
# Cross Section Measurement of a Standard Model Higgs Boson in the H→WW→ $\ell\nu\ell\nu$ Decay Channel with the ATLAS Experiment

Sebastian Moritz

Dissertation
zur Erlangung des Grades
Doktor der Naturwissenschaften

angefertigt am
Institut für Physik der
Johannes Gutenberg-Universität Mainz

# Cross Section Measurement of a Standard Model Higgs Boson in the H→WW→ $\ell\nu\ell\nu$ Decay Channel with the ATLAS Experiment

Sebastian Moritz

Dissertation zur Erlangung des Grades
Doktor der Naturwissenschaften
angefertigt am Institut für Physik im Fachbereich 08
der Johannes Gutenberg-Universität Mainz (D 77)

Datum der mündlichen Prüfung: 29.01.2016

Ich versichere, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie die Zitate kenntlich gemacht habe.

# Abstract

The Standard Model of particle physics successfully describes the elementary particles and their interactions and has been tested to the utmost precision. The discovery of the Higgs boson at the Large Hadron Collider in 2012 marks the detection of the last remaining unobserved phenomenon. Thus, it confirms the mechanism of spontaneous electroweak symmetry breaking which represents the origin of the mass of elementary particles.

The presented thesis examines the properties of a Standard Model Higgs boson in the decay channel $H \rightarrow WW \rightarrow \ell\nu\ell\nu$ with the ATLAS detector at the Large Hadron Collider. The analyzed data has been recorded at a center-of-mass energy of $\sqrt{s} = 8$ TeV and includes the complete 2012 dataset with an integrated luminosity of 20.3 fb$^{-1}$. The investigations focus on the dominant gluon fusion production process.

On the basis of the particular signature of the Higgs boson decay, a general analysis strategy is developed to separate the signal from the background processes. The channel is characterized by two isolated leptons and large missing transverse momentum. With a dedicated event selection the significance of the measurement is successively increased. The accurate determination of the background processes is of essential importance and comprises Monte Carlo simulation as well as data-driven techniques. A statistical fit is used to optimize the complex procedures. This results in an observed signal significance of $Z_0^{obs} = 4.69\ \sigma$ for a Standard Model Higgs boson with a mass of $m_H = 125$ GeV. The corresponding signal strength parameter amounts to $\hat{\mu} = 1.16\ ^{+0.31}_{-0.28}$. Finally, this value leads to the measurement of the Higgs boson production cross section, which is found to be $\sigma^{obs}_{pp\rightarrow H\rightarrow WW} = 5.49\ ^{+1.46}_{-1.37}$ pb for the investigated $\sqrt{s} = 8$ TeV proton-proton collisions at the Large Hadron Collider. Furthermore, exclusion limits on the Standard Model Higgs boson production can be determined for the mass range of 135 – 200 GeV with a confidence level of 95%. All the concluded results are in agreement with the Standard Model predictions.

# Contents

# 1

# Introduction

Since generations, scientists and philosophers wonder about the origin of the universe and the investigation of nature. Within the past 100 years, modern science has achieved enormous progress in observing and describing the phenomena of particle physics and astronomy. While astrophysics deals with the unimaginable dimensions of the universe, particle physics aspires towards smaller and smaller scales to explore the fundamental components of matter. However, both disciplines are closely related to each other. At high energies, particle physics describes the state of the early universe shortly after the big bang, when all matter was concentrated in a most confined space. At the present understanding, it is made up of the so called quarks and leptons whose interactions are mediated by bosons. All conclusions are summarized in the Standard Model of particle physics, which serves as the theoretical basis of this thesis. It unifies the fundamental forces of the electromagnetic, the weak and the strong interaction. Only in the course of baryogenesis, the elementary particles formed bound states. By the means of gravity, those are responsible for the structures of stars and solar systems observed today.

The so called Higgs mechanism plays a particularly important role in these descriptions. It explains how elementary particles gain their mass. Even though this theory has already been developed in 1964, for a long time, it could not be confirmed experimentally. This last missing piece of the puzzle has been dominating the program of high energy physics of the past decades and centered the huge efforts of this field. Not until 2012, the experimental detection of the so called Higgs boson succeeded at the Large Hadron Collider and thus completed the Standard Model of particle physics. Since then, the focus lies on the precise measurement of the newly discovered particle.

The presented thesis studies the evidence of a Standard Model Higgs boson in the $H \to WW \to \ell\nu\ell\nu$ decay channel, which has already been involved in the primal discovery. The main goal is the measurement of the production cross section of Higgs bosons for proton collisions at the Large Hadron Collider. The dataset has been recorded by the ATLAS detector at a center-of-mass energy of $\sqrt{s} = 8$ TeV.

The thesis starts with an introduction into the theoretical foundations in chapter 2, discussing the Standard Model of particle physics and the basic phenomenology of high energy proton collisions. Chapter 3 follows with the experimental composition of the Large Hadron Collider and the ATLAS detector. The reconstruction of the recorded data and the definition of the analysis objects is presented in chapter 4. Thereupon begins the main part of the thesis. Chapter 5 examines the special characteristics of the $H \to WW \to \ell\nu\ell\nu$ decay channel and develops the basic analysis strategy. The particular selection of events aiming to increase the signal significance as well as the methods to determine the background processes are illustrated elaborately in chapter 6. Moreover, the dominant sources of systematic uncertainties are explained within this context. Chapter 7 starts with a discussion of the statistical methods for the evaluation and interpretation of the selected events and reports the results of the main parameters, finally leading to the measurement of the Higgs boson production cross section. Subsequently, further results of the ATLAS collaboration which are directly connected to the provided thesis are presented in chapter 8. The concluding remarks are given by a summary of the substantial results and a short outlook on future developments.

# 2

# Theory of the Standard Model

This chapter provides a brief introduction to the Standard Model of particle physics. After an overview of the phenomenology and the particle content, the discussion turns towards the mathematical formalism of the gauge theory leading to the Lagrangian density of the Standard Model. The introduction of the Higgs mechanism and its necessity within the theoretical baseline will be needed for the interpretation of the analysis later on. While these descriptions are condensed to a comprehensive minimum, more thorough explanations may be found in the text books [1], [2] and [3] as well as the articles [4] and [5]. The recent discovery of the Higgs boson and the status of the searches at the ATLAS experiment brings this chapter to a close.

## 2.1 The Standard Model of Particle Physics

The Standard Model of particle physics (SM) describes the composition of matter and the interactions between elementary particles. With the electromagnetic, the weak and the strong interaction, it unites three of our four known forces of nature to one so called gauge theory. Only gravitation can not be included into this concept. The Standard Model has been confirmed experimentally to a tremendous precision within the last decades. However the discovery of the last missing piece – the Higgs boson – has occured quite recently in 2012. Since then the measurement of its properties is the most pressing topic in particle physics.

The elementary constituents of the matter we know today are quarks and leptons along with the gauge bosons as carrier of the interactions. For being spin-1/2 particles, quarks and leptons are called fermions. While neutrinos only feel the weak interaction, charged leptons like the electron, the muon or the tau-lepton are also interacting electromagnetically. Additionally, quarks carry colour charge and therefore participate in

strong interactions. Depending on their masses, fermions can be grouped into three generations or families. For every fermion there also exists an anti-particle which has the same mass as its partner but apart from that carries opposite quantum numbers. In table 1 the most important properties of the fundamental fermions are summarized.

| Particle | 1. Family | 2. Family | 3. Family | Q/e | Y | $\lvert\vec{I}\rvert$ | $I_3$ |
|---|---|---|---|---|---|---|---|
| Leptons | $e^L$ | $\mu^L$ | $\tau^L$ | –1 | –1 | 1/2 | –1/2 |
| | $\nu_e^L$ | $\nu_\mu^L$ | $\nu_\tau^L$ | 0 | –1 | 1/2 | +1/2 |
| | $e^R$ | $\mu^R$ | $\tau^R$ | –1 | –2 | 0 | 0 |
| Quarks | $u^L$ | $c^L$ | $t^L$ | +2/3 | 1/3 | 1/2 | +1/2 |
| | $d^L$ | $s^L$ | $b^L$ | –1/3 | 1/3 | 1/2 | –1/2 |
| | $u^R$ | $c^R$ | $t^R$ | +2/3 | 4/3 | 0 | 0 |
| | $d^R$ | $s^R$ | $b^R$ | –1/3 | –2/3 | 0 | 0 |

Table 1: Fermions of the Standard Model.

The gauge bosons are acting as mediators of the fundamental forces combined in the Standard Model. They carry integer values of spin. Gluons are responsible for the strong force while the electromagnetic force is mediated by the photons. Both of these sorts of particles are massless. However the gauge bosons of the electroweak sector are exceedingly heavy. Besides the electrically neutral $Z^0$ boson, there are also two charged $W^\pm$ bosons. All of them are listed in table 2.

| Gauge Boson | Symbol | Mass [GeV] | Q [e] | Interaction |
|---|---|---|---|---|
| Photon | $\gamma$ | 0 | 0 | electromagnetic |
| $Z^0$ Boson | $Z$ | 91.187 | 0 | electroweak |
| $W^+$ Boson | $W^+$ | 80.425 | 1 | weak |
| $W^-$ Boson | $W^-$ | 80.425 | –1 | weak |
| Gluon | $g$ | 0 | 0 | strong |

Table 2: The bosons of the Standard Model.

The mathematical foundation of the SM is a gauge invariant field theory. Hence the entire information is included in the Lagrangian density, which has to be invariant under (local) gauge transformations. All the possible transformations fulfilling this criterion are summarized in the symmetry group of the SM.

$$G_{SM} = SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \tag{2.1}$$

The subgroup $SU(3)_C$ contains the transformations in the colour space[1] and therefore describes the strong interactions. A field theory concerning only the strong force is

[1]The index $C$ stands for colour.

called Quantum Chromodynamics (QCD). Transformations of the group $SU(2)_L$ are used to describe the interactions in the weak isospin space $I$. The index $L$ refers to the circumstance that solely left-handed fermions are taking part in weak interactions. The last part of equation 2.1, the group $SU(1)_Y$, represents the gauge transformations of the so called hyper charge $Y$ which are responsible for the electromagnetic force. The electrical charge results from a combination of the hyper charge and the third component of the isospin:

$$Q/e = I^3 + Y/2 \tag{2.2}$$

## 2.2 Lagrangian Density of the SM

The entire information of the Standard Model is accommodated in its Lagrangian density. In order to give a comprehensive overview of the substantial physical conclusions, the Lagrangian is split into its main terms:

$$\mathcal{L}_{SM} = \mathcal{L}_{EW} + \mathcal{L}_{QCD} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa} \tag{2.3}$$

These single terms are now discussed in the following subchapters.

### 2.2.1 The Electroweak Unification

This section provides a summary of the so called Glashow-Weinberg-Salam (GWS) model of electroweak interaction. It unifies the electromagnetic with the weak force. This happens in the frame of a gauge theory based on the $SU(2)_L \times U(1)_Y$ symmetry. The phenomenology of the interaction of weakly charged currents leads to the interpretation that left-handed fermions occur as weak isospin doublets while right-handed fermions appear as singlets:

$$\ell_L = \begin{pmatrix} \nu_{eL} \\ e_L^- \end{pmatrix}, \begin{pmatrix} \nu_{\mu L} \\ \mu_L^- \end{pmatrix}, \begin{pmatrix} \nu_{\tau L} \\ \tau_L^- \end{pmatrix}; \ell_R = e_R,\ \mu_R,\ \tau_R \tag{2.4}$$

For the construction of the Lagrangian of the electroweak interaction it is necessary to clarify which properties it needs to have. In the case of $SU(2)_L$, the gauge transformations among which $\mathcal{L}_{EW}$ should stay invariant are

$$SU(2)_L:\ \psi_L \to \psi'_L \ = \ \exp\left(-i\sum_{k=1}^{3}\alpha_k(x)\frac{\tau_k}{2}\right)\psi_L \tag{2.5}$$

$$\psi_R \to \psi'_R \ = \ \psi_R \tag{2.6}$$

with the space dependent phase $\alpha_k$. The symmetry of the unitary group $U(1)_Y$ additionally allows the following transformations which do not distinguish between left-handed and right-handed fields:

$$U(1)_Y : \psi_{L/R} \to \psi'_{L/R} = \exp\left(-i\beta(x)\frac{Y}{2}\right)\psi_{L/R} \tag{2.7}$$

Here $\tau_k$ represents the Pauli-matrices with the familiar commutation relations

$$\left[\frac{\tau_i}{2}, \frac{\tau_j}{2}\right] = i\epsilon_{ijk}\frac{\tau_k}{2}, \tag{2.8}$$

But these phases violate the gauge invariance when transforming the Lagrangian. Consequently the gauge principle requires the following covariant derivatives to circumvent this problem:

$$D_\mu\psi_L = \left(\partial_\mu + ig\frac{\vec{\tau}\cdot\vec{W}_\mu}{2} + ig'\frac{Y}{2}B_\mu\right)\psi_L \tag{2.9}$$

$$D_\mu\psi_R = \left(\partial_\mu + ig'\frac{Y}{2}\right)\psi_R \tag{2.10}$$

This leads to four gauge fields: The isospin triplet $W_\mu^{1,2,3}$ is coupled to the weak isospin through the coupling constant $g$. On the other hand the single vector field $B_\mu$ is coupled to the hyper charge $Y$ with the strength $g'$. Now the field strength tensor of the $SU(2)$ and the $U(1)$ can be defined via these fields:

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g\epsilon^{ijk}W_{\mu,j}W_{\nu,k} \tag{2.11}$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{2.12}$$

With these components the Lagrangian of the electroweak interaction can now be constructed using the following short notation:

$$\mathcal{L}_{EW} = \overline{\Psi}i\gamma_\mu D^\mu\Psi - \frac{1}{4}\left[W_{\mu\nu,i}W^{\mu\nu,i} + B_{\mu\nu}B^{\mu\nu}\right] \tag{2.13}$$

The fields $\Psi$ symbolize the left and right-handed fields $\psi_{L/R}$. Respectively the correct covariant derivative with the corresponding coupling constant has to be chosen from the definitions above.

But the four vector fields $W_\mu^{1,2,3}$ and $B_\mu$ are not physical fields. The experimentally measured particles of the electromagnetic and the weak force are the photon $\gamma$ and the gauge bosons $Z^0$ and $W^\pm$. They arise from the mixing of the constructed quantities. The charged bosons are the outcome of the following linear combination:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \tag{2.14}$$

They couple exclusively to particles with non-zero isospin, thus to left-handed particles. Mixing the third component of the $SU(2)_L$ gauge field $W_\mu^3$ with the $U(1)_Y$ gauge field $B_\mu$ establishes the electrically neutral photon and the $Z^0$ boson:

$$A_\mu = B_\mu \cos(\Theta_W) + W_\mu^3 \sin(\Theta_W) \tag{2.15}$$
$$Z_\mu = -B_\mu \sin(\Theta_W) + W_\mu^3 \cos(\Theta_W) \tag{2.16}$$

Therefore both these physical fields couple to left-handed as well as right-handed particles. The angle $\Theta_W$ describes the mixing rate and is called weak mixing angle or Weinberg angle. It can be used to specify the following relation of the coupling constants:

$$g \sin(\Theta_W) = g' \cos(\Theta_W) = e \tag{2.17}$$

with the electrical elementary unit $e$ which results from the requirement of the photon field coupling to charged leptons with exactly this strength.

Yet there are no mass terms within the Lagrangian of the electroweak interaction, because they would violate the gauge invariance. But as is well known, the gauge bosons $Z^0$, $W^+$ and $W^-$ are not massless at all. So only the photon seems to fit into the construction so far. The introduction of the Higgs mechanism will solve this problem in the following chapters. Within the frame of this theory a so called spontaneous symmetry breaking takes place and the particles gain their mass.

### 2.2.2 The Strong Interaction

The theory of the strong interaction is entitled Quantum Chromodynamics and describes the dynamics of quarks and gluons. Its name refers to a new quantum number called colour, which serves as "charge". Mathematically spoken, it is a non-abelian gauge theory [2] with a $SU(3)$-symmetry in the colourspace. The construction of the Lagrangian density $\mathcal{L}_{QCD}$ is carried out analogously to the former case of electroweak interactions

[2] Non-abelian gauge theories are often called Yang-Mills theories since the two physicists did investigate the case of $SU(2)$ at first. The label *non-abelian* refers to non-commutating generators.

and therefore will lead to a similar structure. Initially the $SU(3)_C$ transformations under which the Lagrangian needs to stay invariant need to be phrased:

$$SU(3) : q_f = \begin{pmatrix} q_{f,r} \\ q_{f,g} \\ q_{f,b} \end{pmatrix} \rightarrow q'_f = \exp\left(-i \sum_{a=1}^{8} \epsilon(x) \frac{\lambda_a}{2}\right) q_f \tag{2.18}$$

The index f of the quark fields stands for *flavor* and distinguishes between the different quark types (up, down, charm, strange, top and bottom) while the colour charge is labeled with the indices r (red), g (green) and b (blue). The eight generators $\lambda_a/2$ of the $SU(3)$ are realized through the Gell-Mann matrices $\lambda_a$ and obey the commutation relations:

$$\left[\frac{\lambda_a}{2}, \frac{\lambda_b}{2}\right] = i f_{abc} \frac{\lambda_c}{2} \tag{2.19}$$

where the introduced structure constants of the strong interaction $f_{abc}$ (a,b,c = 1,...,8) are anti-symmetric with respect to permutations of two indices. Furthermore the gauge principle again demands a covariant derivative:

$$D_\mu q_f = \left(\partial_\mu + i g_s \sum_{a=1}^{8} \frac{\lambda_a}{2} G_\mu^a\right) q_f \tag{2.20}$$

where $g_s$ is the strong coupling constant. Now the field strength tensor describing the dynamics of the gluon fields can be defined with the eight arising gauge fields $G_\mu^a$:

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_s f^{abc} G_{\mu,b} G_{\nu,c} \tag{2.21}$$

Concluding the Lagrangian density of Quantum Chromodynamics can be written as:

$$\mathcal{L}_{QCD} = \overline{q} i \gamma_\mu D^\mu q - \frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu} \tag{2.22}$$

The characteristics of QCD become obvious once the product $G_{\mu\nu}^a G_a^{\mu\nu}$ is expanded and the different terms are arranged in orders of $g_s$:

$$\mathcal{L}_{QCD} \propto \overline{q} q + g_s \overline{q} q G + (\partial G)^2 + g_s G^2 \partial G + g_s^2 G^4 \tag{2.23}$$

While the first term describes the propagation of a free quark, the second one characterizes the coupling of quarks and gluons and the third term specifies free gluon propagation. However, two additional terms describing the gluon self-interaction occur because of the non-abelian structure of QCD. This feature causes the coupling strength $g_s$ to grow with

increasing distance. Hence free quarks or gluons cannot be observed. They only come in colour neutral states, as mesons ($q\overline{q}$) or baryons ($qqq$). This phenomenon is called confinement. It restricts the strong interaction to the scope of nucleons, although the force carriers are massless.

### 2.2.3 The Higgs Mechanism

Within the development of the theory so far, every emerging gauge boson is massless which does not reflect the observations in nature. The cause of that circumstance lies in the gauge principle. Explicit mass terms of the form $\frac{1}{2}M^2\psi^2$ as they arise in the Klein-Gordon equation violate the gauge invariance of the Lagrangian density. In the Standard Model, this problem is resolved with the introduction of the Higgs mechanism proposed by Higgs [6] in 1964 and simultaneously by Englert and Brout [7] and Guralnik, Hagen and Kibble [8]. The implementation of a Higgs potential leads to a spontaneous symmetry breaking, meaning the transition to a minor symmetry than initially demanded by the primary Lagrangian density. So the following contribution is added to the Lagrangian:

$$\mathcal{L}_{Higgs} = (D_\mu\phi)^\dagger(D^\mu\phi) - V(\phi) \tag{2.24}$$

Here $D_\mu$ is the covariant derivative as defined in 2.9. The field $\phi$ is a two component complex scalar field which can be parameterized such as:

$$\phi = \begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}\phi_1 + i\phi_2\\ \phi_3 + i\phi_4\end{pmatrix} \tag{2.25}$$

Consequently we obtain an isospin doublet with hypercharge $Y = +1$ assigned to it. Therefore the upper component (with $I^3 = +1/2$) comes with charge $Q = +1$ whereas the lower component ($I^3 = -1/2$) has charge $Q = 0$.

Now the exact form of the potential $V(\phi)$ is of utmost importance. The most general ansatz that not only is consistent with the $SU(2)_L \otimes U(1)_Y$ gauge invariance but also remains Lorentz invariant and renormalizable is:

$$V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2 \tag{2.26}$$

Choosing the parameters $\mu^2 > 0$ and $\lambda > 0$ leads to a potential bounded below with the unique ground state at $\phi = 0$ preserving the symmetry. The choice of $\mu^2 < 0$ on the contrary induces a so called mexican hat potential as shown in figure 1. In this case the ground state is degenerate. The minimum results from the derivative.

