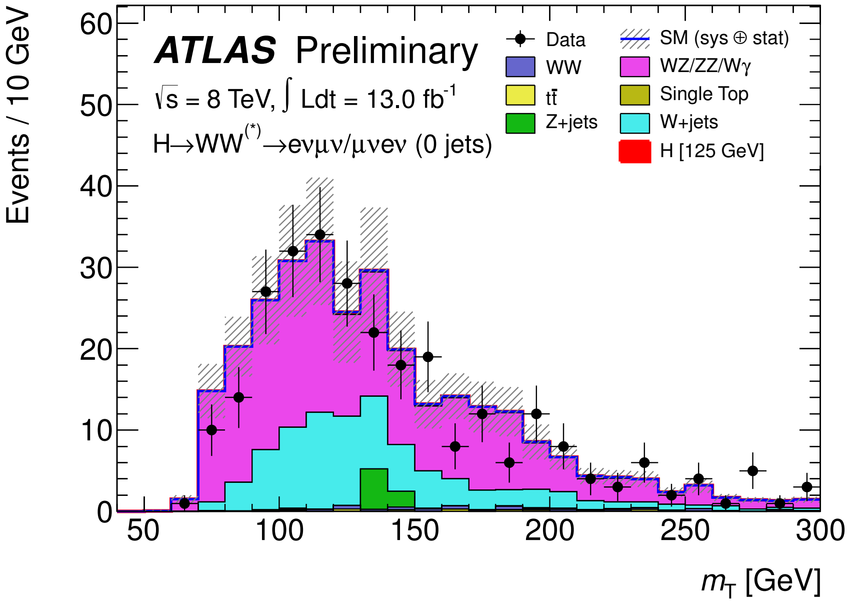
where N_{Lepton} and $N_{\text{Denominator}}$ are the number of events of the fully identified leptons and the number of events of the denominator objects, respectively. This so-called fake factor is defined for electrons and muons, and is determined as a function of the p_T of the anti-identified lepton by using a dijet data sample.

The number of events in the W+jets-enriched control sample is scaled by the fake factor to calculate the W+jets background:

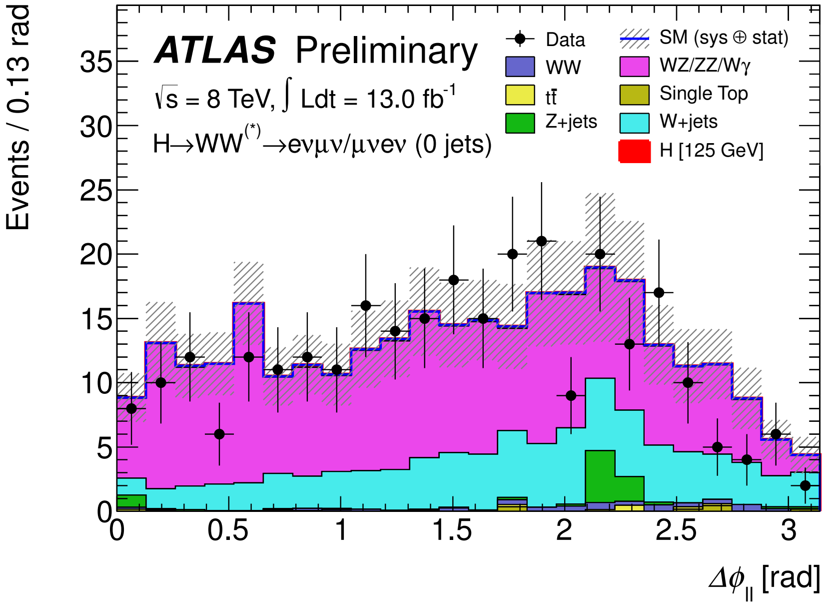
$$N_{W\text{+jets}} = \varepsilon_{\text{fake}} \times N_{\text{Lepton+Denominator}} . \tag{5.9}$$

Residual contributions of leptons originating from W-boson decays and Z-boson decays are subtracted. W+jets, $W\gamma^{(*)}$, $WZ^{(*)}$ and $Z^{(*)}Z^{(*)}$ production are the background processes that produce the majority of same-charge dilepton events. Therefore, the normalisation and kinematic distribution of these backgrounds can be validated in the same-charge validation region as shown in figure 5.15.

Table 5.3 shows the event yield of background and signal processes in the 0-jet channel and in the 2-jet channel after all selection cuts. The signal is split into two rows to distinguish between the light Higgs boson at $m_{\text{h}} = 125$ GeV and the heavy Higgs boson at $m_{\text{H}} = 150$ GeV in the 2HDM. The $t\bar{t}$ control region (top CR) is obtained by requiring at least one b-tagged jet, using an operating point leading to a b-jet identification efficiency of 85%. The diboson control region (WW CR) is defined by events passing the full event selection except the cut on $m(\ell_1\ell_2)$. Rather than $m(\ell_1\ell_2) < 75$ GeV it is required that $m(\ell_1\ell_2) > 80$ GeV. To reduce the Z/γ^* + jets background in this control region, one additionally requires that $p_{\text{T}}(\ell_1\ell_2) > 30$ GeV. In the subsequent statistical analysis, the *a-priori* estimates given in table 5.3 are re-evaluated in the maximum-likelihood fit to the NN-discriminants and the event yield in the top-quark control region.



(a)



(b)

Figure 5.15: m_{T} distribution (a) and $\Delta\phi(\ell_1, \ell_2)$ distribution (b) in the same-charge validation region [139].

Table 5.3: The expected number of signal and background events for 13 fb^{-1} of integrated luminosity in the 0-jet channel and in the 2-jet channel after all selection cuts. The event yield of a potential signal is illustrated for a light Higgs boson with m_h = 125 GeV and a heavy Higgs boson with m_H = 150 GeV in a type-I 2HDM with $\tan\beta = 3$ and $\alpha = \pi/2$. The W+jets background is entirely determined from data [139]. The uncertainties quoted in the table are determined as the sum in quadrature of two sources: the total rate uncertainty, see Table 7.1, and the corresponding cross-section uncertainties. Since the W+jets background is derived from collision data the uncertainties of the data-driven estimation are used. The lowermost row shows the observed numbers of events in the data.

Process	0 jet	2 jet	Top CR	WW CR
Signal (m_h = 125 GeV)	2.55 ± 0.50	5.52 ± 0.71	1.35 ± 0.19	0.76 ± 0.13
Signal (m_H = 150 GeV)	470 ± 140	76 ± 19	20.9 ± 5.7	16.1 ± 3.9
$WW/WZ/ZZ/W\gamma^{(*)}$	1140 ± 290	63 ± 18	22.1 ± 6.2	1170 ± 310
Z/γ^* + jets	41 ± 15	194 ± 72	84 ± 31	15.7 ± 6.4
W+jets	135 ± 58	23.4 ± 9.7	18.3 ± 7.6	78 ± 32
$t\bar{t}/tW/tb/tqb$	175 ± 49	168 ± 77	1760 ± 440	313 ± 97
Total background	1490 ± 420	450 ± 180	1890 ± 480	1580 ± 450
$\mathrm{S}_{h+H}/\mathrm{B}$	0.31	0.18	–	–
Observed	1815	483	1986	1725

The kinematic distributions of few basic variables are shown in the top CR and the WW CR in figure 5.16 to verify the correct modelling, i.e. the agreement between data and MC. The MC prediction is normalised to the number of events observed in data. Striped error bands indicate the total statistical and systematic uncertainty (see chapter 7). For each distribution the χ^2 probability is shown. To determine this value the χ^2 statistic has to be calculated by

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}, \tag{5.10}$$

where n is the number of bins, O_i is the observed number of events in the ith bin, and E_i is the expected number of events of all processes in the ith bin. The χ^2 probability is the probability of getting a χ^2 value equal or higher than the observed one out of a normalised χ^2 distribution with $n-1$ degrees of freedom. The χ^2 distribution with $n-1$ degrees of freedom is defined as the sum of $n-1$ standard normal distributions. The χ^2 probability is only valid for large values of O_i and E_i (typically values larger than 10 are sufficient), and depends strongly on the choice of binning. In addition, the significance of deviations between the model prediction and the observed data is computed bin-by-bin, following [166], and is shown in the lower histograms. A good modelling is found for the leptons and jets as well as for the $E_{\mathrm{T,rel}}^{\mathrm{miss}}$.

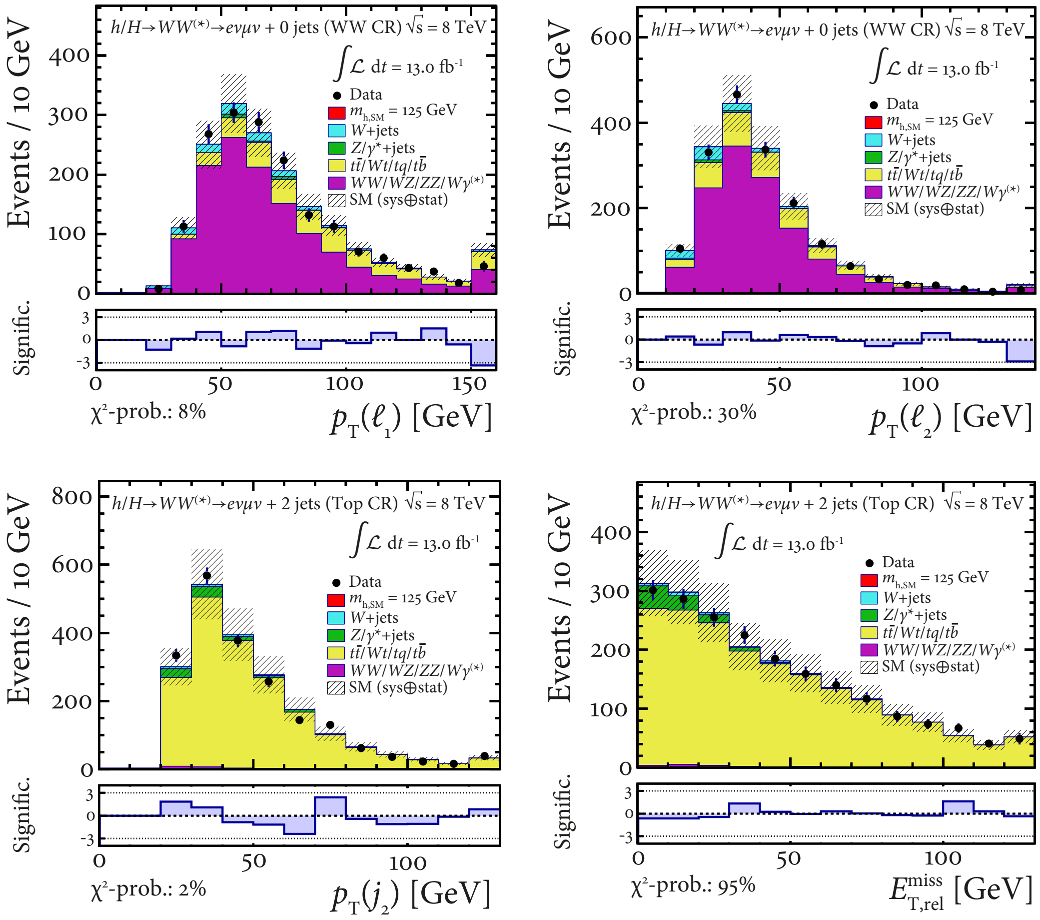


Figure 5.16: Distributions of some basic variables in the diboson control region (top) and in the $t\bar{t}$ control region (bottom) compared to a model based on simulated events that are normalised to data. Striped error bands indicate the total statistical and systematic uncertainty. The significance of deviations between the model prediction and the observed data is shown in the lower histograms, following [166]. The last bin of the histogram represents an overflow bin.

Figure 5.17 shows a Higgs-candidate event in the 0-jet channel containing one identified electron and one identified muon with the following event information:

Run	204026
Event	33133446
Luminosity Block	236
Date	2012-05-28 07:23:47 CEST
m_T	94.4 GeV
$\Delta\phi(\ell_1, \ell_2)$	0.295
$p_\mathrm{T}(e)$	32.8 GeV
$p_\mathrm{T}(\mu)$	29.0 GeV
$p_\mathrm{T}(\ell_1\ell_2)$	61.1 GeV
$E_{\mathrm{T,rel}}^{\mathrm{miss}}$	35.0 GeV
$m(\ell_1\ell_2)$	15.1 GeV

While the muon (red) has a clear signature in the ID and the MS, the reconstructed electron (green) reveals a good track in the ID and a clear signature in the electromagnetic calorimeter. The event shows two characteristic features of a good Higgs candidate: the relatively small angle $\Delta\phi(\ell_1, \ell_2)$ between these two leptons combined with the high missing transverse momentum (magenta) pointing in the opposite direction indicates this event as a good Higgs candidate.

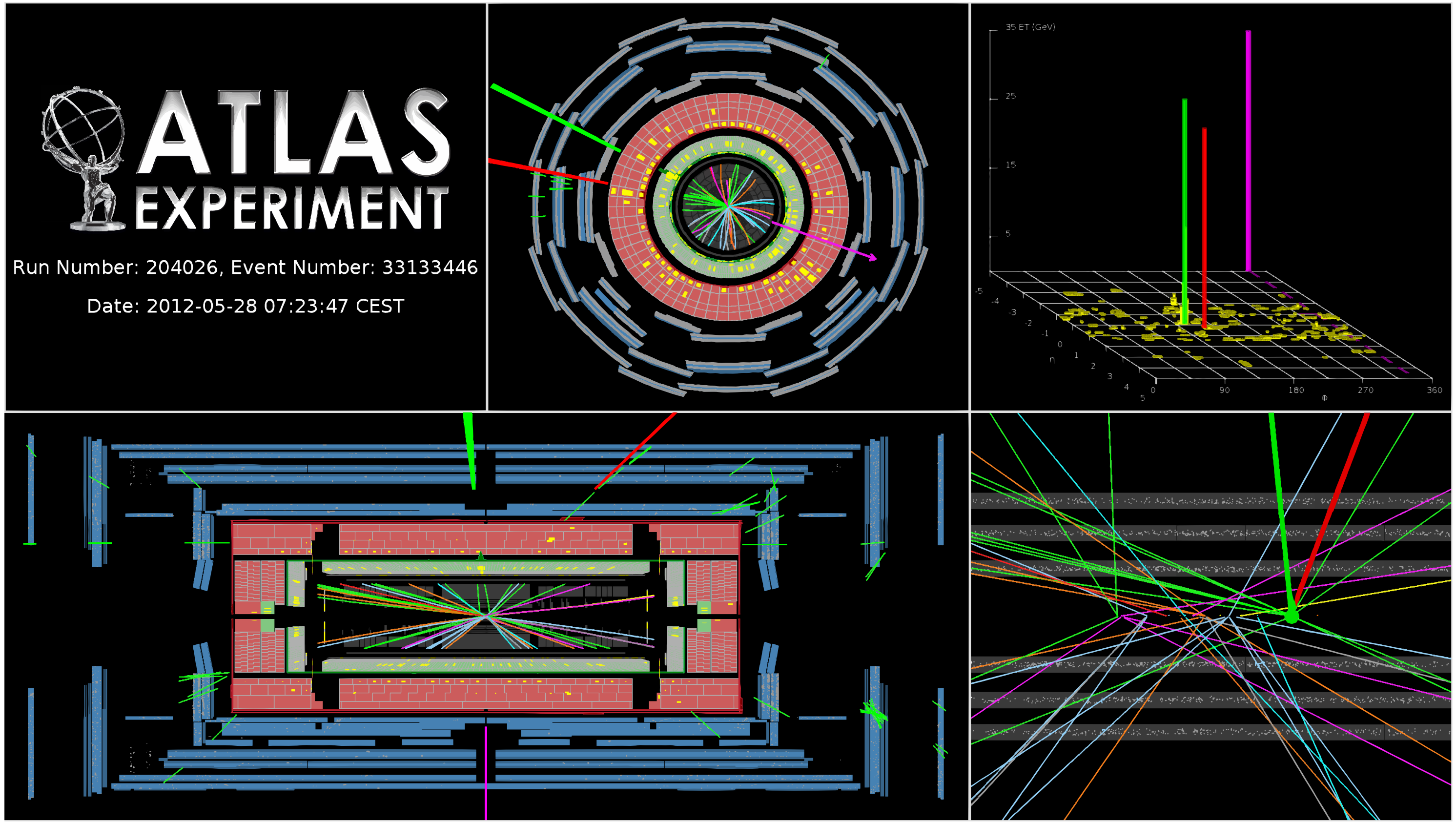


Figure 5.17: Atlantis [167] event display showing a Higgs candidate event. The electron and the muon are represented by a green and a red line, respectively, and $E_{\mathrm{T}}^{\mathrm{miss}}$ is shown by a magenta line starting at the hard subprocess, which is depicted as a green blob. Despite the fact that the electron is stopped in the electromagnetic calorimeter, the green line spread farther outside of the electromagnetic calorimeter to better illustrate the trace of the electron. In the plot on the top right, the dashed magenta line depicted the ϕ value of $E_{\mathrm{T}}^{\mathrm{miss}}$ to illustrate that $E_{\mathrm{T}}^{\mathrm{miss}}$ is independent of the η coordinate.

Chapter 6

Signal and Background Discrimination

In the previous chapter, the identification and reconstruction of electrons, muons, jets, and $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ was explained. The purity, which is defined as the fraction of signal events over background events, shown in table 5.3 for $\tan\beta = 3$ and $\alpha = \pi/2$ is quite good and actually one of the best of all parameter regions. With such a good purity little effort is needed to find the heavy Higgs boson of the 2HDMs at around $m_H = 150$ GeV if there is one, or exclude it if there is none. But there are large regions of the parameter space with a low purity because of the decreasing values of $\sigma \times \mathcal{B}$ for high values of $|\cos(\alpha)|$ as shown in figure 2.6. The described event selection is optimised to keep as much signal events as possible as opposed to a cut-based scenario, where one would use a few variables with some discrimination power and select small interesting signal regions from a possibly large background. By following this path of getting as much information as possible out of the selected data set, multivariate analysis (MVA) techniques are used to combine the information from several variables into one multivariate discriminant. This method is described in this chapter as well as characteristics of the Higgs-boson decay in the WW decay channels.

6.1 Event Reconstruction

The $h/H \to WW^{(*)} \to e\nu\mu\nu$ + 0-jet channel features a signature of two isolated charged leptons, which exhibit large transverse momenta, and a large

$E_{\mathrm{T}}^{\mathrm{miss}}$. The asterisk means that the particular particle which carries the asterisk is virtual, i.e. for Higgs-boson masses lower than 160 GeV one of the W bosons is on-shell (real), while the other one is off-shell (virtual). To be off-shell implies that the particle in question does not obey the energy-momentum relation (in natural units):

$$E^2 = p^2 + m^2 , \tag{6.1}$$

where E is the energy of the particle, and p and m are its momentum and its rest mass, respectively. For higher Higgs-boson masses both W bosons are produced on-shell.

Characteristic topologies of the final-state particles are determined by the properties of their parent particles, i.e. the W bosons. The kinematic distributions of the W bosons, in turn, are determined by the properties of the Higgs boson. In its rest frame, the spin-0 Higgs boson emits two spin-1 W bosons back-to-back due to momentum conservation. The spin directions of the W bosons need to be opposite to satisfy angular momentum conservation. Each W boson decays further into one charged lepton and one neutrino, whereby neutrinos have to have negative, i.e. left-handed, helicity states and antineutrinos have to have positive, i.e. right-handed, helicity states due to the $V - A$ structure of the weak interaction. As a result of the angular momentum conservation, the charged leptons are emitted in the same direction in the inertial reference frame of the Higgs boson as shown in figure 6.1. This feature is preserved in the reference frame of the laboratory as shown in figure 6.2(a). The $E_{\mathrm{T}}^{\mathrm{miss}}$ is large due to the small angle between the two neutrinos (see figure 6.2(b)). Three main observables are the dilepton invariant mass $m(\ell_1\ell_2)$, the transverse mass of the Higgs boson m_{T}, and the dilepton transverse momentum $p_{\mathrm{T}}(\ell_1\ell_2)$, where the last two are highly correlated. $m(\ell_1\ell_2)$ is highly correlated to $\cos\theta(\ell_1, \ell_2)$ (see figure 6.2(a)), since

$$m(\ell_1\ell_2) = 2 \cdot p_{\mathrm{T}}(\ell_1) \cdot p_{\mathrm{T}}(\ell_2) \cdot \{1 - \cos\theta(\ell_1, \ell_2)\} \tag{6.2}$$

if the leptons are considered to be massless.

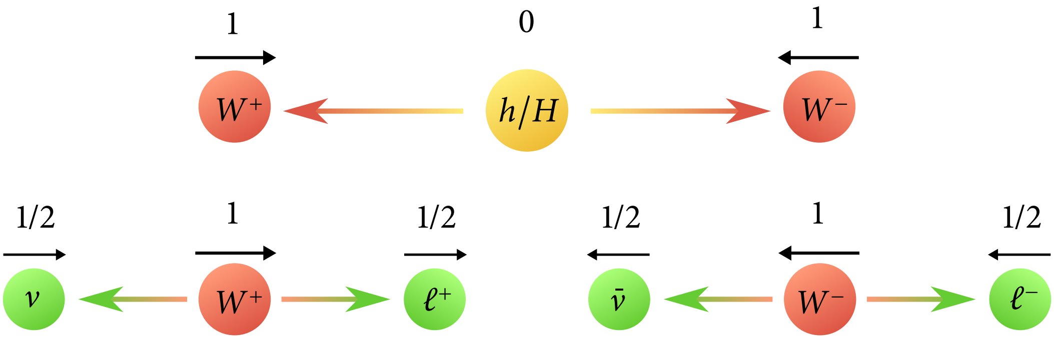
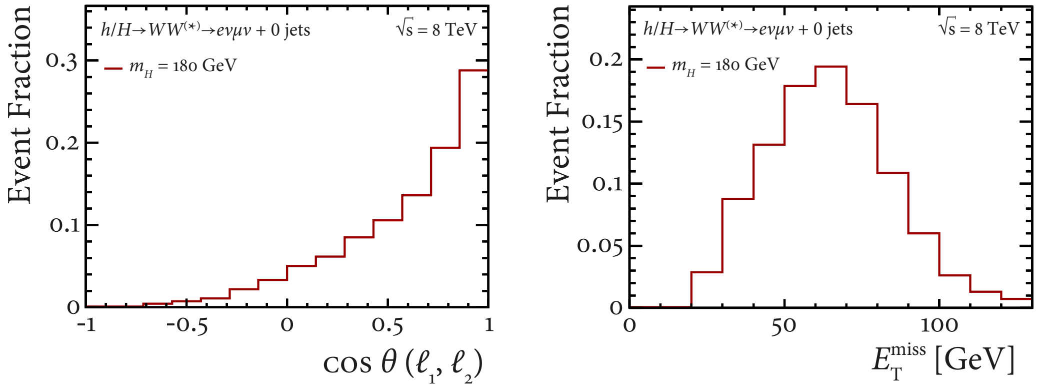
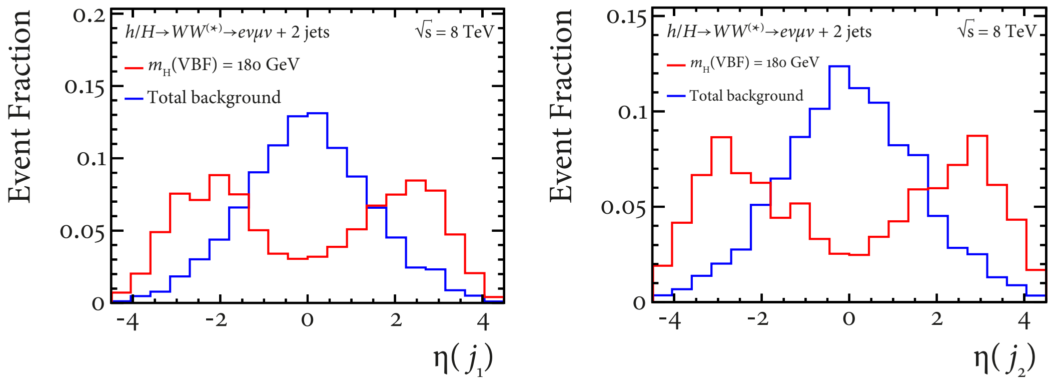


Figure 6.1: Visualisation of the decay of a Higgs boson into two W bosons. The subsequent decay of each W boson into one charged lepton and one neutrino is determined by spin conservation as well as momentum conservation. Black arrows and coloured arrows represent the spin and the momentum of a particular particle, respectively.



