Test of Spin and Parity of the Higgs Boson in the $H \rightarrow WW^* \rightarrow e\nu\mu\nu$ Decay Channel with the ATLAS Detector at the LHC

Dissertation
zur Erlangung des Grades

„Doktor der Naturwissenschaften“

am Fachbereich Physik, Mathematik und Informatik
der Johannes Gutenberg-Universität
in Mainz

Johannes Mattmann
geboren in Bad Kreuznach

Mainz, den 26. Februar 2016

Datum der mündlichen Prüfung: 11.11.2016

Abstract

The Higgs boson takes a key position within the Standard Model and its underlying mechanism is crucial in order to explain why elementary particles and especially the heavy vector bosons have masses. Since its prediction in the early 1960s it took nearly 50 years until the existence of the Higgs boson could be proven. It was on 4th July 2012 that two experiments—ATLAS ("A Toroidal LHC ApparatuS") and CMS ("Compact Muon Solenoid")—at the Large Hadron Collider at CERN, the European Organization for Nuclear research, announced a $5\,\sigma$ excess verifying the existence of the Higgs boson. Its mass amounts about $125\,\mathrm{GeV}/c^2$ making the observation compatible with earlier exclusion limits.

In January 2012 by the time the work on this thesis started, the confirmation of the Higgs boson existence was anticipated and further steps based on this hypothesis were contemplated. For this reason an analysis of the Higgs boson quantum numbers for spin and CP eigenvalue was initiated. These quantities are crucial to verify the agreement between the later discovered resonance and the Standard Model expectation. For the analysis the decay channel $H \to WW^* \to e\nu\mu\nu$ was chosen, which features a high cross section times branching ratio in the anticipated mass regime. In spite of the two neutrinos in the final state the signature of the final state is clearly selectable and can rely on the precise missing transverse momentum reconstruction of the ATLAS detector.

The contributions to the above publications comprise studies and development work for a wide range of aspects of the published analysis such as the definition of sensitive and background-enriched regions in the measured variable space, studies of multivariate techniques to optimise the sensitivity, support of the reweighting studies for the CP-mixing analysis, implementation and impact estimation of systematic uncertainties and further fields.

The major focus though was set on Monte Carlo event generator comparison studies and the statistical evaluation to obtain final results taking background processes and systematic uncertainties into account. A unique feature of the thesis is the evaluation of a sophisticated automated "smart" binning of two-dimensional qualifier outputs from the applied multivariate analysis techniques.

First exclusion limits for non-Standard-Model spin and CP properties of the Higgs boson could be set already in 2013 and further improved in 2015. The analysis presented in this thesis can exclude all tested spin-2 models at a confidence level of 82.5 to 99.1 %. Two non-Standard-Model CP scenarios (CP-odd and CP-even with higher dimensional couplings to Standard Model particles) can be excluded respectively at 97.1 and 64.5 % confidence level. Combining the results from all bosonic decay channels at ATLAS each tested spin-2 and CP scenario can be excluded at more than 99.9 % confidence level.

The newly introduced "smart" binning procedure is shown to be working reliably and takes a stand as a promising tool for future high energy physics analyses.

“Human passions have mysterious ways, in children as well as grown-ups. Those affected by them can’t explain them, and those who haven’t known them have no understanding of them at all.”
- Michael Ende, The Neverending Story

Contents

I. General Introduction and Explanation 11

1. Introduction 13

2. Theoretical Foundations 15
  2.1. Standard Model of Particle Physics . . . 15
    2.1.1. Particles of the Standard Model . . . 15
    2.1.2. Symmetries . . . 16
    2.1.3. Interactions in the Standard Model . . . 19
    2.1.4. Higgs Mechanism . . . 20

3. Experimental Apparatus and Preconditions 25
  3.1. Large Hadron Collider . . . 25
  3.2. ATLAS . . . 28
    3.2.1. Beam Pipe . . . 30
    3.2.2. Inner Detector . . . 30
    3.2.3. Magnet System . . . 31
    3.2.4. Calorimeter System . . . 32
    3.2.5. Muon System . . . 32
    3.2.6. Trigger System . . . 33
    3.2.7. Luminosity . . . 33
  3.3. Event Reconstruction . . . 34

4. Phenomenology 37
  4.1. Modelling of Particle Interactions . . . 37
    4.1.1. Interaction Matrix Element . . . 37
    4.1.2. Parton Distribution Functions . . . 39
    4.1.3. Showering and Hadronization . . . 40
    4.1.4. Detector Simulation . . . 42
  4.2. Monte Carlo Event Generators . . . 43
  4.3. Higgs Phenomenology . . . 45
    4.3.1. Production and Decay Modes . . . 45
    4.3.2. Higgs Discovery and Measurements . . . 47

5. Mathematical Tools 51
  5.1. Boosted Decision Trees . . . 51

5.2. Statistics . . . 54
5.2.1. Likelihood Function . . . 54
5.2.2. Hypothesis Tests . . . 55
5.2.3. Statistical Inference . . . 57

II. Analysis Elucidations: Setup, Execution and Results 61

6. Theoretical Details 65
6.1. Modelling Spin and Parity Variations . . . 65
6.1.1. Spin-2 Model . . . 66
6.1.2. CP Models . . . 67
6.1.3. Sensitive Variables for Spin and CP Test . . . 69
6.2. MC Event Generators . . . 71

7. Event Selection and Backgrounds 75
7.1. Trigger and Offline Criteria . . . 75
7.2. Object Definitions . . . 76
7.2.1. Electrons . . . 76
7.2.2. Muons . . . 77
7.2.3. Common Lepton Criteria . . . 77
7.2.4. Jets . . . 78
7.2.5. Missing Transverse Momentum . . . 79
7.2.6. Overlap Removal . . . 79
7.3. Event Preselection . . . 80
7.4. Event Selection . . . 80
7.4.1. Distributions of Sensitive Variables . . . 80
7.4.2. Signal Region Definition . . . 81
7.4.3. Cut Flow . . . 86
7.5. Backgrounds . . . 89
7.5.1. Control Regions . . . 89
7.5.2. Non-resonant WW Boson Production Process . . . 92
7.5.3. Top Quark Processes Control Region . . . 93
7.5.4. Drell-Yan Backgrounds . . . 95
7.5.5. Treatment of Misidentified Leptons . . . 95

8. Boosted Decision Tree Application 99
8.1. Boosted Decision Tree Overview . . . 99
8.2. Training Parameters and Choice of Variables . . . 99
8.2.1. Spin-2 Analysis Training . . . 100
8.2.2. CP Analysis Training . . . 101

9. Boosted Decision Tree Output Rebinning Approach 105
9.1. Simple Rebinning . . . 105
9.2. Smart Rebinning for Spin and CP Tests . . . 108
9.2.1. Smart Binning Algorithm in Detail . . . 108
9.2.2. Chosen Binning for Spin and CP Analysis . . . 111

10. Systematics 113
10.1. Experimental Uncertainties . . . 113
10.2. Theory Uncertainties . . . 114

10.3. Impact of Systematics 116

11. Results 119

11.1. Event Yields 119
11.2. Hypothesis Test Results 121

12. Summary and Outlook 129

12.1. Results from Other Channels 129
  12.1.1. Bosonic Channels 129
  12.1.2. Higgs to Diphoton Channel Analysis 129
  12.1.3. $H \to WW^* \to \ell\nu\ell\nu$ Decay Channel 130
  12.1.4. $H \to ZZ^* \to 4\ell$ Decay Channel 131
  12.1.5. Combined Diboson Results 131
  12.1.6. Fermionic Channel 132
12.2. Summary 135
12.3. Run 2 Outlook 135

Part I.

General Introduction and Explanation

CHAPTER 1

Introduction

The development of the Standard Model of particle physics is a huge success story of theoretical physics in the 20th century. After findings in early experiments lead to the unveiling of an ever expanding particle zoo, the Standard Model managed to introduce an elegant ordering scheme. Its success is demonstrated by the prediction of particles that were discovered in later experiments, filling in the gaps in the symmetric nature of the theoretical construct. Yet the model was not complete without one keystone which persistently defied detection—the Higgs boson. It plays a key role as it explains why non-composite particles can have mass without leading to contradictions within the theoretical framework. It took 48 years after the first predictions of the underlying mechanism until the predicted Higgs boson was finally discovered at the Large Hadron Collider (LHC) at CERN, the European Organization for Nuclear Research. The discovery was made by two experiments at the same time, by ATLAS ("A Toroidal LHC ApparatuS") and CMS ("Compact Muon Solenoid").

On 4th July 2012 both experiments announced a nearly $5\,\sigma$ excess verifying that an unknown boson at a mass of around $125\,\mathrm{GeV/c^2}$ had been observed. Even though the search conditions to find the Higgs boson already gave an idea of the resonance's very nature, it could not be identified definitely at this point.

It was before the discovery of the Higgs boson in summer 2012 when the preparation of this analysis started at a time when only a small excess had been seen. In order to support the assumption that the looming resonance was compatible with the expectations from the Standard Model, tests of the particle properties were targeted. The presented analysis focuses on two integral quantum numbers, namely spin and parity, and their test in the decay channel to charged leptons and neutrinos via two W bosons. For this, data collected by the ATLAS experiment in 2012 has been analysed. Since the lifetime of the Higgs boson is very small (predicted as about 10^{-22} s [1]), only its decay products can be investigated to infer the particle properties. This analysis focused on the $H \to WW^* \to e\nu\mu\nu$ decay channel where the Higgs boson is produced via gluon-gluon fusion. The channel features two different flavour leptons and corresponding neutrinos in the final state, making lepton identification and missing transverse momentum measurements key demands for the analysis. Even though the two neutrinos in the final state pose a challenge for reconstructing the decay process, it is one of the most promising channels to conduct the tests presented in this thesis. The lack of information on the decay's parameters is overcompensated by the expected event yields[1] in this channel and the nevertheless clear event signature.

---

[1] The event yield equals the luminosity times (predicted) production cross section times branching fraction.

While at first the main focus was to measure the spin— for which results have been published for $H \rightarrow WW$ in [2] and for the combination of all sensitive channels in [3])—the emphasis has since been broadened to not only test the parity but also parity admixture states.

The following chapters in the first part of the document give an overview of the background information covering the theoretical and experimental introduction, the phenomenology as well as mathematical and algorithmic basics. The second part is dedicated to the analysis details starting from the selection of the tested models over the signal specification and background estimation, over the discussion of a multivariate analysis technique and the processing of its output, over the consideration of systematic effects to the presentation of the statistical evaluation and results. Finally, the outcomes from other decay channels are compared with these findings.

All measurements of the discovered particle's properties serve the purpose of validating that the Standard Model expectations are met and to look for possible discrepancies which would directly point to new physics. The design of the tests has been chosen to be as model-independent as possible. This approach guarantees that possible Higgs-like particles which are predicted by extended models (such as Supersymmetry (SUSY) or two-Higgs-doublet models (2HDM)) can be either confirmed or rejected. By testing the Higgs quantum numbers and looking for admixture states, a crucial step in analysing the Higgs boson is taken. At the same time the analysis is a mean to dig deeper into the Standard Model and beyond-the-Standard-Model phenomena which are for the first time observable at the LHC.

The results obtained in the course of the analysis have already been combined with other channels and are published in literature, e.g. by the particle data group (cf. [4]). Dedicated publications of the $H \rightarrow WW^*$ analysis results and the combined results can be found in [5] and [6].

CHAPTER 2

Theoretical Foundations

2.1. Standard Model of Particle Physics

The Standard Model of particle physics (subsequently often abbreviated to *SM*) is the theory which describes all known fundamental particles and their interactions. It was developed in a collaborative effort in the second half of the 20th century by a number of theoretical physicists each of whom contributed parts to the overall model. A first step towards today's model was the combination of the electromagnetic and the weak interaction in 1961 by Sheldon Glashow. Subsequent key contributions were made by Steven Weinberg and Abdus Salam in 1967 who combined the Higgs mechanism with the electroweak model [7]. The theory of the strong interaction anon is based on many individual contributions and has been part of the model in today's form since the early 1970s.

2.1.1. Particles of the Standard Model

The Standard Model describes three different classes of particles. An overview of all discussed particles including a subset of their quantum numbers is shown in figure 2.1. Unless otherwise stated, measurement results such as particle masses are taken from the Particle Data Group publications (e.g. [4]).

All known forms of matter are composed from a set of fundamental fermions with spin $1/2$ which are again broken down into leptons and quarks[1]. The former group is made up of three types of neutrinos as well as three types of charged leptons. While the neutrinos are considered massless in the Standard Model[2] and do not carry electric charge, the charged leptons (each with a charge of -1 elementary charges e) have masses up to the GeV scale. The latter group comprises 6 types of quarks grouped in three doublets which are the constituents of hadrons such as the proton or the pion. In each doublet the *up*-type quark carries a charge of $+2/3$ while the *down*-type quark carries a charge of $-1/3$.

The second class of particles described by the Standard Model are vector bosons with spin-1. Three of the four known fundamental forces (electromagnetic force, gravity, weak interaction and strong interaction) can be understood as the exchange of mediator particles. These mediators, also

---

[1] Throughout the thesis the system of natural units as common in high energy physics is used, where $\hbar = c = k_\mathrm{B} = 1$.

[2] Neutrino oscillation which has been observed requires masses, more details follow in subsequent sections.

referred to as gauge particles, comprise four different types of particles. For each interaction there is a dedicated type of charge that the exchange particles couple to.

The photon is massless and does not carry any charge. It is the exchange boson of the electromagnetic force and couples to the electric charge.

The strong interaction involves three kinds of charges which are called colour charge. The charge type can be *red*, *green* or *blue* as well as the respective anti-colours.

The mediator of the strong interaction is called gluon. Gluons are massless yet they carry a charge such that the exchange bosons of the strong interaction are subject to self-interaction. This effect causes the limited range of the strong interaction and has further consequences that are discussed in section 2.1.3. As a consequence of the underlying charge symmetry there are 8 different gluons.

The weak interaction is mediated by two separate force carriers which differ in a number of aspects. This is on the one hand the W boson which carries an electric charge (± 1) and weighs around 80 GeV and on the other hand the Z boson which is electrically neutral and weighs around 91 GeV. Based on these particle properties, weak interaction processes can be classified as *neutral currents* (Z exchange) and *charged currents* ($W^{\pm}$ exchange). Charged currents affect all quarks and leptons with left handed chirality[3] as well as all antiquarks and antileptons with right handed chirality. Neutral currents additionally affect the right-handed fraction of *charged* fermions. The exchange of a Z boson does not change the flavour (i.e. type) of participating fermions which is the case for processes involving $W^{\pm}$. A more in-depth description of the weak interaction and the extension to the electro-weak interaction is given in the following sections.

The third class of fundamental particles is related to the previous one but solely contains the Higgs Boson as the only known (and predicted) scalar particle. Unlike the mediator particles, the Higgs boson is not a gauge particle and therefore does not mediate an interaction. It was predicted as a mean to assign mass to elementary particles (especially the Z and W boson) within the framework of the Standard Model by breaking the symmetry of the electroweak sector. The details of this mechanism and the concerned symmetry are discussed in section 2.1.4.

2.1.2. Symmetries

The universality of couplings in all interactions described by the Standard Model motivates an approach to the theoretical description which is based on the observed symmetries, by introducing assigned symmetry groups and identifying group elements and generators with observable particles.

As mentioned in the previous section, (left handed) quarks and leptons can be assigned to three families each as doublets of a quantity called the *weak isospin* (T with T_3 as the z component which can take values $\pm 1/2$):

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L \quad \text{and} \quad \begin{pmatrix} u \\ d' \end{pmatrix}_L, \begin{pmatrix} c \\ s' \end{pmatrix}_L, \begin{pmatrix} t \\ b' \end{pmatrix}_L . \tag{2.1}$$

While this applies to the left handed fermions, all right handed fermions—charged leptons and quarks since right-handed neutrinos have not been observed and would not interact—only form singlets on their own with a weak isospin of 0:

$$e_R,\ \mu_R,\ \tau_R \ \text{and}\ u_R,\ d'_R,\ c_R,\ s'_R,\ t_R,\ b'_R,\ . \tag{2.2}$$

---

[3] The chirality indicates if a particle transforms in a left-handed or right-handed manner within the Poincaré space. For massless particles the chirality is identical to the helicity, i.e. the projection of the spin of the particle onto its direction of motion. Right-handed helicity herein means that spin and momentum point into the same direction, vice versa for left-handed helicity.

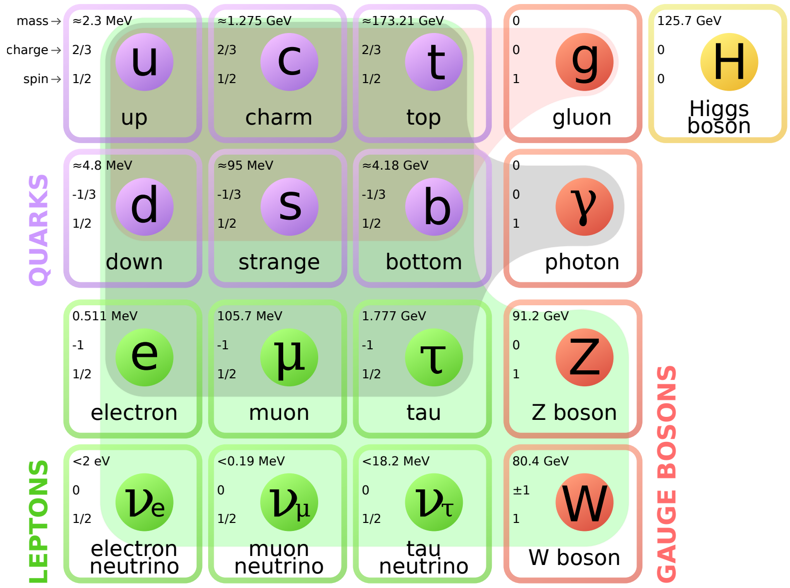


Figure 2.1.: Overview of fundamental particles within the Standard Model of particle physics including their most important properties and their classification. The coloured areas highlight which fermions are subject to the interactions mediated by the corresponding gauge bosons on the right. (derived from [8] using data from [4])

The symmetry associated with the weak isospin is of type $SU(2)_L$ (affecting only left-handed fermions). There is a link between the weak isospin and the charge of a fermion Q via the relation

$$Q = T_3 + \frac{Y_W}{2}. \tag{2.3}$$

The newly introduced quantity Y_W is called the weak hypercharge and corresponds to a $U(1)$ gauge symmetry. In the electroweak model, both symmetries form a $SU(2)_L \otimes U(1)_Y$ gauge group (index Y for the weak hypercharge). The weak isospin and the hypercharge are generators of the group while the particles given in expression 2.1 are base states of the resulting group.

The remaining interaction which is described by the Standard Model is the strong interaction which couples to the colour charge of quarks. The existence of three charge types (*red, blue, green*) implies a $SU(3)_C$ (C for colour) symmetry. In total the gauge group of the Standard Model is

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y. \tag{2.4}$$

The mediators of the strong as well as the electroweak force which are assigned to the $SU(3)_C$ and $SU(2)_L \otimes U(1)_Y$ are gluons (strong) and $W^{\pm}$, Z^0 and γ bosons (electroweak interaction).

The combined electroweak model based on the $SU(2)_L \otimes U(1)_Y$ gauge group introduces four massless gauge bosons. Three of these form a weak isospin triplet and are called A^1 to A^3 in the following. The fourth particle forms a weak isospin singlet and is denoted as B. As a consequence of spontaneous symmetry breaking in the electroweak sector of the Standard Model, linear combinations of these gauge bosons arise as mass eigenstates of then massive fields. This mechanism is discussed in section 2.1.4 on the Higgs mechanism.

The massive mediators of the weak interaction $W^{\pm}$ and Z^0 as well as the massless photon result from the following transformations:

$$\begin{aligned} |\gamma\rangle &= \cos(\theta_W)\,|B\rangle + \sin(\theta_W)\,|A^3\rangle\,, \\ |Z^0\rangle &= -\sin(\theta_W)\,|B\rangle + \cos(\theta_W)\,|A^3\rangle\,, \\ |W^{\pm}\rangle &= \frac{1}{\sqrt{2}}\left(|A^1\rangle \mp \imath\,|A^2\rangle\right). \end{aligned} \tag{2.5}$$

The angle θ_W is the so-called *Weinberg angle* or *weak mixing angle* which parametrises the γ/Z^0 composition from the original gauge bosons. The relation can more clearly be expressed as a rotation matrix and a two dimensional vector holding the respective particles:

$$\begin{pmatrix} Z^0 \\ A \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \cdot \begin{pmatrix} A^3 \\ B \end{pmatrix}. \tag{2.6}$$

In addition the Weinberg angle defines the W/Z mass ratio which is given as

$$\frac{m_{W^{\pm}}}{m_{Z^0}} = \cos(\theta_W). \tag{2.7}$$

The CPT Theorem

There are three fundamental discrete symmetries each of which is represented by one of the letters in this section's title. These comprise

- "C" for charge symmetry indicating whether physical processes are identical if particles are replaced by their respective antiparticles with the opposite charge,
- "P" for parity symmetry indicating whether physical processes are invariant under point reflection at the origin and
- "T" for time reversal symmetry indicating whether physical processes are invariant if the direction of time is reverted.

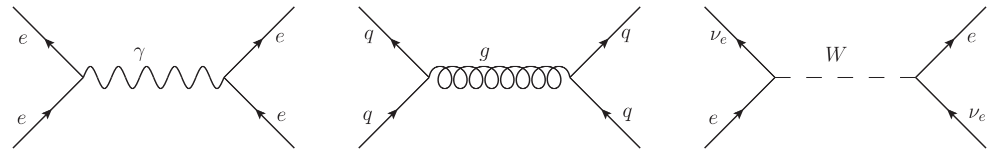
The CPT theorem states that all physical processes are invariant under the combined application of all three transformations while observations have shown that each single symmetry is not preserved for every individual interaction. This also holds for the combined symmetry under a CP conjugation replacing particles with their antiparticles and mirroring their positions which is violated in the weak interaction[4]. The single C and P symmetries are obviously violated in the context of the weak interaction due to the different couplings to chirality components.

Cabibbo-Kobayashi-Maskawa-Matrix

It is important to note, that the down-type quarks listed in the expressions 2.1 and 2.2 are named d' etc. This indicates that for the case of the electroweak interaction the quarks in one doublet component (conventionally the down-type quarks) are not identical to the corresponding quarks participating in the strong interaction. In other words, the mass eigenstates of the quarks (which are also the eigenstates of the strong interaction) are not identical to the eigenstates of the (electro)-

---

[4] The violation of CP symmetry was first discovered at hands of the decay of neutral kaons in 1964 [9].



(a) Electromagnetic interaction **(b)** Strong interaction **(c)** Weak interaction

Figure 2.2.: Overview of tree level diagrams for all three interactions described within the Standard Model. In each case, a scattering process via the exchange of a force carrier boson is shown (i.e. the time axis is intended to point upwards). The respective interactions are indicated below each diagram.

weak interaction. The transition can be expressed as a complex three dimensional matrix, the *Cabibbo-Kobayashi-Maskawa-Matrix* (CKM Matrix):

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.8}$$

At first sight this leaves $2 \cdot 3^2$ free parameters yet their number gets reduced by the requirement of unitarity (i.e. the quark content needs to be conserved) and due to the fact that 5 complex phases cannot be observed. This leaves four free parameters, three of which can be expressed as angles in a three-dimensional rotation matrix. The fourth parameter is a complex phase which is required to accommodate CP violation occurring in the electroweak sector.

The difference between the eigenstates of the weak interaction and the mass eigenstates allows transitions between doublets of the weak isospin through the exchange of $W^{\pm}$ bosons.

2.1.3. Interactions in the Standard Model

Feynman Diagrams

The three interactions described by the Standard Model can be illustrated and calculated based on Feynman diagrams which vividly show the regarded process as well as offer means to calculate the scattering amplitude. Each component of a Feynman diagram has a factor assigned to it such that the scattering amplitude can be calculated by integrating over the composed expression. Due to the structure of the electromagnetic and weak interaction as well as the strong interaction at high momentum transfers, it is possible to apply a perturbative calculation of these scattering amplitudes. By summing up all Feynman diagrams which contain a certain number of vertices (where coupling factors such as the fine-structure constant α therefore appear to identical powers) and have identical signatures, one obtains a descriptive way of calculating a series expansion in the term of the perturbative description of the scattering process regarded.

For each interaction, one simple exemplary process is shown in figure 2.2. Each diagram shows a scattering process via the exchange of a force carrier boson (photon, gluon and W boson) with a coupling vertex at both ends of the boson propagator. For the example of the electromagnetic scattering process, this leads to a scattering amplitude

$$T_{fi} \sim -e^2 \bar{u}(p') \gamma^{\mu} u(p) \frac{1}{q^2} \bar{u}(k') \gamma_{\mu} u(k). \tag{2.9}$$

The quantities p, p', k and k' represent the four-momenta of the incoming (p,k) and outgoing (p', k') electrons (represented by the letter u while the bar above u stands for an outgoing propagator). The momentum transfer is as usually represented by the quantity q^2 with

$$q^\mu = (p' - p)^\mu = (k - k')^\mu. \tag{2.10}$$

The factor e^2 (electron charge) and the gamma matrix elements γ^μ arise from the electromagnetic coupling constant occurring twice in the diagram.

The parity-violating nature of the weak interaction is similarly implemented in the Feynman diagram framework. Each vertex involving a charged current (i.e. a W emission or absorption) is represented by the Feynman factor

$$\imath \frac{g}{\sqrt{2}} \gamma_\mu \frac{1 - \gamma_5}{2}. \tag{2.11}$$

In this expression the $(1 - \gamma_5)$ factor plays the role of a chirality operator projecting a field vector to the corresponding chirality component (which depends on whether a particle or antiparticle field vector is multiplied). The factor g is the coupling constant of the weak interaction.

Coupling Constants

The electromagnetic coupling constant α_{em} (also known as the fine structure constant $\alpha = \frac{e^2}{4\pi} \approx 1/137$ [10]) as well as the coupling constant of the weak interaction $\alpha_W \propto g$ have already been mentioned. Equivalently there is a coupling constant of the strong interaction called α_s. Like the former coupling "constants" it is dependent on the involved momentum transfer q^2 (to a much higher extent, though). It is for this reason that the term "running coupling constant" has been introduced to characterize this behaviour. Since for low values of q^2 the value of α_s grows dramatically, only processes involving high momentum transfers can be described using perturbation theory approaches in QCD (a fact which is referred to as *asymptotic freedom*).

The values of the coupling constants are roughly $1/128$ for $\alpha_{\text{em}}(m_Z)$, $1/30$ for $\alpha_W(m_Z)$ and $1/9$ for $\alpha_s(m_Z)$ (each evaluated at the Z mass scale)[11, 12]. The coupling constant of the weak interaction is in fact higher than the electromagnetic coupling constant which at first may seem counterintuitive. The reason the weak interaction is nevertheless suppressed with respect to the electromagnetic interaction arises from the high mass of the mediator bosons which appears squared in the denominator of their propagator terms. Thus while the coupling term in fact favours the weak interaction over the electromagnetic interaction, it is in total strongly suppressed.

2.1.4. Higgs Mechanism

The Higgs mechanism[5] is a cornerstone of the Standard Model from various perspectives. The Standard Model without the Higgs boson, i.e. without the mechanism of spontaneous symmetry breaking would lack a way of assigning masses to elementary particles without violating gauge invariance of the theory. Furthermore, even after the existence of the W and Z boson was known, there were tree level divergences left to be resolved. The Higgs boson is suited to remove these tree level divergences in the electroweak sector, at the same time imposing an upper Higgs boson mass limit of about 1 TeV. To resolve the unitarity requirements and elementary particle (especially vector boson) mass caveats, the mechanism of spontaneous symmetry breaking of a Higgs field (with an associated Higgs boson) comes into play.

---

[5] Naming all contributors and independent research groups that contributed to the findings, the mechanism has been called the ABEGHHK'tH mechanism (for Anderson, Brout, Englert, Guralnik, Hagen, Higgs, Kibble and 't Hoof) by Peter Higgs [13].

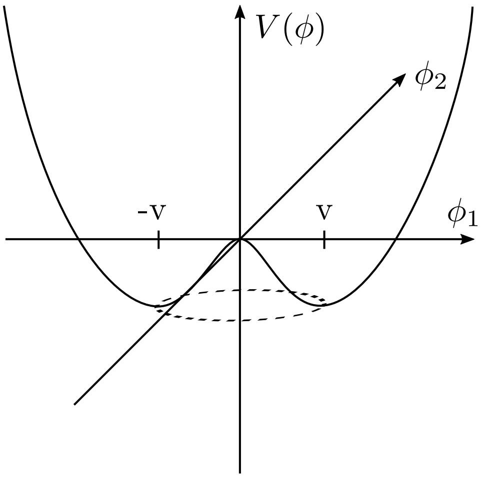


Figure 2.3.: 'Mexican hat' potential illustrating the nature of the broken Higgs potential symmetry with a degenerated ground state.

Spontaneous Symmetry Breaking

The mechanism of spontaneous symmetry breaking combined with local gauge invariance can be illustrated by the example of a QED Lagrangian in a simple way (more details in [10]).

The starting point for this simplified demonstration is a Lagrangian of a complex scalar field $\phi = \frac{1}{\sqrt{2}}(\phi_1 + \imath\phi_2)$ with a kinetic and potential term coupling to itself and an electromagnetic field:

$$\mathcal{L} = -\frac{1}{4}(F_{\mu\nu})^2 + |(D_\mu\phi)^*(D_\mu\phi)| - \underbrace{\left(\mu^2\phi^*\phi + \frac{\lambda}{2}(\phi^*\phi)^2\right)}_{\text{potential } V(\phi)} \tag{2.12}$$

In this expression, D_μ represents the covariant derivative, $D_\mu = \partial_\mu + \imath e A_\mu$ implementing the coupling to the photon.

The Lagrangian is invariant under the following local gauge $U(1)$ transformation known from electrodynamics:

$$\phi(x) \to e^{\imath\alpha(x)}\phi(x), \quad A_\mu(x) \to A_\mu(x) - \frac{1}{e}\partial_\mu\alpha(x). \tag{2.13}$$

If one chooses μ^2 to be positive, this formula describes a potential like the one in figure 2.3 where the state of maximal symmetry (at $(0,0)$) is *not* the ground state, i.e. the state with the least potential energy. Furthermore, the ground state is degenerated in radial direction. W.l.o.g. we can choose a minimum on the real axis:

$$\langle\phi\rangle = \phi_0 = \sqrt{\frac{\mu^2}{\lambda}}. \tag{2.14}$$

It is useful to choose a parametrization of the complex field such that the origin is shifted to the minimum of the potential:

$$\phi(x) = \phi_0 + \frac{1}{\sqrt{2}}(\phi_1(x) + \imath\phi_2(x)). \tag{2.15}$$

Inserting the rewritten field expression into the potential $V(\phi)$ yields

$$V(\phi) = -\frac{1}{2\lambda}\mu^4 + \frac{1}{2} \cdot 2\mu^2 \cdot \phi_1^2 + \mathcal{O}(\phi_i^3). \tag{2.16}$$

This expression shows that the real component of the field, ϕ_1 gets assigned a mass, $m_1 = \sqrt{2}\mu$. The imaginary component of the field, ϕ_2 remains massless and corresponds to the Goldstone boson emerging from the symmetry breaking. Inserting 2.15 into the kinetic term in the Lagrangian (2.12) on the other hand yields

$$\begin{aligned} |D_\mu\phi|^2 =& \tfrac{1}{2}(\partial_\mu\phi_1)^2 + \tfrac{1}{2}(\partial_\mu\phi_2)^2 - 2\phi_0^2\lambda\phi_1^2 + \sqrt{2}e\phi_0 A_\mu\partial^\mu\phi_2 \\ &+ \underbrace{e^2\phi_0^2 A_\mu A^\mu}_{\text{mass term}} - \tfrac{1}{4}(F_{\mu\nu})^2 + \text{interaction terms}. \end{aligned} \tag{2.17}$$

The highlighted term is a mass term for the photon corresponding to $\frac{1}{2}m_A^2 A_\mu A^\mu$ which implies that

$$m_A = \sqrt{2}e\phi_0. \tag{2.18}$$

Thus, the photon (which is of course in fact massless) can acquire mass by applying the Higgs mechanism to QED.

The Higgs Mechanism within the Standard Model

The actual application of the Higgs mechanism within the Standard Model is way more complex but follows the same principles. Here, spontaneous symmetry breaking is applied to the Glashow, Weinberg and Salam (GSW) model of the electroweak interaction which features a $SU(2) \otimes U(1)$ symmetry and requires the mechanism to be used for a non-Abelian gauge theory.

The complex scalar field ϕ is herein replaced with a $SU(2)$ complex doublet (with weak hypercharge $+1$):

$$\boldsymbol{\phi} = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2}\phi^+ \\ \phi^0 + \imath a^0 \end{pmatrix}. \tag{2.19}$$

The potential remains the same as in formula 2.12, wherein only the field gets replaced by the above expression. Equivalently, again the ground state can be chosen to be

$$\langle\boldsymbol{\phi}\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \tag{2.20}$$

where v denotes the vacuum expectation value of the Higgs potential ground state. It is necessary to introduce both, a $SU(2)$ and $U(1)$ gauge symmetry to obtain massless gauge bosons (which can be identified with the massless photon) in the result. The combined gauge transformation therefore looks like

$$\boldsymbol{\phi} \to e^{\imath\alpha^a\sigma^a/2} e^{\imath\beta/2} \boldsymbol{\phi} \tag{2.21}$$

(σ^a are the Pauli matrices). The covariant derivative $\boldsymbol{D}_\mu$ now becomes

$$D_\mu\phi = (\partial_\mu - \imath g A_\mu^a \sigma^a/2 - \tfrac{1}{2}\imath g' B_\mu)\phi. \tag{2.22}$$

In the above expression, A_μ^a and B_μ are the gauge bosons assigned to the $SU(2)$ and $U(1)$ gauge symmetries, each as introduced in section 2.1.2. Inserting expression 2.22 into the kinetic part of the electroweak (Higgs) Lagrangian yields the mass terms of the gauge bosons. The relevant fraction of the Lagrangian after a series of intermediate steps looks like

$$\Delta\mathcal{L} = \frac{1}{2}\cdot\frac{v^2}{4}\left[g^2(A^1_\mu)^2 + g^2(A^2_\mu)^2 + (-gA^3_\mu + g'B_\mu)^2\right]. \tag{2.23}$$

This leads to the three massive and one massless vector bosons by constructing the linear combinations (= mass eigenstates) of the massive boson fields like shown previously (equations 2.5):

$$\begin{aligned} W_\mu\pm &= \frac{1}{\sqrt{2}}(A^1_\mu \mp \imath A^2_\mu), \\ Z^0_\mu &= \frac{1}{\sqrt{g^2+g'^2}}(gA^3_\mu - g'B_\mu), \\ A_\mu &= \frac{1}{\sqrt{g^2+g'^2}}(g'A^3_\mu + gB_\mu). \end{aligned} \tag{2.24}$$

Comparing equations 2.5 to the above expressions leads to the identifications

$$\cos(\theta_W) = \frac{g}{\sqrt{g^2+g'^2}} \quad\text{and}\quad \sin(\theta_W) = \frac{g'}{\sqrt{g^2+g'^2}}. \tag{2.25}$$

The corresponding masses assigned to each boson are

$$m_{W^\pm} = g\cdot\frac{v}{2},\ m_{Z^0} = \sqrt{g^2+g'^2}\cdot\frac{v}{2} \ \text{ and }\ m_A = 0. \tag{2.26}$$

The photon (A_μ) remains massless while the relations in 2.24 introduce the relation between the masses of the W and Z boson as already given in equation 2.7.

Via the above relations the vacuum expectation value v can be determined which has been measured to amount about 246 GeV [14].

It is furthermore possible to relate the field coefficients to the electron charge, it holds that

$$e = \frac{gg'}{\sqrt{g^2+g'^2}}. \tag{2.27}$$

The Higgs Boson

After discussing the mechanism of spontaneous symmetry breaking of a local gauge symmetry with respect to the electroweak symmetries and identifying the occurring field components with massive vector bosons on the one hand and the massless photon on the other hand, it is not immediately obvious where the Higgs boson itself appears. It becomes visible once a parametrisation of the field $\boldsymbol{\phi}(x)$ is chosen where the second component contains the (real-valued) Higgs field as a derivation from the vacuum expectation value v of the Higgs potential ground state:

$$\phi(x) = U(x)\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x)\end{pmatrix}. \tag{2.28}$$

$U(x)$ represents a general $SU(2)$ gauge transformation which can be eliminated likewise for the current consideration.

Inserted into the Higgs part of the Standard Model Lagrangrian (equation 2.22 and analogous to 2.12),

$$\mathcal{L} = |D_\mu\phi|^2 + \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2, \tag{2.29}$$

the kinetic part again leads to a mass term which this time belongs to the newly introduced Higgs field. Carrying out the calculations one obtains the following relation for the Higgs mass:

$$m_h = \sqrt{2}\mu^2 = \sqrt{\frac{\lambda}{2}}\, v. \tag{2.30}$$

Since the value of the self-coupling parameter λ could not be predicted, there was no prediction for the Higgs boson mass (besides constraints e.g. from unitarity considerations).

Fermion Masses

It is not possible to add simple mass terms for fermions to the Lagrangian because in the context of chirality this would lead to the violation of gauge invariance. For this reason, the Higgs mechanism (or the mechanism of spontaneous symmetry breaking) needs to step in here as well to give rise to fermion masses. These do not automatically emerge from the mechanism described above, though. Instead, they can be introduced by adding Yukawa coupling terms to the Lagrangian, linking the various fermions and chirality states and the Higgs potential.

For the example of the electron, this would look like

$$\Delta\mathcal{L} = -\lambda_e \left[(\bar{\nu}_e, \bar{e})_L \boldsymbol{\phi}\, e_R + \bar{e}_R\, \phi^\dagger \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \right] \tag{2.31}$$

with ϕ being the Higgs potential (see 2.28, with $U(x)$ removed via gauge transformation) and λ_e representing a specific coupling factor for the electron. Inserting the expression for ϕ and executing the calculations leads to a mass term for the electron on the one hand (first part of the expression in the field parametrization) and a coupling term to the Higgs field $h(x)$ on the other hand (second part of the named expression). An analogous approach can be applied for quarks (here the CKM matrix needs to be taken into account to relate the different quark eigenstates to one another) yielding mass terms in the Standard Model Lagrangian for all massive fermions.

CHAPTER 3

Experimental Apparatus and Preconditions

The measurements for this analysis have been conducted using the ATLAS detector, one out of 7 experiments (4 main and 3 additional experiments) at the LHC at CERN. The Large Hadron Collider is a proton and heavy ion ring accelerator and collider.

The two biggest experiments *CMS* (see [15]) and *ATLAS* (see [16]) are designed as ‘discovery machines’, i.e. as versatile detectors with the potential of finding the Higgs and looking for a wide range of physics beyond the Standard Model. *LHCb* (“Large Hadron Collider beauty”) focuses on heavy flavour physics. The experiment seeks to investigate CP violation and rare decays of hadrons containing *bottom* and *charm* quarks looking for new physics (see [17]). *ALICE* (“Large Ion Collider Experiment”) is a heavy-ion detector focusing on QCD. It is the primary experiment designed to investigate Pb-Pb collisions as well as other heavy ion interactions and proton-heavy ion interactions. Its goal is to study the created quark-gluon plasma in these collisions at very high energies and temperatures (see [18]).

The three smaller experiments serve specific purposes and have a very limited scope. Their names are *LHCf* (studies neutral particle flux in extreme forward region of one interaction point), *TOTEM* (monitors total *pp* cross section) and *MoEDAL* (search for magnetic monopoles), more information can be found in the particular publications [19], [20] and [21].

3.1. Large Hadron Collider

In order to test the Standard Model of particle physics on the TeV scale and to look for new physics in this regime, the Large Hadron Collider (LHC) was suggested. The first concepts for the LHC go back to a joint workshop between ECFA (European Committee for Future Accelerators) and CERN in 1984 [22, 23]. It was built in the 26.7 km tunnel formerly used by LEP (the *Large Electron–Positron Collider*) and became fully operational by the end of 2009. The LHC was designed to operate with protons and lead ions, reaching particle energies of 7 TeV (*pp* collisions) and 574 TeV (Pb collisions) respectively [24]. The design peak luminosities are $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ for proton-proton and $10^{27}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ for lead collisions.

The luminosity of a collider is given as

$$\mathcal{L} = \frac{N^2 k_b f}{4\pi\sigma_x\sigma_y} F \tag{3.1}$$

with N being the number of particles in each bunch, k_b the number of bunches, f the circulation frequency, $\sigma_{x/y}$ the beam sizes in horizontal and vertical direction at the interaction point and F a reduction factor arising from the angle between the bunches at the interaction point (bunches are oriented parallel to the beam and therefore do not collide exactly frontally) [25] [26].

The choice of protons for both beams was made to avoid limiting the possible luminosity of the LHC due to the necessity of producing a sufficient number of collimated anti-protons in a first step. Other than for the CERN e^+e^- Collider (LEP) and the Fermilab Tevatron Collider where the particles in each beam had opposing charges (e^+e^- and $p\bar{p}$ respectively) and only one beam line was required, this design choice implied the necessity of installing two separate beam lines with inversely oriented magnetic fields. Thus, to accommodate the two counter-rotating proton beams inside the limited tunnel space, a twin-bore superconducting magnet design was implemented.

The downside of this choice is the fact that in pp collisions for quark anti-quark collisions less energy is available on average. This is a consequence of the parton luminosities which differ between valence and sea quarks, peaking around $^1/_3$ for valence quarks but decaying exponentially from 0 for sea quarks. The $p\bar{p}$ collisions at the Tevatron in contrast profited from the fact that always two valence quarks interacted in $q\bar{q}$ processes. Since this downside only affects quark anti-quark processes but not ones involving gluons, its overall impact is acceptable, though.

Even though the LEP tunnel was reused, the fact that protons are accelerated instead of electrons implies that the beam energies can be chosen much higher. This is a consequence of the synchrotron radiation (i.e. radiated energy per circulation ΔE) which grows like

$$\Delta E \propto \frac{1}{\rho}\left(\frac{E}{m}\right)^4 . \tag{3.2}$$

In this expression, E represents the particle energy, m stands for the particle mass and ρ for the radius (see [27]). The ratio of energy and mass of the particle to the fourth power illustrates the overwhelming advantage of proton colliders in the context of synchrotron radiation. In addition the LHC required a profound advancement in the technology of superconducting magnets with respect to LEP technology. The latter can be illustrated at hands of the simplified synchrotron equation for relativistic particles [27],

$$E_{\mathrm{max}} \approx \rho q \boldsymbol{B}. \tag{3.3}$$

The quantities on the right side of this formula are the radius ρ, the particle charge q and the magnetic field $\boldsymbol{B}$. Any increase of achievable energy at constant radii obviously requires a corresponding increase of the magnetic field for bending the synchrotron beams.

The LHC ring itself is not exactly circular but has the shape of a radiused octagon. The straight segments serve diverse purposes. Four of the segments house the four big LHC experiments (ATLAS and CMS at opposed positions). Two out of these segments additionally contain the injection points of the two beams. One segment is dedicated to the accelerator machinery (radio frequency cavities and supportive devices); one other segment holds beam dumps and the remaining two segments are used for beam adjustment and optimisation. The curved ring segments hold arrays of cells which essentially combine quadrupoles in different orientations for beam focusing and dipoles for actually bending the beam via the Lorentz force.[1] All technical details can be found in the LHC design report [28].

An overview of the LHC ring and all connected accelerator facilities is given in figure 3.1. The protons which are to be accelerated originate from a hydrogen bottle. Electrons are stripped from the hydrogen atoms to leave only protons behind. These pass through the linear accelerator

---

[1] In fact also sextupoles, octupoles and quadrupoles with different specific functionalities are included in the bent regions but are not required for the understanding of the basic principle.

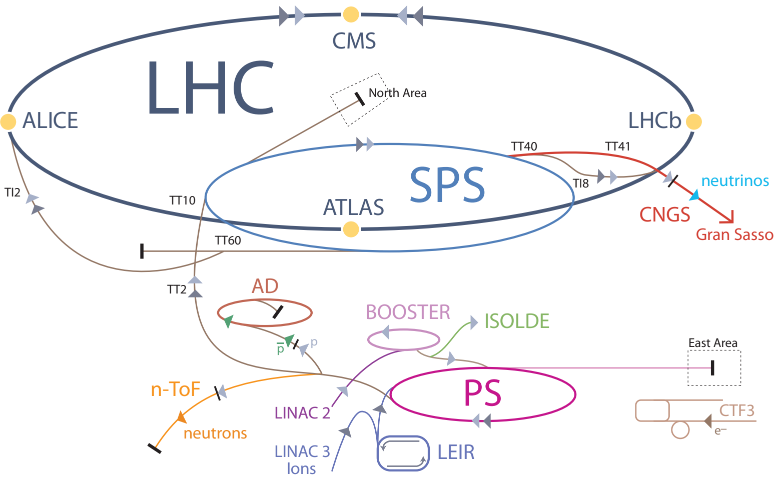


Figure 3.1.: Illustration showing the different accelerators and storage facilities for the LHC and linked experiments. (taken from [29])

LINAC2 accelerating them to 50 MeV. The subsequent booster increases their energy to 1.4 GeV. The next steps are the Proton Synchrotron (*PS*) and the Super Protons Synchrotron (*SPS*) accelerating them to 25 GeV and 450 GeV respectively. At this energy the protons finally get transferred into the LHC where they get injected in both (beam) directions and are accelerated up to their nominal energies [24]. Once the desired beam energy is reached, operation of the LHC is possible over several hours. Protons circulate in bunches within the LHC ring. For 2012 (only data taken during this period has been used for this analysis) each beam held 1374 bunches of $1.6-1.7\times10^{11}$ protons (design values were 2808 bunches of 1.15×10^{11} protons). The temporal distance between orbiting bunches was 50 ns (design value 25 ns) [25].

A plot showing the cumulative luminosity delivered by the LHC for *pp* collisions in 2012 as well as the fraction that was recorded by ATLAS and meets the quality criteria is shown in figure 3.2. The centre-of-mass energy $\sqrt{s}$ was 8 TeV over the entire period.

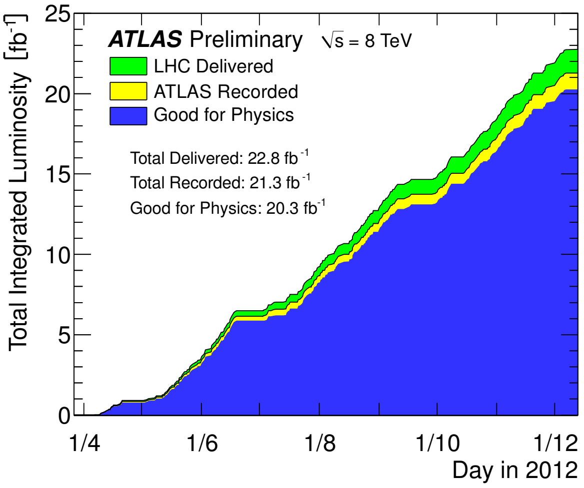


Figure 3.2.: Integrated luminosity of the LHC (2012 run, *pp* collisions with stable beams) and the fraction of recorded/proper data. (taken from [30])

3.2. ATLAS

The ATLAS detector [16] is one of two multi-purpose detectors within the LHC ring. It features a hermetic detector design being nearly cylindrically symmetrical and covering a pseudorapidity range up to $|\eta| \leq 4.9$. The pseudorapidity η is defined as $\eta = -\ln\tan(\theta/2)$ with θ being the azimuthal angle with respect to the beam line[2]. Additionally, a forward-backward symmetry of the detector is given.

The coordinate system used in all geometric concerns is chosen such that the interaction point is the origin of the coordinate system. The positive x axis points to the centre of the LHC ring, the positive y axis points upwards. Both define the plane transverse to the beam direction which is in turn defined as the z axis (see figure 3.3). In many contexts, cylindrical coordinates are used rather than Cartesian ones such that r and ϕ are given the transverse plane.

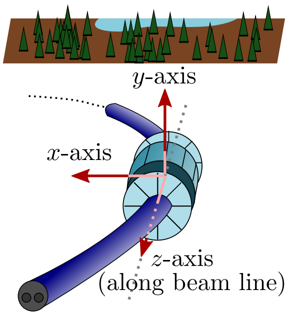


Figure 3.3.: Sketch of the coordinate system used for the analysis with respect to the ATLAS detector and LHC ring.