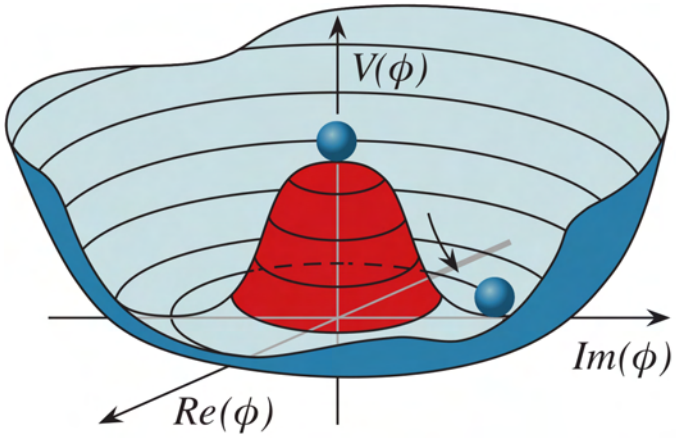



Figure 1: Illustration of the Higgs potential $V(\phi)$ with degenerate ground state.

$$\frac{\partial}{\partial \phi_k} V(\phi^\dagger \phi) = (\mu^2 + \lambda \phi_i^2)\phi_k \stackrel{!}{=} 0 \tag{2.27}$$

$$\Rightarrow \phi^\dagger \phi|_{min} = \frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2)|_{min} = -\frac{\mu^2}{2\lambda} \equiv v^2 \tag{2.28}$$

One possible solution of equation 2.28 is for example $\phi_1 = \phi_2 = \phi_4 = 0$ and $\phi_3 = \sqrt{2}v$. This leads to a vacuum expectation value of the simple form

$$\langle \phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{2.29}$$

A Taylor expansion around this vacuum expectation value eventually yields to a single complex scalar field $H(x)$:

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{2.30}$$

The operators $I_3$ and $Y$ do not annihilate this ground state, solely the application of the charge operator $Q/e = I_3 + Y/2$ can, which means that the remaining symmetry only allows transformations of the following kind:

$$\phi \to \phi' = \exp\left(-i\frac{\epsilon}{2}(1 + \tau_3)\right)\phi \tag{2.31}$$

Those transformations create a subgroub of the original symmetry group and can be interpreted as electromagnetic $U(1)_{em}$ with the photon as gauge boson. The explicit choice of the ground state has "broken" the symmetry

$$SU(2)_L \otimes U(1)_Y \overset{broken}{\longrightarrow} U(1)_{em} \tag{2.32}$$

and thereby yielded the Higgs field $H(x)$. However, the question of the generation of mass of the heavy gauge bosons still needs to be clarified just as the proof of the photon remaining massless. The insertion of the vacuum expectation value 2.29 into the Lagrangian density 2.24 gives the possibility to investigate the resulting structure with respect to terms of the form $\frac{1}{2}M^2\psi^2$. Using the definitions from section 2.2.1 the following terms can be determined from the covariant derivatives:

$$\begin{aligned} |D_\mu \langle\phi\rangle|^2 &= \frac{1}{2}(\partial_\mu H)^2 + \frac{1}{8}g^2 \left|W_\mu^1 - iW_\mu^2\right|^2 (v+H)^2 + \frac{1}{8}\left|gW_\mu^3 - g'B_\mu\right|^2 (v+H)^2 \\ &= \frac{1}{2}(\partial_\mu H)^2 + \frac{1}{4}g^2 W_\mu^+ W^{\mu-}(v+H)^2 + \frac{1}{8}Z_\mu Z^\mu (v+H)^2 \end{aligned} \tag{2.33}$$

where the following fields have been defined:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp iW_\mu^2\right) \ , \quad Z_\mu = \frac{gW_\mu^3 - g'B_\mu}{\sqrt{g^2+g'^2}} \ , \quad A_\mu = \frac{gW_\mu^3 + g'B_\mu}{\sqrt{g^2+g'^2}} \tag{2.34}$$

Besides a kinetic term of the Higgs filed, the spontaneous symmetry breaking leads to three terms quadratic in fields. Those can be associated with the mass terms for the $W^+$, $W^-$ and the $Z^0$ bosons and their mass values can be read from the prefactors after expanding equation 2.33 with the definitions 2.34:

$$M_{W^\pm}^2 = \frac{1}{4}v^2 g^2 \ , \quad M_{Z^0}^2 = \frac{1}{4}v^2(g^2+g'^2) \ , \quad M_A^2 = 0 \tag{2.35}$$

Noticeably there appears no quadratic term of the field $A_\mu$. Thus, the associated photon remains massless. The mixed terms reveal the interactions between the gauge bosons and the Higgs field which appear as $g_{VVH}$ and $g_{VVHH}$ couplings. Plugging the expansion of the vacuum expectation value into the potential 2.26 leads to an additional number of terms containing the Higgs field

$$V(\langle\phi\rangle) = \mu^2 H^2 + \lambda v H^3 + \frac{\lambda}{4}H^4 \tag{2.36}$$

and reveals another mass term, this time for the Higgs field itself. The associated particle is called Higgs-boson and its mass is given as:

$$M_H^2 = -\mu^2 = 2\lambda v^2 \tag{2.37}$$

Moreover, equation 2.36 also shows self-interaction terms of the Higgs field as triple $g_{H3}$ and quartic $g_{H4}$ couplings. The Higgs boson is interpreted as radial excitation of the Higgs field $H(x)$ close to the ground state. But even though the vacuum expectation value $v$ of the Higgs potential is fixed and can be calculated via the coupling constants and the masses of the gauge bosons or the Fermi constant $G_F$ respectively,

$$v^2 = 4\frac{M_W^2}{g^2} = \frac{1}{\sqrt{2}G_F} \approx (250\ \text{GeV})^2 \tag{2.38}$$

the mass of the Higgs boson persists undetermined because the parameter $\lambda$ describing the Higgs self-coupling is not ascertained.

The generation of the fermion masses is also based upon the Higgs mechanism. Technically it is realized by the so called Yukawa couplings which can be included in the Lagrangian density in the following manner:

$$\mathcal{L}_{Yukawa} = -G_f \left( \overline{\Psi}_f^L \phi \Psi_f^R + \overline{\Psi}_f^R \phi^\dagger \Psi_f^L \right) \tag{2.39}$$

Here $G_f$ represents the Yukawa coupling constant of each fermion. The fields labelled with $L$ are left-handed doublets whereas the index $R$ refers to right-handed singlets[3]. Introducing the spontaneous symmetry breaking by the expansion of $\phi$ analogously to equation 2.30 leads to the mass terms of the fermions

$$M_f^2 = \frac{G_f^2 v^2}{2} \tag{2.40}$$

and the coupling of fermions to the Higgs-boson $g_{ffH}$ respectively. Figure 2 illustrates all Standard Model Higgs-boson couplings described above.

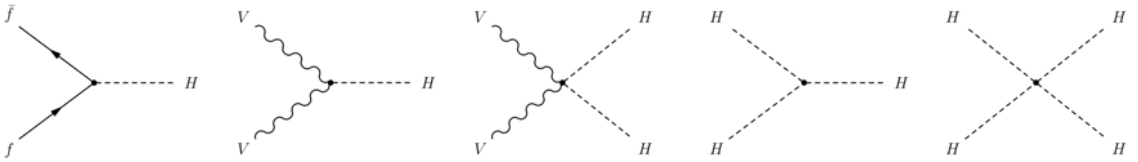



Figure 2: The Higgs boson couplings to fermions ($g_{ffH}$), W/Z gauge bosons ($g_{VVH}$ and $g_{VVHH}$) and the Higgs self-couplings ($g_{H3}$ and $g_{H4}$) in the Standard Model from left to right.

[3]In the case of the electron for example the first term would be: $-G_e \left(\overline{\nu}_e, \overline{e}\right)_L \cdot \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} e_R$

However there is one thing to be considered in the case of quark fields. Their eigenstates do not belong to a fixed family. In order to find the physical fields unitary transformations between the different fields need to be introduced. This mixing of mass eigenstates and the eigenstates of the electroweak interaction are described by the so called Cabibbo-Kobayashi-Maskawa matrix (CKM matrix) [9].

## 2.3 Phenomenology of Proton Collisions

In section 2.2.2 the QCD Lagrangian was describing the strongly interacting elementary particles, the quarks and gluons. But in order to predict the observables of the proton-proton collisions, it is indispensable to reconsider that the protons are not elementary but composite objects. So this section will develop the most important methods approaching realistic descriptions of hadron collisions. Detailed instructions can be found in [10].The first part introduces the parton model used to approximate the composite nature of the proton, leading to the factorization theorem which separates the dynamics, followed by a discussion of the general aspects of cross section calculation and the parton distribution functions. Subsequently the basic concepts of Monte Carlo event generation are presented in the light of the precedent section. This prepares the reader for the final part, which will care about the Higgs phenomenology at the Large Hadron Collider (LHC) and focuses on production and decay of the Standard Model Higgs boson.

### 2.3.1 Parton Model and Factorization Theorem

Since protons are not fundamental particles, it is very useful to go back one step and take a look at the so called parton model developed by Bjorken and Feynman in the late 60's to describe the phenomenological aspects known at that time, driven by early proton-proton collisions and deep inelastic electron-proton scattering (DIS). Here the concepts are described and directly merged with the nomenclature of the later established QCD theory and its particle content.

In this picture, the proton is made up of electrically charged point-like partons which are approximately free particles, only loosely bound by electrically neutral partons. These partons are associated with the valence quarks, sea quarks and gluons. In proton-proton collisions these constituent partons interact with each other. At high energy hadron colliders like the Large Hadron Collider, the processes can be categorized into either hard scattering appearing with large momentum transfer or soft scattering characterized by small momentum transfer between the interacting partons. Even though both categories emerge from QCD, the approaches are very different. Although the soft scattering makes up the majority of the proton-proton interactions while the hard scattering constitutes the rare processes such as Higgs boson, W or Z boson production, the level of insight behaves quite contrary. That is because for hard processes, the strong coupling strength is small and therefore observables can reliably be calculated using perturbative techniques. For soft processes on the other hand, the coupling parameter becomes significantly larger and prohibits the application of perturbation theory and therefore corrupts the calculations.

The fact that the dynamics of these two cases can be separated in terms of different momentum scales can be condensed to the so called factorization theorem [11]. An observable involving strong interactions can be parametrised by a product of two functions:

$$O(Q^2, p_{hadron}) = \mathcal{F}(Q^2, p_{parton} > \mu_F) \otimes \mathcal{D}(Q^2, p_{parton} < \mu_F) \tag{2.41}$$

where $Q^2$ is the Lorentz-invariant square of the transferred momentum and the auxiliary factorization scale $\mu_F$ is introduced. While all the short distance interactions happening at large momentum transfer are described by the function $\mathcal{F}$, the long distance physics happening at low momentum transfer that cannot be handled perturbatively are factorized into the function $\mathcal{D}$. These functions need to be extracted experimentally and can then be multiplied to the perturbative QCD calculations. This is beneficial because usually the same soft functions $\mathcal{F}$ enter several different physics processes and thus are universal. The factorization scale itself is arbitrary, it only marks the metaphorical transition between short and long distance interaction or hard and soft scattering respectively. Technically it arises from infrared (IR) divergences. But once an observable is calculated to all orders in perturbation theory, it remains invariant under changes of the factorization scale. A first successful test of the parton model was the application to the Drell-Yan (DY) process, the production of a massive lepton pair by quark-antiquark annihilation.

### 2.3.2 Cross Section and Parton Distribution Functions

The observable of interest is the cross section of a specific physics process. Figure 3 illustrates the interaction of two partons $a$ and $b$ out of the initial hadrons $A$ and $B$. Via the partonic process $ab \to X$ they produce some final state $X$. Applying the factorization theorem, the hadronic cross section $\sigma_{AB\to X}$ can now be written as:

$$\sigma_{AB} = \sum_{a,b} \int dx_a dx_b f_{a/A}(x_a, Q^2) f_{b/B}(x_b, Q^2) \hat{\sigma}_{ab\to X} \tag{2.42}$$

where $\hat{\sigma}_{ab\to X}$ denotes the partonic cross section and the summation runs over all possible partons $a$ and $b$ that can contribute to the final state $X$. The introduced parton distribution functions (PDFs) $f_{i/I}(x_i, Q^2)$ describe the probability to find a parton $i$ which carries the momentum fraction $x_i$ of the hadron $I$. Because the momentum of all the partons has to add up to the hadron momentum, the PDFs follow the normalization relation:

$$\sum_i \int_0^1 x f_i(x, Q^2) dx = 1 \tag{2.43}$$

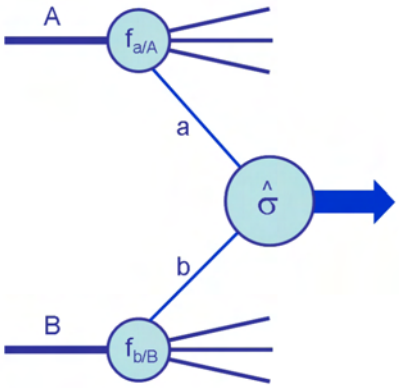



Figure 3: Diagrammatic structure of a generic hard scattering process where the partons (a/b) of the hadrons (A/B) interact and lead to a final state X [10].

independently of $Q^2$. As indicated before, they are not predicted by perturbation theory but have to be experimentally determined from data. However, their evolution depending on the momentum transfer $Q^2$ is accessible with perturbative QCD. The PDF evolution is described by the so called DGLAP[4] equations. So the transaction from high to low energy scales can be written down as:

$$\frac{\partial f_i(x,Q^2)}{\partial \log Q^2} = \frac{\alpha_S(Q^2)}{2\pi} \sum_j \int_x^1 \frac{d\xi}{\xi} P_{i\to j}(\frac{x}{\xi}) f_i(\xi, Q^2) \tag{2.44}$$

with the so called splitting function $P_{i\to j}$ giving the probability to split off a parton $j$ with a momentum fraction $\xi$ of the initial parton $i$[5].

There are several PDF sets made public by different phenomenology working groups based on the available experimental data. Figure 4 shows the MSTW group's PDFs for the proton at two different momentum transfer scales. The proton's valence quarks (up and down) carry the largest momentum fraction followed by the gluons, whereas the sea quarks contribute with lower fractions. The bands show the 68% confidence levels and represent the one-sigma uncertainties associated to the estimations which are largest for the gluon distribution, as it is the least well constrained from data, especially at low momentum transfer. A common technique to estimate these PDF uncertainties is the Hessian method [13] which utilises different orthonormal eigenvectors to propagate the input data uncertainties and subsequent excursions along the + and – directions of each vector. The evaluation of the uncertainties on physical observables such as cross sections can then be obtained by the variation resulting from these error sets.

[4]Acronym for "Dokshitzer-Gribow-Lipatow-Altarelli-Parisi". Often also referred to as Altarelli-Parisi equations.

[5]This describes the case of collinear quark or gluon emission and can be interpreted as correction to the leading order parton cross section.

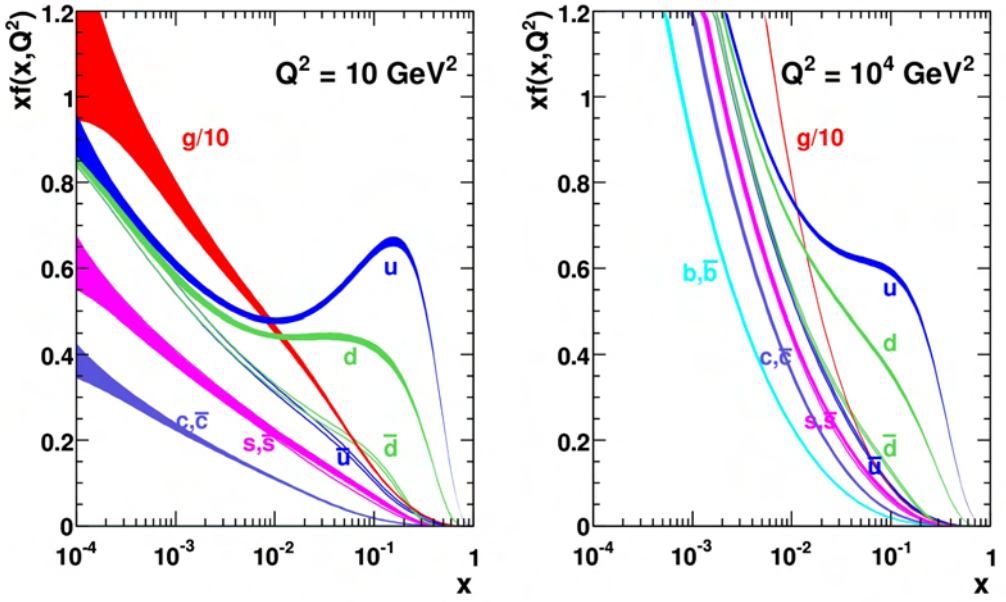



Figure 4: Parton distribution functions for the proton at momentum transfer scales of $Q^2 = 10$ GeV$^2$ (left) and $Q^2 = 10^4$ GeV$^2$ (right) as determined for the MSTW08 PDF set from [12] including the 68% confidence level bands.

Having the PDFs at hand, one can return to equation 2.42 and survey the remaining partonic cross section $\hat{\sigma}_{ab\to X}$. As explained, predictions now can be obtained with perturbative QCD calculations. Since the long distance contributions have been split via the factorization theorem, the partonic cross section can be expanded to a power series of the strong coupling parameter as follows:

$$\hat{\sigma}_{ab\to X} = \hat{\sigma}_0 + \alpha_s(Q^2, \mu_R)\hat{\sigma}_1 + \alpha_s^2(Q^2, \mu_R)\hat{\sigma}_2 + \cdots \tag{2.45}$$

where $\hat{\sigma}_0$ denotes the leading-order (LO) partonic cross section, $\hat{\sigma}_1$ the next-to-leading-order (NLO) and so on. The running coupling constant $\alpha_s(Q^2, \mu_R)$ itself depends on the momentum transfer and the so called renormalization scale $\mu_R$. This scale enters the theory during the regularization of ultra-violet (UV) divergences, singularities for large $Q^2$. However, as stated before, physical quantities have to remain independent of the choice of scale which is expressed with the so called Renormalization group equation (RGE):

$$\mu_R \frac{dR(\alpha_s(\mu_R))}{d\mu_R} = 0 \tag{2.46}$$

So the precision of the prediction can be chosen with the order in $\alpha_s$ taken from the expansion series. The simplest approach is calculating the squared matrix element at leading order represented by the tree level Feynman diagrams for the involved particles

and their couplings over the appropriate phase space. Each higher order leads to more precision, but the calculations are getting more complicated as well, because of additionally introduced singularities and often need to be carried out numerically on restricted space.

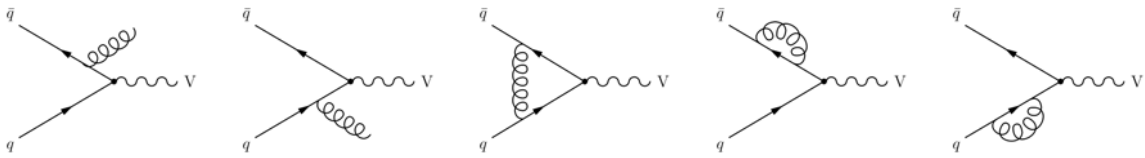


Figure 5: NLO Feynman diagrams with gluon corrections for the process $q\bar{q} \to V$ where $V$ stands for vector boson, representing a W or Z boson. The two left diagrams represent real emissions that cause IR divergencies, while the three diagrams on the right show virtual corrections leading to UV divergencies.

Figure 5 shows an example of additional Feynman diagrams at next-to-leading-order for the process of quark-antiquark annihilation to a vector boson. The first two diagrams on the left hand side describe real (collinear) gluon emissions leading to IR divergences (already mentioned among the splitting functions) arising at small $Q^2$. The three diagrams on the right represent virtual corrections leading to UV divergent contributions happening at large $Q^2$, expressed in renormalized quantities as for example the strong coupling constant. But it could be demonstrated that no matter to which order of the perturbation series the terms are considered, the two different kinds of singularities always cancel each other and therefore lead to a finite result of the cross section. For every higher order taken into account, the dependency on the factorization scale $\mu_F$ and the renormalization scale $\mu_R$ is reduced. It vanishes completely once the calculations are performed to all orders. As long as such a complete set of higher order corrections is absent, a specific choice for both scales is indispensable. Usually values of the typical momentum scale of the hard scattering process are practical. In the example of figure 5 a sensible choice would be $\mu_F = \mu_R = M_Z$.

A useful short notation for the extrapolation from one order to another is the so called K-factor as given by the ratio of cross sections:

$$K_{NLO} = \frac{\sigma_{NLO}}{\sigma_{LO}} \tag{2.47}$$

in this case from leading order to next-to-leading-order. However, this simple approach needs to be treated carefully as the K-factor may vary for different kinematic regions of one and the same process.

### 2.3.3 Monte Carlo Event Generation

In order to compare the experimental data of proton-proton collisions to the SM predictions, it is essential to have an accurate simulation of such physics processes. This is necessary not only to perform the final analysis of event rates and topologies but also to survey an experiment's general feasibility, to study the needed detector requirements and to develop strategies and optimizations before the actual start.

As a result of the quantum mechanical nature of particle physics, the occurrence of a predefined final state with a specific kinematical configuration produced in an inelastic hadron collision can only be predicted on a probabilistic basis. Its probability is proportional to the predicted cross section. To enhance the reliability of these probability distributions, it is necessary to average over large event samples. In particle physics it is most common to use so called Monte Carlo (MC) event generators to simulate large numbers of events with the four momenta of the final state particles.