(a) (b)

Figure 6.2: Normalised distributions of $\cos\theta(\ell_1, \ell_2)$ (a) and $E_{\mathrm{T}}^{\mathrm{miss}}$ (b) in the 0-jet channel. The distribution is shown for a Higgs boson with a mass of 180 GeV.



(a) (b)

Figure 6.3: Normalised distributions of the pseudorapidity of both jets (sorted by p_T) in the 2-jet channel. The two jets point predominantly in the forward and backward direction.

The $h/H \to WW^{(*)} \to e\nu\mu\nu$ + 2-jets channel produces a signature with two isolated charged leptons, two energetic hadronic jets, and a large $E_\mathrm{T}^\mathrm{miss}$. The most distinctive feature of the signal topology in this channel originates from the VBF process, where the two jets point predominantly in the forward and backward directions with respect to the proton beamline (see figure 6.3). This property combined with the absence of QCD colour exchange as an additional feature of VBF leads to a rapidity gap. The two main observables are the dijet invariant mass $m(j_1 j_2)$ and the transverse mass of the Higgs boson m_T.

A complete reconstruction of all final-state particles is not possible because the WW channel involves two neutrinos with unknown momenta. Since the Higgs-boson mass cannot be reconstructed directly the fully determined kinematic variable m_T is used as an approximation. Figure 6.4 shows the normalised distribution of m_T. A Gaussian distribution is fitted to the data. The standard deviation of the fit is 18 GeV and is used as an estimation of the mass resolution.

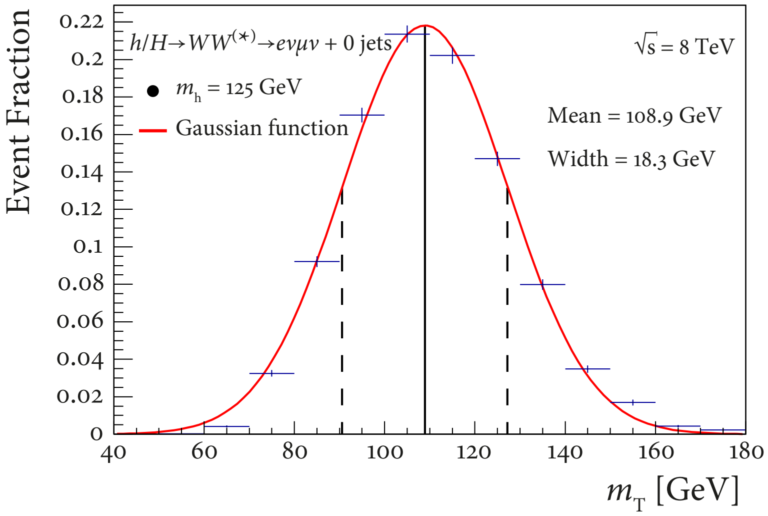


Figure 6.4: Gaussian curve fitted to the reconstructed transverse mass. The standard deviation of the fit is 18 GeV.

6.2 Neural Networks

To separate 2HDM Higgs-boson signal candidate events from background in the WW channel, several kinematic variables are combined to one discriminant in each of the two analysis channels by employing NNs. The `NeuroBayes` [168] tool is used for preprocessing the input variables and for the training of the NNs. A large number of input variables is studied, but only the highest-ranking variables are chosen for the training of the NNs. The ranking of the variables is automatically determined as part of the preprocessing step and is independent of the training procedure [169].

As a result of this optimisation procedure in the 0-jet channel, six kinematic variables are identified that serve as inputs to the NNs. Several NNs are trained since the kinematic distribution of the signal varies at different mass points of the Higgs boson. Three NNs are trained at three different Higgs-boson mass points m_H = 150 GeV, 180 GeV, and 240 GeV to cover the entire mass range between 135 and 300 GeV in which a good sensitivity to the signal can be achieved. For each NN, the same six input variables were used. The NN trained at m_H = 150 GeV is evaluated at the mass points from 135 GeV to 160 GeV, the NN trained at m_H = 180 GeV is applied in the range from 165 GeV to

200 GeV, and the NN trained at $m_H = 240$ GeV is used at the mass points from 220 GeV to 300 GeV. The evaluation of the background samples and the observed data remains the same for each NN, and only the distributions of the 2HDM Higgs-boson samples vary from mass point to mass point. Dedicated studies showed that the NNs at these three mass points achieve the same sensitivity as a large number of NNs that are trained with a spacing of 5 GeV in m_H. The physical reason for this is the relatively low mass resolution in the $h/H \to WW^{(*)} \to \ell^+\nu\ell^-\bar{\nu}$ channel of 18 GeV (see figure 6.4). It was further studied that the addition of more variables to the NNs would lead only to an improvement well below 1% in the expected CLs values across the $\cos\alpha$–m_H plane (see chapter 8). In the same way, three NNs are trained in the 2-jet channel sharing nine input variables.

The selected variables are listed and explained in table 6.1. The three most important input variables for the NN in the 0-jet channel are $m(\ell_1\ell_2)$, m_T, and $p_\mathrm{T}(\ell_1\ell_2)$, while the three most important input variables for the NN in the 2-jet channel are $m(j_1j_2)$, $m(\ell_1\ell_2)$, and m_T.

In the preprocessing step, the variables are first transformed to be distributed between 0 and 1 to reduce the influence of extreme outliers: The correlation matrix of all preprocessed input variables is calculated, including the correlation of all variables to the target [169]. The target function assumes the value 1 for signal events and 0 for background events. The variables are decorrelated and the total correlation of the entire set of variables to the target is computed. To find a ranking of the variables, one variable after the other is omitted to determine the loss of the total correlation to the target caused by its removal. The variable with the smallest loss of correlation is discarded leading to a set of $(n-1)$ variables. The same procedure is repeated with the reduced set of variables to find the least important one of the $(n-1)$ remaining variables. At the end of this iterative process, a list of variables ordered by importance is obtained. Considering the number of simulated events used to compute the ranking, one can compute the significance of the information loss upon removal of a certain variable.

The choice of a certain set of variables is a compromise between discrimination power that increases with the number of variables and keeping the number of variables manageable. The resulting choice of variables are shown in the ranked list of the selected variables for $m_H = 150$ GeV in table 6.2, for $m_H = 180$ GeV in table 6.3, and for $m_H = 240$ GeV in

Table 6.1: Input variables used for the NNs in the 0-jet and 2-jet channels. The definitions of the variables use the terms *leading lepton* and *leading jet*, defined as the lepton/jet with the highest p_{T}.

Variables used in the 0-jet channel and the 2-jet channel	
$\lvert\eta(\ell_1)\rvert$	The absolute value of the pseudorapidity of the leading lepton.
m_{T}	The transverse mass of the lepton-$E_{\mathrm{T}}^{\mathrm{miss}}$ system, as defined in Equation (5.6).
$m(\ell_1\ell_2)$	The invariant mass of the dilepton system.
Variables used in the 0-jet channel only	
$p_{\mathrm{T}}(\ell_1\ell_2)$	The transverse momentum of the dilepton system.
$E_{\mathrm{T,rel}}^{\mathrm{miss}}$	The projection of the calorimeter-based missing transverse momentum onto the direction perpendicular of the nearest object as defined in (5.5).
$\lvert\Delta y(\ell_1,\ell_2)\rvert$	The absolute value of the rapidity differences of the two charged leptons.
Variables used in the 2-jet channel only	
$p_{\mathrm{T}}(\ell_2)$	The transverse momentum of the second-leading lepton.
$p_{\mathrm{T}}(j_1)$	The transverse momentum of the leading jet.
$m(j_1)$	The mass of the leading jet.
$\cos\theta(\ell_1,\ell_2)$	The cosine of the angle between the two charged leptons.
$m(j_1 j_2)$	The invariant mass of the dijet system.
$p_{\mathrm{T}}^{\mathrm{tot}}$	The total transverse momentum, defined as the magnitude of the vector sum of the transverse momenta of the two jets, the two leptons, and the missing transverse momentum: $p_{\mathrm{T}}^{\mathrm{tot}} = \lvert\vec{p}_{\mathrm{T}}^{\,\mathrm{tot}}\rvert = \lvert\vec{p}_{\mathrm{T}}^{\,\ell_1} + \vec{p}_{\mathrm{T}}^{\,\ell_2} + \vec{p}_{\mathrm{T}}^{\,j_1} + \vec{p}_{\mathrm{T}}^{\,j_2} + \vec{E}_{\mathrm{T}}^{\mathrm{miss}}\rvert$.

Table 6.2: Variables used for the training of the NN at m_H = 150 GeV in the 0-jet channel (left) and in the 2-jet channel (right), ordered by their importance. The ordering is given by the loss in total correlation to the target. The total correlation to the target for the entire set of variables is 0.36 and 0.69 respectively.

Variable	Correlation Loss
$m(\ell_1\ell_2)$	0.25
m_T	0.20
$\lvert\eta(\ell_1)\rvert$	0.11
$p_\mathrm{T}(\ell_1\ell_2)$	0.11
$E_\mathrm{T,rel}^\mathrm{miss}$	0.04
$\lvert\Delta y(\ell_1,\ell_2)\rvert$	0.03

Variable	Correlation Loss
$m(j_1j_2)$	0.62
$m(\ell_1\ell_2)$	0.19
m_T	0.18
$p_\mathrm{T}^\mathrm{tot}$	0.12
$p_\mathrm{T}(\ell_2)$	0.12
$\lvert\eta(\ell_1)\rvert$	0.07
$m(j_1)$	0.04
$\cos\theta(\ell_1,\ell_2)$	0.03
$p_\mathrm{T}(j_1)$	0.02

Table 6.3: Variables used for the training of the NN at m_H = 180 GeV in the 0-jet channel (left) and in the 2-jet channel (right), ordered by their importance. The ordering is given by the loss in total correlation to the target. The total correlation to the target for the entire set of variables is 0.40 and 0.73 respectively.

Variable	Correlation Loss
$p_\mathrm{T}(\ell_1\ell_2)$	0.36
$\lvert\eta(\ell_1)\rvert$	0.11
$m(\ell_1\ell_2)$	0.09
m_T	0.07
$E_\mathrm{T,rel}^\mathrm{miss}$	0.03
$\lvert\Delta y(\ell_1,\ell_2)\rvert$	0.02

Variable	Correlation Loss
$m(j_1j_2)$	0.67
m_T	0.20
$p_\mathrm{T}(\ell_2)$	0.12
$m(\ell_1\ell_2)$	0.12
$p_\mathrm{T}^\mathrm{tot}$	0.11
$\lvert\eta(\ell_1)\rvert$	0.05
$\cos\theta(\ell_1,\ell_2)$	0.05
$p_\mathrm{T}(j_1)$	0.05
$m(j_1)$	0.01

Table 6.4: Variables used for the training of the NN at m_H = 240 GeV in the 0-jet channel (left) and in the 2-jet channel (right), ordered by their importance. The ordering is given by the loss in total correlation to the target. The total correlation to the target for the entire set of variables is 0.40 and 0.73 respectively.

Variable	Correlation Loss
$p_{\mathrm{T}}(\ell_1\ell_2)$	0.35
$m(\ell_1\ell_2)$	0.15
$\lvert\Delta y(\ell_1,\ell_2)\rvert$	0.12
$E_{\mathrm{T,rel}}^{\mathrm{miss}}$	0.08
$\lvert\eta(\ell_1)\rvert$	0.02
m_{T}	0.01

Variable	Correlation Loss
$m(j_1j_2)$	0.69
m_{T}	0.21
$p_{\mathrm{T}}^{\mathrm{tot}}$	0.11
$p_{\mathrm{T}}(\ell_2)$	0.08
$\lvert\eta(\ell_1)\rvert$	0.06
$\cos\theta(\ell_1,\ell_2)$	0.03
$m(j_1)$	0.01
$m(\ell_1\ell_2)$	0.01
$p_{\mathrm{T}}(j_1)$	0.01

table 6.4.

6.2.1 Control Regions

Once a set of variables has been chosen based on the criteria outlined above, the modelling of the input variables is checked in two control regions that are enriched with the dominating backgrounds: diboson production and $t\bar{t}$ production. The diboson control sample is used to check the modelling of the input variables that are fed into the NNs of the 0-jet channel, while the $t\bar{t}$-enriched control sample is used to check the modelling of the input variables that are fed into the NNs of the 2-jet channel. The input variables of the two analysis channels are shown in figure 6.5 in the diboson control region, and in figure 6.6 and 6.7 in the $t\bar{t}$ control region. Hatched error bands indicate the total statistical and systematic uncertainty (see chapter 7). Good modelling of the important diboson and $t\bar{t}$ backgrounds is found.

The significance of deviations between the model prediction and the observed data is shown in the lower histograms, following [166]. When com-

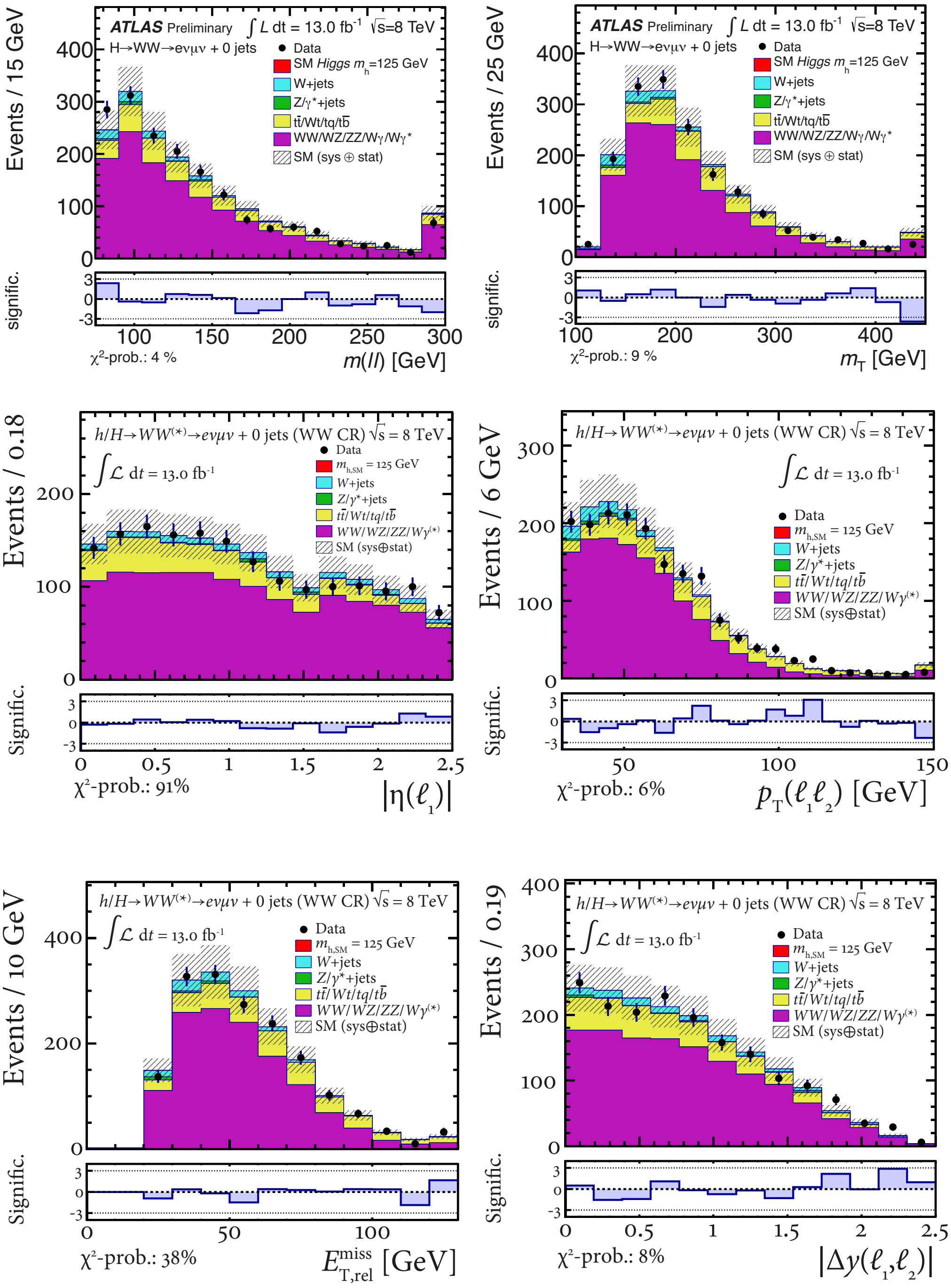


Figure 6.5: Distributions of all input variables of the NNs in the diboson control region. The distributions observed in collision data are compared to a model based on simulated events for which the contribution of each process is normalised to the event yields in table 6.5. The sufigures (top) with the label ATLAS Preliminary are taken from [164].

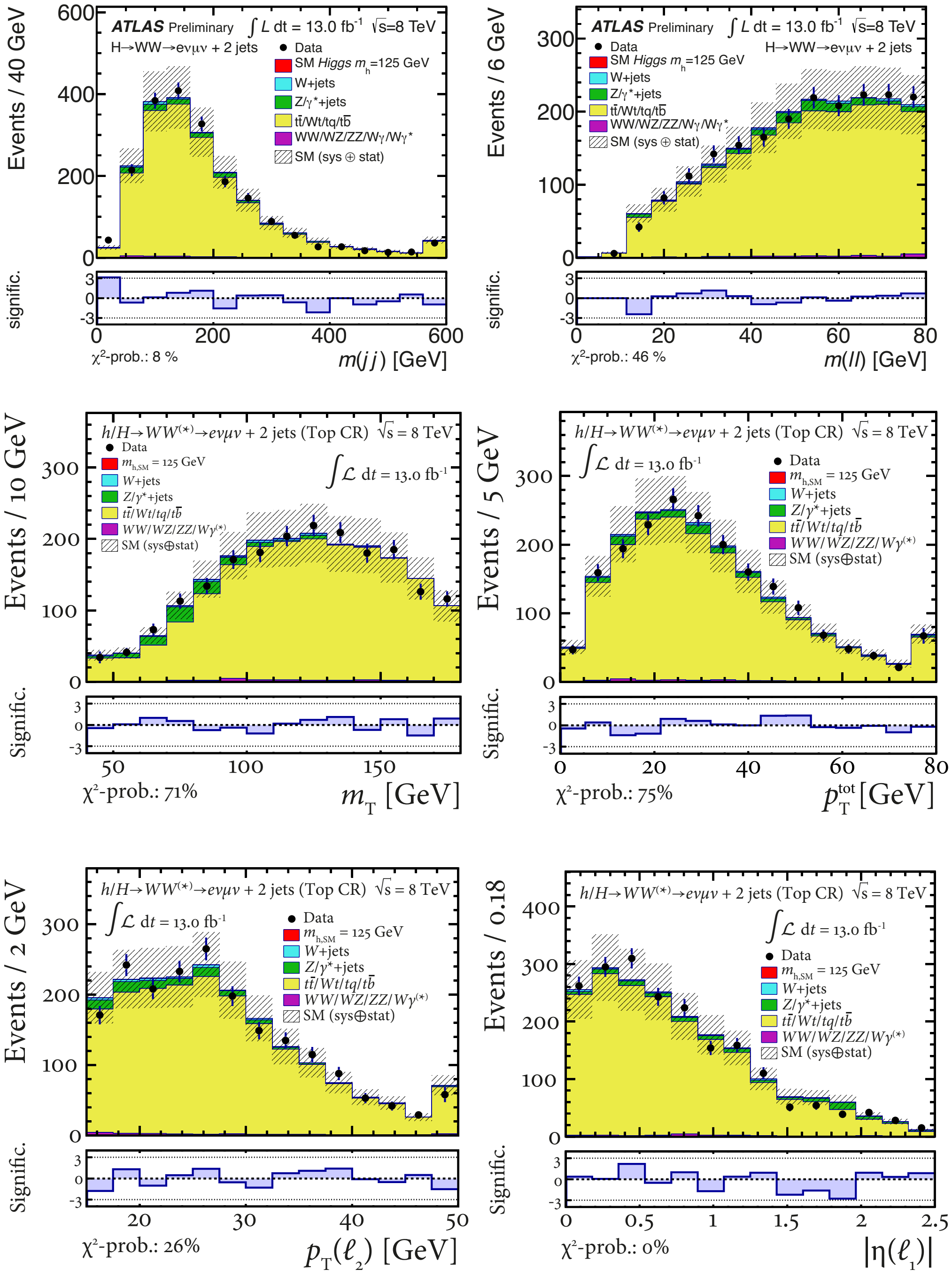


Figure 6.6: Distributions of the six highest-ranking input variables of the NNs in the $t\bar{t}$ control region. The distributions observed in collision data are compared to a model based on simulated events for which the contribution of each process is normalised to the event yields in table 6.5. The sufigures (top) with the label ATLAS Preliminary are taken from [164].

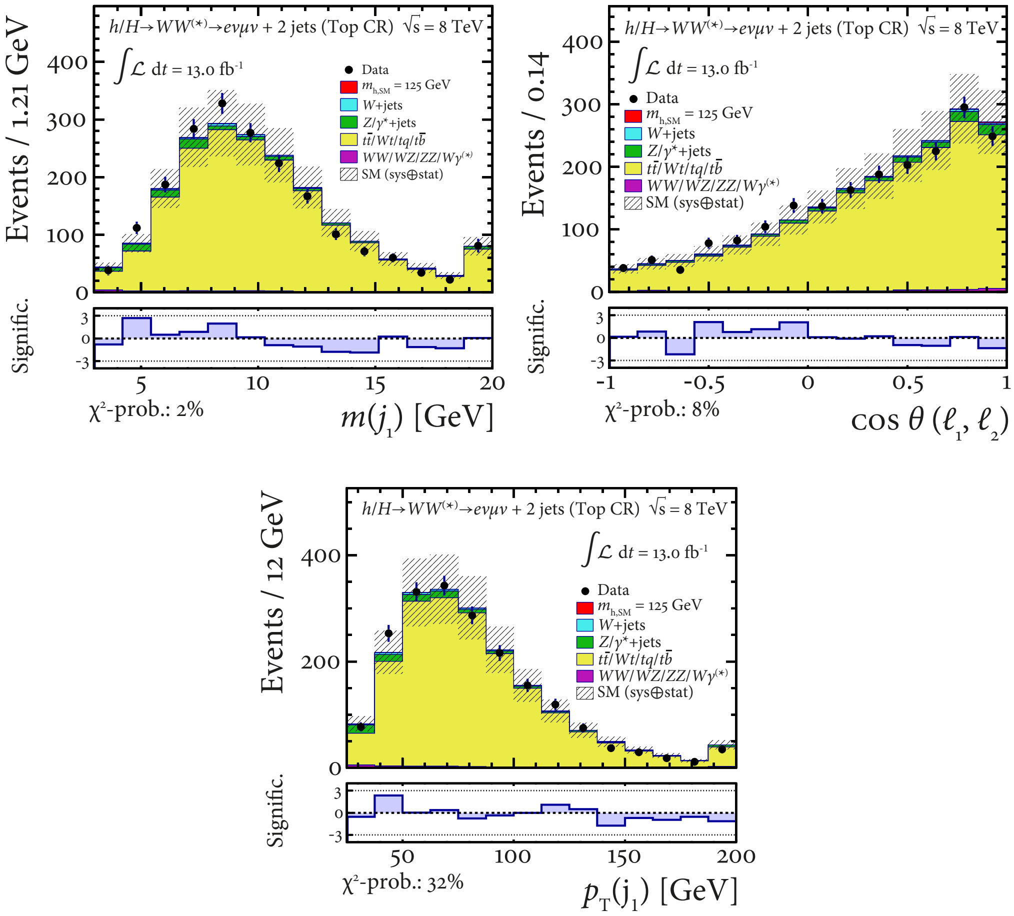


Figure 6.7: Distributions of the three lowest-ranking input variables, which are fed into the NNs of the 2-jet signal channel, in the $t\bar{t}$ control region. The distributions observed in collision data are compared to a model based on simulated events for which the contribution of each process is normalised to the event yields in table 6.5.

Table 6.5: The estimated number of signal and background events after the maximum-likelihood fit for 13 fb^{-1} of integrated luminosity in the 0-jet channel and in the 2-jet channel after all selection cuts. The W+jets background is entirely determined from data [139]. The uncertainties quoted in the table are determined as the sum in quadrature of two sources: the total rate uncertainty, see table 7.1, and the corresponding cross-section uncertainties. Since the W+jets background is derived from collision data the uncertainties of the data-driven estimation are used. The lowermost row shows the observed numbers of events in the data.