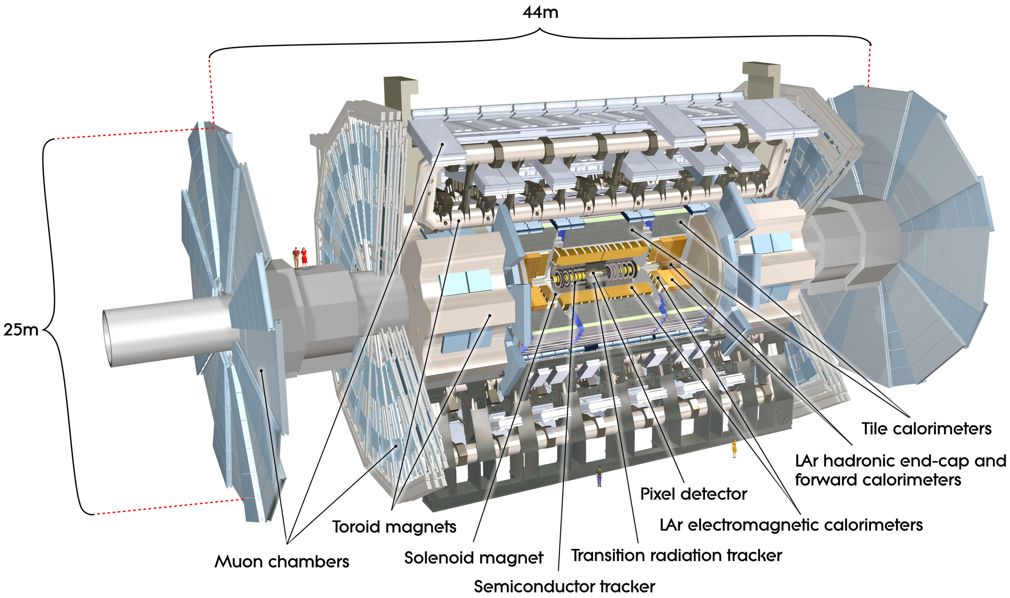


Figure 3.4.: Cutaway view of the ATLAS detector with different sub detectors and magnet components highlighted. (taken from [33])

---

[2] By using the pseudorapidity instead of the angle θ one profits from two characteristics: In proton-proton collisions, the multiplicity of particles over the pseudorapidity features plateaus ($\sqrt{s}$ dependent) over a wide η range. In addition, $\Delta\eta$ is nearly Lorentz invariant for boosts along the beam axis. [31, 32].

Inside ATLAS, several different detector types are aligned in concentric layers (see figure 3.4). On the innermost level there are three tracking sub-detectors (each subdivided into a cylindrical central part and disks in the end caps). The two innermost ones are silicon detectors, the third one is a transition radiation tracker made from straw tubes. The tracking system is surrounded by the electromagnetic and hadronic calorimeters. The outermost detector layers form the muon system.

The motivation for this cylindrical onion design is the nature of the collision products to be investigated. Very close to the interaction point a high tracking resolution is necessary to determine the primary vertex (e.g. to identify b quark jets which emerge from displaced vertices from the delayed b quark decay). The silicon detectors which surround the beam pipe in close proximity have a very high track resolution which additionally allows identifying the charge of particles from the curve direction and the momentum from the curvature of the track within the solenoidal magnetic field. The surrounding transition radiation tracker built from straw tubes is designed differently since a much larger volume needs to be covered. Beyond its application as a particle tracker, the transition radiation furthermore helps to identify particles that pass and specifically contributes to the electron identification.

The electromagnetic calorimeter primarily serves the purpose of identifying photons and electrons and measuring their direction and energy. For this reason its granularity is still relatively high compared to the surrounding hadronic calorimeter. There, hadrons (e.g. pions) create showers and are usually entirely stopped. Their energy and direction is also measured yet the granularity is reduced. The outermost part is the muon system including its own toroidal magnetic system. Hardly any other particles besides muons pass the hadronic calorimeter such that almost only muons are detected. The large volume allows a good momentum resolution with a smaller magnetic field strength. Since muons are crucial for many processes (such as the signal process in this analysis), a lot of effort has been put into the muon system to optimise its characteristics. The design requirements for all detector systems are listed in table 3.1, listing also the η acceptance ranges for the trigger and reconstruction stage.

Table 3.1.: Table of resolution and acceptance requirements for ATLAS sub-detectors. All energies and momenta are given in units of GeV. (taken from [16])

Detector component	Required resolution	η coverage	
		Measurement	Trigger
Tracking (Inner Detector)	$\frac{\sigma_{p_T}}{p_T} = 0.05\,\% p_T \oplus 1\,\%$	± 2.5	
EM calorimeter	$\frac{\sigma_E}{E} = \frac{10\,\%}{\sqrt{E}} \oplus 0.7\,\%$	± 3.2	± 2.5
Hadronic calorimetry (jets)			
barrel and end-cap	$\frac{\sigma_E}{E} = \frac{50\,\%}{\sqrt{E}} \oplus 3\,\%$	± 3.2	± 3.2
forward	$\frac{\sigma_E}{E} = \frac{100\,\%}{\sqrt{E}} \oplus 10\,\%$	$3.1 < \lvert\eta\rvert < 4.9$	$3.1 < \lvert\eta\rvert < 4.9$
Muon spectrometer	$\frac{\sigma_{p_T}}{p_T} = 10\,\%$ at $p_T = 1\,\mathrm{TeV}$	± 2.7	± 2.4

All single detector parts will be covered in more detail in the subsequent sections. The discussion of the subdetectors refers to the configuration of the experiment which was used during the 2012 data taking, i.e. without the modifications applied during the following shutdown.

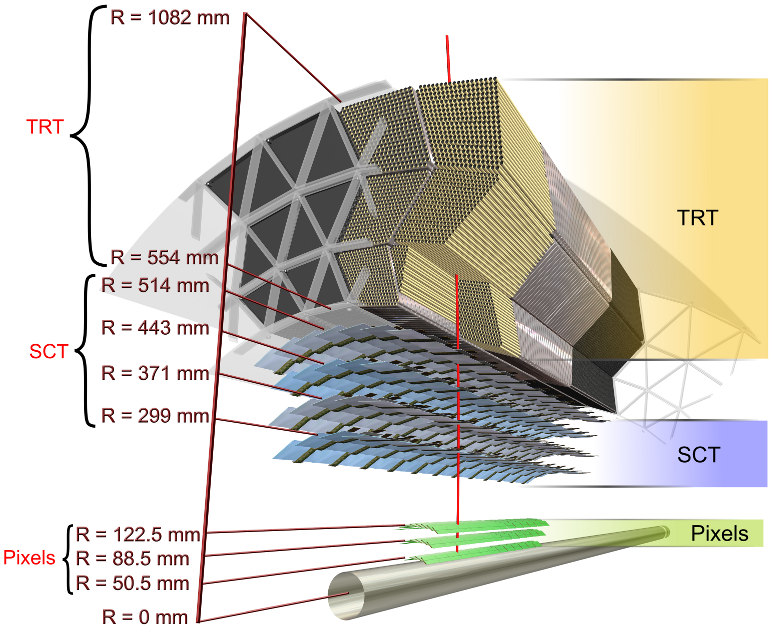


Figure 3.5.: ATLAS Inner Detector during Run 1, central (barrel) region with Pixel Detector, silicon strip detector (SCT) and transistion radiation tracker (TRT) highlighted. (taken from [35])

3.2.1. Beam Pipe

The innermost part of the ATLAS detector is the beam pipe which forms part of the experiment over a length of 38 meters. Its central part is integrated with the pixel detector which adjoins the beam pipe in radial direction. This part of the beam pipe has an inner diameter of 58 mm and a wall thickness of 0.8 mm and is made from beryllium. The choice of beryllium is motivated by its properties as a light element with a low density yet high stiffness and thermal stability (see [34]). The surrounding parts of the beam pipe are made from stainless steel.

3.2.2. Inner Detector

The overall layout of the barrel region of the Inner Detector is shown in figure 3.5. The main purpose of the innermost layers is the identification and reconstruction of the tracks of charged particles. High tracking precision is achieved in the region within $|\eta| < 2.5$ using the high granularity silicon tracker components. Within $|\eta| < 2.0$ a significant amount of information is added by the transition radiation tracker supporting track pattern recognition as well as momentum determination and electron identification.

The Pixel Detector

The Pixel Detector aligns itself with the beam pipe and consists of concentric cylindrical layers of pixel detector modules in the central barrel region and circular disk layers in the surrounding end cap regions. There are 3 layers in the barrel and endcaps region, respectively.

The general layout of the cylindrical detector layers in the centre is chosen such that single panels of active material overlap. This design has been chosen to make sure that at least one

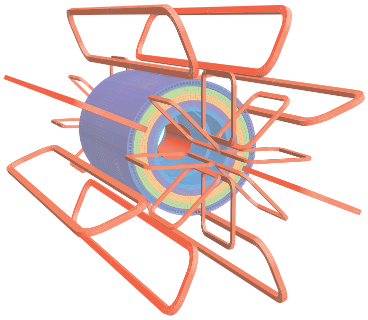


Figure 3.6.: Magnet layout of the ATLAS detector. Most prominently the 8 toroid coils in barrel and end cap region are displayed. The solenoid magnet is integrated with the calorimeter and displayed as four layers with different magnetic properties plus one return yoke in this image (forward shielding not shown). (taken from [16])

detector element is passed by any particle per layer. In the end cap region the rectangular silicon modules are arranged radially on both sides of the disk layers. The pixel detector has a total of about 80.4 million readout channels.

The Silicon-Microstrip Detector

The Silicon-Microstrip Detector (usually referred to as *SCT* for "SemiConductor Tracker") surrounds the pixel detector and is also a semiconductor detector. In the central region it is composed of four double layers studded with silicon strip modules which are aligned parallel to the beam axis. On each double layer the inner and outer panels are twisted by a stereo angle of 40 mrad to provide some localization information along the strip length (z coordinate). In the end cap regions the SCT consists of 9 annuli with radially mounted detector panels on one side and slightly twisted modules on the other side which again serves the purpose of gaining information (radial in this case) along the strip length. The number of readout channels in the SCT is at around 6.3 million.

The Transition Radiation Tracker

The SCT modules are surrounded by the transition radiation tracker (TRT) in radial direction. The TRT comprises a central part and two end cap parts which cover the length of the SCT end cap region. It is made from straw tubes with a diameter of 4 mm[3] which are aligned parallel to the beam line in the barrel region and aligned radially in the end caps. Besides in the transition region between end caps and central region each track passes at least 36 straws adding between 20 (transition between barrel and higher granularity end cap) and 40 hits (transition region between higher and lower granularity end cap) to the track [36]. Even though only coordinates in $R-\phi$ can be recorded, the high number of points significantly improves momentum resolution and enables additional track pattern recognition approaches. The TRT has a total of about 351000 readout channels.

3.2.3. Magnet System

The magnet system inside the ATLAS detector consists of four superconducting magnetic coil arrays which comprise on the one hand the solenoid magnet around the inner (tracking) detector and on the other hand the all-encompassing toroid magnets and two end cap toroid magnets. The

---

[3] A straw chamber is a bunch of densely packed tubular drift chambers equipped with anode wires and very thin aluminium-coated walls. The tubes are filled with a gas mixture optimised with respect to drift conditions.

layout as part of the detector structure is shown in figure 3.4; a breakdown showing only the relevant magnet components is shown in figure 3.6.

The solenoid magnet creates a nearly homogeneous field in the inner detector space with a magnetic field strength of 2 T. The field direction is aligned with the beam direction to allow for measurements of particle charge and transverse momentum from the track curvature in the transverse plane.

The toroid magnets are part of the muon system. The layout was chosen to achieve a very high transverse momentum resolution over a wide η range. It was thus helpful to choose an open structure of air-core toroidal magnets to control the amount of multiple scattering ([37], p. 20 et seqq.). The muon system itself will be discussed in section 3.2.5.

3.2.4. Calorimeter System

One of the main design goals of the ATLAS Calorimeter System was the coverage of a wide η range. Nesting of calorimeter components based on different technologies allows for a range of $|\eta| < 4.9$ to be covered. The calorimeter granularity in the extension of the inner detector ($|\eta| < 2.5$ region) is chosen sufficiently fine to allow for precision measurements of electrons and photons. Outside this regime, a coarser granularity allows to reconstruct jets and the missing transverse energy sufficiently well. Additionally, adequate dimensions of the calorimeter system (at least $\sim$ 10 interaction lengths over the entire η range) were chosen in order to keep punch-through into the muon system on a very low level and guarantee a good containment of both electromagnetic and hadronic showers.

The various subsystems can be studied under reference of figure 3.4. The inner part of the calorimeter system is based on liquid argon as an active medium, granting linear response behaviour and a high durability and stability. This part comprises the entire electromagnetic calorimeter (i.e. barrel and end cap region) as well as the inner part of the hadronic calorimeter end caps (hadronic end cap and within this the hadronic forward calorimeter). The absorber material in this liquid argon part of the calorimeter system is lead in the electromagnetic calorimeter and respectively copper (hadronic end cap) and copper-tungsten (hadronic forward calorimeter). For the electromagnetic calorimeter an accordion layout of the material layers was chosen to avoid gaps and improve uniformity. The hadronic liquid argon calorimeter end caps feature a flat-plate sampling design.

The required cooling is provided within three cryostats housing the electromagnetic barrel calorimeter and the two end caps including the joint hadronic end cap segments.

All above calorimeter structures are surrounded by the tile calorimeter which is the main part of the hadronic calorimeter. It is again subdivided into a barrel and two end cap regions. The active material components here are scintillating tiles with steel being used as absorber material. The tiles are oriented radially.

3.2.5. Muon System

The Muon System surrounds the calorimeters and consists of the dedicated toroidal air-core magnets (see subsection 3.2.3 and figure 3.6) and tracking chambers. The dimensions of the toroidal magnets and the design as such aim for the creation of a high-quality field which is mostly orthogonal to the muon trajectories while avoiding multiple scattering arising from additional material within the detector. In both, barrel and end cap region, there are three layers of tracking chambers which are aligned cylindrically in the barrel or in plates in the end caps (see “Muon chambers” in figure 3.4). For most of the $|\eta|$ range multiple layers of drift tubes within the single chambers are used for measuring track coordinates. Only in a smaller (surrounding) region with $2 < |\eta| < 2.7$ multi-wire proportional chambers are used because of higher rates. The muon system furthermore includes a dedicated embedded trigger system. It features special types of tracking chambers which were chosen concerted with trigger rate and timing requirements. The trigger system serves not only for triggering but also for introducing well-defined p_T thresholds as well as adding additional coordinate information to the main tracking chamber measurements.

3.2.6. Trigger System

The trigger system consists of a hierarchy of three trigger levels with increasing available processing time and information. The first level ("L1") is implemented fully in hardware and has a time span of less than 2.5 μs. It reduces the output event rate from a design collision rate of 40 MHz[4] to about 75 kHz. The trigger selects events with interesting features, e.g. (pseudo-)particles with high transverse momenta such as muons, jets or missing transverse energy. It therefore uses a rather coarse detector representation and simplified algorithms. The available input is processed in a central trigger processor. The definition of interesting features is controlled by a so-called trigger 'menu' which is processed by this trigger processor. The menu holds several hundred trigger chains which define the selection criteria for each trigger step [38]. The L1 trigger identifies regions of interest (RoI) in (η, ϕ) which contain selected features and get passed to higher level triggers.

The second level trigger ("L2") investigates the RoIs defined by the L1 trigger and has access to the full detector granularity and precision within these regions. The L2 trigger and higher trigger levels are software-based and run on standard computer hardware. On average, it has 40 ms processing time available and its menu is designed to reduce the output event rate to 3.5 kHz.

The third trigger level ("Event Filter") has around four seconds of processing time and reduces the output event rate below the kHz level on average (design rate is 200 Hz, in fact by the end of Run 1 around 600 Hz were reached, time-averaged over stable beam periods). The events passing the last selection criteria get transmitted to the CERN computer centre for storage.

3.2.7. Luminosity

The measurement of the luminosity plays a key role to determining accurate cross sections of observed processes. By combining information from a set of subdetectors, a higher measurement precision can be achieved and systematic uncertainties can be estimated more precisely. The main detector systems which are also capable of measuring the luminosity per bunch, i.e. the instantaneous luminosity are the *BCM* (Beam Conditions Monitor) and *LUCID* (LUminosity measurement using Cerenkov Integrating Detector) systems. Additional information and cross checks are contributed by calorimeters, the *MPX*[5] detectors (see [39]) and partly the *ALFA* (Absolute Luminosity For ATLAS) detector which measure the integrated luminosity over all bunches.

Additional information about the detectors, the algorithms for determining the online luminosity and Van Der Meer scan results can be found in [40] for the time up to 2011. For the 2012 data taking the same principles have been followed. The underlying Van Der Meer scans for a preliminary luminosity calibration have been performed in November 2012 (plots are available online, cf. [41]).

Beam Conditions Monitor

The BCM consists of two stations on both sides of the interaction point at a distance of 1.84 m corresponding to $|\eta| = 4.2$. Each of the stations is equipped with four modules holding radiation-hard diamond sensors with an effective area of 1 cm^2 and high speed readout electronics. It serves

---

[4] The design LHC bunch spacing was not used in Run 1 but rather a bunch spacing of 50 ns, i.e. the collision rate was 20 MHz.

[5] MPX is short for Medipix2, meaning pixel detectors measuring the overall radiation at different spots within ATLAS.

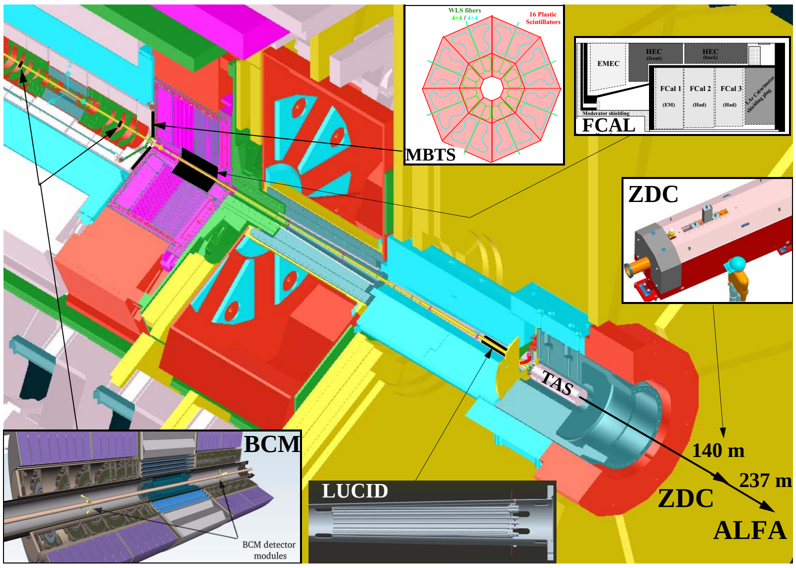


Figure 3.7.: Overview image of the ATLAS detector with highlighted luminosity detector components; remote forward detector parts are indicated along the extended beam line with their distances to the interaction point. (taken from [42])

the purpose of triggering an abort in case the beams delivered by the LHC get too unstable as well as the luminosity measurement.

LUCID

LUCID is specifically designed for luminosity measurements. Like the BCM it comprises two stations on each side of the interaction point at a distance of 17 m which corresponds to $|\eta| \approx 5.8$ (the radial distance is $\approx$ 10 cm). Each station is composed of 16 aluminium tubes pointing to the interaction point that are filled with C_4F_{10} gas and equipped with a photo multiplier tube (PMT) to collect Cerenkov radiation. Also the LUCID detector is designed to be radiation hard and have a fast timing response.

3.3. Event Reconstruction

The event reconstruction process is described in the ATLAS Computing Technical Design Report [44]. A brief overview of the reconstruction steps implemented in the ATLAS software *Athena* is presented in this section.

From the event reconstruction point of view, the detector is split up into three subsystems: the inner detector, the calorimeters and the muon system. Each detector system stores recorded data in *Raw Data Objects* (RDO) which serve as inputs to the first event reconstruction steps in each subsystem. The reconstruction steps such as track finding and energy clustering operate on abstracted hit information for which calibrations and corrections are taken into account. Details about the algorithms used and about the reconstruction performance can be found in [45] for the inner detector track reconstruction, [46] for the calorimeter clustering process and [47] for the muon system track reconstruction and combination with the inner detector.

The tagging, i.e. the assignment of the particle type to a reconstructed track and energy deposition depends on the analysis-specific algorithms which build the final objects for physics analyses. The definitions used for the analysis presented here are discussed in 7.2.

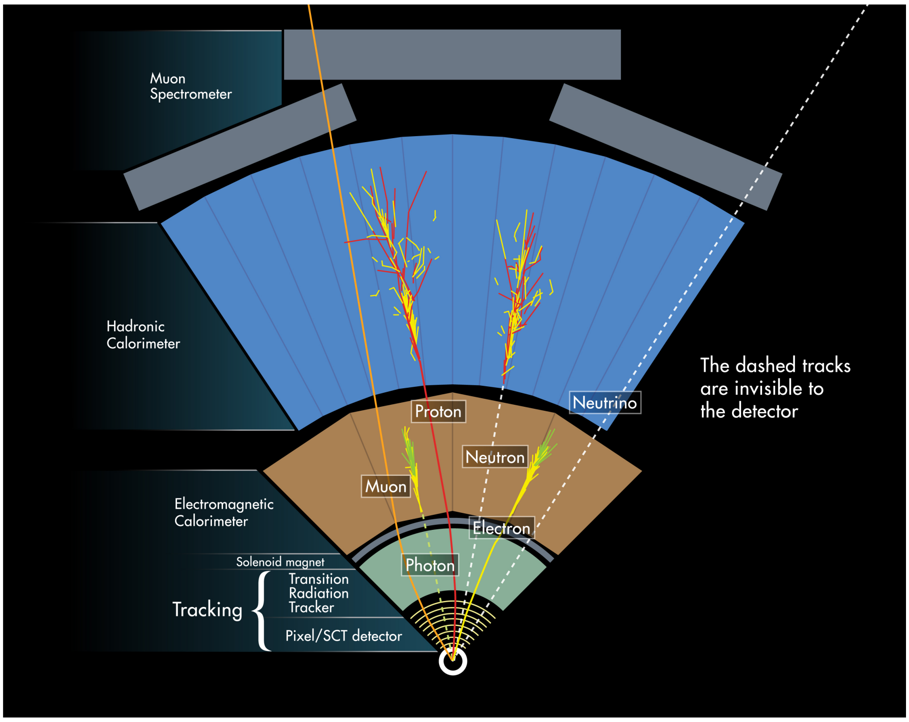


Figure 3.8.: Schematic view of the ATLAS detector, showing the subdetector parts and typical tracks of relevant kinds of particles in though the detector volume. Dotted lines indicate that the particles fly along the drawn path but do not interact with the specific detector part(s). (taken from [43])

Inner Detector Event Reconstruction

At first, *space points* (i.e. abstracted hit information after clustering of active pixels/strips) are formed in the two silicon subdetectors. The starting point for track and vertex reconstruction in the inner detector is the innermost pixel detector with 3 layers of pixel detectors. Combinations of hits on these innermost layers in the barrel and/or end cap region (extended by the first silicon strip detector layer) are checked if they can be starting points of valid tracks. These *seeds* are propagated to the silicon strip layers (8 layers in the barrel, 18 layers in the end caps) along the track estimate. Subsequent hits are added to the track candidate via a Kalman filter algorithm merging measured hit and track parameters (see [48] for an introduction). In parallel, a Hough transformation (feature extraction algorithm, see [49] for details) is applied to the hits in the transition radiation tracker (TRT) surrounding the silicon part of the detector. The outputs are drift circle segments. The silicon detector track candidates are extrapolated to the TRT where they get matched with the drift circle segments. An ambiguity solving step selecting only the most likely track candidates in cases of overlaps and a final track fit produce the final tracks of charged particles from the inner detector. Additionally, the tracks are extended to the beam line to determine the primary vertices.

In the course of this thesis, a GPU (*Graphics Processing Unit*) based implementation of the first steps of the inner detector track reconstruction comprising seed search and extrapolation through the silicon strip detector has been evaluated and integrated into the Athena framework (see [50]).

Calorimeter Reconstruction

The corrected calorimeter cell energy entries (e.g. taking defects into account) are further processed by one out of two clustering algorithms:

- The *Sliding-Window Clustering* algorithm introduces a grid in the $\eta \times \phi$ plane and combines all calorimeter cells in one grid segment on all longitudinal layers into so-called calorimeter towers. A fixed size window is moved across this plane with every significant local maximum making up a cluster seed. Depending on the particle type, surrounding calorimeter cells within a specific range in η and ϕ are added to the cluster. In this step, the calorimeter layer determines the range assignment.

- For the *Topological Clustering*, a loop identifies calorimeter cells with a significant energy deposition to noise ratio. Starting from these seeds, surrounding cells (in lateral and longitudinal direction) are added with a reduced signal to noise threshold. Completed clusters are reprocessed in a subsequent step to resolve single particles in overlap scenarios by making use of local maxima within one cluster.

The choice of clustering algorithm depends on the object of interest (e.g. τ, jets, ...).

Muon Spectrometer Reconstruction

The muon spectrometer is composed of layers of drift chambers which are separated by air-filled volumes and span a range of several meters. The track reconstruction in this outer part therefore requires a combination of methods. On the one hand there is the track segment finding within the drift chamber volumes (partly via Hough transformation) which is also used in the TRT of the inner detector. On the other hand, these track segments are extrapolated between the layers along a curved track which is defined by the traversed material and magnetic field as in the silicon part of the inner detector. To actually build the track candidates, the outermost track segments are used as seeds and get extrapolated towards the detector centre. The compatibility of the extended tracks with the inner segments of the muon spectrometer layers is measured by fitting the (curved) track to the segments. Finally the track candidate is extrapolated to the beam line to obtain its parameters at the nearest position to the beam line.

Depending on the availability of compatible track segments from the calorimeter and/or the inner detector, four different quality classes of muons are defined:

- *Combined* muons make up the majority of reconstructed muons and are formed by combining tracks from the inner detector and the muon spectrometer where possible.

- *Stand-Alone* muons have only been detected in the muon spectrometer. Track parameters are obtained from the extrapolation to the beam axis taking material interactions into account.

- *Segment-tagged* muons are constructed from tracks in the inner detector for which at least one matching track segment in the muon spectrometer can be found.

- *Calorimeter-tagged* muons are created from inner detector tracks with a matching minimizing ionizing particle signature in the calorimeter. The purity of these muons is the least of all types but it can help compensate uncovered space in the muon spectrometer.

CHAPTER 4

Phenomenology

4.1. Modelling of Particle Interactions

The transfer from the purely theoretical Standard Model predictions to measurable quantities plays a huge role for the vast majority of analyses. This process involves a number of steps where different theoretical aspects come into play. The final result of this process which is described in this chapter are Monte Carlo simulation samples that have undergone the same event reconstruction mechanism that is also applied to measured data. These samples are generated for each background and signal hypothesis separately and combined to reproduce the expected data distributions.

4.1.1. Interaction Matrix Element

The first step when simulating particle interactions is the calculation of the matrix element which describes the transition from the particular initial to final state. This matrix element refers to the *hard process*, i.e. the actual process of interest without taking into account any additional decays or radiated particles. One simple example of a Feynman graph representing a scattering process is shown in figure 4.1, depicting an electron-positron annihilation process (x-axis corresponds to time). A virtual photon is created which in turn decays to a muon anti-muon pair. Evaluating the features of the diagram, i.e. incoming, outgoing and internal lines with their respective line styles as well as the vertices (see any quantum field theory textbook for details, e.g. [51]) leads to the matrix element for this leading-order process:

$$\mathcal{M} = -\frac{e^2}{Q^2} g_{\mu\nu} \left[\bar{v}(p_2)\gamma^\mu u(p_1)\right] \left[\bar{u}(p_3)\gamma^\nu v(p_4)\right] . \tag{4.1}$$

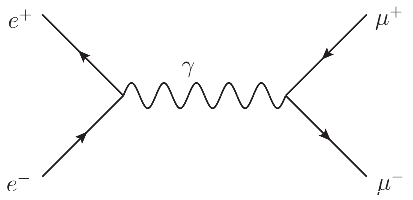
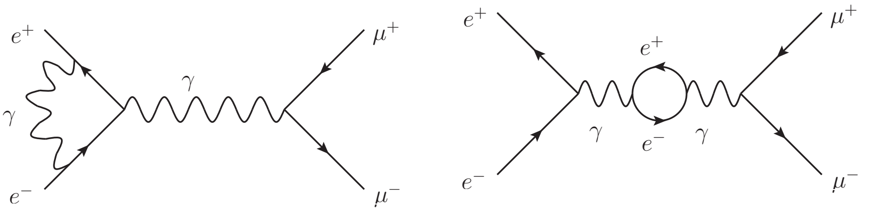


Figure 4.1.: Exemplary Feynman graph of e^+e^- annihilation to a photon decaying to two muons.



(a) Diagram including an additional photon exchange in the initial state.

(b) Diagram including a virtual electron positron loop within the photon propagator.

Figure 4.2.: Exemplary NLO (next-to-leading order) Feynman graphs of e^+e^- annihilation to a photon decaying to two muons.

In this expression, Q is the momentum transfer in the process, $g_{\mu\nu}$ the Minkowski tensor (both originate from the photon *propagator* in the centre), $u/v(p_x)$ are (incoming) initial state particle/antiparticle with assigned momenta p_x and $\bar{u}/\bar{v}(p_x)$ are the corresponding (outgoing) final state particles/antiparticles. The squared electron charge e and the Dirac matrices γ^ζ are contributions from the QED interactions when the gamma is created and destroyed.

Obviously the shown diagram in figure 4.1 is not the only process which yields the same signature. It only represents the leading order of processes with the given characteristics and requires corrections from higher order processes in nature. One example is shown in figure 4.2 showing a *next-to-leading order (NLO)* process with one additional photon exchange in the initial state. Thus, the technique of using Feynman diagrams for illustrative purposes as well as calculations also extends to the mathematical description of particle interactions as a series expansion of ascending orders. The number of diagrams grows combinatorially and makes computation of higher order corrections very complex and computationally expensive. In most cases (except soft QCD processes, see section 2.1.3) contributions from higher orders get smaller in each order since the coupling factors for the involved interactions appear with higher magnitudes. Like this, the series expansion approach is a legitimate treatment of describing the interactions and furthermore allows estimating the impact of higher order calculations or applying coarse corrections, respectively. Nevertheless, progress made in the theoretical tools as well as computational power makes it possible to produce Monte Carlo samples at higher orders which allows using NLO samples for the signal samples in the analysis presented in this thesis.

Based on the matrix elements whose derivation is often complex and requires a lot of mathematical subtleties (an introduction is given for example in [10]) it is possible to proceed to the initial generation of simulated particles. The relation between the (differential) cross section of a process and the matrix element is given based on *Fermi's Golden Rule No. 2* [52], [51]:

$$\Gamma_{fi} = 2\pi |T_{fi}|^2 \rho(E_i). \tag{4.2}$$

Γ_{fi} is the total transition rate, T_{fi} is to first order identical to $\langle f|\hat{H}'|i\rangle$ with $\hat{H}'$ as the perturbation Hamilton operator of the interaction (furthermore $T_{fi} \propto \mathcal{M}_{fi}$) and $\rho(E_i) = \left|\frac{\mathrm{d}n}{\mathrm{d}E_f}\right|_{E_i}$ is the density of states in the final state. In the latter expression, $\mathrm{d}n(E_f)$ is the differential of the number of states in the given final energy regime. Applied for the e^+e^- annihilation process, the final expression for the differential cross section in the centre-of-mass frame after integrating all terms of the matrix element can be obtained as [51]

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{1}{64\pi^2 s} \frac{p_f}{p_i} |\mathcal{M}_{fi}|^2. \tag{4.3}$$

In this expression, s is the centre-of-mass energy squared and p_{fi} are the magnitudes of the momenta of the initial- and final-state particles in the centre-of-mass frame. A simple extension to take into account some of the NLO corrections when a NLO matrix element calculation is not (yet) possible are the so-called K-factors. These are calculated as the ratio of the NLO over the LO cross section. Nevertheless, this method imposes additional consistency requirements.

The cross section is linked to the rate at which the specific process occurs via the luminosity (see 3.1) in the following way [26]:

$$\frac{\mathrm{d}N}{\mathrm{d}t} = \mathcal{L} \cdot \sigma. \tag{4.4}$$

For the differential cross section this expression becomes

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{1}{\mathcal{L}} \frac{\mathrm{d}N}{\mathrm{d}\Omega \cdot \mathrm{d}t}. \tag{4.5}$$

Basically it is possible to create events using the Matrix element and luminosity in the allowed phase space (which may be constrained by cuts already on generator level) as long as the incoming particles are well-defined. Since the LHC collides protons containing a sub-structure of gluons and quarks, the initial state is not known a-priori and needs to be reproduced corresponding to the constituents' distributions.

4.1.2. Parton Distribution Functions

Parton Distribution Functions (PDFs) describe the distribution of probabilities of finding specific longitudinal momenta x of different kinds of partons (gluons and the various quark flavours) as a function of x and of the momentum transfer (or energy scale) Q. There are different sets of PDFs available which rely on different experimental inputs and on a variety of mechanisms for fitting, interpolation and uncertainty estimation.

Exemplary, figure 4.3 shows a set of distributions for the different partons of the particular momentum fraction weighted by the corresponding PDF value (= probability density of finding the specific parton at the given x value for a fixed Q scale) over the given x range for the *MSTW2008* set (see also [53]). Since protons are composed of two *up* and one *down* quark, this structure needs to reflect in the PDF functions, yielding the requirement that

$$\int_0^1 \mathrm{d}x \left[f_u(x,Q) - f_{\bar{u}}(x,Q)\right] = 2 \quad \text{and} \quad \int_0^1 \mathrm{d}x \left[f_d(x,Q) - f_{\bar{d}}(x,Q)\right] = 1. \tag{4.6}$$

Furthermore since the sum of all momentum fractions needs to yield the whole momentum the following must apply:

$$\sum_i \int_0^1 x f_i(x,Q) \mathrm{d}x = 1. \tag{4.7}$$

Due to the non-perturbative nature of QCD with respect to the description of partons inside a hadron and the limitations in concurrent lattice QCD calculations, parton distribution functions cannot be predicted entirely. In fact they are obtained as fits of parton models to experimental data that has been collected at various collider experiments as well as fixed target experiments in deep inelastic scattering (among others at the Fermilab Tevatron Collider, the DESY *ep* Collider (HERA) and most recently at the LHC itself).

Since measurements are only possible for certain Q^2 scales, it is necessary to extrapolate to the regime of interest. This can be achieved via the *DGLAP* equations, a set of differential integral equations which allow the calculation of $f_i(x, Q'^2)$ for any Q'^2 of interest, as long as the strong

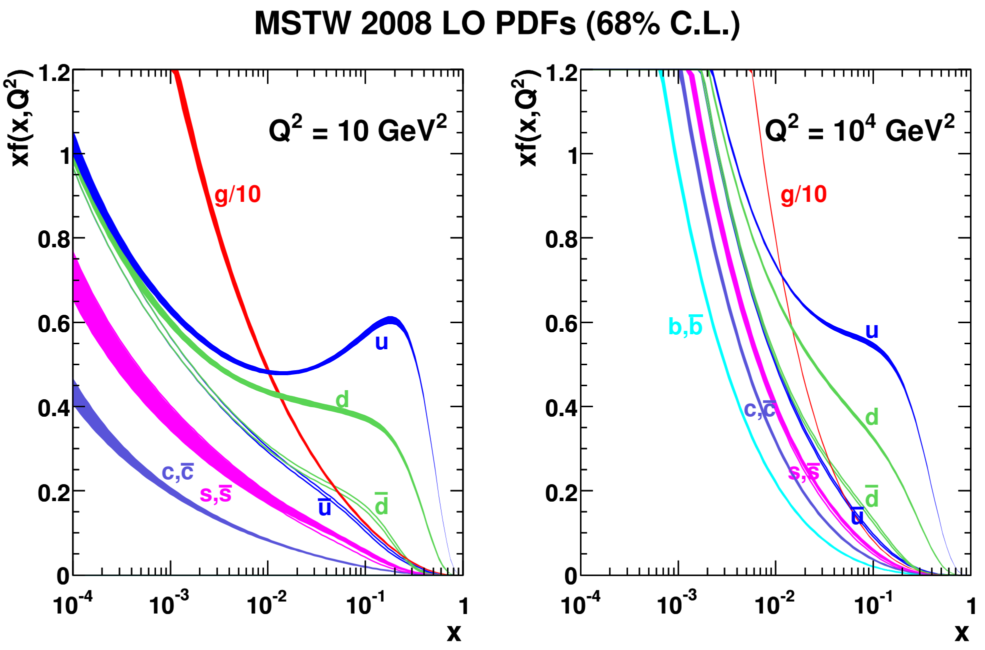


Figure 4.3.: Leading order parton distribution functions from "MSTW2008" PDF set for protons, showing the momentum fraction x weighted by its PDF value $f(x, Q^2)$ over a wide range of x for all kinds of partons at two different energy scales Q^2. The band width indicates the 68 % confidence level uncertainty. [55]

coupling constant $\alpha_s(Q'^2)$ is known. As a consequence of the QCD factorization theorem (originally suggested by Drell and Yan) which states that an overall hadronic cross section could be factorized into the hard process and the parton distribution functions which would be treated as weights. At leading order this can be expressed schematically[1] as

$$\sigma = \int \mathrm{d}x_a \mathrm{d}x_b f_{a,A}(x_a, Q^2) f_{b,B}(x_b, Q^2) \sigma_{ab \to X}. \tag{4.8}$$

In the above expression a, b, A and B represent the original hadrons (capital letters) and the colliding partons (lower case letters) respectively. $\sigma_{ab \to X}$ is the cross section of the hard process with a,b as incoming partons. More details also for higher order calculations can be found in [54].

4.1.3. Showering and Hadronization

The results of the hard process as described by the matrix element (see 4.1.1) exist on parton level. As a consequence of the colour charge of the outgoing partons, above a hadronization scale of order 1 GeV, additional gluons or quark antiquark pairs get emitted in collinear direction forming a parton shower. This is usually simulated by randomly creating these emissions whilst the scale Q is reduced in each subsequent step using a perturbative QCD description. Since this description is only valid for soft and collinear emissions, it is necessary to include the treatment of hard wide-angle emissions into the original matrix element of the hard process. Additional care must be taken in this case and in the case of an NLO matrix element to make sure the parton showers

---

[1] More precisely, Q^2 as the parameter of the parton distribution functions $f_{i,I}(x_i, Q^2)$ should be the factorization scale μ_f here and $\sigma_{ab \to X}$ the first term of a perturbation series in $\alpha_S(\mu_R^2)^n$ where μ_R is the renormalisation scale; both μ_R and μ_F can be chosen as $M_x \approx Q^2$ though.

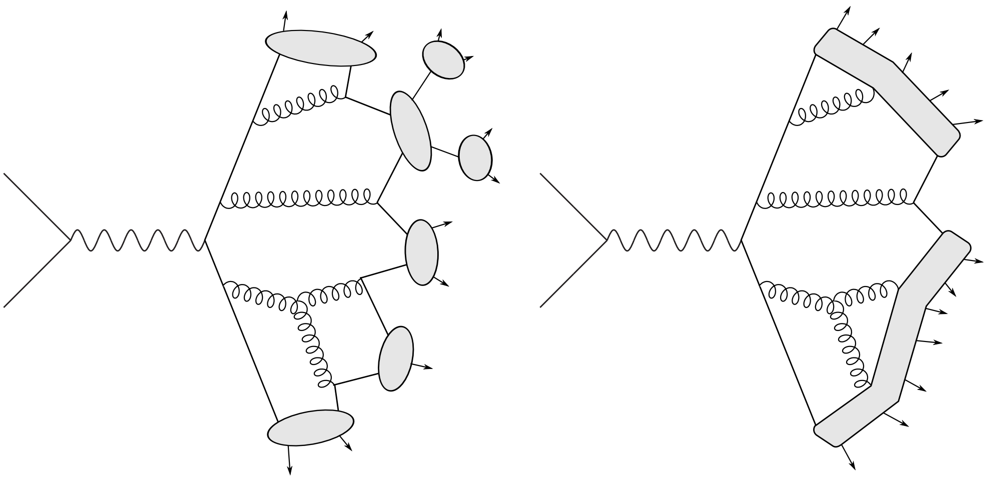


Figure 4.4.: Illustration of the two hadronization models presented, cluster model on the left and Lund string model on the right. (inspired by [58])

match the (multi-)partons in the matrix element calculation by removing the overlap in phase space between matrix element jets and the parton shower. Once the hadronization scale of $\sim$ 1 GeV is reached, the shower constituents start to form colourless hadrons. This energy regime also marks the transition to non-perturbative QCD processes. To simulate this process, mainly two phenomenological models are available which are on the one hand string models (most widespread the Lund string model, see e.g. [56]) and on the other hand cluster models (an early version described in [57]). While the string model is implemented in the PYTHIA generator, the cluster model is deployed in HERWIG and SHERPA. The subsequent description of both models is based on the descriptions in [14] and [58].

The cluster model is based on the concept of *preconfinement* (which can be motivated e.g. as shown in [59]). This means that the observation that colour-singlet $q\bar{q}$ states have a universal mass spectrum which does not depend on the current scale Q of the regarded process. In a first step all gluons in the showering output are split up into quark antiquark pairs meaning that there are discrete starting points for the subsequent process of forming hadrons (compare figure 4.4, left side). Afterwards, all quark antiquark pairs are combined into colour-singlet clusters. If the mass of the resulting cluster lies above a certain threshold (usually $\sim 3-4$ GeV) it gets split up into two clusters (i.e. two quark antiquark pairs) along the axis between both constituents. Some clusters with the lowest masses directly decay to single hadrons whilst spare momentum is transferred to neighbouring clusters. In the last step the remaining low-mass clusters (ignoring the already formed hadrons) are treated as excited mesons which isotropically decay to two hadrons each wherein spin and available phase space steer the final configuration. In this final hadronization step, all uds hadron multiplets including their excited states are formed corresponding to the respective branching fractions to maintain the expected $SU(3)$ flavour symmetry.

Downsides of the cluster model are problems that occur when dealing with very heavy clusters at the beginning. Furthermore the fraction of baryons and heavy quarks produced is overestimated.

The (Lund) string model employs a more continuous description of the hadronization process. The starting points for this algorithm are all colour-connected $q\bar{q}$ pairs emerging from the shower in the previous step (see 4.4, right side). For each of these, as the distance between the partons grows a virtual colour flux 'tube' is stretched against a potential $V(r) \propto \kappa r$ between the partons. Once the distance between the two partons is large enough such that the potential energy is sufficient to create a new $q\bar{q}$ pair, the string breaks up and the new pair is formed. In this model, gluons from the shower are treated as kinks in the strings at a transversal offset. Due to causality, the string breakups can be treated independently, wherein the Lund model starts forming the outermost hadrons and alternately iterating towards the string centre. For each breakup vertex the resulting quark masses and transverse momenta are set based on two Gaussian distributions whose parameters can be adjusted. In order to enable the production of baryons, it is possible to allow the creation of *diquarks*[2] in string breaks. In the last step, all quarks and diquarks are combined into hadron multiplets. The downsides of this model are problems related to the formation of baryons which are smaller than for the cluster model, though.

Underlying Event

Besides the (dominating) hard process and its associated initial and final state radiation, further interactions take place that are caused by the proton remnants. This effect is referred to as the *underlying event*. It is assumed that further colour exchanges between the colliding particles play a key role in the genesis of underlying event activity. Comparisons to minimum bias events (which are triggered at very low activity) at the same centre of mass energy but without a hard process occurring have shown that the activity from the underlying event outruns the one in these minimum bias events which is referred to as the "jet pedestal“ effect. Since the description of the underlying event relies on low energy QCD calculations the modelling is also performed solely on a phenomenological basis.

4.1.4. Detector Simulation

The last step in the event simulation process is the detector simulation. After this step the generated events can be treated like recorded data and be processed with the standard event reconstruction software (see section 3.3). For simulation runs a simulated first level trigger is applied for categorization purposes. In the subsequent simulation step the interaction of all generated "stable"[3] particles with the detector material and the active components is simulated. In order to achieve a high accuracy, a precise model of the material within the detector volume and of the response of every detector component is required. The actual simulation of the propagation through the detector and the interactions with the detector material and its active components is performed using GEANT4[60].

The simulated particles passing through the active detector components create virtual energy deposits. These are the base for the generation of detector "hits" for which additional effects are considered to reproduce real world data taking conditions. This comprises particles from different additional origins (see section on pile-up below) as well as detector noise. Taking all these overlaid effects into account the detector response is digitized and passed on to the reconstruction algorithms.

An overview of the overall process including the GEANT4 application is given in [61]. This document also goes into details about the ATLAS software framework Athena which controls the data flow and calls the relevant tools for every simulation and reconstruction step.

Pile-up

The term *pile-up* means additional activity inside the detector which is not associated to the primary vertex, i.e. the most energetic proton-proton collision in an event. The most prominent sources of pile-up are further proton-proton collisions taking place during the same bunch crossing. The average number of interactions per bunch crossing in 2012 was 20.7, ranging up to about

---

[2] A diquark is a hypothetical sub-constituent of baryons which would accordingly consist of a quark and a diquark (i.e. a group of two quarks).

[3] The term *stable* particle refers to a particle that does not decay promptly and whose decay is therefore not handled directly by the event generators.

40 interactions in rare cases[30]. As this type of pile-up occurs within the same bunch crossing it is referred to as *in-time pile-up*.

A second prominent pile-up contribution is called *out-of-time pile-up* and arises from collisions in preceding or succeeding bunch crossings. Different types of detectors have dead times, acceptance times above the bunch crossing time and longer integration times such that collisions in neighbouring bunch crossings affect the recorded hits and energies in different ways.

Both, the in-time and out-of-time pile-up are added during the Monte Carlo sample production and considered during the detector reconstruction process to take the arising distortions into account.

Further effects contributing to the total amount of pile-up that are often not part of the simulated samples (e.g. because they only affect specific processes or analyses) are cavern background (thermal neutrons and protons creating random hits), beam halo events (particles from interactions with up-stream collimator elements) and beam gas events (interactions of protons with residual gas in the beam pipe)[62].

4.2. Monte Carlo Event Generators

The following list features a brief description of all Monte Carlo event generators that have been used to produce event samples or input datasets (such as normalisations or distributions of specific variables like the Higgs boson p_{T}).

AcerMC [63] is an event generator providing Standard Model background sample generation functionality for pp collisions at the LHC. It is solely used to generate $tq\bar{b}$ background samples for the analysis and only provides leading-order accuracy.

The POWHEG BOX [64] is an implementation of the *POWHEG* (Positive Weight Hardest Emission Generator) method allowing to automatically combine NLO (QCD) calculations with parton shower simulations. It provides a number of processes and is widely used both for the SM Higgs signal process and numerous background processes.

Alpgen [65] focuses on the creation of multi-parton hard processes. It only supports leading order calculations for QCD and electroweak interactions. For the presented analysis Alpgen is used to model the $W\gamma$ process with one misidentified photon and the $Z/\gamma*$ Drell-Yan process each with up to five additional jets in the (tree level) matrix element.

Sherpa [66] ("Simulation of High-Energy Reactions of PArticles") is a general purpose event generator written in C++ for a wide range of collision types including the LHC configuration of two high-energetic proton beams. It comprises two matrix element calculation tools and supports (besides SM processes) the MSSM (Minimal Supersymmetric Standard Model), Anomalous Gauge Couplings and Minimal UED (Universal Extra Dimensions). It furthermore contains tools for parton showering plus merging, hadronization (using the cluster model) and subsequent hadron and tau decays, treatment of beam remnants, addition of QED radiation and multi-parton interactions.

GG2VV [67] is a specialized generator focusing only on gluon fusion processes with two vector bosons in the final state, optionally via a Higgs boson. In this analysis it is used to generate the non-resonant $gg \to WW$ process at leading order including the interference with $gg \to ZZ$.

Pythia [68, 69] is a general purpose event generator covering a lot of aspects of the production process. Besides the hard processes that are provided it can simulate initial and final state radiation and showering, matching and merging of the hard process and the shower and furthermore supports multi-parton interactions, the description of beam remnants, string fragmentation and particle decays. These features make Pythia one of the generators of choice for processing parton level outputs generated from more specialized hard process event generators to run the showering and hadronization. Since the transition between Pythia versions 6 (written in FORTRAN77) and 8 (C++) is ongoing and the feature set differs between both versions, both versions have been employed for a number of background samples in this analysis.

Herwig [70] is a general-purpose event generator capable of modelling hard scattering processes of leptons or hadrons (also combined). It comes with a subset of processes already implemented and the possibility to interface to more specialized matrix element generators. BSM models are partly available in the generator or can be added by hand. A key feature is the included parton shower module for simulating initial and final state QCD radiation, especially that from heavy particles. It supports the simulation of the underlying event, of the hadronization process (applying the cluster model) and it models subsequent hadronic decays as accurately as possible via the respective matrix elements and further corrections.

Jimmy [71] is an extension to Herwig specialized on modelling underlying event particles. For this, it supports multiple parton scattering simulations e.g. in hadron-hadron collisions.

MadGraph5_aMC@NLO [72] is the successor of previous MADGRAPH5 and aMC@NLO versions combining the respective development efforts. It is a framework consisting of different modules for cross section calculations, for the generation of events from hard processes including the parton shower matching, matrix element computation at tree and one loop level and furthermore provides auxiliary tools. For all Standard Model processes as well as the set of study cases of non-Standard-Model Higgs resonances NLO calculations of QCD corrections are supported. For any user-defined Lagrangian leading order computation is possible.