Monte Carlo techniques describe statistical simulation methods, where the key feature is the utilization of sequences of random numbers. Typically they are used for problems with coupled degrees of freedom, multi-dimensional numerical integration or phenomena with uncertainties in the inputs. As a major ingredient, the physical problem must be described by a set of probability density functions[6] specifying the evolution of the system. Even if no stochastic content of a problem is apparent, it might however be transformed to be expressed by probability density functions, which opens up a wide scope to the application of Monte Carlo techniques. Fast and effective generation of random numbers uniformly distributed over the unit interval must be available to draw values from the probability density functions. The prescription for such a sampling of possible states from a specific probability density function is called sampling rule. Furthermore the statistical uncertainty (variance) as a function of trials needs to be determined. Thus, averaging over the outcomes of these multi-trials, it is possible to measure the precision of the predictions or rather the uncertainty of the outcome. The technical implementation of Monte Carlo techniques usually implies methods for variance reduction, parallelization and vectorization to achieve efficient use of computation resources. Due to the similarity with games of chance, the name Monte Carlo refers to the capital of Monaco, well known for its casino and gambling background. In this context, the game is realized as a physical (or mathematical) problem and the outcome is the solution to the problem, in this case rather the expected outcome of the particular experiment. Therefore the simulated events are often referred to as pseudo-data.

The complex structure of particle physics events requires the simulation to be split up into several simulation stages as illustrated by figure 6. Even if the processes involve electromagnetic and weak interactions, in case of hadron collisions the most challenging aspect of Monte Carlo simulation is the description of the QCD phenomenology as

[6]Note that probability density functions are often abbreviated as PDF, which is ambiguous to the abbreviation of parton distribution functions introduced earlier in this thesis and therefore avoided here.

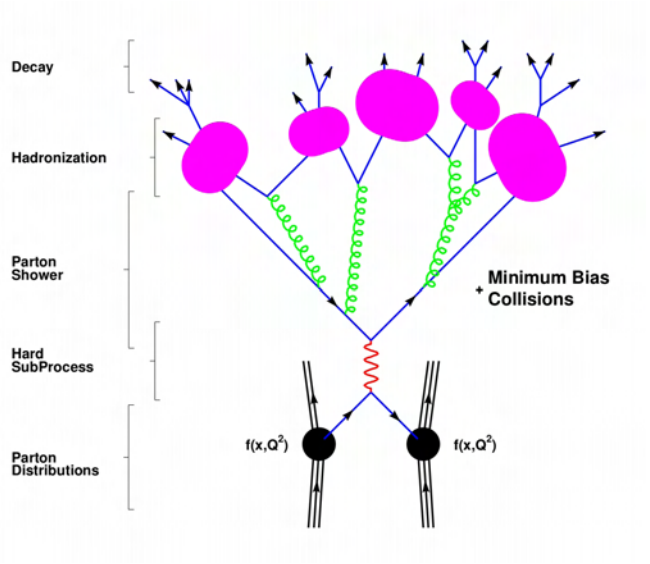



Figure 6: Basic structure of an event simulated by a showering and hadronization generator [14].

explained above. It starts with the hard interaction of the beam particles in the parton picture. The following branching of quarks and gluons is called parton showering and contains higher order effects. It temporarily leads to a number of elementary particles in the event which do not necessarily have a net non-zero colour charge. A phenomenological model called hadronization regroups the coloured partons into composite colour neutral hadrons as expected from confinement. These resulting particles can further decay into the final observable objects. Additionally, features of the underlying event and pile-up are considered because of the desired high collision rate at hadron colliders. All these steps will be briefly summarized in the following sections while detailed descriptions can be found in [15], [14] and [16]. Another stage of the event simulation is the response of the detector when fed with four momenta of the final state particles. This will be described after the introduction of the ATLAS detector in chapter 3.

At present, there are several simulation programs available utilizing the steps mentioned above. They are often referred to as general purpose Monte Carlo generators. While the exact implementation differs of course, their organisation is similar. The programs used throughout this thesis are called Pythia [17], Herwig [18], MC@NLO [19] and Powheg [20]. The first two are leading order generators including only the tree level Feynman graphs, whereas the others involve next-to-leading order QCD corrections. They will get revisited in section 5.4 where the simulation of the background and signal processes used for this thesis are described. But for now, the focus lies on their common features and the key steps of the simulation.

### Hard Interaction

The hard scattering process is the starting point of event generation. At parton level, the momentum of the colliding constituents is selected by sampling the hadron PDFs at the energy scale of the hard interaction. Once this is done, the convolution with the differential cross section of the hard subprocess can be calculated. It is proportional to the probability of the occurrence of the event. The average over many of these candidate event weights approximates the integral over the phase space, converging to the production cross section as noted in equation 2.42.

Typically the events are unweighted to match the distribution of their theoretical prediction by the so called hit-and-miss technique. For each candidate event, a random number is compared to the ratio of the event weight over the maximum event weight. If the ratio exceeds the random number the event is accepted or otherwise gets rejected, so that the accepted events occur at the expected frequency with the theoretically predicted distributions.

While most of the currently available parton level Monte Carlo generators compute only the tree level matrix elements, some include so called multi-leg processes with multiple final state partons. Only very few actually describe higher orders in perturbation theory.

### Parton Showering

In the hard subprocess the partons get strongly accelerated due to the large momentum transfer. The colour-charged partons emit radiation in form of gluons, just like electrical charges do with photons. Moreover, these gluons emit further radiation, because they are colour-charged themselves. So the information about the colour flow from the hard subprocess needs to be handed over to the parton shower algorithms. The showers represent higher order effects in perturbation theory. Evolving the event is an iterative process involving the so-called Sudakov form factors closely related to the DGLAP splitting functions as presented in equation 2.44. The Sudakov form factors describe the probability for a parton to evolve from a hard scale to a softer scale without splitting and specify the range where the branching such as $g \to q\bar{q}$, $g \to qg$ or $g \to gg$ is actually resolvable. An example of these branching processes is shown in figure 7.

The simulation of final state radiation (FSR) starts the branching from the energy scale of the hard interaction. To include also initial state radiation (ISR) processes, the algorithms are evolving backwards in time, whereas it needs to be ensured that the energy distribution of the incoming partons is compatible with the PDFs at the corresponding scale. But these iterative methods are only justified for collinear parton splitting and soft gluon radiation where the final outcome of the successive branching leads to a number of particles moving roughly in the same direction as the initial partons. The scale of the transverse momenta after showering is defined by a cutoff, typically

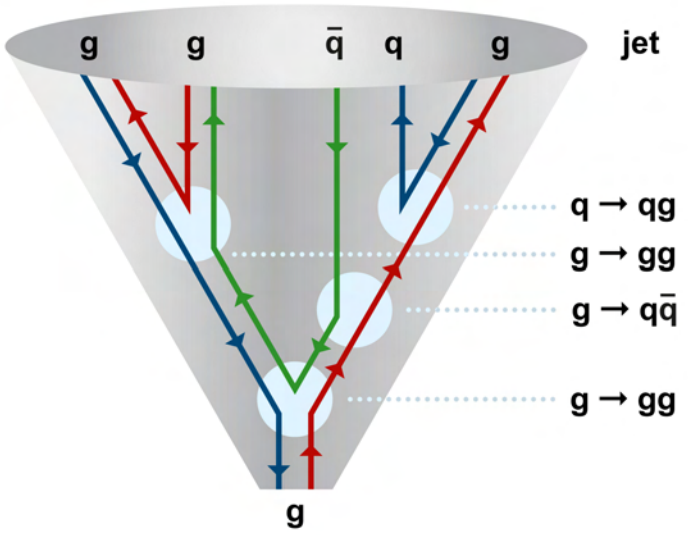



Figure 7: Sketch of parton shower evolution forming a hadronic jet.

chosen to be 1 GeV. Advanced methods to further stretch the algorithms to hard and wide-angled emissions are explained in [15]. Programs such as PYTHIA or HERWIG for example contain sophisticated parton shower algorithms. A number of other Monte Carlo generators do not come with their own implementations but interface those of the two mentioned before.

### Hadronization

While the simulation steps considered before work at parton level, only colour-neutral hadrons have been observed experimentally. The phase where the outgoing partons get confined into hadrons is called hadronization. Because perturbation theory is not applicable at long distances, this part cannot be described from first principles. Phenomenological models are used to approximate such processes under the assumption that the hadronization scale is much smaller than the scale of the hard interaction. Therefore, hadronization is independent of perturbative processes and does not change the original parton information.

Widely used are the so called (Lund) string model and the cluster model as shown in figure 8. Based on observations from lattice QCD that at large distances the potential energy of colour charges increases linearly with their separation, the string model uses string dynamics to describe the colour flux between quarks. A gluonic string is stretched between them until the potential energy reaches the order of hadron masses. Then the string breaks creating a new quark-antiquark pair leading to two shorter strings of colour singlet states.

The cluster model rests on the so called pre-confinement property of the branching process. Partons in a shower are clustered in colourless groups. Therefore gluons remaining from the shower algorithms are forced to split into quark-antiquark pairs which are regrouped with other quarks in colour singlet clusters, typically decaying into two hadrons. To improve these models and optimize the agreement, their parameters are tuned to data from former experiments.

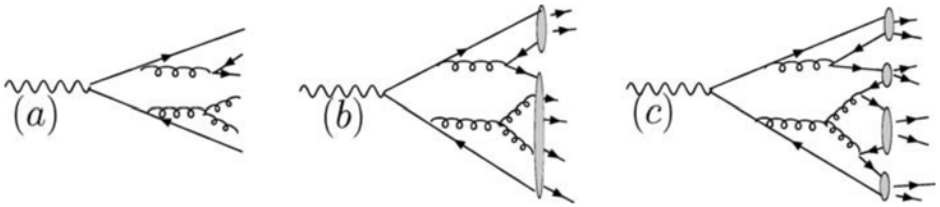



Figure 8: Possible radiation pattern from a $q\bar{q}$ pair (a) and illustration of string fragmentation (b) and cluster hadronization (c) models [16].

### Underlying Event and Pileup

The preceding steps describe the final state of the hard interaction sufficiently by extracting the high energy partons from the incoming hadrons. But one must not forget about the hadron remnants that do also evolve, hadronize and interact with each other as described in figure 9. Since they are colour-connected with the partons involved in the hard subprocess but by definition can only contribute low momentum transfer, such additional activity can not be described by perturbative QCD. Phenomenological models need to account for these underlying events and have to be tuned to the experimental circumstances.

Moreover, multiple interactions do occur as collider beams never only contain single hadrons (see section 3.1.2) but rather bunches of billions of protons per beam, like in the case of the Large Hadron Collider. So there is a background of additional mainly soft inelastic hadronic interactions that is called minimum-bias events. They also contain so called pileup contributions, as all those multiple interactions literally pile up in the detector. If this happens simultaneously to the hard process, the additional interactions are categorized as in-time pileup. In case the detector response is larger than the bunch crossing time, remnants from previous collisions can overlay with the hard process leading to so called out-of-time pileup. These contributions are measured in data and fed to the underlying event models.

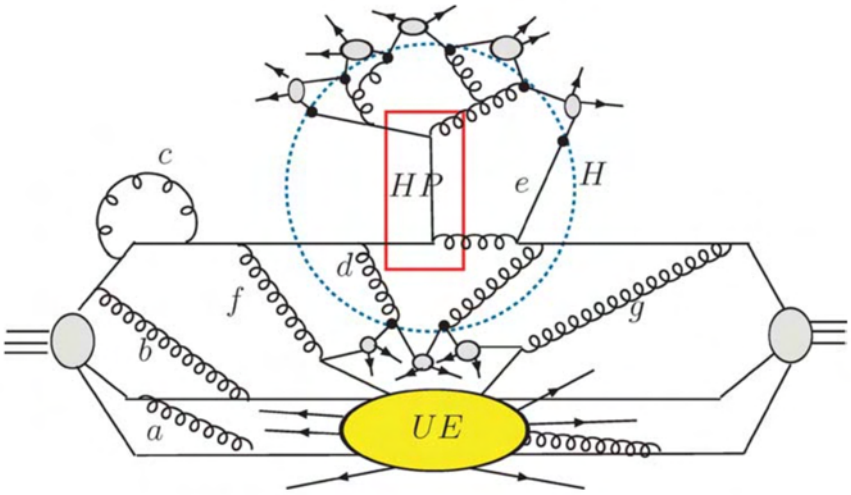



Figure 9: General structure of a hard proton-proton collision [16]. The hard process (HP) is shown with the factorization scale (H) and the underlying event (UE).

## 2.4 SM Higgs Boson at LHC

The techniques described above have been used to predict the cross sections of several important Standard Model processes at hadron colliders. Figure 10 shows an overview of the benchmark processes at the LHC and the Tevatron accelerator. The Tevatron accelerator is an proton-antiproton collider at the Fermi National Accelerator Laboratory in Chicago (USA) running until late 2011 with a center-of-mass energy of $\sqrt{s} = 1.96$ TeV. The comparison to the LHC (see section 3.1) illustrates the transition to even higher energies. Evidentially, the total inelastic proton-(anti)proton cross section is about six orders of magnitude larger compared to the production cross section of W or Z bosons and more than ten orders compared to Higgs production (depending strongly on the mass of the Higgs boson). High luminosities are required to produce such rare processes at sufficiently high rates. The obvious increase of the Higgs production cross section with the center-of-mass energy was one driving argument in the conception of the LHC, as the discovery of the Higgs boson and the search for new physics phenomena was declared the main goal of the project.

### 2.4.1 Constraints on the SM Higgs Boson

As already stated in section 2.2.3, the mass of the Standard Model Higgs boson is not fixed by theory because the self-interaction coupling constant $\lambda$ can take any positive and real value. However, there are several theoretical considerations constraining the value of $m_H$. Additionally, experimental limits have been deduced from both direct searches as well as indirect measurements. These topics are covered in [21], [22] and [23] and now get summarized in the following passages.

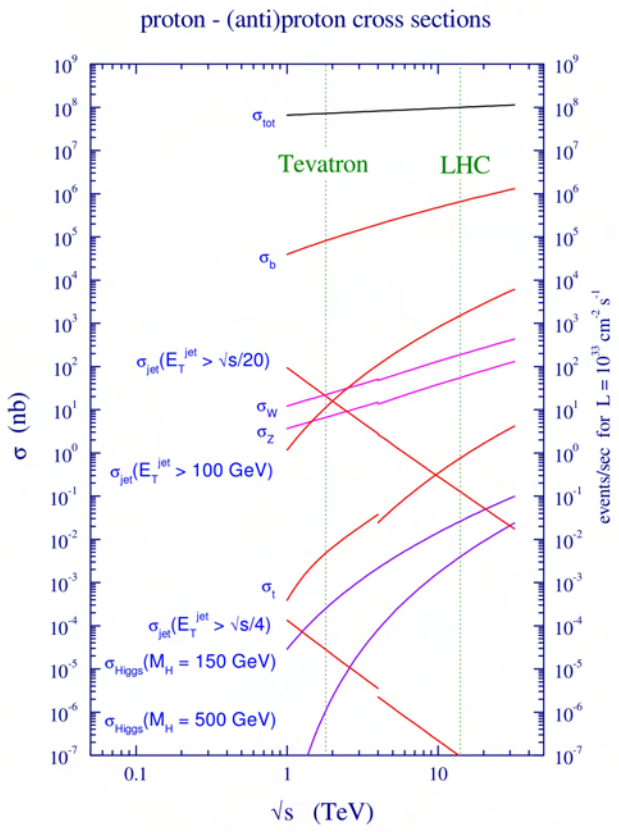



Figure 10: Standard Model cross sections at the Tevatron and LHC colliders as a function of the center-of-mass energy [10].

An upper bound on the mass of the Higgs boson can be derived from the so called perturbative unitarity preservation. Considering the amplitudes for elastic scattering of W-bosons with longitudinal polarizations at high energies, unitarity requires finite values for the cross section. The inclusion of Higgs boson contributions to the scattering process leads to the unitarity condition $m_H \leq 870$ GeV. Such large values for the mass imply also a large value for the coupling and the perturbative approach needs to be reconsidered as it becomes unreliable. So under the assumption that the Standard Model remains perturbative even for Higgs masses above the boundary, new phenomena would have to appear at the TeV scale to prevent unitarity violation.

Even tighter bounds on $m_H$ can be extracted from the examination of the Higgs boson quartic self-coupling (see equation 2.36). The variation of the coupling with the energy scale is given by the renormalization group equation. It varies logarithmically with $Q^2$. So for very small energies, the coupling vanishes and the theory is said to be trivial since it is non-interacting. In the opposite limit, way above the electroweak scale, the quartic coupling becomes infinte at the so called Landau pole. But this again means, that the coupling (and thus the Higgs mass) has to be zero for the theory to remain perturbative at all scales. To avoid the Landau pole, one can establish an energy cut-off

$\Lambda_C$ below which the coupling $\lambda$ remains finite and interpret it as the point up to which the Standard Model is valid. This upper boundary is therefore often called triviality bound. If this cut-off scale is large the Higgs boson has to be rather light, while low values of $\Lambda_C$ allow the Higgs boson mass to grow up to about 1 TeV. This behaviour is illustrated in the upper band in figure 11.

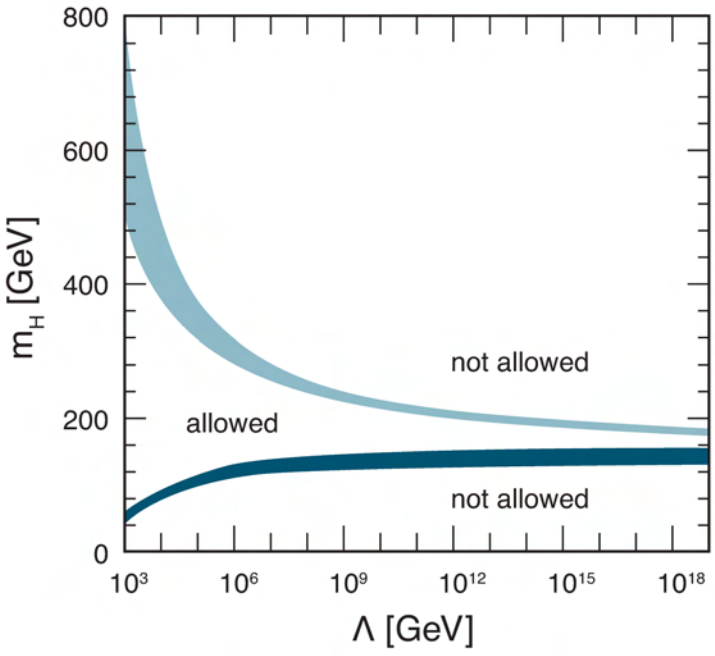



Figure 11: Theoretical boundaries on the Higgs mass [22]. The upper bound (triviality) and the lower bound (vacuum stability) on the Higgs mass as a function of the cut-off or new physics scale $\Lambda$ for $m_t = 175$ GeV and $\alpha_S = 0.118$. The allowed region lies in between the bands which represent the theoretical uncertainties.

Additionally, there is another class of theoretical boundaries which arises from the behaviour of the Higgs self-coupling under renormalization, the so called stability bound. It is shown as the lower band in figure 11. In the regime where $\lambda$ is small, one needs to include also the contributions from fermions and gauge bosons into the running of the quartic coupling. Their impact depends of course on the mass and therefore the top quark contribution has the most sizeable effect. It becomes dominant when the coupling gets too small and could finally drive $\lambda$ to negative values, leading to an effective potential that is not bound from below anymore. Thus, without a minimum, the vacuum becomes unstable. To conserve vacuum stability, again a cut-off scale can be introduced which keeps the coupling positive and therefore constrains the mass of the Higgs boson.

The experimental limits on the mass of the Standard Model Higgs boson in the past were based on the abstinence of direct observations. A lower limit has been set by the studies of electron-positron collisions at LEP[7] searching for Higgs bosons produced in

[7]Large Electron Positron collider

association with a Z boson and decaying into a heavy fermion-antifermion pair. Up to a value of $m_H \geq 114.4$ GeV, the Higgs boson has been experimentally excluded as documented in [24]. Later, the mass range could be constrained even further. The searches in proton-antiproton collisions at the Tevatron did elude direct observations as well, leading to the exclusion of the mass range 156 GeV $\leq m_H \leq$ 177 GeV as summarized in [25].

An indirect way of probing the Higgs mass in the Standard Model is provided by precision measurements of laboratories like SLC[8] or the aforementioned LEP and Tevatron. The value of $m_H$ affects other SM observables through radiative corrections. Although the impact is rather small, the great accuracy achieved by the experiments allows to obtain a preferred Higgs boson mass from a global fit of the electroweak observables. The so called blue-band plot shown in figure 12 illustrates the $\Delta\chi^2$ of this fit as a function of the Higgs mass constraining the preferred value to $m_H = 94\ ^{+29}_{-24}$ GeV as performed in [26] and updated in [27]. The exclusion from direct searches is indicated in yellow. So both, experimental and theoretical considerations prefer a rather light Standard Model Higgs boson.