Process	0 jet	2 jet	Top CR	WW CR
Signal ($m_{h_{\mathrm{SM}}}$)	154 ± 30	20.6 ± 2.6	6.08 ± 0.86	1.21 ± 0.21
$WW/WZ/ZZ/W\gamma^{(*)}$	1290 ± 330	71 ± 20	24.9 ± 7.0	1320 ± 350
Z/γ^{*} + jets	41 ± 15	193 ± 72	84 ± 31	15.6 ± 6.4
W+jets	135 ± 58	23.4 ± 9.7	18.3 ± 7.6	78 ± 32
$t\bar{t}/tW/tb/tqb$	184 ± 51	176 ± 81	1850 ± 460	330 ± 100
Total background	1650 ± 450	460 ± 180	1980 ± 510	1740 ± 490
$\mathrm{S}_{h_{\mathrm{SM}}}/\mathrm{B}$	0.09	0.04	–	–
Observed	1815	483	1986	1725

puting the significance, only statistical uncertainties are taken into account. Beneath the significance histogram the χ^2 probability is shown. The χ^2 probability is explained in section 5.3.

In the subsequent statistical analysis, the a-priori estimates are re-evaluated in the maximum-likelihood fit to the NN-discriminants and the event yield in the $t\bar{t}$ control region. In this fit, the top-quark background is entirely free to float, while the diboson and Z/γ^{*}+ jets backgrounds are floating within the uncertainties of the a-priori estimates given in table 5.3. The MC prediction in figure 6.5 – 6.7 is normalised to the event yields in table 6.5. The last bin of the histograms represents an overflow bin. The normalisation factor β for each process obtained by the maximum-likelihood fit is shown in table 6.6. The fitted value of the SM Higgs-boson event yield is very well compatible with that measured in the $H \to WW^{*}$ SM analysis [139].

Table 6.6: Normalisation factor β for each process obtained by the maximum-likelihood fit to the NN discriminant optimised at m_H = 150 GeV. Only the statistical uncertainties on the normalisation factors are shown.

Process	$\beta \pm \Delta\beta$ (stat)
SM(ggF)	1.6 ± 0.66
SM(VBF)	0.9 ± 1.04
$Z/\gamma^* + \text{jets}$	1.0 ± 0.10
$t\bar{t}/tW/tb/tqb$	1.1 ± 0.03
$WW/WZ/ZZ/W\gamma^{(*)}$	1.1 ± 0.07

6.2.2 Training of Neural Networks

The analysis proceeds with the training of the NNs. A three-layer feed-forward architecture with six hidden nodes is used for all NNs (see figure 6.8). Bayesian regularisation techniques are applied for the training process to damp statistical fluctuations in the training sample and to avoid overtraining. The ratio of signal events to background events in the training is chosen to be 1:1, while the different background processes are weighted relative to each other according to the number of expected events.

The input layer contains one input node for each input variable plus one bias node, which is used to shift the weighted sum of the input variables to the sensitive region of the sigmoid function to avoid saturation effects. The output value o of each network is calculated by

$$o = S\left(\sum_m w_m \times S\left(w_{m,\text{Bias}} + \sum_l w_{ml}\tilde{i}_l\right)\right), \tag{6.3}$$

where $\tilde{i}_l$ is the value of the decorrelated input variable l after the preprocessing. w_{ml} is the weight that connects the node l of the input layer to the node m of the hidden layer, while the weight w_m connects the node m of the hidden node to the output node. $w_{m,\text{Bias}}$ is the weight that connects the bias node to the node m of the hidden layer. The indices l and m are running over all input nodes and all hidden nodes, respectively. $S(x)$ is the sigmoid function, which

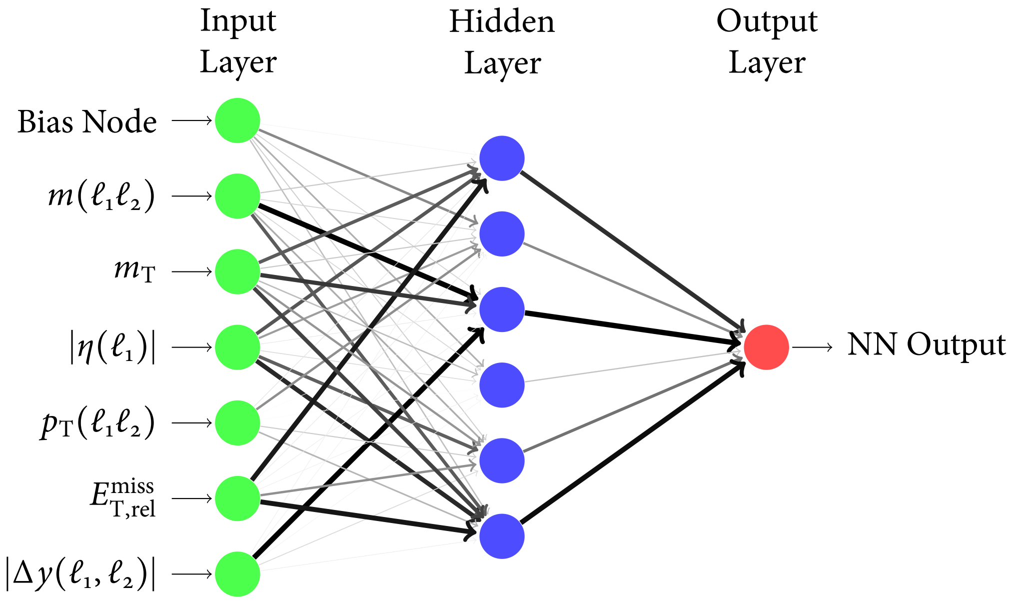


Figure 6.8: Overview of the architecture of a feed-forward NN with three layers and six hidden nodes optimised in the 0-jet channel at $m_H = 150$ GeV. The input variables are sorted by importance. Thickness as well as blackness of the arrows visualise the absolute values of the weights, which represent the single significance of each variable.

is defined by

$$S : (-\infty, +\infty) \to (-1, 1),\ x \mapsto \frac{2}{1 + \exp(-x)} - 1\,. \tag{6.4}$$

w_{ml} and w_{nm} are iteratively adjusted to minimise the deviation of the actual network output to the desired output, i.e. the specified target. In `Neurobayes`, the entropy loss function E is used as the numerically minimised function:

$$E = \sum_i \log\left(\frac{1}{2}\left(1 + T_i o_i\right) + \varepsilon\right)\,. \tag{6.5}$$

T_i and o_i represent the network target value and the output value of event i, respectively. To avoid numerical problems at the beginning of the training, ε is used as a small regularisation factor, which is reduced in each training

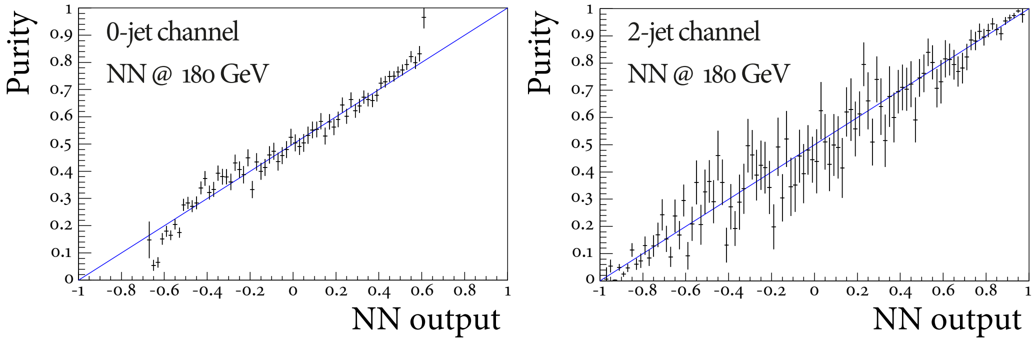


Figure 6.9: The expected linear relation (solid line) and the actual signal purity as a function of the output of the NN optimised at m_H = 180 GeV in the 0-jet channel (left) and in the 2-jet channel (right).

iteration. A completely wrong classification, i.e. $T_i = 1$ and $o_i = -1$ or vice versa, leads to an infinitely large value of the entropy loss function, $E \to \infty$, while a completely correct classification, $T_i = o_i = 1$ or $T_i = o_i = -1$, leads to $E = 0$. The maximum number of training iterations is set to one hundred and the momentum of the training, which adds a fraction of the previous weight update to the current one, is specified to be 0.5. The purity p should be a linear function of the NN output if the network is fully trained:

$$P = \frac{o_i + 1}{2} . \tag{6.6}$$

Figure 6.9 shows this relation for the NN optimised at m_H = 180 GeV in the 0-jet channel (left) and in the 2-jet channel (right).

A linear transformation is used to scale the NN output values from the interval [-1, 1] to [0, 1]. The resulting distributions of the 2HDM signal and the sum of all pre-fit background estimates normalised to unit area are shown in figure 6.10 for the NNs optimised at m_H = 150 GeV, 180 GeV, and 240 GeV. The light Higgs boson h in a 2HDM at a given mass has the same kinematic distribution as the SM Higgs boson at the same mass. Only the predicted rate is changed as explained in section 2.2.2. Therefore, the shape of the NN discriminant of the light Higgs boson h in a 2HDM is the same as the one of a pure SM Higgs boson. In the same way the shape of the NN discriminant of the heavy Higgs boson H in a 2HDM corresponds to a pure SM Higgs boson

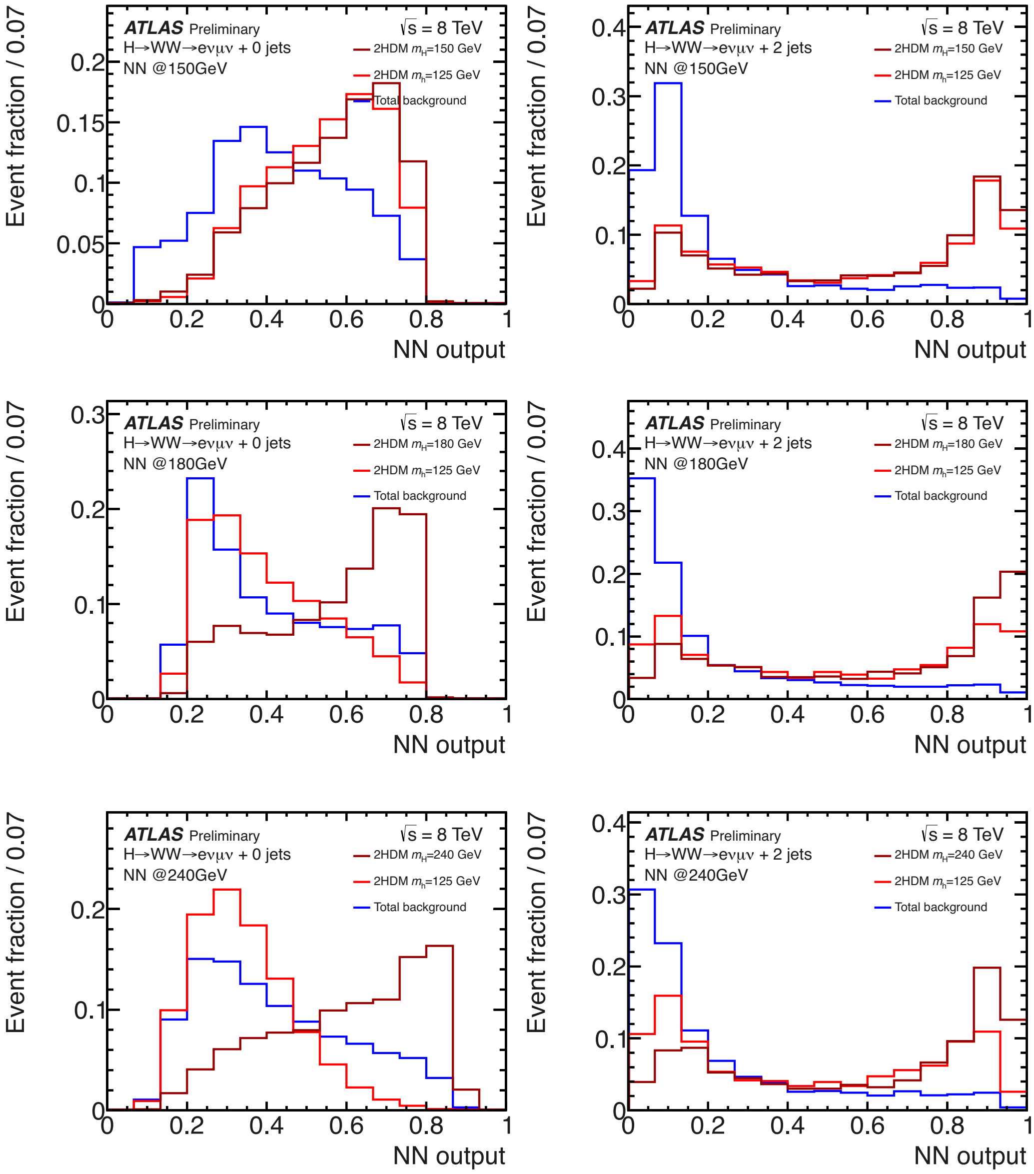


Figure 6.10: Normalised discriminant distributions obtained for the NNs optimised at m_H = 150 GeV, 180 GeV, and 240 GeV (from top to bottom) for the 2HDM signal and SM background processes, on the left in the 0-jet channel and on the right in the 2-jet channel [164]. The distributions are normalised to unit area.

at the respective mass.

After the training is completed, the modelling of the NN-discriminant distributions is checked in the $t\bar{t}$ control region and in the diboson control region. These distributions are shown in figure 6.11. It is important to note that the NNs have to be evaluated in the same phase-space region as used during the training. Since the diboson control region is defined by $m(\ell_1\ell_2) > 80$ GeV, the value of $m(\ell_1\ell_2)$ can not be used to evaluate the NNs in this region. That is why the input variable $m(\ell_1\ell_2)$ of each event is randomly chosen between 10 and 75 GeV to check the modelling of the NNs in the WW CR. The MC prediction in figure 6.11 is normalised to the event yields in table 6.5.

The agreement between the model and the observed data is very good, and the next step of the analysis is the application of the NNs to the events in the signal samples. The corresponding distributions of the NN discriminants are shown in figure 6.12. The distributions observed in collision data are compared to a model based on simulated events for which the contribution of each process is normalised to the event yields in table 6.5 including the SM Higgs boson. The fit is performed to the NN-discriminant distributions trained at $m_H = 150$ GeV (on the top of figure 6.12), simultaneously. The obtained fit values are used to normalise the other NN-discriminant distributions as well. The statistical analysis of the resulting discriminant distributions is discussed in chapter 8.

The distributions of the input variables are checked in the signal regions by comparing the data to the distributions predicted by the model of simulated events normalised to the event yields in table 6.5, as described above. The distributions of the all input variables in the 0-jet and 2-jet channels are shown in figure 6.13 – 6.15. A good agreement between data and the SM prediction is found.

Multivariate analysis techniques are in general prone to overtraining, which means learn the kinematic configuration by heart. Overtraining compromises the result and leads to a bias if the network is applied to analyse collision data. The overtraining in `NeuroBayes` is thoroughly checked during the training process. Only 80% of the MC events are used to train the NN by minimising the entropy loss function, while the other 20% of MC events are used as overtraining test sample. The values of the entropy loss function for the training and test samples are shown in figure 6.16. An increase of the value of the entropy loss function on the test sample, would indicate overtraining.

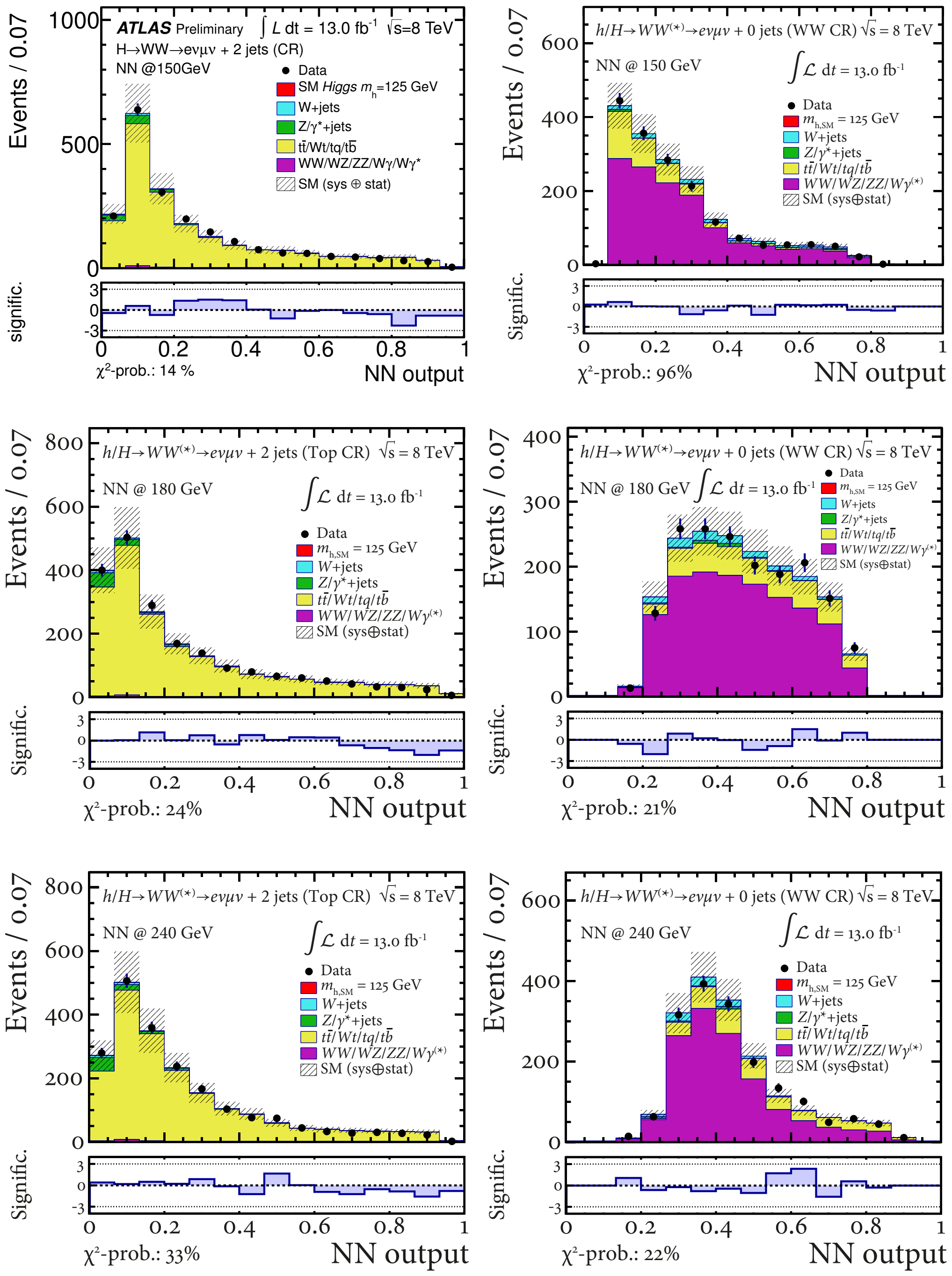


Figure 6.11: Discriminant distributions obtained with the NNs optimised for m_H = 150 GeV, 180 GeV, and 240 GeV with the standard training sample, but evaluated in the $t\bar{t}$-enriched CR (left) and in the diboson CR (right). The contribution of each process is normalised to the event yields in table 6.5. The last bin of the histogram represents an overflow bin. The sufigure (top left) with the label ATLAS Preliminary is taken from [164].

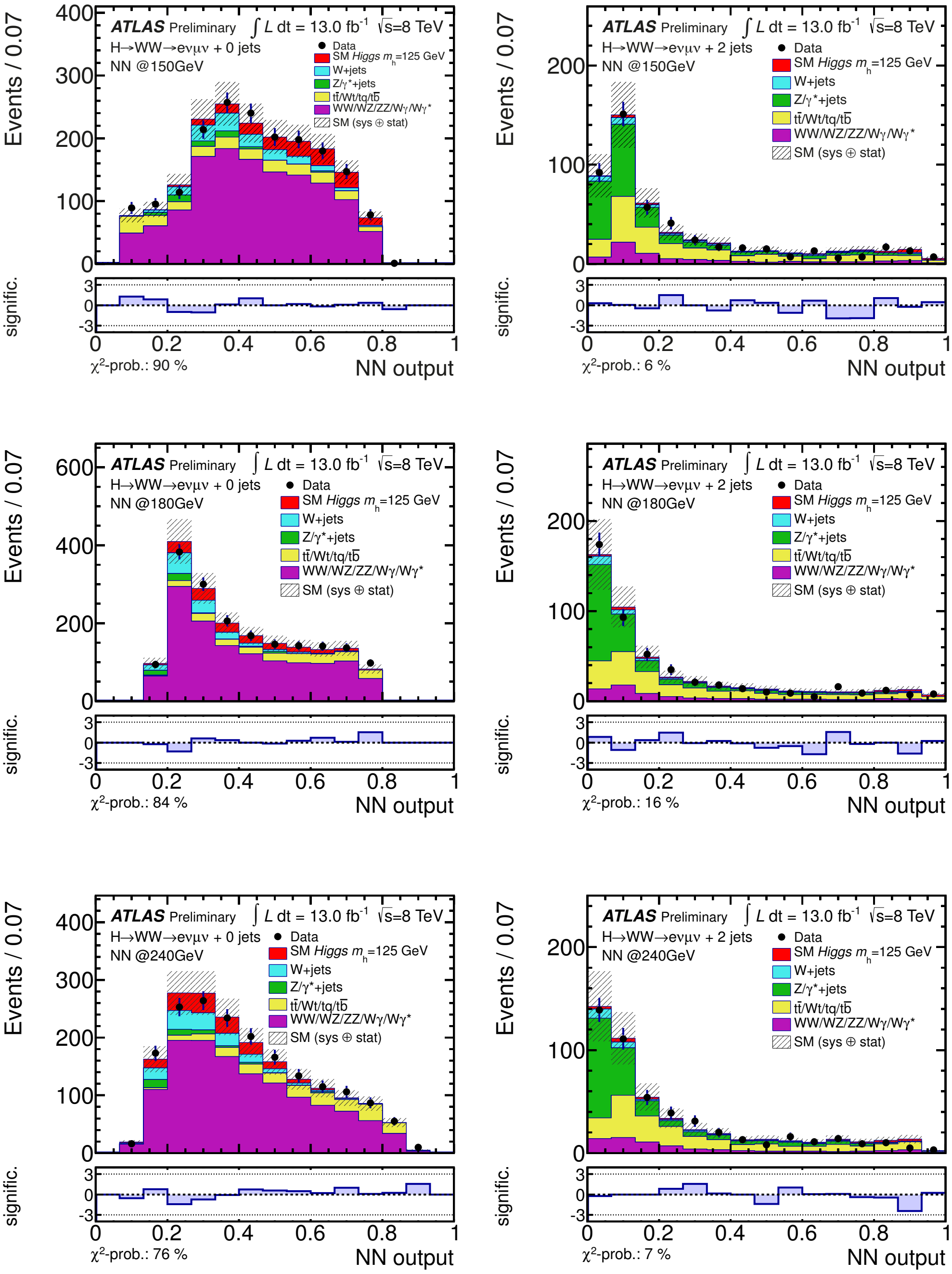


Figure 6.12: Discriminant distributions obtained from the NNs for three different Higgs-boson mass points in the 0-jet channel (left) and in the 2-jet channel (right) [164]. The contribution of each process is normalised to the event yields of the 0-jet signal channel and 2-jet signal channel in table 6.5. The last bin of the histogram represents an overflow bin.