HRes2.1 [73, 74] is a tool dedicated to the calculation of the Standard Model Higgs boson cross section. It features calculations in NNLO in QCD perturbation theory and NNLL precision in for low-p_{T} contribution resummation. Its accuracy in modelling the Higgs p_{T} spectrum exceeds the one of POWHEG or MADGRAPH5_aMC@NLO such that the events generated with the latter are reweighted using HRES results.

MCFM [75] is a tool for computing cross sections for a set of processes at hadron colliders. It supports NLO calculations with full spin correlation modelling.

Top++2.0 [76] provides methods to numerically compute the total inclusive cross section for the $t\bar{t}$ production at hadron colliders. It offers very precise modelling results with NNLO accuracy in the hadronic cross section calculation and NNLL accuracy for the resummation.

DYNNLO [77] is a specialized Monte Carlo tool providing NNLO (in QCD perturbation theory) cross section computations for vector boson production at pp and $p\bar{p}$ colliders. Among its features are the modelling of the Z/γ interference, of finite-width effects and of vector bosons decaying to leptons including their spin correlation effects.

JHU [78–80] is a Monte Carlo event generator specifically targeting the simulation of different spin and parity models of resonances. As a software package it also contains the so-called "MELA" module allowing for stand-alone matrix element calculations as applicable for matrix element analyses.

4.3. Higgs Phenomenology

An introduction to the Higgs mechanism and the theoretical background is given in 2.1.4. This section describes the practical aspects of the Higgs physics for the current analysis. It covers the predicted production and decay modes with a special emphasis on gluon fusion and the WW decay channel. Afterwards the additional considerations about non-SM Higgs models that have been tested (i.e. with different spin and/or parity) are covered.

4.3.1. Production and Decay Modes

Higgs Production

According to the Standard Model, gluon-gluon fusion is the dominant Higgs production mode at the LHC [81]. This is the case even though the direct coupling of the gluons to the Higgs is not possible such that a heavy-quark loop (clearly dominated by the top quark) is required to bridge the gap. An overview of Feynman diagrams describing the four most prominent production modes is given in figure 4.5. These four production modes are listed below:

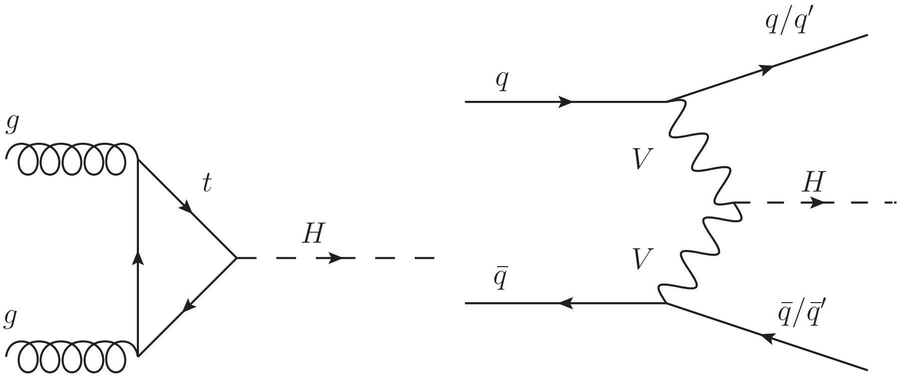
- **Gluon-gluon fusion** (ggF) is as mentioned above the dominant production mode and requires two incoming gluons to interact via a heavy quark loop which may generate a Higgs boson. Since the coupling strength of the Higgs boson depends on the fermion mass, the top quark is strongly favoured and lighter quarks flavours are suppressed proportional to m_q^2.

- **Vector boson fusion** (VBF) has the second highest cross section at the LHC and occurs when two (anti)quarks each radiate an oppositely charged W boson or a Z boson which merge into a Higgs boson. The quarks which radiated the vector bosons get deflected as a consequence of the momentum transfer and form hard forward jets (i.e. jets boosted parallel to the beam axis) in the calorimeters which lead to a rather clear signature of the process.

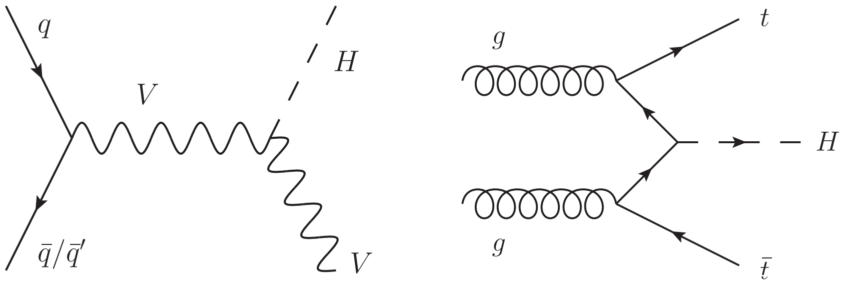
- **Higgsstrahlung** (VH) (or associated production with a vector boson) is the third most likely production mode. A vector boson created in a scattering process of an incoming quark anti-quark pair radiates a Higgs boson. Due to the high mass of the vector boson which needs to be produced in addition to the Higgs boson, this channel is suppressed with respect to the previous one.

- **Associated production** (ttH, bbH) with a $t\bar{t}$ (or $b\bar{b}$) pair is structurally similar to the gluon-gluon fusion process since here gluons form the incoming particles as well. Since the top quarks exist not only in a virtual loop but also as two outgoing real particles, this process is strongly suppressed for kinematic reasons. The required energies to reach the production threshold for this process are nearly three times higher than for the gluon-gluon fusion process.

Table 4.1 lists the predicted production cross sections and the associated uncertainties from PDF plus α_S and QCD scale uncertainties for the dominant modes for a Higgs mass of 125.0 GeV at $\sqrt{s} = 8$ TeV.



(a) Gluon-gluon fusion **(b)** Vector boson fusion



(c) Higgsstrahlung **(d)** Associated production

Figure 4.5.: Overview of exemplary leading order diagrams for the four most prominent Higgs production modes at the LHC, arranged in descending order of their cross section. V may stand for W or Z; the change of the quark flavour in case of W bosons is indicated by an apostrophe.

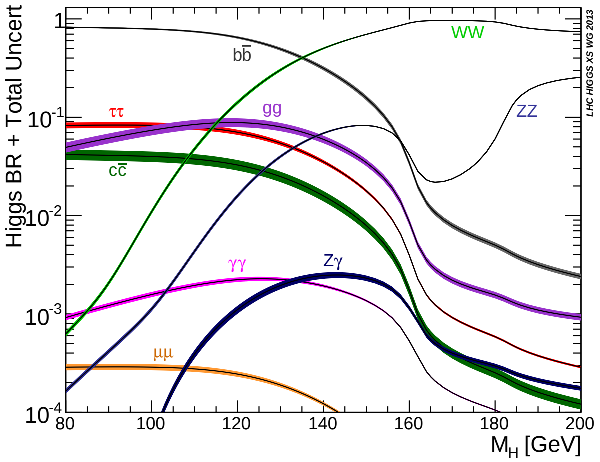


Figure 4.6.: Branching ratios for the Standard Model Higgs boson decay including total uncertainties over a hypothetical Higgs mass range of 80–200 GeV. (from [83])

Table 4.1.: Table of predicted Higgs production cross sections for a Higgs mass of 125.0 GeV at $\sqrt{s} = 8\,\text{TeV}$. Each row lists the process, the accuracy of the computation and the cross section with its absolute error which is composed from PDF plus α_S and QCD scale uncertainties. The terms 4FS and 5FS refer to four- and five-flavour scheme calculations, see section 12.2.2. in [1] for details. All numbers are taken from [82].

Process	Accuracy	Cross section / pb
ggF	NNLL QCD and NLO EW	$19.27^{+2.00}_{-2.01}$
VBF	NNLO QCD and NLO EW.	$1.58^{+0.04}_{-0.04}$
WH	NNLO QCD and NLO EW	$0.70^{+0.02}_{-0.02}$
ZH	NNLO QCD in qq/gg→ZH and NLO EW	$0.42^{+0.02}_{-0.02}$
ttH	NLO QCD	$0.13^{+0.01}_{-0.02}$
bbH	5FS (NNLO) and 4FS (NLO)	$0.20^{+0.02}_{-0.03}$

Higgs Decay Modes

Figure 4.6 shows the corresponding theoretical branching ratios for the Higgs decay over a hypothetical Higgs mass range between 80 and 200 GeV. The coloured bands indicate the total uncertainty of the theory prediction for each channel. As the coupling strength of other particles to the Higgs boson increases with their mass, decays to heavy fermions and vector bosons dominate in spite of the a smaller available phase space. At the reconstructed Higgs mass of 125 GeV the dominant observable decay mode is $b\bar{b}$ with a branching fraction of 57.7 % followed by WW^* with 21.5 %, $\tau\tau$ with 6.3 %, ZZ^* with 2.6 %, $\gamma\gamma$ with 0.23 % and further even rarer decay modes[81].

Since the Higgs boson lifetime is very short (expected to be 1.6×10^{-22} s with a measured upper bound of 1.9×10^{-13} s at 95 % confidence level, [84]) and also the lifetimes of the direct decay products is short, in most cases only the subsequent decay products can be reconstructed. This applies as well for the $H \to WW^*$ channel where three configurations in the final state are possible: $WW \to \ell\nu\ell\nu$, $WW \to \ell\nu qq$ and $WW \to qqqq$. A large QCD background makes the (semi-)hadronic W decays hard to reconstruct such that in spite of the neutrinos in the final state the first option is most suitable for Higgs analyses.

4.3.2. Higgs Discovery and Measurements

Since the initial discovery of the new resonance in 2012 which was later confirmed to be the Higgs boson, a large amount of analyses have been conducted to investigate its properties. These include the Higgs mass, the couplings to other particles, the lifetime (i.e. total resonance width) and the particle quantum numbers.

The best available measurements of the Higgs mass and its couplings to the other Standard Model particles are given in two combined publications by ATLAS and CMS. From these, the Higgs mass has been determined as $m_\text{H} = 125.09 \pm 0.24\,\text{GeV}$. The single contributions that were considered to obtain this result are shown in figure 4.7. The couplings to fermions and vector bosons were found to be compatible with the Standard Model expectation, as shown in figures 4.8a and 4.8b.

Initial studies of the spin and in parts of the parity of the new boson were published soon after the discovery starting by the end of 2012. This comprises first CMS publications in the WW and ZZ channel[85, 86] and a combined ATLAS paper[3] also including first diphoton results. At this point nearly all of the tested non-SM models have been excluded at high confidence levels.

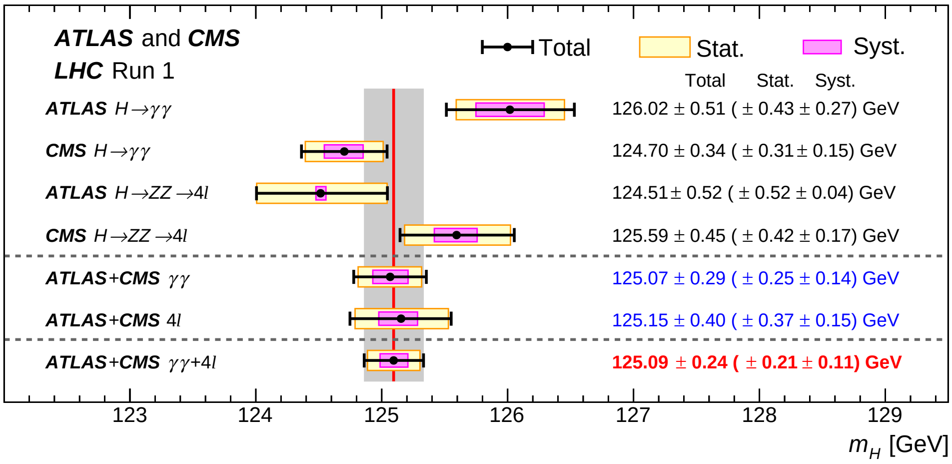
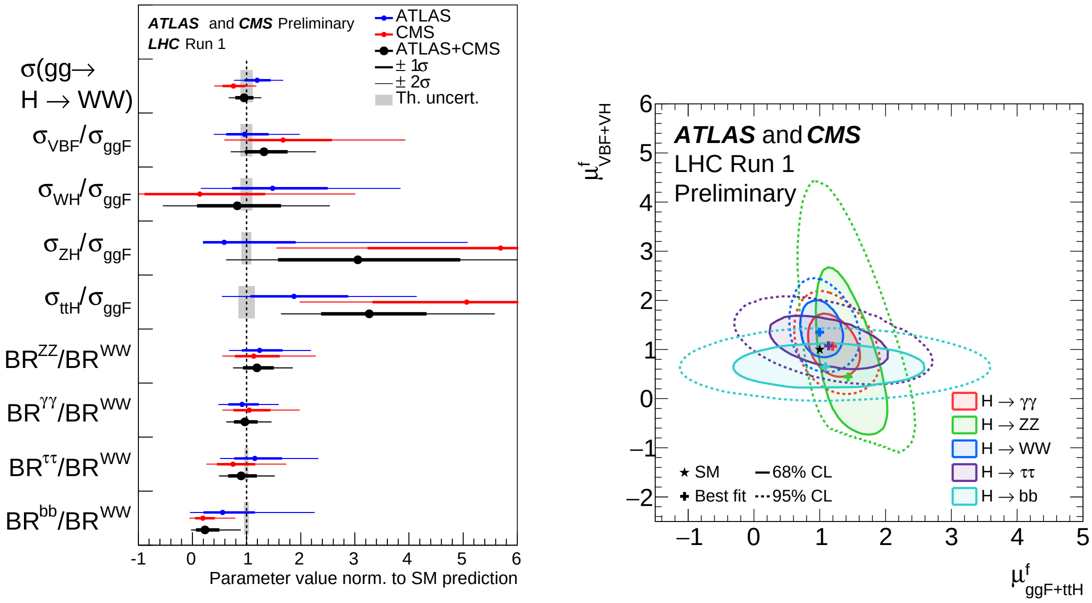


Figure 4.7.: Mass measurements from sensitive channels in ATLAS and CMS plus their combination. The vertical red line with the surrounding grey area corresponds to the central value and the total uncertainty of the combined measurement (bottom row). For each data point the respective statistical (yellow) and systematic (magenta) error bands are shown below the data points and the total uncertainty (black). (from [87])

The current analysis is part of the second iteration within ATLAS broadening the analysis scope in terms of the tested models and improving the analysis strategy.



(a) Plot of cross sections and branching ratios with respect to the $gg \rightarrow H \rightarrow WW$ decay channel combined from ATLAS and CMS measurements. In addition the single experiment results are shown. Each point indicates the respective best fit values with their one σ (thick line) and two σ (thin line) uncertainty ranges. Every data point is given relative to the Standard Model expectations for which the grey areas display the corresponding theory uncertainties.

(b) Higgs couplings plot from ATLAS and CMS. Shown are 68 % (solid line) and 95 % (dashed line) likelihood contours in the $\mu^{f}_{\mathrm{ggF+ttH}} \times \mu^{f}_{\mathrm{VBF+VH}}$ plane with their respective best fit values (cross). The μ values correspond to the signal strengths observed in fermionic (ggf, ttH) and vector boson (VBF, VH) production modes. The index f represents the decay channels the coupling parameters have been tested in. The Standard Model expectation (asterisk) lies within the 68 % confidence level contours for each decay mode.

Figure 4.8.: Couplings measurement results from ATLAS and CMS combined (from [88]).

CHAPTER 5

Mathematical Tools

In order to make deductions from measurements, the expectations from theory in the form of Monte Carlo event samples for background and signal processes need to be matched to experimental data. In this process, the impact of systematic uncertainties and Monte Carlo sample normalizations are considered. By optimizing the agreement between the modelled scenario and data, a best fit result including an uncertainty can be obtained.

Classically the final fit of Monte Carlo samples to data would rely on one or more sensitive variables that have gone through a list of cuts to optimize the background rejection with respect to signal events. However, multivariate analysis techniques allow for the usage of one single quantity in the final fit which condenses the information contained in a set of input variables including their respective correlations. Additionally the task of optimizing cuts to increase the signal to background ratio can be supported by a multivariate algorithm. One tool that is widely used in high energy physics analyses and is therefore well-tested and understood are *boosted decision trees* (BDTs). They combine simple decision trees with limited capabilities and the mechanism of boosting; both of which will be discussed in the following section.

5.1. Boosted Decision Trees

Decision trees are a popular tool developed by machine learning researchers with applications in a wide range of fields, such as finance, marketing, engineering, medicine (see [89], p. 6) and (high energy) physics. The decision trees used here are binary trees with one criterion (cutting value for one variable) applied per node. Starting from a *root node*, each node can have up to two children (*left* and *right*) which again hold cut decisions such that each *leaf node* (nodes without children) can be reached via one specific sequence of cuts that are represented by preceding internal nodes of the tree. When constructing the decision tree, the best separation criterion between background and signal events (i.e. the best combination of a specific cutting value and variable) is evaluated and used for each node. Starting from the first single (= root) node, this optimization is performed in each subsequent step building up the entire decision tree. This process is illustrated with the help of a simple example following the analysis use case in figure 5.1 showing one resulting tree and its correspondence in the (simplified) underlying variable space.

The process terminates when either the number of events in a node falls below a minimum or the maximal tree depth is reached. One common quantity (used by default in the TMVA framework [90]) for measuring the separation power is the Gini index. It is defined as $p \cdot (1 - p)$ with $p =$

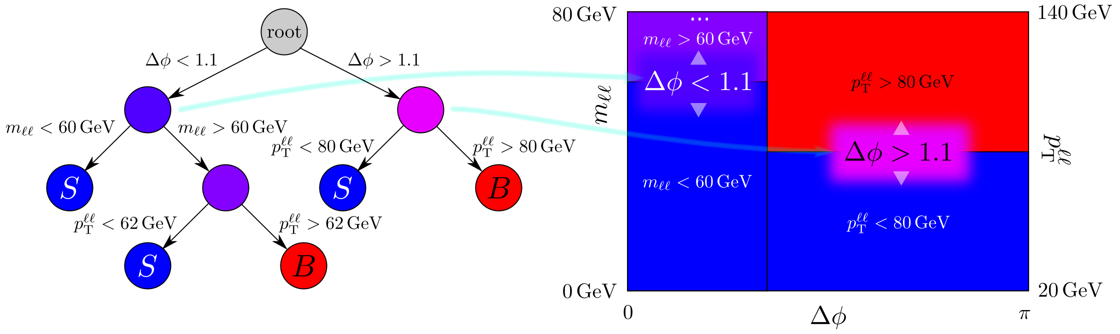


Figure 5.1.: Simplified example of one single decision tree with the tree structure including cut decisions and signal/background markers on the left and an illustration of the 'variable plane' on the right. The light blue arrows highlight the correspondence between tree nodes and variable ranges. Additionally corresponding nodes (left) and separators/regions (right) are coloured identically. The third layer of nodes is not represented in the right plane for clarity reasons. All labels on the right indicate the cut which led to the corresponding node/selection. While the root node would correspond to the entire variable plane (no cut applied), the first step splits up the events by their $\Delta\phi$ value (see horizontal axis on the right). The subsequent cut is applied on different variables in both branches. The left child node holds a cut on $m_{\ell\ell}$ (see left vertical axis of the plane) which corresponds to the split on the left region of the variable plane. The right child holds a cut on $p_T^{\ell\ell}$ (see right vertical axis) which corresponds to the split in the right region of the plane. The leaf nodes are labelled according to the majority of events falling into the respective event class.

$\frac{\text{\# signal events}}{\text{\# all events}}$ (both per node) representing the signal purity. More precisely the gain in separation power comparing the Gini index of the parent node to the sum of the indices of the child nodes weighted by their event fractions is considered for the decision on variable and cut choice.

Various methods can be applied in order to cope with instabilities arising from statistical fluctuations in the training sample which is used to construct the decision tree and in order to increase the classification performance for more complex signatures. In a comparison of different techniques, for cases of a sufficiently clear classification (as signal or background) it has been shown that boosting is able to properly fix this issue [91].

Boosting means the creation of not only one decision tree but hundreds of trees with assigned weights. All these trees contribute (according to their weight) to the classification of events. Generally speaking *boosting* means applying the decision tree construction algorithm to reweighted training samples.

There is a variety of boosting methods and variations available within the TMVA framework. The widely used default option is called *AdaBoost* for *Adaptive Boost*. Its main principle is to reweight the events of the training sample such that those events which are misclassified by all existing trees are assigned a higher weight for the next iteration of tree-building. More details on this algorithm can be found in [92] in the respective section. For the analysis presented it was decided to use the *Gradient Boosting* algorithm instead which is illustrated in the following.

Basically one can think of the mechanism of gradient boosting as a function expansion[1] with the single decision trees representing the expansion terms. In every step a loss function comprising the collection of existing trees and the new testing tree gets minimized to determine the optimal tree configuration. In this sense the gradient boosting algorithm can be categorized as a 'greedy' algorithm [93], i.e. an algorithm which attempts to approximate a global minimum by choosing the optimal solution in each optimisation step.

---

[1] see class description in `http://root.cern.ch/root/htmldoc/TMVA__MethodBDT.html`

The evaluation of any tree that is to be added to the ensemble of existing trees is based on a loss function, which is by default given as

$$L(F, y) = \ln\left(1 + e^{-2F(x)y}\right). \tag{5.1}$$

Here, $F(x)$ is the model response for the data vector x belonging to one event and y is the true (classification) value. Herein a value of -1 represents a pure background classification, $+1$ a pure signal classification.

The key part for the boosting step is the differentiation of this function with respect to the classification function comprising the ensemble of existing trees (here for one particular event with data vector x):

$$\left.\frac{\mathrm{d}L(F(x), y)}{\mathrm{d}F(x)}\right|_{F=F_{\text{ensemble}}} = \frac{2y}{1 + e^{-2yF_{\text{ensemble}}(x)}}. \tag{5.2}$$

The evaluation of the derivative for the ensemble of existing trees yields a target value for each training event in the dataset. In a subsequent step, a new *regression* tree—a tree with real number qualifier values assigned to leaves instead of a clear signal or background classification—is grown and fitted to this target value. The response values of each leaf of the newly created tree are then adjusted to the mean gradient value in the specific region which is defined by the leaf hyperrectangle. This approach has the effect that the tree ensemble tends to enrich signal or background in the corresponding regions. The final BDT response can be generated simply from the normalized sum of all regression values in the ensemble. A detailed description of the *Gradient Boosting* algorithm is given in [94].

According to the TMVA manual, Gradient Boosting works best when combined with *weak classifiers*, which means for the case of BDTs (the same boosting technique can be applied to other analysis techniques), small trees with a very limited depth (2-3). Using a sufficient number of trees (i.e. boosting steps) these perform sufficiently well while the danger of overtraining is minimized.

Overtraining is an effect which can occur especially with small data sets compared to the number of degrees of freedom of a BDT (or any machine learning algorithm). In such a case the training may use specific artefacts of the training sample as separating characteristics. Overtraining manifests itself as an apparent increase of the BDT classification power with respect to the real effectiveness. In order to test the extent of overtraining, the available data set is split into a training sample and a test sample (both with identical sizes for this analysis). The BDT response is evaluated for both samples and the resulting distributions are compared e.g. by computing the KS test value (Kolmogorov-Smirnov test, see [95]) and applying an acceptance limit.

Overtraining avoidance can still be improved by introducing a `Shrinkage` parameter which gets multiplied with the average gradient that is stored in each leaf. Like this the shrinkage parameter slows down the learning process which can help to introduce more stability in certain settings.

Compared to the *AdaBoost* algorithm one gains more stability with respect to outliers and misclassification from wrong labelling. Additionally it would be possible to use any differentiable loss function with the algorithm. The Gradient Boost algorithm is on the other hand less straightforward and has a higher computational complexity.

5.2. Statistics

The final analysis results are obtained via a frequentist statistical method. An introduction to the statistical foundations required for the analysis is given in this section.

The main focus is put on hypothesis testing, i.e. considering two hypotheses and calculating the agreement or rejection, comparing them to measured or simulation data. One defines a *null hypothesis* (H_0) which corresponds to the known (or expected) outcome (i.e. only backgrounds or in this case the Standard Model signal plus background expectation). The alternative hypothesis (H_1) is defined as the alternative signal model for this case or for other analyses as the background plus signal hypothesis which is to be tested. To quantify the outcome of a hypothesis test, a p-value can be computed. It represents the probability of obtaining an experimental outcome which is less compatible with the tested hypothesis than the (measured) value used to compute the p-value. In other words a low p-value corresponds to a high rejection significance of the assigned hypothesis.

The starting point for the statistical evaluation of the analysis outcome is the formulation of a likelihood which incorporates all aspects of the measurement (i.e. all contributions of signal and background samples in all regions and all systematics). Once this likelihood has been constructed, a viable test statistic is established. In the subsequent step the test statistic in combination with the overall probability density function (which corresponds to the likelihood function defined in the first step) is used to obtain expected and observed sensitivities and limits.

An in-depth overview of the methodology can be found in [96]. Given the nature of the analysis presented the interpretation of the parameter is modified with respect to this description; instead of the signal strength over the sum of backgrounds, the relative fraction of two signal samples is used. The statistical treatment is analogous to the one described in the above publication, though.

5.2.1. Likelihood Function

Generally speaking the overall likelihood can be expressed as a product over Poisson distributions. The parameter that is common among most analyses is the signal normalisation μ, so the simplest version of a likelihood function describing a measurement with N observed events can be written as

$$\mathcal{L}(\mu) = P(N|\mu S + B) \tag{5.3}$$

with S and B representing the *number* of expected signal and background events. The signal normalisation μ is defined such that $\mu = 0$ corresponds to the background-only hypothesis and $\mu = 1$ to the nominal signal hypothesis (often the expected signal strength from Standard Model predictions). The above formula is associated with the so-called *signal region* which is the main region of each analysis since it contains the relevant signal contribution. To improve the understanding of backgrounds and obtain an indicator on their normalisation a common approach is to introduce *control regions*. These additional regimes within the variable space are chosen such that a certain background is largely enriched within one control region. Their selection is chosen to be orthogonal to the signal region. This means that it mostly corresponds to the one of the signal region besides the adjustments with respect to the specific background enrichment, i.e. by inverting one cut with respect to the signal region. Extended by control regions the above expression can look like

$$\mathcal{L}(\mu, \theta) = P(N|\mu S + \theta B) \cdot P(N_{\mathrm{CR}}|\theta B_{\mathrm{CR}}). \tag{5.4}$$

The first part represents as before the signal enriched regime. The newly introduced variable θ which appears in both parts of the expression now is a nuisance parameter controlling the background normalisation. It is constrained by the second part of the expression via the additional measurement in the control region. N_{CR} and B_{CR} stand for the measured number of events and the background expectation in the control region.

Expressed as a product of Poisson distributions for binned datasets instead of single values (and implicitly extending the parameter θ to a vector of nuisance parameters $\boldsymbol{\theta}$) the above expression leads to

$$\mathcal{L}(\mu, \theta) = \prod_{j=1}^{N} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} \cdot e^{-(\mu s_j + b_j)} \cdot \prod_{k=1}^{M} \frac{u_k^{m_k}}{m_k!} \cdot e^{-u_k}. \tag{5.5}$$

In this example, each of the parameters s_j, b_j and u_k depends on a set of nuisance parameters within $\boldsymbol{\theta}$. N denotes the number of bins in the signal region histogram with entries n_j and the expectation value

$$E[n_j] = \mu s_j + b_j. \tag{5.6}$$

Equivalently, M stands for the number of bins in the control region histogram with expectation values

$$E[m_k] = u_k(\boldsymbol{\theta}). \tag{5.7}$$

In the case of two hypotheses the signal strength μ is treated only as a nuisance parameter among others. The *parameter of interest* (POI) in this case is the relative fraction of the two tested signal hypotheses. This fraction is parametrized as ε such that $S_{\text{test}} = \varepsilon \cdot S_{SM} + (1-\varepsilon)S_{\text{ALT}}$. Like this, expression 5.4 becomes

$$\mathcal{L}(\varepsilon, \mu, \theta) = P(N|\mu(\varepsilon \cdot S_{SM} + (1-\varepsilon)S_{\text{ALT}}) + \theta B) \cdot P(N_{\text{CR}}|\theta B_{\text{CR}}). \tag{5.8}$$

5.2.2. Hypothesis Tests

Profile Likelihood and Test Statistic

For all subsequent steps, it is sensible to define a profile likelihood ratio as a test statistic to test specific values of the parameter of interest. For the usual case of fitting the signal strength μ the corresponding expression is

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\hat{\theta}})}{\mathcal{L}(\hat{\mu}, \hat{\theta})}. \tag{5.9}$$

Herein, the numerator holds the likelihood where for the given value of μ the function is maximized by finding the best value of $\boldsymbol{\theta}$. On the other hand, in the denominator both, μ and θ are varied to maximize the likelihood function, which means they both are their maximum likelihood estimators.

While other functions would be suitable as test statistics as well, the Neyman-Pearson lemma ([97]) states that under conditions as in the present case this sort of likelihood ratio has the highest statistical power (when combined with an acceptance or rejection limit).[2]

Since for the properties analysis the POI is not the signal strength μ but the signal model fraction ε, the likelihood ratio here is defined as

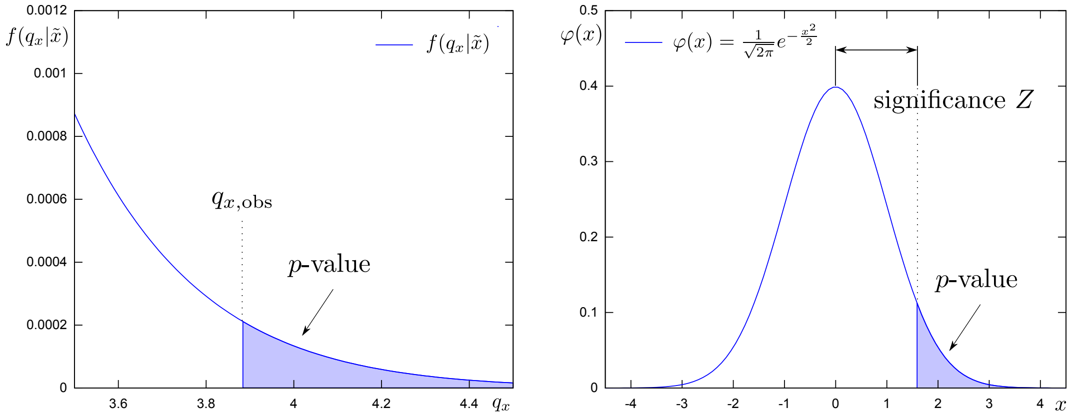
$$\lambda(\varepsilon) = \frac{\mathcal{L}(\varepsilon = 1, \hat{\hat{\mu}}_{\varepsilon=1}, \hat{\hat{\theta}}_{\varepsilon=1})}{\mathcal{L}(\varepsilon = 0, \hat{\hat{\mu}}_{\varepsilon=0}, \hat{\hat{\theta}}_{\varepsilon=0})}. \tag{5.10}$$

For technical reasons given below it makes sense to define the test statistic

$$q_x = -2 \ln \lambda(x), \quad x \in \{\mu, \varepsilon\} \tag{5.11}$$

---

[2] Optimality in this context means that the choice of the cut-off limit implicitly sets a probability of rejecting the null hypothesis (H_0) even though it is correct while the probability of accepting it although it is wrong is minimal.



(a) Plot illustrating the role of the probability density function and the p-value in equation 5.12.

(b) Plot of the standard Gaussian distribution showing the relation between the significance Z and the p-value.

Figure 5.2.: Illustrations of the p-value showing probability density functions and the area corresponding to the p-value highlighted in blue. The p-value represents the probability of obtaining a measurement result which is less compatible with the tested hypothesis than a certain measurement $q_{x,\mathrm{obs}}$. It can equivalently be defined on the lower side of a distribution. (cf. [96])

based on the profile likelihood ratio. On the one hand introducing the logarithm changes the product in the likelihood formula to a sum making it easier to handle and to compute. On the other hand, the interpretation and optimization becomes simpler since now, smaller values of the test statistic indicate a better agreement between the data and the tested value of μ or ε. For the following paragraphs, μ and ε are represented by x where applicable.

Significance and p-value

A standardized quantity to describe the level of disagreement between the measurement data and the hypothesis $\tilde{x}$ is the *p-value*:

$$p_{\tilde{x}} = \int_{q_{x,\mathrm{obs}}}^{\infty} f(q_x|\tilde{x})\,\mathrm{d}q_x. \tag{5.12}$$

In this expression, $q_{x,\mathrm{obs}}$ is the observed value from data of the test statistic (equation 5.11). The function $f(q_x|\tilde{x})$ denotes the probability density function associated with the test statistic q_x under the assumption of a specific value $\tilde{x}$ for the variable x. Figure 5.2a illustrates the quantities in the above expression in an exemplary plot.

It is possible (and often more descriptive) to express the compatibility with the hypothesis to be tested as a significance. The significance (Z) is obtained by computing the distance from the mean of a standard Gaussian distribution to the upper tail region of the Gaussian curve corresponding to the p-value. This value is divided by one standard deviation (see 5.2b). In short the calculation can be written as

$$Z = \Phi^{-1}(1 - p). \tag{5.13}$$

Φ^{-1} denotes the *quantile* of the standard Gaussian distribution, which is the inverse of the Gaussian's cumulative distribution.

$\mathrm{CL_s}$ Limit Setting

To express the final results, a so-called $\mathrm{CL_s}$ (from *confidence level*) exclusion limit is calculated. The $\mathrm{CL_s}$ value is a common quantity for setting limits in high energy physics and is motivated by a frequentist perspective of statistics inasmuch as the imposed limit is characterised by error probabilities.

Compared to a regular frequentist rejection limit it imposes stronger requirements, though. Its concept reduces the susceptibility to overestimating the signal exclusion capability from an experiment by normalisation to the exclusion capability of the alternative (e.g. background only) hypothesis (see e.g. [98]). Its usage for this analysis follows a suggestion from [99], [100]. The general definition that is used within the ATLAS Higgs analysis groups is

$$\mathrm{CL_s} = \frac{p_s}{1 - p_b}. \tag{5.14}$$

The p-values in the expression correspond to the p-values obtained from the signal (p_s) and background (p_b) distributions corresponding to the two alternative hypotheses. For the properties analysis, this definition has been adapted accordingly so that the rejection and acceptance functionality of the tested hypotheses remains identical (the “signal” corresponds to the alternative signal model plus backgrounds while the “background” is the Standard Model signal plus backgrounds). The expression used for this analysis is

$$\mathrm{CL_s} = \frac{p_{\mathrm{obs}}^{\mathrm{ALT}}}{1 - p_{\mathrm{obs}}^{\mathrm{SM}}}. \tag{5.15}$$

5.2.3. Statistical Inference

In order to derive a physical interpretation from measurements involving two hypotheses which are tested against each other, the significance of rejection of either case is required. Additionally, it is helpful to compute a confidence interval (or band) for the rejection power to give best and worst case exclusion limits.

It is important to not only compute the observed significance of an analysis. Also, the knowledge of expected sensitivities is crucial to have a benchmark of the analysis. In other words the expected significance sets the scale for the observation and serves as a cross check. It is obtained as the median significance using the probability density functions for both hypotheses tested. The p.d.f. itself is constructed from systematic uncertainties and signal and background normalisations as explained above. Median significance here means the significance computed with the median of the other hypothesis’ probability density function as observation $t_{\mu,\mathrm{obs}}$. In order to calculate the corresponding p-value and significance, sampling distributions of the p.d.f.’s need to be obtained, since the usage of the p.d.f.’s is not possible (among other reasons they are not analytic).

Two ways are presented in this chapter, while the approach relying on asymptotic formulae was only used provisionally.

Asymptotic Formulae and Asimov Data Set

One way to obtain the relevant quantities in a time-efficient way is via an *Asimov data set*[3]. A defining feature of the Asimov dataset is that the evaluation of the parameter yields returns the

---

[3] The term Asimov data set was chosen inspired by the short story *Franchise* by Isaac Asimov in which presidential elections are held by choosing one single representative person who actually votes (introduced in [96]).

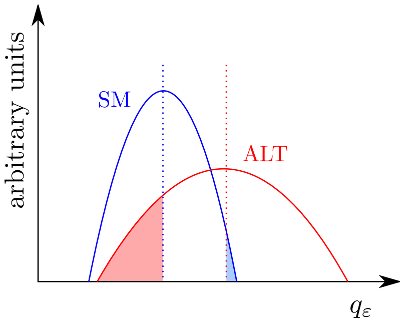


Figure 5.3.: Illustration of toys for two exemplary hypothesis test statistics distributions. The coloured regions correspond to the p-values (or $(1-p)$-value in the left case).

true values which were used to create the dataset in the first place. With respect to expressions 5.6 and 5.7 this implies that

$$n_{i,A} = E[n_i], \tag{5.16}$$

$$m_{i,A} = E[m_i], \tag{5.17}$$

where A indices denote Asimov dataset variables. Starting from the Asimov dataset it is now possible to derive the standard deviation of the $\hat{x}$ (the maximum likelihood estimator of the variable x) distribution which is assumed do be Gaussian with a mean x' (note the Asimov dataset construction with $x = x'$). The result is a simple formula:

$$\sigma_A^2 = \frac{(x-x')^2}{q_{x,A}}. \tag{5.18}$$

In this approximated expression, $q_{x,A}$ is the likelihood ratio of the Asimov dataset likelihood functions:

$$q_{x,A} = -2\ln\lambda_A(x) = \frac{\mathcal{L}_A(x,\hat{\hat{\theta}})}{\mathcal{L}_A(x',\theta)}. \tag{5.19}$$

Based on these relations and the assumptions that have been used to construct the above expressions, it is possible to derive an approximation of the probability density function of the test statistic for a given value of x'. Approximations of significance and p-value can be computed using these distributions. Details about the derivation and application can be found in the referenced publication [96].

Toy Method

Another method to calculate significances from the overall p.d.f.'s is the usage of Monte Carlo toys. These are a straight-forward approach of generating sampling distributions from parametrized p.d.f.'s using a Monte Carlo method. The application of this method in the context of particle physics analyses goes back to a publication by Feldman and Cousins ([101]). While it is more flexible and more precise than the asymptotic approach above, the downside of the toys method is the high computing time requirement. Depending on the number of uncertainties making up the probability density function the generation of a sufficient number of toys requires a high number of high-dimensional Monte Carlo integrations.

Figure 5.3 shows an example of the test statistics distributions for two hypotheses with highlighted regions corresponding to the respective p-values. The more entries for the test statistics distribution (i.e. toys) can be generated the more precisely the p-values or significances can be computed. Given the prepared distributions, these can be read off the plot immediately.

Part II.

Analysis Elucidations: Setup, Execution and Results

After the introduction of the foundations and background information required for the analysis, this part of the document discusses the steps taken in the analysis of the Higgs boson spin and parity. The following chapters rely on the general introduction given in the first part of the document assuming that the reader is familiar with the concepts introduced there.

The starting point for the analysis of both spin and parity is the choice of a theory model describing the states to be tested. Based on these predictions Monte Carlo simulation samples of all models including the Standard Model are generated. These samples pass the detector reconstruction software in the same way as recorded data does such that they reflect the features of each simulated sample in the way they would be reconstructed in the detector. This step of taking detector and reconstruction inefficiencies into account is crucial for the later comparison with recorded data.

By using a multivariate analysis technique, in this case in fact a set of boosted decision trees, several sensitive variables can be combined to one single qualifier. Like this it is possible to use the discrimination power of different variables, each of which reflects a fraction of the physical properties of the original decay.

A specific event selection—loosened with respect to the Higgs boson couplings analysis in the decay channel to two W bosons [102]—is applied to data and Monte Carlo samples. It defines a signal enriched region where the actual tests are performed as well as several background enriched regions. The latter allow constraining specific backgrounds from data. The selection of the signal enriched region takes the differences between the models into account, i.e. it is chosen such that specific features occurring only in non-Standard-Model samples are preserved for the hypothesis tests. Uncertainties arising from modelling the different aspects of the particle interactions and their reconstruction in the detector are discussed in a dedicated chapter and are taken into account in the hypothesis tests.

Before investigating the data in the signal region an estimate of the expected significances has been obtained using a “blinded” signal region while taking all uncertainties into account. Finally, for each hypothesis a test is performed yielding exclusion limits on the Standard Model and alternate model being tested.

In the end an overview of results from other decay channels at the ATLAS experiment and their respective counterparts from the CMS experiment is given where analogous tests have been performed. Furthermore, the perspectives for Run 2 of the CERN Large Hadron Collider (LHC) are given discussing which questions come within reach at higher luminosities and higher beam energies.

CHAPTER 6

Theoretical Details

This chapter discusses the theory aspects directly related to the analysis. General remarks on the Higgs mechanism and the Standard Model are given in the introduction in chapter 2.

Since neither a spin-2 Higgs boson nor a CP-odd (or CP-even but deviating from the Standard Model predictions) Higgs boson fit into the framework of the Standard Model, an extension of the Standard Model is crucial to accommodate these potential states. The models used to achieve this aim are presented in the subsequent section followed by an overview of the generators and tools used to generate the Monte Carlo simulation samples used in the analysis.

6.1. Modelling Spin and Parity Variations

The models tested with the recorded dataset comprise different spin-2 configurations, a spin-0 CP-odd state, a (beyond-the-Standard-Model) spin-0 CP-even state and an admixture of CP states. The latter would imply that the mass eigenstate of the discovered Higgs boson is not a CP eigenstate and therefore establish that there is CP violation in the Higgs sector. The effect of such a possible CP violation might even suffice to explain the matter to antimatter imbalance observed in the universe.

The starting point is an *effective field theory* (EFT) approach formulation which by design is only valid up to a cut-off scale Λ. Λ is assumed to be 1 TeV here, taking the exclusion ranges of beyond-the-Standard-Model physics of the LHC and of previous colliders into account.

Technically, the EFT approach is implemented by adding higher dimensional operators to the Standard Model Lagrangian which describe the interactions of beyond-the-Standard-Model Higgs candidates with other Standard Model particles. From a perturbation theory perspective one can think of this approach as introducing higher order corrections to the Lagrangian which are suppressed by a factor of Λ^{-1} with respect to the Standard Model coupling terms.

In the following, wherever applicable only the terms ruling the couplings to W bosons are considered and shown. The omitted additional coupling terms (see [103] for the full Lagrangian and further details) do not affect the kinematic results of the decay in the $H \to WW^*$ channel and would only affect the production mode and the signal normalisation, where the latter is no parameter of interest for this analysis.

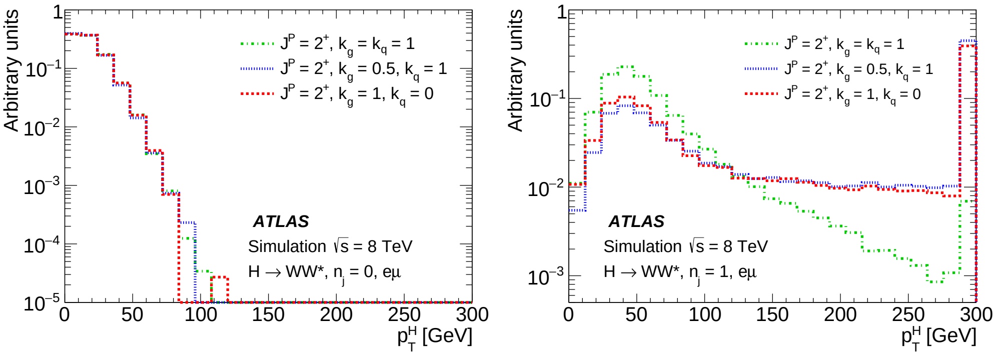


Figure 6.1.: Plots of p_T^H, the Higgs transverse momentum of a spin-2 Higgs boson for different scenarios of couplings to quarks and gluons (coupling factors κ_g and κ_q); in the 0-jet channel hardly any dependence on couplings is visible while in the 1-jet channel a tail appears in the case of non-universal couplings. (cf. [6])

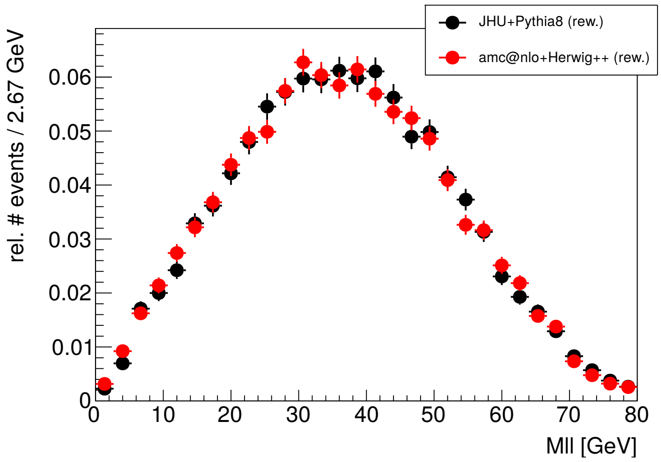


Figure 6.2.: Normalised diboson mass comparison plot as part of the generator cross checks. The distributions shown are for pure CP-odd states in the 1-jet channel comparing JHU and aMC@NLO (for this test showering has been performed with Pythia 8 and Herwig++ respectively). Both samples are reweighted to reproduce the p_T^H-distibution from HRES. This approach assimilates the showered LO and NLO samples to the extent visible in the plot.

6.1.1. Spin-2 Model

Out of the variety of options to model a hypothetical spin-2 Higgs boson, a graviton-inspired approach (similar to the Randall-Sundrum model suggestion [104]) with minimal couplings to Standard Model particles has been chosen. As production modes, gluon-gluon fusion and quark antiquark scattering is considered. The Lagrangian terms added for the EFT approach are:

$$\mathcal{L}_2^p = \sum_{p=V,f} -\frac{1}{\Lambda}\kappa_p T_{\mu\nu}^p X_2^{\mu\nu}. \tag{6.1}$$

In this expression $T_{\mu\nu}^p$ is the energy-momentum tensor and $X_2^{\mu\nu}$ is the spin-2 particle field. V and f denote all vector bosons and all fermions (quarks and leptons) respectively. The κ_p parameters indicate the coupling strengths to particle p, e.g. κ_g denoting the coupling strength to gluons and κ_q denoting the same for quarks.

For producing the non-Standard-Model Higgs Monte Carlo samples, the MADGRAPH5_aMC-@NLO generator was used which provides higher order calculations in α_S. This is a novelty with respect to former group publications [2, 3] where the JHU generator was used which provided only leading order accuracy.

The extension of the spin analysis to also include events with one reconstructed jet is a crucial improvement with respect to earlier studies. For this reason, studies testing a generator providing higher order calculations in α_S (i.e. processes with additional partons) are one cornerstone of the extended analysis.

Figure 6.2 shows an example plot for a $m_{\ell\ell}$ cross check plot of both generators (for cross check purposes: 1-jet channel, pure CP-odd) after reweighting by $p_{\mathrm{T}}^{\mathrm{H}}$. By applying this reweighting a good agreement in most variables can be achieved for the CP analysis test cases. Nevertheless, for the spin-2 analysis a complication arises in the 1-jet channel due to the $p_{\mathrm{T}}^{\mathrm{H}}$ distribution originating from next-to-leading order terms.

This difference with respect to leading order becomes clear when considering production processes of the spin-2 (Higgs) boson. At leading order, both production modes (gluon-gluon fusion and $q\bar{q}$ scattering) are independent. However, in processes with additional partons in the final state a term proportional to the squared difference between both coupling strengths, $(\kappa_q - \kappa_g)^2$ comes into play. It scales with the centre of mass energy s like $\frac{s^3}{m_H^4 \Lambda^2}$ and gives rise to a nearly flat tail in the Higgs p_{T} distribution in the case of non-universal couplings, i.e. for $\kappa_g \neq \kappa_q$. Due to the additional partons involved, this effect hardly affects the 0-jet channel, yet the impact in the 1-jet channel is obvious.

Because the high p_{T} tail is reflected in some distributions of kinematic variable as a distinctive feature it helps to increase the sensitivity when testing the non-universal couplings scenarios. Since the high $p_{\mathrm{T}}^{\mathrm{H}}$ tail inevitably would cause unitarity violation at a certain scale two $p_{\mathrm{T}}^{\mathrm{H}}$ cut-off values have been chosen. This is a work-around motivated by the nature of the EFT approach: The effective field theory model is only valid up to a certain scale so in nature the expectation is to find new physics resolving the unitarity violation around this scale.

Two benchmark points have been studied, one of which is conservative, by applying a $p_{\mathrm{T}}^{\mathrm{H}}$ cut on the Higgs resonance mass which by construction is valid within the EFT model. The other benchmark is chosen at $p_{\mathrm{T}}^{\mathrm{H}} < 300\,\mathrm{GeV}$ which is slightly below one third of the previously mentioned scale of new physics Λ of 1 TeV. For all tests performed here, this point is still considered safe but it is up to the reader to decide which benchmark to rely on. The exact configurations tested are:

- universal couplings, i.e. $\kappa_g = \kappa_q$, no $p_{\mathrm{T}}^{\mathrm{H}}$ cut-off needed,
- non-universal couplings, $kg = 1$, $\kappa_q = 0$ with the two upper cuts on $p_{\mathrm{T}}^{\mathrm{H}}$ at 125 and 300 GeV,
- non-universal couplings, $kg = 0.5$, $\kappa_q = 1$, again with the two upper cuts on $p_{\mathrm{T}}^{\mathrm{H}}$ at 125 and 300 GeV.