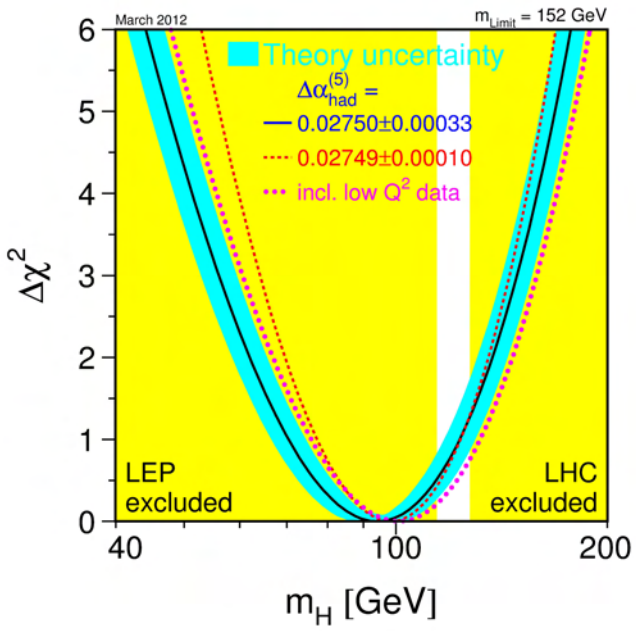
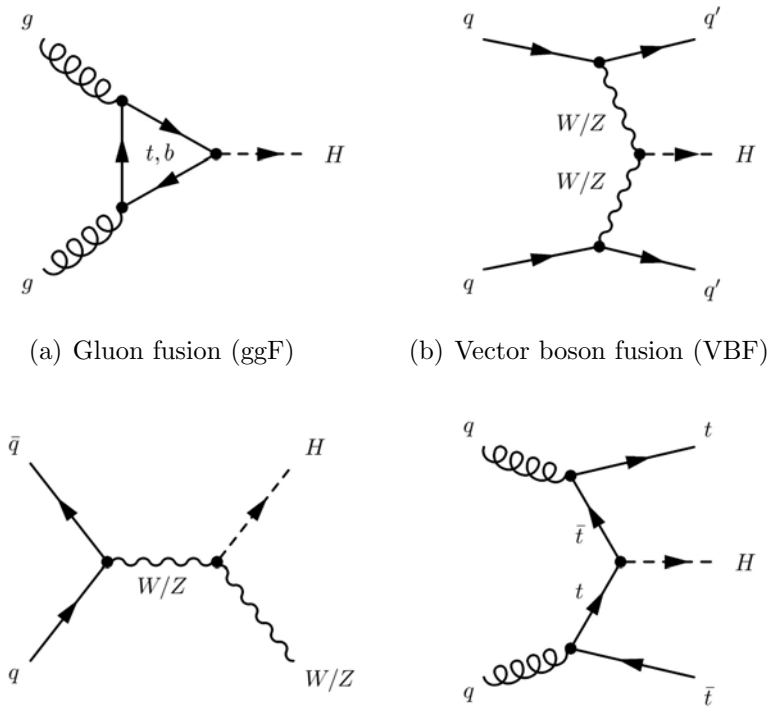



Figure 12: Experimental constraints on the mass of the Higgs boson [27]. $\Delta\chi^2$ vs. $m_H$ curve. The line is the result of the fit using all high-$Q^2$ data, the band represents an estimate of the theoretical error due to missing higher order corrections. The vertical band shows the 95% CL exclusion limit on $m_H$ from the direct searches at LEP and the LHC.

[8]Stanford Linear Collider

### 2.4.2 Higgs Boson Production at the LHC

The proton-proton collisions at LHC offer four relevant production processes for the Standard Model Higgs boson. The leading modes are the gluon fusion ($gg \to H$) process and the vector boson fusion ($qq \to q'q'H$). Furthermore, there are the associated productions with a vector boson ($qq \to W/ZH$), often called Higgs-Strahlung, or with top quarks ($gg \to t\bar{t}H$). Figure 13 shows the leading order Feynman graphs of these production processes.




(a) Gluon fusion (ggF)

(b) Vector boson fusion (VBF)

(c) Higgs-Strahlung (WH/ZH)

(d) Associated production (ttH)

Figure 13: Feynman graphs of different Higgs boson production processes. (a) Gluon fusion, (b) vector boson fusion, (c) Higgs-Strahlung and (d) ttH production.

The gluon fusion production (ggF) via an intermediate quark loop dominates with the highest cross section for Higgs boson masses up to 1 TeV. Its experimental measurement is the main goal of this thesis. Since the coupling is proportional to the mass of the quarks, the top and bottom quark contributions are the most relevant ones. The NLO QCD corrections to the cross section have been calculated in [28] and are rather large at the order of 80 – 100%. To include even higher order corrections, it is possible to reduce the complexity of the calculations by working in the infinite top mass limit and thereby shrinking the process to an effective tree-level vertex. Then the NNLO corrections are found to provide an additional increase of the cross section of $\sim 25\%$ as conducted in [29].

Additionally, soft gluon contributions have been resummed up to NNLL[9] in [30]. The size of two-loop electroweak corrections has been determined in [31], where the minor impact is depending on the Higgs boson mass.

The final result is shown in figure 14 for a center of mass energy of $\sqrt{s} = 8$ TeV, where the bands indicate the uncertainties arising from the PDFs and the choice of the factorization and renormalization scales. As a subdominant production process, the vector boson fusion (VBF) cross section is already about one order of magnitude smaller than the gluon fusion cross section. The associated modes are negligible for the following analysis. All these examinations are summed up in [32] and the corresponding online resource [33].

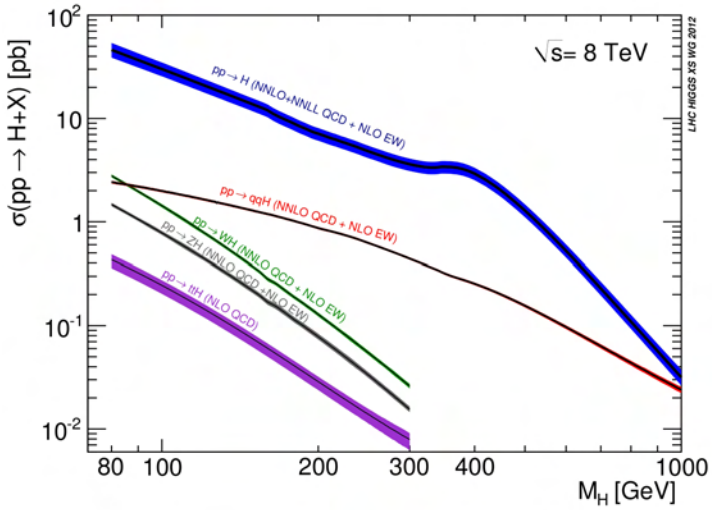



Figure 14: Standard Model Higgs boson production cross sections at $\sqrt{s} = 8$ TeV [33].

### 2.4.3 Higgs Boson Decay Modes and Search Channels

Once produced in a proton-proton collision, the Higgs boson predominantly decays immediately into the heaviest particles that are kinematically available. So the search channels at the LHC are focusing on the different decay modes as they lead to different signatures in the final state. For a fixed Higgs boson mass, the couplings to fermions and gauge bosons are given by the Standard Model. So the cross section, the decay width and the branching ratios are fully determined as well. The branching ratios are defined as ratio of the decay width of the process in demand and the sum of all possible decays as in

$$BR(H \to XX) = \frac{\Gamma(H \to XX)}{\sum_i \Gamma(H \to X_i X_i)} \tag{2.48}$$

[9] next-to-next-leading-logarithm

They are shown in figure 15 for the range between 80 and 200 GeV. The dominant modes are the $H \to b\bar{b}$ decay and the decay into two W-bosons $H \to W^+W^-$. At the threshold of $m_H = 2m_W$ the branching fraction into a W-pair is close to 100%.

Below that threshold, at least one W-boson needs to be virtual and the hadronic decay channels dominate the lower region. The $H \to t\bar{t}$ becomes sizeable for very large Higgs boson masses, opening up at 350 GeV. But these search channels suffer from the overwhelming QCD background produced in the hadronic collisions at the LHC. They become feasible for production modes like VBF and associated production as they provide further typical characteristics, but are of course afflicted by the low production rates.

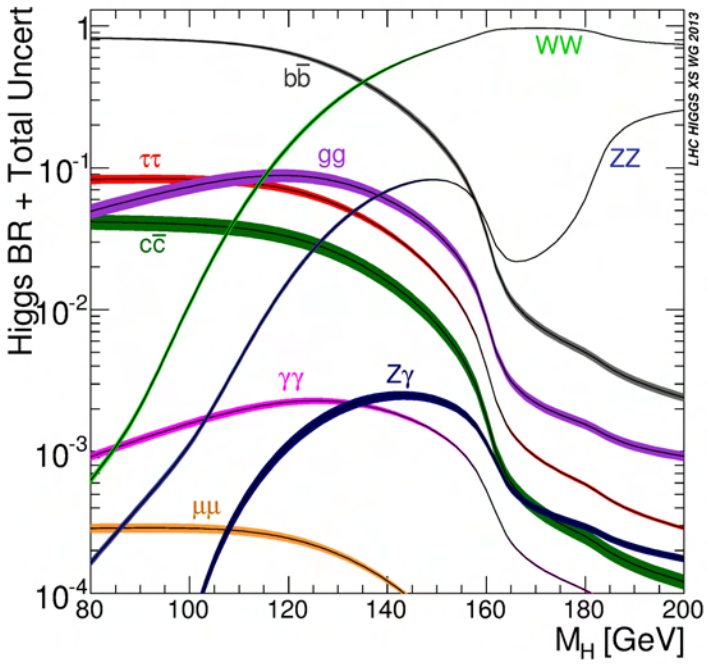



Figure 15: Standard Model Higgs boson decay branching ratios [33].

More promising are indeed the channels $H \to ZZ$ and $H \to \gamma\gamma$. They are of particular interest because the Higgs boson mass can be fully reconstructed from the final state leptons and photons, which are beneficial in the hadronic environment. Unfortunately the branching ratio to two photons is very small as the photons are massless and so it is only relevant for low Higgs masses. The low branching ratio of the further decaying Z-pair $ZZ \to \ell^+\ell^-\ell^+\ell^-$ on the other hand reduces the expected rate significantly.

So the search channel $H \to WW \to \ell^+\nu\ell^-\nu$ provides the highest sensitivity across a wide mass range. The decay products of the two W-bosons lead to a particular signature with two oppositely charged isolated leptons and missing transverse momentum due to the two neutrinos, which do not interact with the detector materials. Therefore it is predestined for the measurement of the cross section and thus, the decay channel investigated in

this thesis. So the following chapters will first introduce the experimental setup for the analysis and then guide to the characteristics of the WW-channel afterwards.

### 2.4.4 Status of Observations and Predictions

As the framework of the Standard Model has been presented in previous sections, it is now time to compare its predictions to the experimental observations. This part focuses on a number of analyses of benchmark processes performed with the ATLAS detector to illustrate the incredible accuracy, that this powerful instrument already delivered at the early stages of data taking with the dedication of an ambitious collaboration. A summary of the most interesting processes is shown in figure 16.

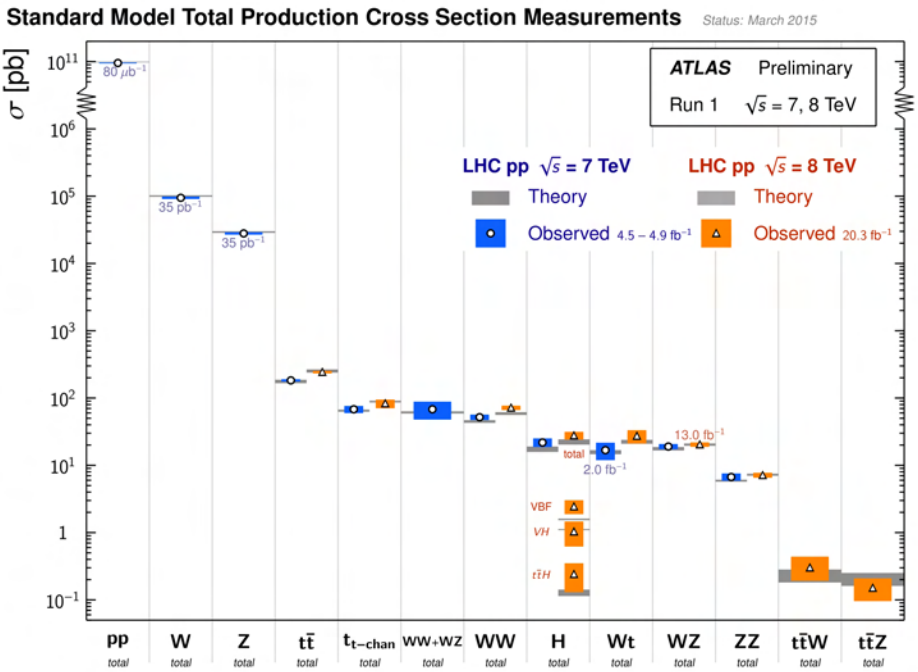



Figure 16: Summary of several Standard Model total production cross section measurements [34], corrected for leptonic branching fractions, compared to the corresponding theoretical expectations. All theoretical expectations were calculated at NLO or higher. Uncertainties for the theoretical predictions are quoted from the original ATLAS papers.

With the 35 pb$^{-1}$ of collected data in 2010 for example, the cross section of the inclusive Drell-Yan processes $Z \to \ell^+\ell^-$ and $W \to \ell\nu$ ($\ell = e, \mu$) have been measured in the proton-proton collisions at a center of mass energy of $\sqrt{s} = 7$ TeV. The total integrated cross sections for $W$ and $Z/\gamma*$ are obtained by integrating over a fiducial kinematic region and extrapolating it to the full range. Compared to the predictions in NNLO perturbation theory, the experimental results show a good agreement at high precision.

The $t\bar{t}$ production cross section has been measured for two different experimental setups. A dataset of 1.1 fb$^{-1}$ collected in 2011 at $\sqrt{s} = 7$ TeV was analysed as well as 20.3 fb$^{-1}$ in 2012 at $\sqrt{s} = 8$ TeV. Both results show reasonably good agreement with the NNLO predictions, underlining the success even for the hadronic processes.

The WW production cross section is of particular interest for this thesis, as it is the dominant background source in the analysis of the $H \to WW$ decay channel. It has been measured with the full 2011 dataset of 4.6 fb$^{-1}$ integrated luminosity at $\sqrt{s} = 7$ TeV collected by the ATLAS detector. Compared to the NLO prediction, the result corresponds very well with the theory of the Standard Model.

### Higgs Observation

In July 2012, the ATLAS and CMS collaboration simultaneously published sensational results: The discovery of a new Higgs-like boson at statistically significant level ([35] and [36]). The driving analyses were the search channels $H \to ZZ \to 4\ell$ and $H \to \gamma\gamma$ with the best mass resolution, supported by the $H \to WW \to \ell\nu\ell\nu$ channel for each of the two competing experiments. Figure 17 shows the local probability of the outcome of the ATLAS analyses to be accidental versus the Higgs boson mass. When this distribution falls below the 5 $\sigma$ level, experimentalists speak of an actual discovery because the probability to observe the result by chance are below one in three million.

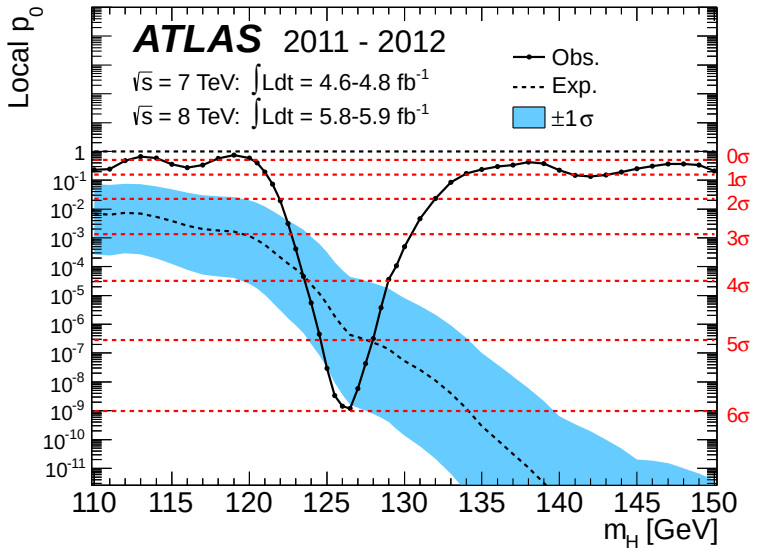



Figure 17: The observed (solid) local $p_0$ as a function of $m_H$ in the low mass range [35]. The dashed curve shows the expected local $p_0$ under the hypothesis of a SM Higgs boson signal at that mass with its plus/minus one sigma band. The horizontal dashed lines indicate the p-values corresponding to significances of 1 to 6 standard deviations.

In the meantime, dedicated analyses further investigated the properties of the new particle. So far, every measurement is in good agreement with the SM Higgs boson. It is favoured to be a $J^P = 0^+$ particle, although other possibilities have not yet been excluded completely. Now that the full dataset of about 25 fb$^{-1}$ has been considered, the combination of the ATLAS and CMS search channels results in a mass measurement of $m_H = 125.09 \pm 0.21(\text{stat.}) \pm 0.11(\text{syst.})$ GeV [37]. Values below 123 GeV and between 127 – 710 GeV could be excluded at a 95% confidence level. These observations serve as the starting point of the following thesis.

Moreover, with the discovery of the Higgs boson, its mass can be included into the global fit to the electroweak precision data described above. With all fundamental parameters fixed, it allows to actually overconstrain the Standard Model at the electroweak scale and therefore improves the other predictions as well, setting new benchmarks for the direct measurements. One example performed by the GFitter group [38] is given in figure 18. It shows the impressive consistency of the SM by comparing the masses of the Higgs boson, the W-boson and the top quark.

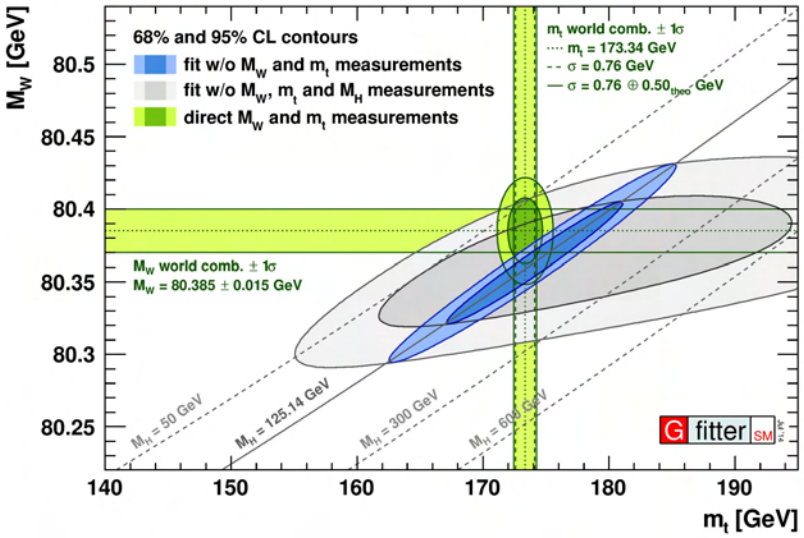



Figure 18: Consistency test of the Higgs sector of the Standard Model [38]. The blue (grey) areas illustrate the fit results when including (excluding) the $M_H$ measurements. The direct measurements of $M_W$ and $m_t$ are always excluded in the fit. The vertical and horizontal bands (green) indicate the 1 $\sigma$ regions of the direct measurements.

# 3

# Accelerator and Detector

The first part of this section will give an introduction into the basic concepts of hadron colliders and their application at the LHC following the prescriptions in [39] and [40]. The treatment of the particle beams is presented along their way from the source to the interaction points where the detectors are located. Closing this part is the overview of the running conditions until 2012 and the corresponding dataset delivered. Therefore the second part of this section is dedicated to the ATLAS detector collecting the data used for the final analysis of this thesis. It makes use of a variety of different detection methods and materials to collect as much information as possible about the particles passing through. A breakdown of its main components will explain the experimental setup.

## 3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a collider ring located at CERN[10] near Geneva. Its 27 km long tunnel is placed about 100 m underground and spans beneath the swiss-french border. Formerly accommodating the Large Electron Positron Collider (LEP) from 1989 until 2000, the site was extensively rebuilt until the LHC started operations in September 2008. The ring system consists of eight sectors of 2.4 km long arcs divided by eight straight sections of about 530 m length for the interaction points as sketched in figure 19.

At four of these interaction points (IP) the main experiments are located. The two high luminosity experiments ATLAS and CMS focusing on Higgs physics and physics beyond the SM are located at IP1 and IP5 which are diametrically opposite straight sections. LHCb investigates the decay of b-hadrons and the occurrent CP violation at

[10] Conseil Européen pour la Recherche Nucléaire

low luminosities while the ALICE experiment is designed for heavy ion collisions. This is another operations mode of the LHC system accelerating and colliding lead ions to create conditions as they have occurred shortly after the big bang and to investigate the so called quark-gluon plasma. The latter two octants also host the beam injection systems whereas the remaining four octants do not contain beam crossings. They are used for beam cleaning and dumping. Point four contains two independent radio frequency (RF) cavities, one for each beam.

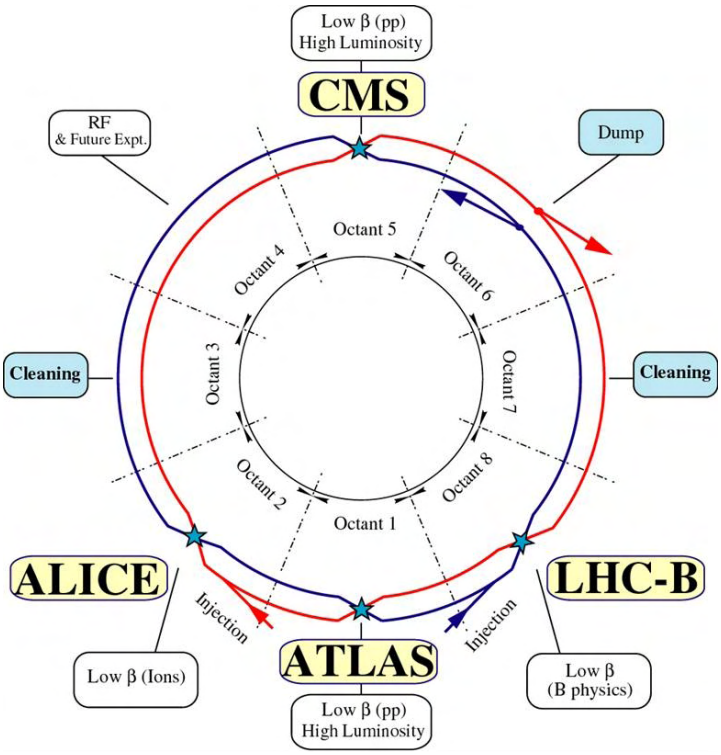



Figure 19: LHC scheme with the different interaction points [41].