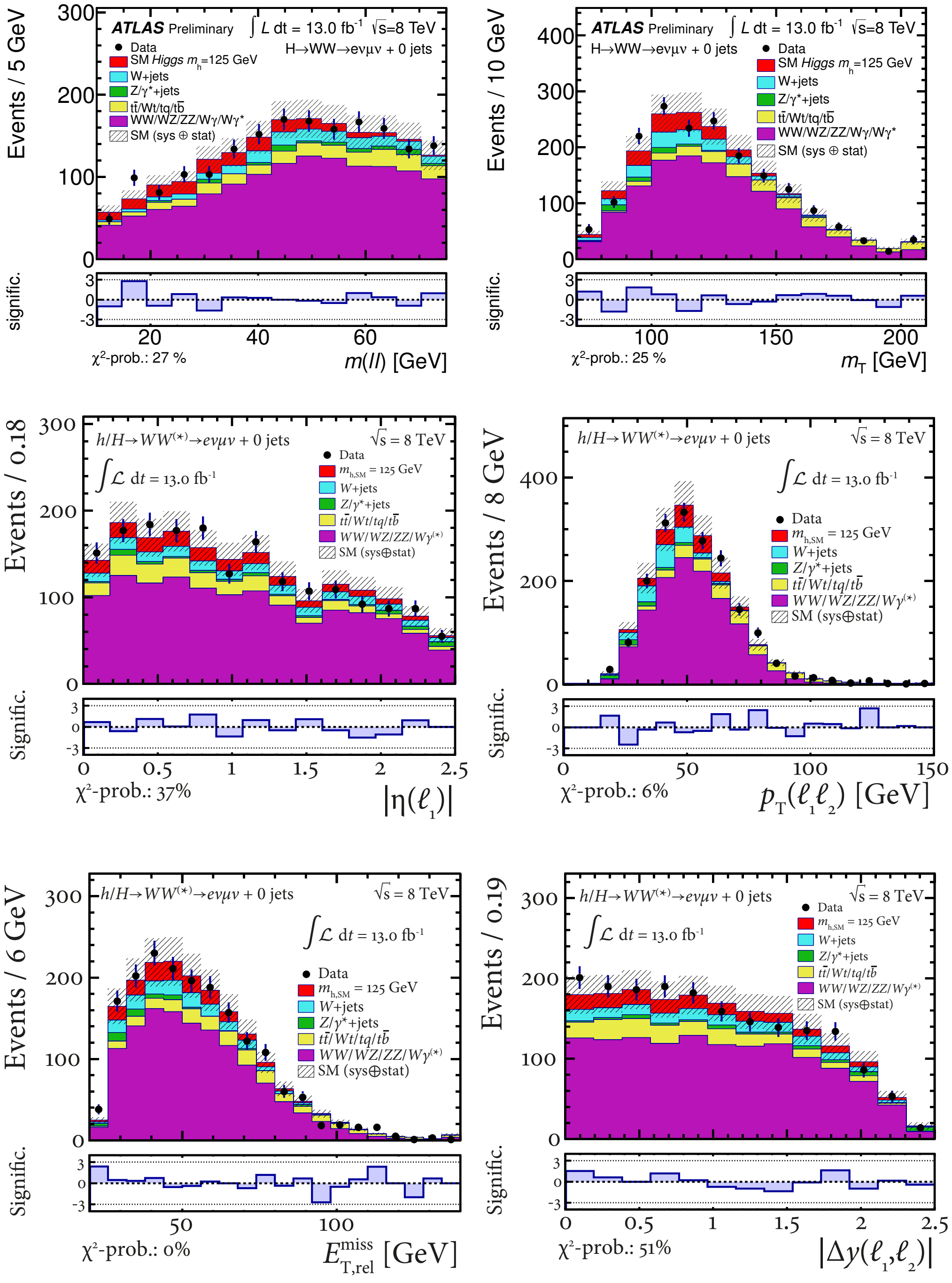


Figure 6.13: Distributions of the input variables of the NN in the signal region of the 0-jet channel. The contribution of each process is normalised to the event yields of the 0-jet signal channel in table 6.5. The last bin of the histogram represents an overflow bin. The sufigures (top) with the label ATLAS Preliminary are taken from [164].

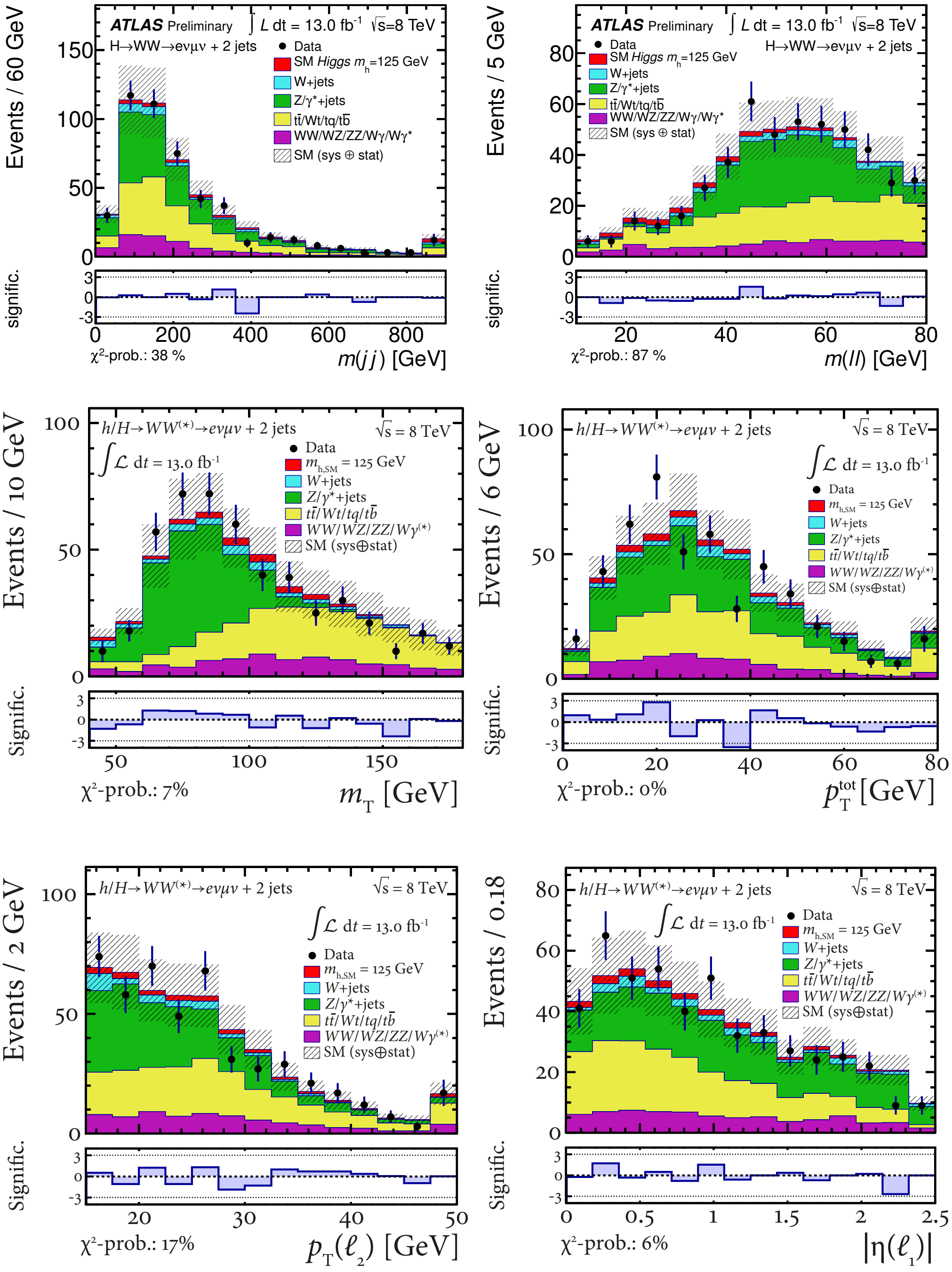


Figure 6.14: Distributions of the six highest-ranking input variables of the NN in the signal region of the 2-jet channel. The contribution of each process is normalised to the event yields of the 2-jet signal channel in table 6.5. The last bin of the histogram represents an overflow bin. The sufigures (top) with the label ATLAS Preliminary are taken from [164].

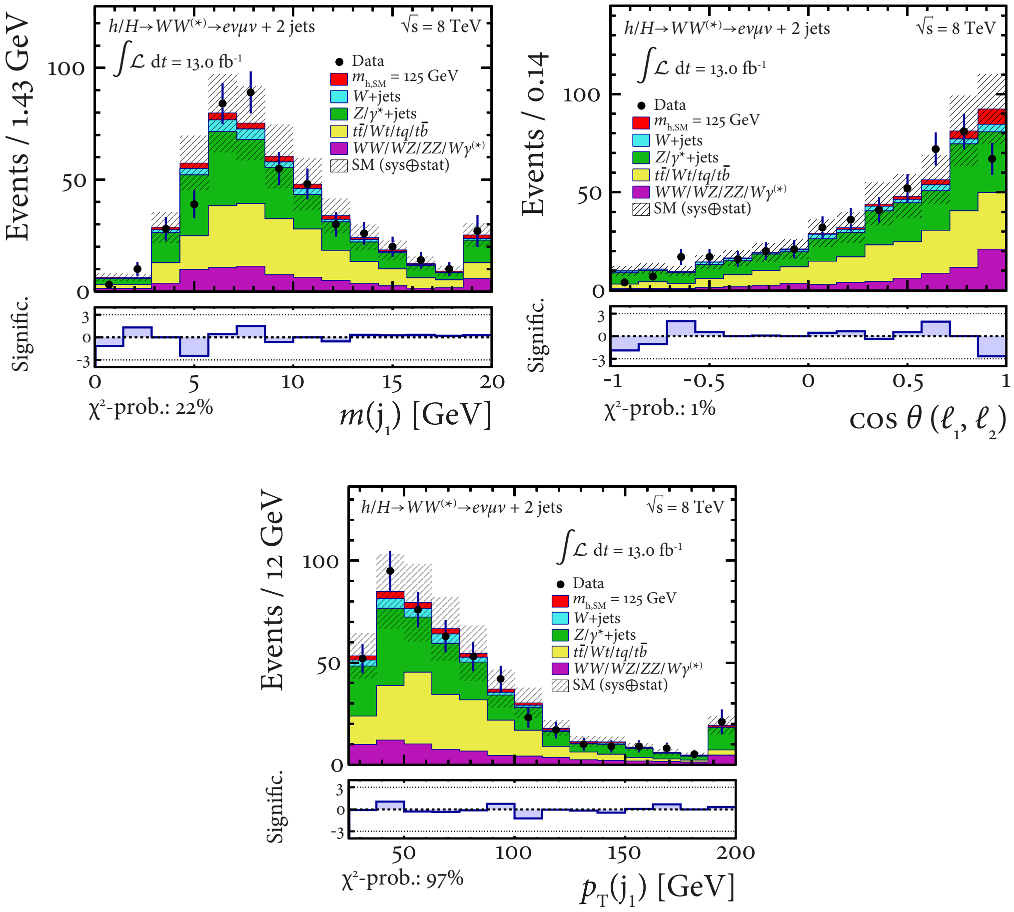


Figure 6.15: Distributions of the three lowest-ranking input variables of the NN in the signal region of the 2-jet channel. The contribution of each process is normalised to the event yields of the 2-jet signal channel in table 6.5. The last bin of the histogram represents an overflow bin.

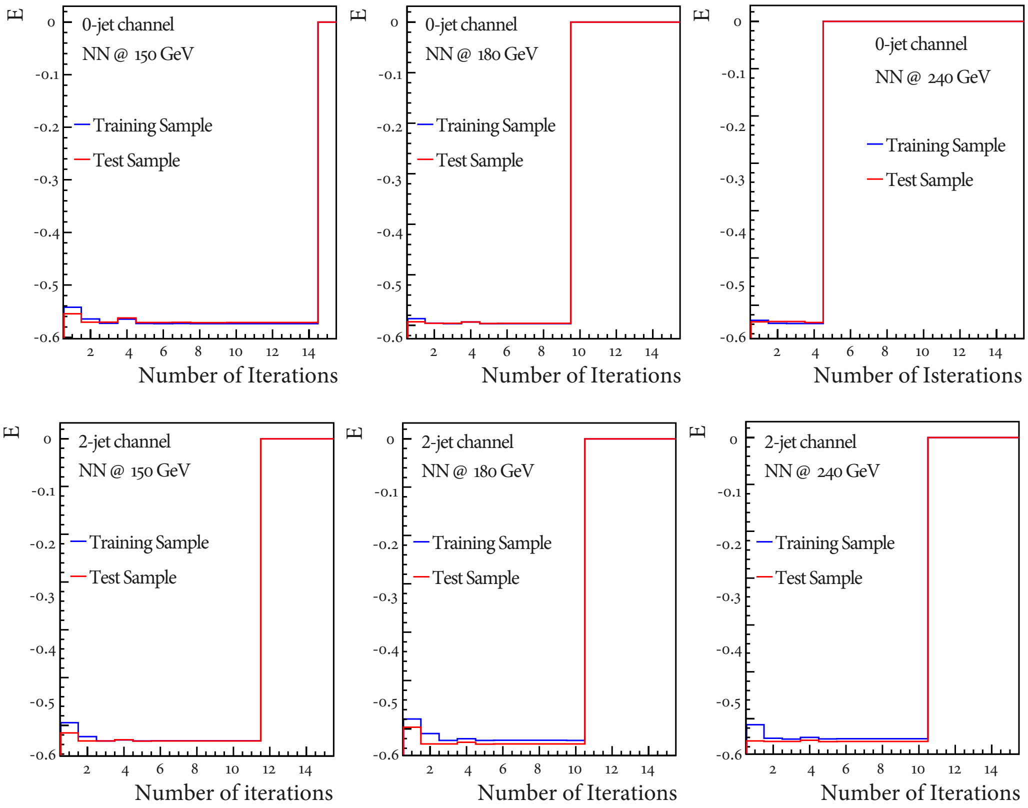


Figure 6.16: Entropy loss function in the 0-jet channel (top) and in the 2-jet channel (bottom) during the training of the neural networks. In black is the training sample and in red the overtraining test sample. The x-axis shows the number of steps in the learning process. The abrupt rise of the entropy loss function to 0 indicates the end of the training process.

No overtraining is observed. Figure 6.16 also shows that `NeuroBayes` requires only 4 to 14 iterations to finish the training.

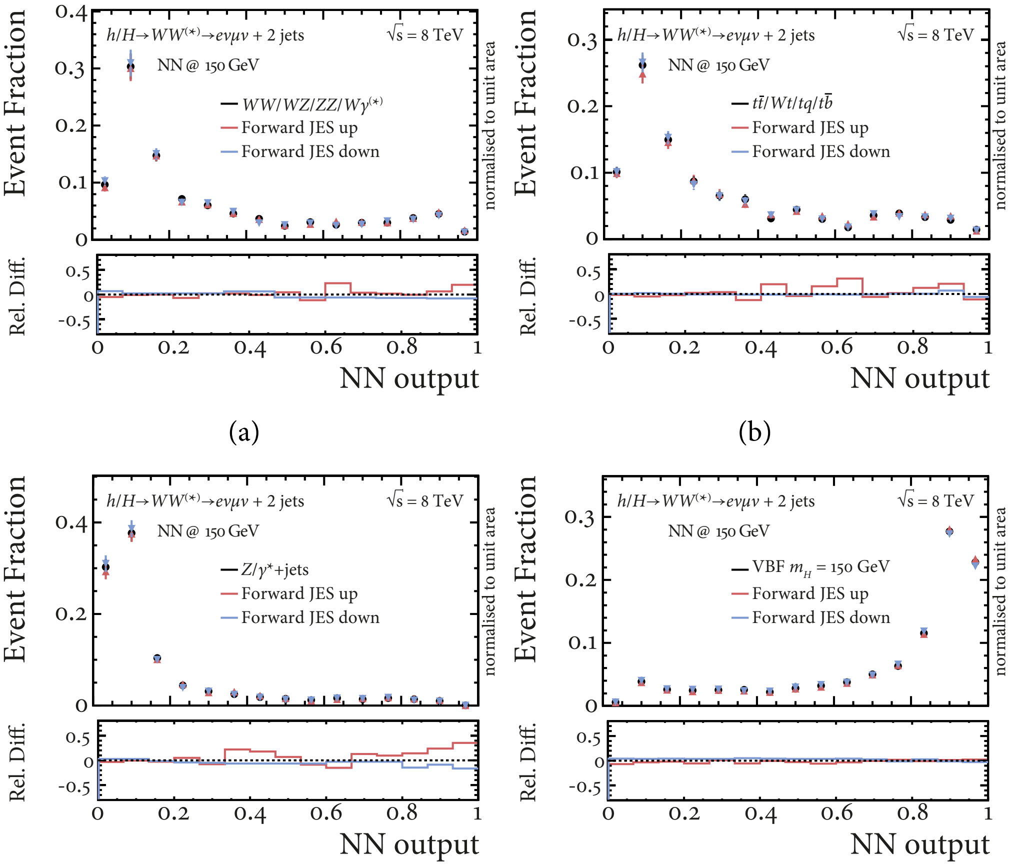
Chapter 7

Systematic Uncertainties

Systematic uncertainties on the normalisation of the different backgrounds, on the signal acceptance, and on the shape of the NN-discriminant distributions for signal and background processes reduce the sensitivity of the search for Higgs-boson production. Systematic uncertainties arise due to the limited understanding of the residual differences between data and Monte Carlo simulations for the reconstruction and energy calibration of jets, electrons, and muons. Other sources of systematic uncertainties are due to the modelling of physics processes in simulations. All considered uncertainties are fully propagated through the entire analysis. The following categories of systematic uncertainties are considered:

Jet modelling The main source of uncertainty on the modelling of jets comes from the jet energy scale (JES) [150]. For the jet definition applied in this analysis, the JES uncertainty varies from 2% to 14% as a function of p_{T} and η. Figure 7.1 shows the uncertainties on the shape of the NN-discriminant distributions due to the uncertainty on the JES of the forward jets, i.e. in the region $|\eta| > 2.5$ where the JVF criterion in equation (5.4) is not effective. An additional contribution to the JES uncertainty is caused by pile-up effects and ranges from less than 1% to 4% as a function of jet p_{T} and η. For b-quark-induced jets an additional JES uncertainty of 0.8% to 2.5%, depending on the jet p_{T}, is added in quadrature to the JES uncertainty. The flavour of the b-quark-induced jets is identified by truth information.

Scale factors, determined from collision data [170], are applied to correct the b-tagging performance in simulated events to match the data. The uncertainties on the scale factors vary from 10% to 20% for b-jets. For light-quark



(a) (b) (c) (d)

Figure 7.1: Uncertainties on the shape of the NN discriminant due to the uncertainty on the JES of the forward jets for different groups of physics processes: (a) Diboson production (WW, WZ, ZZ, $W\gamma^{(*)}$), (b) top-quark production ($t\bar{t}$, Wt, tq, $t\bar{b}$), (c) Drell-Yan, Z+jets, and W+jets production, and (d) VBF Higgs-boson production. The black dots represent the nominal NN-output distribution. The red and blue markers represent the up and down shape variations, respectively, which are obtained by evaluating the NN on the systematically altered samples. In the lower histogram, the relative difference between the nominal shape and the systematically shifted shapes is shown. The uncertainty due to the JES is causing the largest uncertainties on the shape of the NN discriminant of all considered sources of uncertainty.

jets, the mis-tagging uncertainty ranges from 20% to 50% as a function of jet p_{T} and η.

Uncertainties on the modelling of the jet energy resolution (JER) also contribute and are estimated from collision data. Additional uncertainties arise from the modelling of jets with $p_{\mathrm{T}} < 20$ GeV as well as from soft energy deposits in the calorimeter that are not associated with reconstructed physics objects. Other minor uncertainties are assigned to the reconstruction of $E_{\mathrm{T}}^{\mathrm{miss}}$ and to account for the impact of pile-up collisions on $E_{\mathrm{T}}^{\mathrm{miss}}$.

Lepton modelling Uncertainties in the electron and muon reconstruction, identification, and trigger efficiencies are estimated using tag-and-probe methods on samples enriched with $Z \to \ell\ell$, $J/\psi \to \ell\ell$, or $W^{\pm} \to \ell\nu$ ($\ell = e, \mu$) events [155–157]. Other components include the electron and muon energy scale.

Luminosity The uncertainty on the integrated luminosity is 3.6%. The preliminary calibration of the luminosity scale is derived from beam-separation scans performed in April 2012, as described in [171].

PDFs The systematic uncertainties related to the parton distribution functions are taken into account for all processes using simulated events. The events are reweighted according to each of the PDF-uncertainty eigenvectors, using the uncertainties for the `CT10` [91] PDF sets, the `MSTW2008nlo` [172] PDF sets, and the `NNPDF2.0` [173] sets. The uncertainties are calculated using the formula given in equation (43) of [174]. Following the `PDF4LHC` recommendation, the final PDF uncertainty is calculated as the envelop of the uncertainty bands of the three PDF sets. Figure 7.2 shows the efficiencies of the WW and $t\bar{t}$ processes, respectively, for the different eigenvectors. The resulting envelop is also shown.

Generator and parton shower Uncertainties in the modelling of the event kinematics with Monte Carlo generators are taken into account for the top-quark background and the diboson background. The uncertainty due to the choice of `POWHEG` as generator for the diboson background is estimated by comparing the event rate and the shape of the NN discriminants to the ones obtained with `MC@NLO`. The generator uncertainty on the $t\bar{t}$ background

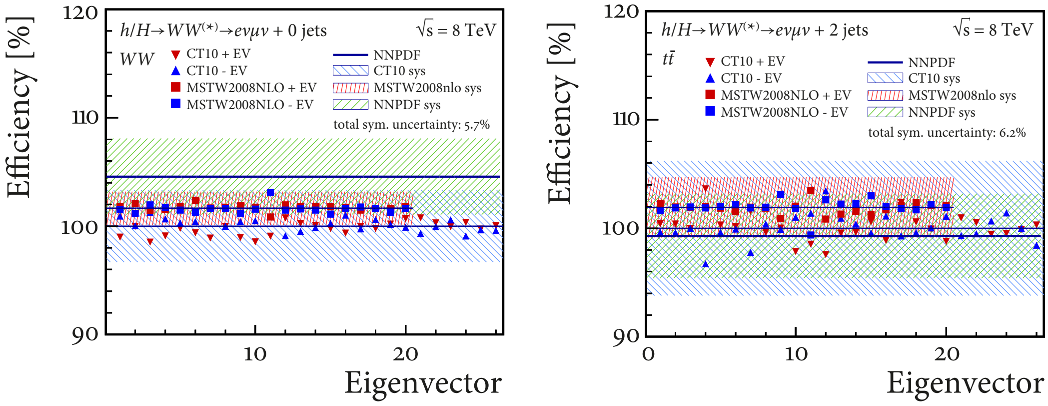


Figure 7.2: Estimated selection efficiencies relative to the nominal selection and their uncertainties for the `CT10` [91] PDF set, the `MSTW2008nlo` [172] PDF set, and the `NNPDF2.0` [173] set as a function of the eigenvector for the WW process in the 0-jet channel (left) and the $t\bar{t}$ process in the 2-jet channel (right).

is estimated by comparing `MC@NLO` and `ALPGEN`. In addition, uncertainties are assigned to the modelling of the parton showers and hadronisation by interchanging the modelling between `PYTHIA` and `HERWIG`.

Pile-up The uncertainty on the modelling of pile-up events is determined by varying the average number of collisions per bunch crossing in simulated events.

MC statistics The impact of using simulation samples of limited size is also taken into account. This is done by generating pseudo-experiments, in which the distributions of the NN discriminants is varied within their statistical uncertainty.

Background rates The rate of the top-quark background is normalised using the event yield in a control region as one component in the maximum-likelihood fit. However, when generating the pseudo-experiments for the statistical analysis the theoretical cross-section uncertainty of 22% is taken into account. The W+jets background arises from a jet being misidentified as a lepton. The uncertainty on understanding the misidentification rate and the

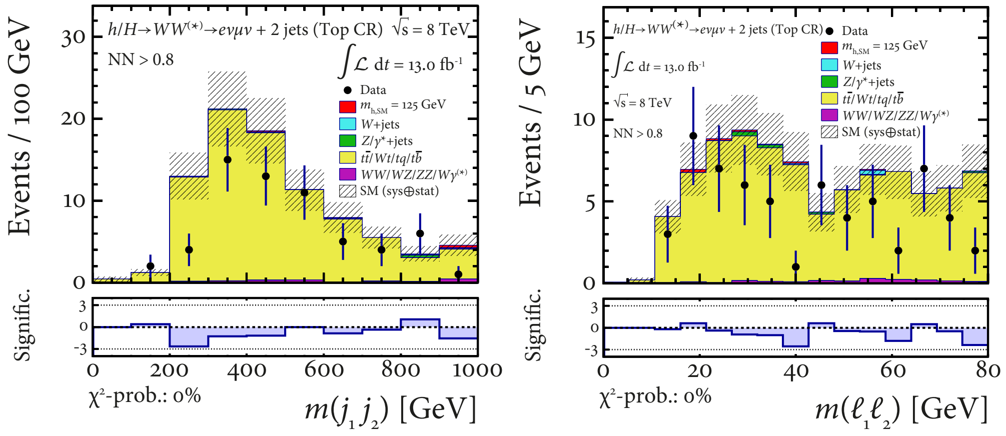


Figure 7.3: Distribution of $m(j_1 j_2)$ (left) and $m(\ell_1 \ell_2)$ (right) in the top CR after the the following cut on the NN output is applied: NN < 0.8. The contribution of each process is normalised to the fitted event rate (see table 6.6). The last bin of the histogram represents an overflow bin.

resulting W+jets event rate is 43% in the 0-jet channel and 41% in the 2-jet channel (see section 5.3). The diboson and Z+jets backgrounds are normalised according to the cross sections predicted by theoretical computations and the corresponding uncertainties are quoted in table 7.1.