The assumption that there is no coupling to gluons (i.e. $\kappa_g = 0$, $\kappa_q = 1$) has not been tested since the obtained $p_{\mathrm{T}}^{\mathrm{H}}$ distribution in such a scenario contradicts observations in the $H \to \gamma\gamma$ and $H \to ZZ$ channels (see analysis results from differential cross section measurements in the respective channels, [105, 106]).

6.1.2. CP Models

As stated in the exposition of this section there are different possible CP models to be tested including pure beyond-the-Standard-Model CP-odd/even states and admixtures. Theory frameworks proposing Higgs bosons with these properties comprise the 2HDM (two-Higgs-doublet models

[107]) and the MSSM (Minimal Supersymmetric Standard Model [108]). An in-depth elaboration of the models within reach of the LHC is presented in [109]. The non-Standard-Model Higgs boson with deviating CP behaviour is also modelled via an effective field theory approach. It is assumed to show strong similarity to the Standard Model Higgs boson when it comes to its other properties. The Lagrangian including the relevant coupling terms to W bosons (see [110]) is given as

$$\begin{aligned}\mathcal{L}_0^W = \Bigg\{ & c_\alpha \kappa_{\mathrm{SM}} \left[g_{HWW}\, W_\mu^+ W^{-\mu} \right] - \frac{1}{2}\frac{1}{\Lambda} \left[c_\alpha \kappa_{HWW}\, W_{\mu\nu}^+ W^{-\mu\nu} + s_\alpha \kappa_{AWW}\, W_{\mu\nu}^+ \widetilde{W}^{-\mu\nu} \right] \\ & - \frac{1}{\Lambda} c_\alpha \left[\left(\kappa_{H\partial W}\, W_\nu^+ \partial_\mu W^{-\mu\nu} + h.c. \right) \right] \Bigg\} X_0 \,. \end{aligned} \tag{6.2}$$

In this formula, the newly introduced tensors are defined as $W_{\mu\nu} = \partial_\mu W_\nu^\pm - \partial_\nu W_\mu^\pm$ and $\widetilde{W}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma} W^{\rho\sigma}$. In the latter expression, $\epsilon_{\mu\nu\rho\sigma}$ denotes the Levi-Civita tensor. The (pseudo-)scalar Higgs boson field is called X_0. The letters c_α and s_α are abbreviations of $\cos(\alpha)$ and $\sin(\alpha)$, where α is a mixing angle parametrizing the admixture of the operators. The coloured terms represent coupling strengths to different scenario Higgs states. The term $c_\alpha \kappa_{\mathrm{SM}}$ represents the coupling of the Standard Model Higgs boson to W bosons. $c_\alpha \kappa_{HWW}$ represents the beyond-the-Standard-Model CP-even Higgs coupling term to W bosons and $s_\alpha \kappa_{AWW}$ represents the respective CP-odd Higgs state coupling term.

The term in the second row proportional to $\kappa_{H\partial W}$ is not regarded in the following since its impact on kinematic distributions is way below the ones of the other terms in Monte Carlo simulations. In fact, the changes to kinematic variables due to this derivative operator make up at most 20 % of the changes originating from the κ_{HWW} term. Since the latter are small themselves given the amount of collected data to date, the derivative operator is omitted for now.

From the occurrences of the angle α in the Lagrangian above, it is obvious that it parametrizes the ratio of the couplings to a CP-even Higgs state (regardless of SM or BSM) compared to the couplings to a CP-odd Higgs state. On the other hand the ratio of κ_{HWW} over κ_{SM} for $c_\alpha = 1, s_\alpha = 0$ parametrizes the admixture of both Standard Model and beyond-the-Standard-Model contributions to a combined scalar, i.e. CP-even Higgs state. As for the spin, the higher order operators mediating the coupling of the BSM Higgs to the Standard Model W boson are suppressed by a factor of Λ^{-1}. As for the spin analysis (see section 6.1.1), a set of benchmarks has been selected which are tested in this analysis. This comprises two fixed hypothesis tests and two range scans:

- a hypothesis test of the Standard Model Higgs boson versus a pure beyond-the-Standard-Model CP-even state for which the different parameters are set to the following values: $\kappa_{SM} = 0\,, \kappa_{AWW} = 0\,, \kappa_{HWW} = 1\,, c_\alpha = 1$,
- a hypothesis test of the Standard Model Higgs boson versus a pure CP-odd state whose parameters are set to $\kappa_{SM} = 0\,, \kappa_{AWW} = 1\,, \kappa_{HWW} = 0\,, c_\alpha = 0$,
- a scan of the tangent of the mixing angle α testing an admixture of a Standard Model Higgs boson and a purely CP-odd BSM state with parameters set to $\kappa_{SM} = 1\,, \kappa_{AWW} = 1\,, \kappa_{HWW} = 0$. More precisely, the mixing parameter is $\frac{\tilde{\kappa}_{AWW}}{\kappa_{SM}} \tan\alpha$, where $\tilde{\kappa}_{AWW} = \frac{1}{4}\frac{v}{\Lambda}\kappa_{AWW}$, with v denoting the vacuum expectation value of the Higgs field. While the fraction in the mixing parameter remains constant, $\tan(\alpha)$ is the only parameter which is varied. The range of variation corresponds to $\cos(\alpha)$ ranging between -1 and 1,
- a scan of the admixture within the CP-even regime, i.e. the ratio of the SM Higgs state over the BSM CP-even Higgs state is scanned with coupling parameters set to $\kappa_{SM} = 1\,, \kappa_{AWW} = 0\,, c_\alpha = 1$. The mixing parameter is chosen to be $\frac{\tilde{\kappa}_{HWW}}{\kappa_{SM}}$, where $\tilde{\kappa}_{HWW} = \frac{1}{4}\frac{v}{\Lambda}\kappa_{HWW}$. The scan

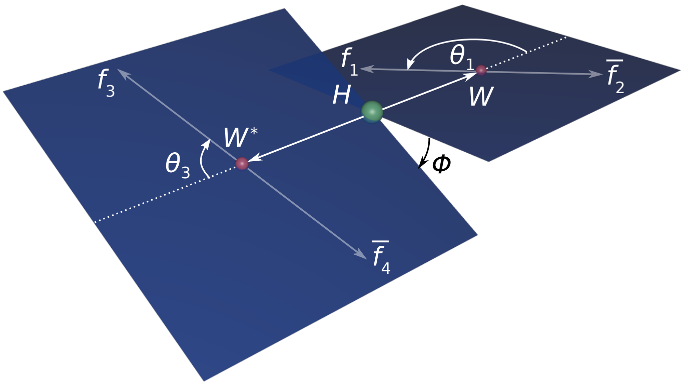


Figure 6.3.: Higgs decay illustration, highlighting the Cabibbo-Maksymowicz variables [111] characterizing the kinematics of the decay via two to four particles in the final state. The variables of interest are the two W masses, the azimuthal angle Φ between the two decay planes of the W bosons and the angles between the leptons f_i originating from the W decays and the Higgs decay axis θ_1 and θ_3.

is performed by varying κ_{HWW} such that the mixing parameter varies between -2.5 and 2.5. The last scan point is chosen to be a purely CP-even BSM state which is approached asymptotically for higher mixing parameter values.

While for the fixed hypothesis benchmarks dedicated Monte Carlo samples have been produced, the scans rely on reweighted samples derived from a single starting point. The latter is chosen at the following parameters: $\kappa_{SM} = 1\,, \kappa_{AWW} = 2\,, \kappa_{HWW} = 2\,, c_\alpha = 0.3$. The choice has been made such that the full range of kinematic configurations reachable in all test cases is indeed accessible from events included in this sample. The reweighting procedure relies on the matrix element computed from the CP-mixing Lagrangian (equation 6.2) for the specific scan point. Dedicated tests have shown that the accuracy is sufficient compared to samples that are directly generated using the scan parameters of interest and deviations are below the percent level.

A scan varying both mixing parameters at a time is not reasonably possible given the collected dataset to date. With higher integrated luminosities this combined scan will be accessible and worth considering.

6.1.3. Sensitive Variables for Spin and CP Test

The kinematics of the Higgs decay in its rest frame (omitting the relation to the incoming partons) can be fully expressed via the five Cabibbo-Maksymowicz variables [111]. These comprise the masses of the direct decay products - here the W and W^* mass, the azimuthal angle Φ between the W decay planes as well as the angles between the leptons originating from the $W^{(*)}$ decay and the Higgs decay axis θ_1 (θ_3). A corresponding illustration of all quantities is shown in figure 6.3. Figure 6.4 extends the scheme by non-rest-frame variables introducing a relation to the incoming partons. The resulting distributions of the azimuthal angle Φ between the decay planes in the Higgs rest frame are plotted in figure 6.5. The plots compare a Standard Model scalar Higgs boson to a spin-2 Higgs boson as well as to a CP-odd Higgs boson with spin-0. The resulting distributions reproduce the ones shown in theory publications dealing with alternative Higgs models (compare e.g. [79] [112] [78]). Due to the two neutrinos in the final state, none of these variables is directly accessible and the Higgs rest frame can only be approximated (see e.g. [113]). The corresponding system of equations is underdetermined and can rely solely on the reconstructed vectorial missing transverse momentum which reproduces the neutrino transverse momentum sum in an idealized scenario. The challenge for all subsequent steps of the analysis is

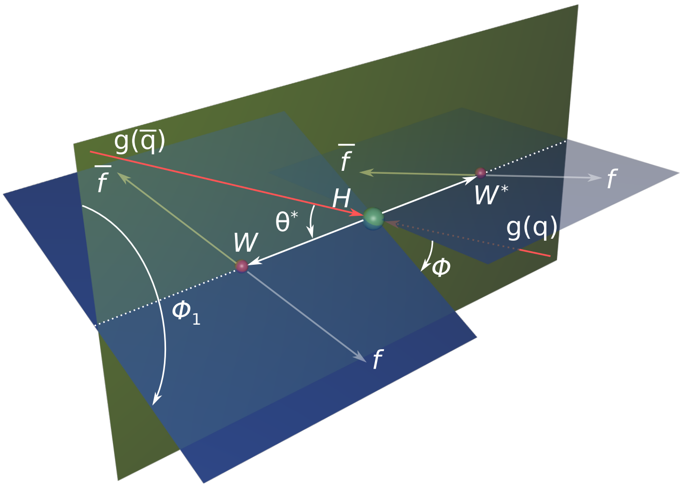


Figure 6.4.: Higgs decay illustration as in figure 6.3 but with additional variables introducing relations to the incoming particles (gluons (g) or quarks (q)). As before, the angle ϕ is the angle between the W boson decay planes and f denotes the leptons from the W boson decays. The additinoal (green) plane is defined by the Higgs boson decay axis and the axis of the incoming partons. The newly introduced angle θ^* denotes the angle between the incoming partons and the Higgs boson decay axis. The angle ϕ_1 is the angle between the decay plane of the on-shell W boson and the added green plane.

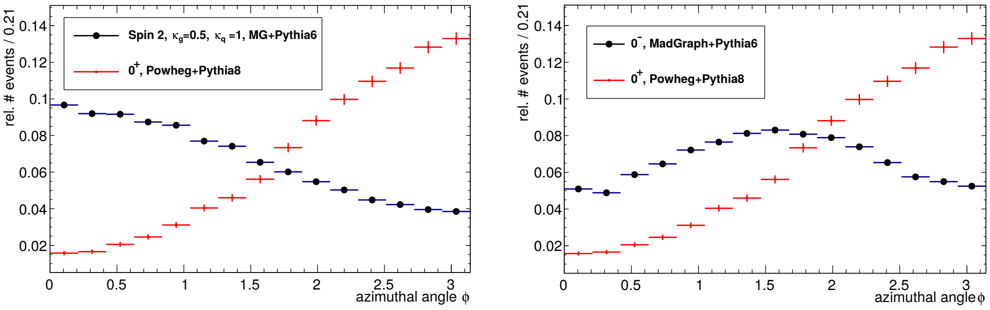


Figure 6.5.: Truth level (i.e. before detector simulation but after showering and hadronization) distributions of the azimuthal angle in the Higgs rest frame Φ as displayed in figure 6.3 for the spin-2, $\kappa_g = 0.5, \kappa_q = 1$ (left) and BSM CP-odd (right) comparison (the red dots indicate a Standard Model Higgs distribution). The neutrinos are available in the Monte Carlo samples and are used here for the computation of the plotted variables for illustration purposes. The non-Standard-Model samples have been generated with MADGRAPH5_aMC@NLO.

the identification of sensitive variables carrying the most significant fraction of information originally contained in the Cabibbo-Maksymowicz variables. The final choice of variables is detailed in section 7.4 and relies substantially on the polar angle between both reconstructed leptons which is correlated to the angle Φ.

6.2. MC Event Generators

The samples used for the spin and CP analysis are mostly identical to those used for the couplings measurements of Higgs decays to WW^* which are described in [102]. Only the signal samples representing the alternative models (spin-2 and CP analysis models) have been produced specifically for this analysis. In-depth studies of the most suitable generator have preceded the choice of the MadGraph5_aMC@NLO generator, replacing the formerly used JHU generator.

Besides the W+jets background which is entirely estimated from data (see 7.5.5), Monte Carlo simulation samples have been generated for every relevant background process. Table 6.1 lists all samples with their process, generator, cross section times branching ratio and optional filters. The latter have been applied in the generation process where necessary or sensible. Short descriptions of all generators showing up in the table and in the subsequent explanation are given in the introductory remarks on Monte Carlo generators in section 4.2.

The majority of samples are produced with the Powheg generator providing corrections at next-to-leading order in α_s for these processes. For all processes where more partons (up to 5) need to be simulated, Alpgen or Sherpa have been employed. acermc and gg2vv have been used where necessary when no higher order generator is available (e.g. $gg \rightarrow WW$ is not included in Powheg).

For the modelling of showering, hadronization and underlying event different generators have been used: Pythia versions 6 and 8 (depending on the required interface and feature set), Herwig (with Jimmy for modelling the underlying event) and Sherpa. Different PDF sets have been used with different generators, namely ct10 combined with Powheg and Sherpa and cteq6L1 with Alpgen and acermc. For the $Z/\gamma^* \rightarrow \tau\tau$ sample the mrst PDF set has been deployed for reweighting.

All pile-up contributions have been modelled with Pythia 8, comprising the underlying event and additional interactions in the same or adjacent bunch crossings.

The signal samples are generated using either Powheg + Pythia 8 for the Standard Model Higgs at $m_H = 125\,\text{GeV}$ or MadGraph5_aMC@NLO+ Pythia 6 for all other models. Additionally, a Standard Model Higgs sample has been generated with MadGraph5_aMC@NLO to be used for the CP analysis (in order to avoid purely generator-related differences between different CP states or admixtures). The PDF set used for this is cteq6L1. The approach of using a dedicated spin-0 Standard Model sample produced with the same generator was dropped for the spin-2 scenario since the kinematic differences are more prominent making the common (ggF) Standard Model Higgs sample sufficiently reliable in this case.

The MadGraph5_aMC@NLO samples for the spin-2 studies are created as combinations of processes where a Higgs boson plus zero to two partons are produced, each at leading order. This is mostly important for the 1-jet channel as explained earlier in section 6.1.1. Comparisons with real NLO generation of inclusive samples (only investigating the 0 and 1-jet channel though) have shown that there is no actual benefit from using NLO level accuracy in this case. The Higgs p_{T} distributions of *all* spin-0 Higgs models have been reweighted to match the NNLL and NNLO (next-to-next-to-leading-logarithms and next-to-next-to-leading-order) calculations from hres2.1 to achieve the best modelling accuracy possible. While higher order calculations were not available for the BSM spin-0 models, this approach avoids introducing an artificially increased separation power originating from the p_{T} reweighting.

For the calculation of the inclusive WW cross section at next-to-leading order MCFM was applied while the calculation (at leading order in α_S) for $gg \rightarrow WW$ was also performed with GG2VV (as for the sample generation). The latter takes account of the production of WW and ZZ as well as their interference.

For the $t\bar{t}$ normalisation TOP++2.0 is used providing NNLO calculations in α_S and resummation with NNLL accuracy. For single top processes the normalisation is computed to NNLL based on dedicated calculations, see e.g. [114]. The samples listed in the table for the $W\gamma^*$ and WZ processes on the one hand and the $Z\gamma^*$ and ZZ processes on the other hand belong together as associated $W + Z/\gamma^*$ and $Z + Z/\gamma^*$ production. The reason for splitting these processes is that POWHEG +PYTHIA 8 cannot model the lower $m_{\ell\ell}$ regions which are nevertheless of interest. Among other measures the SHERPA samples generated in this low-$m_{\ell\ell}$ regime are corrected by reweighting the total cross section to MCFM NLO calculations.

For both processes modelled with ALPGEN+HERWIG a normalisation to higher order calculations is performed: the Drell-Yan sample is normalised using DYNNLO for NNLO calculations and the $W\gamma$ sample (i.e. $W \rightarrow \ell\nu$ decay plus an asymmetric $\gamma \rightarrow e^+e^-$ pair production) was normalised to NLO MCFM calculations. The utilization of ALPGEN with its multi-parton modelling capabilities is crucial in this context for generating accurate multijet samples for both processes.

Finally, the background normalisation for the $Z\gamma$ process (with the Z decaying to two leptons which need to fulfil a $m_{\ell\ell} > 10\,\mathrm{GeV}$ requirement) is adjusted to match the NLO predictions from MCFM. Both, the $Z\gamma$ and $W\gamma$ sample require a minimum p_T of the radiated photon of at least 8 GeV.

Table 6.1.: List of Monte Carlo samples used for the analysis. Listed are the processes with the respective generator, filters where applied during the generation and cross section times branching fraction for a centre of mass energy of $\sqrt{s} = 8\,\text{TeV}$. The latter includes decays of $t \to Wb$, $W \to \ell\nu$ and $Z \to \ell\ell$ (except for $ZZ \to \ell\ell\nu\nu$ where ℓ represents all charged leptons). $W\gamma^*$ and WZ as well as $Z\gamma^*$ and ZZ refer to the same processes each, split up into a low and high dilepton mass regime (corresponding to γ^* or Z as the second decay particle). The coupling parameters in the signal sample section are explained in the corresponding preceding sections 6.1.1 and 6.1.2. (table from [6])

Process	MC generator	Filter	$\sigma \cdot \mathcal{B}$ (pb)
Signal samples used in $J^P = 2^+$ analysis			
SM $H \to WW^*$	Powheg +Pythia 8		0.435
$k_g = k_q$	MadGraph5_aMC@NLO+Pythia 6		-
$k_g = 1, k_q = 0$	MadGraph5_aMC@NLO+Pythia 6		-
$k_g = 0.5, k_q = 1$	MadGraph5_aMC@NLO+Pythia 6		-
Signal samples used in CP-mixing analysis			
$c_\alpha = 0.3$, $k_{SM} = 1$ $k_{HWW} = 2$, $k_{AWW} = 2$	MadGraph5_aMC@NLO+Pythia 6		-
SM $H \to WW^*$	MadGraph5_aMC@NLO+Pythia 6		0.435
Background samples			
WW			
$q\bar{q} \to WW$ and $qg \to WW$	Powheg +Pythia 6		5.68
$gg \to WW$	gg2vv+Herwig		0.196
Top quarks			
$t\bar{t}$	Powheg +Pythia 6		26.6
Wt	Powheg +Pythia 6		2.35
$tq\bar{b}$	acermc+Pythia 6		28.4
$t\bar{b}$	Powheg +Pythia 6		1.82
Other dibosons (VV)			
$W\gamma$	Alpgen +Herwig	$p_\mathrm{T}^\gamma > 8$ GeV	369
$W\gamma^*$	Sherpa	$m_{\ell\ell} \leq 7$ GeV	12.2
WZ	Powheg +Pythia 8	$m_{\ell\ell} > 7$ GeV	12.7
$Z\gamma$	Sherpa	$p_\mathrm{T}^\gamma > 8$ GeV	163
$Z\gamma^*$	Sherpa	min. $m_{\ell\ell} \leq 4$ GeV	7.31
ZZ	Powheg +Pythia 8	$m_{\ell\ell} > 4$ GeV	0.733
$ZZ \to \ell\ell\,\nu\nu$	Powheg +Pythia 8	$m_{\ell\ell} > 4$ GeV	0.504
Drell-Yan			
Z	Alpgen +Herwig	$m_{\ell\ell} > 10$ GeV	16500

CHAPTER 7

Event Selection and Backgrounds

The following sections discuss core aspects of the analysis, starting from the trigger configuration over the object definition and event preselection to the definition of the signal region and control regions. The last part of the chapter sheds light on the most prominent, irreducible backgrounds in some more detail.

7.1. Trigger and Offline Criteria

In general all events used for the analysis need to fulfil quality requirements, demanding that all relevant detector components were functioning correctly during data taking. Applying this condition to the overall recorded dataset of the year 2012 at $\sqrt{s} = 8\,\text{TeV}$ yields an integrated luminosity of $20.3\,\text{fb}^{-1}$ of data to be analysed. The 2011 dataset recorded at $\sqrt{s} = 7\,\text{TeV}$ is not considered, because many technical differences would require a dedicated analysis for this dataset which has been discarded with respect to the small expected gain in significance.

For this analysis a single lepton and a dilepton trigger have been applied. This means that events get selected if one or both of these trigger conditions are met. The quality and p_{T} requirements for the single lepton trigger are chosen tighter than for the dilepton trigger. This is due to the fact that the single lepton triggers yield a higher background rate such that a higher rejection is required for compensation. The precise trigger thresholds used for the 2012 data taking are listed in table 7.1.

The trigger efficiencies have been studied using a tag-and-probe method[1] with electron or muon candidates from $Z/\gamma^* \to ee, \mu\mu$ decays. The efficiencies range from $\sim 70\,\%$ for muons in the central region (for $|\eta| < 1.05$) to $95\,\%$ for the dilepton trigger in $e\mu$ events, i.e. events with electrons as the leading and muons as the subleading lepton (both in the 0-jet channel). In general the efficiencies lie around $90\,\%$ for most configurations (cf. [102]).

---

[1] The tag-and-probe method is described in [115]. Briefly summarized, one lepton from the Z decay is selected after applying *strict* selection criteria and used as the “tag” lepton. The candidate for the corresponding second lepton from the decay is the “probe” lepton. Both leptons need to pass combined selection and invariant mass criteria. The efficiency to be estimated is obtained as the fraction of “probe” leptons passing the tested requirements. The expected contamination from backgrounds is taken into account for the efficiency calculation.

Table 7.1.: The table below lists the trigger thresholds (all numbers in GeV) used in the 2012 data taking. The values followed by the letter 'i' imply an isolation requirement (looser than for the subsequent offline reconstruction, cf. 7.2.3). Thus, for the single lepton trigger either the isolation criteria or a higher p_T threshold are required. The numbers in parentheses for the single electron trigger denote that the condition is either ($p_\mathrm{T} \geq 18$ GeV [L1]$\wedge p_\mathrm{T} \geq 24$ GeV, isol. [HL]) or ($p_\mathrm{T} \geq 30$ GeV [L1]$\wedge p_\mathrm{T} \geq 60$ GeV [HL]). The pair of thresholds in the dilepton section correspond to the leading and subleading lepton respectively. (adapted from [102])

Name	Level-1 trigger threshold / GeV	High-level trigger threshold / GeV
Single lepton		
e	18 (30)	24i (60)
μ	15	24i $\vee$ 36
Dilepton		
e, μ	10 $\wedge$ 6	12 $\wedge$ 8

The vertex of interest within each event (out of about 21 inelastic collisions which on average occurred per bunch crossing during the 2012 data taking[2]) is determined from the transverse track momenta associated with each vertex. The vertex for which

$$\sum_{\substack{\text{track from}\\\text{vertex}}} (p_{\mathrm{T,track}})^2 \tag{7.1}$$

is maximal is regarded as the *primary* vertex which is investigated in subsequent steps. This primary vertex is required to have at least three associated tracks whose p_T is above 400 MeV.

7.2. Object Definitions

The definition of physics objects that are analysed to deduce the Higgs boson properties (i.e. electrons, muons, jets and missing transversal energy) is crucial and needs to be adapted specifically for each analysis. Since the Standard Model $H \to WW^*$ process with the same final state ($e\mu$, μe) as used for the properties analysis is investigated in detail in the couplings analysis (cf. [102]), the same object selection is applied here. The following section gives an overview of the criteria and algorithms applied for the object definition. The input variables for this step of the analysis are provided by the event reconstruction software which is described in 3.3. The various object definitions are given in the respective subsections below.

7.2.1. Electrons

Objects identified as electrons require an inner detector track with an associated calorimeter energy deposit (see [115]). To account for the energy loss due to material interactions on the way through and between both subdetectors, a Gaussian sum fit (cf. [116]) is performed. This way, the fit improves the precision of the track parameters and of the impact parameter. The transversal energy

---

[2] An average number of 20.7 interactions per bunch crossing is given in the respective performance studies, see the "Number of Interactions per Crossing" plot in the "2012 pp Collisions" section on the `LuminosityPublicResults` website [30].

of the electron is obtained from the energy deposit in the calorimeter combined with the angular information from the inner detector track extrapolated to the interaction point.

Electrons in the energy range $10 < E_{\mathrm{T}} < 25\,\mathrm{GeV}$ need to satisfy the "very tight" requirement (class definitions can be found in [115]). This definition is likelihood-based and comprises nearly all available requirements with respect to sensitive variables for electrons. In total 18 requirements need to be met covering the inner detector track (silicon and TRT region) as well as the energy deposits in both the electromagnetic and hadronic calorimeter (i.e. a cross check ratio of the energy deposit in the first layer of each calorimeter type to suppress hadronic energy deposits) and the track to calorimeter cluster matching. Further crucial quantities characterize the shower shape in longitudinal and transversal direction. The "very tight" requirement is chosen in response to the high 'fake electrons' contribution from various backgrounds which increases strongly with lower transverse momenta. The origin of the fake electrons is the misidentification of photons and light quark jets.

Above the threshold of 25 GeV, the distortion from backgrounds is sufficiently low to adopt a "medium" quality requirement. This criterion is derived from a cut-based identification test and misses different quantities, especially with respect to the matching between calorimeter and inner detector tracks. The "medium" selection as defined in the reference is extended slightly in order to reduce photon conversions being identified as electrons. To achieve this, a hit on the innermost pixel detector layer is required and tracks associated with conversion vertices are excluded. A loosened definition naturally increases the acceptance of higher E_{T} electrons in the reconstruction which increases the available dataset for the analysis.

The acceptance range in η for the electron candidates is $|\eta| < 2.47$ with an excluded range of $1.37 < |\eta| < 1.52$. The latter is the transition region between the barrel and end cap region of the electromagnetic calorimeter which is excluded because of lacking calorimeter measurements.

7.2.2. Muons

Objects are identified as muons if they have an inner detector track as well as a muon spectrometer track assigned. More precisely, the classification as a "combined muon" (cf. section 3.3) is required, including a combined fit in both regions. The tracks in the muon spectrometer are required to hold track segments in all three muon spectrometer layers and at least a minimal number of hits in each inner detector part. Details about the muon reconstruction and the identification performance can be found in [47]. Additionally, muons to be used for the analysis are required to lie within $|\eta| < 2.5$.

7.2.3. Common Lepton Criteria

In addition to the requirements listed in the two previous subsections which are applied respectively for electrons and muons there are common criteria which need to be met. This comprises constraints on the impact parameter in longitudinal and transversal direction as well as on the lepton isolation parameters[3]. The latter are computed as track-based and calorimeter-based quantities. In both cases, limits on the ratio of the summarized p_{T} (E_{T}) of tracks (calorimeter clusters) surrounding the lepton within a ΔR cone divided by the lepton p_{T} (E_{T}) are imposed. In doing so, the computation of the calorimeter isolation parameter takes the effect of pile-up, underlying event and wrongly assigned contributions from the electron itself into account. The track-based isolation parameter on its part imposes quality requirements on the surrounding tracks to reduce erroneous rejection.

---

[3] For the transverse impact parameter d_0, an uncertainty condition of $|d_0|/\sigma_{d_0} < 3.0$ is required. The cut on the longitudinal impact parameter z_0 is $|z_0 \sin\theta| < 0.4\,\mathrm{mm}$ for electrons and 1.0 mm for muons. The isolation criterion for electrons is described in detail in reference [115] while a brief description for both cases is also given in [102].

The efficiency of the two additional requirements (conc. isolation and impact parameter) has been estimated using a tag-and-probe method applied to $Z/\gamma^* \rightarrow ee, \mu\mu$ candidates from data and ranges up to 90 % for electrons and 96 % for muons, each with $p_T > 25\,\mathrm{GeV}$.

7.2.4. Jets

Jets are particle showers originating from partons (hadronic jet) or photons/electrons (electromagnetic jet) that are radiated from the interaction or previous decay products. In the reconstruction process they are reconstructed using the anti-k_T algorithm [117] with a radius parameter of $R = 0.4$. The anti-k_T algorithm is a sequential clustering algorithm which satisfies collinear and infrared safety requirements imposed by the factorization and hadronization modelling in theory. It iterates over all 'particles' (tracks and clusters of energy deposits in the calorimeters) in an event and sequentially combines the associated four vectors within a geometrical range specified via the R (radius) parameter which translates to the distance in rapidity and azimuthal angle. The corresponding formula to compute the "distance" between particles i and j is

$$d_{ij} = \min\left(\frac{1}{k_{\mathrm{Ti}}^2}, \frac{1}{k_{\mathrm{Tj}}^2}\right) \cdot \frac{\Delta_{ij}^2}{R^2} \tag{7.2}$$

with $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, i.e. Δ_{ij}^2 stands for the geometrical distance in rapidity and azimuthal angle. The terms k_{Ti} denote the transversal momentum of the i-th particle of the event. The consequence of the first factor is that high-p_T particles appear closer to other particles than soft radiation particles. In this way, soft particles in proximity of hard particles get assigned to the hard particles (i.e. the jet) rather than to other soft particles nearby. The results are conical jets created around the hardest particles of the event. In the case of two nearby jets (hard particles with $R < \Delta_{ij} < 2R$) an intermediate bounding volume is formed, the shape of which depends on the ratio of magnitudes of the two particle's transverse momenta.

Energy deposits in calorimeter cells need to lie above a noise threshold to be considered in the formation of clusters. As a measure to compensate different effects (non-compensating design of the calorimeters, signal and energy losses for technical reasons), energy calibrations are applied to the three dimensional input calorimeter clusters before they enter the jet reconstruction. The so-called local cluster weighting (LCW) calibration [118, 119] applies these corrections depending on the type of the energy deposits, i.e. whether they originate from electromagnetic or hadronic showers.

Pile-up effects and the primary vertex position are accounted for in a subsequent step. After the correction a data-driven calibration is applied taking into account the transversal momentum and the pseudorapidity of the jet. The associated systematic uncertainties are obtained from the data sample.

In order to suppress jets from pile-up vertices (i.e. from other collisions in the same or adjacent bunch-crossings), the *jet vertex fraction* (JVF) is introduced. This quantity relates the jets in the event to the primary vertex (abbreviated as "p.v." below). It reproduces the share of the summarized p_T of tracks from the primary vertex in the combined p_T of all jets in the event,

$$\mathrm{JVF} = \frac{\sum\limits_{\substack{\text{tracks, from p.v.,}\\ \Delta R<0.4}} |p_{\mathrm{T,track}}|}{\sum\limits_{\substack{\text{tracks,}\\ \Delta R<0.4}} |p_{\mathrm{T,track}}|}, \tag{7.3}$$

where ΔR is to be understood as relative to the jet axis. For the current analysis, the JVF for jets with $p_T < 50\,\mathrm{GeV}$ in the range of $|\eta| < 2.4$ is required to be above 50 %.

The jets used to categorize events in jet bins (for this analysis either zero jets or one jet are required) need to fulfil special p_{T} requirements. In the central region of $|\eta| < 2.4$ jets need to have a p_{T} above 25 GeV while in the outer region of $2.4 \leq |\eta| < 4.5$ at least 30 GeV are required to suppress pile-up jets.

B-Quark Jets

The identification of heavy quark flavour jets (bottom and charm type) is crucial for the suppression of top-quark backgrounds. The MV1 algorithm is used for this purpose (cf. [120]). It combines the most discriminating outputs of three primary algorithms and evaluates them using a neural network which has been trained with b jets against light jet flavours. The three primary algorithms make use of the impact parameter, secondary vertex information and a decay vertex tracing algorithm which uses a Kalman filter approach. Each jet gets assigned a *tag weight* rating the likelihood of the jet to belong to a heavy flavour jet. For the current analysis, a working point of 85 % of b tagging efficiency is used. In general, b tagging is only possible in the η range covered by the inner detector ($|\eta| < 2.5$); furthermore jets to be tagged need to have a $p_{\mathrm{T}} > 20$ GeV.

7.2.5. Missing Transverse Momentum

Events like the signal process of this analysis which contain neutrinos in the final state will have signatures with "missing" energy and momentum entries in the transversal plane, since the energy and momentum carried away by the neutrinos is not accessible. From an idealized perspective, the sum of all recorded detector activity, i.e. energy deposits in the calorimeter and reconstructed tracks, should exactly balance the sum of energy and momentum 'lost' through the neutrinos. Thus, in order to gain as much information as possible about the overall direction and magnitude of missing momentum, all entries of observed objects are summarized (cf. [121]). Shortly one can summarize this as

$$\boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}} = -\left(\sum_{\text{selected}} \boldsymbol{p}_{\mathrm{T}} + \sum_{\text{soft}} \boldsymbol{p}_{\mathrm{T}} \right), \tag{7.4}$$

where the bold symbols denote four-momenta. *Selected* refers to the identified objects such as leptons and jets while *soft* refers to low-p_{T} objects that do not fall in any pre-defined categories. The $p_{\mathrm{T}}^{\mathrm{miss}}$ can be corrected by making use of calorimeter entries for the soft objects on the one hand or low-p_{T} tracks (less prone to pile-up effects) assigned to the primary vertex on the other hand. These different optimizations yield two (or more) different varieties of $p_{\mathrm{T}}^{\mathrm{miss}}$ which can be chosen depending on the desired use case. For this analysis, a mixture of the two approaches has been utilized by combining the track-based $p_{\mathrm{T}}^{\mathrm{miss}}$ with jet corrections based on the respective calorimeter entries.

7.2.6. Overlap Removal

Since objects in the detector may be misidentified or misinterpreted, a number of rules have been implemented to remove spurious objects in cases where objects are identified very close to one another in ΔR. The rules applied in this analysis are as follows:

- An electron candidate object is removed if its track extends to the muon spectrometer. Also if a muon and an electron candidate are less than 0.1 sr apart in the spherical ($\eta \times \phi$) plane, the electron is removed. This follows the assumption that in these cases the electron candidate originated from bremsstrahlung emitted by a muon.
- Since high-p_{T} electrons are always additionally reconstructed as electromagnetic jets, a reconstructed jet that overlaps with an electron ($\Delta R \leq 0.3$) is always removed.

- The analogue case involving a muon and a jet (also $\Delta R \leq 0.3$) is treated the other way around since in that case the assumption that the muon originates from a non-prompt decay within the jet is more likely, so the muon gets removed.

- Two electron candidates that are at most $\Delta R = 0.1$ apart may originate from bremsstrahlung in one of the innermost detector parts and therefore belong together. In this case only the electron candidate with the higher p_T is regarded for precautionary reasons.

7.3. Event Preselection

The common event preselection criteria which are applied for each region and jet channel of each analysis are discussed in this section. These cuts serve the purpose of generally reducing the contributions from dominant backgrounds as described in 7.5.

In the offline event preselection only events with oppositely charged different flavour leptons (μ and e) fulfilling the isolation criteria and originating from the primary vertex are accepted. The same-flavour channel (i.e. with ee and $\mu\mu$ in the final state) has been omitted due to the large background contribution from $Z/\gamma^* \to ee/\mu\mu$ to be expected in this channel. The leading lepton (i.e. the lepton with the highest p_T, denoted as ℓ_1) is required to have a $p_\mathrm{T} \geq 22\,\mathrm{GeV}$ and the subleading lepton (ℓ_2) needs to have a $p_\mathrm{T} \geq 15\,\mathrm{GeV}$. Additionally the pair of leptons must have an invariant mass (computed from added lepton four-momenta) above 10 GeV. The missing transverse momentum $p_\mathrm{T}^\mathrm{miss}$ within the event is required to be greater than 20 GeV. Also, at most one jet is allowed in the event. To be counted as a jet, each jet candidate must satisfy $p_\mathrm{T} > 25\,\mathrm{GeV}$ in the central region ($|\eta| \leq 2.4$) while in the forward region ($2.4 < |\eta| < 4.5$) the condition is $p_\mathrm{T} > 30\,\mathrm{GeV}$ (the latter to reduce pile-up effects).

In subsequent steps of the analysis, the 0-jet and 1-jet channel are regarded separately since the background composition and magnitudes differ notably.

7.4. Event Selection

7.4.1. Distributions of Sensitive Variables

Starting from the distributions of the Cabibbo-Maksymowicz variables and based on sensitivity tests comparing BDT results within the TMVA framework, a set of sensitive variables for both analyses was chosen. For the spin-2 analysis, this set comprises the transverse momentum ($p_\mathrm{T}^{\ell\ell}$) and invariant mass ($m_{\ell\ell}$) of the dilepton system (obtained by adding the lepton four vectors and computing their p_T and invariant mass), the angle between both leptons $\Delta\phi_{\ell\ell}$ and the transverse mass m_T in the 0 and 1-jet channel. The latter variable is defined as the invariant mass of the dilepton system and the missing transverse momentum [122] and can be obtained as

$$m_\mathrm{T} = \sqrt{(E_\mathrm{T}^{\ell\ell} + p_\mathrm{T}^{\nu\nu})^2 - |\boldsymbol{p}_\mathrm{T}^{\ell\ell} + \boldsymbol{p}_\mathrm{T}^{\nu\nu}|^2}. \tag{7.5}$$

In this expression, $E_\mathrm{T}^{\ell\ell} = \sqrt{(p_\mathrm{T}^{\ell\ell})^2 + (m_{\ell\ell})^2}$, $\boldsymbol{p}_\mathrm{T}^{\ell\ell}$ and $\boldsymbol{p}_\mathrm{T}^{\nu\nu}$ denote the vector sum of the spatial vector components of the two leptons and neutrinos (the latter approximated as $p_\mathrm{T}^\mathrm{miss}$). Accordingly, $p_\mathrm{T}^{\nu\nu}$ (like $p_\mathrm{T}^{\ell\ell}$) is the modulus of the transverse momentum of both neutrinos, i.e. approximated from the modulus of the missing transverse momenta. Even though the invariant mass of the final state cannot be reconstructed, this transverse mass is a valuable quantity which among others has the feature that its upper limit at generator level is the Higgs mass (after reconstruction the smearing via the detector weakens this constraint). Like this it is very well suited to suppress backgrounds, especially the WW background due to its preference of higher invariant masses of the intermediate quark loop.

The distributions of all four variables obtained from Monte Carlo simulations in the 0 and 1-jet channel are shown in figure 7.1. The plot displays the Standard Model (0^+) distribution as well as the distributions for all three coupling configurations for a spin-2 Higgs boson. Additionally the sum of all backgrounds is shown in the plots.

For the CP analyses two slightly different sets of variables have shaped up as being most sensitive. Again, tests have been performed to evaluate different sets of variables also varying the number of variables entering the BDT. For the BSM CP-even analysis, the list of variables is $p_{\mathrm{T}}^{\ell\ell}$, $m_{\ell\ell}$, $\Delta\phi_{\ell\ell}$ and $p_{\mathrm{T}}^{\mathrm{miss}}$, i.e. $p_{\mathrm{T}}^{\mathrm{miss}}$ replaces the transverse mass m_{T} by comparison with the spin analysis.

In the case of the CP-odd versus Standard Model analysis, the list is more divergent: $m_{\ell\ell}$, $E_{\ell\ell\nu\nu}$, Δp_{T} and $\Delta\phi_{\ell\ell}$ enter the BDT in this case. The newly introduced quantity is

$$E_{\ell\ell\nu\nu} = p_{\mathrm{T}}^{\ell_1} - \frac{1}{2}p_{\mathrm{T}}^{\ell_2} + \frac{1}{2}p_{\mathrm{T}}^{\mathrm{miss}}, \tag{7.6}$$

where $\frac{1}{2}p_{\mathrm{T}}^{\ell_{1,2}}$ denotes the transverse momentum of the leading or subleading lepton.

7.4.2. Signal Region Definition

Following the pre-selection cuts, table 7.2 lists the cuts defining the 0 and 1-jet signal regions. Basically these resemble the ones applied in the common $H \to WW^*$ publication [102] for the gluon-gluon fusion production mode in the different flavour channel. Yet since the latter are chosen in order to select a Standard Model Higgs, some adjustments have been made for the properties analysis in order to preserve the differences between the tested models (as shown in figures 7.1, 7.2 and 7.3). Those cuts which are intended primarily for background suppression (see common publication for categorization) are only slightly adjusted.

The 0-jet selection contains cuts on three kinematic variables, namely $p_{\mathrm{T}}^{\ell\ell}$, $m_{\ell\ell}$ and $\Delta\phi_{\ell\ell}$. Since these variables are also among the most sensitive variables with respect to the separation of alternative models and the Standard Model Higgs, a compromise has been found which reduces the background contribution while preserving the kinematic features of all models. The upper cut on $m_{\ell\ell}$ primarily serves the purpose of reducing the WW background contribution (which is enriched at higher $m_{\ell\ell}$ values while the Higgs resonance sets an upper bound to the dilepton mass). The cuts on $p_{\mathrm{T}}^{\ell\ell}$ and $\Delta\phi_{\ell\ell}$ help reduce the contribution from Drell-Yan ($Z/\gamma^* \to \tau\tau$) backgrounds, mostly by removing nearly back-to-back lepton configurations which would be preferred by the $Z/\gamma^* \to \tau\tau$ background process. Figure 7.4 shows extended range plots of $m_{\ell\ell}$ and $\Delta\phi_{\ell\ell}$, in which the cutting value is highlighted and the respective motivation of each choice becomes evident.

All three cuts are loosened with respect to the Standard Model Higgs analysis. The importance of this change becomes clear when considering the spin-2 Higgs model which implies different helicities of the W. These lead to a larger angular separation ($\Delta\phi_{\ell\ell}$) between the electron and muon which needs to be accommodated by the signal region definition. At the same time, this fact makes the spin-2 samples harder to distinguish from the WW background since it resembles the WW decay configuration.

The 1-jet selection uses the same cuts on $m_{\ell\ell}$ and $\Delta\phi_{\ell\ell}$ as the 0-jet selection and adds some extra cuts e.g. to reject the top background which has a significantly larger impact in the 1-jet channel. For this, events containing b-tagged jets with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ are excluded. Additionally an upper cut on the transverse mass m_{T} (see formula 7.5) at 150 GeV is introduced which reduces top and WW background contributions.

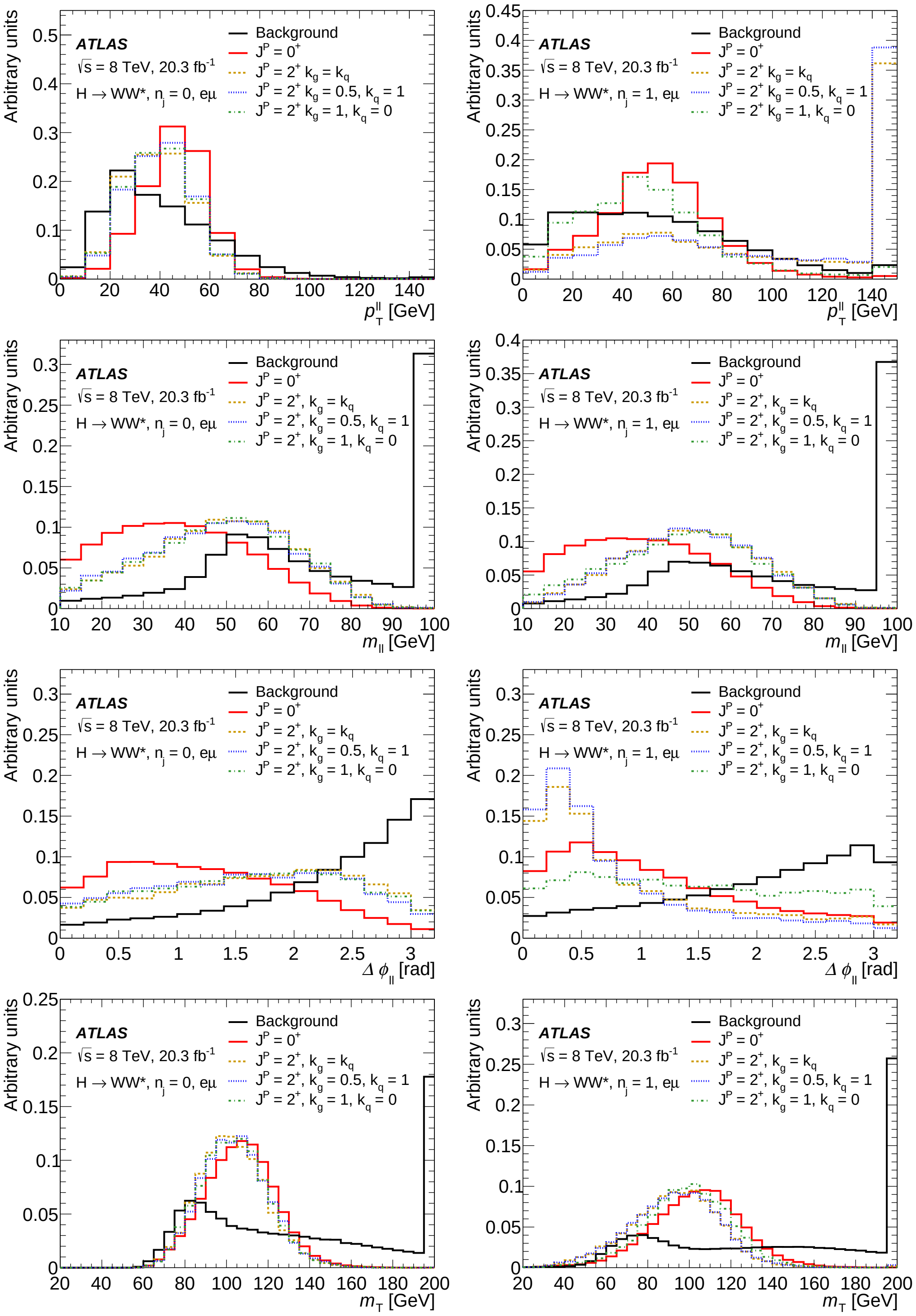


Figure 7.1.: Plots of most sensitive variables to discriminate between Standard Model (red line) and the three different spin-2 models (dotted lines with ochre, blue and green colour), namely $p_T^{\ell\ell}$ (p_T of the dilepton system), $m_{\ell\ell}$ (mass of the dilepton system), $\Delta\phi_{\ell\ell}$ (polar angle between both leptons) and m_T (transverse mass). Plots in the left column show the 0-jet channel distributions, in the right column the 1-jet channel distributions. The sum of all backgrounds is depicted as a black line. In each plot, the last bin represents the overflow content. All variables are plotted after applying the preselection criteria as listed in table 7.2. (plots from [6])

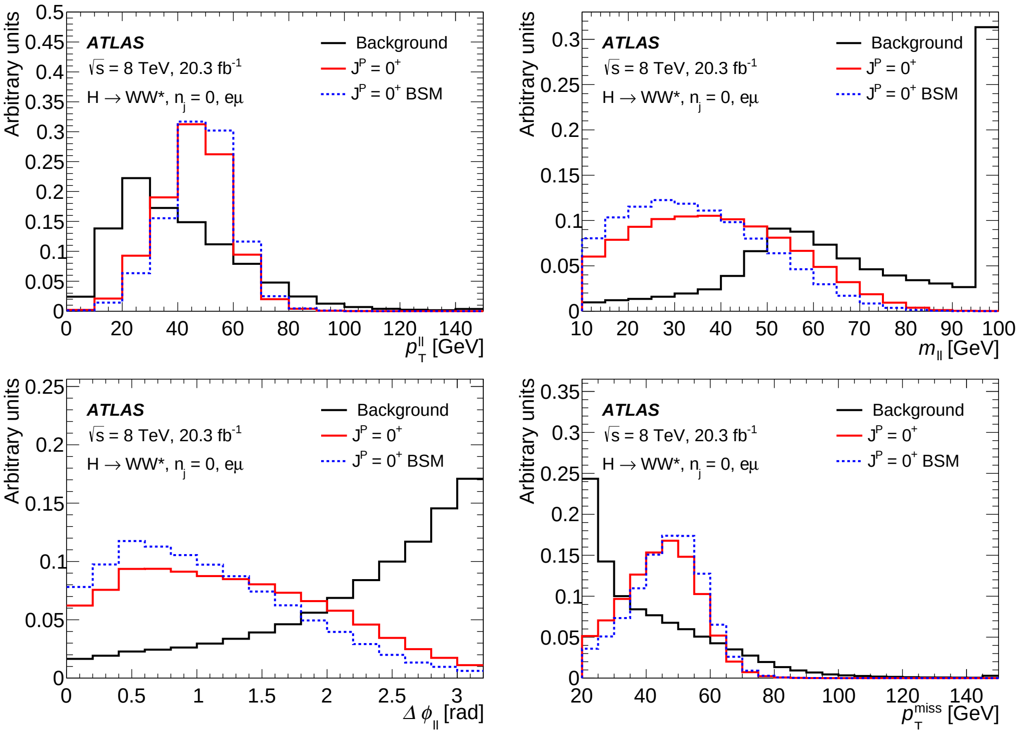


Figure 7.2.: Plots of sensitive variables to discriminate between a BSM CP-even state and the Standard Model Higgs boson. All plots are for the 0-jet channel and every variable entering the analysis BDT is shown. These are $p_{\mathrm{T}}^{\ell\ell}$, $m_{\ell\ell}$, $\Delta\phi_{\ell\ell}$ (as in figure 7.1) and $p_{\mathrm{T}}^{\mathrm{miss}}$ (missing transverse momentum). All plots represent the state after applying the preselection cuts, cf. table 7.2. (plots from [6])

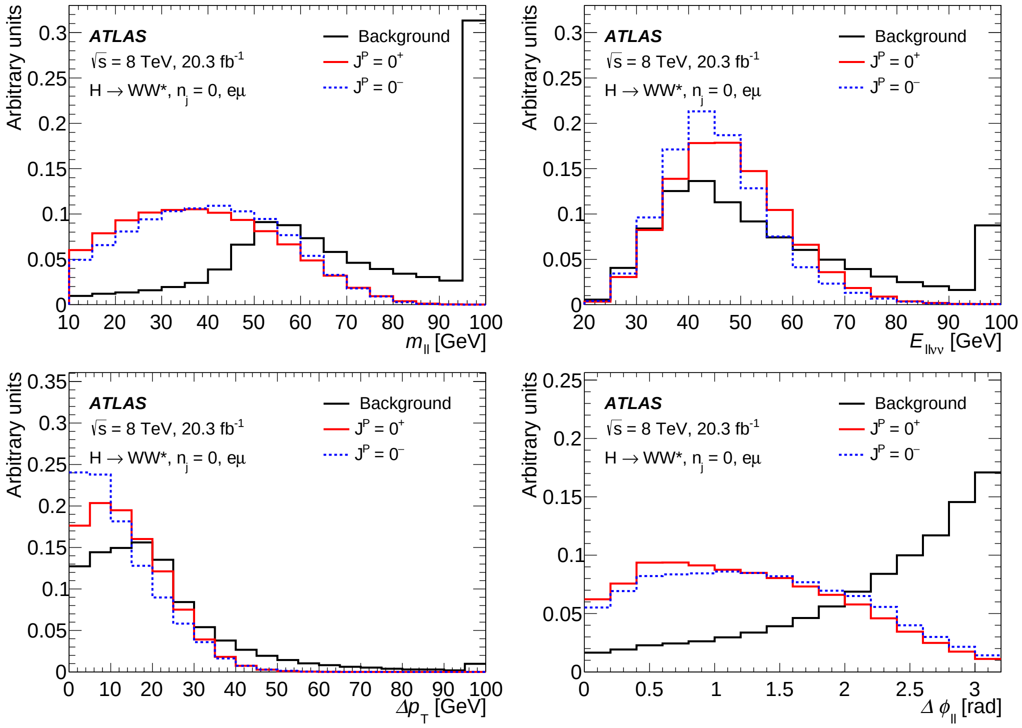


Figure 7.3.: Plots of sensitive variables for the BSM CP-odd model versus Standard Model test in the 0-jet channel. These comprise $m_{\ell\ell}$, $E_{\ell\ell\nu\nu}$ (see formula 7.6), Δp_{T} (lepton p_{T} difference) and $\Delta\phi_{\ell\ell}$. Every plot has been made after applying the preselection cuts, cf. table 7.2. (plots from [6])

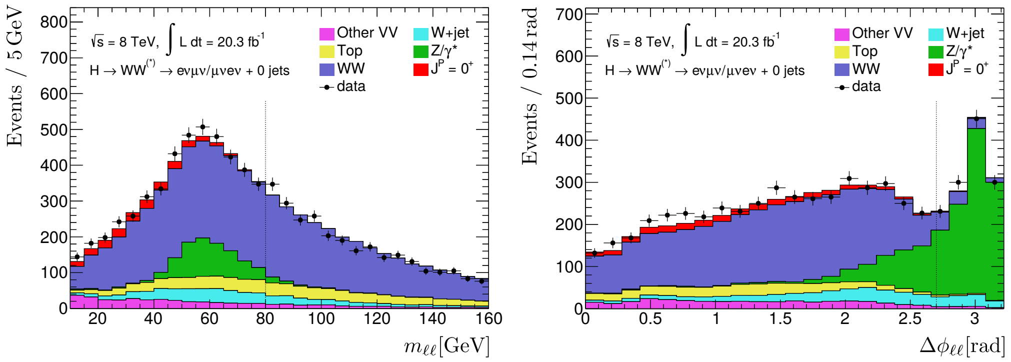


Figure 7.4.: Plots of two sensitive variables in the 0-jet category with all cuts applied besides the one on the plotted variable. The included signal (red) is a Standard Model Higgs boson sample with $m_H = 125\,\mathrm{GeV}$ at the nominal signal strength. Where available, control regions have been used for normalisation. No variations from systematics are shown. The dotted lines indicate the cut values in the signal region (cf. table 7.2).