To accelerate two proton beams in opposite directions, two separate beam pipes with antipodal magnetic dipole fields share one tube hosting the vacuum vessel and the iron yoke for the superconducting magnets due to the limited space in the tunnel diameter of 3.7 m. This installation is shown in figure 20.

### 3.1.1 Proton Acceleration Chain

Actually the LHC is the last step in a chain of smaller pre-accelerators, each increasing the proton energy typically by a factor 10-20. The full scheme with the various stages and their locations is shown in figure 21.

It all starts with a bottle of hydrogen from where the protons take up their journey to the final collision. The proton extraction is realised by a so called duoplasmatron placed in a Faraday cage because of the static high-voltage it utilises. It works via gas

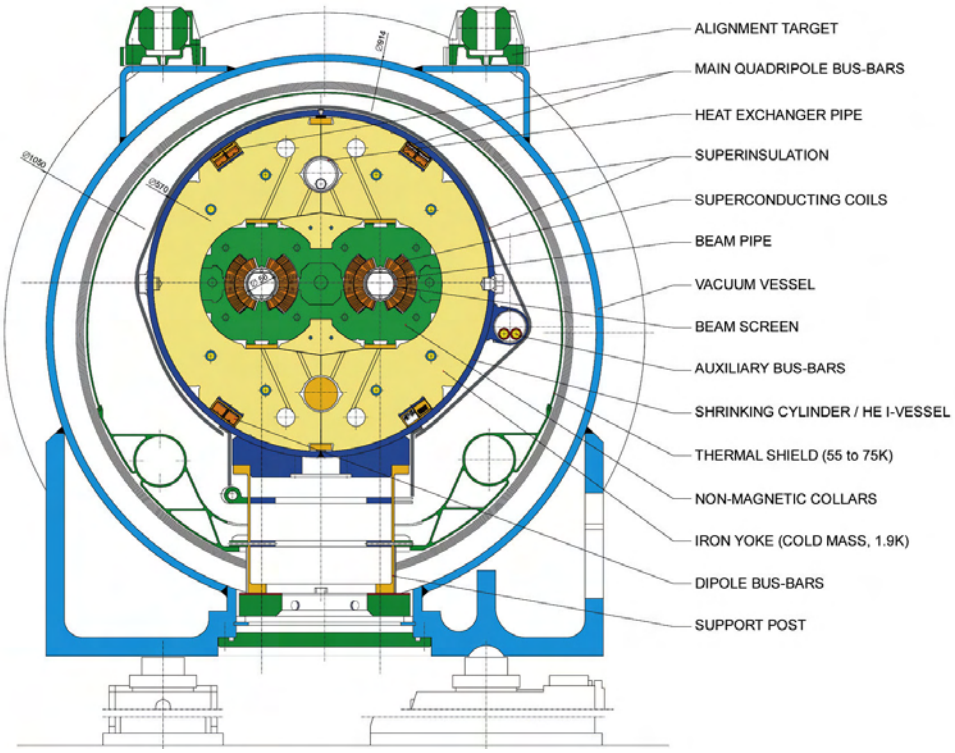



Figure 20: Cross section of an LHC dipole [41].

discharge of the hydrogen and supplies the subsequent acceleration stages with protons of the energy of 90 keV.

Then follows an acceleration via an RF quadrupole to 750 keV, feeding 6 bunches to the LINAC2, which is an 30 m long linear drift tube accelerator increasing the energy to 50 MeV. The next step is the so called Proton Synchrotron Booster (PBS), an RF cavity synchrotron literally boosting the protons up to 1.4 GeV. It consists of four beam pipes of 157 m length to increase the number of simultaneously accelerated particles. Those are injected into the 628 m long Proton Synchrotron (PS) not only to increase their energy to 25 GeV but also to split up the 6 proton bunches following a special scheme. This ends up in 72 consecutive bunches always leaving a 320 ns gap in the train structure for the rise time of the ejection kicker magnet.

The last stage before the final LHC ring is the Super Proton Synchrotron (SPS). This is a 7 km long circular accelerator lying about 50 m underground which can be filled with either two, three or at maximum four PS batches of 72 bunches whereas their structure is not further modified. With an energy of 450 GeV, the protons are then injected into the LHC ring close to IP2 for beam 1 (clockwise) and IP8 for beam 2 (counter clockwise).

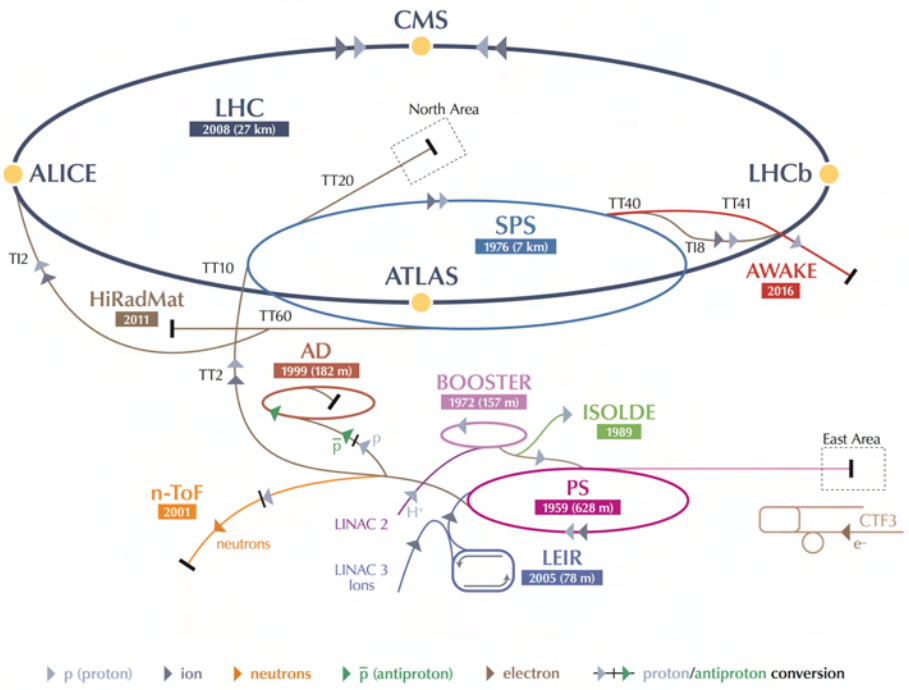



Figure 21: Large Hadron Collider Scheme including all pre-accelerators [42].

### 3.1.2 Beam Parameters

As in every synchrotron accelerator, the particles travelling along the ring get accelerated by a high amplitude alternating electric field which applies at each passage. This is achieved by RF cavities generating a longitudinal oscillating voltage along the beam axis with the frequency $f_{RF}$. The beams are focused and bent by transverse magnetic fields which keep the particles on the designed path, called the reference orbit or the equilibrium orbit. In order to create a constant force, the particle needs to cross the electric field at a constant phase at every turn. Such a synchronous particle would always stay at the centered orbit. This limits the frequency of the RF cavities to a constant integer of the revolution frequency of a synchronous particle to achieve a stable acceleration:

$$f_{RF} = i \cdot f_{rev} \tag{3.1}$$

However in reality there occur small disturbances which cause the particles to deviate from the reference orbit and therefore cause interference with the acceleration scheme. These small energy offsets lead to a longitudinal oscillation around a synchronous particle. Particles that are travelling slightly faster (slower) than the synchronous particle will experience less (more) deflection by the bending magnets and describe a longer (shorter) orbit. Therefore they will arrive later (earlier) than expected at the cavity on

their next turn and the incident phase shift leads to more (less) acceleration stabilizing the beam. This effect is called phase focusing and is limited to phase deviations below $\pi$ as otherwise the particles would receive acceleration that moves them away from the synchronous particle.

The boundary between stable and unstable conditions defines a so called RF bucket spanned by the frequency of the alternating field. A collection of particles sharing the same bucket is defined as bunch. So the RF frequency forms a chain of buckets to be filled with a bunch of particles. The longitudinal shape of these bunches is thus configured by $f_{RF}$ while the number of buckets in the accelerator is defined by the ratio $f_{RF}/f_{rev}$.

A similar transverse particle motion arises while bending and focusing the beams with transverse magnetic fields. Usually guided along the foreseen orbit with dipole magnets, the beam size is controlled with quadrupole magnets. But depending on the polarity of the quadrupole, it has a focusing effect in one plane but a defocusing effect on the plane orthogonal to it. Particles travelling on the reference orbit which are passing through the center of the quadrupole field do not experience any force, whereas the bending of the trajectory rises linearly with the distance from the center. A net focusing effect can be achieved by arranging several alternating elements as long as the drift space between the quadrupoles is small compared to the focal length of the magnets[11]. So any beam extend in the transverse plane causes the particles afar from the center to perform oscillations in the horizontal and the vertical plane around the reference orbit, the so called betatron motion. It is described by the Hill equation:

$$\frac{d^2x}{ds^2} + K_x(s)x = 0 \tag{3.2}$$

here for simplicity shown for only one transverse dimension $x$. It resembles a harmonic oscillator equation with the magnet focusing function $K_x(s)$, which depends on the longitudinal particle position $s$ as explained above. The general solution can be written in the following form:

$$x(s) = \sqrt{\beta_x(s)\epsilon} \cdot cos[\phi_x(s) + \phi_0] \tag{3.3}$$

The betatron function $\beta_x(s)$ is proportional to the local amplitude of the oscillation and particularly of interest at the interaction point, then labelled as $\beta^*$ and usually given in units of meter. It is defined by the arrangement of the magnets as is the phase $\phi_x(s)$, both are important parameters for stable beam operations. The factor $\epsilon$ is called emittance and the area $\pi\epsilon$ defines the effective area of the beam in phase space by analysing the particle position $x$ versus its divergence $dx/ds$. While storing the beam in the synchrotron, the emittance area is conserved according to Liouville's theorem. But

[11] Such a design is often referred to as FODO lattice (F=focusing, O=drift through other instrumentation, D=defocusing)

accelerating the protons does violate this condition because the longitudinal momentum increases whereas the transverse momentum stays the same. The desired result is a shrinking beam often referred to as adiabatic damping. To state a factual conserved quantity, it is common to define the normalized emittance:

$$\epsilon_N = \gamma_r \beta_r \epsilon \tag{3.4}$$

with $\beta_r = v/c$ and $\gamma_r = 1/\sqrt{1-\beta_r}$. Under the assumption that the particles are Gaussian distributed in the horizontal and the vertical plane, the emittance can also be defined via the cross-sectional size of the particle bunch $\sigma_{x,y}$ and the relation:

$$\epsilon_{x,y} = \frac{\sigma_{x,y}^2}{\beta_{x,y}(s)} \tag{3.5}$$

Now that the main beam parameters have been introduced, it is time to discuss what happens when the two beams are finally brought to collision. It is of crucial importance to know the rate of particle collisions, the so called luminosity. In case of any given process $pp \to X$ producing a particle $X$ in a proton collision, the luminosity is given by the ratio of the rate of this process $R_{pp\to X}$ and its production cross section $\sigma_{pp\to X}$:

$$\mathscr{L}(t) = \frac{R_{pp\to X}}{\sigma_{pp\to X}} \tag{3.6}$$

It is also referred to as instantaneous luminosity given in $cm^{-2}s^{-1}$, the cross section typically in barn[12]. As the running conditions vary with time, it is common to perform the integral with respect to time $L = \int \mathscr{L}(t)dt$, then called integrated luminosity given in units of $b^{-1}$.

To understand how this quantity looks like for the case of bunched beams at the LHC, one can derive an expression related to the beam parameters introduced above as in [39]. Two beams passing through each other as shown in figure 22 can be described by the overlap integral of their (normalised) particle density functions $\rho_{1,2}$ depending on the longitudinal position $s$ of the bunches and their distance to the collision point written as $s_0 = c \cdot t$:

$$\mathscr{L} = n_1 n_2 f_{rev} K \int\int\int\int_{-\infty}^{+\infty} \rho_1(x,y,s,-s_0)\rho_2(x,y,s,s_0)dxdydsds_0 \tag{3.7}$$

Here, $n_{1,2}$ are the number of particles per bunch and $f_{rev}$ is the revolution frequency as described above. The kinematic factor $K$ is called Møller luminosity factor [43]:

[12] $1barn = 1b = 10^{-24}cm^{-2}$

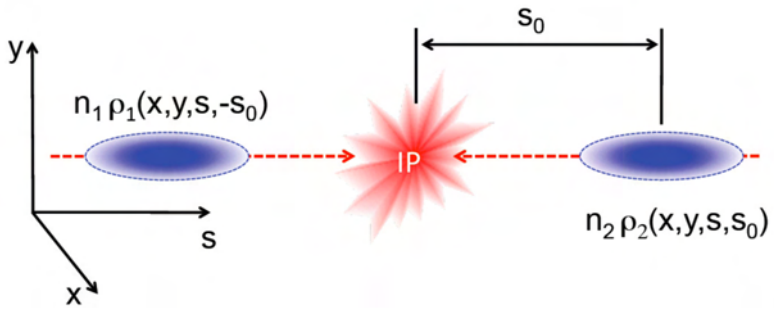



Figure 22: Schematic view of two bunches colliding at the IP [44].

$$K = \sqrt{(\vec{v}_1 - \vec{v}_2)^2 - \frac{(\vec{v}_1 \times \vec{v}_2)^2}{c^2}} \tag{3.8}$$

taking into account the velocities $\vec{v}_{1,2}$ of the two beams and converging to $K = 2c$ for $\vec{v}_1 = -\vec{v}_2$ in the relativistic approximation. If one assumes uncorrelated Gaussian density profiles for collinear beams colliding head-on of the form:

$$\rho_{iu}(u) = \frac{1}{\sqrt{2\pi}\ \sigma_{iu}}\ \exp\left(-\frac{u^2}{2\sigma_{iu}^2}\right) \tag{3.9}$$

where $i = 1, 2$ and $u = x, y$, the integral can be solved analytically and leads to the following expression:

$$\mathscr{L} = \frac{n_1 n_2 f_{rev}}{2\pi\ \sqrt{\sigma_{1x}^2 + \sigma_{2x}^2}\ \sqrt{\sigma_{1y}^2 + \sigma_{2y}^2}} \tag{3.10}$$

Recalling equation 3.5, this can also be expressed in terms of the beam parameters:

$$\mathscr{L} = \frac{n_1 n_2 f_{rev} n_b}{2\pi\sqrt{(\epsilon_{1x}\beta_{1x}^* + \epsilon_{2x}\beta_{2x}^*)(\epsilon_{1y}\beta_{1y}^* + \epsilon_{2y}\beta_{2y}^*)}} \tag{3.11}$$

with the number of colliding bunches $n_b$. So in order to achieve high luminosities it is necessary to collide many highly populated bunches at high frequency with low transverse beam sizes at the interaction points.

However, these ideal circumstances are of course not completely fulfilled in reality. To mention only one example, the LHC collisions are performed under a small crossing angle $\Theta_c$ to avoid parasitic collisions throughout the drift space before and after the experiment locations. In this case the longitudinal beam size must not be neglected

leading to a reduced luminosity. Detailed information on such effects can be found in [39] and [45].

### 3.1.3 Proton Collisions and Data Taking

In December 2009 the LHC produced first collisions with the injection energy of 450 GeV per beam. At the beginning of 2010 the energy was ramped up slowly to 3.5 TeV. From March until November 48.1 $\text{pb}^{-1}$ of integrated luminosity were delivered. In 2011 the energy was kept at 3.5 TeV but the beams were further squeezed at the IPs and the number of bunches circulating in the LHC was increased. That lead to an integrated luminosity of $5.5 \pm 0.1$ $\text{fb}^{-1}$.

The LHC pp run in 2012 started after a short technical stop at the end of March and lasted until late December. Its delivered integrated luminosity of 22.8 $\text{fb}^{-1}$ will serve as basis for the following analysis. During the ongoing long shutdown until 2015, preparations for higher beam energies towards the design values are made.

A list of beam parameters for the different periods is given in table 3. The evolution can be compared to the design values.

| Parameter | 2010 | 2011 | 2012 | Nominal |
|---|---|---|---|---|
| Beam Energy [TeV] | 3.5 | 3.5 | 4 | 7 |
| Bunch intensity [$10^{11}$p] | 1.1 | 1.49 | 1.6 | 1.15 |
| Number of bunches $n_b$ | 368 | 1380 | 1380 | 2808 |
| Bunch spacing [ns] | 50 | 50 | 50 | 25 |
| $\mathscr{L}$ at IP 1/5 [$10^{33}\text{cm}^{-2}\text{s}^{-1}$] | 0.2 | 3.6 | 7.7 | 10 |
| $\Theta_c$ at IP1/5 [$\mu$ rad] | 200 | 240 | 290 | 285 |
| $\beta^*$ [m] | 3.5 | 1.0 | 0.6 | 0.55 |
| $\epsilon_N$ [$\mu$m] | 2.0 | 2.6 | 2.5 | 3.75 |

Table 3: Peak performance of the LHC parameters for the different running periods 2010, 2011 and 2012 on the left, design values on the right.

Figure 23 shows integrated luminosity for the data taking in 2012. The dataset recorded by the ATLAS detector is slightly smaller than what the LHC delivered because of inefficiencies of the data aquisition and delays in the ramping of different detector components. Further data quality requirements reduce the final dataset which is used for physics analyses. Nevertheless the total efficiency is about 90%.

The different run conditions do affect the constitution of the collected data. For example the higher beam energy in 2012 in addition to the increased bunch intensity and lower $\beta^*$ lead to much larger instantaneous luminosities compared to 2011. As a result, the mean number of interactions per bunch crossing went up by a factor of 2. This is also shown in figure 23.

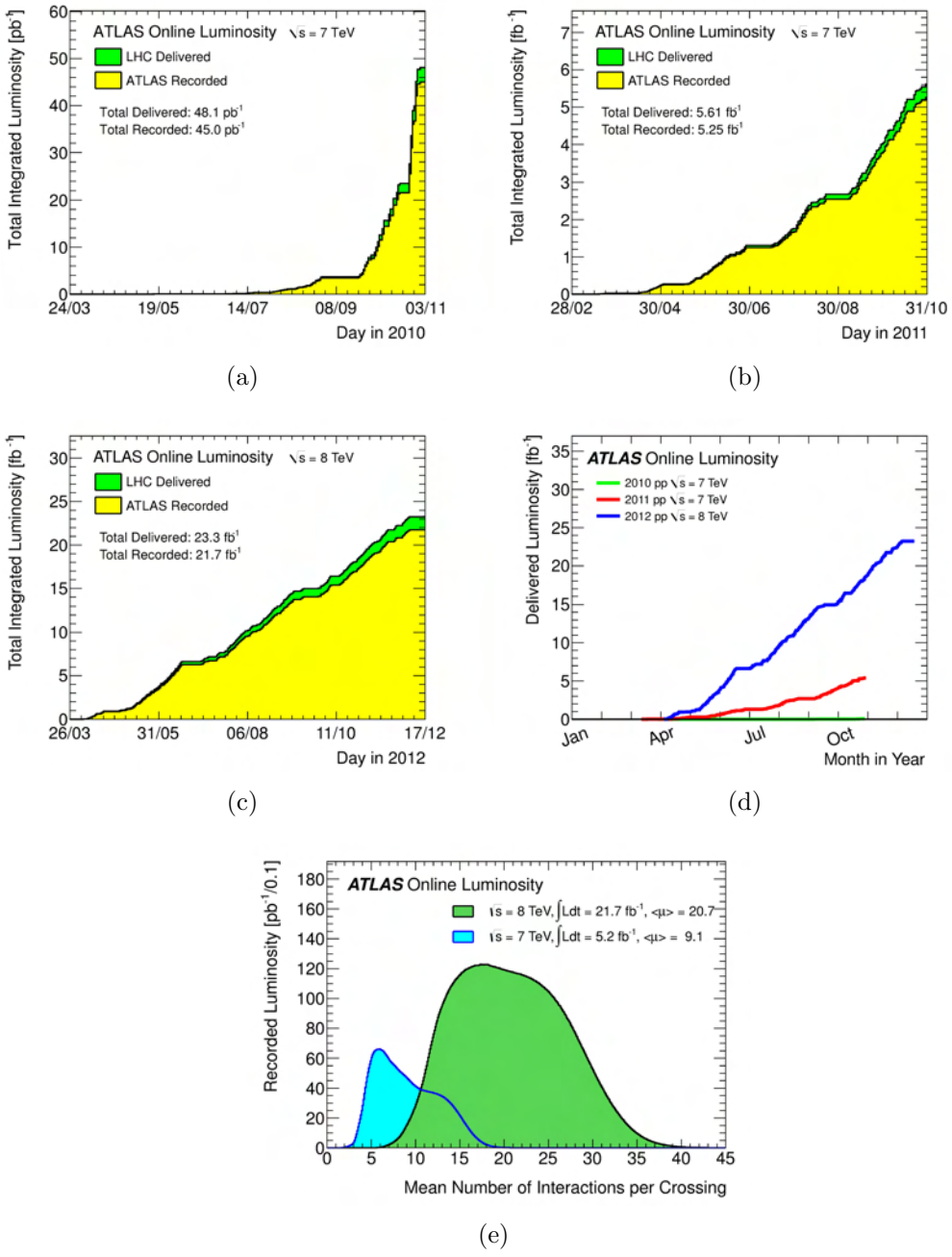



Figure 23: Cumulative luminosity versus day delivered to (green) and recorded by (yellow) ATLAS during stable beams for pp collisions for the different run periods (a) 2010, (b) 2011, (c) 2012 and (d) with respect to each other. Figure (e) shows the luminosity-weighted distribution of the mean number of interactions per crossing $\mu$ for 2011 and 2012.