In the previous chapter, it was shown that two forward jets are characteristic for the VBF topology. The ranking of the variables in the 2-jet channel (see table 6.2–6.4), which underlines the importance of the forward η region, indicates $m(j_1 j_2)$ to be the most import variable. However, the events in the top CR, which contain at least one b-tagged jet, are located more often in a different phase-space region, i.e. in the region $|\eta| \leq 2.5$ because of the geometrical acceptance of ATLAS. A cut on the NN discriminant optimised at $m_H = 150$ GeV (see figure 6.11) has been applied to check the distributions of the input variables in the top CR with jets close to the signal phase-space region, i.e. in the region between $\eta = 2.5$ and $\eta = 4.5$. It is required that NN > 0.8. Besides a good modelling of all input variables in the top CR that contains signal-like events, the distributions reveal a slight offset in the normalisation, which is shown in figure 7.3. This rate difference is covered by a large cross-section uncertainty of 22%, which is assigned to top processes, i.e. $t\bar{t}/tW/tb/tqb$.

Signal cross section The relative uncertainty on the signal cross sections (gluon fusion, VBF, and WH/ZH) is determined following [46, 133] including uncertainties on the QCD renormalisation and factorisation scales, and uncertainties on the underlying event and parton shower. The uncertainties on the QCD renormalisation and factorisation scales are evaluated by independent up and down variation of the scale by a factor of two. The relative cross-section uncertainty of these processes is assumed to be 25% for gluon fusion in the 0-jet channel and 30% in the 2-jet channel, while the uncertainty on the combined VBF/WH/ZH category is 10%. These uncertainties also account for uncertainties in the modelling of the underlying event. The cross-section uncertainties in the gluon-fusion process are treated uncorrelated between the 0-jet channel and the 2-jet channel, since they arise from different sources.

The rate and cross-section uncertainties of the background processes that are modelled based on simulated events are summarised in table 7.1. In addition to the rate and cross-section uncertainties, shape uncertainties for jet, lepton, and pile-up modelling as well as uncertainties for generators are taken into account.

Table 7.1: Relative systematic rate uncertainties for background processes in the 0-jet channel and the 2-jet channel. The uncertainties are rounded to full per cent, but all uncertainties that are smaller than 1% are rounded up. As described in the main text, the exact value of the luminosity uncertainty is 3.6% for each process. DY stands for Drell-Yan processes and the single top-quark processes are denoted $Wt/tq/t\bar{b}$.

Process	Channel	WW/WZ $ZZ/W\gamma^{(*)}$	$t\bar{t}/Wt$ $tq/t\bar{b}$	DY Z+jets
Jet modelling	0-jet	3%	14%	10%
	2-jet	11%	37%	12%
Lepton modelling	0-jet	2%	2%	6%
	2-jet	2%	2%	2%
Lumi	0-jet	4%	4%	4%
	2-jet	4%	4%	4%
PDF	0-jet	6%	6%	6%
	2-jet	5%	7%	5%
Generator	0-jet	1%	3%	–
	2-jet	2%	22%	–
Pile-up modelling	0-jet	2%	1%	2%
	2-jet	1%	1%	1%
Parton Shower	0-jet	–	7%	–
	2-jet	–	13%	–
Total	0-jet	8%	18%	14%
	2-jet	13%	46%	14%
Cross Section		25%	22%	34%

Chapter 8

Hypothesis Tests

To estimate a possible signal content of the selected samples, a binned maximum-likelihood fit is performed to the NN-output distributions and to the event yield in the $t\bar{t}$ control region. Including all bins of the NN-output distributions in the fit has the advantage of making maximal use of all signal events remaining after the event selection, and, in addition, allows for the determination the background in-situ. The likelihood fit is performed simultaneously in both signal channels, that is the 0-jet channel and the 2-jet channel, and in the $t\bar{t}$ control region. In the $t\bar{t}$ control region only the event yield is considered, since it is very pure and discrimination between different processes is not needed. The sensitivity to the background rates is given by the background dominated region of the NN output distribution, that is the region close to zero, and the event yield in the $t\bar{t}$ control region.

The binned likelihood function $\mathcal{L}$ used in the fit is given by the product of Poisson likelihoods $\mathcal{P}$ for the individual histogram bins k and Gaussian priors $\mathcal{N}$ with a mean of one and a width of Δ_j:

$$\mathcal{L}\left(\beta_i^{\mathrm{s}}, \beta_j^{\mathrm{b}}\right) = \prod_k \mathcal{P}\left(n_k, \mu_k\left(\beta_i^{\mathrm{s}}, \beta_j^{\mathrm{b}}\right)\right) \times \prod_{j=1}^{B-1} \mathcal{N}\left(\beta_j^{\mathrm{b}}, 1, \Delta_j\right) \ . \tag{8.1}$$

Within the likelihood function the predicted signal and background rates are multiplied by scale factors β_i^{s} for the signal processes and β_j^{b} for the background processes, respectively. The indices i and j running over all signal and background processes, respectively. Gluon fusion and VBF are considered as different processes, so that each Higgs boson of a 2HDM has two signal processes. Therefore, i is running over four signal process. n_k is the number of observed events in bin k, and μ_k is the mean value of the estimated number of expected events

$$\mu_k\left(\beta_i^{\mathrm{s}},\beta_j^{\mathrm{b}}\right) = \sum_i^S \mu_k\left(\beta_i^{\mathrm{s}}\right) + \sum_{j=1}^B \mu_k(\beta_j^{\mathrm{b}}) \tag{8.2}$$

$$= \sum_i^S \beta_i^{\mathrm{s}}\,\tilde{\nu}_i^{\mathrm{s}}\,\alpha_{k,i}^{\mathrm{s}} + \sum_{j=1}^B \beta_j^{\mathrm{b}}\,\tilde{\nu}_j^{\mathrm{b}}\,\alpha_{k,j}^{\mathrm{b}}\,, \tag{8.3}$$

where $\tilde{\nu}_i^{\mathrm{s}}$ and $\tilde{\nu}_j^{\mathrm{b}}$ represent the predicted number of events of the signal process i and the background process j, respectively. $\alpha_{k,i}^{\mathrm{s}}$ and $\alpha_{k,j}^{\mathrm{b}}$ denote the relative fraction in bin k of events of the signal process i and the background process j, respectively. The Gaussian priors constrain the rate of each background process $j \in \{1,\ldots,B-1\}$, except for the $t\bar{t}$ background $j = B$, and are defined by

$$\mathcal{N}\left(\beta_j^{\mathrm{b}},1,\Delta_j\right) = \frac{1}{2\pi\Delta_j^2}\exp\left(\frac{-\left(\beta_j^{\mathrm{b}}-1\right)^2}{2\Delta_j^2}\right), \tag{8.4}$$

where Δ_j represents the relative uncertainty on the cross section of the background process j. The Poisson likelihood is defined by

$$\mathcal{P}(n_k,\mu_k) = \frac{e^{-\mu_k}\,\mu_k^{n_k}}{n_k!}\,. \tag{8.5}$$

The meaning of all ingredients of the likelihood function are explained in table 8.1.

Table 8.1: Notations and meaning of all functions, variables, and indices used in the likelihood function.

Function	Meaning
$\mathcal{L}$	Likelihood function
$\mathcal{P}$	Poisson likelihood
$\mathcal{N}$	Gaussian priors
Variable	**Meaning**
β_i^s	Scale factor of the signal process i
β_j^b	Scale factor of the background process j
n_k	Number of observed events
μ_k	Mean value of the estimated number of expected events
Δ_j	Relative uncertainty on the cross section of the background process j
$\tilde{\nu}_i^\mathrm{s}$	Predicted number of events of the signal process i
$\tilde{\nu}_j^\mathrm{b}$	Predicted number of events of the background process j
$\alpha_{k,i}^\mathrm{s}$	Relative fraction in bin k of events of the signal process i
$\alpha_{k,j}^\mathrm{b}$	Relative fraction in bin k of events of the background process j
S	Number of signal processes
B	Number of background processes
Index	**Meaning**
k	Histogram bin
i	Signal process
j	Background process

8.1 The q-value Test Statistic

The compatibility of the observed data with different signal hypotheses, which depend on the parameters of the 2HDM, and the background-only hypothesis is evaluated by performing frequentist hypotheses tests based on pseudo-experiments. Two hypotheses are compared. The null hypothesis H_0 assumes, that there is no Higgs boson at all. The signal hypothesis H_1 assumes a Higgs-boson signal as predicted by a specific 2HDM, depending on the values of $\cos\alpha$, $\tan\beta$, and m_H. The light Higgs-boson h of a 2HDM is part of the signal hypothesis, and the mass of the h boson is assumed to be 125 Gev. In this way the light Higgs-boson is compatible with the Higgs-like boson observed at 125 Gev. The signal strength of the h boson used in the statistical test is the one predicted by the 2HDM under consideration. In addition, the signal hypothesis includes the contribution of the heavy Higgs-boson H of the considered 2HDM. Therefore, both Higgs-bosons of a 2HDM are treated collectively and the combined signal model corresponds to $\sum_i^S \mu_k\left(\beta_i^{\mathrm{s}}\right)$ events in bin k in the likelihood function. For both scenarios ensemble tests, that is large sets of pseudo-experiments, are performed as described in [175]. As a test-statistic the so-called q-value is used. The q-value is defined as

$$q \equiv -2\ln\left(\frac{\mathcal{L}\left(\beta^{\mathrm{s}}=1,\hat{\beta}_j^{\mathrm{b}}\right)}{\mathcal{L}\left(\beta^{s}=0,\hat{\hat{\beta}}_j^{\mathrm{b}}\right)}\right), \tag{8.6}$$

where $\mathcal{L}\left(\beta^{\mathrm{s}}=1,\hat{\beta}_j^{\mathrm{b}}\right)$ is the likelihood function evaluated when fixing the signal cross sections to their predicted values in the 2HDM, i.e. $\beta^{\mathrm{s}}=1$, while $\mathcal{L}\left(\beta^{s}=0,\hat{\hat{\beta}}_j^{\mathrm{b}}\right)$ is the likelihood function evaluated by setting the signal cross sections to zero, i.e. $\beta^{\mathrm{s}}=0$. The background scale factors β_j^{b} are left free to float in both cases within their uncertainties. The two fits yield the two maxima $\mathcal{L}(\beta^{\mathrm{s}}=1,\hat{\beta}_j^{\mathrm{b}})$ and $\mathcal{L}(\beta^{s}=0,\hat{\hat{\beta}}_j^{\mathrm{b}})$, respectively. The hats indicate the estimators of the scale factors, which are determined in the fit. The q-value is computed for each of the two hypotheses of each pseudo experiment. Each time the pseudo data are fitted twice to obtain the two values of the likelihood function.

By running a large set of pseudo experiments one obtains the q-value distribution of the null hypothesis, denoted by $g_0(q)$, and the q-value distribution of the signal hypothesis $g_1(q)$. By normalising these distributions to unit area, one obtains the probability densities $\hat{g}_0(q)$ and $\hat{g}_1(q)$.

To test the compatibility of the observed data with the two hypotheses H_0 and H_1, one computes the observed q-value q^{obs} by fitting the discriminant distributions in data and compares q^{obs} with the two probability densities $\hat{g}_0(q)$ and $\hat{g}_1(q)$.

8.2 Exclusion Limits

The CL_s method [176, 177] is used to derive confidence levels (CL) for a certain value of q^{obs}. Based on the q-value probability densities $\hat{g}_0(q)$ and $\hat{g}_1(q)$ one first computes the background p value, also referred to as p_b, and the signal plus background $p_{\text{s+b}}$ value as follows (see figure 8.1):

$$\begin{aligned} p_\text{b}^{\text{obs}} &= \quad p_\text{b}\left(q^{\text{obs}}\right) &= \int_{-\infty}^{q^{\text{obs}}} \hat{g}_0\left(q'\right) dq' \qquad (8.7) \\ p_{\text{b+s}}^{\text{obs}} &= \quad p_{\text{b+s}}\left(q^{\text{obs}}\right) &= \int_{q^{\text{obs}}}^{-\infty} \hat{g}_1\left(q'\right) dq'\,. \qquad (8.8) \end{aligned}$$

The expected p values p_b^{exp} and $p_{\text{b+s}}^{\text{exp}}$ are calculated by using the median of the $\hat{g}_1$ and $\hat{g}_0$ distributions, $\tilde{q}_{\hat{g}_1}^{\text{exp}}$ and $\tilde{q}_{\hat{g}_0}^{\text{exp}}$, as the upper and lower bound, respectively (see figure 8.1):

$$\begin{aligned} p_\text{b}^{\text{exp}} &= \quad p_\text{b}\left(\tilde{q}_{\hat{g}_1}^{\text{exp}}\right) &= \int_{-\infty}^{\tilde{q}_{\hat{g}_1}^{\text{exp}}} \hat{g}_0\left(q'\right) dq' \qquad (8.9) \\ p_{\text{b+s}}^{\text{exp}} &= \quad p_{\text{b+s}}\left(\tilde{q}_{\hat{g}_0}^{\text{exp}}\right) &= \int_{\tilde{q}_{\hat{g}_0}^{\text{exp}}}^{-\infty} \hat{g}_1\left(q'\right) dq'\,. \qquad (8.10) \end{aligned}$$

The definition of CL_s is given by:

$$\text{CL}_\text{s} = \frac{p_{\text{b+s}}}{1 - p_\text{b}}. \qquad (8.11)$$

Due to the denominator, the CL_s method includes the sensitivity to distinguish between the considered hypotheses and thus avoids spurious exclusions. A particular signal hypothesis H_1, determined by a triplet of the 2HDM parameters $\cos\alpha$, $\tan\beta$, and m_H, is excluded at the 95% CL if one finds $\text{CL}_\text{s} < 0.05$.

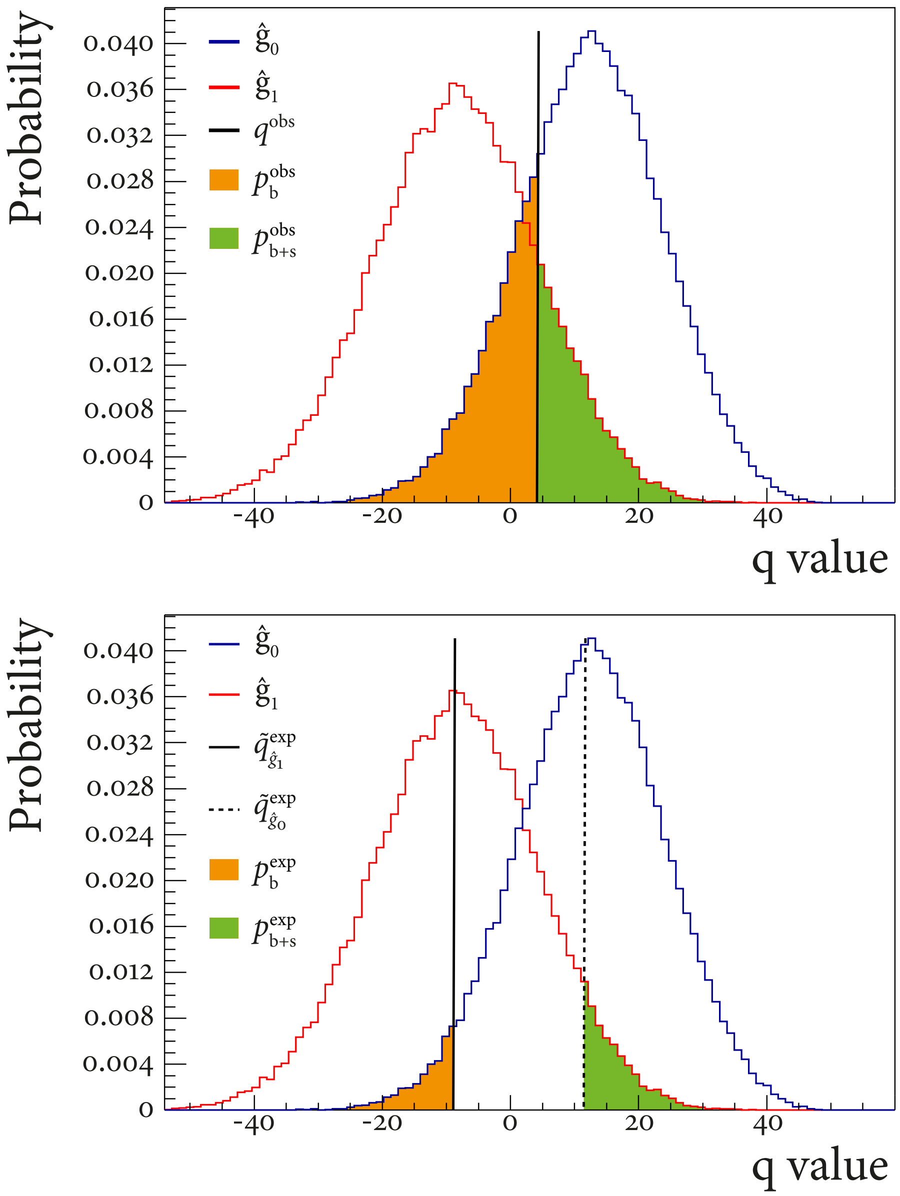


Figure 8.1: Example distribution of $\hat{g}_0$ (blue) and $\hat{g}_1$ (red). The observed p values are shown on the top, where the area under $\hat{g}_0$ from minus infinity to q^{obs} represents the observed p value in the null Hypothesis: p_b^{obs}. The area under $\hat{g}_1$ from q^{obs} to plus infinity represents the observed p value in the signal Hypothesis: p_{b+s}^{obs}. The expected p value in the null Hypothesis p_b^{exp} is depicted on the bottom by the area under the $\hat{g}_0$ distribution from minus infinity to $\tilde{g}_1$, i.e. the median of the $\hat{g}_1$ distribution. The expected p value in the signal Hypothesis p_{b+s}^{exp} is depicted by the area under the $\hat{g}_1$ distribution from $\tilde{g}_0$ to plus infinity.

8.3 Comparison of Sensitivity

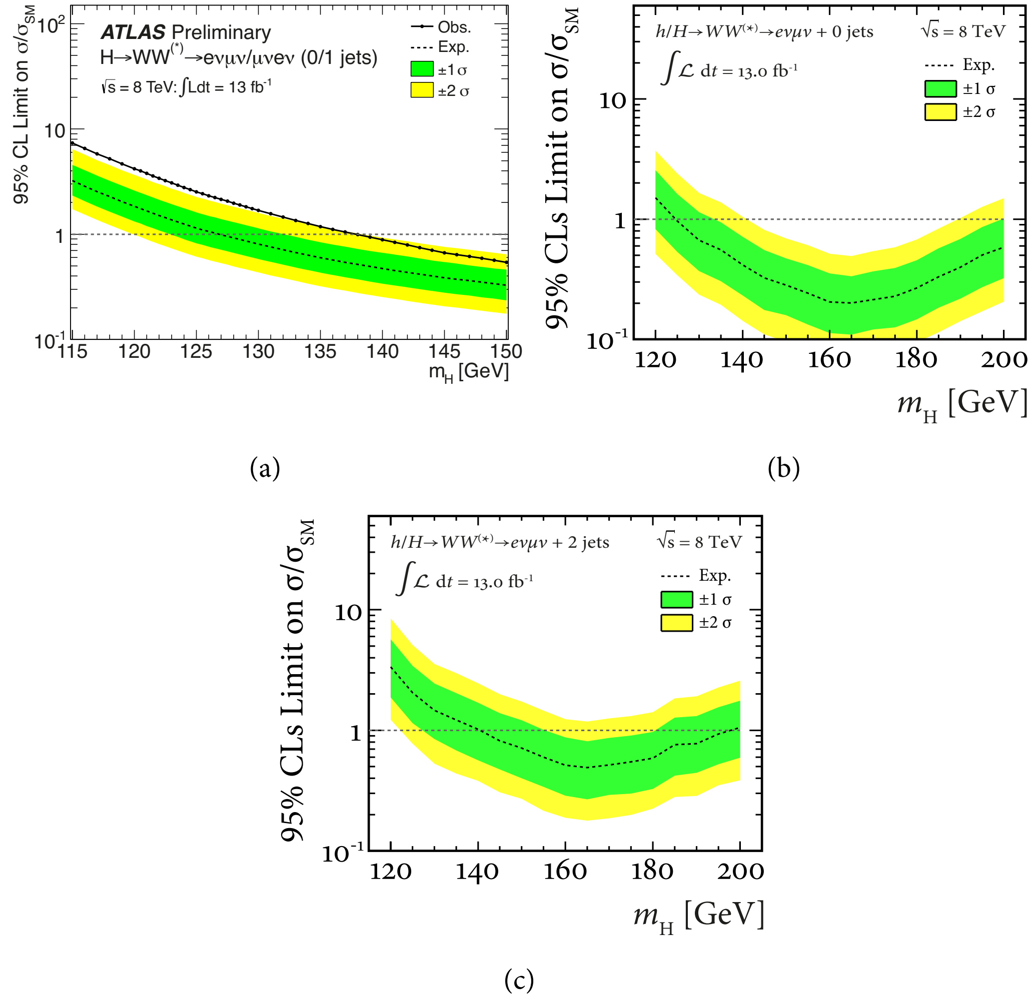
In this section, the sensitivity of the analysis described in this thesis is compared to what the SM $H\rightarrow WW^{(*)}\rightarrow \ell\nu\ell\nu$ analysis published in [139] can achieve. In both analyses, the null hypothesis is identical, but the signal hypothesis of this analysis has to be changed to accommodate the SM Higgs boson. Therefore, in this section, the signal hypothesis consists of all background processes plus the SM Higgs boson.

The expected 95% CL upper limits on the cross section divided by the SM Higgs-boson cross-section expectation have been calculated to test whether the sensitivity of this analysis is comparable with the SM $H\rightarrow WW^{(*)}\rightarrow \ell\nu\ell\nu$ analysis. To calculate the expected 95% CL upper limits the SM cross section is varied. If the computed $\mathrm{CL_s}$ value is larger than 0.05, we successively generate ensembles of pseudo experiments with increasing cross-section of the signal process and compute the corresponding $\mathrm{CL_s}$ value (see equation (8.11)) until it reaches 0.05. The observed limit is determined applying the same procedure using the observed q value q^{obs}.

The production cross section that reaches the $\mathrm{CL_s}$ value of 0.05 is excluded at 95% CL and denotes the expected limit. Figure 8.2(a) shows the exclusion limit of the cut-based SM analysis from [139], which combines the 0-jet channel and the 1-jet channel. The cut-based SM analysis expected to exclude a SM Higgs boson with a mass larger than 127 GeV. The lower limit of the exclusion power of this analysis, using the 0-jet channel only, is at 125 GeV (see figure 8.2(b)). The sensitivity of the presented analysis is quite comparable to what the SM $H\rightarrow WW^{(*)}\rightarrow \ell\nu\ell\nu$ analysis can achieve.

The $\pm 1\sigma$ uncertainty bands are calculated by changing the upper limit of the integral in (8.9) and the lower limit of the integral in (8.10) to the $\pm 1\sigma$ quantiles of the $\hat{g}_0(q)$ and $\hat{g}_1(q)$ distributions. In the same way, the $\pm 2\sigma$ uncertainty band are calculated by changing the limits of the integrals to the corresponding $\pm 2\sigma$ quantiles of $\hat{g}_0(q)$ and $\hat{g}_1(q)$.

The expected 95% CL upper limits on the SM Higgs-boson cross-section expectation in the 2-jet channel are shown in figure 8.2(c). These limits have been compared internally in ATLAS with the expectation of the SM analysis and feature similar exclusion power.



(a)

(b)

(c)

Figure 8.2: Expected 95% CL upper limits on the cross section normalised to the SM Higgs-boson cross section of the SM cut-based analysis (a) from [139], in which the 0-jet and 1-jet channel are combined, in comparison with the SM exclusion limit of this analysis in the 0-jet channel (b). Expected 95% CL upper limits on the cross section in the 2-jet channel normalised to the SM Higgs-boson cross section is shown in figure (c). The dashed and full black line show the expected and observed limit as a function of m_H, respectively. The green and yellow bands indicate the regions of $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty, respectively, in which the limit is expected to lie in the absence of a signal.