Table 7.2.: List of selection criteria for the event preselection and the signal region definition which are applied for all presented analyses. The cuts marked with an asterisk (*) are applied only in the spin-2 analysis for the benchmarks featuring non-universal couplings to quarks and gluons (see section 6.1.1 for details). (cf. [6])

Variable	Requirements
	Preselection
N_{leptons}	Exactly 2 with $p_{\text{T}} > 10$ GeV, $e\mu$, opposite sign
$p_{\text{T}}^{\ell_1}$	> 22 GeV
$p_{\text{T}}^{\ell_2}$	> 15 GeV
$m_{\ell\ell}$	> 10 GeV
$p_{\text{T}}^{\text{miss}}$	> 20 GeV
	0-jet selection
$p_{\text{T}}^{\ell\ell}$	> 20 GeV
$m_{\ell\ell}$	< 80 GeV
$\Delta\phi_{\ell\ell}$	< 2.8
p_{T}^{H}	< 125 or 300 GeV (*)
	1-jet selection
$b-$veto	No b-jets with $p_{\text{T}} > 20$ GeV
$m_{\tau\tau}$	$< m_{\text{Z}} - 25$ GeV
m_{T}^{ℓ}	> 50 GeV
$m_{\ell\ell}$	< 80 GeV
$\Delta\phi_{\ell\ell}$	< 2.8
m_{T}	< 150 GeV
p_{T}^{H}	< 125 or 300 GeV (*)

In order to improve the Drell-Yan background suppression, a new quantity is considered. Based on the principle of collinear approximation [123], the missing transverse momentum is used to reconstruct a hypothetical τ pair mass, $m_{\tau\tau}$. The value of this variable is required to be below ($m_Z - 25$ GeV) in order to evade the Z peak and thus reduce the $Z/\gamma^* \to \tau\tau$ background. Finally, as a means to suppress W+jets contamination, a modified transverse mass combining one lepton and the missing transverse momentum is constructed as

$$m_{\text{T}}^{\ell_i} = \sqrt{2\, p_{\text{T}}^{\ell_i}\, p_{\text{T}}^{\text{miss}} \cdot (1 - \cos(\Delta\phi'))}, \quad \Delta\phi' = \angle\, p_{\text{T}}^{\ell_i}\, p_{\text{T}}^{\text{miss}}. \tag{7.7}$$

The actual variable of interest is the maximum $m_{\text{T}}^{\ell} = \max(m_{\text{T}}^{\ell_1}, m_{\text{T}}^{\ell_2})$ which is required to be above 50 GeV. This cut suppresses the W+jets background and at the same time helps reducing the Drell-Yan background. This separation of backgrounds is possible since in case of real W bosons occurring in the process (which is i.e. the case for the Higgs decay and WW background), at least for one lepton $m_{\text{T}}^{\ell_i}$ takes a large value.

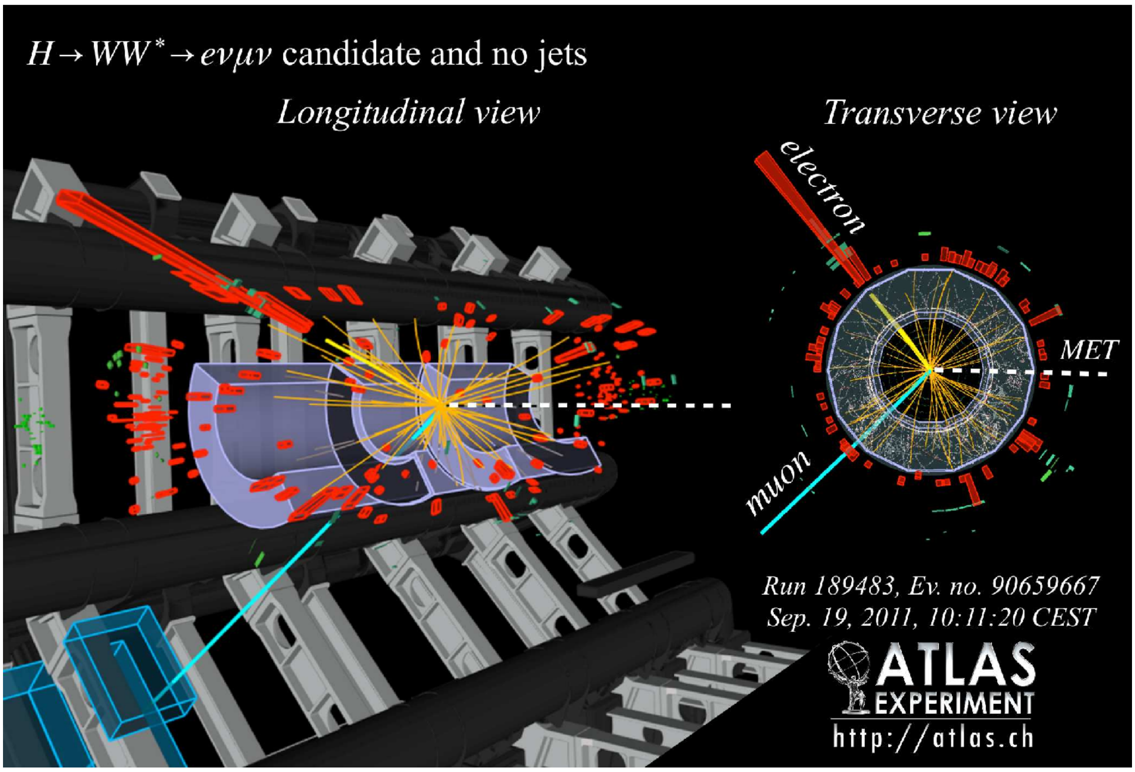


Figure 7.5.: Event display of a $H \to WW^* \to e\nu\mu\nu$ candidate event without additional jets, showing a very clear signature of the sought for decay (from the 2011 data taking with $\sqrt{s} = 7\,\text{TeV}$). On the left side a simplified cross section of the detector with indicated tracks, calorimeter entries and missing transverse momentum (dashed line) is shown. The right illustration shows a transverse projection of the event with the relevant features highlighted. The reconstructed quantities in the event are $p_{\mathrm{T}}^{e} = 33\,\text{GeV}, p_{\mathrm{T}}^{\mu} = 24\,\text{GeV}, m_{\ell\ell} = 48\,\text{GeV}, \Delta\phi_{\ell\ell} = 1.7, p_{\mathrm{T}}^{\text{miss}} = 37\,\text{GeV}$ and $m_{\mathrm{T}} = 98\,\text{GeV}$. (from [102])

In addition to the listed cuts in both the 0 and 1-jet channel, a cut on $p_{\mathrm{T}}^{\mathrm{H}}$ (the estimated transverse momentum of the Higgs boson) is applied for tests of the spin-2 model with non-universal couplings. The value of $p_{\mathrm{T}}^{\mathrm{H}}$ is computed as follows:

$$p_{\mathrm{T}}^{\mathrm{H}} = |\boldsymbol{p}_{\mathrm{T}}^{\ell\ell} + \boldsymbol{p}_{\mathrm{T}}^{\text{miss}}|. \tag{7.8}$$

An event satisfying all requirements of the (Standard Model) Higgs signal event selection that at the same time shows the characteristics in a very clear way is depicted in figure 7.5.

The resulting distributions of all sensitive variables for the spin and CP analysis are shown in figures 7.6 and 7.7. In the majority of cases a reasonable agreement of data and Monte Carlo backgrounds including a (Standard Model) signal sample is reached.

7.4.3. Cut Flow

Table 7.3 lists the number of events for the Standard Model Higgs sample (produced via gluon-gluon fusion) as well as for each background sample individually and in total. Additionally, the corresponding event numbers in data and the ratio of data over the sum of all backgrounds are listed. In each column the numbers are given after applying the full event selection. Every category contains the corresponding numbers in the 0 and 1-jet channel, including separate values for each couplings benchmark of the spin-2 analysis. As expected the number of signal events in the 1-jet

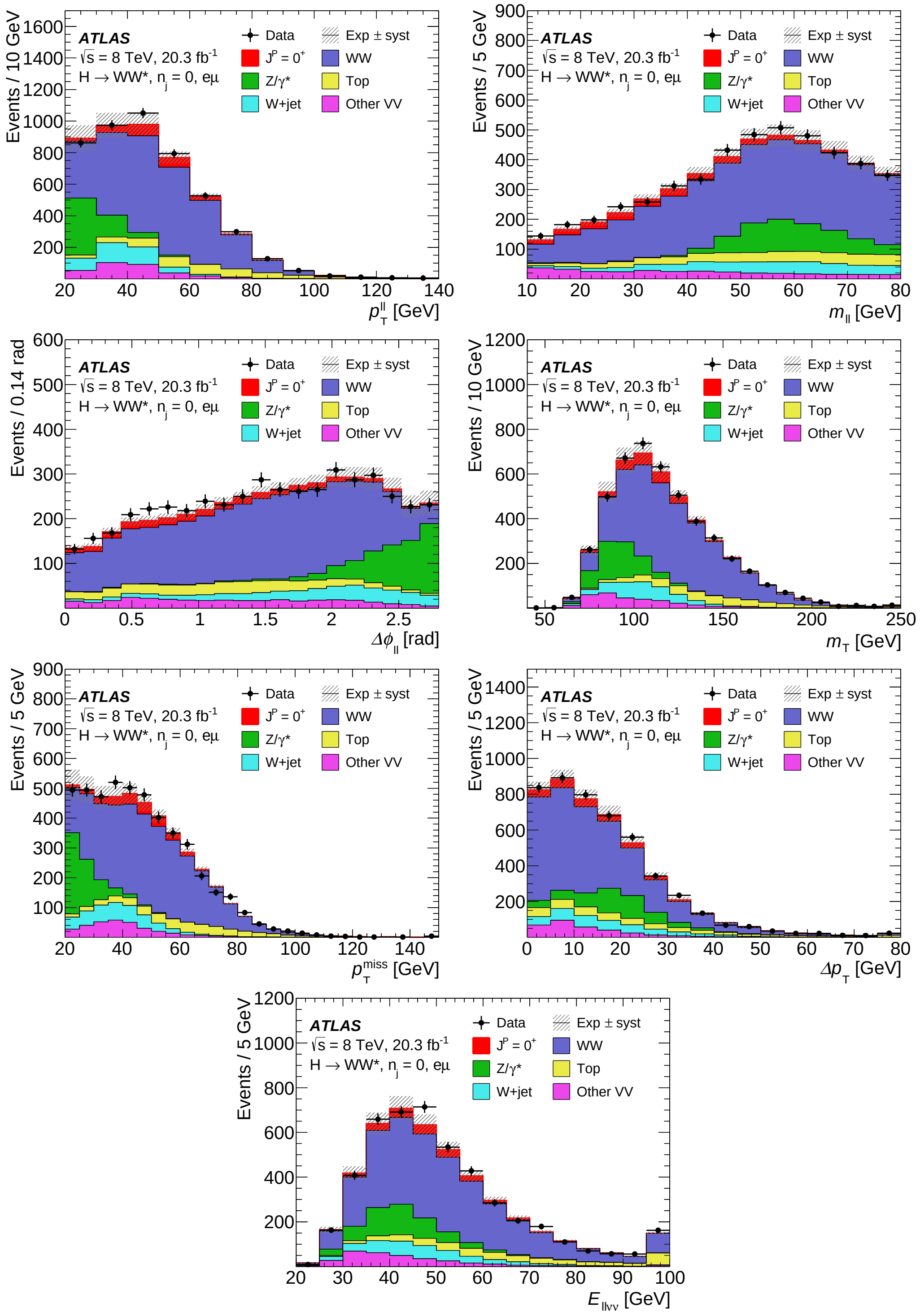


Figure 7.6.: Plots of the sensitive variables used for BDTs in the different analyses, all in the 0-jet category. The included signal (red) is a Standard Model Higgs boson sample with $m_H = 125$ GeV at the nominal signal strength. The shaded areas indicate the up and down variation from systematics. Where available, control regions have been used for normalisation. The rightmost bin in each plot contains the overflow entries. (from [6])

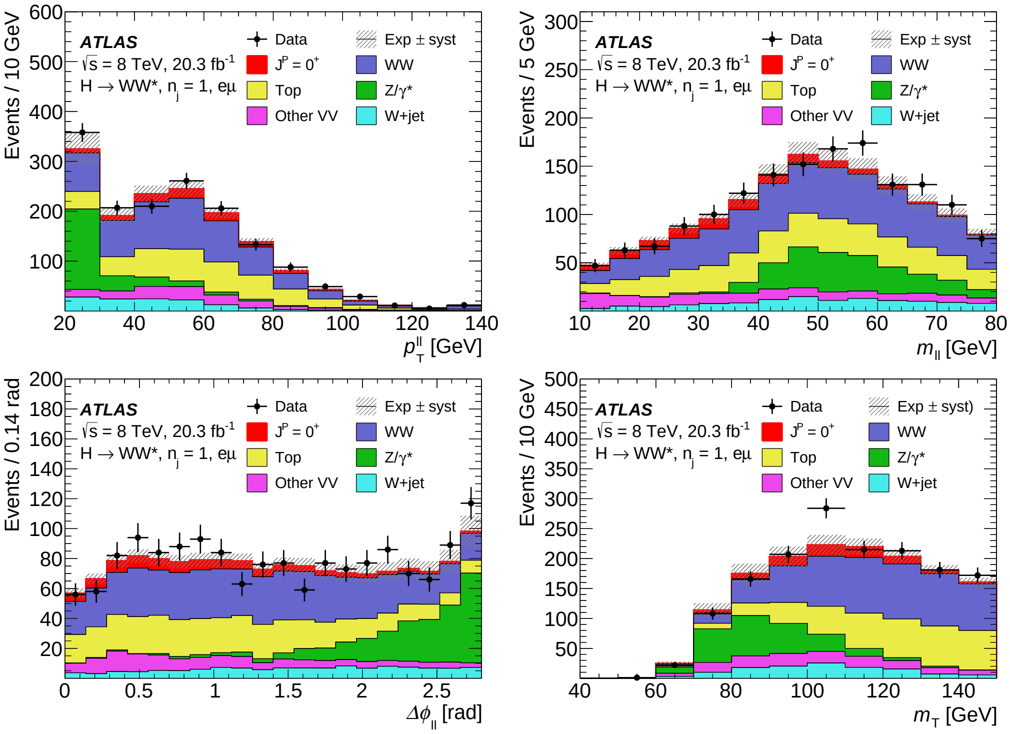


Figure 7.7.: Plots of the sensitive variables used for BDTs in the spin analysis, all in the 1-jet category. The included signal (red) is a Standard Model Higgs boson sample with $m_H = 125$ GeV at the nominal signal strength. The shaded areas indicate the up and down variation from systematics. Where available, control regions have been used for normalisation. The rightmost bin in each plot contains the overflow entries. (from [6])

Table 7.3.: List of the expected event yields in the signal regions (including the p_T^H cut variations which have a negligible impact in the 0-jet channel and are therefore omitted in the table). Background normalisations from control regions are applied for WW, $Z/\gamma^* \to \tau\tau$ and top quark backgrounds (see section 7.5 for details). The first table column contains the expected number of signal events (Standard Model Higgs boson from gluon-gluon fusion, N_{ggF}). The subsequent columns contain background event yields that are broken down into the single processes and the total number of background events (N_{bkg}). The single categories are non-resonant WW production (N_{WW}), $t\bar{t}$ and single top production ($N_{t\bar{t}}$ and N_{t}), Drell-Yan processes with Z/γ^* decaying to either a τ pair ($N_{\mathrm{DY},\tau\tau}$) or another same-flavour pair ($ee, \mu\mu$) ($N_{\mathrm{DY,SF}}$), misidentified leptons ($N_{\mathrm{W+jets}}$) and diboson (WZ, ZZ and $W\gamma$) processes. Two separated columns contain the number of events in data and the data to (expected) background ratio. The errors indicated in the last column are only the statistical uncertainties of the expected and observed event numbers so the total uncertainty including systematics would be larger. (from [6])

	N_{ggF}	N_{WW}	$N_{t\bar{t}}$	N_{t}	$N_{\mathrm{DY},\tau\tau}$	$N_{\mathrm{W+jets}}$	N_{VV}	$N_{\mathrm{DY,SF}}$	N_{bkg}	Data	Data/N_{bkg}
0j SR	218	2796	235	135	515	366	311	32	4390	4730	1.08 ± 0.02
1j SR:	77	555	267	103	228	123	131	5.8	1413	1569	1.11 ± 0.03
1j SR: $p_T^H < 300$ GeV	77	553	267	103	228	123	131	5.8	1411	1567	1.11 ± 0.03
1j SR: $p_T^H < 125$ GeV	76	530	259	101	224	121	128	5.8	1367	1511	1.11 ± 0.03

channel is way lower than in the 0-jet channel, it amounts to about 1/3 of the 0-jet events. Nearly the same ratio applies to the total backgrounds. The contribution from the 1-jet channel to the overall significance is therefore significantly lower than from the 0-jet channel. The sum of signal and background events in the table and the number of events in data agree well taking into account that the quoted errors comprise only the statistical uncertainty of the number of expected and observed events. All systematic variations need to be added on top of the errors.

7.5. Backgrounds

This section discusses the backgrounds to be considered for the properties analysis. In most cases, relevant processes have similar or identical final states as the signal process; in some cases a background plays a role due to limited identification efficiencies. The most prominent backgrounds which are discussed in this section are:

- non-resonant WW pair production,
- $t\bar{t}$ (top quark pair) and single top quark production,
- $\tau\tau$ production from Z/γ^* decays,
- W boson production with an associated jet (or multiple jets, each misidentified as a lepton).

To illustrate the extent to which each background resembles the signal process an overview of Feynman diagrams of these processes is depicted in figure 7.8. Further less prominent processes like other diboson decays ($W\gamma, W\gamma^*, WZ$ or ZZ, henceforth denoted as VV) make up a small total background contribution in the signal region. The same goes for misidentified contributions from same-flavour ($ee, \mu\mu$) final states from Z/γ^* decays. Both kinds of backgrounds are estimated from theory predictions implemented in the different Monte Carlo sample generators.

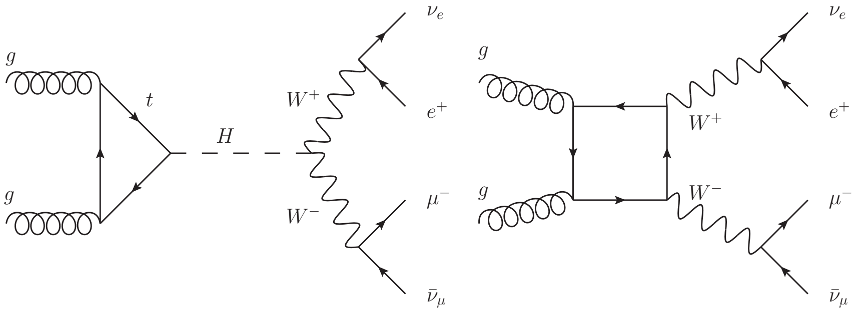
For most of the above backgrounds ($Z/\gamma^* \rightarrow \tau\tau$, single top quark/$t\bar{t}$ and non-resonant WW background), their respective contribution to the signal region is estimated from control regions. This is achieved by choosing an appropriate control region for each one where the regarded background is enriched. Fitting the Monte Carlo samples to data in the control region yields the respective normalisation which is extrapolated to the signal region (partly as well to other control regions). Details on the definition and usage of the control regions are given in section 7.5.1. For the W+jets background, a fully data-driven estimate in the signal region is available which is discussed in section 7.5.5.

The next sections discuss the defined control regions as well as their concept in general. The subsequent sections address one of these main backgrounds each.

7.5.1. Control Regions

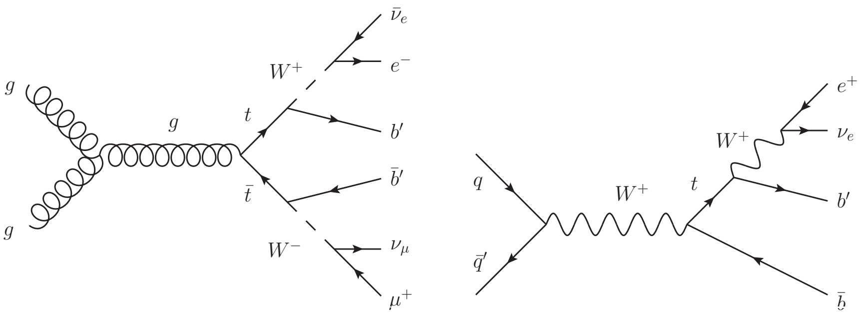
The definition of each control region is laid out orthogonal to the signal region while the alterations have been chosen such that each control region is as similar as possible to the signal region. At the same time the definition is optimized such that the purity of the background in question is optimized in its control region and a large number of events fall into the region. The procedure of laying out the control region as similar as possible to the signal region serves the purpose of keeping the extrapolation uncertainties from the control regions to the signal region as low as possible.

In each control region, a background normalisation factor is obtained by comparing the modelled contribution of the respective background to the observed yields minus the sum of all other back-



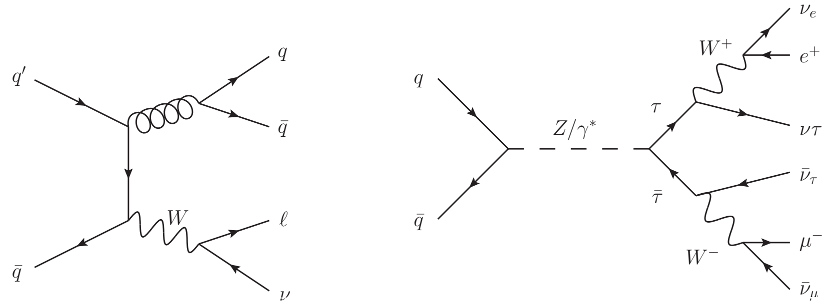
(a) Signal process

(b) Non-resonant (irreducible) WW background process



(c) $t\bar{t}$ background process

(d) Single top-quark background process



(e) W+jets background process

(f) Drell-Yan ($Z/\gamma^* \to \tau\tau$) process

Figure 7.8.: Feynman diagrams of the signal process and the most dominant background processes. In some cases specific configurations have been chosen to highlight the similarity of the respective final states.

Table 7.4.: Control region definitions for WW, single top/$t\bar{t}$ and $Z/\gamma^* \to \tau\tau$ backgrounds in the 0 and 1-jet channel (see table 7.2 for the preselection cuts). (cf. [6])

Control region	Selection
WW CR 0-jet	Preselection, $p_{\mathrm{T}}^{\ell\ell} > 20$ GeV, $80 < m_{\ell\ell} < 150$ GeV
WW CR-1-jet	Preselection, b-veto, $m_{\tau\tau} < m_{\mathrm{Z}} - 25$ GeV $m_{\mathrm{T}}^{\ell} > 50$ GeV, $m_{\ell\ell} > 80$ GeV
Top CR 0-jet	Preselection, $\Delta\phi_{\ell\ell} < 2.8$, all jets inclusive
Top CR 1-jet	At least one b-jet, $m_{\tau\tau} < m_{\mathrm{Z}} - 25$ GeV
$Z/\gamma^* \to \tau\tau$ CR 0-jet	Preselection, $m_{\ell\ell} < 80$ GeV, $\Delta\phi_{\ell\ell} > 2.8$
$Z/\gamma^* \to \tau\tau$ CR 1-jet	Preselection, b-veto, $m_{\mathrm{T}}^{\ell} > 50$ GeV, $m_{\ell\ell} < 80$ GeV, $\lvert m_{\tau\tau} - m_{\mathrm{Z}}\rvert < 25$ GeV

ground (and signal) contributions. This normalisation factor is denoted as β. The expected contribution from the respective background in the signal region can thus be modelled as

$$B_{\mathrm{SR}}^{\mathrm{est.}} = B_{\mathrm{SR}} \cdot \underbrace{\frac{N_{\mathrm{CR}}}{B_{\mathrm{CR}}}}_{\equiv \beta} = N_{\mathrm{CR}} \cdot \underbrace{\frac{B_{\mathrm{SR}}}{B_{\mathrm{CR}}}}_{=:\alpha} . \tag{7.9}$$

In this expression, N_{CR} denotes the observed number of events in the control region while B_{CR} and B_{SR} denote the expected event number from Monte Carlo predictions in the control and signal region. The ratio $N_{\mathrm{CR}}/B_{\mathrm{CR}} \equiv \beta$ is the normalisation factor mentioned above. By rearranging the expression one can introduce the ratio $B_{\mathrm{SR}}/B_{\mathrm{CR}} =: \alpha$. This variable α is the extrapolation factor from the observed event yield in the control region to the signal region. In this way the relatively large error on the background Monte Carlo sample in the signal region from theoretical uncertainties can be replaced by the relatively small extrapolation error plus the statistical error of the observed number of events in the control region. This applies as long as a sufficient number of entries fall into the control region. For the Monte Carlo samples the usual sources of uncertainties are taken into account:

- Higher orders of perturbation theory are not considered and would alter shapes and normalisations of the Monte Carlo samples. In order to estimate the arising uncertainty for the QCD calculations, the renormalisation and factorization scales are multiplied or divided by a factor of 2.

- The choice of the PDF set applied in the generation process also affects the generated sample. In order to have an estimate of this effect, the largest difference between the respective nominal sample and two other PDF sets is determined. This value and the uncertainty of the sample PDF set itself are added up in quadrature to obtain a PDF uncertainty.

- The subsequent step of the event generation is similarly fraught with uncertainties since the hadronization and parton shower process as well as the underlying event cannot be modelled exactly. This is taken into account by comparing the results from different generators employing different modelling approaches.

The control regions used for the current analysis are listed in table 7.4. As stated above they have been defined as close as possible to the signal region (cf. table 7.2). Nevertheless, in each case at least one cut is inverted or modified such that every control region is orthogonal to the signal region. As would seem natural, the choice of variables and cuts to be altered is based on insights from the signal region definition. Since the latter is optimized to suppress the diverse background-rich regions in the variable space, the starting points to enrich the backgrounds are derived in a straightforward manner.

Table 7.5.: Theoretical extrapolation factor (α) uncertainties (in percent) for the three control regions. Where applicable an extrapolation factor uncertainty to other control regions is given as well. The single columns list independent (i.e. uncorrelated) sources of uncertainty which are added up in quadrature to obtain the total uncertainty given in the rightmost column. From left to right the columns are 'Scale' which denotes the QCD factorization and renormalisation scale uncertainties, 'PDF' for PDF uncertainties, 'Gen' for uncertainties from matching the hard process matrix element to the parton showering, 'EW' to account for the missing electroweak corrections, 'UE/PS' for the uncertainties from modelling the parton shower and the underlying event (PS/UE) themselves and finally p_T^Z which denotes the uncertainty due to the respective reweighting applied in the 0-jet channel (only $Z/\gamma^* \rightarrow \tau\tau$). The negative signs in the table mark values which are anti-correlated to positive entries within the same column. Entries with identical signs are correlated. (taken from [6])

Category	Scale	PDF	Gen	EW	UE/PS	p_T^Z	Total
WW extrapolation factor							
SR 0-jet	0.9	3.8	6.9	-0.8	-4.1	–	8.2
SR 1-jet	1.2	1.9	3.3	-2.1	-3.2	–	5.3
Top-quark background extrapolation factor							
SR 1-jet	-0.8	-1.4	1.9	–	2.4	–	3.5
WW CR 1-jet	0.6	0.3	-2.4	–	2.0	–	3.2
$Z/\gamma^* \rightarrow \tau\tau$ extrapolation factor							
SR 0-jet	-7.1	1.3	–	–	-6.5	19	21.3
SR 1-jet	6.6	0.66	–	–	-4.2	–	7.9
WW CR 0-jet	-11.4	1.7	–	–	-8.3	16	21.4
WW CR 1-jet	-5.6	2.2	–	–	-4.8	–	7.7

7.5.2. Non-resonant WW Boson Production Process

The non-resonant WW production is the main, irreducible background in this analysis. Its final state is mostly in perfect agreement with the sought for Higgs decay final state (compare figure 7.8b vs. 7.8a). The only leverage point to distinguish between both processes are those kinematic variables that are affected by the difference between the resonant and non-resonant process. This difference becomes manifest in two ways: On the one hand the charged leptons originating from the Higgs decay have a smaller opening angle than the WW decay products. On the other hand the dilepton system from the non-resonant WW decay tends to have a higher mass (in the Higgs decay the invariant mass is constrained by the Higgs mass and forces one W boson to be far off-shell). Both effects reflect in the $m_{\ell\ell}$ distribution by favouring higher values. This causes this variable to be suitable for defining the corresponding control region. Herein, all cuts correspond to the signal region up to the cut on $m_{\ell\ell}$ (see table 7.2) where the adjacent region towards higher values of $m_{\ell\ell}$ has been chosen. The $\Delta\phi_{\ell\ell}$ cut following the $m_{\ell\ell}$ cut in the signal region is omitted.

For the 1-jet channel, the $m_{\ell\ell}$ cut was simply inverted, i.e. $m_{\ell\ell} > 80\,\mathrm{GeV}$ is required. In the 0-jet channel it has turned out that introducing an additional upper cut on $m_{\ell\ell}$ still increases the purity in the control region such that $80 < m_{\ell\ell} < 150\,\mathrm{GeV}$ is required. The purity of the WW background within the WW control region is supposed to be 69 % in the 0-jet channel and 43 % in the 1-jet channel. Two variables crucial for the analysis, $m_{\ell\ell}$ and $\Delta\phi_{\ell\ell}$ are plotted in figure 7.9 for the zero and one jet channel. From the plots, the mentioned background composition may be inferred.

With regard to the scale of contributions from top quark and $Z/\gamma^* \rightarrow \tau\tau$ backgrounds within the WW control region, both are incorporated into a global fit of the background normalisations. This fit makes use of the assigned control region where the respective background is enriched. This approach implicitly constrains the background normalisation in all other control regions and the signal region. The same holds true the other way around for the WW background in the other control regions.

With respect to the extrapolation uncertainties associated with the extrapolation factor α (cf. formula 7.9) two additional contributions are considered for the WW control region:

- A generator uncertainty due to deviating models that are applied in different Monte Carlo sample generators is considered, estimated by comparing results from a POWHEG+HERWIG combination on the one hand and aMC@NLO+HERWIG on the other hand.
- Predictions for higher-order electroweak corrections are available that are not incorporated into the Monte Carlo production process. They are taken into account by reweighting the Monte Carlo samples to the NLO results and deriving and uncertainty from the observed variation.

7.5.3. Top Quark Processes Control Region

Different processes contribute to the top-quark background, two of which are displayed in figures 7.8c ($t\bar{t}$) and 7.8d (single top quark). The normalisation in the top control region is obtained and applied regardless of the precise production processes.

The 0-jet control region is different from the other control regions in as much as it only serves to compute a normalisation factor a-priori without including it into the global fit. This simplified approach is chosen since in the 0-jet channel the top background contributions in the signal region and in other control regions are small so the impact in the global fit would be negligible. The region is defined on the basis of the common preselection cuts plus the additional $\Delta\phi_{\ell\ell}$ cut which also gets applied in the signal region to constrain the contributions from the Drell-Yan ($Z/\gamma^* \rightarrow \tau\tau$) background. Furthermore, the control region is inclusive in terms of jet channels, i.e. all numbers of jets in an event are accepted. The purity within this control region is 74 %. The extrapolation factor α (cf. 7.9) is obtained as explained in the respective section but additional data-driven corrections are applied (see [102] for the application and [124] for an in-depth description of the technique). The approach is summarized briefly in the following: By requiring one b-tagged jet (i.e. a jet assumed to originate from a b quark decay) it is possible to select a nearly pure top background region in data and Monte Carlo. The extrapolation factor α is then corrected by the factor

$$\gamma_{1b} = \left(\frac{\alpha^{1b}_{\mathrm{data}}}{\alpha^{1b}_{\mathrm{MC}}} \right)^2 . \tag{7.10}$$

In this expression, $\alpha^{1b}_{\mathrm{sample}}$ denotes the rate of events without additional jets beyond the b-tagged jet in the respective sample (data or Monte Carlo). This factor helps reduce systematic effects which

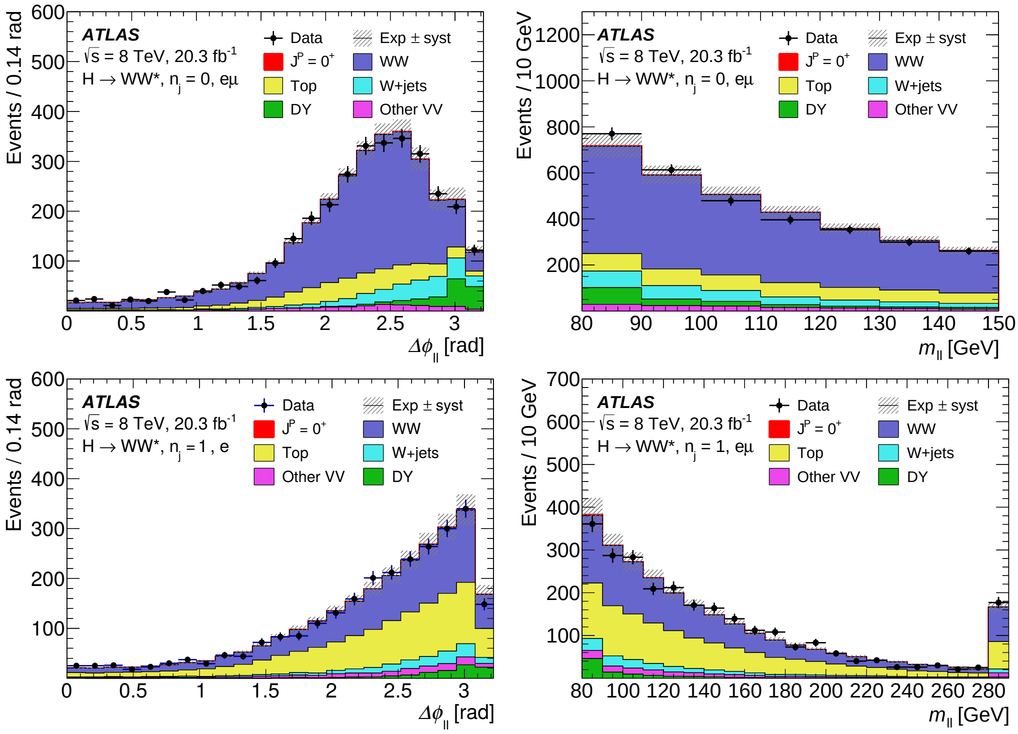


Figure 7.9.: WW control region plots of two variables of interest, $m_{\ell\ell}$ and $\Delta\phi_{\ell\ell}$ in the 0 and 1-jet channel. The signal process included in the plots is a Standard Model Higgs sample with $m_H = 125\,\text{GeV}$. (cf. [6])

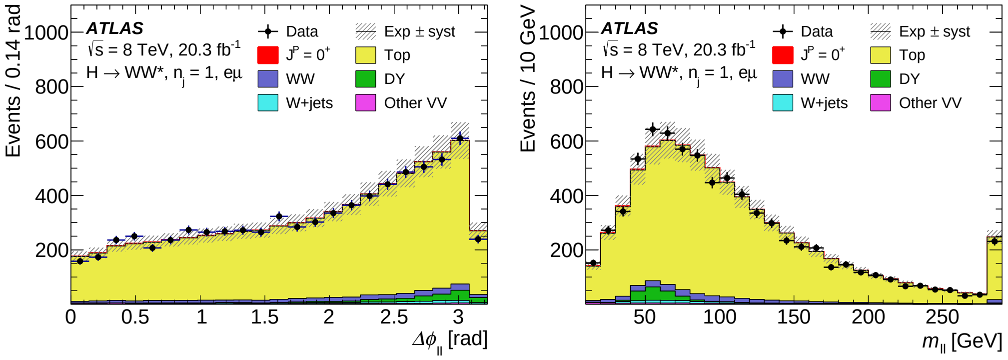


Figure 7.10.: Plots of control region for $t\bar{t}$ and single top quark processes. The control region exists only for the 1-jet channel; the variables shown are $\Delta\phi_{\ell\ell}$ and $m_{\ell\ell}$. The signal process included in the plots is a Standard Model Higgs sample with $m_H = 125\,\text{GeV}$. (cf. [6])

occur in both the inclusive control region and the b-tagged region, such as the jet energy scale and jet resolution. At the same time a number of (small) additional systematic uncertainties are added through this correction factor which originate from different topologies of the control region and the 1 b-tag region. This comprises among others the relative cross sections of the processes contributing to the top background.

The normalisation is computed independently from the final global fit and yields a factor of 1.08 ± 0.02 (stat.) with a total uncertainty of 8.1 % including all relevant uncertainties.

The top control region in the 1-jet category basically follows the same mechanisms as the WW control region. The cuts defining the 1-jet top control region require exactly one b-tagged jet where the signal region implements a b veto. The $m_{\tau\tau}$ and m_{T}^{ℓ} cuts are retained but all subsequent cuts from table 7.2 are omitted. Figure 7.10 shows the same sensitive variables as before for WW, namely $\Delta\phi_{\ell\ell}$ and $m_{\ell\ell}$ for the current control region.

The top quark background contribution in the signal region is the second largest and spoils the WW control region (about 40 % of all events in the 1-jet WW CR are from top processes) as well as the signal region. For this reason two extrapolation factors are defined: α_{SR} denotes the extrapolation factor to the signal region and α_{WW} the one to the respective control region. The uncertainties associated with the extrapolation factors α are listed in table 7.5 and arise from the usual sources (cf. table columns) plus specific contributions. These comprise the cross sections of the $t\bar{t}$ and single top processes and their interference.

7.5.4. Drell-Yan Backgrounds

The specific Drell-Yan background adding the largest contribution is the $Z/\gamma^* \to \tau\tau$ decay where both τ leptons decay leptonically (see 7.8f). This final state features identical particles as the signal process plus two additional neutrinos which only appear as modifications to the missing transverse energy. In the 0-jet channel, the $Z/\gamma^* \to \tau\tau$ control region (as given in table 7.4) is defined by applying the preselection cuts and the signal region $m_{\ell\ell}$ cut ($m_{\ell\ell} < 80\,\mathrm{GeV}$). On top of these, an inverted $\Delta\phi_{\ell\ell}$ cut with respect to the signal region definition is applied. The reason the $\Delta\phi_{\ell\ell}$ variable is crucial for the discrimination of this background is the fact that the τ leptons are emitted back-to-back leading to a large angular separation also between the decay products. Thus, as can be seen in the plots of the two variables $\Delta\phi_{\ell\ell}$ and $m_{\ell\ell}$ for the $Z/\gamma^* \to \tau\tau$ control region in figure 7.11, one achieves a very pure control region with an expected purity of 90 %. For 1-jet the contribution from other backgrounds gets a little more prominent and the purity is decreased to 80 %. Up to the m_{T}^{ℓ} cut the cuts are identical to the signal region (b-jet veto, $m_{\mathrm{T}}^{\ell} > 50\,\mathrm{GeV}$) besides the reversed cut on $m_{\tau\tau}$, i.e. $|m_Z - m_{\tau\tau}| < 25\,\mathrm{GeV}$ is required.

In the case of the $Z/\gamma^* \to \tau\tau$ background as for the top quark background two extrapolation factors with their respective uncertainties need to be considered because of a contamination of the WW control region with $Z/\gamma^* \to \tau\tau$ events ($\sim 4-5\,\%$). Thus, there are again two extrapolation factors, α_{SR} and α_{WW}.

One specifically notable correction needs to be applied to the Monte Carlo samples of the $Z/\gamma^* \to \tau\tau$ background. Especially in the 0-jet channel, the ALPGEN + HERWIG generator does not model the p_{T} of the Z boson correctly, i.e. the reconstructed $p_{\mathrm{T}}^{\ell\ell}$ distribution is distorted. In order to fix this issue the sample has been reweighted using weights from data-to-MC comparisons at the Z peak. The arising additional uncertainty is accounted for in the list of uncertainties in table 7.5.

7.5.5. Treatment of Misidentified Leptons

As mentioned above the W+jets background is the last of the four main backgrounds for this analysis. The reason this background is sometimes misinterpreted as signal is the misidentification

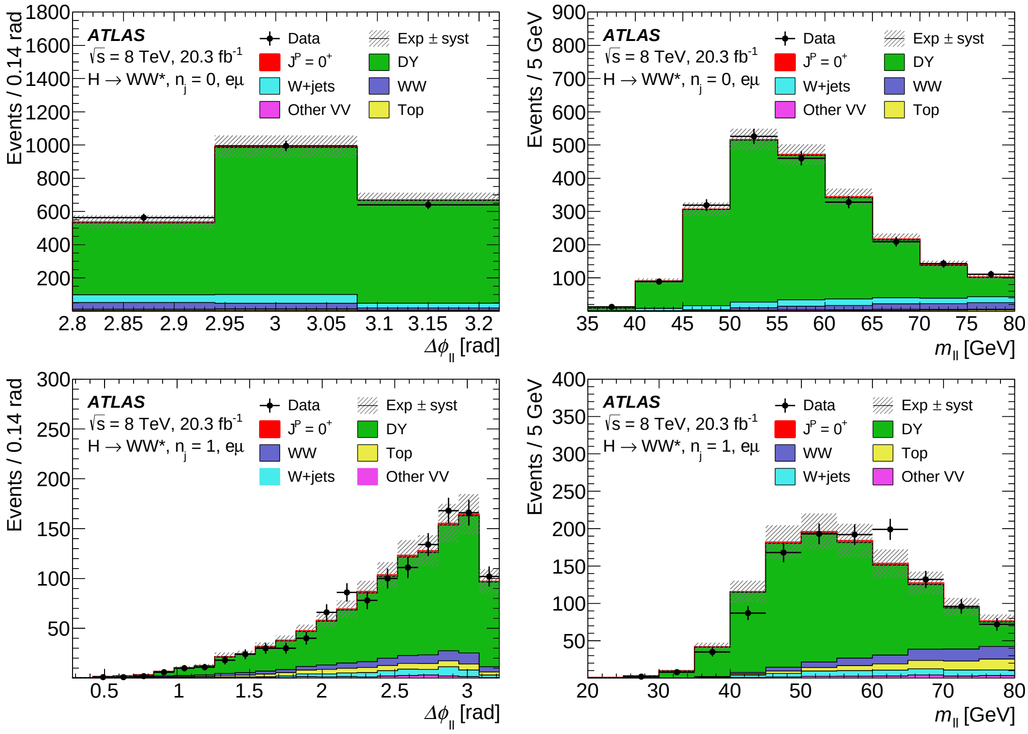


Figure 7.11.: $Z/\gamma^* \to \tau\tau$ (Drell-Yan) control region plots of two variables of interest, $m_{\ell\ell}$ and $\Delta\phi_{\ell\ell}$ in the 0 and 1-jet channel. The signal process included in the plots is a Standard Model Higgs sample with $m_H = 125\,\text{GeV}$. (cf. [6])

of the radiated jet or of a lepton within this jet originating not from the primary vertex but from a decay within the jet itself as a 'primary' lepton. The modelling of this background does not rely on a control region but is estimated in a fully data-driven manner. The exact approach is described in [102] while only a brief description is given in the following.