## 3.2 The ATLAS Detector

The ATLAS[13] experiment is a multi purpose particle detector located at IP1 of the LHC. Its physics programme focuses on the Higgs boson and new phenomena beyond the Standard Model. Therefore it is designed to cope with the high interaction rates and particle multiplicities expected from proton-proton collisions. That means efficient reconstruction and identification of basic objects such as electrons, photons, muons and $\tau$-leptons as well as light mesons and hadronic jets. These requirements translate into high resolution tracking of charged particles and vertex reconstruction plus excellent calorimetry to be achieved with high granularity elements in hermetic $4\pi$ coverage. Moreover the huge particle flux demands radiation hard electronics and sensors combined with fast information read-out and triggering to filter the events of interest and also to reduce the amount of data to an acceptable size. This could only be achieved by an outstanding collaboration of thousands of physicists, engineers and technicians who dedicated their work to the design, fabrication, installation and maintainance for about 20 years.

The result can be seen in figure 24 which shows a cut-away view of the ATLAS detector. Its cylindrical layout of several concentric layers called barrel regions is completed with end cap disks on each side to close the acceptance to nearly full coverage, leading to a total length of 44 m with a diameter of 25 m. The overall weight of the detector is approximately 7000 t. From inside to outside the main components are the inner detector, the calorimeters and the muon system, each composed of various subdetectors which are briefly discussed in the following sections based on the descriptions in [46], [47] and [48] and the fact sheet [49]. Fundamental interaction of particles with matter and its application to detector design are described in detail in [50] and [51].

The coordinate system used throughout this analysis is a right-handed system with its origin in the nominal interaction point. The z-axis is defined in the direction of the counter clockwise circulating beam. Positive z values are referred to as A-side, negative z values define the C-side. The x-axis is pointing towards the center of the LHC ring and the y-axis upwards from the interaction point spanning the transversal plane. So the azimuthal angle $\phi$ is measured around the beam axis whereas the polar angle $\Theta$ is measured between the z-axis and the transversal plane (x,y):

$$\phi = \arctan\left(\frac{y}{x}\right) \tag{3.12}$$

$$\Theta = \arctan\left(\frac{\sqrt{x^2+y^2}}{z}\right) \tag{3.13}$$

[13] Acronym for: A Toroidal Lhc ApparatuS

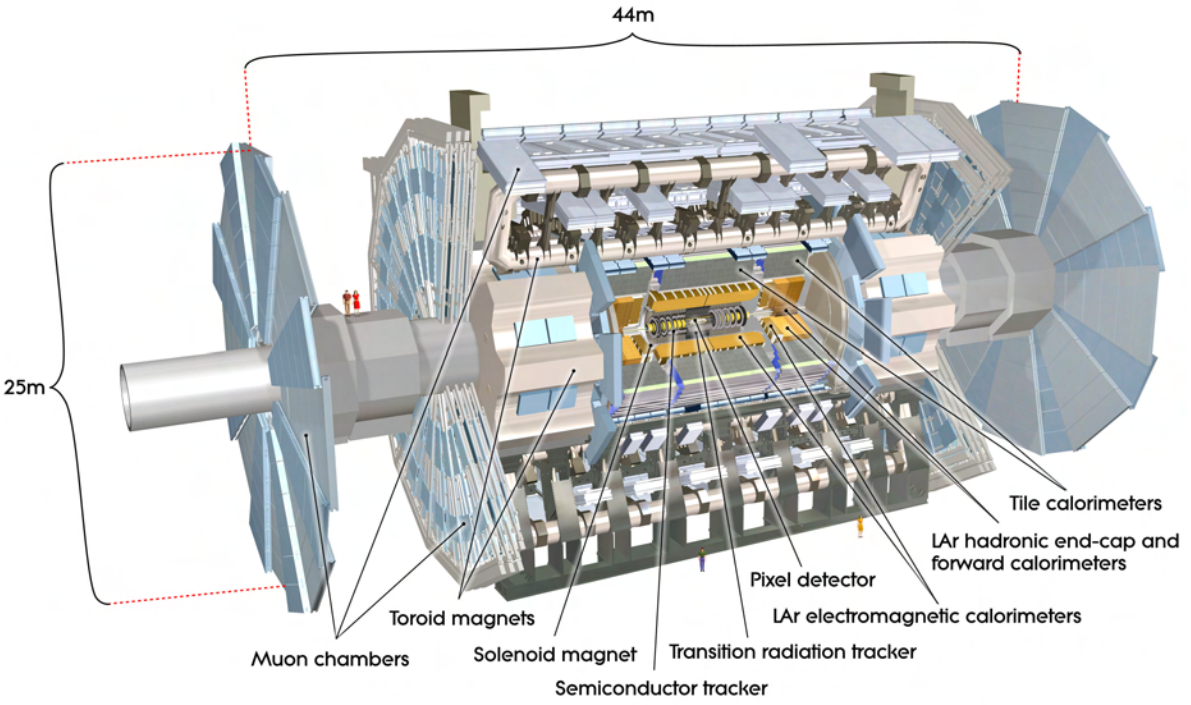



Figure 24: ATLAS detector scheme [48].

But instead of using the polar angle it is more common to define the pseudorapidity

$$\eta = -\ln[\tan(\Theta/2)] \tag{3.14}$$

In case of massless particles this is equal to the rapidity $y = -\ln[(E + p_z)/(E - p_z)]$. The distance $\Delta R$ of two objects in the ATLAS coordinate system is then given by:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \tag{3.15}$$

where the $\Delta\eta$ and $\Delta\phi$ label the difference of the two objects coordinates in $\eta$ and $\phi$ respectively.

### 3.2.1 Inner Detector

The ATLAS inner detector (ID) is built as close as possible to the interaction point to provide robust pattern recognition, precise tracking and accurate momentum and charge measurement. It is operating in a nearly homogeneous magnetic field of 2 T created by a superconducting solenoid magnet forcing charged particle trajectories to bend depending on their charge and momentum. Made up of three independent subcomponents with

decreasing spacial resolution, the inner detector consists of the pixel detector followed by the semiconductor tracker (SCT) and the transition radiation tracker (TRT) as shown in figures 25 and 27. Each component is set up in cylindrical layers around the beam pipe and wound up by perpendicular end cap disks. The length of the complete inner detector is 6.2 m with a radius of 1.15 m.

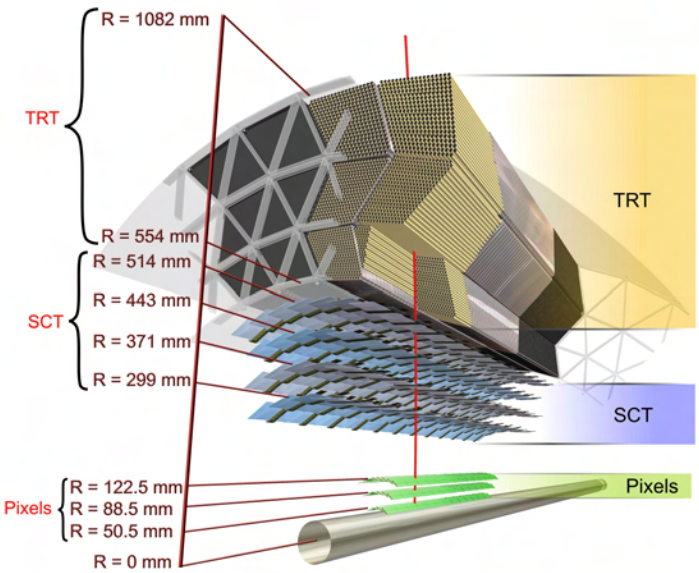



Figure 25: Drawing of the ATLAS Inner detector components showing the sensors and structural elements traversed by a charged track [48].

This leads to a large amount of individual modules. The accurate position of each component needs to be monitored in order to correctly reconstruct the particle tracks. This procedure is called alignment. The ATLAS inner detector alignment [52] uses a track-based approach to remove detector deformations and to update the magnetic field map. Technically, this is performed by minimizing the track-hit residuals with a $\chi^2$ method, iterating over the various subsystems to reduce the number of degrees of freedom. The result is an alignment of the individual sensing devices with a precision of about 1 $\mu$m.

### Pixel Detector

The Pixel Detector itself consists of three cylindrical layers and three end cap disks on both sides made up of semiconductor detectors segmented in two dimensions called pixels. Each of the pixels is 250 $\mu$m thick and covers an area of 50×400 $\mu$m providing the highest granularity of the ATLAS subsystems with a total of approximately 80 million readout channels. This leads to a resolution of 14 $\mu$m in R-$\phi$ and 115 $\mu$m in z for the barrel and in R for the end caps respectively. At the LHC design luminosity, approximately 1000 particles emerge from the collision point every 25 ns within the

region $|\eta| < 2.5$ covered by the pixel detector. The innermost layer is located only 50.5 mm from the nominal interaction point to maximize the accuracy of the primary and secondary vertex reconstruction. It is also called b-layer or vertex-layer since it is crucial for the identification of the different vertices, especially those of hadronic jets emerging from b-quarks.

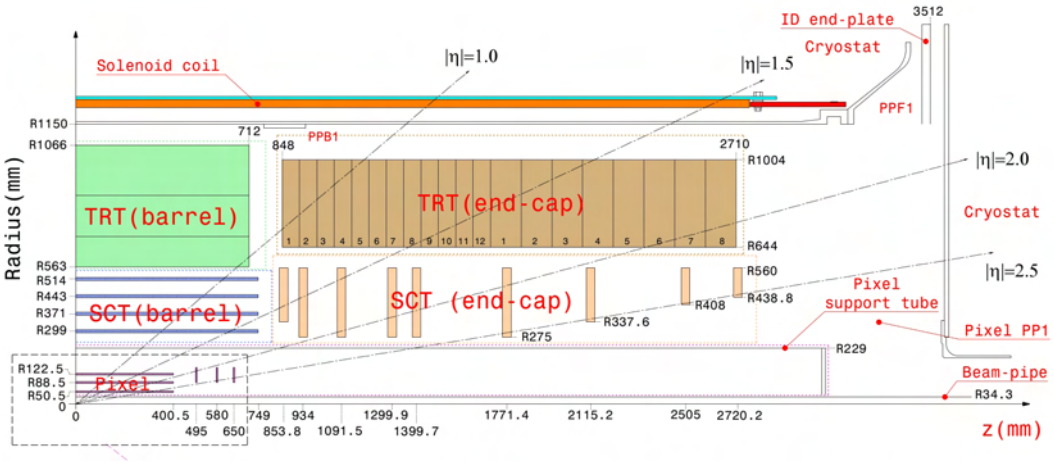



Figure 26: Plan view of a quarter section of the ATLAS inner detector showing each of the major detector elements with its active dimensions and envelopes [48].

### Semiconductor Tracker

The Semiconductor Tracker is a silicon strip detector which comprises four cylindrical layers in the barrel region plus nine end cap disks on each side. Starting about 300 mm from the collision point, it also covers up to $|\eta| < 2.5$. Each layer consists of two strips glued together back to back. They are rotated against each other by $\pm 20$ $\mu$rad to provide the required space-point resolution in R-$\phi$ and R. In the barrel one side of the stereo strips is aligned axially to the beam and the other therefore at a stereo angle of 40 $\mu$rad. In the disks, the same arrangement is made in radial direction respectively. Their mean pitch is 80 $\mu$m. The SCT and the Pixel Detector sensors are operating at low temperatures from $-5$°C to $-10$°C to maintain the noise performance after damage due to radiation.

With its larger lever arm the SCT provides better momentum resolution compared to the Pixel Detector, although the granularity is generally lower with about 6.3 million readout channels. Charged particles are usually creating eight hit points in the stereo layers resulting in four 3-d track points. The resolution is 17 $\mu$m in R-$\phi$ and 580 $\mu$m in z for the barrel as well as in R for the end caps.

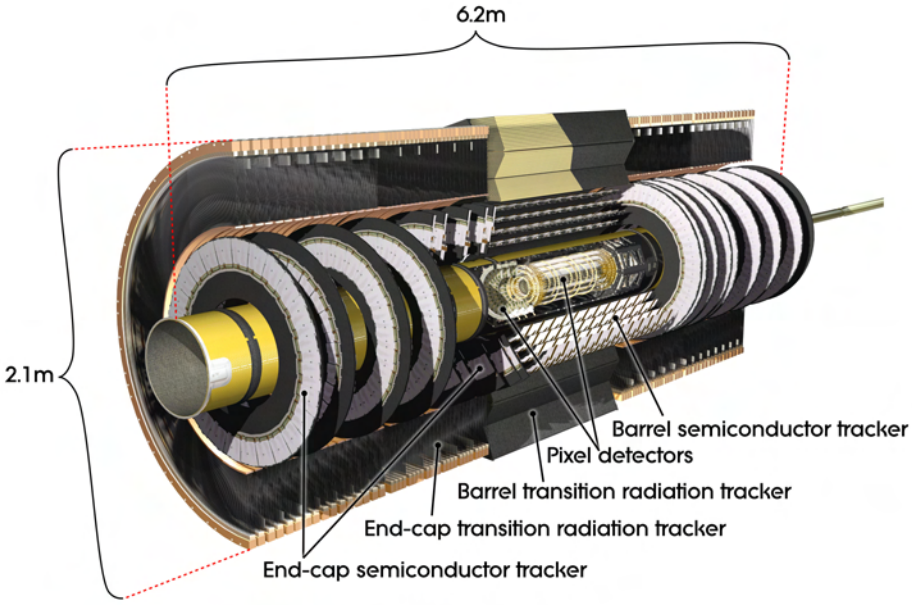



Figure 27: Cut-away view of the ATLAS Inner detector [53].

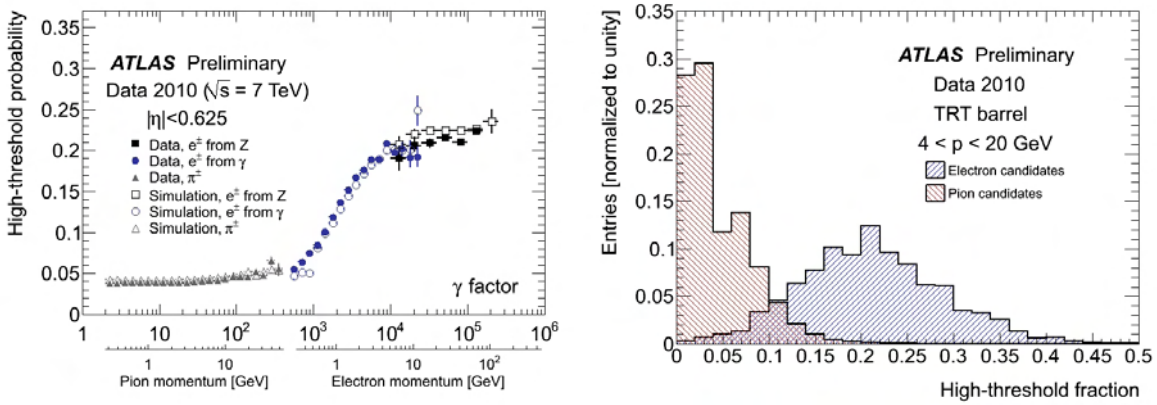
### Transition Radiation Tracker

The Transition Radiation Tracker is made up of about 300000 drift tubes interleaved in transition radiation material. The drift tubes have a diameter of 4 mm and are operated as proportional counters. The tube walls consist of a 25 $\mu$m thick polyimide mantle coated with a 0.2 $\mu$m film of aluminium which serves as cathode kept at a voltage of –1.5 kV. The construction is mechanically stabilized by carbon fibres. A gold-plated 31 $\mu$m tungsten wire in the tube center acts as anode kept at ground potential. In the barrel region, the series of 144 cm long drift tubes are installed parallel to the beam axis. The electronic read-out happens right at the front end of the modules. A glass capillary in the middle splits up each tube into two 71 cm long sensitive parts. In the end cap wheels, tubes of 37 cm length are inserted radially. The drift tubes are filled with a mixture of 70% xenon, 27% carbon dioxide and 3% oxygen as detection gas. The TRT covers the region of $|\eta| < 2.0$. In contrast to the Pixel detector and the SCT it is operating at room temperature. The resolution in R-$\phi$ is 130 $\mu$m. However, its lack of intrinsic precision is compensated by the large number of detection points.

The drift tube structure is embedded in a matrix of 19 $\mu$m diameter polypropylene-polyethylene fibres in the barrel region and foils in the end caps. These layers serve as transition radiation material. Relativistic charged particles crossing the borderline of two different dielectric media emit transition radiation, typically in the roentgen range of several keV. The emitted radiation is then absorbed in the gas of the straw tubes while its intensity is proportional to the dilation factor $\gamma = E/m$. This can be used to distinguish particles of different mass and therefore enhances the detector's identification

capabilities, in this case primarily electrons versus pions. The xenon efficiently absorbs photons in the energy range of about 10 keV. Carbon dioxide is added as so called quench gas, absorbing ultraviolet photons produced in the ionisation avalanche and thus confining it spatially. The oxygen in the mixture enhances this property and stabilizes the amplification process for the detection read-out.

Consequently, the Transition Radiation Tracker provides two types of information which are the positional tracking and the particle identification through the characteristic transition radiation. That is why the TRT read-out comprises a dual threshold discriminator, one low threshold hit for tracking and a second high threshold hit to detect larger energy deposits due to transition radiation absorption. Figure 28 (a) shows the so called turn-on curve for the barrel region. It illustrates the probability to measure a high threshold hit depending on the Lorentz factor $\gamma$ of the detected particle. The turn-on starts at $\gamma \sim 10^3$, which corresponds to a pion momentum of about 100 GeV. Its plateau is reached at $\gamma \sim 10^4$, yielding a high threshold hit probability of about 20%. But this corresponds to electron momenta of only a few GeV. So electrons have a much larger probability to produce high threshold hits than pions of comparable momentum, which can be used as identification information. On average, there are 36 low threshold hits per track along with 10 high threshold hits for electrons above 2 GeV. The fraction of high threshold hits versus the total number of hits associated to the particle track is shown in figure 28 (b) for electron and pion candidates in the barrel region. In combination with the calorimeter information described in the following section, this ratio can be used to separate electrons from heavier particles.




(a) (b)

Figure 28: (a) Probability of a high threshold TRT hit as function of Lorentz factor for the central region $|\eta| < 0.625$ and (b) the fraction of high threshold hits for electron and pion candidates [54]. The value of the Lorentz factor is calculated using the assumed mass of the candidate.

### 3.2.2 Calorimeter

The calorimeter measures the amount and position of energy deposited in the detection material by traversing particles. All calorimeters of the ATLAS detector are so called sampling calorimeters, which are made up of alternating layers of absorber and detection media. When a particle hits the absorber, it produces a shower of secondary particles which are detected in the active material. Therefore it loses part of its energy at each step of the sampling layers until it is finally stopped leaving a cascade of calorimeter signals along its track. Different calorimeter components using different materials come into operation to provide good resolution of the electromagnetic and hadronic showers emerging from the various particles produced in the hadron collisions. All these particles must be stopped completely in the calorimeter to make sure that the total energy is collected and to avoid punch through into the following muon system leading to the final thickness of the detector components.

The ATLAS calorimeter system can be subdivided into three independent parts: The Electromagnetic Calorimeter, the Hadronic Calorimeter and the Forward Calorimeter. Together they cover the region $|\eta| < 4.9$. An overview is shown in figure 29.

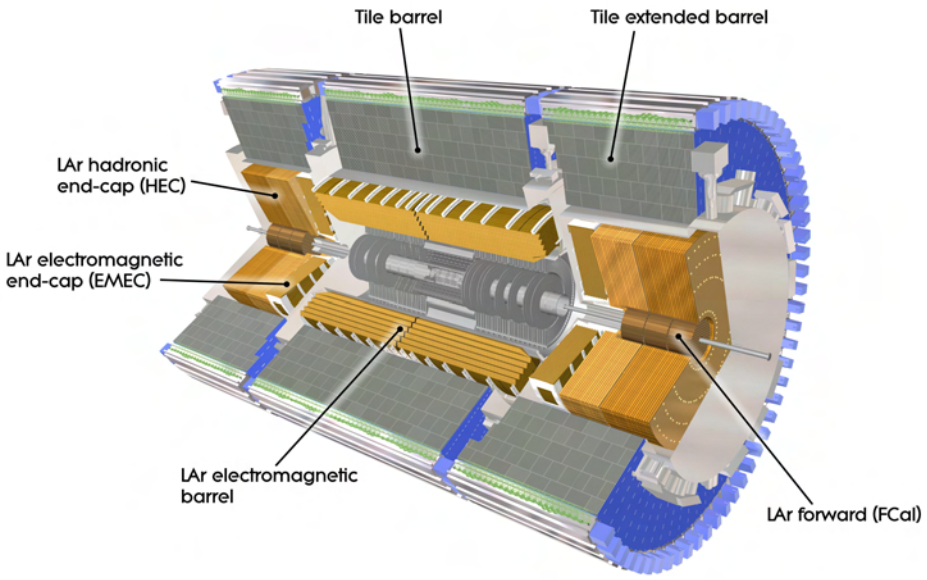



Figure 29: ATLAS Calorimeter [55].