Chapter 9

Results

The simulation of the light and the heavy Higgs boson of the 2HDMs is based on the MC samples generated for the SM Higgs boson. The coupling constants of the SM Higgs boson are changed in a consistent way according to table 2.4. As a consequence of the differing coupling constants, the total decay width of both Higgs bosons, $\Gamma_{\mathrm{tot}}^{h/H}$, is also changed relative to the decay width of the SM Higgs boson $\Gamma_{\mathrm{tot}}^{h_{\mathrm{SM}}}$. Therefore, $\Gamma_{\mathrm{tot}}^{h/H}$ has to be checked before calculating the confidence limits as explained in the last chapter. Figure 9.1 shows $\Gamma_{\mathrm{tot}}^{h_{\mathrm{SM}}}$ as a function of $m_{h_{\mathrm{SM}}}$ ranging between 90 GeV and 1000 GeV. In the mass region in which a good sensitivity to the signal can be achieved, 135 GeV–300 GeV, the maximal value of $\Gamma_{\mathrm{tot}}^{h_{\mathrm{SM}}}$ is 10 GeV. If the difference between $\Gamma_{\mathrm{tot}}^{h_{\mathrm{SM}}}$ and $\Gamma_{\mathrm{tot}}^{h/H}$ is small compared to the mass resolution of about 18 GeV in the $h/H \to WW^{(*)} \to e\nu\mu\nu$ channel (see section 6.2), the change of kinematic topologies of the MC samples of the SM Higgs boson is negligible compared to the H boson of the 2HDMs. This check has to be done for all considered masses of the two Higgs bosons as well as for all values of α and $\tan\beta$ to decide in which regions a limit can be calculated reliably.

$\Gamma_{\mathrm{tot}}^{h/H}$ is calculated in the $\cos\alpha$–$m_{h/H}$ plane for various values of $\tan\beta$ as follows:

$$\Gamma_{\mathrm{tot}}^{h/H} = \Gamma_{\mathrm{tot}}^{h_{\mathrm{SM}}} \sum_X \mathcal{B}\left(h_{\mathrm{SM}} \to XX\right) \cdot \left(\xi_{h/H}^X\right)^2 , \tag{9.1}$$

where the coupling constants $\xi_{h/H}^X$ are taken from table 2.4 and $X = W, Z, b$. For $\tan\beta = 3$ and $\tan\beta = 50$ the total decay width of the heavy Higgs boson are shown in figure 9.2. In the type-I 2HDM, the total decay width is almost independent of $\tan\beta$ and with maximum values around 8 GeV, noncritical

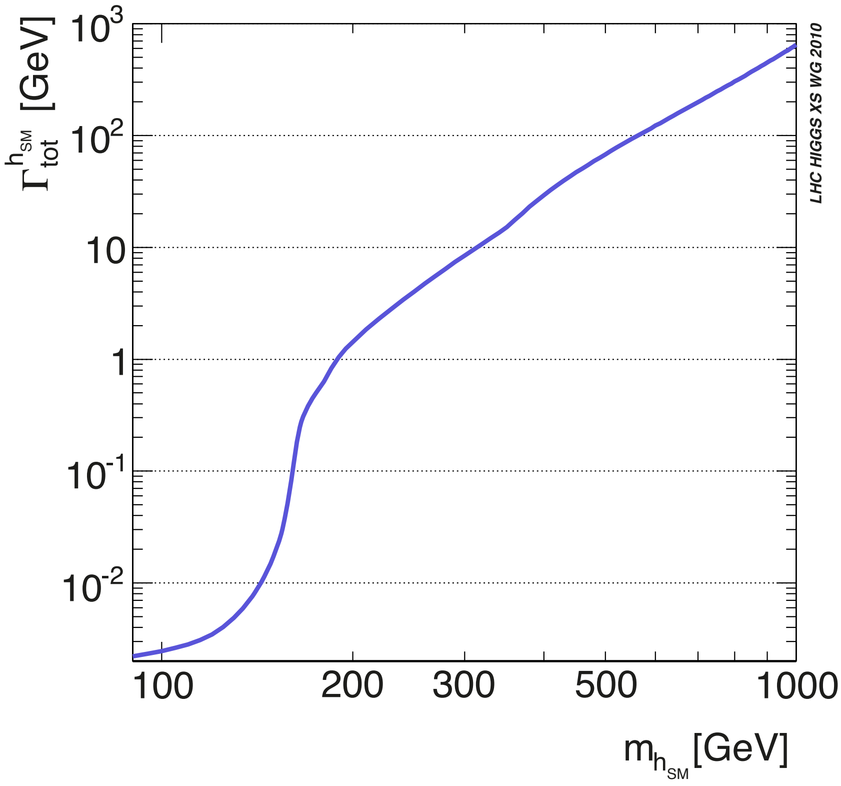
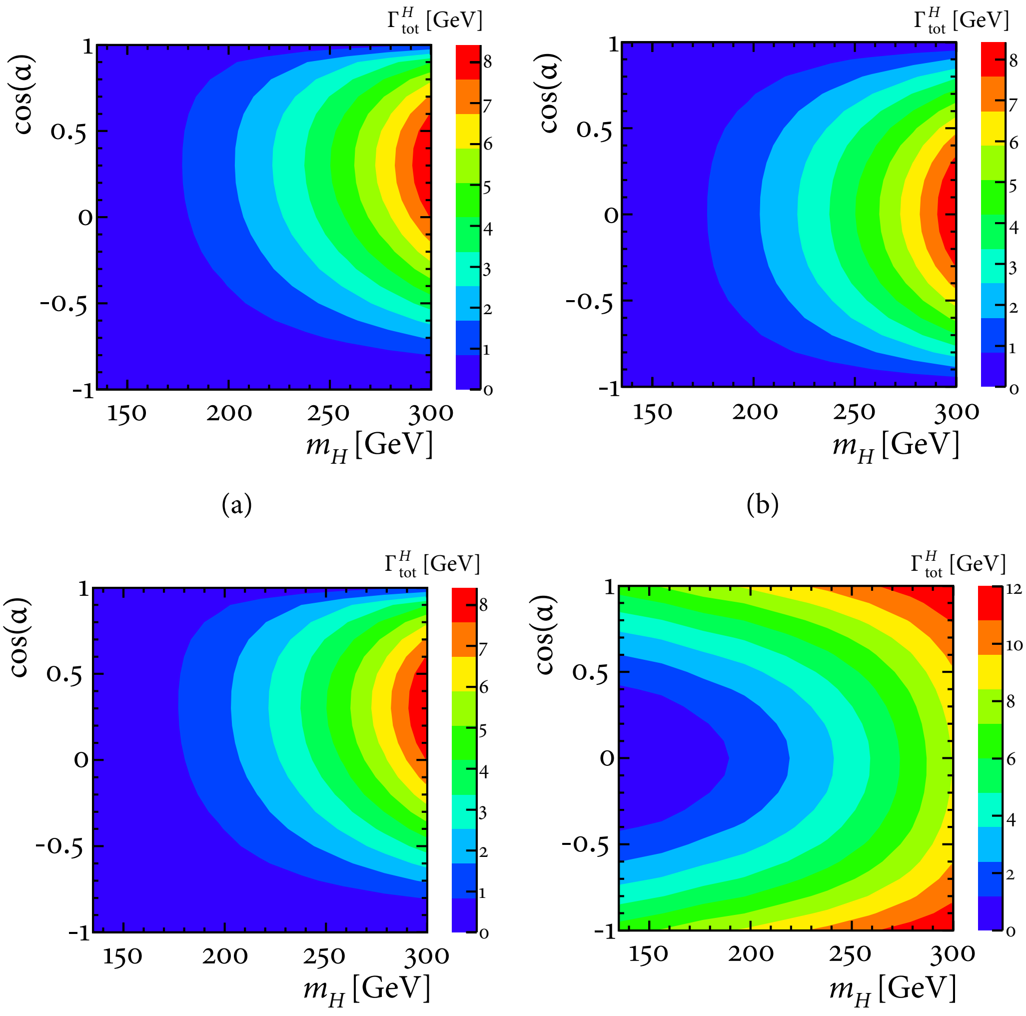


Figure 9.1: Total decay width of the SM Higgs boson as a function of its mass [46].

compared to the total decay width of the SM (see figure 9.2(a) and 9.2(b)). In the type-II 2HDM, the shape of the total decay width changes much between $\tan\beta = 3$ and $\tan\beta = 50$. But nonetheless, with a maximum value around 12 GeV the total decay width is non-critical for small values of $\tan\beta$ (see figure 9.2(c)) as well as for high values of $\tan\beta$ (see figure 9.2(c)) compared to the total decay width of the SM Higgs boson. Therefore, all values of $\tan\beta$ can be investigated in the mass range from 135 GeV to 300 GeV.

The search for a heavy Higgs boson is performed for a large part of the 2HDM parameter space, considering type-I as well as type-II models. The mixing angle α is scanned in $\cos\alpha$ in steps of 0.1. The following values of $\tan\beta$ are considered: 1, 3, 6, 20 and 50, for the type-I model as well as for the type-II model. The mass range of $135 < m_H < 300$ GeV is considered for the mass of the CP-even Higgs boson H, avoiding the mass region close to the light Higgs boson at $m_h = 125$ Gev. The mass range is scanned in steps of 5 GeV from



(a)

(b)

(c)

(d)

Figure 9.2: The total decay width of the heavy Higgs boson in the type-I 2HDM for $\tan\beta = 3$ (a) and $\tan\beta = 50$ (b), and in the type-II 2HDM for $\tan\beta = 3$ (c) and $\tan\beta = 50$ (d).

m_H = 135 GeV to 200 GeV and in steps of 20 GeV in the range from 20 GeV to 300 GeV. For each combination of these parameters the $\mathrm{CL_s}$ values are determined and exclusion contours are drawn in the $\cos\alpha$–m_H plane at the 95% CL. The results are given in figure 9.3 for type-I 2HDM and in figure 9.4 for type-II 2HDMs. For further illustration exclusion contours at the 99% CL are also displayed.

To get an understanding of the *a-priori* sensitivity of the analysis, the expected exclusion contours are computed. In this calculation, the median of the $\hat{g}_0(q)$ and $\hat{g}_1(q)$ distribution are used as limits of the integral in equation (8.10) and equation (8.9), respectively. In general, one finds a very good agreement between the expected and the observed exclusion. In most cases, the observed exclusion is slightly smaller than the expected one, since one finds an excess of events over the expectation. This is particularly true when approaching the mass region close to the observed Higgs boson at 125 GeV, which was not included in the null hypothesis. Not including it reduces the background estimate and diminishes the region, in which the heavy Higgs boson is expected to be excluded at a CL of 95% and 99%, respectively.

In the mass region from 135 to 200 GeV, a large part of the $\cos\alpha$–m_H plane can be excluded. In type-I models with $\cos\alpha \approx 0$ and $\tan\beta = 1$, masses up to 250 GeV are excluded. For type-II models the excluded region shrinks strongly with increasing $\tan\beta$, since the branching fraction to the WW final state decreases. The reason for this is that the couplings to third generation fermions, namely the bottom quark and the tau lepton, scale with $\tan\beta$ in type-II models. As the branching fraction to $b\bar{b}$ and $\tau^+\tau^-$ final states increases, the branching fraction of the WW channel decreases.

The Higgs boson observed at a mass of approximately 125 GeV features SM-like characteristics as shown in up-to-date results by ATLAS [178–180] and CMS [181–183]. To exclude the heavy Higgs boson of a 2HDM in the context of a SM-like Higgs boson the null hypothesis has to be changed, while the signal hypothesis still remains the same representing a 2HDM. In this approach, the Higgs boson with a SM cross section at 125 GeV is included in the null hypothesis as an additional process.

To check that the analysis presented in this thesis is sensitive to discover a heavy 2HDM Higgs boson compared to a null hypothesis representing the SM, the expected p value, $p_\mathrm{b}^\mathrm{exp}$, in Gaussian standard deviations are shown in figure 9.5 for $\tan\beta = 1$ in the type-I and type-II 2HDM. Figure 9.5 shows that this

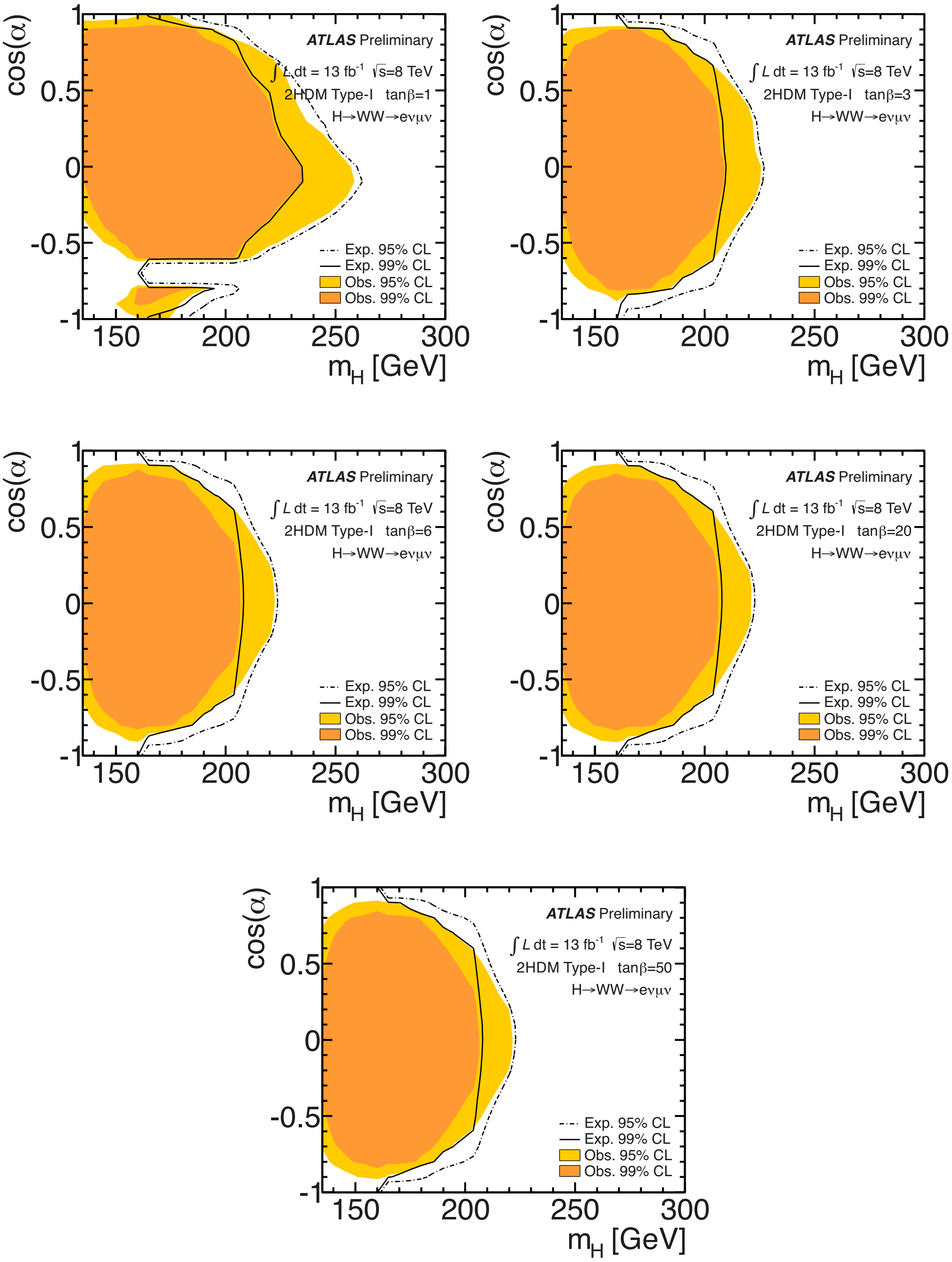


Figure 9.3: Exclusion contours in the $\cos\alpha$–m_H plane for type-I 2HDMs [164].

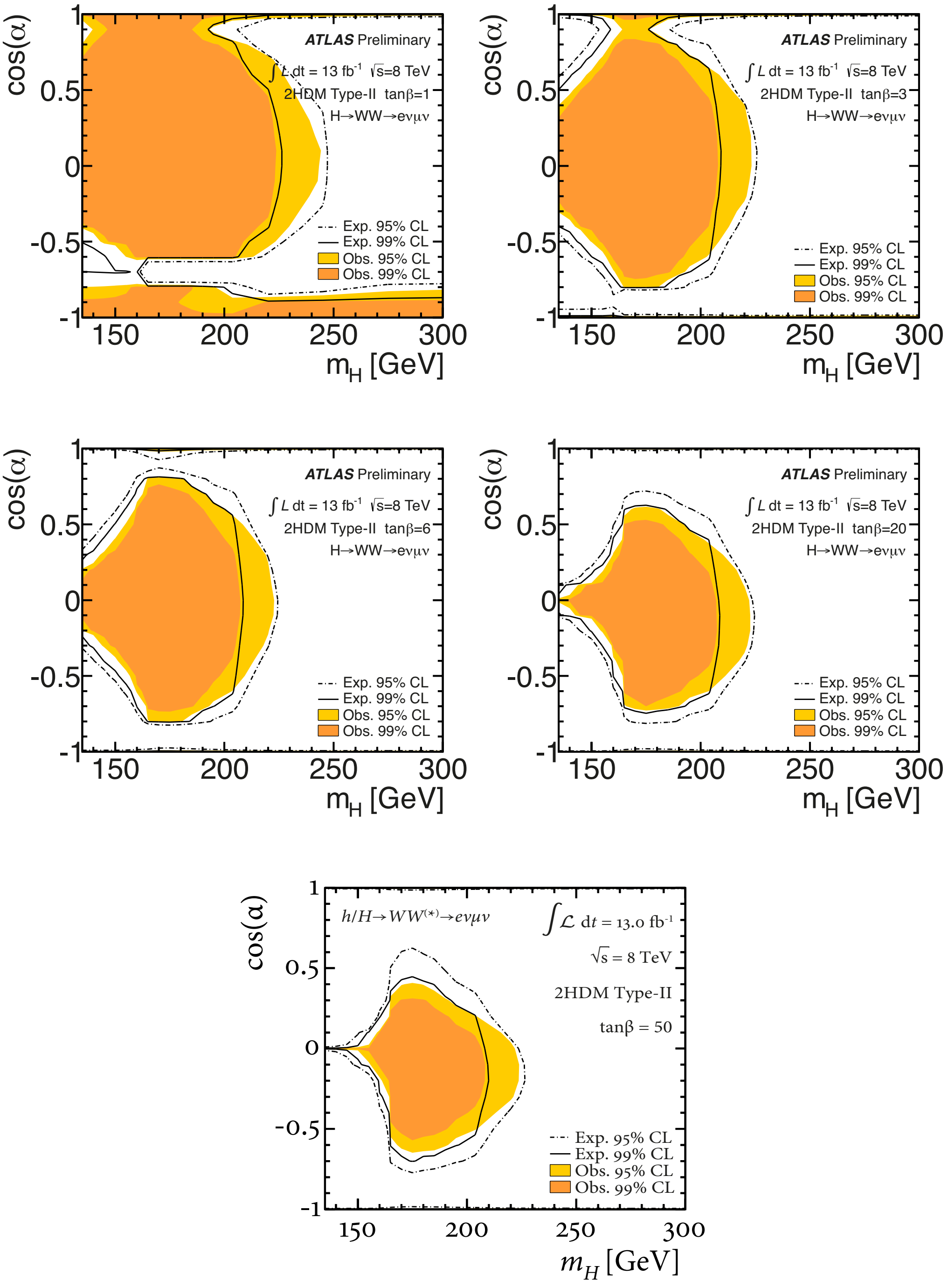
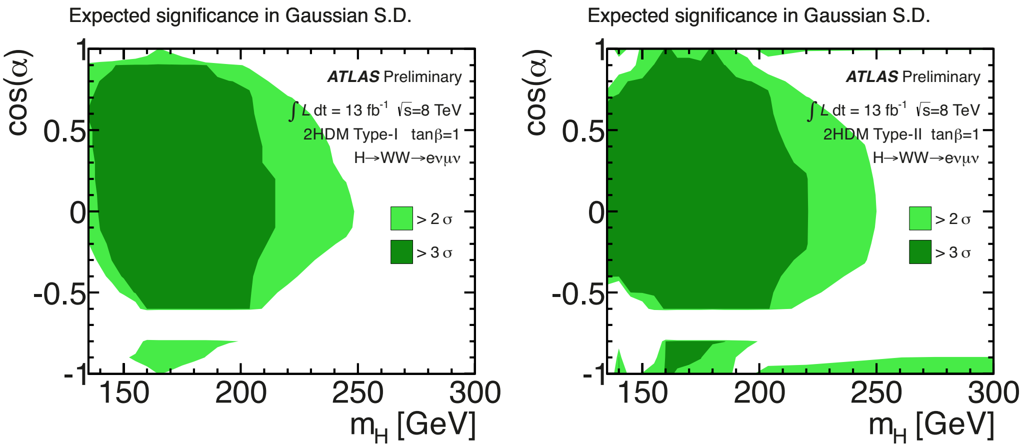


Figure 9.4: Exclusion contours in the $\cos\alpha$–m_H plane for type-II 2HDMs. The subfigures with the label ATLAS Preliminary are taken from [164].



(a) (b)

Figure 9.5: Contour plots of expected significance in Gaussian standard deviations for the type-I model (a) and the type-II model (b). The SM Higgs boson is included as part of the null hypothesis [164].

analysis reaches a sensitivity larger than 2σ for masses up to almost 250 GeV and a sensitivity larger than 3σ for masses slightly above 200 GeV.

Figure 9.6 and figure 9.7 show the calculated exclusion contours for the type-I and type-II 2HDM, respectively, when including the Higgs boson with a SM cross section at 125 GeV in the null hypothesis as an additional process. Differences between the expected and observed exclusion regions are apparent in the low mass region in figure 9.3 and figure 9.4, whereas in figure 9.6 and figure 9.7 these differences vanish because of the SM Higgs boson in the null hypothesis. In the mass region from 135 to 200 GeV, a large part of the $\cos\alpha$–m_H plane can be excluded, even when the SM Higgs boson is taken into account in the null hypothesis.

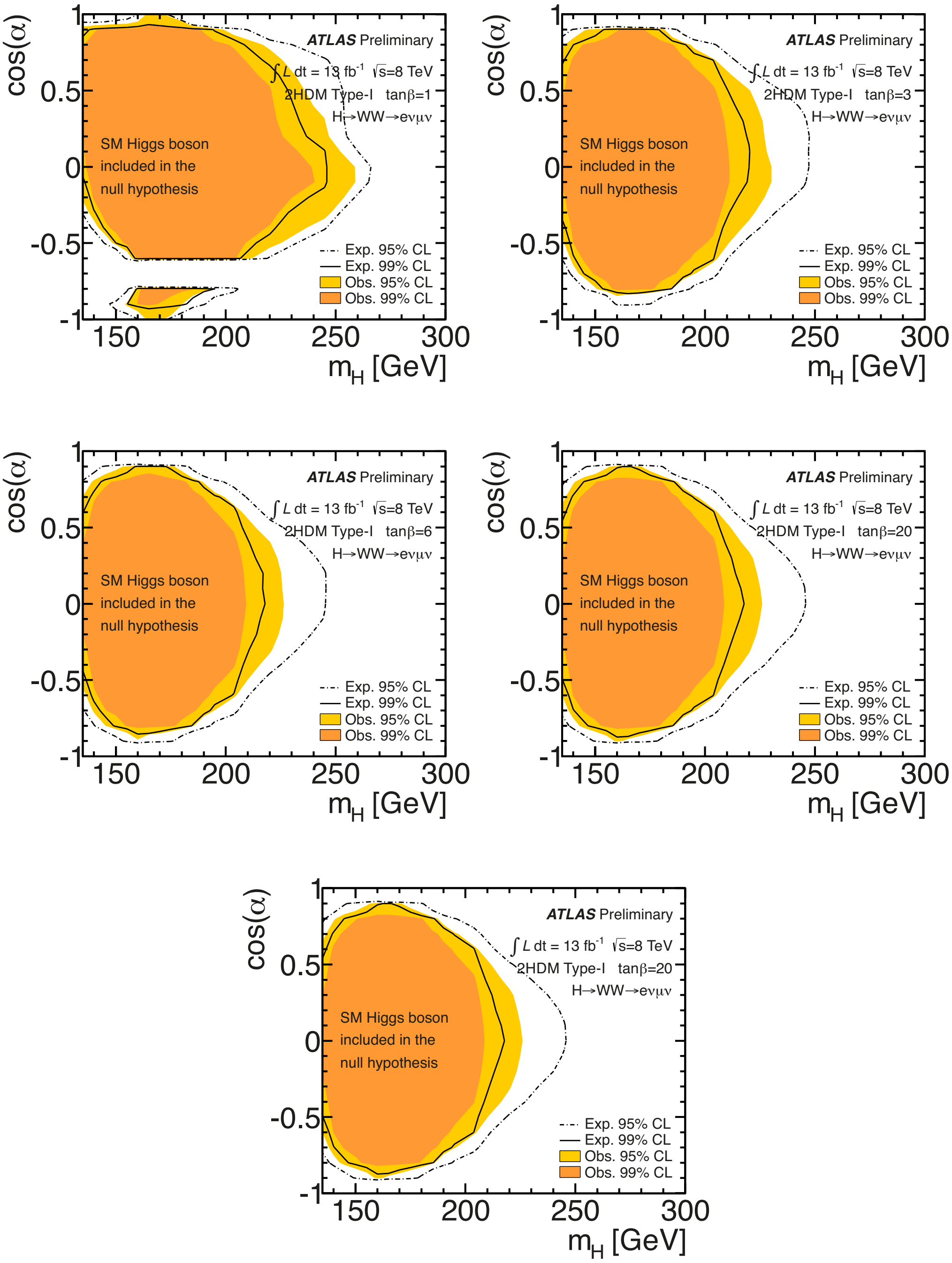


Figure 9.6: Exclusion contours in the $\cos\alpha$–m_H plane for type-I 2HDMs [164]. The SM Higgs is included as part of the null hypothesis.