For the purpose of deriving a background estimate, a W+jets control sample is constructed from data by requiring one lepton satisfying all lepton reconstruction requirements (w.r.t. identification and isolation) and one lepton which only fulfils looser classification requirements. The latter leptons are called "anti-identified" leptons. The resulting dataset is mostly composed of W+jets background events, i.e. events where a jet has been mis-identified as a lepton. This type of event makes up 85–90 % of the selected sample. In order to obtain an approximation of the W+jets contribution in the signal region the event content of this enriched area needs to be extrapolated to the signal region. The corresponding extrapolation factor is estimated from a Z+jets control sample in data, where a Z boson decays to an ee or $\mu\mu$ final state plus additional jets. This approach is chosen since the Z+jets process is rather similar to the W+jets process from most aspects but can be easily identified via the reconstructed boson masses. From the Z+jets control sample the extrapolation factor is obtained as the ratio of the number of leptons passing all selection criteria over the number of anti-identified leptons. Remaining contaminations are accounted for by applying corrections estimated from Monte Carlo simulation. Since the factor is highly η- and p_T-dependent, it is always applied in bins of η and "anti-identified" lepton p_T.

Differences arising from the different nature of Z+jets and W+jets sample which mostly concern the jet composition are compensated by deriving approximate corrections from Monte Carlo simulations. The latter are as well used to estimate the related uncertainties. The effect of limited statistics of the Z+jets sample and the uncertainties of the subtracted additional physics processes within the sample (as mentioned above) are equivalently accounted for by adding corresponding uncertainties.

The combined uncertainties vary for muons and electrons and are furthermore dependent on the p_{T} of the anti-identified lepton. For electrons they range from 29–61 % while for muons they are a bit lower, ranging between 25–46 %. Overall it is the uncertainty associated with the (corrected) extrapolation factor and not the one of the input W+jets sample itself which dominates the total uncertainty of the W+jets background.

CHAPTER 8

Boosted Decision Tree Application

8.1. Boosted Decision Tree Overview

In both the spin and CP analysis a combination of two boosted decision trees (see section 5.1 for details) is used to form the separating variable for the final binned likelihood fit of the data to the respective models. The BDTs are trained to either separate one signal from all backgrounds (spin and CP analysis) or one signal hypotheses from the other (CP analysis). The final discriminant is generated by merging the two single BDT output variables. Technically, the training inputs are on the one hand the various signal Monte Carlo samples and on the other hand the stacked samples of all backgrounds weighted by their production cross sections. The implementation of the BDT algorithm including the training, application and cross checks (over-training test, ROC curve (*receiver operating characteristic*) etc.) is provided by the TMVA framework [90].

For the training of the BDTs, a slightly altered event selection is applied with respect to the nominal cuts listed in table 7.2. The preselection cuts are applied in both jet channels but some further cuts are omitted or loosened. This is done in light of the fact that higher statistics improve the results from the BDT training step. To the required extent the background suppression should still be given of its own accord since the training is anyway meant to yield a discriminant telling signal and background apart. Thus, the $m_{\ell\ell}$ cut is loosened to 100 GeV (instead of requiring $m_{\ell\ell} < 80\,\mathrm{GeV}$ in both channels) and the $\Delta\phi_{\ell\ell}$ cut is omitted in the 0-jet channel. The cuts on $p_{\mathrm{T}}^{\mathrm{H}}$ for non-universal couplings are applied as in the signal region. The same goes for the lower cut on $p_{\mathrm{T}}^{\ell\ell}$ in the 0-jet channel, i.e. $p_{\mathrm{T}}^{\ell\ell} > 20\,\mathrm{GeV}$ is also required for the BDT training. In the 1-jet channel the $\Delta\phi_{\ell\ell}$ cut ($\Delta\phi_{\ell\ell} < 2.8$) is applied as in the signal region to reduce the large $Z/\gamma^* \rightarrow \tau\tau$ background.

8.2. Training Parameters and Choice of Variables

The training of all BDTs was done with identical settings which mostly correspond to the default settings in the TMVA framework. For the boosting of the decision trees, the Gradient Boosting algorithm described in the introductory section 5.1 is deployed. Gradient Boosting is used with an additional option to run the tree growing on random subsets of the training dataset instead of the full set of events. This technique is called *bagging* and serves the purpose of stabilizing the tree generation against overtraining effects. In each bagging step, half the number of events with respect to the original training dataset were used.

In total for each training scenario, 400 trees have been grown. The growth of each tree branch was terminated when either a maximum depth of the overall three or a minimal node size (i.e. number of events in a leaf node after the preceding cuts) of 10 % of the original training sample size is reached. Consequently, this corresponds to the default value of 5 % of the original training dataset required in each leaf when taking the bagging step into account. Finally, the shrinkage parameter is set to 0.1.

Each training is done twice using either the odd numbered events as training and the even numbered events as cross check samples or vice versa. The resulting weights from both runs are combined into one BDT response by taking the maximum of both qualifier values when applying the BDT to data. The Monte Carlo samples are classified using the training performed on the other half of the Monte Carlo dataset in each case.

8.2.1. Spin-2 Analysis Training

For the spin-2 analysis two BDTs, henceforth denoted as BDT_0 and BDT_2, are trained. BDT_0 is trained using the (spin-0) Standard Model Higgs sample as the signal sample input and the sum of all backgrounds as the background sample input. BDT_2 is trained with the spin-2 model sample versus all backgrounds. In preparation of the first publication regarding spin tests in the $H \to WW$ channel [2], some studies were conducted regarding the optimal BDT configuration. These have shown that the described configuration of two BDTs, trained each with a signal hypothesis versus backgrounds, outperforms a single BDT as well as using one BDT for background suppression and one to discriminate between both signal hypotheses.

The training is done independently for all spin-2 benchmark models and in both jet channels. The variables offering the best separating power in the BDTs are $m_{\ell\ell}$, $p_{\mathrm{T}}^{\ell\ell}$, $\Delta\phi_{\ell\ell}$ and m_{T} which was validated in both jet channels.

Plots of the BDT outputs for BDT_0 and BDT_2 are shown in figures 8.1 and 8.2 with the Standard Model signal on top of all backgrounds, including uncertainties and data points. Figure 8.1 shows the comparison of the Standard Model distribution to the universal couplings scenario ($\kappa_g = \kappa_q$) in both jet channels. Figure 8.2 shows the same kind of distribution for the two non-universal couplings benchmarks as indicated in the plots. In all cases the $p_{\mathrm{T}}^{\mathrm{H}} < 125\,\mathrm{GeV}$ cut has been applied (the corresponding distributions for the 300 GeV upper cut on $p_{\mathrm{T}}^{\mathrm{H}}$ look similar). The Standard Model output distribution in figure 8.2 differs from the one in figure 8.1 due to the additionally applied upper cut on $p_{\mathrm{T}}^{\mathrm{H}}$. In general, the separation of the Standard Model signal from all backgrounds tends to be more powerful than for any alternative signal. This is a consequence of the higher similarity of the spin-2 models to the WW background, making the discrimination between both less powerful.

A number of cross checks were conducted to support the applicability and correct functioning of the above BDT analysis approach. These cover the following aspects each of which was evaluated successfully:

- Exclude possible overtraining of the BDTs by evaluating the Kolmogorov-Smirnov test values computed from the comparison of the BDT response of training events and remaining (test) events. The resulting KS probabilities (numbers closer to 100 % indicate a better agreement) are listed in table 8.1. Similar results were obtained for the other spin-2 benchmarks. All test results can be regarded sufficient for the safe BDT application.

- Verify the correct modelling of correlations in the Monte Carlo samples compared to data. For this reason, mean values of input variable correlations as well as profile plots of variable pairs were checked. In the vast majority of cases the modelling of correlations agrees well between the simulation samples and data. A minor divergence in the $Z/\gamma^* \to \tau\tau$ control region is covered by the uncertainty band and can therefore be tolerated.

Table 8.1.: Overview of Kolmogorov-Smirnov test results to check for BDT overtraining, listed for a set of exemplary benchmarks.

Spin 2, minimal couplings $\kappa_g = \kappa_q$		
0-jets	Signal	Background
BDT_0	77 %	32 %
BDT_2	95 %	73 %
1-jet		
BDT_0	76 %	13 %
BDT_2	66 %	11 %
CP-odd/even		
0-jets	Signal	Background
BDT_0	42 %	28 %
BDT_{CP}	15 %	100 %

- Check the correct modelling especially of the WW background in the BDT response. The entries in the background-enriched control regions populate only few bins of the BDT output which is why the correct modelling cannot be checked in detail. This is why for the WW (0 and 1-jet channel) and top control region (1-jet channel) an additional test was conducted. In each case a BDT was trained omitting the signal sample and using the specific background sample (corresponding to the control region) as the signal. This signal was trained against the remaining backgrounds and the resulting distributions were compared. The resulting KS test values from comparisons to data (applying control region normalisation factors) are 0.58 and 0.23 (WW control region, 0 and 1-jet channel) and 0.39 (top control region, 1-jet channel). In summary, the agreement was confirmed to be good over the entire phase space.

8.2.2. CP Analysis Training

The CP analysis BDT training is essentially identical to the approach for the spin analysis with some minor adjustments discussed in the following. Instead of training both BDTs versus backgrounds, studies have shown that given the smaller kinematic differences between different CP scenarios, the separation power can be optimized by using one BDT to solely discriminate between the two signal hypotheses and one BDT to discriminate between a signal and the sum of backgrounds. The necessity to optimize the discrimination between the alternative signal and a specific similar background (like WW for spin-2) is not given for the CP analysis, either. The signal versus background BDT is henceforth denoted as BDT_0 as in the above case, while the BDT discriminating between both signal hypotheses is denoted as $\mathrm{BDT}_{\mathrm{CP}}$.

Individual BDTs were trained for testing the CP-odd model state versus the Standard Model state on the one hand and for testing the non-Standard-Model CP-even state versus the Standard Model state. The training parameters for the $\mathrm{BDT}_{\mathrm{CP}}$ are slightly different to the parameters presented above in as much as the minimal leaf node size is reduced to 5 % since more subtle shape differences need to be resolved. The training setup and choice of variables of BDT_0 is identical to the spin analysis. The approach of training and applying the BDT on odd-numbered and even-numbered events independently, which is described in the spin-2 section above, was pursued identically for the CP analysis.

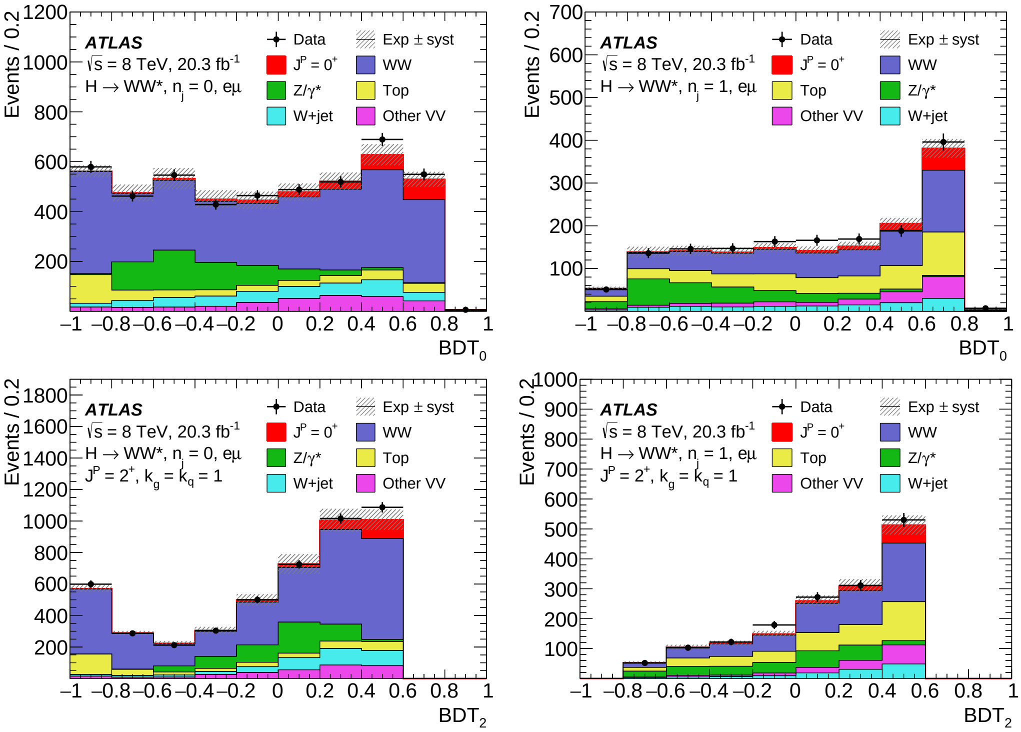


Figure 8.1.: Outputs of BDT_0 and BDT_2, both applied to signal region events (respectively spin-0 and 2 hypothesis vs. backgrounds) for universal couplings ($\kappa_g = \kappa_q$) in the 0 (left column) and 1-jet (right column) channel. The signal hypothesis in the plot is the Standard Model Higgs boson with a mass of 125 GeV at the nominal signal strength. The shaded areas indicate the up and down variation from systematics. Where available, control regions have been used for normalisation. (from [6])

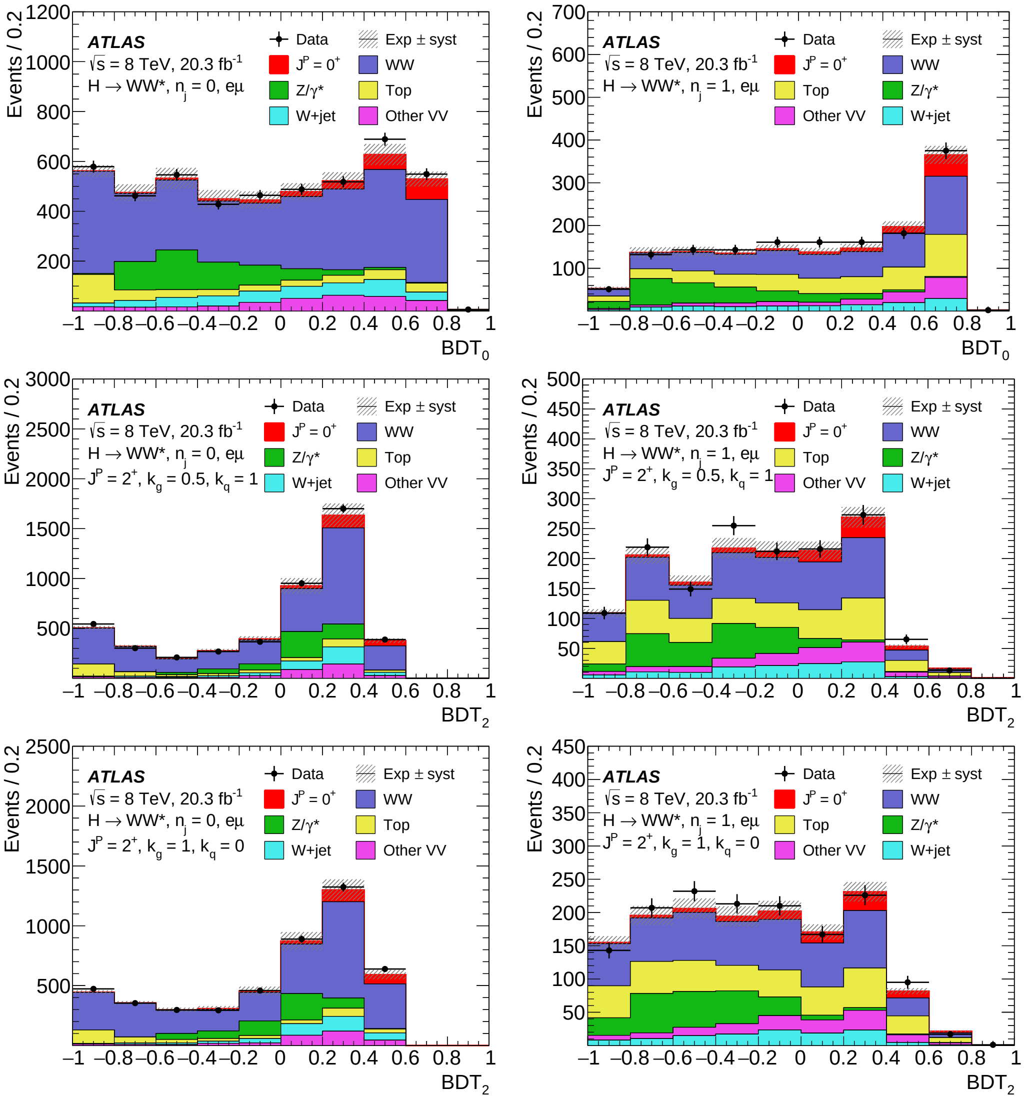


Figure 8.2.: Outputs of BDT_0 and BDT_2, both applied to signal region events for non-universal couplings ($\kappa_g = 0.5, \kappa_q = 1$ and $\kappa_g = 1, \kappa_q = 0$) respectively in the 0 (left column) and 1-jet (right column) channel. The top row plots show outputs from BDT_0 which is trained with a spin-0 (SM) model versus backgrounds. The middle and bottom rows show outputs from BDT_2 for two spin-2 benchmark models trained versus backgrounds each. The signal hypothesis in the plot is the Standard Model Higgs boson with a mass of 125 GeV at the nominal signal strength. The shaded areas indicate the up and down variation from systematics. Where available, control regions were used for normalisation. (from [6])

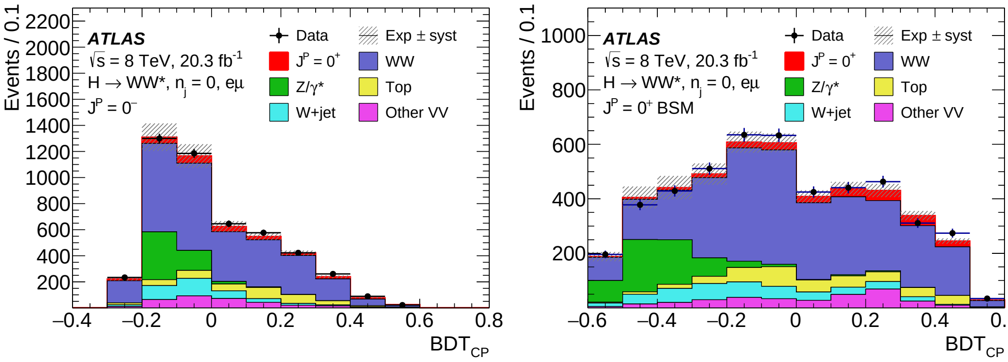


Figure 8.3.: Outputs of BDT_{CP} which is applied to signal region events and has been trained to tell apart the Standard Model and alternate CP models. The plots show the two different analysis scenarios (CP-odd and BSM CP-even) in the 0-jet channel. The signal hypothesis in the plot is the Standard Model Higgs boson with a mass of 125 GeV. (cf. [6])

Investigating the 1-jet channel for the CP analysis has been omitted due to technical difficulties concerning the implementation of the matrix element reweighting for this channel. Furthermore, the spin-0 states lack a discriminating feature such as the high Higgs-p_{T} tail as seen in the spin-2 model case so the expected gain in sensitivity is lower in comparison. For these reasons the CP analysis BDTs are only trained for the 0-jet channel.

The choice of variables differs between the two CP scenarios:

- For the CP-odd analysis, the most powerful variables are $m_{\ell\ell}$, $\Delta\phi_{\ell\ell}$, $E_{\ell\ell\nu\nu}$[1] and Δp_{T}. The latter denotes the absolute value of the difference between the leading and subleading lepton momenta.
- For the BSM CP-even analysis slightly different variables have been found ideal (close to the BDT_0 selection): $m_{\ell\ell}$, $\Delta\phi_{\ell\ell}$, $p_{\text{T}}^{\text{miss}}$ and $p_{\text{T}}^{\ell\ell}$.

The overtraining check and correlation tests mentioned in the spin section have also been performed for the CP analysis BDTs yielding satisfying results here as well.

The resulting distributions of the BDT_{CP} classifier for both CP analysis modes applied to the stacked backgrounds plus a Standard Model Higgs signal are shown in figure 8.3.

---

[1] Reminder: $E_{\ell\ell\nu\nu} = p_{\text{T}}^{\ell_1} - \frac{1}{2}p_{\text{T}}^{\ell_2} + \frac{1}{2}p_{\text{T}}^{\text{miss}}$ where $p_{\text{T}}^{\ell_1}$ and $p_{\text{T}}^{\ell_2}$ denote the transverse momentum of the leading (ℓ_1) and subleading lepton (ℓ_2) (equation 7.6).

CHAPTER 9

Boosted Decision Tree Output Rebinning Approach

As explained in the previous chapter, two BDT classifiers are employed in parallel for the analysis. Since the resulting two-dimensional plane is sparsely populated in some regions and the availability of simulated Monte Carlo events is limited, a procedure for mapping and rebinning the BDT outputs to a one dimensional distribution is applied. This is done in order to avoid empty bins in the fit and to reduce the fit complexity. Two approaches have been tested and are explained in more detail. The first section in the following explains the BDT output rebinning approach as used for the publications in 2015 [6] which serves the purpose of avoiding fit instabilities due to background fluctuations. In the second section, a more complex rebinning algorithm is explained, which has been designed to improve the achieved sensitivity while also maintaining the fit stability.

From a technical perspective the newly added support for both approaches has made the background shape systematics treatment easier and more flexible with respect to earlier analysis iterations.

9.1. Simple Rebinning

The simple rebinning approach is fairly straight-forward and merges adjacent bins in cases where the sum of all backgrounds in a bin falls below a specified threshold. The procedure has been applied independently for every benchmark of the spin-2 analysis as well as for both CP (mixing) tests. Separate binning configurations are necessary because the event numbers and BDT training results differ in each case.

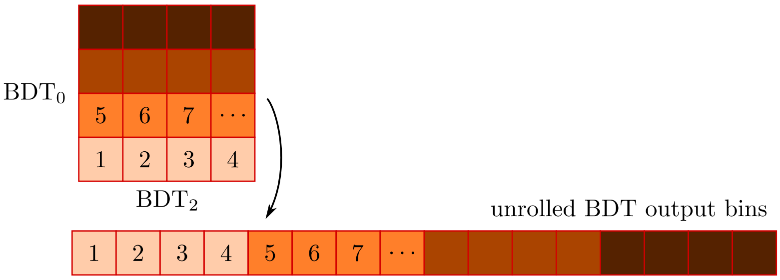
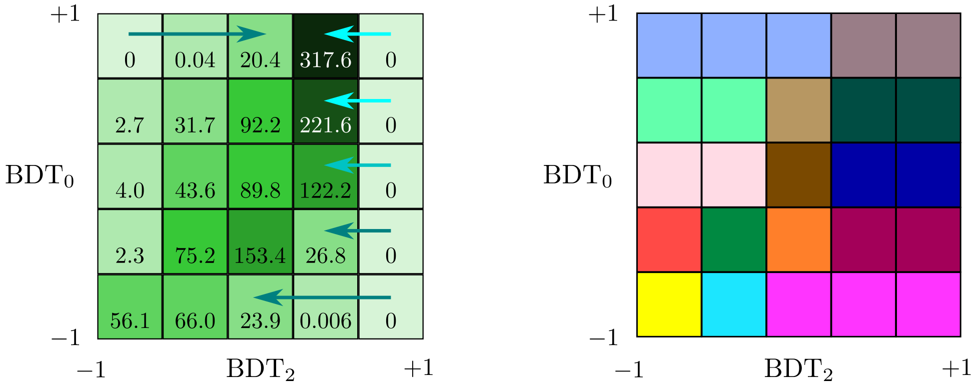


Figure 9.1.: Unrolling of the two dimensional BDT output bin grid to one dimension as used for the simple binning approach.



(a) Schematic illustration of the simple rebinning algortihm in the 1-jet channel (using 5 × 5 bins for the BDT output plane) as used for the results in [6].

(b) Actual binning configuration for the minimal couplings benchmark model in the 1-jet channel. Merged bins are highlighted in identical colours in the 2D plane.

Figure 9.2.: Simple rebinning example illustrations.

The starting point for this algorithm is the two dimensional plane filled with all events in the signal region at their respective BDT response value coordinate (BDT_2, BDT_0). In a first step, a grid is introduced splitting the two dimensional BDT qualifier output plane into single bins. The number of bins is chosen differently for the single analyses and jet channels. For the spin tests a 10×5 ($\mathrm{BDT}_0 \times \mathrm{BDT}_2$) grid has been chosen for the 0-jet channel and a 5×5 grid has been chosen for the 1-jet channel.

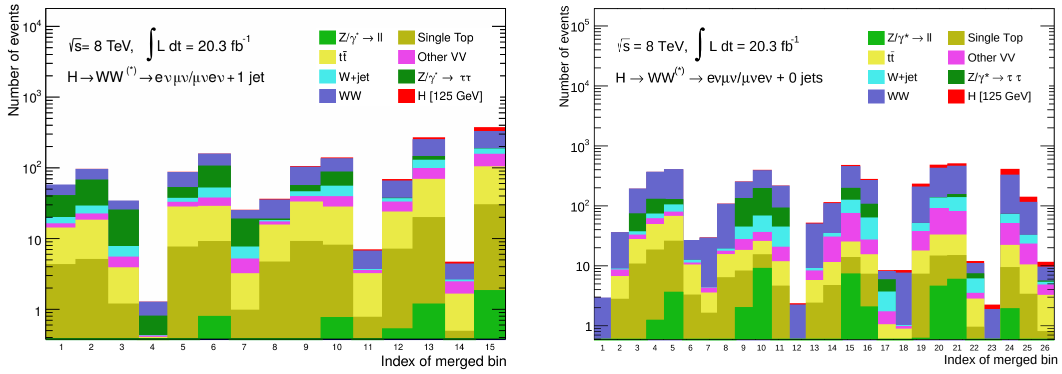
Given the different characteristics of the second BDT (labelled $\mathrm{BDT}_{\mathrm{CP}}$) in the CP analyses, it has proven adequate to use a 5×15 ($\mathrm{BDT}_0 \times \mathrm{BDT}_{\mathrm{CP}}$) grid for these cases.

The basic remapping from a two dimensional grid of bins to a one dimensional histogram is illustrated exemplarily in figure 9.1. The subsequent combination of bins starts from this unrolled histogram. The criterion for merging two adjacent bins is a threshold put on the weighted sum of all backgrounds in the specific bin. This background sum takes the respective cross sections and the luminosity into account. Given a sufficiently high merging threshold this condition guarantees a safeguard against unexpected fluctuations from systematic variations. If the resulting bin content is below one event in a specific bin, it gets merged with an adjacent bin. This requirement of at least one event per bin has been chosen in accordance with a recommendation from the HWW statistics forum for the specific use case.

Where possible, a nearly empty bin gets merged with a bin on its left side (i.e. in the same BDT_0 row); an example is given in figure 9.2a. Subsequent bins in the same row with too few entries get merged to the same bin. If the leftmost bin in a BDT_0 row is not filled sufficiently, the merging happens to the right side until a bin with sufficient event content can be formed or is reached. A row which in sum does not suffice the threshold condition would entirely get added to the rightmost bin in the row above. In case this happens in the top row, the merging is performed with the leftmost bin in the row below.

Two examples of one dimensional output distributions after the simple rebinning procedure are shown in figure 9.3.

Figure 9.4 shows the resulting signal distributions (0 and 1-jet) for the spin-2 hypothesis test with universal couplings, displaying the final differences in both distributions. Overlaid with these are the stacked backgrounds (analogous to those shown in 9.3a). The bins in these plots are reordered by the sum of backgrounds to better emphasize the signal characteristic without favouring one signal model over the other.



(a) Spin 2 test, minimal couplings model, 1-jet channel.

(b) CP-even BSM test, 0-jet channel.

Figure 9.3.: Unrolled BDT outputs after merging for two different test cases. All backgrounds are shown in the signal region with a Standard Model Higgs signal on top (red).

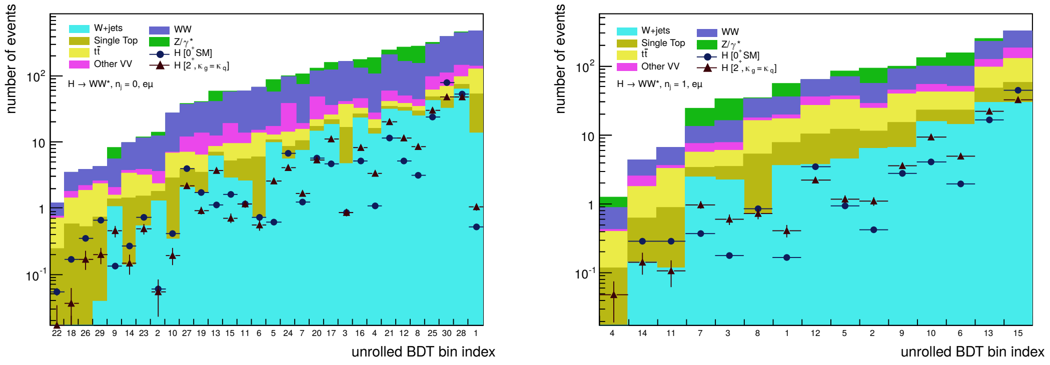


Figure 9.4.: Unrolled BDT output (0-jet channel on the left, 1-jet channel on the right) for the spin-2 hypothesis test with minimal couplings, illustrating the final shape differences between both signal models in the combined BDT output. The bins have been reordered such that the total background increases from left to right. The stacked backgrounds are overlaid with the two signal hypotheses displayed as dark red and dark blue points with statistical errors.

9.2. Smart Rebinning for Spin and CP Tests

As for the simple rebinning algorithm, the smart rebinning approach has been developed to eliminate fit instabilities arising from background fluctuations in sparsely populated regions of the $\mathrm{BDT}_0 \times \mathrm{BDT}_2$ output plane. In addition, the objective during the algorithm development has been the optimization of the output binning yielding a higher expected significance with given constraints such as a maximum bin number. To achieve the former, a requirement for the remapped BDT output variable bins is enforced, making sure that every bin holds a pre-defined minimal number of background entries just as in the simple binning case. The latter is achieved by merging bins based on their signal to background ratio in order to avoid spoiling signal-enriched grid regions. The optimized binning can thus be seen as a means to reduce the bin number while maintaining a level of separating power or as a mean to increase the sensitivity given a fixed bin number. As for the simple binning the bin re-ordering scheme (i.e. the mapping from the two dimensional BDT output plane to the one dimensional fit input variable) is created once for each hypothesis test and benchmark as well as the individual jet channels.

For the binning configuration production all samples are scaled to their post-fit normalisations which can be obtained e.g. using a fixed binning or an external plotting tool.

The algorithm requires 3 floating point parameters plus 2 integral values defining the initial (fine) binning of the 2D grid ($\mathrm{BDT}_0 \times \mathrm{BDT}_2$). This fixes the maximal granularity in the following process, i.e. no subdividing of these initial bins is allowed. The exact choice of parameters is motivated by their respective impact on the per-bin signal purity, the final number of bins and the avoidance of nearly empty bins. For this purpose a number of tests have been conducted varying a single parameter or a set of parameters at once and testing the resulting expected sensitivity from toy experiments taking background normalisations and their respective uncertainties into account. The aim has been to obtain the configuration yielding the highest sensitivity while maintaining a nearly identical number of output bins as used for the simple binning. Like this the possible improvements in significance can be studied. In principle smaller numbers of bins may be sufficient to obtain the same (optimal) sensitivity. However, since earlier tests have shown that an increased number of bins tends to increase the achievable sensitivity and in order to have a more precisely defined comparison framework, (nearly) identical numbers are demanded.

Since the fits for evaluating toy experiments including all nuisance parameters are very computationally expensive, an alternative optimization goal could have been to minimize the number of bins while keeping the expected sensitivity unchanged. The optimal configurations that have been found for the spin and the CP hypothesis test are listed in section 9.2.2.

9.2.1. Smart Binning Algorithm in Detail

The basic 2D to 1D mapping scheme that has been used in the analysis before is illustrated in figure 9.5. Here, the algorithm loops over the two-dimensional BDT output grid ($\mathrm{BDT}_0 \times \mathrm{BDT}_2$) and maps the grid bins to a 1D histogram by concatenating subsequent columns in the 2D plane. This plain unrolled BDT output is the starting point of the smart rebinning algorithm.

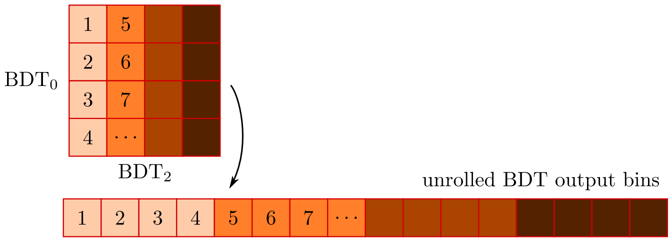


Figure 9.5.: Simple remapping from $BDT_2 \times BDT_0$ plane to 1D histogram/list.

The first subsequent step loops over the one dimensional distribution created in the initial step and merges neighbouring bins until every bin contains a minimal number (`threshold`) of events (i.e. including event weights). For this all backgrounds are summed up per bin and compared to the threshold as done for the simple rebinning scheme. The process is illustrated in figure 9.6.

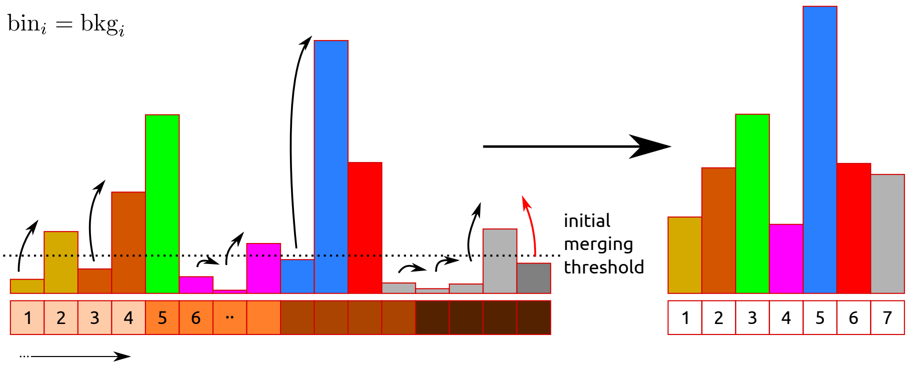


Figure 9.6.: Merging bins in unrolled BDT histogram applying first threshold. Whenever the background sum in one bin falls below the initial merging threshold, it gets merged to the subsequent bin. The resulting background distribution is shown in the right part of the image.

This first threshold is chosen as low as possible to avoid spoiling rather pure signal bins by adding nearby bins with a high fraction of background events. As for the simple binning, a minimum of one background event (including weights) per bin is chosen for each test. The implementation in this step is chosen ‘naively’ as explained to avoid introducing overtraining effects. Every bin of the output histogram is required to fulfil the threshold requirement. The merging starts with the lowest index of the unrolled BDT output (starting at $(-1.0, -1.0)$ for BDT_0 and BDT_2 output values).

When the very last bin of the unrolled BDT histogram does not reach the threshold it is added to the previous bin, even if the latter already fulfils the threshold requirement (also in case this last bin already is a ‘merged’ bin).

Smart Binning Example

In the following, as a small example a grid with 8×8 initial bins in the $\text{BDT}_0 \times \text{BDT}_2$ plane is chosen. For demonstration purposes, the initial threshold here is set to 10.0 instead of 1.0 background events per bin (including event weights). The two other parameters which are introduced in the following named `maxDiff` and `mergeCutoff` have been assigned arbitrary values for demonstration purposes of the algorithm. The resulting histogram after the first merging step (figures 9.5 and 9.6) is shown in figure 9.7.

For the second merging step of the algorithm an additional histogram is created in order to store a signal purity qualifier. The latter is chosen to be $\left(\overline{s}/b\right) := \frac{1}{2}\left(0^+ + 2^+\right) / \text{bkg}$, i.e. a signal average over the sum of backgrounds. The distribution obtained from the previous step is shown in plot 9.7. Corresponding bins in both plots (9.7, left and right) are marked with identical colours. This colouring scheme is applied identically for plots 9.8 and 9.9, such that the merging process and final grid regions for each bin can be read from the five plots. For the second merging step all bins are ordered by their signal average divided by the sum of backgrounds. Starting from the rightmost bin in the resulting histogram (figure 9.7, left plot) the algorithm loops over all bins and applies the following criteria for merging bins i and $i-1$:

- `maxDiff` parameter: $\Delta\overline{s}/b < 0.002$,
- `mergeCutoff` parameter: $\overline{s}/b < f \cdot \max(\overline{s}/b),\ \ f = 0.5$),

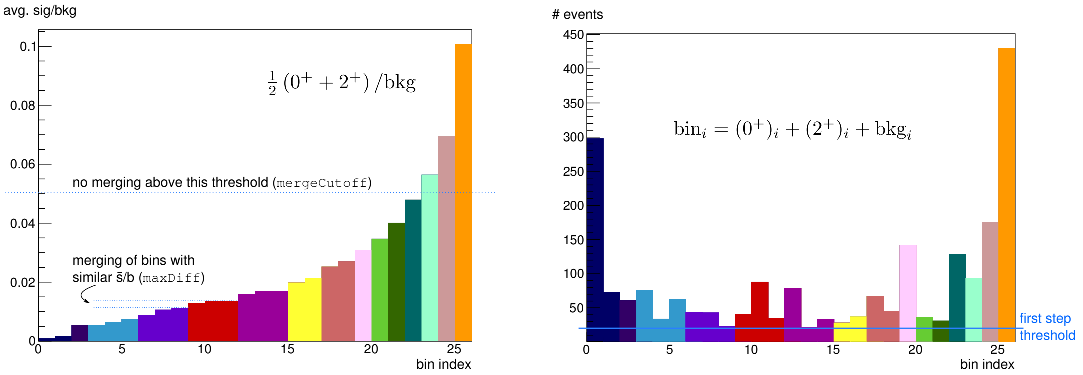


Figure 9.7.: Example binning (8 × 8 grid, 1-jet channel) after first merging step; identical colours indicate corresponding bins in both plots. The initial threshold (10.0 background events per bin) applied in the previous step is indicated in the right plot to give an idea of its effect. The two criteria applied in the current step are illustrated in the left panel.

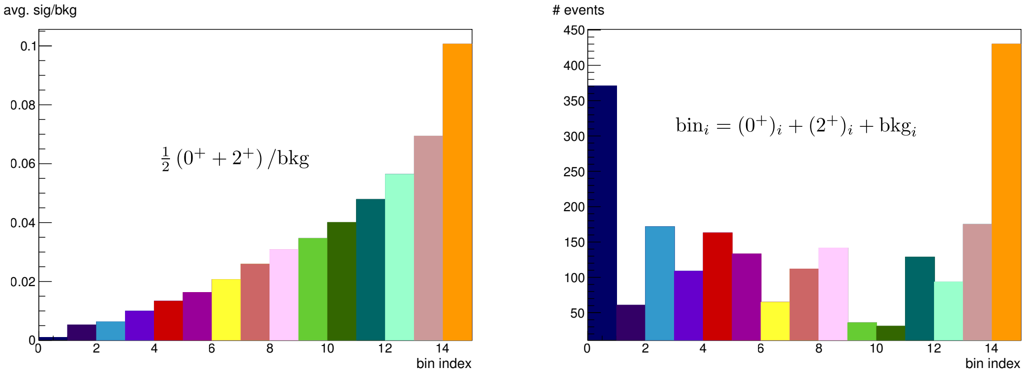


Figure 9.8.: Example binning (8 × 8 grid, 1-jet channel) after second merging step; identical colours indicate corresponding bins in both plots.

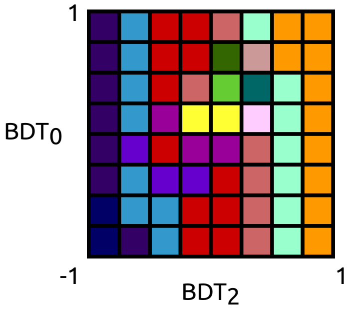


Figure 9.9.: Final binning for 8 × 8 grid. The colours correspond to figures 9.7 and 9.8.

When this merging step succeeds, the same (merged) bin is checked again in the next iteration. This means that several neighbouring bins can be merged into one output bin in subsequent steps. The resulting bin map obtained using the listed parameters is shown in figure 9.9. BDT grid regions marked with identical colours are thus mapped to the output bin indices with identically coloured histogram bars in figures 9.7 and 9.8.

9.2.2. Chosen Binning for Spin and CP Analysis

The actual sets of parameters used to generate the bin remapping schemes as used for the analysis are listed in table 9.1. For comparison the respective bin numbers used for the simple binning are listed in addition. When creating the smart binning configurations a comparable number of bins has been aimed for, to directly compare the achievable expected sensitivities. Since the initial merging threshold (`threshold` in the table) has been fixed to 1.0 as for the simple binning, the only parameters that have been varied during the tests are the input grid size, the maximum $\overline{s}/b$ difference (`maxDiff`) and the limit in $\overline{s}/b$ above which no merging is performed (`mergeCutoff`). For each benchmark case a number of tests varying these parameters has been conducted out of which the best configuration (i.e. the configuration yielding the highest expected sensitivity) has been chosen. The initial grid sizes (in $\mathrm{BDT}_0\times\mathrm{BDT}_2$) have been varied and found to be optimal as chosen with respect to the number of events in both jet channels. A lower number as well as a higher number of bins imposes stricter constraints on the other parameters to achieve the predefined bin number without adding to the resulting expected sensitivities. The two merging parameters `mergeCutoff` and `maxDiff` behave partly like opposite numbers, i.e. lower cut-off values allow for looser maximum differences. This again increases the chance of merging bins which mostly contain background events while on the other hand the lower merging cut-off increases the number of bins that remain untouched. Taking these dependencies and preconditions into account, the eligible parameter space is rather constrained. Like this the set of test cases can be assumed to be sufficient to find a near optimal configuration in each case.

As a cross check of the optimized binning, an overtraining check was performed for the $2^+, \kappa_g = 0.5, \kappa_q = 1, p_\mathrm{T}^\mathrm{H} < 300\,\mathrm{GeV}$ benchmark in both jet channels. This benchmark was chosen as a sort of worst case scenario since here the expected sensitivity relies in large parts on the specific signal distribution in the one jet channel. For this reason binning changes are expected to have a higher impact than for other benchmarks if the rather pure signal bins get merged or rearranged. Two smart binning configurations have been created using half of the events per sample each (scaled by a factor of 2.0). Comparing the resulting expected sensitivities from toys experiments (again with only background and signal normalisations as free parameters) no significant deviation has been found which is also in line with earlier tests that have been conducted for all benchmark models.

Table 9.1.: List of smart rebinning configurations for all benchmarks, listed per jet channel. The final number of bins for the unrolled BDT output histogram is listed in the second to last column. Where no number is given in the respective column, the `mergeCutoff` criterion is not applied. The rightmost column lists the bin numbers for the respective simple binning configurations for comparison.

jet channel	input grid	`threshold`	`maxDiff`	`merge-Cutoff`	merged bins	simple binning
Spin 2, $\kappa_g = 1,\ \kappa_q = 0,\ p_T^H < 125\,\text{GeV}$						
0	12×12	1.0	0.003	0.5	29	31
1	8×8	1.0	0.0022	-	22	19
Spin 2, $\kappa_g = 1,\ \kappa_q = 0,\ p_T^H < 300\,\text{GeV}$						
0	12×12	1.0	0.003	0.5	29	31
1	8×8	1.0	0.003	0.7	21	22
Spin 2, $\kappa_g = 0.5,\ \kappa_q = 1,\ p_T^H < 125\,\text{GeV}$						
0	12×12	1.0	0.002	0.5	30	32
1	8×8	1.0	0.0032	0.5	21	19
Spin 2, $\kappa_g = 0.5,\ \kappa_q = 1,\ p_T^H < 300\,\text{GeV}$						
0	12×12	1.0	0.002	0.5	30	30
1	8×8	1.0	0.0028	0.7	20	20
Spin 2, $\kappa_g = \kappa_q$						
0	12×12	1.0	0.002	-	27	30
1	8×8	1.0	0.0018	-	17	15
BSM CP-even						
0	12×12	1.0	0.001	0.7	39	40
BSM CP-odd						
0	12×12	1.0	0.0015	0.5	34	34

CHAPTER 10

Systematics

Two categories of systematic uncertainties are considered in the analysis: uncertainties arising from theory and experimental uncertainties. Both types are explained in more detail in the two subsequent sections.

The uncertainties may affect only certain backgrounds or the signal or all samples in the same way. Additionally some systematics affect only the normalisation of sample distributions, while others also or only distort the shape.

Both kinds of systematics (normalisation and shape variation) are treated independently in the modelling of the likelihood to be fitted. Where applicable the correlations between shape and normalisation uncertainties are taken into account.

10.1. Experimental Uncertainties

The most dominant experimental uncertainties arise from the jet reconstruction. These comprise the b-tagging efficiency and the jet energy scale and resolution. The jet energy scale (JES) and jet energy resolution (JER) model the uncertainty of the resolution (comparing MC and data) of the jet (p_{T}) reconstruction as well as the uncertainty arising from corrections of the reconstructed jet energy from calorimeter clusters with respect to the modelled depositions obtained from Monte Carlo simulations [118].

Additional prominent uncertainties are the lepton resolution, identification and trigger efficiencies plus the reconstruction of the missing transverse momentum. Uncertainties can be omitted if their impact is negligible. This is the case if for a shape uncertainty the variation is smaller than 5 % in each bin within the region and channel regarded. The limit has been chosen considering the statistical uncertainties and after cross checking for obvious shape distortions below this limit in the omitted variations.

Normalisation uncertainties are omitted below a 0.1 % variation. This is a conservative cut-off; in a test case ($J^P = 2^+$, minimal couplings benchmark) a systematic variation of this magnitude affecting all samples in the signal region yields a variation of roughly $1/3$ of the least significant quoted digit in the resulting $\mathrm{CL_s}$ value.

Table 10.1.: List of important uncertainties assigned with different aspects of the experiment and event reconstruction. The impacts and handling within the analysis is given for each uncertainty. Where not stated differently, the uncertainty only affects the normalisation. (taken from [6])

Source of uncertainty	Treatment in the analysis and magnitude
Jet Energy Scale (JES)	1 – 7% in total as a function of jet η and p_{T}
Jet Energy Resolution (JER)	5 – 20% as a function of jet η and p_{T} Relative uncertainty on the resolution is 2 – 40%
b-tagging	b-jet identification: 1 – 8% decomposed in p_{T} bins Light-quark jet misidentification: 9 – 19% as a function of η and p_{T} c-quark jet misidentification: 6 – 14% as a function of p_{T} *(also treated as shape variation)*
Leptons	Reconstruction, identification, isolation, trigger efficiency: below 1% except for electron identification: 0.2 – 2.7% depending on η and p_{T} *(also treated as shape variation)* Momentum scale and resolution: $< 1\%$
Missing Transverse Momentum	Propagated jet-energy and lepton-momentum scale uncertainties Resolution (1.5 – 3.3 GeV) and scale variation (0.3 – 1.4 GeV)
Pile-up	The amount of pile-up events is varied by 10%
Luminosity	2.8% [125]

10.2. Theory Uncertainties

The uncertainties linked to the extrapolation factors from control regions have already been listed in section 7.5. Yet there are further uncertainties arising from modelling the event generation process. Because the WW process constitutes the dominant background, its associated uncertainties play an extensive role for the overall analysis and are discussed in more detail.

A very important aspect is the modelling of the WW background shape in the signal region. Basically the same sources of uncertainties as for the extrapolation parameter α from the control region are considered. Also the same computation mechanism is applied. Thus the PDF (parton density function) uncertainty is obtained by combining the uncertainties of the PDF used to generate the WW background samples (CT10 PDF set) with the difference between NNPDF2.3 and CT10.

The scale uncertainty is estimated by varying the factorization and renormalisation upwards and downwards by a factor of two. To evaluate the generator modelling uncertainty, aMC@NLO+ HERWIG and POWHEG +HERWIG are compared (i.e. two generators for the hard process interfaced to the same showering algorithm are compared). Analogously, to estimate the parton shower uncertainty, samples showered with either HERWIG or PYTHIA are compared. A set of plots showing the resulting systematic variations for one exemplary benchmark (minimal couplings) is shown in figure 10.1.

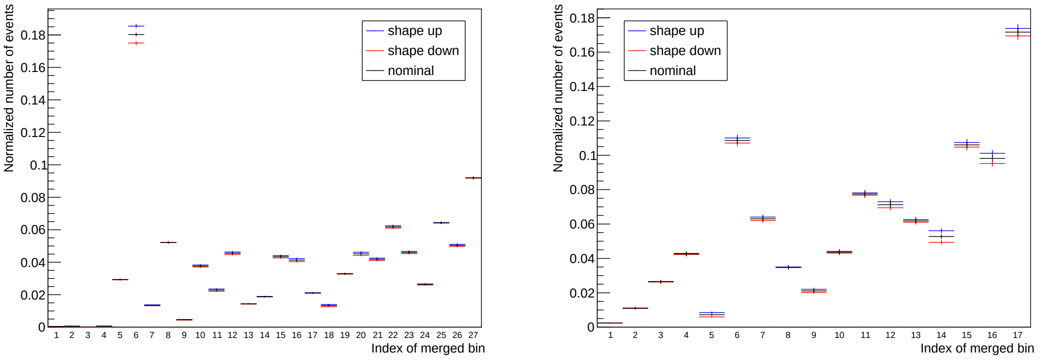
All these variations are evaluated for each benchmark model in both the spin and CP analysis and applied as bin-by-bin shape uncertainties if the impact is above the limits given above. This is only the case for the generator and parton shower uncertainties while the PDF and scale variations of the distributions' shapes are found to be negligible and only considered as an additional normalisation uncertainty.