#### Electromagnetic Calorimeter

The electromagnetic (EM) calorimeter is made of lead absorbers, spacers and copper readout electrodes constructed in a so called accordion shape, which provides optimal $\phi$-symmetry without azimuthal cracks and therefore gives uniform response functions

independent of the direction of the incoming particles. The active material is liquid argon (LAr), chosen for its radiation hardness, signal speed and linearity. Traversing particles create secondary particle showers in the lead absorbers ionizing the liquid argon and the induced 2 kV electric field collects and reads out the calorimeter signal through the electrodes.

The EM calorimeter consists of a barrel part and two end cap wheels enclosing the Inner Detector. The barrel is split into two halves of 3.2 m length separated by a 4 mm gap at $z = 0$, covering $|\eta| < 1.475$. It is 53 cm thick translating into 22 $X_0$ radiation lengths. The end caps expand the acceptance within $1.35 < \eta < 3.2$, each wheel having a thickness of 63 cm and a radius of 2 m. Both parts come with their own cryostats to keep the liquid argon at its operation temperature of $-183°$ C. To correct for the energy loss in the dead material emerging from another cryostat of 1.5 $X_0$ placed in front of the calorimeter, a 11 mm thick liquid argon presampler is installed as well.

For the extensive survey of the shower shapes, the barrel EM calorimeter is composed of three longitudinal layers with decreasing granularity shown in figure 30. The first layer is called strip layer since it is made up of very fine slices in $\eta$ with the granularity of $\Delta\eta \times \Delta\phi = 0.003 \times 0.1$ and a radiation length of 4.3 $X_0$. The second layer collects most of the energy of electrons and photons with its radiation length of 16 $X_0$. Therefore its segmentation is uniform in the barrel with a granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. The following third layer is only 2 $X_0$ thick and slightly coarser with a granularity of $\Delta\eta \times \Delta\phi = 0.05 \times 0.025$. It is used as a kind of veto layer to distinguish electromagnetic from hadronic showers. The EM calorimeter end caps consist of only two such layers with reduced granularity. Overall there are about 173000 readout channels.

### Hadronic Calorimeter

The hadronic calorimeter is surrounding the EM calorimeter and has considerably larger dimensions. Its design with an interaction length $\lambda > 10$ is sufficient for hadronic jets.In the barrel region it is made of a sandwich structure of steel absorbers and scintillating plastic tiles as active material. Therefore it is usually referred to as Tile Calorimeter consisting of one 5.8 m long barrel plus two additional extended barrels of 2.6 m length each, spanning from an inner radius of 2.28 m to an outer radius of 4.25 m. The light of the scintillators is collected at the edges of each tile and read out into two photomultiplier tubes by wavelength-shifting fibres. All of the 500000 tiles are oriented radially perpendicular to the beam axis providing almost seamless azimuthal coverage up to $|\eta| < 1.7$. Eventually, the three longitudinal sampling segments have a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in the first two layers and $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ in the third.

The hadronic end cap calorimeter (HEC) consists of two wheels on each side made up of copper absorbers and liquid argon similar to the EM calorimeter. It covers the region between $1.5 < |\eta| < 3.2$ whereas each wheel is divided into two longitudinal segments of granularity $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ below $\eta = 2.5$ and $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ above $\eta = 2.5$

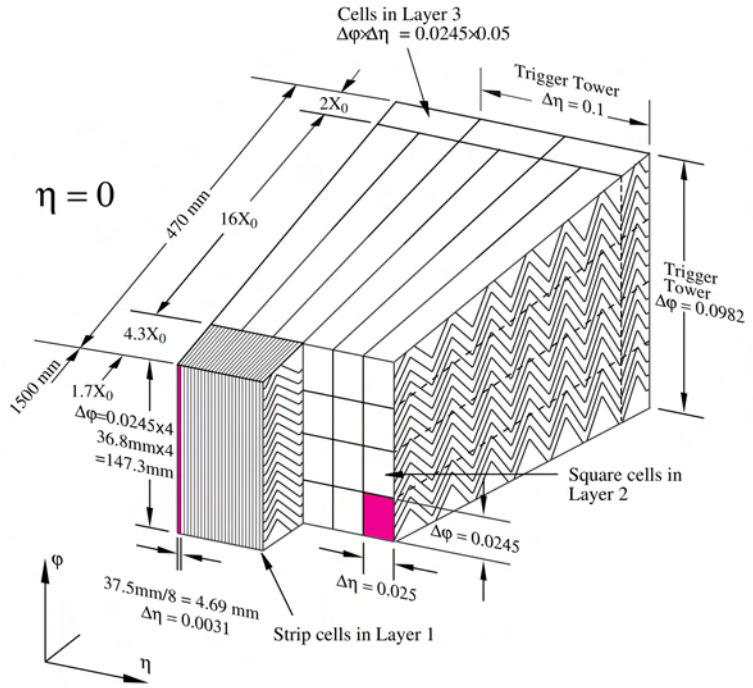



Figure 30: Sketch of a Barrel Module of the ATLAS Calorimeter [48]. The granularity in $\eta$ and $\phi$ of the cells of each of the three layers and of the trigger towers is also shown.

respectively. Both parts of the hadronic calorimeter have about 15500 readout channels. However there is a so called crack region, the transition region between barrel and end caps at $1.37 < |\eta| < 1.52$, reserved for the service of the inner detector and the LAr calorimeters. Because of the reduced resolution within this sector, it is excluded from studies with electrons.

### Forward Calorimeter

The Forward Calorimeter (FCal) is a three layer liquid argon calorimeter covering the high $3.1 < |\eta| < 4.9$ range with an interaction length of $\lambda \sim 10$. Starting at a distance of about 4.7 m from the collision point, each 45 cm long layer is filling the center of the end cap disks with a radius of 50 cm around the beam axis. While the first layer using copper absorbers is dedicated to electromagnetic showers, the second and third layer use tungsten absorbers for the detection of hadronic showers. They are structured in tubes parallel to the beam axis to provide fast read-out at high density with considerably small liquid argon gaps, to cope with high particle flux in the forward directions making out more than 3500 channels. The material distribution of all calorimeter subdetectors can be seen in figure 31. It shows the interaction length of the various layers depending on the pseudorapidity $\eta$ and the contributions from the surrounding material.

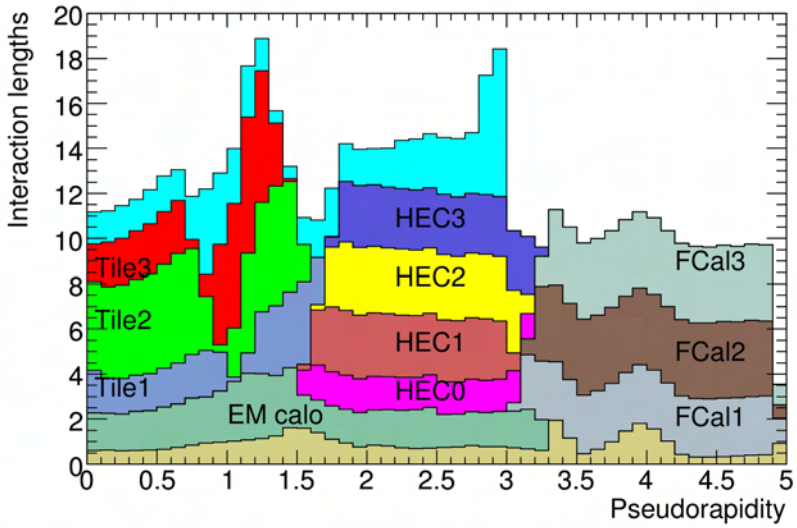



Figure 31: Cumulative amount of material, in units of interaction length, as a function of $\eta$, in front of the electromagnetic calorimeters, in the electromagnetic calorimeters themselves, in each hadronic layer, and the total amount at the end of the active calorimetry [48].

### 3.2.3 Muon Detector

Due to the high interaction length of the calorimeters, only muons are expected to pass through the detectors described in the previous sections. That is why the muon system constitutes the final part of the ATLAS detector, making up most of its enormous size. Covering the region of $|\eta| < 2.7$, the muon system provides precise identification and momentum measurement for muons above $p_T > 3$ GeV, independent from the inner detector tracking. Due to its long lever arm the resolution is very large, designed to reach 10% even at transverse momenta of 1 TeV. This of course needs again a magnetic field in which the muon trajectories are bent, in this case provided by a toroidal air core magnet system.

The air core magnets have no magnetic yoke but consist of a non-magnetic structure fixing the superconducting cables with a total length of about 70 km. This minimizes the degradation of momentum resolution due to multiple scattering which would occur in a typical dense material yoke. In the barrel region the magnetic field is created by eight race-track shaped coils with 3 – 7 T. They cover up to $|\eta| < 1.4$ and are housed in large cryostats. Two end cap magnets with eight coils each are spanning from $1,6 < |\eta| < 2.7$ and are rotated about 22.5° with respect to the barrel coils to provide radial overlap of the fields and optimize the bending power in the transition region. This structure of the coils is shown in figure 32. The magnetic field is pointing in $\phi$ direction, thus bending the muons trajectories in the $\Theta$ direction. With a nominal current of 20000 Ampere the stored energy is about 1 GJ.

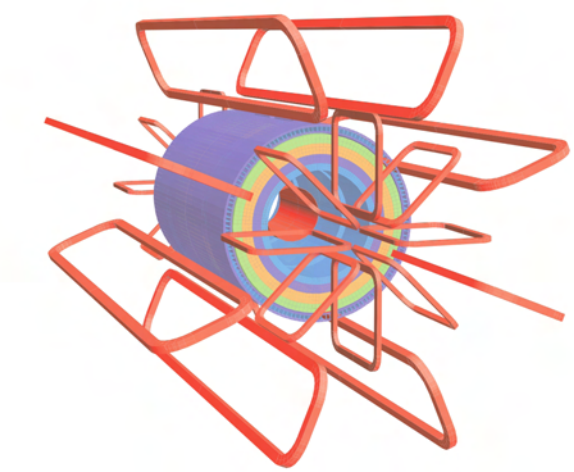


Figure 32: Geometry of the toroidal magnet windings. The eight barrel toroid coils surrounding the tile calorimeter, with the end-cap coils interleaved [48].

The muon detector modules are located between and on the the barrel coils, as well as in front and behind the end cap coils respectively. While the three concentric cylindrical layers of the barrel region are surrounding the beam axis at a radius of $r = 5/7.5/10$ m, the four end cap wheels on each side are set up perpendicular to the beam axis at $|z| = 7.4/10.8/14$ and 21.5 m.

The main part of the muon system is made up of 1088 monitored drift chambers (MDT) perpendicular to the beam axis, spread over the whole $\eta$-range. Each chamber is carrying two multi-layers of three to four drift tubes layers. Those tubes with a diameter of 3 cm are filled with a gas mixture of 93% argon and 7% carbon dioxide at a pressure of 3 bar and a 50 $\mu$m diameter tungsten rhenium wire serving as anode. When the passing muons are ionizing the gas, the induced electron avalanche is collected at the anode and the signal timing as function of the drift radius gives the muon's closest approach to the wire and therefore the position in one dimension.

In the first end cap wheel on each side of the ATLAS detector, there is one exception to the general structure. The range $2.0 < |\eta| < 2.7$ of the muon system is covered by 32 high granularity cathode strip chambers (CSC) which are multiwire proportional chambers. Their wires are radially oriented while the cathodes on both sides are segmented into strips, one set perpendicular to the wires and the other in parallel. The CSC modules provide the high rate capabilities and time resolution needed for the operation in the forward boundary of the muon system.

An additional set of trigger chambers rounds off the muon system. In the barrel region they are realised by 544 resistive plate chambers (RPC) located below and above the MDTs. Their 2 mm separated electrode plates are filled with a special gas mixture combining low operating voltage, non-flammability and low cost with comfortable plateau.

The end caps are covered by 3588 thin gap chambers (TGC) which are multiwire proportional chamber where the wire-to-cathode distance is smaller than the wire-to-wire distance. Those trigger chambers are providing not only trigger information but also bunch crossing information and last but not least azimuthal muon coordinates, orthogonal to the tacking chambers in the $|\eta| < 2.4$ range. The set-up of all the different modules can be seen in figure 33.

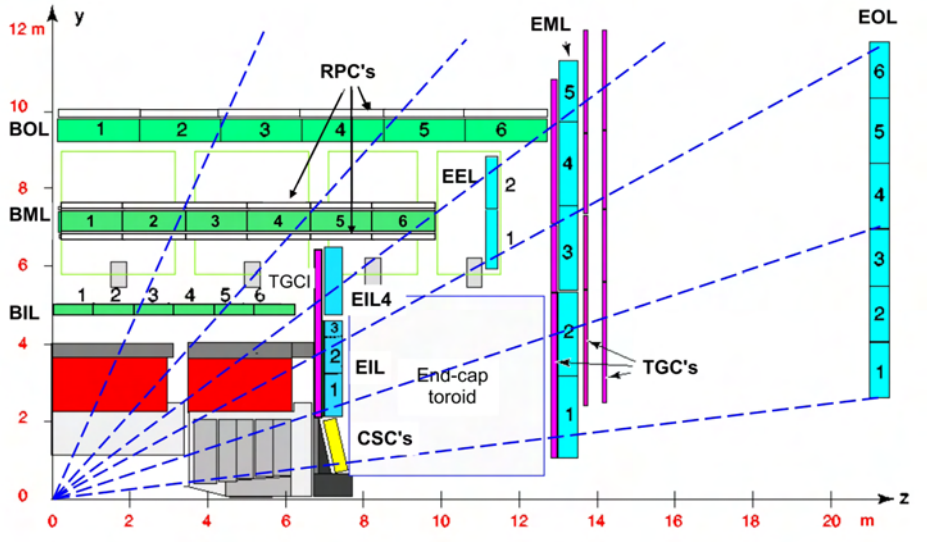



Figure 33: Cross section of the muon system in a plane containing the beam axis (bending plane). Infinite-momentum muons would propagate along straight trajectories which are illustrated by the dashed lines and typically traverse three muon stations [48].

With more than 1 million readout channels the muon system detectors are covering about 5500 $m^2$. But a very important ingredient to achieve the resolution of about 40 $\mu$m is the alignment of this huge amount of large modules. To obtain the grade of internal deformation and the relative position of the modules to each other, a number of 12000 precision mounted alignment sensors is installed to monitor optically the deviations from a straight line.

### 3.2.4 Trigger

As the number of readout channels listed in the previous sections already indicate, the total amount of information provided by the ATLAS detector is tremendously huge. The number of events collected at high instantaneous luminosity and high collision rate is simply impossible to record and analyse, even with today's computing resources and technologies. Fortunately, the vast majority of collisions are uninteresting as the focus lies on rare processes with very specific signatures. So the ATLAS trigger system [48] is designed to reduce the information by rapidly selecting the wanted topologies online.

In a three-level approach, the initial event rates are reduced from 40 MHz to $\sim$ 200 Hz, which is compatible with the storage management and the offline computing power. At each level, the trigger decisions get refined by applying advanced selection criteria. Figure 34 illustrates the three distinct levels.

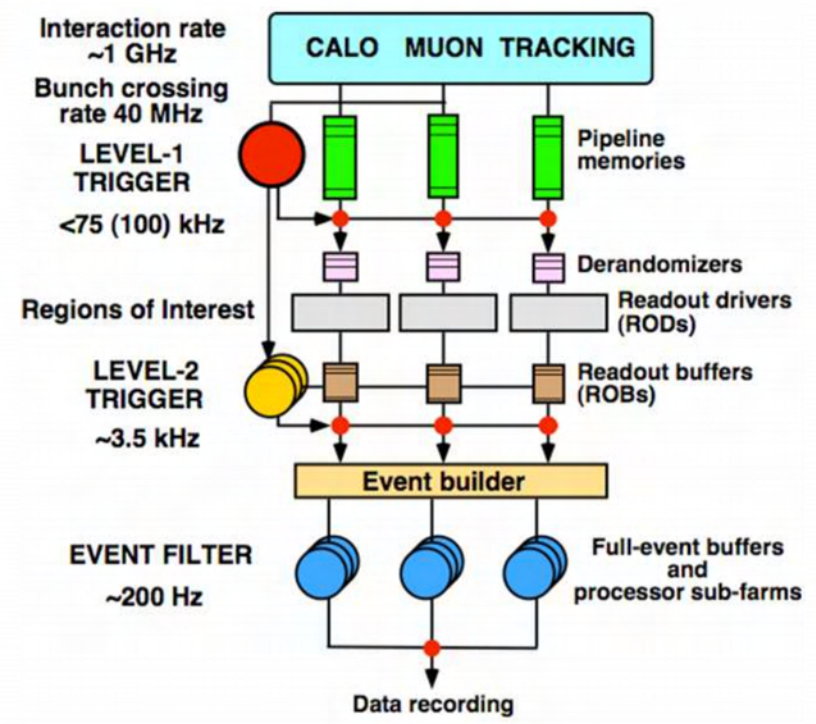



Figure 34: Schematic view of the ATLAS three-level trigger system [56].

The Level 1 (L1) trigger stage is hardware based, using only a limited amount of information from the calorimeters and the muon spectrometer to decide whether to keep or refuse the event within 2.5 $\mu$s. The raw data is cached in Read Out Buffer (ROB) queues directly in the detector electronics, which get grouped together into a Read Out System (ROS) from which the trigger logic can pull the information.

The calorimeter trigger (L1Calo) [56] searches for high transverse momentum signatures emanating from electrons, photons, jets and hadronically decaying tau-leptons as well as missing and total transverse momentum. It sums the deposited energy in so called towers of coarse granularity, separately for the electromagnetic and the hadronic calorimeters. Two processors run in parallel to find a decision. The cluster processor is used for e, $\gamma$ and $\tau$ based triggering and works on the basis of so called regions of interest (ROIs), identified by clusters of $2 \times 2$ EM calorimeter towers. Those are defined by their geographical coordinates $\eta$ and $\phi$. To fire the trigger, they need to pass programmable thresholds. The surrounding ring of towers is considered for isolation criteria as indicated in figure 35. Similarly, the energy-sum processor identifies larger tower ROIs of $4 \times 4$, $6 \times 6$ or $8 \times 8$ centered around a local maximum found in the EM and the hadronic calorimeters.

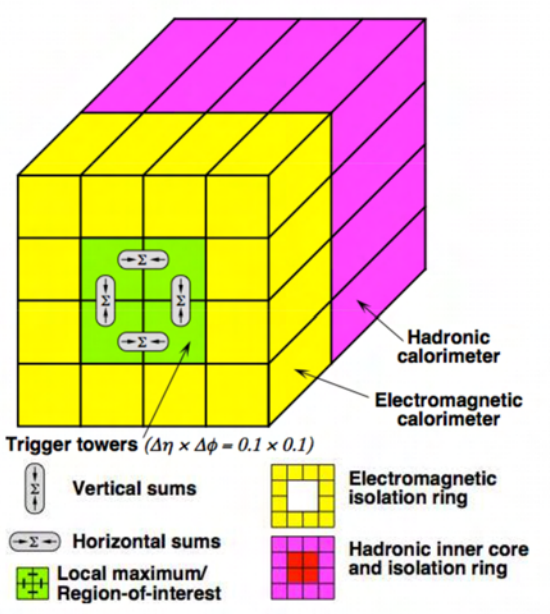



Figure 35: Illustration of the trigger towers as used in the e/$\gamma$ algorithm of the L1 hardware trigger [56].

Additionally, the L1 muon system uses information from the RPC in the barrel region and the TGC in the end caps. Those are composed of planes of two to four chambers. Figure 36 shows how the hits are used to reconstruct rough estimates of $p_T$, $\eta$ and $\phi$ for the passing muons. In the end, the L1 trigger reduces the event rate from the initial 40 MHz to about 75 kHz which are transmitted to the next stage.

The Level 2 trigger (L2) is software-based. It uses nearly the full granularity information of all subdetectors for the transmitted ROIs. Therefore the L2 trigger and the subsequent third stage are often referred to as High Level Trigger (HLT). A sequence of trigger algorithms is executed to compute refined event characteristics in order to determine if the candidate should be retained. With an average processing time of 40 ms, the event rate gets cut down to approximately 3.5 kHz.

The final stage of the online selection is performed by the so called Event Filter (EF). Typically, it uses the same algorithms as the offline reconstruction described in the following chapter. For this purpose, the full detector information is taken into account which takes a few seconds per examined event. This leads to a rate reduction to ∼200 Hz in the end. Every event passing this series of tightening decisions is finally stored in the CERN computer center for further offline processing and available for analysis. The bandwidth reaches about 500 MB/s.

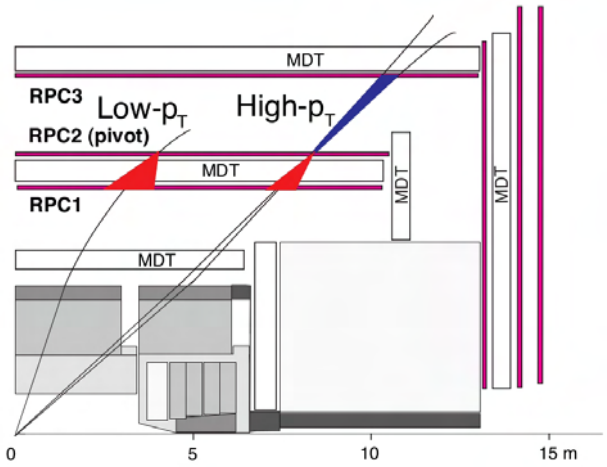



Figure 36: Scheme of the L1 muon barrel trigger illustrating the $p_T$ estimation based on hits in the trigger chambers [48].