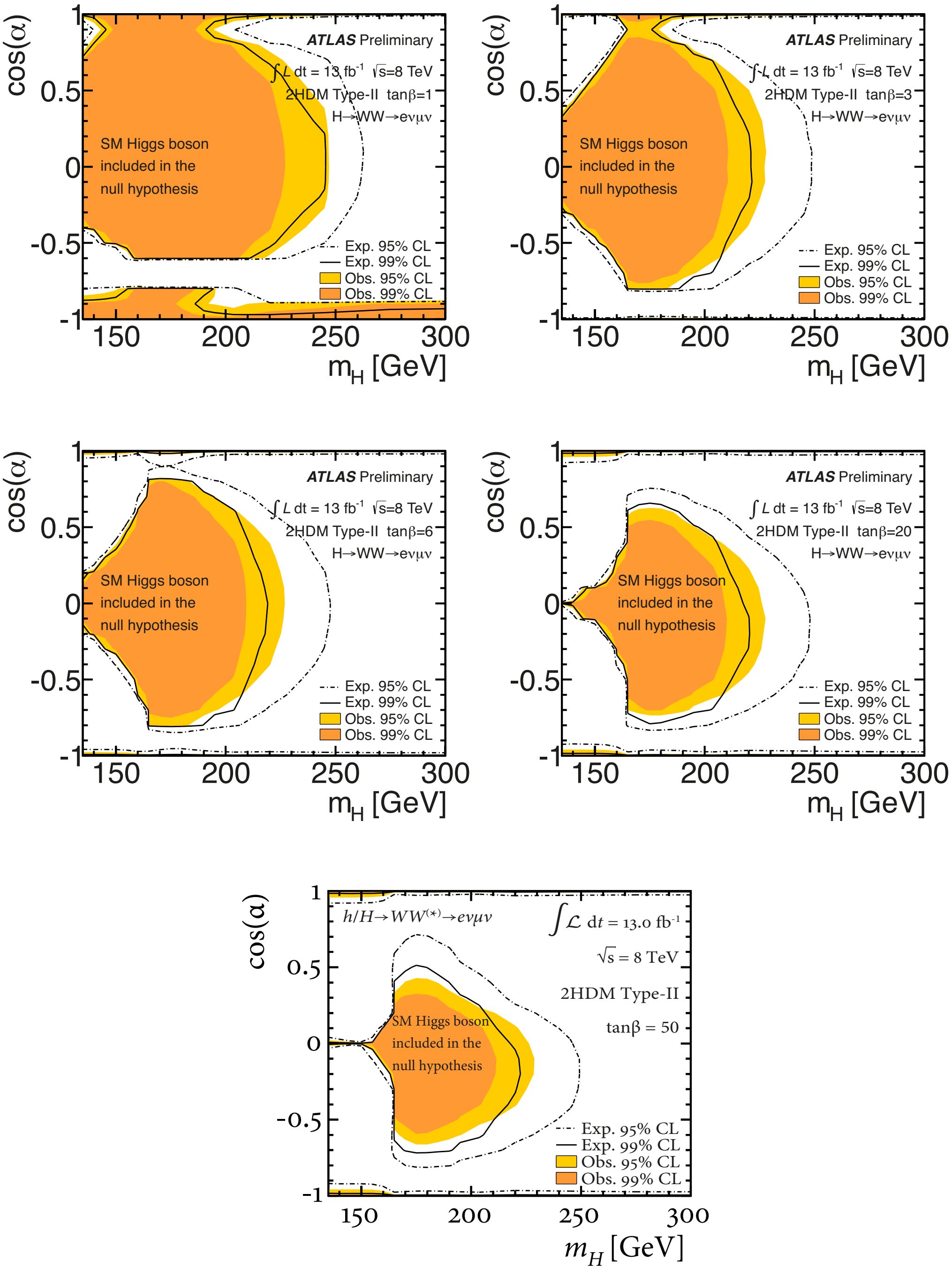


Figure 9.7: Exclusion contours in the $\cos\alpha$–m_H plane for type-II 2HDM. The SM Higgs is included as part of the null hypothesis. The subfigures with the label ATLAS Preliminary are taken from [164].

9.1 Fermiophobic Higgs Boson

The 2HDMs exhibit several interesting phenomenological limits, when considering specific parameter settings. One of these special parameter regions defines the fermiophobic limit, in which $\alpha = 0$ so that the heavy Higgs boson H has a zero coupling to fermions at tree level and becomes a fermiophobic Higgs boson (see table 2.4). In the fermiophobic model (FPM) the VBF cross sections are the same as in the SM, but the branching ratios differ, in particular in the low mass region. One consequence of the change of the branching ratios is that the $H \to WW^{(*)}$ decay mode outweighs all other decay modes in the entire mass range from 135 GeV to 300 GeV. The product of the cross section times branching ratio ranges from 1.25 pb at $m_H = 135$ GeV to 0.31 pb at $m_H = 300$ GeV.

In figure 9.8, the resulting expected and observed limits on the production cross section of a fermiophobic Higgs boson is shown as a function of m_H. The existence of a fermiophobic Higgs boson is excluded at the 95% CL in the mass range from 147.3 – 184.4 GeV, while the expected exclusion in the absence of a signal ranges from 135 – 194.8 GeV.

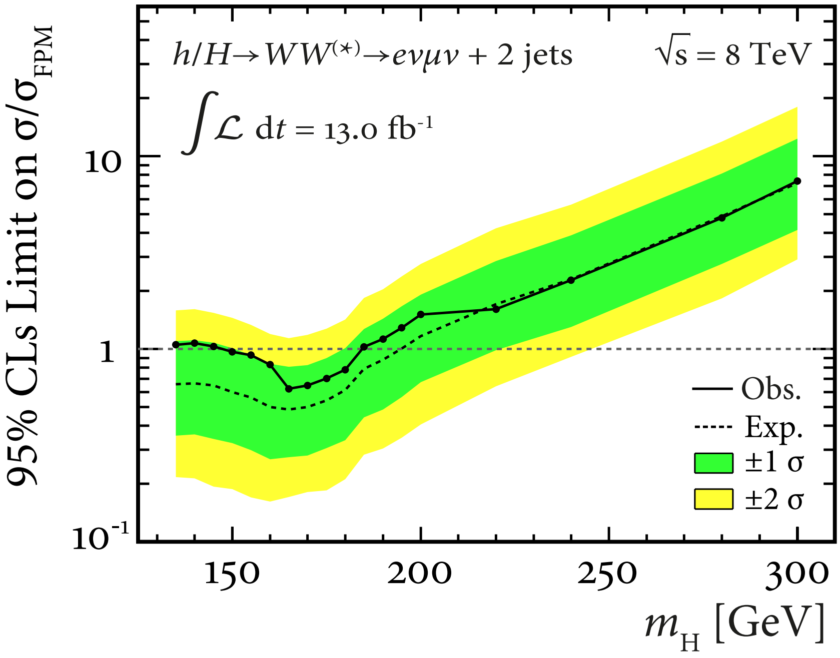


Figure 9.8: Expected and observed 95% CL upper limits on the production cross section of a fermiophobic Higgs boson as a function of m_H. The dashed and full black line show the expected and observed limit as a function of m_H, respectively. The green and yellow bands indicate the regions of $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty, respectively, in which the limit is expected to lie in the absence of a signal.

Chapter 10

Conclusion

The SM was tested over the last decades at particle accelerators with high precision. The Higgs boson was the last missing piece of the SM before the observation of a new elementary particle was announced on the 4th of July 2012. This new particle is yet known as a Higgs boson with a mass of approximately 125 GeV. To date, all measurements concerning the production rates, the branching ratios and kinematic distributions are compatible with this particle being the Higgs boson predicted by the SM. Alternative models, which are capable of explaining this Higgs boson, arise naturally by extending the Higgs sector with a second complex Higgs doublet. These 2HDMs are favoured models by theorists because of their simplicity and their capability to explain important problems in particle physics, including the origin of the observed asymmetry between baryons and antibaryons in the universe by new sources of CP symmetry violation.

The analysis presented in this thesis is the first direct search dedicated to exclude a heavy 2HDM Higgs boson in the context of the observation of a Higgs boson found at 125 GeV. This analysis hypothesises that the Higgs boson h observed by the LHC experiments at $m_h \approx 125$ GeV is part of a 2HDM. Following this hypothesis, a search for a second, heavier, CP-even scalar boson H, has been performed, including the new boson h as part of the 2HDM signal model and comparing it to a background-only model as well as to the SM including the SM Higgs boson. Both bosons of the 2HDMs are reconstructed in the $h/H \to WW^{(*)} \to e\nu\mu\nu$ decay channel and contribute to the signal rate. The collision data used in the analysis were recorded with the ATLAS detector at a centre-of-mass energy of 8 TeV and correspond to an integrated luminosity of 13.0 fb^{-1}. No evidence for an additional Higgs boson

H predicted by the type-I and type-II 2HDM is found in the investigated mass range of $135 < m_H < 300$ GeV. Exclusion contours at the 95% and 99% CL are determined in both 2HDMs in the $\cos\alpha$-m_H plane for different values of $\tan\beta$ ranging from 1 to 50. In the mass region from 135 to 200 GeV a large part of the $\cos\alpha$-m_H plane can be excluded for the type-I 2HDM as well as for the type-II 2HDM.

List of Figures

List of Tables

Bibliography

[1] L. Evans and P. Bryant. 'LHC Machine'. In: *JINST* 3 (2008) (cit. on p. 1).

[2] ATLAS Collaboration. 'The ATLAS Experiment at the CERN Large Hadron Collider'. In: *JINST* 3 (2008) (cit. on pp. 1, 27).

[3] CMS Collaboration. 'The CMS experiment at the CERN LHC'. In: *JINST* 3 (2008) (cit. on pp. 1, 27).

[4] ATLAS Collaboration. 'Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC'. In: *Phys. Lett.* B716 (2012), pp. 1–29. arXiv: `1207.7214 [hep-ex]` (cit. on pp. 1, 14).

[5] CMS Collaboration. 'Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC'. In: *Phys. Lett.* B716 (2012), pp. 30–61. arXiv: `1207.7235 [hep-ex]` (cit. on p. 1).

[6] J. E. Kim. 'Light pseudoscalars, particle physics and cosmology'. In: *Phys. Rept.* 150 (1987), pp. 1–177 (cit. on p. 1).

[7] C. Chen and S. Dawson. 'Exploring two Higgs doublet models through Higgs production'. In: *Phys. Rev. D* 87 (2013). arXiv: `1301.0309 [hep-ph]` (cit. on p. 2).

[8] J. Chang et al. 'Implications on the heavy CP-even Higgs boson from current Higgs data'. In: *Phys. Rev. D* 87 (2013). arXiv: `1211.3849 [hep-ph]` (cit. on p. 2).

[9] A. Drozd et al. 'Two-Higgs-doublet models and enhanced rates for a 125 GeV Higgs'. In: *JHEP* 2013 (2013), pp. 1–26 (cit. on p. 2).

[10] W. Altmannshofer, S. Gori and G. D. Kribs. 'A minimal-flavor-violating 2HDM at the LHC'. In: *Phys. Rev.* D86 (2012). arXiv: `1210.2465 [hep-ph]` (cit. on p. 2).

[11] D. S. M. Alves, P. J. Fox and N. J. Weiner. *Higgs signals in a type-I 2HDM or with a sister Higgs*. 2012. arXiv: `1207.5499 [hep-ph]` (cit. on p. 2).

[12] T. D. Lee. 'A theory of spontaneous T violation'. In: *Phys. Rev.* D8 (1973), pp. 1226–1239 (cit. on p. 2).

[13] P. M. Ferreira et al. 'Implications of the LHC two-photon signal for two-Higgs-doublet models'. In: *Phys. Rev.* D85 (2012). arXiv: `1112.3277 [hep-ph]` (cit. on p. 2).

[14] CDF Collaboration. 'Search for a two-Higgs-boson doublet using a simplified model in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV'. In: *Phys. Rev. Lett.* 110 (2013). arXiv: `1212.3837 [hep-ex]` (cit. on p. 2).

[15] CDF Collaboration. 'Search for Higgs Bosons Predicted in Two-Higgs-Doublet Models via Decays to Tau Lepton Pairs in 1.96 TeV $p\bar{p}$ Collisions'. In: *Phys. Rev. Lett.* 103 (2009). arXiv: `0906.1014 [hep-ex]` (cit. on p. 2).

[16] P. W. Higgs. 'Broken symmetries and the masses of gauge bosons'. In: *Phys. Rev. Lett.* 13 (1964), pp. 508–509 (cit. on p. 5).

[17] P. W. Higgs. 'Spontaneous Symmetry Breakdown without Massless Bosons'. In: *Phys. Rev.* 145 (1966), pp. 1156–1163 (cit. on p. 5).

[18] F. Englert and R. Brout. 'Broken symmetry and the mass of gauge vector mesons'. In: *Phys. Rev. Lett.* 13 (1964), pp. 321–322 (cit. on p. 5).

[19] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble. 'Global conservation laws and massless particles'. In: *Phys. Rev. Lett.* 13 (1964), pp. 585–587 (cit. on p. 5).

[20] T. W. B. Kibble. 'Symmetry breaking in nonAbelian gauge theories'. In: *Phys. Rev.* 155 (1967), pp. 1554–1561 (cit. on p. 5).

[21] ATLAS Collaboration. 'Evidence for the spin-0 nature of the Higgs boson using ATLAS data'. In: *Phys. Lett.* B726 (2013), pp. 120–144. arXiv: `1307.1432 [hep-ex]` (cit. on p. 6).

[22] CMS Collaboration. 'Study of the Mass and Spin-Parity of the Higgs Boson Candidate via Its Decays to Z Boson Pairs. On the mass and spin-parity of the Higgs boson candidate via its decays to Z boson pairs'. In: *Phys. Rev. Lett.* 110 (2012). arXiv: `1212.6639 [hep-ex]` (cit. on p. 6).

[23] ATLAS Collaboration. 'Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC'. In: *Phys. Lett.* B726 (2013), pp. 88–119. arXiv: `1307.1427 [hep-ex]` (cit. on p. 6).

[24] CMS Collaboration. *Measurement of the properties of a Higgs boson in the four-lepton final state*. 2013. arXiv: `1312.5353 [hep-ex]` (cit. on p. 6).

[25] CMS Collaboration. 'Measurement of Higgs boson production and properties in the WW decay channel with leptonic final states'. In: *JHEP* 1401 (2014). arXiv: `1312.1129 [hep-ex]` (cit. on p. 6).

[26] Wikipedia, The Free Encyclopedia. *Standard Model of Elementary Particles*. URL: `https://upload.wikimedia.org/wikipedia/commons/0/00/Standard_Model_of_Elementary_Particles.svg` (cit. on p. 7).

[27] Tevatron Electroweak Working Group, CDF Collaboration and D0 Collaboration. *Combination of CDF and D0 results on the mass of the top quark using up to* 8.7 fb^{-1} *at the Tevatron*. 2013. arXiv: `1305.3929 [hep-ex]` (cit. on p. 7).

[28] Particle Data Group Collaboration. 'Review of Particle Physics (RPP)'. In: *Phys. Rev.* D86 (2012) (cit. on pp. 7, 11).

[29] Wikipedia, The Free Encyclopedia. *Standard Model of Elementary Particles*. URL: `https://upload.wikimedia.org/wikipedia/commons/4/4c/Elementary_particle_interactions.svg` (cit. on p. 8).

[30] S. L. Glashow. 'Partial Symmetries of Weak Interactions'. In: *Nucl. Phys.* 22 (1961), pp. 579–588 (cit. on p. 10).

[31] A. Salam and J. C. Ward. 'Weak and electromagnetic interactions'. In: *Nuovo Cim.* 11 (1959), pp. 568–577 (cit. on p. 10).

[32] S. Weinberg. 'A Model of Leptons'. In: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266 (cit. on p. 10).

[33] T. Nakano. 'Charge Independence for V-particles'. In: *Progress of Theoretical Physics* 10 (1953), pp. 581–582 (cit. on p. 10).

[34] K. Nishijima. 'Charge Independence Theory of V Particles'. In: *Progress of Theoretical Physics* 13 (1955), pp. 285–304 (cit. on p. 10).

[35] M. Gell-Mann. 'The Interpretation of the New Particles as Displaced Charged Multiplets'. In: *Il Nuovo Cimento* 4 (1956) (cit. on p. 10).

[36] D. J. Gross and F. Wilczek. 'Asymptotically Free Gauge Theories. 1'. In: *Phys. Rev.* D8 (1973), pp. 3633–3652 (cit. on p. 11).

[37] D. J. Gross and F. Wilczek. 'Asymptottically Free Gauge Theories. 2'. In: *Phys. Rev.* D9 (1974), pp. 980–993 (cit. on p. 11).

[38] C. P. Burgess and G. D. Moore. *The standard model: A primer*. Cambridge University Press, 2007 (cit. on p. 11).

[39] W. N. Cottingham and D. A. Greenwood. *An introduction to the standard model of particle physics*. Cambridge University Press, 2007 (cit. on p. 11).

[40] M. K. Gaillard, P. D. Grannis and F. J. Sciulli. 'The standard model of particle physics'. In: *Rev. Mod. Phys.* 71.2 (1999) (cit. on p. 11).

[41] I. J. R. Aitchison. *Gauge Theories in Particle Physics*. Taylor and Francis, 2004 (cit. on p. 11).

[42] F. Wilczek. 'Quantum field theory'. In: *Rev. Mod. Phys.* 71 (1999), S85–S95. arXiv: `hep-th/9803075 [hep-th]` (cit. on p. 11).

[43] G. 't Hooft. 'Renormalization of Massless Yang-Mills Fields'. In: *Nucl. Phys.* B33 (1971), pp. 173–199 (cit. on p. 12).

[44] G. 't Hooft. 'Renormalizable Lagrangians for Massive Yang-Mills Fields'. In: *Nucl. Phys.* B35 (1971), pp. 167–188 (cit. on p. 12).

[45] A. Djouadi. 'The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model'. In: *Phys. Rept.* 457 (2008), pp. 1–216. arXiv: `hep-ph/0503172 [hep-ph]` (cit. on p. 14).

[46] LHC Higgs Cross Section Working Group. *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*. CERN, 2011. arXiv: `1101.0593 [hep-ph]` (cit. on pp. 16, 48, 110, 122).

[47] LHC Higgs Cross Section Working Group. *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*. CERN, 2013. arXiv: `1307.1347 [hep-ph]` (cit. on p. 14).

[48] LHC Higgs Cross Section Working Group. *Recommended values on SM Higgs cross sections and branching ratios*. URL: `http://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections` (cit. on p. 14).

[49] R. G. Sachs. 'CP Violation in K^0 Decays'. In: *Phys. Rev. Lett.* 13 (1964), pp. 286–288 (cit. on p. 19).

[50] A. D. Sakharov. 'Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe'. In: *Pisma Zh. Eksp. Teor. Fiz.* 5 (1967), pp. 32–35 (cit. on p. 19).

[51] M. Trodden. *Electroweak baryogenesis: A brief review*. 1998. arXiv: `hep-ph/9805252 [hep-ph]` (cit. on p. 19).

[52] ATLAS Collaboration. *Search for single top-quark production via FCNC in strong interaction in* $\sqrt{s} = 8$ TeV *ATLAS data*. ATLAS-CONF-2013-063. 2013 (cit. on p. 19).

[53] W. Wang. $b \to s\gamma$ *and* $b \to d\gamma$ *(B factories)*. 2011. arXiv: `1102.1925 [hep-ex]` (cit. on p. 19).

[54] S. L. Glashow and S. Weinberg. 'Natural Conservation Laws for Neutral Currents'. In: *Phys. Rev.* D15 (1977) (cit. on p. 20).

[55] E. A. Paschos. 'Diagonal Neutral Currents'. In: *Phys. Rev.* D15 (1977) (cit. on p. 20).

[56] G. C. Branco et al. 'Theory and phenomenology of two-Higgs-doublet models'. In: *Phys. Rept.* 516 (2012), pp. 1–102. arXiv: `1106.0034 [hep-ph]` (cit. on pp. 20–23).

[57] Y. L. Wu and L. Wolfenstein. 'Sources of CP violation in the two Higgs doublet model'. In: *Phys. Rev. Lett.* 73 (1994), pp. 1762–1764. arXiv: `hep-ph/9409421 [hep-ph]` (cit. on p. 22).

[58] N. Craig et al. 'Multi-Lepton Signals of Multiple Higgs Bosons'. In: *JHEP* 2013 (2013), pp. 1–39. arXiv: `1210.0559 [hep-ph]` (cit. on p. 24).

[59] R. V. Harlander, S. Liebler and H. Mantler. 'SusHi: A program for the calculation of Higgs production in gluon fusion and bottom-quark annihilation in the Standard Model and the MSSM'. In: *Comput. Phys. Commun.* 184 (2013), pp. 1605–1617. arXiv: `1212.3249 [hep-ph]` (cit. on pp. 23, 25–26, 48).

[60] R. V. Harlander and W. B. Kilgore. 'Next-to-next-to-leading order Higgs production at hadron colliders'. In: *Phys. Rev. Lett.* 88 (2002). arXiv: `hep-ph/0201206 [hep-ph]` (cit. on pp. 23, 48).

[61] C. Lefèvre. *The CERN accelerator complex. Complexe des accélérateurs du CERN*. 2008 (cit. on p. 28).

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

[63] LHCb Collaboration. 'The LHCb Detector at the LHC'. In: *JINST* 3 (2008) (cit. on p. 27).

[64] LHCf Collaboration. 'The LHCf detector at the CERN Large Hadron Collider'. In: *JINST* 3 (2008) (cit. on p. 27).

[65] MoEDAL Collaboration. *Technical Design Report of the MoEDAL Experiment*. 2009 (cit. on p. 27).

[66] TOTEM Collaboration. 'The TOTEM experiment at the CERN Large Hadron Collider'. In: *JINST* 3 (2008) (cit. on p. 27).

[67] C. E. Hill. *Ion and electron sources*. 1994 (cit. on p. 29).

[68] M. Benedikt et al. *LHC Design Report*. CERN, 2004 (cit. on p. 29).

[69] K. Schindl. *The PS Booster as Pre-Injector for LHC*. 1997 (cit. on p. 29).

[70] R. Cappi. *The PS in the LHC injector chain*. 1997 (cit. on p. 29).

[71] T. P. R. Linnecar. *Preparing the SPS for LHC*. 1997 (cit. on p. 29).

[72] O. S. Brüning et al. *LHC Design Report*. CERN, 2004 (cit. on p. 29).

[73] M. Lamont. 'Status of the LHC'. In: *Journal of Physics: Conference Series* 455.1 (2013) (cit. on p. 30).

[74] D. Denisov. *Moriond QCD 2013 Experimental Summary*. 2013. arXiv: `1306.6908 [hep-ex]` (cit. on p. 30).