An additional set of shape uncertainties are the variations of the fake rates for both muons and electrons from W+jets events, which are split up into different components: the correction factor from Z+jets (needed for the misidentified lepton treatment, see 7.5.5), a statistical uncertainty and a contribution from other backgrounds.

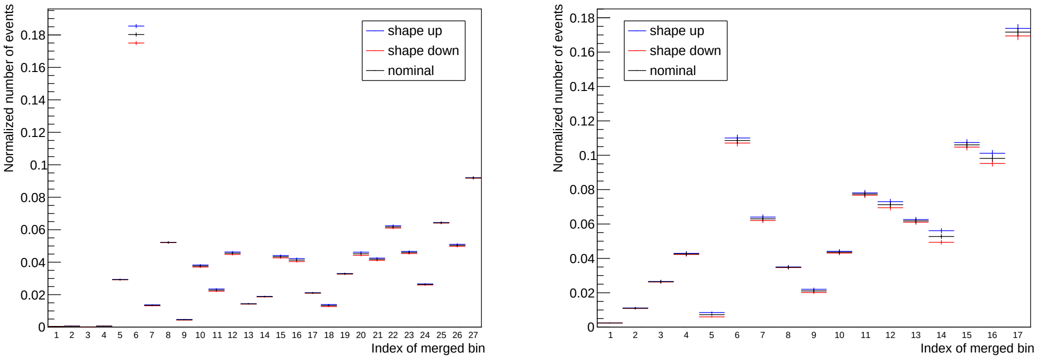
For the modelling of the Higgs p_{T} a common model of the p_{T} distribution expected for the Standard Model Higgs (see 6.2) is used for all (including BSM) spin-0 Higgs samples. This is done in order to avoid artificially introduced artefacts influencing the expected sensitivity of the parity analysis. Since at least for the Standard Model the p_{T} distribution obtained from HRES2.1 (at NNLL+NNLO) can be assumed as the best available description, no additional shape uncertainty is derived from the deviations between the original signal generator (POWHEG and MADGRAPH5_aMC@NLO) outputs.

For the spin-2 models no $p_{\mathrm{T}}^{\mathrm{H}}$ uncertainties are considered. In particular, the effects of the $p_{\mathrm{T}}^{\mathrm{H}}$ cut-off value for the non-universal couplings scenarios on the BDT output distributions clearly exceed such an uncertainty making it negligible for the current tests.

Interference effects of the $gg \to WW$ and the $gg \to H$ processes are as well omitted due to their negligible impact.



(a) WW UEPS uncertainty, 0-jet channel **(b)** WW UEPS uncertainty, 1-jet channel



(c) WW modelling uncertainty, 0-jet channel **(d)** WW modelling uncertainty, 1-jet channel

Figure 10.1.: The plots show systematic WW background shape variations from Monte Carlo studies for underlying event plus parton showering (UEPS) and modelling uncertainties. The minimal couplings benchmark ($\kappa_g = \kappa_q$) using the smart binning approach has been chosen as an example.

10.3. Impact of Systematics

The impact of the different nuisance parameters on the fitted $\hat{\varepsilon}$ (variable parametrizing the relative fraction of both tested signal hypotheses, i.e. Standard Model and alternative model) has been studied. The results obtained using the smart binning for three benchmarks (minimal couplings exemplary for the spin-2 studies plus CP-odd and BSM CP-even) are listed in table 10.2. Two plots showing the pulls[1] and impacts of systematics for the CP-odd benchmark are shown in figure 10.2. The corresponding pull plots for the other benchmarks look comparable with respect to pulls and fit constraints.

The impact on the observed $\mathrm{CL_s}$ values is lower (cf. [6]) than the impact on the fitted $\hat{\varepsilon}$ yet the ranking is expected to be similar in both cases. The uncertainties from the treatment of misidentified leptons (specifically with respect to the extrapolation from the Z+jetssample to the W+jetsbackground) play a dominant role beside the WW modelling and underlying event/parton shower uncertainties. Other systematic uncertainties appear more prominently in single cases such as the top scale factor uncertainty or the statistical uncertainty of the $Z/\gamma^* \to \tau\tau$ background sample. The QCD scale uncertainty for inclusive jet production in the 0-jet channel (all jet inclusive and ≥ 1 jet inclusive) affects two of the benchmarks as well to a significant extent.

Table 10.2.: Ranking of systematics based on their relative impact of the fitted spin hypothesis mixing parameter $\hat{\varepsilon}$ for the smart binning approach.

spin-2, $\kappa_g = \kappa_q$	up	down	BSM CP-even	up	down
Misid. rate (elec. stats)	18.2 %	-16.5 %	QCD scale ggH (incl.)	14.4 %	-10.2 %
Misid. rate (elec. flavour)	17.2 %	-16.0 %	Misid. rate (elec. stats)	12.4 %	-12.0 %
WW generator	12.6 %	-11.6 %	QCD scale ggH (1-jet incl.)	9.1 %	-11.7 %
Misid. rate (muon flavour)	11.3 %	-11.0 %	Misid. rate (elec. flavour)	9.2 %	-9.4 %
Misid. rate (muon stats)	10.9 %	-10.9 %	WW generator	8.6 %	-7.6 %
Misid. rate (elec. other)	10.3 %	-10.5 %	Misid. rate (muon flavour)	6.9 %	-8.2 %
QCD scale ggH (incl.)	9.7 %	-6.9 %	Higgs gg pdf uncert.	7.7 %	-6.9 %
QCD scale ggH (1-jet incl.)	3.8 %	-16.6 %	Misid. rate (elec. other)	6.4 %	-7.7 %
WW UE/PS extrapolation	4.5 %	-7.0 %	Misid. rate (muon stats)	5.9 %	-7.6 %
$Z/\gamma^* \to \tau\tau$ stats.	5.9 %	-5.5 %	WW UE/PS	4.7 %	-5.3 %

BSM CP-odd	up	down
WW generator	18.1 %	-14.7 %
Misid. rate (elec. stats)	9.8 %	-9.8 %
WW UE/PS	9.8 %	-9.0 %
Misid. rate (elec. flavour)	8.6 %	-8.6 %
Misid. rate (elec. other)	6.6 %	-6.7 %
top scale factor theo., 0-jet	4.1 %	-5.5 %
Misid. rate (muon flavour)	3.0 %	-3.0 %
Misid. rate (muon stats)	2.6 %	-3.0 %
top scale factor other bkg., 0-jet	2.2 %	-3.2 %
b-jet JES	2.9 %	-1.8 %

---

[1] The pull is defined as the fraction $\frac{\hat{\theta}-\theta_0}{\hat{\sigma}(\hat{\theta})}$ with $\hat{\theta}$ denoting the nuisance parameter estimator and θ_0 the expected value used for the generation. The denominator holds the estimated uncertainty $\hat{\sigma}(\hat{\theta})$. Noticeable variations may indicate modelling or fitting issues.

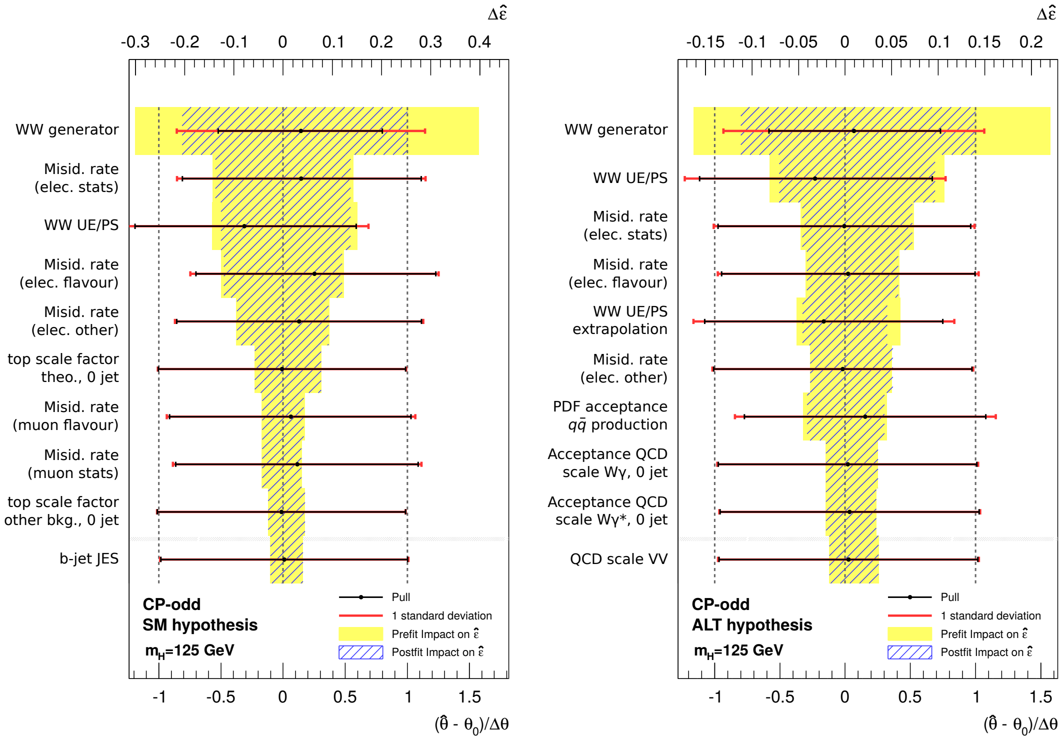


Figure 10.2.: Plots of pulls and systematics impact on $\hat{\varepsilon}$ for the CP-odd benchmark using the smart binning approach and assuming one of the two signal hypotheses each. In each case 10 systematics with the highest impact on the signal model fraction parameter $\hat{\varepsilon}$ are shown. The top axis indicates the absolute impact on the fitted $\hat{\varepsilon}$ while the bottom axis illustrates the relative scaling of the nuisance parameter ranges with respect to their estimate. While the black lines indicate the post-fit value and variation range of the nuisance parameters, the red line shows the 1 σ range around the fitted value. The dashed and coloured regions indicate the pre-fit and post-fit impact of each nuisance parameter on the fitted value of $\hat{\varepsilon}$.

CHAPTER 11

Results

This chapter presents the results obtained from the hypothesis tests of the presented benchmarks for spin-2 and spin-0 with both the simple and the smart binning approach. The first section discusses the post-fit yields and uncertainties while the exclusion limits are given in the next but one section.

11.1. Event Yields

The post-fit event yields for the seven benchmarks using the smart binning configurations are listed in table 11.1. The signal yield for the different flavour (e,μ) final state under the Standard Model assumption (lower table part) is 280 ± 110 which is compatible with the expectation based on the gluon-gluon fusion signal strength measurement of 238 events[102]. The difference between the two upper $p_{\mathrm{T}}^{\mathrm{H}}$ cuts of 125 and 300 GeV can be explained by the characteristic high p_{T} tail in the $p_{\mathrm{T}}^{\mathrm{H}}$ spectrum which alters the unrolled BDT output spectrum. In a similar yet less obvious way the impact of the specific model on the output BDT distributions is reflected in the variations of the signal and background yields between all benchmark models.

Two post-fit unrolled BDT output distributions displaying the stacked backgrounds plus signal and the respective data are shown in figures 11.1 and 11.2. The former figure depicts the spin-2 $\kappa_g = 0.5$, $\kappa_q = 1$, $p_{\mathrm{T}}^{\mathrm{H}} < 125\,\mathrm{GeV}$ benchmark (simple binning) and the latter depicts the spin-0 CP-odd benchmark (smart binning). In both cases (as for the other benchmarks not shown here) a good agreement between post-fit Monte Carlo predictions and data is achieved.

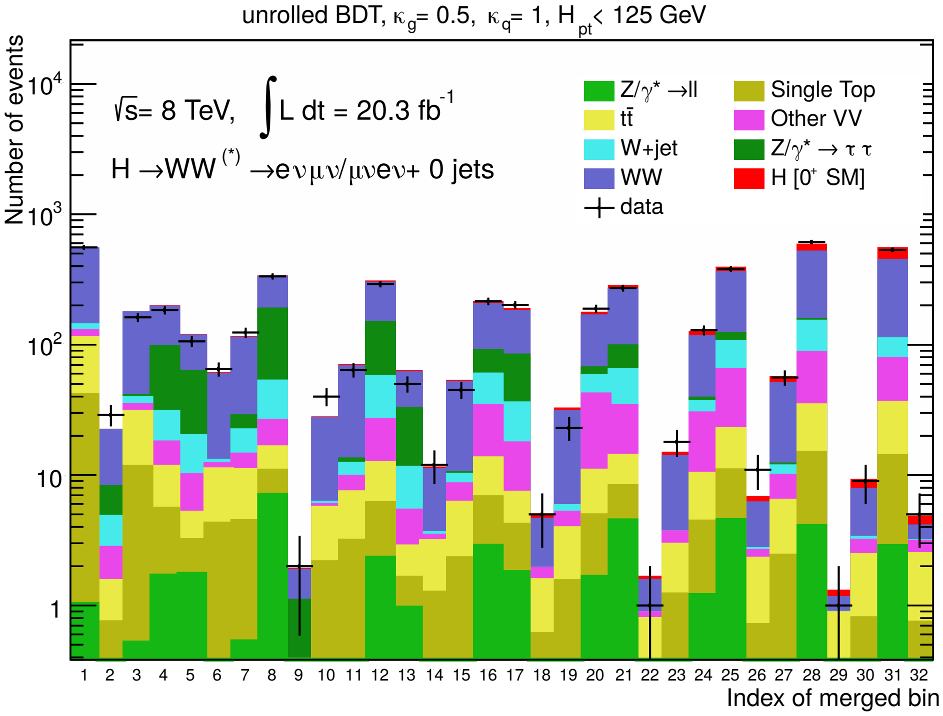


Figure 11.1.: Post-fit unrolled BDT output plot with bin indices of merged bins indicated. The model regarded is spin-2, $\kappa_g = 0.5$, $\kappa_q = 1$, $p_{\mathrm{T}}^{\mathrm{H}} < 125\,\mathrm{GeV}$, data points are shown with assigned errors for the simple binning.

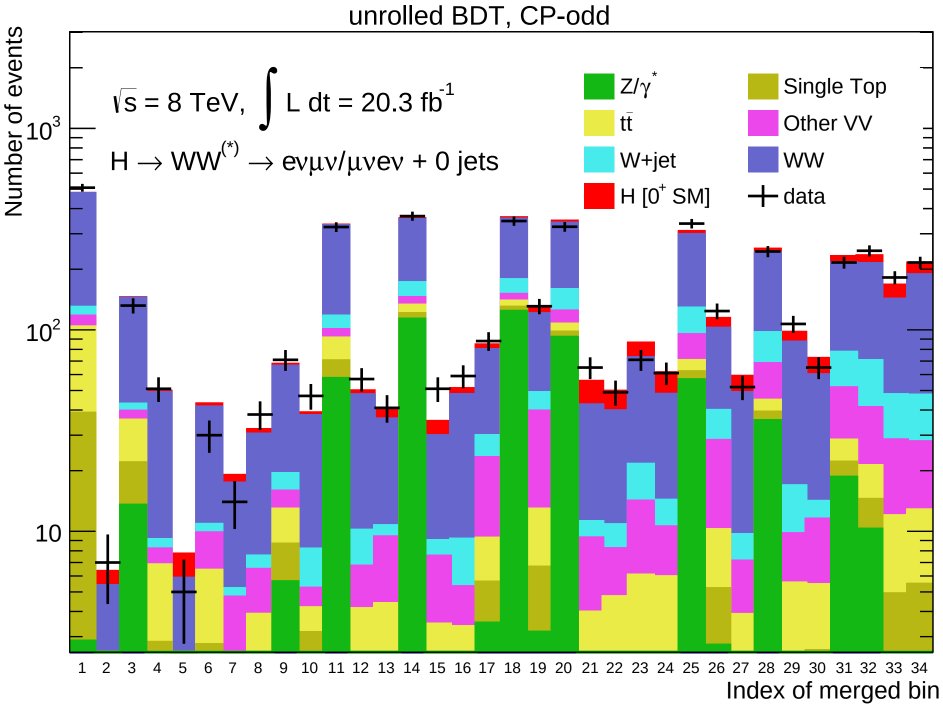


Figure 11.2.: Post-fit unrolled BDT output plot with bin indices of merged bins indicated. The model regarded is CP-odd (BSM), data points are shown with assigned errors for the smart binning.

Table 11.1.: The top part of the table lists the post-fit event yields and uncertainties for all benchmarks, under the assumption of the respective alternative model in the fit. In each case the total background and the signal yield is shown in each jet channel. The bottom part of the table lists results assuming a Standard Model Higgs signal for the fit, obtained for the minimal couplings ($\kappa_g = \kappa_q$) case. For this example the data yields and relative background composition is specified in addition. All results have been computed using the respective smart binning configuration. All event numbers are rounded to the nearest tens place.

Benchmark	Signal		Total background	
	0-jet	1-jet	0-jet	1-jet
Spin-2, $\kappa_g = 1,\ \kappa_q = 0,\ p_{\mathrm{T}}^{\mathrm{H}} < 125\,\mathrm{GeV}$	320 ± 130	110 ± 40	4450 ± 170	1390 ± 80
Spin-2, $\kappa_g = 1,\ \kappa_q = 0,\ p_{\mathrm{T}}^{\mathrm{H}} < 300\,\mathrm{GeV}$	230 ± 80	80 ± 30	4400 ± 170	1430 ± 90
Spin-2, $\kappa_g = 0.5,\ \kappa_q = 1,\ p_{\mathrm{T}}^{\mathrm{H}} < 125\,\mathrm{GeV}$	370 ± 150	130 ± 50	4430 ± 170	1360 ± 80
Spin-2, $\kappa_g = 0.5,\ \kappa_q = 1,\ p_{\mathrm{T}}^{\mathrm{H}} < 300\,\mathrm{GeV}$	320 ± 120	110 ± 40	4400 ± 170	1430 ± 90
Spin-2, $\kappa_g = \kappa_q$	360 ± 150	120 ± 50	4450 ± 180	1440 ± 90
BSM CP-odd	230 ± 170	-	4430 ± 200	-
BSM CP-even	110 ± 30	-	4350 ± 190	-

	Data	Signal	Tot. bkg.	WW	Top	DY	W+jets	Other
SM 0-jet	4730 ± 70	280 ± 110	4380 ± 190	64 %	8 %	13 %	8 %	7 %
SM 1-jet	1570 ± 40	100 ± 40	1440 ± 90	41 %	26 %	16 %	8 %	9 %

11.2. Hypothesis Test Results

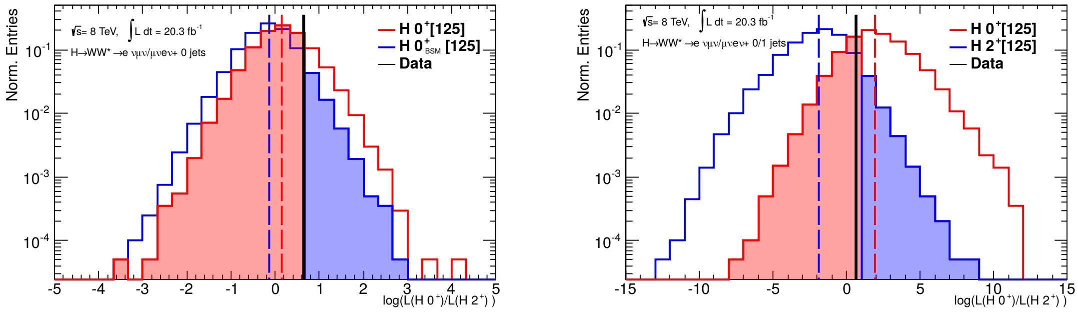
For the hypothesis test results, Monte Carlo toy experiments on the basis of the probability density function including all systematic uncertainties and normalisations have been run. The procedure for the generation and evaluation of toy experiments and the derivation of exclusion limits and $\mathrm{CL_s}$ values follows the description in sections 5.2.2 and 5.2.3.

Two exemplary toy distributions are shown in figure 11.3 for the CP-even and the minimal couplings spin-2 benchmark. In both cases as for the other benchmarks that are not shown here the resulting test statistic distributions of the profile likelihood q (cf. equation 5.11) are relatively symmetrical and no outliers outside the plotting range are observed. The coloured regions are only indicators of the integrated regions; in fact the p values are calculated on the unbinned data on the left and right of the data entry.

To illustrate the fit outcome figure 11.4 displays a post-fit comparison of data minus all backgrounds and the two signal models (Standard Model and spin-2 with minimal couplings for this example). Both hypotheses are shown for 0 and 1-jet. The compatibility between data (minus backgrounds) and the respective signal hypothesis is given as the KS test value.

Tables 11.2 and 11.3 list the resulting expected and observed p-values and $\mathrm{CL_s}$ exclusion limits of the respective alternative models. All exclusion limits are given for both binning configurations, i.e. the simple binning (table 11.2) and the smart binning (table 11.3).

A direct comparison of the expected sensitivities instead of p-values is given in table 11.4 in order to increase the clarity.



(a) Distribution of toys for spin-0, BSM CP-even hypothesis test (using smart binning), including all systematic uncertainties.

(b) Distribution of toys for spin-2, $\kappa_g = \kappa_q$ hypothesis test (using smart binning), including all systematic uncertainties.

Figure 11.3.: Exemplary distributions of toys of the profile likelihood estimator q (cf. equation 5.11).

All obtained CL_s values and their respective uncertainties including all systematic uncertainties are shown in figures 11.6 (simple binning) and 11.7, again for both binning configurations.

All results are perfectly compatible with the Standard Model expectation of a CP-even Higgs boson with spin 0. At the same time, all alternative spin-2 model benchmarks are excluded at 82.5 % (0.94 σ) to 99.1 % confidence level (2.37 σ) (based on the smart binning results). The exclusion limits of the tested CP models amount to 64.5 % C.L. (0.37 σ) in the CP-even case with higher dimensional operators and 97.1 % C.L. (1.90 σ) in the CP-odd case. In summary, for the vast majority of tests a clear exclusion of the alternative model has been observed while the agreement with the Standard Model expectation can be confirmed for every tested benchmark.

With respect to the results broken down by the single jet channels in the result tables, it is important to note that the way the single channels contribute to the combined result is not exactly straight-forward. There are even cases where the combined exclusion limit lies below the limit obtained in a single channel. This effect can occur in few cases for the observed limits—nevertheless the uncertainty on the exclusion limit for the combined results (as depicted in figures 11.6 and 11.7) is smaller. By comparing the *expected sensitivities* in the single jet channels to the combined ones (cf. table 11.4) the correct modelling can be verified for both binning configurations.

When assessing the quality of the smart binning method, the expected sensitivity is the yardstick to consider. The observed sensitivity on the other hand is only obtained *after* fixing the binning and is supposed to follow a distribution predetermined by the distributions of toys. These reflect the expected probability density function which should emerge if multiple datasets were recorded and tested, i.e. if a large number of real world experiments was conducted.

Comparing the expected sensitivities obtained with both binning approaches, the smart binning clearly outperforms the simple binning in terms of expected sensitivities in nearly all tested benchmarks and jet channels.

A worse performance with respect to the exclusion sensitivities of both models is only observed for the CP-odd case. Since only square initial bin configurations have been tested for the smart binning while the number of bins for the simple binning has been chosen very asymmetric (5 × 15 bins in the $\text{BDT}_0{\times}\text{BDT}_{CP}$ plane for spin-0 tests) further studies using asymmetric initial smart binning grids may be of interest.

As a cross check of the most divergent observed result in the BSM CP-even benchmark, additional tests have been performed. For this, sets of toy datasets have been created from Monte Carlo events following a Poisson distribution with the post-fit background normalisation used as the distribution's mean. These (unbinned) datasets have been binned according to the simple and smart

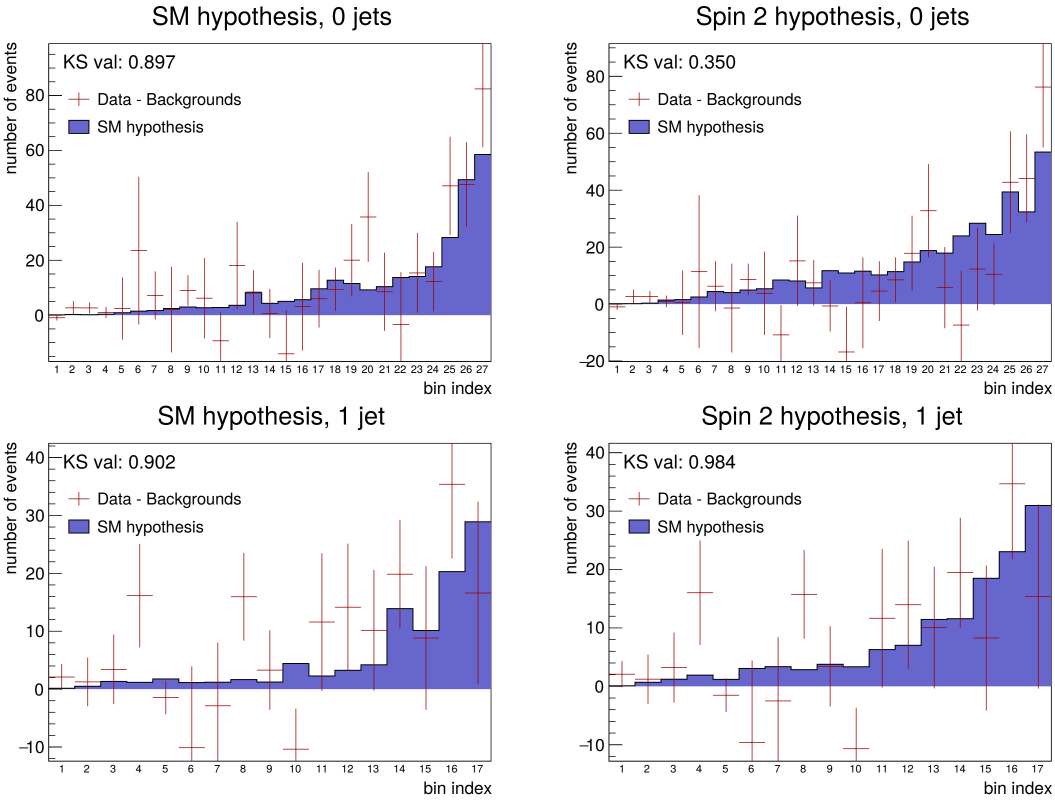


Figure 11.4.: Post-fit distribution of data minus all backgrounds compared to the two signal hypothesis distributions in the spin-2 minimal couplings ($\kappa_g = \kappa_q$) case using the smart binning configuration. Both hypotheses are shown in the 0 and 1-jet channel each. The compatibility between each hypothesis and the data minus background is indicated in the plots as the Kolmogorov-Smirnov test value[95].

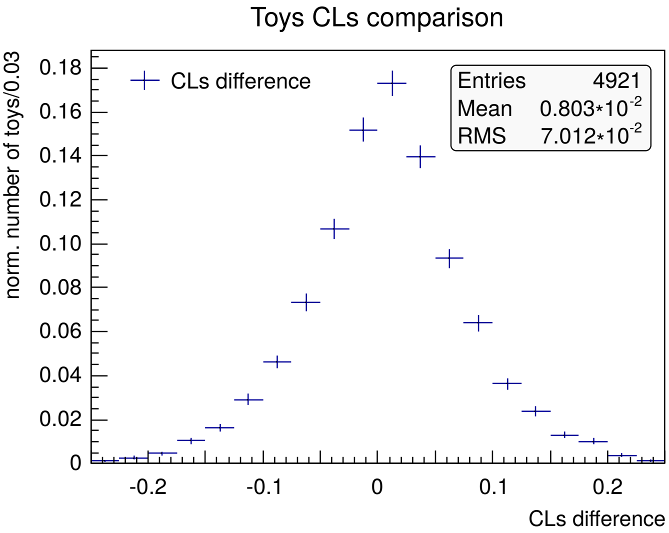


Figure 11.5.: Comparison plot for BSM CP-even hypothesis test cross check. Plotted are the CLs differences obtained from simulated data sets using the simple binning versus the smart binning.

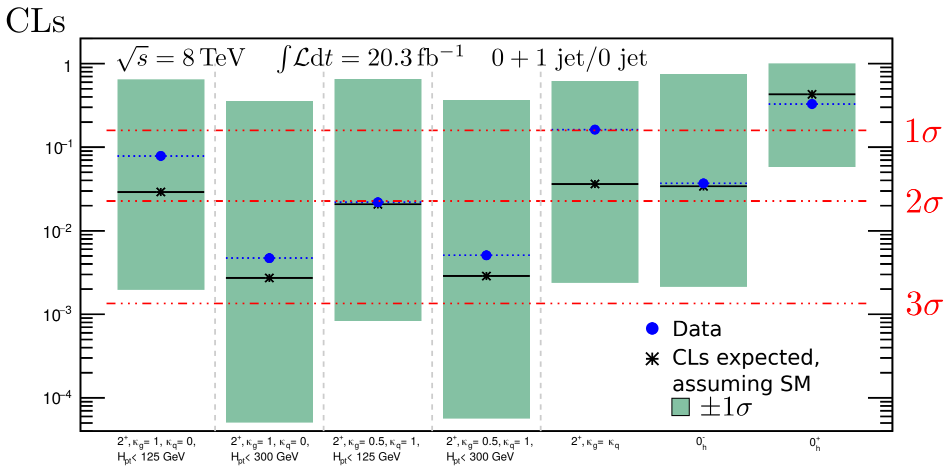


Figure 11.6.: Exclusion limits for the five spin-2 benchmark models and the spin-$0^\pm$ model taking the respective sensitivity uncertainties into account. The results have been produced using the simple binning.

binning configurations and have been processed in the subsequent analysis chain where they have been treated as measured data. The resulting $\mathrm{CL_s}$ differences comparing identical events in both binning scenarios are plotted in figure 11.5. The difference in the observed $\mathrm{CL_s}$ values between the simple binning ($1 - \mathrm{CL_s} = 67.1\,\%$) and the smart binning ($1 - \mathrm{CL_s} = 64.5\,\%$) amounts to 0.026 which is still very close to the peak of the distribution (approximating the distribution with a Gaussian it would be way below the $1\,\sigma$ level). It is therefore obvious that also for this case both binning configurations provide absolutely compatible results.

Table 11.2.: Table of expected and observed sensitivities (p values and observed 1-CL$_\mathrm{s}$ values, see section 5.2.2) for the various benchmarks of spin-2$^+$ and spin-0$_h^\pm$ hypotheses. The results have been obtained taking all systematic variations into account with the respective simple binning configuration. For the spin-2 tests, both single jet channel results as well as the combined results (fitted with a common signal strength μ) are given

channel	$p^{\mathrm{SM}}_{\mathrm{exp},\,\mu=\hat{\mu}}$	$p^{\mathrm{ALT}}_{\mathrm{exp},\,\mu=\hat{\mu}}$	$p^{\mathrm{SM}}_{\mathrm{obs}}$	$p^{\mathrm{ALT}}_{\mathrm{obs}}$	1-CL$_\mathrm{s}$
Spin-2, $\kappa_g = 1,\ \kappa_q = 0,\ p_\mathrm{T}^\mathrm{H} < 125\,\mathrm{GeV}$					
0+1-jet	0.040	0.028	0.410	0.046	92.1 %
0-jet	0.065	0.074	0.153	0.305	64.0 %
1-jet	0.157	0.092	0.740	0.025	90.5 %
Spin-2, $\kappa_g = 1,\ \kappa_q = 0,\ p_\mathrm{T}^\mathrm{H} < 300\,\mathrm{GeV}$					
0+1-jet	0.009	0.003	0.543	0.002	99.5 %
0-jet	0.065	0.071	0.153	0.301	64.4 %
1-jet	0.063	0.012	0.681	0.003	99.1 %
Spin-2, $\kappa_g = 0.5,\ \kappa_q = 1,\ p_\mathrm{T}^\mathrm{H} < 125\,\mathrm{GeV}$					
0+1-jet	0.039	0.020	0.630	0.008	97.8 %
0-jet	0.064	0.051	0.488	0.054	89.4 %
1-jet	0.149	0.087	0.749	0.021	91.5 %
Spin-2, $\kappa_g = 0.5,\ \kappa_q = 1,\ p_\mathrm{T}^\mathrm{H} < 300\,\mathrm{GeV}$					
0+1-jet	0.010	0.003	0.520	0.002	99.5 %
0-jet	0.075	0.057	0.510	0.054	89.0 %
1-jet	0.049	0.013	0.581	0.007	98.3 %
Spin-2, $\kappa_g = \kappa_q$					
0+1-jet	0.040	0.035	0.249	0.122	83.8 %
0-jet	0.051	0.047	0.243	0.156	79.4 %
1-jet	0.153	0.140	0.519	0.129	73.3 %
BSM CP-odd					
0+1-jet	0.058	0.032	0.648	0.013	96.3 %
BSM CP-even					
0+1-jet	0.299	0.283	0.884	0.038	67.1 %

Table 11.3.: Table of expected and observed sensitivities (p values and observed 1-$\mathrm{CL_s}$ values, see section 5.2.2) for the various benchmarks of spin-2^+ and spin-$0_h^\pm$ hypotheses. The results have been obtained taking all systematic variations into account with the respective smart binning configuration. For the spin-2 tests, both single jet channel results as well as the combined results (fitted with a common signal strength μ) are given.

channel	$p^{\mathrm{SM}}_{\mathrm{exp},\ \mu=\hat{\mu}}$	$p^{\mathrm{ALT}}_{\mathrm{exp},\ \mu=\hat{\mu}}$	$p^{\mathrm{SM}}_{\mathrm{obs}}$	$p^{\mathrm{ALT}}_{\mathrm{obs}}$	1-$\mathrm{CL_s}$
Spin-2, $\kappa_g = 1,\ \kappa_q = 0,\ p_{\mathrm{T}}^{\mathrm{H}} < 125\,\mathrm{GeV}$					
0+1-jet	0.027	0.021	0.363	0.045	92.9 %
0-jet	0.046	0.037	0.383	0.066	89.2 %
1-jet	0.183	0.190	0.231	0.433	43.7 %
Spin-2, $\kappa_g = 1,\ \kappa_q = 0,\ p_{\mathrm{T}}^{\mathrm{H}} < 300\,\mathrm{GeV}$					
0+1-jet	0.001	0.001	0.220	0.007	99.1 %
0-jet	0.048	0.037	0.379	0.066	89.3 %
1-jet	0.050	0.049	0.275	0.147	79.7 %
Spin-2, $\kappa_g = 0.5,\ \kappa_q = 1,\ p_{\mathrm{T}}^{\mathrm{H}} < 125\,\mathrm{GeV}$					
0+1-jet	0.022	0.022	0.161	0.147	82.5 %
0-jet	0.041	0.035	0.308	0.090	87.0 %
1-jet	0.108	0.131	0.123	0.466	46.8 %
Spin-2, $\kappa_g = 0.5,\ \kappa_q = 1,\ p_{\mathrm{T}}^{\mathrm{H}} < 300\,\mathrm{GeV}$					
0+1-jet	0.001	0.001	0.048	0.054	94.3 %
0-jet	0.044	0.038	0.303	0.096	86.2 %
1-jet	0.008	0.022	0.035	0.294	69.6 %
Spin-2, $\kappa_g = \kappa_q$					
0+1-jet	0.024	0.020	0.249	0.081	89.2 %
0-jet	0.044	0.030	0.446	0.040	92.8 %
1-jet	0.133	0.161	0.071	0.640	31.1 %
BSM CP-odd 0_h^-					
0+1-jet	0.071	0.036	0.724	0.008	97.1 %
BSM CP-even 0_h^+					
0+1-jet	0.284	0.273	0.796	0.072	64.5 %

Table 11.4.: Table of expected sensitivities (significances Σ, see section 5.2.2) for the various benchmarks of spin 2^+ and spin $0_h^\pm$ hypotheses, comparing the simple and smart binning expectation. The results have been obtained taking all systematic variations into account with the respective binning configurations. For the spin-2 tests, both single jet channel results as well as the combined results (fitted with a common signal strength μ) are given.

	simple binning		smart binning	
channel	Σ^0_{exp}	Σ^2_{exp}	Σ^0_{exp}	Σ^2_{exp}
Spin-2, $\kappa_g = 1,\ \kappa_q = 0,\ p_{\text{T}}^{\text{H}} < 125\,\text{GeV}$				
0+1-jet	1.75	1.91	1.92	2.03
0-jet	1.52	1.45	1.68	1.79
1-jet	1.01	1.33	0.9	0.88
Spin-2, $\kappa_g = 1,\ \kappa_q = 0,\ p_{\text{T}}^{\text{H}} < 300\,\text{GeV}$				
0+1-jet	2.37	2.78	3.11	3.14
0-jet	1.51	1.47	1.67	1.79
1-jet	1.53	2.26	1.65	1.66
Spin-2, $\kappa_g = 0.5,\ \kappa_q = 1,\ p_{\text{T}}^{\text{H}} < 125\,\text{GeV}$				
0+1-jet	1.76	2.06	2.01	2.02
0-jet	1.53	1.63	1.74	1.81
1-jet	1.04	1.36	1.24	1.12
Spin-2, $\kappa_g = 0.5,\ \kappa_q = 1,\ p_{\text{T}}^{\text{H}} < 300\,\text{GeV}$				
0+1-jet	2.31	2.77	3.14	3.12
0-jet	1.44	1.58	1.7	1.77
1-jet	1.66	2.23	2.4	2.01
Spin-2, $\kappa_g = \kappa_q$				
0+1-jet	1.76	1.81	1.98	2.05
0-jet	1.63	1.67	1.71	1.88
1-jet	1.02	1.08	1.11	0.99
BSM CP-odd				
0+1-jet	1.57	1.85	1.47	1.8
BSM CP-even				
0+1-jet	0.53	0.57	0.57	0.6

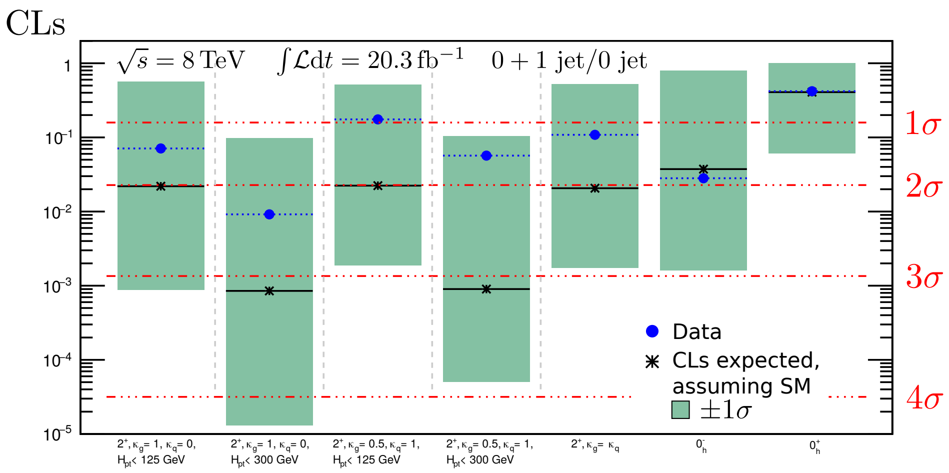


Figure 11.7.: Exclusion limits for the five spin-2 benchmark models and the spin-$0^{\pm}$ model taking the respective sensitivity uncertainties into account. The results have been produced using the smart rebinning approach.

CHAPTER 12

Summary and Outlook

12.1. Results from Other Channels

12.1.1. Bosonic Channels

The hypothesis tests described in this thesis have analogously been conducted in two other Higgs decay channels by ATLAS and CMS, namely in $H \to ZZ^* \to 4\ell$ and in parts in $H \to \gamma\gamma$. Like for the $H \to WW^*$ process, both channels have their respective advantages and disadvantages.

The $H \to ZZ^* \to 4\ell$ channel features four charged leptons in the final state each of which can be reconstructed. Like this, all angular variables of the Higgs decay are fully accessible, as well as the mass of the reconstructed particle. The downside of the $H \to ZZ^*$ channel is the significantly lower cross section compared to $H \to WW^*$ which reduces the number of events in the recorded dataset (cf. the branching ratio plot in the phenomenology section 4.6, which displays the processes, though regardless of the specific final state). The number of events with four leptons in the final state is way lower still because the branching fraction of $Z \to \ell^+\ell^-$ (where ℓ stands for all lepton flavours) amounts to only around $1/10$ while hadronic final states make up the largest fraction (cf. [4]).

For $H \to \gamma\gamma$ the available number of events is sufficient but the backgrounds are very prominent. In addition, since the process is a two body decay only few sensitive variables are available for the discrimination between different spin and CP states. In particular, the quantity of angular information reflected in the final state variables is very limited which makes some analysis questions inaccessible with the dataset recorded yet.

The analyses and outcome for both ATLAS and CMS in the other bosonic channels as well as the results of the $H \to WW^*$ analysis of the CMS experiment are presented briefly in the following.

12.1.2. Higgs to Diphoton Channel Analysis

For the $H \to \gamma\gamma$ channel, both experiments substantially rely on one variable specifically sensitive to the spin of the Higgs boson. The variable in question is the modulus of the polar angle between the photons and the z-axis in the Collins-Soper frame [126], which is defined as

$$|\cos\theta^*_{\mathrm{CS}}| = \frac{|\sinh(\eta_{\gamma_1} - \eta_{\gamma_2})|}{\sqrt{1 + \left(p_{\mathrm{T}}^{\gamma\gamma}/m_{\gamma\gamma}\right)^2}} \cdot \frac{2p_{\mathrm{T}}^{\gamma_1} p_{\mathrm{T}}^{\gamma_2}}{m_{\gamma\gamma}^2}. \tag{12.1}$$

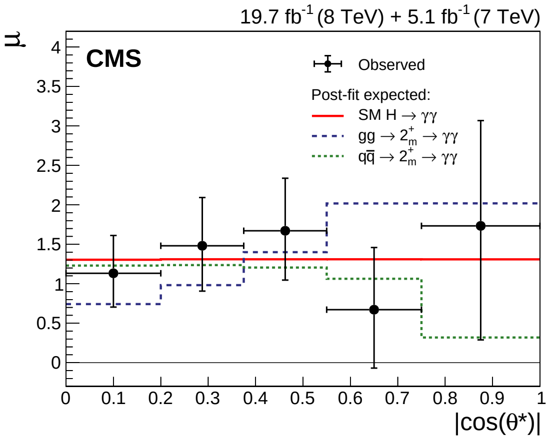


Figure 12.1.: CMS analysis plot showing the expected signal strength (μ) distributions in bins of $|\cos\theta^*_{\mathrm{CS}}|$ for a Standard Model Higgs as well as two alternate spin-2 hypotheses. The measured data is shown as black points. (from [127])

In this expression $p_{\mathrm{T}}^{\gamma\gamma}$ and $m_{\gamma\gamma}$ denote the transverse momentum and invariant mass of the diphoton system, while η_{γ_i} and $p_{\mathrm{T}}^{\gamma_i}$ denote the pseudorapidity and transverse momentum of the two single photons.

For the ATLAS analysis [5] the diphoton p_{T} is used as a second variable with discrimination power. A fit of the hypotheses' distribution shapes in eleven categories yields the result for the five tested spin-2 benchmarks. Ten out of eleven categories spread out equally over the $|\cos\theta^*_{\mathrm{CS}}|$ range with $p_{\mathrm{T}}^{\gamma\gamma} < 125\,\mathrm{GeV}$ and the remaining additional category holds all events with $p_{\mathrm{T}}^{\gamma\gamma} \geq 125\,\mathrm{GeV}$. For non-universal couplings the upper cut-off values of 125 and 300 GeV are applied as in the $H \to WW^*$ channel analysis.

The CMS analysis [127] follows a slightly different approach by defining four event classes based on the photon η value and a shower shape qualifier. Each of these classes is subdivided into five $|\cos\theta^*_{\mathrm{CS}}|$ bins in which the signal strength is fitted. This procedure is applied independently for the 7 and 8 TeV datasets. A likelihood ratio fit evaluated for five different production mode admixtures of the hypothetical spin-2 Higgs (0, 25, 50, 75, 100 % fraction of the $q\bar{q}$ production mode) yields the rejection of either model.

For CMS the resulting signal strength distribution in bins of $|\cos(\theta^*)|$ comparing the data to the Standard Model as well as to two alternative hypotheses is shown in figure 12.1.

12.1.3. $H \to WW^* \to \ell\nu\ell\nu$ Decay Channel (CMS)

The analysis approach in the $H \to WW^* \to e\nu\mu\nu$ decay channel chosen by CMS [86, 128] differs from the one pursued by ATLAS presented in this thesis. The hypothesis test is based on a two dimensional fit of $m_{\mathrm{T}} \times m_{\ell\ell}$. For this purpose a slightly different definition of m_{T} is used, neglecting the invariant mass of the dilepton pair. Also the CMS analysis makes use of the 0 and 1-jet categories.

Plots displaying both quantities for one spin-2 model and the Standard Model in the 0-jet channel are shown in figure 12.2. Since here m_{T} does not include $m_{\ell\ell}$ both quantities are independent and have been shown to provide high discrimination power between signals and backgrounds on the one hand and between the different signal models on the other hand.

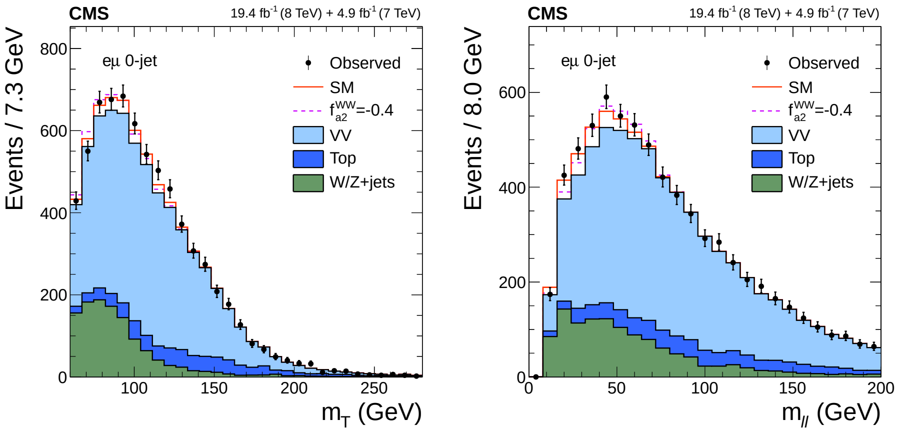


Figure 12.2.: CMS plots of m_{T} (neglecting $m_{\ell\ell}$) and $m_{\ell\ell}$ showing data compared to one alternative model (dotted violet line) and the Standard Model (red line) on top of backgrounds in the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ decay mode for the 0-jet channel. For this the Standard Model cross section and a Higgs boson mass of 125.6 GeV is assumed. (taken from [128])

12.1.4. $H \rightarrow ZZ^* \rightarrow 4\ell$ Decay Channel

The $H \rightarrow ZZ^* \rightarrow 4\ell$ channel is remarkable for a fully reconstructable final state combined with a similar decay signature as the $H \rightarrow WW^*$ process such that a high number of sensitive variables are in principle accessible. Nevertheless, the decay rate to ZZ^* is significantly lower (see figure 4.6) such that the number of events available for the analysis with the recorded dataset of the first LHC run is by far lower than for the $H \rightarrow WW^*$ channel. Information about spin and parity of the resonance is carried by all variables highlighted in figure 6.4 and by the reconstructed $m_{4\ell}$. This comprises angular variables associated with the production and decay as well as the mass spectra of the intermediate Z bosons since not both of them can be on-shell. The resulting distributions for two parity hypotheses are shown in figure 12.3a.

Main backgrounds of the four lepton channel are the non-resonant ZZ production, Z+jets and $t\bar{t}$. Both, the CMS [128, 129] and ATLAS analysis [5] perform a template based maximum likelihood fit of multidimensional quantities derived from the observables mentioned above.

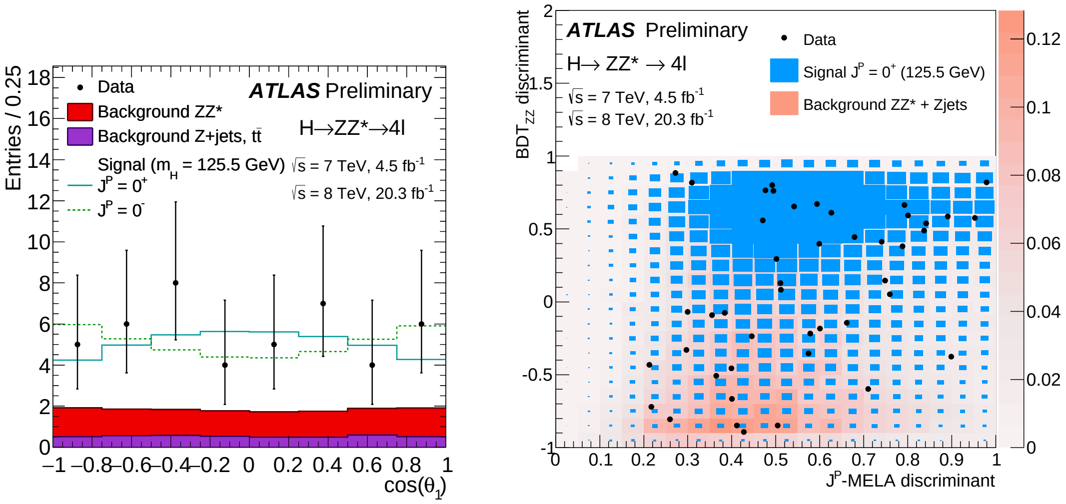
In the ATLAS case shown in figure 12.3b, a BDT (denoted as BDT_{ZZ}) is trained to tell signal and backgrounds apart. The signal hypothesis separation is provided by a specific matrix element likelihood qualifier (J^{P}-MELA variable).