[75] ATLAS Collaboration. *ATLAS inner detector: Technical design report. Vol. 1.* 1997 (cit. on p. 31).

[76] ATLAS Collaboration. *ATLAS inner detector: Technical design report. Vol. 2.* 1997 (cit. on p. 31).

[77] ATLAS Collaboration. 'Ultimate Performance of the ATLAS Superconducting Solenoid'. In: *IEEE Trans. Appl. Supercond.* (2007) (cit. on p. 34).

[78] ATLAS Collaboration. 'ATLAS pixel detector electronics and sensors'. In: *JINST* 3 (2008) (cit. on p. 34).

[79] ATLAS Collaboration. 'The silicon microstrip sensors of the ATLAS semiconductor tracker'. In: *Instrum. Meth.* (2008) (cit. on p. 34).

[80] ATLAS Collaboration. 'The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: design and performance'. In: *JINST* 3 (2008) (cit. on p. 34).

[81] ATLAS Collaboration. *ATLAS calorimeter performance: Technical Design Report*. Technical Design Report ATLAS. CERN, 1996 (cit. on p. 35).

[82] ATLAS Collaboration. *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN, 1996 (cit. on p. 35).

[83] ATLAS Collaboration. *ATLAS tile calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN, 1996 (cit. on p. 35).

[84] ATLAS Collaboration. 'Construction, assembly and tests of the ATLAS electromagnetic barrel calorimeter'. In: *Nucl. Instrum. Meth.* A (2006), pp. 558–388 (cit. on p. 35).

[85] ATLAS Collaboration. 'The ATLAS hadronic tile calorimeter: from construction toward physics'. In: *IEEE Trans. Nucl. Sci.* (2006) (cit. on p. 35).

[86] ATLAS Collaboration. *ATLAS muon spectrometer: Technical Design Report*. Technical Design Report ATLAS. CERN, 1997 (cit. on p. 36).

[87] ATLAS Collaboration. *ATLAS first level trigger: Technical design report.* 1998 (cit. on p. 37).

[88] ATLAS Collaboration. *ATLAS high-level trigger, data acquisition and controls: Technical design report*. 2003 (cit. on p. 37).

[89] ATLAS Collaboration. *The baseline dataflow system of the ATLAS trigger and DAQ*. 2003 (cit. on p. 37).

[90] M. Landwehr. *Monte Carlo generator for the LHC*. URL: `http://www.bw-grid.de/uploads/pics/hd_schumann_fig.jpg` (cit. on p. 41).

[91] H. L. Lai et al. 'New parton distributions for collider physics'. In: *Phys. Rev.* D82 (2010). arXiv: `1007.2241 [hep-ph]` (cit. on pp. 42, 107–108).

[92] G. Altarelli and G. Parisi. 'Asymptotic Freedom in Parton Language'. In: *Nucl. Phys.* B126 (1977) (cit. on p. 40).

[93] J. C. Collins, D. E. Soper and G. F. Sterman. 'Factorization of Hard Processes in QCD'. In: *Adv. Ser. Direct. High Energy Phys.* 5 (1988), pp. 1–91. arXiv: `hep-ph/0409313 [hep-ph]` (cit. on p. 42).

[94] CTEQ Collaboration. 'Handbook of perturbative QCD'. In: *Rev. Mod. Phys.* 67 (1995), pp. 157–248 (cit. on p. 42).

[95] D. Amati and G. Veneziano. 'Preconfinement as a property of perturbative {QCD}'. In: *Physics Letters B* 83.1 (1979), pp. 87–92 (cit. on p. 42).

[96] M. L. Mangano et al. 'ALPGEN, a generator for hard multi-parton processes in hadronic collisions'. In: *JHEP* 0307 (2003) (cit. on p. 43).

[97] S. Catani et al. 'QCD matrix elements + parton showers'. In: *JHEP* 0111 (2001). arXiv: `hep-ph/0109231 [hep-ph]` (cit. on p. 43).

[98] F. Krauss. 'Matrix elements and parton showers in hadronic interactions'. In: *JHEP* 0208 (2002). arXiv: `hep-ph/0205283 [hep-ph]` (cit. on p. 43).

[99] M. L. Mangano, M. Moretti and R. Pittau. 'Multijet matrix elements and shower evolution in hadronic collisions: $Wb\bar{b} + n$ jets as a case study'. In: *Nucl. Phys.* B632 (2002), pp. 343–362. arXiv: `hep-ph/0108069 [hep-ph]` (cit. on p. 43).

[100] S. Höche et al. 'QCD matrix elements and truncated showers'. In: *JHEP* 0905 (2009). arXiv: `0903.1219 [hep-ph]` (cit. on p. 43).

[101] J. Alwall et al. 'Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions'. In: *Eur. Phys. J. C* C53 (2008), pp. 473–500 (cit. on p. 43).

[102] T. Binoth et al. 'Gluon-induced W-boson pair production at the LHC'. In: *JHEP* 12 (2006). arXiv: `hep-ph/0611170v1` (cit. on pp. 43, 49).

[103] J. Alwall et al. 'MadGraph/MadEvent v4: The new web generation'. In: *JHEP* 0709 (2007). arXiv: `0706.2334 [hep-ph]` (cit. on p. 45).

[104] J. Alwall et al. 'MadGraph 5 : Going Beyond'. In: *JHEP* 06 (2011). arXiv: `1106.0522 [hep-ph]` (cit. on p. 45).

[105] R. C. Gray et al. *Backgrounds To Higgs boson searches from* $W\gamma^* \to l\nu l(l)$ *asymmetric internal conversion*. 2011. arXiv: `1110.1368 [hep-ph]` (cit. on p. 45).

[106] S. Frixione and B. Webber. 'Matching NLO QCD and parton showers in heavy-flavor production'. In: *JHEP* 0308 (2003) (cit. on p. 45).

[107] T. Kinoshita. 'Mass singularities of Feynman amplitudes'. In: *J.Math.Phys.* 3 (1962), pp. 650–677 (cit. on p. 45).

[108] T. D. Lee and M. Nauenberg. 'Degenerate Systems and Mass Singularities'. In: *Phys. Rev.* 133 (1964), B1549–B1562 (cit. on p. 45).

[109] S. Frixione, P. Nason and C. Oleari. 'Matching NLO QCD computations with Parton Shower simulations: the POWHEG method'. In: *JHEP* 0711 (2007). arXiv: `0709.2092 [hep-ph]` (cit. on p. 45).

[110] B. P. Kersevan and E. Richter-Was. 'The Monte Carlo event generator AcerMC versions 2.0 to 3.8 with interfaces to PYTHIA 6.4, HERWIG 6.5 and ARIADNE 4.1 '. In: *Comput. Phys. Commun.* 184 (2013), pp. 919–985. arXiv: `hep-ph/0405247 [hep-ph]` (cit. on p. 46).

[111] E. Byckling and K. Kajantie. 'N-particle phase space in terms of invariant momentum transfers'. In: *Nucl. Phys.* B9 (1969), pp. 568–576 (cit. on p. 46).

[112] G. P. Lepage. 'A new algorithm for adaptive multidimensional integration'. In: *Journal of Computational Physics* 27.2 (1978), pp. 192–203 (cit. on p. 46).

[113] B. P. Kersevan and I. Hinchliffe. 'A Consistent prescription for the production involving massive quarks in hadron collisions'. In: *JHEP* 0609 (2006). arXiv: `hep-ph/0603068 [hep-ph]` (cit. on p. 46).

[114] M. A. G. Aivazis, F. I. Olness and W. Tung. 'Leptoproduction of heavy quarks. 1. General formalism and kinematics of charged current and neutral current production processes'. In: *Phys. Rev.* D50 (1994), pp. 3085–3101. arXiv: `hep-ph/9312318 [hep-ph]` (cit. on p. 46).

[115] M. A. G. Aivazis et al. 'Leptoproduction of heavy quarks. 2. A Unified QCD formulation of charged and neutral current processes from fixed target to collider energies'. In: *Phys. Rev.* D50 (1994), pp. 3102–3118. arXiv: `hep-ph/9312319 [hep-ph]` (cit. on p. 46).

[116] S. Alioli et al. 'NLO Higgs boson production via gluon fusion matched with shower in POWHEG'. In: *JHEP* 0904 (2009). arXiv: `0812.0578 [hep-ph]` (cit. on p. 46).

[117] P. Nason and C. Oleari. 'NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG'. In: *JHEP* 1002 (2010). arXiv: `0911.5299 [hep-ph]` (cit. on p. 46).

[118] G. Corcella et al. 'HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including super-symmetric processes)'. In: *JHEP* 0101 (2001) (cit. on p. 46).

[119] J. M. Butterworth, J. R. Forshaw and M. H. Seymour. 'Multiparton interactions in photoproduction at HERA'. In: *Z. Phys.* C72 (1996), pp. 637–646. arXiv: `hep-ph/9601371 [hep-ph]` (cit. on pp. 47–48).

[120] T. Sjostrand, S. Mrenna and P. Z. Skands. 'PYTHIA 6.4 physics and manual'. In: *JHEP* 0605 (2006) (cit. on p. 47).

[121] C. Anastasiou and K. Melnikov. 'Higgs boson production at hadron colliders in NNLO QCD'. In: *Nucl. Phys.* B646 (2002), pp. 220–256. arXiv: `hep-ph/0207004 [hep-ph]` (cit. on p. 48).

[122] V. Ravindran, J. Smith and W. L. van Neerven. 'NNLO corrections to the total cross-section for Higgs boson production in hadron hadron collisions'. In: *Nucl. Phys.* B665 (2003), pp. 325–366. arXiv: `hep-ph/0302135 [hep-ph]` (cit. on p. 48).

[123] U. Aglietti et al. 'Two loop light fermion contribution to Higgs production and decays'. In: *Phys. Lett.* B595 (2004), pp. 432–441. arXiv: `hep-ph/0404071 [hep-ph]` (cit. on p. 48).

[124] S. Actis et al. 'NLO electroweak corrections to Higgs boson production at hadron colliders'. In: *Phys. Lett.* B670 (2008), pp. 12–17. arXiv: `0809.1301 [hep-ph]` (cit. on p. 48).

[125] S. Catani et al. 'Soft gluon resummation for Higgs boson production at hadron colliders'. In: *JHEP* 0307 (2003). arXiv: `hep-ph/0306211 [hep-ph]` (cit. on p. 48).

[126] G. Degrassi, S. Di Vita and P. Slavich. 'On the NLO QCD corrections to the production of the heaviest neutral Higgs scalar in the MSSM'. In: *Eur. Phys. J. C* C72 (2012). arXiv: `1204.1016 [hep-ph]` (cit. on p. 48).

[127] M. Ciccolini, A. Denner and S. Dittmaier. 'Strong and electroweak corrections to the production of Higgs + 2 jets via weak interactions at the LHC'. In: *Phys. Rev. Lett.* 99 (2007). arXiv: `0707.0381 [hep-ph]` (cit. on p. 48).

[128] M. Ciccolini, A. Denner and S. Dittmaier. 'Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC'. In: *Phys. Rev.* D77 (2008). arXiv: `0710.4749 [hep-ph]` (cit. on p. 48).

[129] K. Arnold et al. 'VBFNLO: A parton level Monte Carlo for processes with electroweak bosons'. In: *Comput. Phys. Commun.* 180 (2009), pp. 1661–1670. arXiv: `0811.4559 [hep-ph]` (cit. on p. 48).

[130] P. Bolzoni et al. 'Higgs production via vector-boson fusion at NNLO in QCD'. In: *Phys. Rev. Lett.* 105 (2010). arXiv: `1003.4451 [hep-ph]` (cit. on p. 48).

[131] O. Brein, A. Djouadi and R. V. Harlander. 'NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders'. In: *Phys. Lett.* B579 (2004), pp. 149–156. arXiv: `hep-ph/0307206 [hep-ph]` (cit. on p. 48).

[132] M. L. Ciccolini, S. Dittmaier and M. Kramer. 'Electroweak radiative corrections to associated WH and ZH production at hadron colliders'. In: *Phys. Rev.* D68 (2003). arXiv: `hep-ph/0306234 [hep-ph]` (cit. on p. 48).

[133] LHC Higgs Cross Section Working Group. *Handbook of LHC Higgs cross sections: 2. Differential distributions*. CERN, 2012. arXiv: `1201.3084 [hep-ph]` (cit. on pp. 48, 110).

[134] M. Cacciari et al. 'Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation'. In: *Phys. Lett.* B710 (2012), pp. 612–622. arXiv: `1111.5869 [hep-ph]` (cit. on p. 49).

[135] P. Bärnreuther, M. Czakon and A. Mitov. 'Percent Level Precision Physics at the Tevatron: First Genuine NNLO QCD Corrections to $q\bar{q} \to t\bar{t} + X$'. In: *Phys. Rev. Lett.* 109 (2012). arXiv: `1204.5201 [hep-ph]` (cit. on p. 49).

[136] M. Czakon and A. Mitov. 'NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels'. In: *JHEP* 1212 (2012). arXiv: `1207.0236 [hep-ph]` (cit. on p. 49).

[137] M. Czakon and A. Mitov. 'NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction'. In: *JHEP* 1301 (2013). arXiv: `1210.6832 [hep-ph]` (cit. on p. 49).

[138] M. Czakon, P. Fiedler and A. Mitov. 'The total top quark pair production cross-section at hadron colliders through $O(\alpha_S^4)$'. In: *Phys. Rev. Lett.* 110 (2013). arXiv: `1303.6254 [hep-ph]` (cit. on p. 49).

[139] ATLAS Collaboration. *Update of the* $H \to WW^{(*)} \to e\nu\mu\nu$ *Analysis with 13* fb^{-1} *of* $\sqrt{s} = 8$ *TeV Data Collected with the ATLAS Detector*. ATLAS-CONF-2012-158. 2012 (cit. on pp. 49, 67, 70, 73–74, 91, 119–120).

[140] ATLAS Collaboration. 'The ATLAS Simulation Infrastructure'. In: *Eur. Phys. J. C* C70 (2010), pp. 823–874. arXiv: `1005.4568 [hep-ex]` (cit. on p. 49).

[141] S. Agostinelli et al. 'GEANT 4, A Simulation Toolkit'. In: *Nucl. Instrum. Meth.* A506 (2003), pp. 250–303 (cit. on p. 49).

[142] ATLAS Collaboration. 'The ATLAS experiment at the CERN Large Hadron Collider'. In: *JINST* 3 (2008) (cit. on p. 50).

[143] ATLAS Collaboration. *Performance of the Electron and Photon Trigger in p-p Collisions at* $\sqrt{s} = 7$ TeV. ATLAS-CONF-2011-114. 2011 (cit. on p. 51).

[144] 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. 2012 (cit. on p. 51).

[145] ATLAS Collaboration. *Performance of the ATLAS Muon Trigger in p-p collisions at $\sqrt{s} = 7$ TeV*. ATLAS-CONF-2010-095. 2010 (cit. on p. 51).

[146] ATLAS Collaboration. *Performance of the ATLAS muon trigger in 2011*. ATLAS-CONF-2012-099. 2012 (cit. on p. 51).

[147] ATLAS Collaboration. *Multiple Year Collision Plots* (cit. on p. 53).

[148] ATLAS Collaboration. 'Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector at the LHC'. In: *Eur. Phys. J. C* C71 (2011). arXiv: `1101.2185` (cit. on p. 53).

[149] ATLAS Collaboration. *Calorimeter Clustering Algorithms: Description and Performance*. ATL-LARG-PUB-2008-002. 2008 (cit. on pp. 54, 56).

[150] ATLAS Collaboration. 'Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV'. In: *Eur. Phys. J. C* 73 (2013). arXiv: `1112.6426 [hep-ex]` (cit. on pp. 54, 63, 105).

[151] R. E. Kalman. 'A new approach to linear filtering and prediction problems'. In: *Transactions of the ASME – Journal of Basic Engineering, Series D* 82 (1960), pp. 35–45 (cit. on p. 54).

[152] P. Billoir and S. Qian. 'Fast vertex fitting with a local parametrization of tracks'. In: *Nucl. Instrum. Meth.* A311 (1992), pp. 139–150 (cit. on p. 55).

[153] ATLAS Collaboration. 'Primary vertex reconstruction in the ATLAS experiment at LHC'. In: *J. Phys.: Conf. Ser.* 119 (2008) (cit. on p. 55).

[154] ATLAS Collaboration. 'A framework for vertex reconstruction in the ATLAS experiment at LHC'. In: *J. Phys.: Conf. Ser.* (2010) (cit. on p. 55).

[155] ATLAS Collaboration. 'Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data'. In: *Eur. Phys. J. C* C72 (2012). arXiv: `1110.3174 [hep-ex]` (cit. on pp. 59, 107).

[156] ATLAS Collaboration. *Measurements of the photon identification efficiency with the ATLAS detector using 4.9 fb−1 of pp collision data collected in 2011*. ATLAS-CONF-2012-123. 2012 (cit. on pp. 59, 107).

[157] ATLAS Collaboration. *Preliminary results on the muon reconstruction efficiency, momentum resolution, and momentum scale in ATLAS 2012 pp collision data*. ATLAS-CONF-2013-088. 2013 (cit. on pp. 59, 61, 107).

[158] ATLAS Collaboration. *Electron efficiency measurements for 2012 and 2011 data* (cit. on p. 60).

[159] M. Cacciari, G. P. Salam and G. Soyez. 'The anti-k_t jet clustering algorithm'. In: *JHEP* 0804 (2008). arXiv: `0802.1189 [hep-ph]` (cit. on p. 61).

[160] ATLAS Collaboration. *Update on the jet energy scale systematic uncertainty for jets produced in proton-proton collisions at $\sqrt{s} = 7$ TeV measured with the ATLAS detector*. ATLAS-CONF-2011-007. 2011 (cit. on p. 63).

[161] V. M. Abazov et al. 'Measurement of the $p\bar{p} \to t\bar{t}$ production cross section at $\sqrt{s} = 1.96$ TeV in the fully hadronic decay channel.' In: *Phys. Rev.* D76 (2007). arXiv: `hep-ex/0612040 [hep-ex]` (cit. on p. 64).

[162] ATLAS Collaboration. *Measurement of the b-tag efficiency in a sample of jets containing muons with 5 fb^{-1} of data from the ATLAS detector*. ATLAS-CONF-2012-043. 2012 (cit. on p. 64).

[163] ATLAS Collaboration. *Performance of the missing transverse energy reconstruction and calibration in proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 7$ TeV with the ATLAS detector*. ATLAS-CONF-2010-057. 2010 (cit. on p. 66).

[164] ATLAS Collaboration. *Search for Higgs bosons in Two-Higgs-Doublet models in the $H \to WW \to e\nu\mu\nu$ channel with the ATLAS detector*. ATLAS-CONF-2013-027. 2013 (cit. on pp. 69, 88–89, 95, 97–100, 125–129).

[165] ATLAS Collaboration. 'Measurement of the WW cross section in $\sqrt{s} = 7$ TeV pp collisions with ATLAS'. In: *Phys. Rev. Lett.* 107 (2011). arXiv: `1104.5225 [hep-ex]` (cit. on p. 70).

[166] D. Casadei and G. Choudalakis. 'Plotting the differences between data and expectation'. In: *Eur. Phys. J. Plus* 127 (2012) (cit. on pp. 75–76, 87).

[167] ATLAS Collaboration. *ATLANTIS, Event Display for ATLAS*. URL: `http://www.cern.ch/atlantis` (cit. on p. 78).

[168] M. Feindt and U. Kerzel. 'The NeuroBayes neural network package'. In: *Nucl. Instrum. Meth. A* 559 (2006), pp. 190–194 (cit. on p. 83).

[169] ATLAS Collaboration. 'Measurement of the t-channel single top-quark production cross section in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector'. In: *Phys. Lett.* B717 (2012), pp. 330–350. arXiv: `1205.3130 [hep-ex]` (cit. on pp. 83–84).

[170] ATLAS Collaboration. *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*. ATLAS-CONF-2011-102. 2011 (cit. on p. 105).

[171] ATLAS Collaboration. *Improved luminosity determination in pp collisions at* $\sqrt{s}$ = 7 TeV *using the ATLAS detector at the LHC*. 2013. arXiv: `1302.4393 [hep-ex]` (cit. on p. 107).

[172] A. D. Martin et al. 'Parton distributions for the LHC'. In: *Eur. Phys. J. C* C63 (2009), pp. 189–285. arXiv: `0901.0002` (cit. on pp. 107–108).

[173] R. D. Ball et al. 'Impact of heavy-quark masses on parton distributions and LHC phenomenology'. In: *Nucl. Phys.* B849 (2011), pp. 296–363. arXiv: `1101.1300` (cit. on pp. 107–108).

[174] J. M. Campbell, J. W. Huston and W. J. Stirling. 'Hard interactions of quarks and gluons: A primer for LHC physics'. In: *Rept. Prog. Phys.* 70 (2007), pp. 89–193. arXiv: `hep-ph/0611148` (cit. on p. 107).

[175] G. Sartisohn. 'Higgs Boson Search in the $H \to WW^{(*)} \to \ell\nu\ell\nu$ Channel using Neural Networks with the ATLAS Detector at 7 TeV'. CERN-THESIS-2012-072. PhD thesis. Universität Wuppertal, 2012 (cit. on p. 116).

[176] T. Junk. 'Confidence level computation for combining searches with small statistics'. In: *Nucl. Instrum. Methods* A434 (1999), pp. 435–443 (cit. on p. 117).

[177] A. L. Read. 'Presentation of search results: the $\mathrm{CL_s}$ technique'. In: *J. Phys. G: Nucl. Part. Phys.* 28 (2002), pp. 2693–2704 (cit. on p. 117).

[178] ATLAS Collaboration. *Combined measurements of the mass and signal strength of the Higgs-like boson with the ATLAS detector using up to 25 fb*$^{-1}$ *of proton-proton collision data*. ATLAS-CONF-2013-014. 2013 (cit. on p. 124).

[179] ATLAS Collaboration. *Combined coupling measurements of the Higgs-like boson with the ATLAS detector using up to 25 fb*$^{-1}$ *of proton-proton collision data*. ATLAS-CONF-2013-034. 2013 (cit. on p. 124).

[180] ATLAS Collaboration. *Study of the spin of the new boson with up to* 25 fb^{-1} *of ATLAS data*. ATLAS-CONF-2013-040. 2013 (cit. on p. 124).

[181] CMS Collaboration. *Properties of the observed Higgs-like resonance using the diphoton channel*. CMS-PAS-HIG-13-016. 2013 (cit. on p. 124).

[182] CMS Collaboration. *Search for a standard model like Higgs boson in the* $H \to ZZ \to 2\ell 2q$ *decay channel at* $\sqrt{s} = 8$ TeV. CMS-PAS-HIG-12-024. 2013 (cit. on p. 124).

[183] CMS Collaboration. *Update of the search for the Standard Model Higgs boson decaying into WW in the vector boson fusion production channel*. CMS-PAS-HIG-13-022. 2013 (cit. on p. 124).

Danksagung

Mein ganz besonderer Dank gilt Professor Dr. Wolfgang Wagner für die freundliche Aufnahme in die Arbeitsgruppe und die sehr gute Betreuung. Die Zusammenarbeit an dem überaus interessanten Thema war für mich stets motivierend und inspirierend. Ich habe besonders die vielfältigen Möglichkeiten genossen mein eigenes Wissen zu erweitern und den Verlauf der Analyse zu gestalten.

I gratefully acknowledge Professor Dr. Marumi Kado for his support of the 2HDM analysis during the review of the ATLAS conference note and his kind agreement to be the second referee of my thesis.

Ich möchte mich ganz herzlich bei Dr. Dominic Hirschbühl bedanken. Seine zu jeder Tageszeit präsente Hilfsbereitschaft und Unterstützung, sowie die sehr gute Zusammenarbeit und der niemals endende Erfahrungsschatz waren für diese Arbeit essentiell.

Mein Dank gilt außerdem allen Mitgliedern der Arbeitsgruppe, insbesondere Dr. Georg Sartisohn für die Einarbeitung in das Thema und Gunar Ernis für seine Unterstüzung bei der Analyse. Des Weiteren danke ich Kathrin Becker, Phillipp Tepel und Philipp Sturm für ihren Beistand, ihre Motivation und die interessanten Gespräche. Außerdem möchte ich mich bei Julia Fischer und Christian Riegel für die gute Stimmung in unserem Büro und ihren vielen guten Ratschlägen bedanken. Beim Sekretariat und den Computing-Administratoren bedanke ich mich für die gute Zusammenarbeit.

Meinen Eltern, die mir mein Studium und meine Promotion überhaupt erst ermöglicht haben, möchte ich für ihre langjährige Unterstützung danken.