The CMS analysis tests a wider set of alternative spin hypotheses and uses two to three dimensional templates built from matrix element based variables. In each case one of the quantities is designed to suppress backgrounds while the other one or two serve the purpose of identifying the signal in question.

12.1.5. Combined Diboson Results

Both experiments, ATLAS and CMS have combined the results of the three channels to obtain combined exclusion limits. For CMS, figure 12.4 shows the limits for various hypotheses. For ATLAS, figure 12.5 shows the latest combined results. While for CMS the spin-1 hypothesis with odd or even parity is included in the plot, the ATLAS results for these hypotheses are omitted and can be found in [3].

In summary the conclusion from both experiments clearly rules out the alternate models. For CMS all non-Standard-Model hypotheses with spin-1 or 2 in the fixed hypothesis tests are excluded at more than 99.9 % confidence level. The ATLAS results can exclude all tested hypotheses—



(a) Plot of $\cos(\theta_1)$ (cf. figure 6.4) as one exemplary sensitive variable for the spin/-parity analysis. The distribution is shown for signal region events with $115\,\mathrm{GeV} < m_{4\ell} < 130\,\mathrm{GeV}$. The Standard Model and the purely CP-odd (0^+) state expectatios are shown on top of the backgrounds.

(b) Two dimensional distribution of the BDT_{ZZ} trained to discriminate non-resonant ZZ background contributions versus the matrix element based qualifier ($\mathrm{J^P}$-MELA) discriminating between a SM and 0^- hypothesis. The data points are overlaid as black dots.

Figure 12.3.: Two plots from the ATLAS $H \to ZZ^* \to 4\ell$ channel analysis, both from [5].

non-SM spin-0 models, spin-1 (from previous publication) and spin-2 models—at more than 99 % confidence level.

The additional parameter scans of a mixture between the Standard Model Higgs and a BSM spin $0^{\pm}$ yield no significant deviations from the Standard Model expectation for both experiments (see figure 12.6).

12.1.6. Fermionic Channel

The CDF and D0 experiments at the Tevatron have run two hypothesis tests and a $0^+/0^-$ mixing study in the $H \to b\bar{b}$ decay channel with their combined datasets[130], contributing a properties study relying on fermions instead of bosons as the Higgs decay products. Under the assumption that the object of inquiry - the Standard Model Higgs boson - is identical for the LHC and Tevatron tests, and starting from the idea that the cross section times branching ratios are identical to the Standard Model expectation, the tested $J^P = 2^+$ and $J^P = 0^-$ hypotheses can each be excluded with 4.9 and 5.0 standard deviations from $H \to b\bar{b}$.

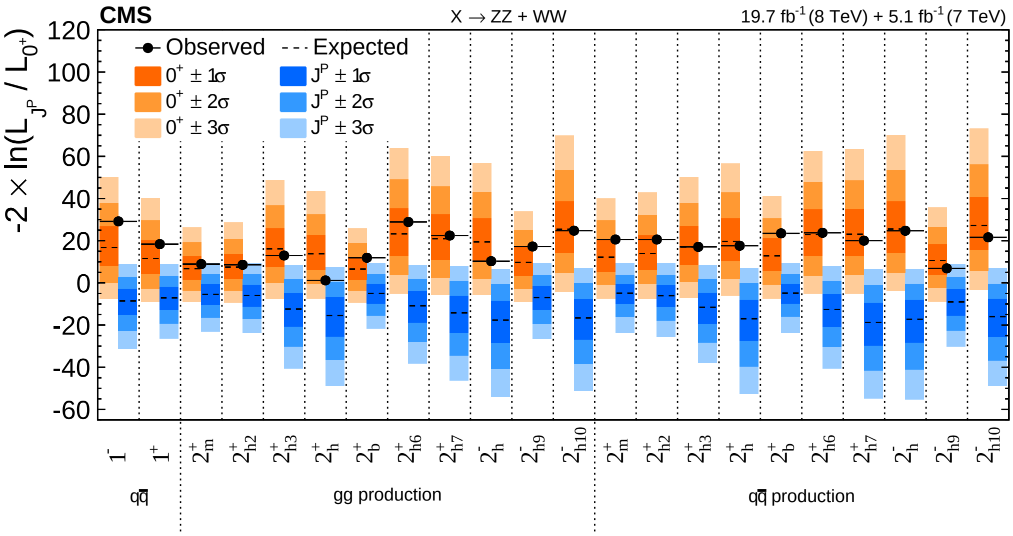


Figure 12.4.: Summary plot showing the combination of the $H \to ZZ^*$ and $H \to WW^*$ channel results from CMS. The different models (including higher order operators for spin-2) are listed on the horizontal axis. For each hypothesis the test statistic distributions are indicated where the Standard Model Higgs is shown in shades of orange and the alternative models in shades of blue. The coloured regions correspond to a 1–3 σ deviation from the median expected value (dashed lines). The corresponding results from data are shown as black dots. (from [128])

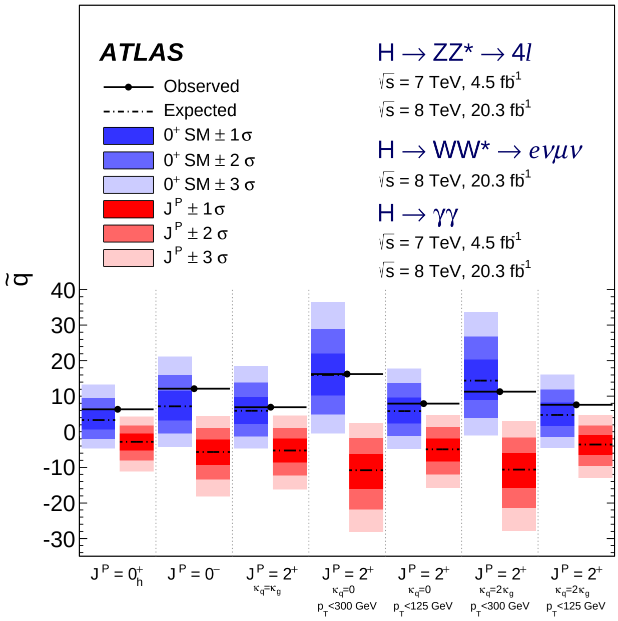


Figure 12.5.: Summary plot showing the combined ATLAS results from the $H \to WW^* \to e\nu\mu\nu$, $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ channels. For each hypothesis test listed on the horizontal axis, the observed and expected values of the test statistic $\tilde{q} = \log\left(\mathcal{L}(J^P_{\rm SM})/\mathcal{L}(J^P_{\rm alt})\right)$ are shown. The different shades of blue (SM) and red (alternative hypothesis) indicate the $1-3\,\sigma$ compatibility regions with the median expectation value of each hypothesis indicated as a dotted line. The solid black dots show the corresponding data points. (from [5])

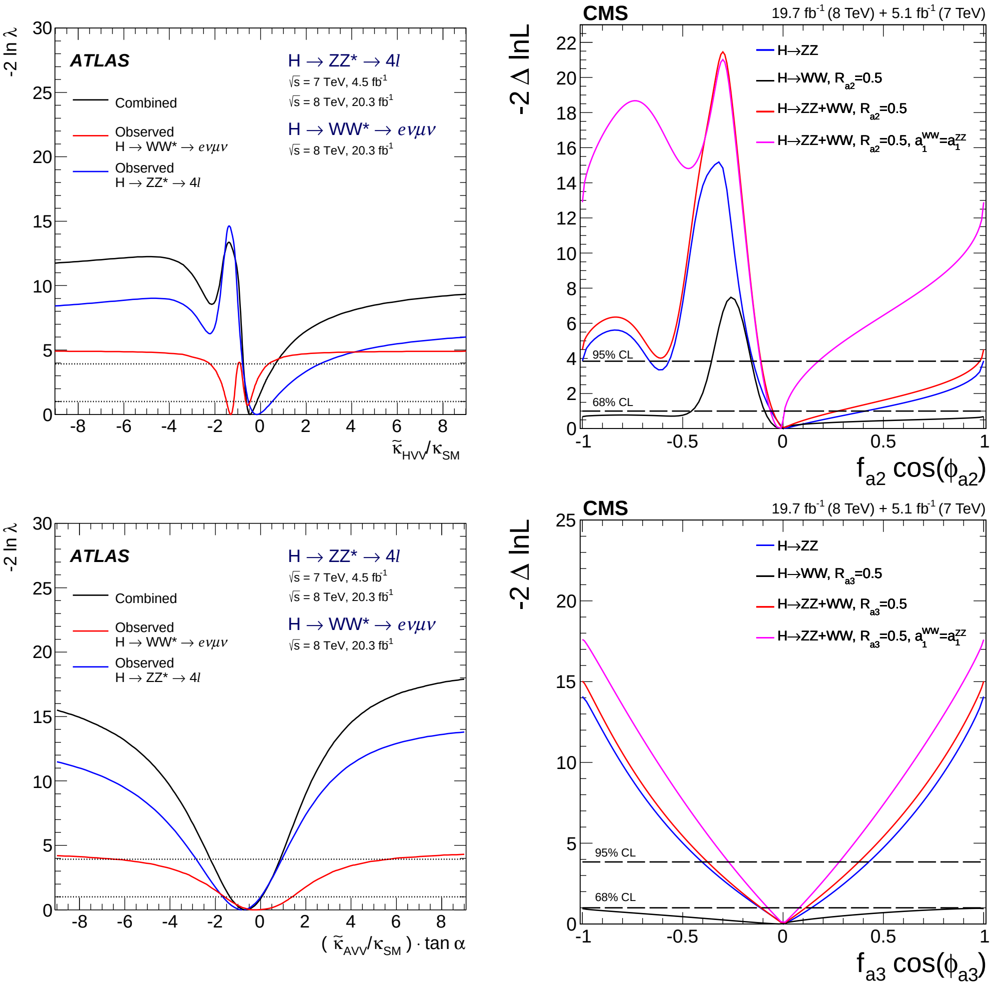


Figure 12.6.: Combined CP-mixing test results from the $H \to ZZ^* \to 4\ell$ and $H \to WW^* \to e\nu\mu\nu$ channels from ATLAS (left column) and CMS (right column). Each plot shows the logarithmic likelihood (difference) over the scan range of the respective mixing parameter. The ATLAS plots in the left column show the observed results from the two single channels as well as their combination. The CMS plots in the right column show the results of the single channels and two combination scenarios. The top row contains plots for the BSM CP-even mixing scan while the bottom row shows the CP-odd mixing scan results. The dashed lines correspond to one (and two) sigma exclusion limits. A value of 0 on the x-axis corresponds to a pure Standard Model state. (from [5] (ATLAS) and [128] (CMS), consult for more details)

12.2. Summary

The discovery of the Higgs boson has rung in a new era in particle physics as the last missing piece of the puzzle of Standard Model particles has finally been unveiled. Over the course of working on this thesis it was possible to observe and support the transition from an observed excess in the invariant mass spectrum of different decay channels, to the evidence for a new resonance, to the confirmation of the Higgs discovery. Where a long sought for goal has been reached, a new world opens up and will enthral particle physicists in the foreseeable future. In this sense the questions discussed in this analysis light the way for further studies checking the Standard Model for hints to new physics. Even though all current findings indicate that the discovered resonance is in perfect agreement with the Standard Model predictions (which of course is once more a huge triumph of this theoretical framework), a widened focus in the already started second run of the LHC will bring new insights. A number of questions can only be answered with a much larger dataset than the one available today, as discussed in the following outlook section.

Summarising the results of this analysis, one can state that the hypothesis that the discovered 'Higgs' particle has spin-2 is strongly disfavoured and all spin-2 models can be excluded at a confidence level of 82.5 to 99.1 % from the $H \to WW^* \to e\nu\mu\nu$ perspective alone. Adding the other channels, the rejection increases to more than 99.9 % C.L.

A similar picture arises for the spin-0 tests in the $H \to WW^*$ channel with either higher dimensional operators for a CP-even or a CP-odd Higgs boson. While the CP-odd hypothesis can be excluded at roughly the same confidence level as the spin-2 hypothesis (97.1 % CL), the exclusion of the 0_h^+ hypothesis falls short of the former result (64.5 %). Nevertheless, for the CP tests, the combination of exclusion limits from both sensitive channels ($H \to WW^*$ and $H \to ZZ^*$) also yields an exclusion way above 99.9 % confidence level.

Only narrow ranges around the pure Standard Model state have not yet been excluded in the admixture scans conducted in the $H \to WW^* \to e\nu\mu\nu$ and $H \to ZZ^* \to 4\ell$ channels.

After discussing the final results, going into the details of the analysis deserves some consideration. The smart rebinning approach as it has been introduced for this analysis is a unique feature of this thesis within the analysis group results. It has been shown to provide a beneficial way of automatically optimizing one or two dimensional bin arrangements to increase the expected sensitivities given predefined boundary conditions. This allows to optimise the run time of statistical evaluations with respect to the non-optimal simple rebinning approach without adding much complexity from an analyst's perspective. In addition it automatically helps avoid unfavourable binning choices which would otherwise unnecessarily decrease the obtained significance from hypothesis tests.

During the work on this thesis, contributions to the analysis as presented in the group's common publications [2], [3], [6] and [5] have been made in different regimes. This comprises substantial contributions to the entire analysis with an emphasis on the Monte Carlo studies and the statistical evaluation. Additionally it has been possible to contribute to the combined publication of all bosonic channels in the latter regime.

All in all one can close by stating that there is still a promising and compelling way to go. Nevertheless, the combined ATLAS and CMS measurements based on the recorded data of the first LHC run have managed to take a large step ahead in understanding the Higgs boson.

12.3. Run 2 Outlook

The total integrated luminosity expected to be collected by the ATLAS detector by the end of Run 2 of the LHC[1] amounts about $100\,\text{fb}^{-1}$ of data at centre of mass energies between 13 and 14 TeV

---

[1] To date $3.9\,\text{fb}^{-1}$ of data have been recorded by ATLAS during Run 2 at a centre of mass energy of 13 TeV[131].

[132]. This exceeds the combined Run 1 dataset of about 25 fb^{-1} [30] by far and permits to study rarer processes and an extended invariant mass range.

Additionally, at the higher centre of mass energy of 13 TeV and above, higher cross sections for all Higgs production processes are predicted, e.g. the production cross section for gluon-gluon fusion is supposed to be increased by a factor of around 2.3 assuming a Higgs mass of 125.0 GeV[133]. Since the cross sections of various backgrounds such as the top quark background increase with higher cross sections as well, the need for optimising the background suppression becomes an even more important challenge.

The higher yields of Higgs bosons allow to study additional aspects of the extended Higgs Lagrangian. This comprises among other options, taking the further couplings parameters besides the coupling to WW into account (see section 6.1.2 for the Lagrangian) which optimizes the current modelling and covers a wider range of possible non-SM hypotheses. Also studying the differential operator in the Lagrangian, which has been omitted for the current analysis, may be within reach of statistics. Additional studies of the CP-mixing hypotheses with larger datasets will be possible as well as an extension of the scans that have already been performed. Instead of varying only one parameter at a time (ratio of BSM/SM 0^+ couplings on the one hand and 0^+ vs 0^- couplings on the other hand) it will be possible to conduct simultaneous multi dimensional scans.

Bibliography

[1] LHC Higgs Cross Section Working Group. *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*. Tech. rep. CERN-2012-002. Geneva, 2012. arXiv: `1201.3084`. URL: `http://cds.cern.ch/record/1416519`.

[2] The ATLAS Collaboration. *Study of the spin properties of the Higgs-like particle in the $H \to WW^{(*)} \to e\nu\mu\nu$ channel with 21 fb^{-1} of $\sqrt{s} = 8$ TeV data collected with the ATLAS detector*. Tech. rep. ATLAS-CONF-2013-031. Geneva: CERN, Mar. 2013. URL: `http://cds.cern.ch/record/1527127`.

[3] The ATLAS Collaboration. "Evidence for the spin-0 nature of the Higgs boson using ATLAS data". In: *Physics Letters B* 726.1–3 (2013), pp. 120–144. ISSN: 0370-2693. DOI: `10.1016/j.physletb.2013.08.026`. URL: `http://www.sciencedirect.com/science/article/pii/S0370269313006527`.

[4] Particle Data Group. "Review of Particle Physics". In: *Phys. Rev. D* 86 (2012), p. 010001. URL: `http://pdg.lbl.gov/2013/reviews/rpp2013-rev-higgs-boson.pdf`.

[5] The ATLAS Collaboration. "Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector". In: *The European Physical Journal C* 75.10, 476 (2015). ISSN: 1434-6044. DOI: `10.1140/epjc/s10052-015-3685-1`. URL: `https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/HIGG-2013-17/`.

[6] The ATLAS Collaboration. "Determination of spin and parity of the Higgs boson in the $WW^{*} \to e\nu\mu\nu$ decay channel with the ATLAS detector". In: *The European Physical Journal C* 75.5, 231 (2015). ISSN: 1434-6044. DOI: `10.1140/epjc/s10052-015-3436-3`.

[7] R. Oerter. *The Theory of Almost Everything: The Standard Model, the Unsung Triumph of Modern Physics*. Pi Press, 2006. ISBN: 9780132366786.

[8] MissMJ, Nasfarley88 et al. *Standard Model of Elementary Particles*. 2014. URL: `https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles_modified_version.svg`.

[9] J. H. Christenson et al. "Evidence for the 2π Decay of the K_2^0 Meson". In: *Phys. Rev. Lett.* 13 (4 July 1964), pp. 138–140. DOI: `10.1103/PhysRevLett.13.138`. URL: `http://link.aps.org/doi/10.1103/PhysRevLett.13.138`.

[10] M. E. Peskin and D. V. Schroeder. *An introduction to quantum field theory*. 5., (corr.) printing. Reading, Mass.: Addison Wesley, 1997, XXII, 842 S. ISBN: 0-201-50397-2.

[11] Particle Data Group. “Review of Particle Physics”. In: *Chin. Phys.* C38 (2014), p. 090001. DOI: `10.1088/1674-1137/38/9/090001`.

[12] D. Green. *Lectures in Particle Physics*. World Scientific lecture notes in physics. World Scientific, 1994. ISBN: 9789810216832.

[13] F. Close. *The Infinity Puzzle: The Personalities, Politics, and Extraordinary Science Behind the Higgs Boson*. OUP Oxford, 2013. ISBN: 9780199673308.

[14] Particle Data Group. “Review of Particle Physics”. In: *Phys. Rev. D* 86 (1 July 2012), p. 010001. DOI: `10.1103/PhysRevD.86.010001`. URL: `http://link.aps.org/doi/10.1103/PhysRevD.86.010001`.

[15] The CMS Collaboration. “The CMS experiment at the CERN LHC”. In: *Journal of Instrumentation* 3.08 (2008), S08004. URL: `http://stacks.iop.org/1748-0221/3/i=08/a=S08004`.

[16] The ATLAS Collaboration. “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *Journal of Instrumentation* 3.08 (2008), S08003. URL: `http://stacks.iop.org/1748-0221/3/i=08/a=S08003`.

[17] The LHCb Collaboration. “The LHCb Detector at the LHC”. In: *Journal of Instrumentation* 3.08 (2008), S08005. URL: `http://stacks.iop.org/1748-0221/3/i=08/a=S08005`.

[18] The ALICE Collaboration. “The ALICE experiment at the CERN LHC”. In: *Journal of Instrumentation* 3.08 (2008), S08002. URL: `http://stacks.iop.org/1748-0221/3/i=08/a=S08002`.

[19] The LHCf Collaboration et al. “The LHCf detector at the CERN Large Hadron Collider”. In: *Journal of Instrumentation* 3.08 (2008), S08006. URL: `http://stacks.iop.org/1748-0221/3/i=08/a=S08006`.

[20] The TOTEM Collaboration. “The TOTEM Experiment at the CERN Large Hadron Collider”. In: *Journal of Instrumentation* 3.08 (2008), S08007. URL: `http://stacks.iop.org/1748-0221/3/i=08/a=S08007`.

[21] The MoEDAL Collaboration. *Technical Design Report of the MoEDAL Experiment*. Tech. rep. CERN-LHCC-2009-006. MoEDAL-TDR-001. Geneva: CERN, June 2009. URL: `https://cds.cern.ch/record/1181486`.

[22] G. Brianti et al. “Summary report”. In: *ECFA-CERN Workshop on large hadron collider in the LEP tunnel, Lausanne and CERN, Geneva, Switzerland, 21-27 Mar 1984: Proceedings. 1*. Sept. 1984, pp. 1–18. URL: `http://cds.cern.ch/record/151593`.

[23] A. M. Asner et al. “Chapter II: a feasibility study of possible machine options”. In: *ECFA-CERN Workshop on large hadron collider in the LEP tunnel, Lausanne and CERN, Geneva, Switzerland, 21-27 Mar 1984: Proceedings. 1*. Sept. 1984, pp. 49–142. URL: `http://cds.cern.ch/record/866884`.

[24] L. Evans and P. Bryant. “LHC Machine”. In: *Journal of Instrumentation* 3.08 (2008), S08001. URL: `http://stacks.iop.org/1748-0221/3/i=08/a=S08001`.

[25] M. Lamont. “Status of the LHC”. In: *Journal of Physics: Conference Series* 455.1 (2013), p. 012001. URL: `http://stacks.iop.org/1742-6596/455/i=1/a=012001`.

[26] W. Herr and B. Muratori. “Concept of luminosity”. In: *CAS - CERN Accelerator School: Intermediate Course on Accelerator Physics* (2006), pp. 361–378. DOI: `10.5170/CERN-2006-002.361`. URL: `http://cds.cern.ch/record/941318`.

[27] K. Wille. *Physik der Teilchenbeschleuniger und Synchrotronstrahlungsquellen: Eine Einführung*. German. Stuttgart: Teubner Verlag, 1992. ISBN: 3-519-03087-X.

[28] O. S. Bruning et al. “LHC Design Report Vol.1: The LHC Main Ring”. In: (2004). Ed. by O. S. Bruning. URL: `http://inspirehep.net/record/656250`.

[29] C. Lefevre. “LHC: the guide (English version). Guide du LHC (version anglaise)”. (accessed 03.03.2015). Feb. 2009. URL: `http://cds.cern.ch/record/1165534/`.

[30] The ATLAS Collaboration. *Integrated luminosity summary plots for 2011-2012 data taking*. Website. (accessed 23.02.2016). URL: `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults#Integrated_luminosity_summary_pl`.

[31] Y. Hama. “A Note on Lorentz Transformation and Pseudo-Rapidity Distributions”. In: *Journal of the Physical Society of Japan* 50.1 (1981), pp. 21–23. DOI: `10.1143/JPSJ.50.21`.

[32] C. Berger. *Elementarteilchenphysik : Von den Grundlagen zu den modernen Experimenten*. German. 2006. URL: `http://scans.hebis.de/HEBCGI/show.pl?18442934_toc.html`.

[33] J. Pequenao. “Computer generated image of the whole ATLAS detector”. (accessed 14.04.2015). Mar. 2008. URL: `http://cds.cern.ch/record/1095924`.

[34] BeST. *Beryllium Beam Pipe Used in Centrepiece of CERN Large Hadron Collider*. Website. (accessed 12.05.2015). URL: `http://beryllium.eu/news/beryllium-beam-pipe-used-centrepiece-cern-large-hadron-collider/`.

[35] J. Pequenao. “Computer generated image of the ATLAS inner detector”. (accessed 12.05.2015). Mar. 2008. URL: `http://cds.cern.ch/record/1095926`.

[36] T. Petersen. “Public TRT Plots for Collision Data: Notes of explanation”. (accessed 07.01.2016). July 2014. URL: `https://twiki.cern.ch/twiki/pub/AtlasPublic/TRTPublicResults/TRTetaPID.png`.

[37] The ATLAS Collaboration. *ATLAS muon spectrometer: Technical design report*. Tech. rep. 1997. URL: `https://inspirehep.net/record/455675`.

[38] C. Bernius. “The ATLAS trigger menu: Design, performance and monitoring”. In: *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2012 IEEE*. Oct. 2012, pp. 1382–1385. DOI: `10.1109/NSSMIC.2012.6551337`.

[39] The ATLAS and Medipix-2 Collaborations. “Luminosity Monitoring in ATLAS with MPX Detectors”. In: *JINST* 9.01 (2014), p. C01027. DOI: `10.1088/1748-0221/9/01/C01027`. arXiv: `1312.6296 [physics.ins-det]`.

[40] The ATLAS Collaboration. “Improved luminosity determination in pp collisions at $\sqrt{(s)} = 7\,\mathrm{TeV}$ using the ATLAS detector at the LHC”. In: *Eur. Phys. J.* C73.8 (2013), p. 2518. DOI: `10.1140/epjc/s10052-013-2518-3`. arXiv: `1302.4393 [hep-ex]`.

[41] The ATLAS Collaboration. “LuminosityPublicResults: Preliminary plots on luminosity distribution evolution during 2012 VdM scans”. (accessed 21.01.2016). Mar. 2013. URL: `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults#Preliminary_plots_on_luminosity`.

[42] The ATLAS luminosity measurement taskforce. “Photos And Drawings Of The Forward Detectors”. (accessed 07.01.2016). Feb. 2010. URL: `http://atlas.web.cern.ch/Atlas/TCOORD/Activities/CommonSys/Shielding/photos/atlas/lumi_detectors.pdf`.

[43] J. Pequenao and P. Schaffner. “An computer generated image representing how ATLAS detects particles”. (accessed 28.07.2015). Jan. 2013. URL: `https://cds.cern.ch/record/1505342`.

[44] The ATLAS Collaboration. *ATLAS Computing: technical design report*. Technical Design Report ATLAS. Geneva: CERN, 2005. URL: `https://cds.cern.ch/record/837738`.

[45] The ATLAS Collaboration. *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*. Tech. rep. ATLAS-CONF-2012-042. Geneva: CERN, Mar. 2012. URL: `http://cds.cern.ch/record/1435196`.

[46] W. Lampl et al. *Calorimeter Clustering Algorithms: Description and Performance*. Tech. rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003. Geneva: CERN, Apr. 2008. URL: `https://cds.cern.ch/record/1099735`.

[47] The ATLAS Collaboration. “Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data”. In: *The European Physical Journal C* 74.11, 3130 (2014). ISSN: 1434-6044. DOI: `10.1140/epjc/s10052-014-3130-x`.

[48] H. Sorenson. “Least-squares estimation: from Gauss to Kalman”. In: *IEEE Spectrum* 7.7 (July 1970), pp. 63–68. ISSN: 0018-9235. DOI: `10.1109/MSPEC.1970.5213471`.

[49] R. O. Duda and P. E. Hart. “Use of the Hough Transformation to Detect Lines and Curves in Pictures”. In: *Commun. ACM* 15.1 (Jan. 1972), pp. 11–15. ISSN: 0001-0782. DOI: `10.1145/361237.361242`. URL: `http://doi.acm.org/10.1145/361237.361242`.

[50] J. Mattmann and C. Schmitt. “Track finding in ATLAS using GPUs”. In: *Journal of Physics: Conference Series* 396.2 (2012), p. 022035. URL: `http://stacks.iop.org/1742-6596/396/i=2/a=022035`.

[51] M. Thomson. *Modern Particle Physics*. Modern Particle Physics. Cambridge University Press, 2013, XVI, 554 p. ISBN: 9781107034266.

[52] E. Fermi. *Nuclear physics : a course given by Enrico Fermi at the Univ. of Chicago*. Revised ed., 10. impr. Chicago: Univ. of Chicago Press, 1963, IX, 248 S.

[53] A. Martin et al. “Parton distributions for the LHC”. In: *The European Physical Journal C* 63.2 (2009), pp. 189–285. ISSN: 1434-6044. DOI: `10.1140/epjc/s10052-009-1072-5`.

[54] J. M. Campbell, J. W. Huston and W. J. Stirling. “Hard interactions of quarks and gluons: a primer for LHC physics”. In: *Reports on Progress in Physics* 70.1 (2007), p. 89. URL: `http://stacks.iop.org/0034-4885/70/i=1/a=R02`.

[55] G. Watt. *MSTW 2008 LO PDFs (68% C.L.)* Website. (accessed 09.07.2015). URL: `http://mstwpdf.hepforge.org/plots/plots.html`.

[56] B. Andersson et al. “Parton fragmentation and string dynamics”. In: *Physics Reports* 97.2–3 (1983), pp. 31–145. ISSN: 0370-1573. DOI: `10.1016/0370-1573(83)90080-7`. URL: `http://www.sciencedirect.com/science/article/pii/0370157383900807`.

[57] R. D. Field and S. Wolfram. “A QCD model for e^+e^- annihilation”. In: *Nuclear Physics B* 213.1 (1983), pp. 65–84. ISSN: 0550-3213. DOI: `10.1016/0550-3213(83)90175-X`. URL: `http://www.sciencedirect.com/science/article/pii/055032138390175X`.

[58] B. R. Webber. “Fragmentation and hadronization”. In: *Int. J. Mod. Phys.* A15S1 (2000), pp. 577–606. DOI: `10.1142/S0217751X00005334`. arXiv: `hep-ph/9912292 [hep-ph]`. URL: `http://inspirehep.net/record/511385`.

[59] D. Amati and G. Veneziano. “Preconfinement as a property of perturbative QCD”. In: *Physics Letters B* 83.1 (1979), pp. 87–92. ISSN: 0370-2693. DOI: `10.1016/0370-2693(79)90896-7`.

[60] The GEANT4 Collaboration. “GEANT4: A Simulation toolkit”. In: *Nucl. Instrum. Meth. A* 506 (2003), pp. 250–303. DOI: `10.1016/S0168-9002(03)01368-8`.

[61] The ATLAS Collaboration. “The ATLAS Simulation Infrastructure”. In: *Eur. Phys. J. C* 70 (2010), pp. 823–874. DOI: `10.1140/epjc/s10052-010-1429-9`. arXiv: `1005.4568 [physics.ins-det]`.

[62] Z. Marshall and the Atlas Collaboration. “Simulation of Pile-up in the ATLAS Experiment”. In: *Journal of Physics: Conference Series* 513.2 (2014), p. 022024. URL: `http://stacks.iop.org/1742-6596/513/i=2/a=022024`.

[63] 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. DOI: `10.1016/j.cpc.2012.10.032`. arXiv: `hep-ph/0405247 [hep-ph]`.

[64] P. Nason. “A New method for combining NLO QCD with shower Monte Carlo algorithms”. In: *JHEP* 0411 (2004), p. 040. DOI: `10.1088/1126-6708/2004/11/040`. arXiv: `hep-ph/0409146 [hep-ph]`.

[65] M. L. Mangano et al. “ALPGEN, a generator for hard multiparton processes in hadronic collisions”. In: *JHEP* 0307 (2003), p. 001. DOI: `10.1088/1126-6708/2003/07/001`. arXiv: `hep-ph/0206293 [hep-ph]`.

[66] T. Gleisberg et al. “Event generation with SHERPA 1.1”. In: *JHEP* 0902 (2009), p. 007. DOI: `10.1088/1126-6708/2009/02/007`. arXiv: `0811.4622 [hep-ph]`.

[67] N. Kauer. “Interference effects for $H \to WW/ZZ \to \ell\bar{\nu}_\ell\bar{\ell}\nu_\ell$ searches in gluon fusion at the LHC”. In: *JHEP* 1312 (2013), p. 082. DOI: `10.1007/JHEP12(2013)082`. arXiv: `1310.7011 [hep-ph]`.

[68] T. Sjostrand, S. Mrenna and P. Z. Skands. “PYTHIA 6.4 Physics and Manual”. In: *JHEP* 0605 (2006), p. 026. DOI: `10.1088/1126-6708/2006/05/026`. arXiv: `hep-ph/0603175 [hep-ph]`.

[69] T. Sjostrand, S. Mrenna and P. Z. Skands. “A Brief Introduction to PYTHIA 8.1”. In: *Comput. Phys. Commun.* 178 (2008), pp. 852–867. DOI: `10.1016/j.cpc.2008.01.036`. arXiv: `0710.3820 [hep-ph]`.

[70] G. Corcella et al. “HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)”. In: *JHEP* 0101 (2001), p. 010. DOI: `10.1088/1126-6708/2001/01/010`. arXiv: `hep-ph/0011363 [hep-ph]`.

[71] J. Butterworth, J. R. Forshaw and M. Seymour. “Multiparton interactions in photoproduction at HERA”. In: *Z. Phys. C* 72 (1996), pp. 637–646. DOI: `10.1007/s002880050286`. arXiv: `hep-ph/9601371 [hep-ph]`.

[72] J. Alwall et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In: *JHEP* 1407 (2014), p. 079. DOI: `10.1007/JHEP07(2014)079`. arXiv: `1405.0301 [hep-ph]`.

[73] M. Grazzini and H. Sargsyan. “Heavy-quark mass effects in Higgs boson production at the LHC”. In: *JHEP* 1309 (2013), p. 129. DOI: `10.1007/JHEP09(2013)129`. arXiv: `1306.4581 [hep-ph]`.

[74] D. de Florian et al. “Higgs boson production at the LHC: transverse momentum resummation effects in the $H \to \gamma\gamma$, $H \to WW \to l\nu l\nu$ and $H \to ZZ \to 4l$ decay modes”. In: *JHEP* 06 (2012), p. 132. DOI: `10.1007/JHEP06(2012)132`. arXiv: `1203.6321 [hep-ph]`. URL: `https://inspirehep.net/record/1095242`.

[75] J. M. Campbell, R. K. Ellis and C. Williams. “Vector boson pair production at the LHC”. In: *JHEP* 1107 (2011), p. 018. DOI: `10.1007/JHEP07(2011)018`. arXiv: `1105.0020 [hep-ph]`.

[76] M. Czakon and A. Mitov. “Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders”. In: *Comput. Phys. Commun.* 185 (2014), p. 2930. DOI: `10.1016/j.cpc.2014.06.021`. arXiv: `1112.5675 [hep-ph]`.

[77] S. Catani et al. “Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO”. In: *Phys. Rev. Lett.* 103 (2009), p. 082001. DOI: `10.1103/PhysRevLett.103.082001`. arXiv: `0903.2120 [hep-ph]`.

[78] Y. Gao et al. “Spin determination of single-produced resonances at hadron colliders”. In: *Phys. Rev. D* 81 (7 Apr. 2010), p. 075022. DOI: `10.1103/PhysRevD.81.075022`. arXiv: `1001.3396`. URL: `http://link.aps.org/doi/10.1103/PhysRevD.81.075022`.

[79] S. Bolognesi et al. “On the spin and parity of a single-produced resonance at the LHC”. In: *Phys. Rev.* D86 (2012), p. 095031. DOI: `10.1103/PhysRevD.86.095031`. arXiv: `1208.4018 [hep-ph]`.

[80] I. Anderson et al. “Constraining anomalous HVV interactions at proton and lepton colliders”. In: *Phys. Rev. D* 89 (3 Feb. 2014), p. 035007. DOI: `10.1103/PhysRevD.89.035007`. URL: `http://link.aps.org/doi/10.1103/PhysRevD.89.035007`.

[81] LHC Higgs Cross Section Working Group. *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group*. Tech. rep. CERN-2013-004. Geneva, 2013. arXiv: `1307.1347`. URL: `https://cds.cern.ch/record/1559921`.

[82] LHC Higgs Cross Section WG (Rei Tanaka, Chiara Mariotti). “SM Higgs production cross sections at $\sqrt{s} = 8\,\text{TeV}$ (update in CERN Report3)”. (accessed 07.01.2016). Nov. 2014. URL: `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt8TeV`.

[83] LHC Higgs Cross Section WG (Rei Tanaka, Chiara Mariotti). “Higgs production cross sections”. (accessed 22.07.2015). Oct. 2012. URL: `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWGCrossSectionsFigures`.

[84] The CMS Collaboration. “Limits on the Higgs boson lifetime and width from its decay to four charged leptons”. In: *Phys. Rev. D* 92 (7 Oct. 2015), p. 072010. DOI: `10.1103/PhysRevD.92.072010`. URL: `http://link.aps.org/doi/10.1103/PhysRevD.92.072010`.

[85] The CMS Collaboration. “Study of the Mass and Spin-Parity of the Higgs Boson Candidate Via Its Decays to Z Boson Pairs”. In: *Phys. Rev. Lett.* 110.8 (2013), p. 081803. DOI: `10.1103/PhysRevLett.110.081803`. arXiv: `1212.6639 [hep-ex]`.

[86] The CMS Collaboration. “Measurement of Higgs boson production and properties in the WW decay channel with leptonic final states”. In: *Journal of High Energy Physics* 2014.1, 96 (2014). DOI: `10.1007/JHEP01(2014)096`.

[87] The ATLAS and CMS Collaborations. “Combined Measurement of the Higgs Boson Mass in *pp* Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments”. In: *Phys. Rev. Lett.* 114 (2015), p. 191803. DOI: `10.1103/PhysRevLett.114.191803`. arXiv: `1503.07589 [hep-ex]`. URL: `https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/HIGG-2014-14/`.

[88] The ATLAS and CMS Collaborations. *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV*. Tech. rep. 2015. URL: `http://cds.cern.ch/record/2052552`.

[89] L. Rokach. *Data Mining with Decision Trees: Theory and Applications*. Series in machine perception and artificial intelligence. World Scientific Publishing Company Pte Limited, 2007. ISBN: 9789812771728.

[90] P. Speckmayer et al. “The toolkit for multivariate data analysis, TMVA 4”. In: *Journal of Physics: Conference Series* 219.3 (2010), p. 032057. URL: `http://stacks.iop.org/1742-6596/219/i=3/a=032057`.

[91] T. G. Dietterich. “An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization”. In: *Machine Learning* 40.2 (2000), pp. 139–157. ISSN: 0885-6125. DOI: `10.1023/A:1007607513941`.

[92] A. Hoecker et al. “TMVA - Toolkit for Multivariate Data Analysis”. In: *ArXiv Physics e-prints* (Mar. 2007). arXiv: `physics/0703039`.

[93] P. E. Black. *”greedy algorithm“*. in *Dictionary of Algorithms and Data Structures* [online], Vreda Pieterse and Paul E. Black eds. (accessed 11.06.2015). Feb. 2005.

[94] M. Trottier-McDonald and D. O’Neil. “Evidence for the Standard Model Higgs Boson Decaying to One Semi-hadronically Decaying Tau and One Leptonically Decaying Tau at ATLAS”. PhD thesis. Burnaby, British Columbia, Canada: Simon Fraser U., June 2014. URL: `http://cds.cern.ch/record/1709068`.

[95] I. M. Chakravarti, R. G. Laha and J. Roy. *Handbook of methods of applied statistics: 1. Techniques of computation, descriptive methods, and statistical inference*. New York: Wiley, 1967.

[96] G. Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: *The European Physical Journal C* 71.2, 1554 (2011). ISSN: 1434-6044. DOI: `10.1140/epjc/s10052-011-1554-0`.

[97] J. Neyman and E. S. Pearson. “On the Problem of the Most Efficient Tests of Statistical Hypotheses”. In: *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 231 (1933), pp. 289–337. ISSN: 02643952. URL: `http://www.jstor.org/stable/91247`.

[98] O. Behnke et al. *Data Analysis in High Energy Physics: A Practical Guide to Statistical Methods*. Weinheim: Wiley-VCH, 2013. ISBN: 9783527653430.

[99] A. L. Read. “Modified frequentist analysis of search results (the CL_s method)”. In: *Workshop on confidence limits, CERN, Geneva, Switzerland, 17-18 Jan 2000: Proceedings*. 2000. URL: `http://cds.cern.ch/record/451614`.

[100] A. L. Read. “Presentation of search results: The CL(s) technique”. In: *J. Phys. G* 28 (2002), pp. 2693–2704. DOI: `10.1088/0954-3899/28/10/313`.

[101] G. J. Feldman and R. D. Cousins. “Unified approach to the classical statistical analysis of small signals”. In: *Phys. Rev. D* 57 (7 Apr. 1998), pp. 3873–3889. DOI: `10.1103/PhysRevD.57.3873`. arXiv: `physics/9711021`. URL: `http://link.aps.org/doi/10.1103/PhysRevD.57.3873`.

[102] The ATLAS Collaboration. “Observation and measurement of Higgs boson decays to WW^* with the ATLAS detector”. In: (2014). arXiv: `1412.2641 [hep-ex]`. URL: `http://inspirehep.net/record/1333228`.

[103] P. Artoisenet et al. “A framework for Higgs characterisation”. In: *Journal of High Energy Physics* 2013.11, 43 (2013). DOI: `10.1007/JHEP11(2013)043`.

[104] L. Randall and R. Sundrum. “A Large mass hierarchy from a small extra dimension”. In: *Phys. Rev. Lett.* 83 (1999), pp. 3370–3373. DOI: `10.1103/PhysRevLett.83.3370`. arXiv: `hep-ph/9905221 [hep-ph]`.

[105] The ATLAS Collaboration. “Measurements of fiducial and differential cross sections for Higgs boson production in the diphoton decay channel at $\sqrt{s} = 8$ TeV with ATLAS”. In: *JHEP* 1409 (2014), p. 112. DOI: `10.1007/JHEP09(2014)112`. arXiv: `1407.4222 [hep-ex]`.

[106] The ATLAS Collaboration. “Fiducial and differential cross sections of Higgs boson production measured in the four-lepton decay channel in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector”. In: *Phys. Lett. B* 738 (2014), p. 234. DOI: `10.1016/j.physletb.2014.09.054`. arXiv: `1408.3226 [hep-ex]`.

[107] J. F. Gunion et al. “The Higgs Hunter’s Guide”. In: *Front. Phys.* 80 (2000), pp. 1–448. URL: `http://inspirehep.net/record/279039`.

[108] M. S. Carena et al. “CP violating MSSM Higgs bosons in the light of LEP-2”. In: *Phys. Lett. B* 495 (2000), pp. 155–163. DOI: `10.1016/S0370-2693(00)01215-6`. arXiv: `hep-ph/0009212 [hep-ph]`.

[109] E. Accomando et al. “Workshop on CP Studies and Non-Standard Higgs Physics”. In: (2006). arXiv: `hep-ph/0608079 [hep-ph]`.

[110] P. Artoisenet et al. “A framework for Higgs characterisation”. In: *JHEP* 1311 (2013), p. 043. DOI: `10.1007/JHEP11(2013)043`. arXiv: `1306.6464 [hep-ph]`.

[111] N. Cabibbo and A. Maksymowicz. “Angular Correlations in K_{e4} Decays and Determination of Low-Energy $\pi - \pi$ Phase Shifts”. In: *Phys. Rev.* 137 (2B Jan. 1965), B438–B443. DOI: `10.1103/PhysRev.137.B438`. URL: `http://link.aps.org/doi/10.1103/PhysRev.137.B438`.

[112] 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. DOI: `10.1016/j.physrep.2007.10.004`. arXiv: `hep-ph/0503172 [hep-ph]`.

[113] R. Z. Aben, S. C. M. Bentvelsen and P. M. Kluit. “Spinning the Higgs: Spin and parity measurement of the discovered Higgs-like boson in the $H \to WW \to \ell\nu\ell\nu$ decay mode”. PhD thesis. Amsterdam U., Apr. 2015. URL: `https://cds.cern.ch/record/2012081`.

[114] N. Kidonakis. “Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production”. In: *Phys. Rev. D* 83 (2011), p. 091503. DOI: `10.1103/PhysRevD.83.091503`. arXiv: `1103.2792 [hep-ph]`.

[115] The ATLAS Collaboration. *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data*. Tech. rep. ATLAS-CONF-2014-032. Geneva: CERN, June 2014. URL: `https://cds.cern.ch/record/1706245`.

[116] The ATLAS Collaboration. *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*. Tech. rep. ATLAS-CONF-2012-047. Geneva: CERN, May 2012. URL: `https://cds.cern.ch/record/1449796`.

[117] M. Cacciari, G. P. Salam and G. Soyez. "The Anti-k(t) jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. DOI: `10.1088/1126-6708/2008/04/063`. arXiv: `0802.1189 [hep-ph]`. URL: `http://inspirehep.net/record/779080`.

[118] The ATLAS Collaboration. "Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *Eur. Phys. J.* C75 (2015), p. 17. DOI: `10.1140/epjc/s10052-014-3190-y`. arXiv: `1406.0076 [hep-ex]`.

[119] G. Aad et al. "Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{\mathrm{s}} = 7$ TeV". In: *The European Physical Journal C* 73.3, 2304 (2013). ISSN: 1434-6044. DOI: `10.1140/epjc/s10052-013-2304-2`.

[120] The ATLAS Collaboration. *Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data*. Tech. rep. ATLAS-CONF-2014-046. Geneva: CERN, July 2014. URL: `http://cds.cern.ch/record/1741020`.

[121] "Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 7$ TeV with ATLAS". In: *The European Physical Journal C* 72.1, 1844 (2012). ISSN: 1434-6044. DOI: `10.1140/epjc/s10052-011-1844-6`.

[122] A. J. Barr, B. Gripaios and C. G. Lester. "Measuring the Higgs boson mass in dileptonic W-boson decays at hadron colliders". In: *Journal of High Energy Physics* 2009.07 (2009), p. 072. URL: `http://stacks.iop.org/1126-6708/2009/i=07/a=072`.

[123] The ATLAS Collaboration. "Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics". In: *ArXiv e-prints* (Dec. 2009). arXiv: `0901.0512 [hep-ex]`.

[124] B. Mellado, X. Ruan and Z. Zhang. "Extraction of Top Backgrounds in the Higgs Boson Search with the $H \to WW^\star \to \ell\ell + E_{\mathrm{T}}^{\mathrm{miss}}$ Decay with a Full-Jet Veto at the LHC". In: *Phys. Rev.* D84 (2011), p. 096005. DOI: `10.1103/PhysRevD.84.096005`. arXiv: `1101.1383 [hep-ph]`.

[125] The ATLAS Collaboration. "Improved luminosity determination in *pp* collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC". In: *Eur. Phys. J. C* 73.8 (2013), p. 2518. DOI: `10.1140/epjc/s10052-013-2518-3`. arXiv: `1302.4393 [hep-ex]`.

[126] J. C. Collins and D. E. Soper. "Angular distribution of dileptons in high-energy hadron collisions". In: *Phys. Rev. D* 16 (7 Oct. 1977), pp. 2219–2225. DOI: `10.1103/PhysRevD.16.2219`. URL: `http://link.aps.org/doi/10.1103/PhysRevD.16.2219`.

[127] V. Khachatryan et al. "Observation of the diphoton decay of the Higgs boson and measurement of its properties". In: *The European Physical Journal C* 74.10, 3076 (2014). ISSN: 1434-6044. DOI: `10.1140/epjc/s10052-014-3076-z`.

[128] The CMS Collaboration. "Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV". In: *Phys. Rev. D* 92 (1 July 2015), p. 012004. DOI: `10.1103/PhysRevD.92.012004`. URL: `http://cds.cern.ch/record/1969386`.

[129] The CMS Collaboration. "Measurement of the properties of a Higgs boson in the four-lepton final state". In: *Phys. Rev. D* 89 (9 May 2014), p. 092007. DOI: `10.1103/PhysRevD.89.092007`. URL: `http://link.aps.org/doi/10.1103/PhysRevD.89.092007`.

[130] CDF Collaboration and D0 Collaboration. "Tevatron Constraints on Models of the Higgs Boson with Exotic Spin and Parity Using Decays to Bottom-Antibottom Quark Pairs". In: *Phys. Rev. Lett.* 114 (15 Apr. 2015), p. 151802. DOI: `10.1103/PhysRevLett.114.151802`. URL: `http://link.aps.org/doi/10.1103/PhysRevLett.114.151802`.

[131] The ATLAS Collaboration. *Total Integrated Luminosity in 2015*. Website. (accessed 15.02.2016). URL: `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2#Luminosity_summary_plots_for_201`.

[132] "Physics at a High-Luminosity LHC with ATLAS". In: *Community Summer Study 2013: Snowmass on the Mississippi (CSS2013) Minneapolis, MN, USA, July 29-August 6, 2013*. 2013. arXiv: `1307.7292 [hep-ex]`. URL: `http://inspirehep.net/record/1245017`.

[133] LHC Higgs Cross Section WG (Rei Tanaka, Chiara Mariotti). "SM Higgs production cross sections at $\sqrt{s} = 13\text{-}14\,\text{TeV}$". (accessed 16.12.2015). June 2015. URL: `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt1314TeV#Parton_luminosity_ratio`.

Danksagung