# 4

# Reconstruction of Physics Objects

This section will briefly introduce the main concepts of event reconstruction. The transition from raw detector information to physical objects is a major prerequisite for any physics analysis. Therefore, the methods and algorithms in the reconstruction stage have to be reliably robust to maintain optimal performance. A substantial description of the following chapter is given in [46], but updated studies will be referenced where applicable. While the first part deals with the reconstruction of tracks and vertices, the second part passes on to the identification of the physics objects, including their four-momenta and the complete event kinematics.

## 4.1 Event Reconstruction

Before covering the reconstruction of physics objects, there are two basic event parameters that need to be clarified. The track determination is a fundamental ingredient of the object reconstruction of charged particles or jets presented later on. But it is also essential for the identification of the decay vertices, which set the fundamental frame for the physics analyses. So the vertex reconstruction is highly correlated with the track reconstrunction. The reconstruction of so called energy clusters from the calorimeter information is described later on for electrons and jets.

### 4.1.1 Track Reconstruction

Charged particles follow a circular trajectory in the transverse plane due to the mounted magnetic field. At the high instantaneous luminosities achieved in 2012, up to $\sim 1000$ tracks are observed per bunch crossing. As explained in section 3.2.1, the inner detector is designed to measure such large numbers and densities of particle tracks. So the

tracking is limited to its coverage of $|\eta| < 2.5$. The reconstruction of charged particle tracks is explained in [57]. It makes use of a set of five parameters to determine the trajectories from the single hit points and to estimate the quality of the observed tracks. These are the charge over the transverse momentum $q/p_T$, the azimuthal angle $\phi$ and the polar angle $\Theta$ plus the transverse impact parameter $d_0$ and the longitudinal impact parameter $z_0$. The latter two define the distance from the point of closest approach to the reference point. As long as the primary vertex is not yet known, the reference point is given by the center of the beamspot.

There are two separate track finding algorithms. The first one works inside-out, which means that it starts from the innermost tracking layers and extrapolates to those sitting further outside. Track seeds are built from the three dimensional space points of the high granularity pixel hits and the SCT cluster pairs. They define the so called roads, on which further track points are expected. The track fit is performed with a Kalman filter algorithm. Hits are added iteratively while the trajectory is refitted at each step. A $\chi^2$ threshold separates the actual hit points from the so called outliers. The performance of the silicon detector has been tested in $\sqrt{s} = 7$ TeV data and was found to be well described by Monte Carlo simulation [58]. Several cuts on the track quality criteria are applied in order to solve ambiguities in the cluster-to-track association and to reject fake tracks. The surviving track candidates are extrapolated to the TRT. Additional TRT hits are searched for and included into a combined fit. This procedure leads to the final track parameters. Every track candidate above $p_T > 400$ MeV is written to the database and available for further event and object reconstruction. A complementary outside-in algorithm is seeded from unused segments in the TRT. It predominantly recovers tracks from decays of long lived particles such as kaons and also photon conversions, because their traces consist of less hits in the innermost detector layers. This is often referred to as back-tracking and improves the efficiency for secondary tracks.

### 4.1.2 Vertex Reconstruction

The vertex finding and fitting procedure needs to cope with the simultaneous reconstruction of multiple interactions, taking into account the sharing of tracks between close-by vertices. Therefore, it has to be adapted to the predominant pileup conditions. It constitutes an essential input not only for a precise track reconstruction but for the general categorization of the analyzed event.

The method and performance of the vertex reconstruction are described in [59]. First, the vertex finding considers all reconstructed tracks as defined in the previous section. It searches for a global maximum of track z-coordinates computed at the point of closest approach to the beam spot center, which is determined every few minutes during collisions. This leads to exactly one seed vertex. In the next step, an adaptive vertex fitter [60] determines its position with a robust $\chi^2$-based method. Tracks that are incompatible by more than seven standard deviations with this first vertex are seeding the next vertex to be fitted. This procedure is repeated until no more tracks are left. Then all tracks are

refitted with the newly assigned constraint of the associated vertex. The reconstructed vertex with the largest sum of squared transverse momenta of the associated tracks is called hard scatter vertex or simply primary vertex (PV). It defines the reference point for the further analysis and therefore identifies objects that are originating from the collision of interest. All other vertices are assumed to result from underlying interactions, beam gas interactions or cosmic particles. Therefore, the primary vertex is required to be reconstructed from at least three tracks. The reconstruction efficiency for $n^{PV}_{tracks} \geq 3$ is nearly 100%.

## 4.2 Object Definition and Identification

This section covers the main ingredients for the presented analysis. The definition of the physical objects that are expected in the decay signature of the $H \to WW \to \ell\nu\ell\nu$ process are introduced one by one. As there are different versions tweaked to the different event topologies, the focus lies on the actual working points used in the final analysis.

The first part deals with electrons and muons. Since particles and anti-particles interact equally with the detector material, there is no need to distinguish. The only difference is their electrical charge, which has to be determined, of course. But in the habitual language use, they will be called electrons and muons from now on. Note that positrons and anti-muons are always included in this description, unless explicitly stated otherwise. The following part pictures the so called jets, which are made of strongly interacting particles generating a shower in the calorimeter material that are summed up to one distinct object, avoiding a flush of information from their substructure. Finally, in the last section, the missing transverse energy is defined. Since it describes the balance of the actually observed particles, it obviously relies on the definition of the other objects.

### 4.2.1 Electrons

Electrons are reconstructed using information from both the calorimeter as well as the inner detector [61]. The deposition of energy in the calorimeter alone is not sufficient to distinguish electrons from photons. Therefore, a matching track has to be detected to uniquely identify the charged electrons. The algorithm providing a high and uniform efficiency over a wide range in $p_T$ and $\eta$ is designed to reject real isolated electrons against background objects. These are for example misidentified jets or electrons emerging from photon conversions and heavy flavour decays.

The electron reconstruction algorithm consists of two separate parts. First, the clustering of neighbouring energy deposits in the EM calorimeter and second, the association of a reconstructed track to a particular cluster. This approach is called outside-in, since the calorimeter is located further away from the interaction point than the tracker. As a starting point, the clusters are formed by a sliding window algorithm scanning the EM calorimeter for local maxima in the transverse energy above 2.5 GeV. The window spans

over $3 \times 5$ cells, each corresponding to the granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ in the second calorimeter layer. The efficiency is almost 100% for $E_T > 20$ GeV [62]. The second step is the track association. It is limited to the acceptance of $\eta < 2.47$, which is the coverage of the inner detector. Tracks above 0.5 GeV are extrapolated to the middle layer of the EM calorimeter. They are required to match the cluster position within $|\Delta\eta| < 0.05$. The $\Delta\phi$ requirement on the other hand is enlarged to 0.1, to account for the Bremsstrahlung loss which results in increased transverse bending in the magnetic field. A Gaussian Sum Filter (GSF) technique [63] is employed to refit the tracks and improve the bending plane parameters, such as the transverse impact parameter or the track angular direction. One or more tracks associated to an energy cluster define a successful track-cluster match and thus an electron. In case of several track candidates, those with hits in the pixel detector or the SCT are prioritized before choosing the one which is closest in $\Delta R$. Without a track-cluster match, the object is identified as an unconverted photon. If the matching track does not refer to the primary vertex, the candidate electron has emerged from a photon conversion.

As last step, the cluster sizes are optimized to take into account the overall energy distribution over the different calorimeter regions. Therefore, the window units are enlarged to $3 \times 7$ in the barrel region and $5 \times 5$ in the end caps. The total reconstructed energy is the sum of four different contributions. The first is of course the estimated energy within the calorimeter, corrected for the sampling fraction. Furthermore, the energy deposit in the material before entering the calorimeter needs to be accounted. This correction is based on the presampler signal. Additionally, the estimated energy outside the cluster (lateral leakage) and the deposit beyond the EM calorimeter (longitudinal/hadronic leakage) have to be determined. Those corrections are derived from Monte Carlo simulations. So in the end, the resulting electron four momenta are given by the cluster energy, while the $\eta$ and $\phi$ directions are taken from the track parameters. After all these steps, the reconstructed objects are called electron candidates and are passed on to a further identification procedure.

The electron candidates are classified into three different quality categories, each representing different degrees of true electron efficiency and background rejection. Depending on the amount and rigidity of the selection criteria, they are called *loose*, *medium* and *tight*. An optimized cutting scheme binned in $\eta$ and cluster $E_T$ has been applied to define the working points.

- The ***loose*** identification requirements include the hadronic leakage information and the shower shape variables.
- The ***medium*** identification criteria additionally make use of the strip layer of the EM calorimeter, the track quality and track-to-cluster matching constraints.
- The ***tight*** identification requirements demand even tighter track-matching and track-quality constraints. Furthermore, TRT information is used to provide additional information and hadron suppression. By requiring a hit in the innermost tracking layer (b-layer), photon conversions can be reduced.

Beyond that, the cut-based scheme has recently been appended with a multivariate analysis (MVA) technique [64]. These are extensively used in physics analyses to separate signal from background. Their advantage is the simultaneous evaluation of multiple properties, rather than cutting it down one by one. In this case, a likelihood (LH) method makes use of signal and background probability density functions for the discriminating variables. Table 4 shows a list of variables that are used in the identification procedure. From top to bottom, every new constraint includes the above. A checkmark flags the variables that are employed by the likelihood method.

| Type | Name | Description | Cut | LH |
|---|---|---|---|---|
| **Loose selection** | | | | |
| Hadronic leakage | $R_{Had1}$ | Ratio of $E_T$ in the first layer of the hadronic calorimeter to $E_T$ of the EM cluster (used over the range $\|\eta\| < 0.8$ and $\|\eta\| > 1.37$) | ✓ | ✓ |
| | $R_{Had}$ | Ratio of $E_T$ in the first layer of the hadronic calorimeter to $E_T$ of the EM cluster (used over the range $\|\eta\| > 0.8$ and $\|\eta\| < 1.37$) | ✓ | ✓ |
| Third layer of EM calorimeter | $f_3$ | Ratio of the energy in the third layer to the total energy | ✓ | ✓ |
| Middle layer of EM calorimeter | $R_\eta$ | Ratio of the energy in $3 \times 7$ cells over the energy in $7 \times 7$ cells centered at the electron cluster position | ✓ | ✓ |
| | $R_\phi$ | Ratio of the energy in $3 \times 3$ cells over the energy in $3 \times 7$ cells centered at the electron cluster position | | ✓ |
| | $W_{\eta 2}$ | Lateral width of the shower | ✓ | ✓ |
| **Medium selection** | | | | |
| Strip layer of EM calorimeter | $W_s^{tot}$ | Total shower width | ✓ | |
| | $E_{ratio}$ | Ratio of the difference between the largest and second largest energy deposits in the cluster over the sum | ✓ | ✓ |
| | $f_1$ | Ratio of the energy in the strip layer to the total energy | | ✓ |
| Track quality | $nPixHits$ | Number of hits in the pixel detector | ✓ | ✓ |
| | $nSiHits$ | Number of total hits in the pixel and SCT detectors | ✓ | ✓ |
| | $d_0$ | Transverse impact parameter | ✓ | ✓ |
| | $\sigma_{d_0}$ | Significance of transverse impact parameter | | ✓ |
| Track-cluster matching | $\Delta\eta_1$ | $\Delta\eta$ between the cluster position in the strip layer and the extrapolated track | ✓ | ✓ |
| **Tight selection** | | | | |
| Track-cluster matching | $\Delta\phi_2$ | $\Delta\phi$ between the cluster position in the middle layer and the extrapolated track | ✓ | |
| | $E/p$ | Ratio of the cluster energy to the track momentum | ✓ | |
| TRT | $nTRTHits$ | Total number of hits in the TRT | ✓ | |
| | $F_{HT}$ | Ratio of the number of high-threshold hits to the total number of hits in the TRT | ✓ | ✓ |
| Conversions | $nBlayerHits$ | Number of hits in the B-layer | ✓ | ✓ |
| Bremsstrahlung (GSF output) | $\Delta p/p$ | Momentum lost by the track between the perigee and the last measurement point divided by original momentum | | ✓ |
| | $\Delta\phi_{Res}$ | Same as $\Delta\phi_2$, but the track momentum is rescaled to the cluster energy before extrapolating to the middle layer | | ✓ |

Table 4: Definition of electron discriminating variables that are used in the 2012 electron cut-based menus ("Cut") and in the likelihood ("LH") taken from [64].

The likelihood method always uses all the marked variables. But several operating points have been defined suiting the efficiency benchmarks from the cut-based menu to fit the needs of different analyses. Therefore, these working points are labelled Loose LH, Medium LH and Tight LH. The Very tight LH point was additionally constructed

to deal with electrons below 20 GeV. The resulting particle identification (PID) has an improved rejection of hadrons and conversions. Table 5 shows a list of the described menus comparing their efficiencies, estimated on signal and background from a $Z \to ee$ tag-and-probe method.

| Menu | $20 < E_T < 50$ GeV | |
|---|---|---|
| | Data Efficiency ($Z \to ee$) | Data Efficiency Background |
| Loose Cuts | $95.68 \pm 0.17$ | $5.414 \pm 0.025$ |
| Loose LH | $92.82 \pm 0.18$ | $0.963 \pm 0.011$ |
| Medium Cuts | $88.09 \pm 0.22$ | $1.133 \pm 0.012$ |
| Medium LH | $87.79 \pm 0.25$ | $0.535 \pm 0.008$ |
| Tight LH | $84.15 \pm 0.27$ | $0.396 \pm 0.007$ |
| Tight Cuts | $77.48 \pm 0.24$ | $0.463 \pm 0.008$ |
| Very Tight LH | $76.97 \pm 0.29$ | $0.278 \pm 0.006$ |

Table 5: Signal and background efficiencies for likelihoods and cut-based menus, averaged over $\eta$ and $E_T$ between 20 and 50 GeV taken from [64]. Efficiencies are quoted in % ; errors on the signal include statistical and systematic uncertainties, while uncertainties in background are statistical only.

The purpose of the tag-and-probe method is to provide a clean sample of unbiased *probe* objects. So the selection cuts are mainly focusing on the *tag* object. In case of the electron efficiency measurements, the method is applied to $Z \to ee$, $J/\Psi \to ee$ and $W \to e\nu$. So one well identified electron or large missing transverse energy are required to tag the event and to select the probe electrons. The $Z \to ee$ events provide a very clean environment to study the performance, while $J/\Psi \to ee$ allows to exploit the lower transverse energy region. To reduce the contamination of the probe sample, a side-band fit is performed on the dielectron invariant mass of same sign and opposite sign pairs and the background contributions are subtracted. This procedure is illustrated in figure 37 while figure 38 shows the improvement of the likelihood approach versus the cut-based menu. The identification efficiency is defined as the fraction of probe electron candidates passing a specific set of cuts after the background subtraction.

The analysis of the $W \to e\nu$ events adds statistical power to these studies. The background contribution is evaluated with a template fit method based on isolation variables. As the isolation criteria depend strongly on the specific analysis needs, they are not included in the identification procedure described so far. They specifically help to discriminate misidentified jets or real leptons emerging from semi-leptonic heavy quark decays, both providing additional charged particles manifesting in collateral low quality tracks and energy contributions. The track-based isolation is sensitive to those accompanying tracks or pileup contributions. It is defined as the scalar sum of the transverse momenta of the selected tracks that lie within a cone of radius $R_{iso}$ around the electron candidate:

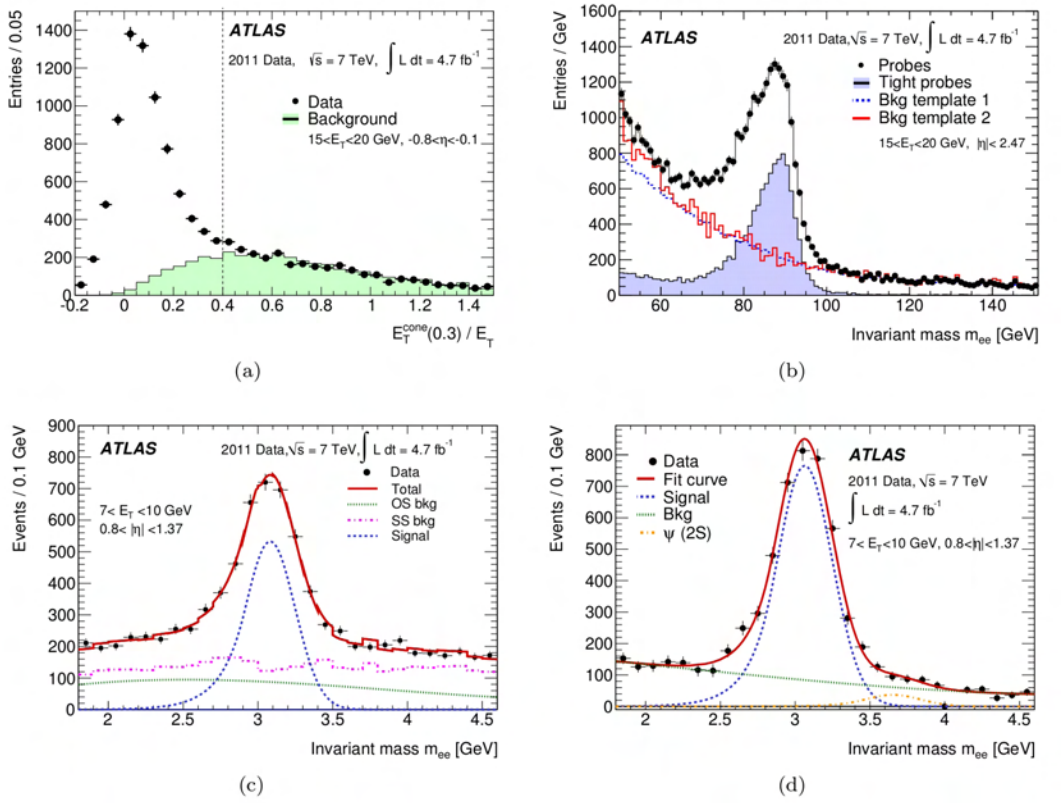



Figure 37: Examples of discriminating variables and background-subtraction techniques for illustrative $(E_T, \eta)$ bins [61]. (a) The $E_T^{cone}(0.3)/E_T$ distribution of probes in the $W \to e\nu$ sample superimposed with the normalised background template. The black dashed line indicates the threshold chosen to delineate the signal and background regions. (b) Invariant mass distribution in the $Z \to ee$ sample. The normalised shapes of two different background templates are also shown. (c) Invariant mass distribution for the $J/\Psi \to ee$ sample in the short-lifetime range. (d) Invariant mass distribution for the $J/\Psi \to ee$ sample using the lifetime-fit method.

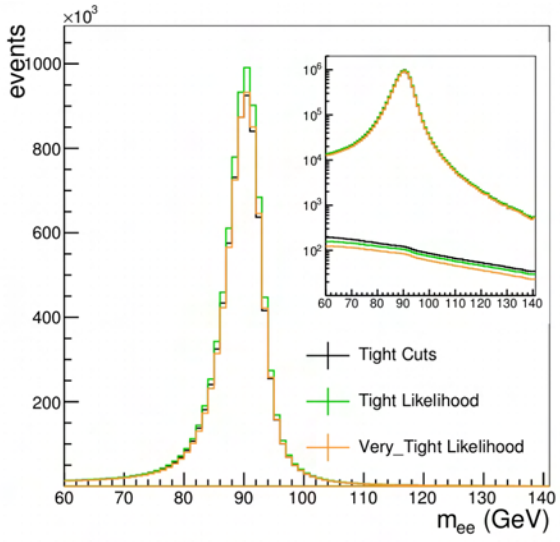



Figure 38: A comparison of the $m_{ee}$ spectrum of electron probes in data, selected using the tag-and-probe method and passing different identification criteria [64].

$$P_T^{cone,R_{iso}} = \sum_{\substack{tracks \\ w/o\ electron\ cand. \\ \Delta R<R_{iso}}} p_T^{track} \tag{4.1}$$

where the electron track itself is excluded from the summation. The amount of energy deposited in cells around the cluster helps to discover additional energy contributions of additional charged and neutral particles. The calorimeter isolation is therefore defined as the sum of energy deposited in a cone of size $R_{iso}$ around the $3 \times 5$ cells of the electron candidate:

$$E_T^{cone,R_{iso}} = \sum_{\substack{cells \\ w/o\ electron\ cand. \\ \Delta R<R_{iso}}} E_T^{cells} \tag{4.2}$$

There are two main effects contributing to the isolation determination. One is the lateral leakage into the isolation cone, the other is the energy deposit from pileup collisions before and during the bunch crossing. Dedicated corrections are described in [65]. The choice of the cone size represents a trade-off between high discrimination power and robustness against pileup. In case of the presented analysis, the value of $R_{iso} = 0.3$ was found to be performing best.

The electron energy scale can also be determined from the dielectron mass in $Z \to ee$ and $J/\Psi \to ee$ events or alternatively from the measurement of the ratio of the cluster energy over the track momentum $E/p$ in $W \to e\nu$ events. With the parameterization:

$$E_i^{meas} = E_i^{true}(1 + \alpha_i) \tag{4.3}$$

the residual miscalibration $\alpha_i$ in some $\eta$ bin $i$ can be factorized. Here, $E^{true}$ is the true electron energy, while $E^{meas}$ is the energy measured by the calorimeter (including corrections determined from Monte Carlo simulation). With a fit of the distributions mentioned above, the correction factors $\alpha$ can be used to calibrate the energy scale, resulting in effects of $\pm$ a few percent for the different detector regions.

The fractional energy resolution of the calorimeter can be parameterized as:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \tag{4.4}$$

The parameter $a$ is the sampling term, $b$ represents the noise term and $c$ is the constant term. All of them are $\eta$ dependent. But while the first two are evaluated from Monte Carlo simulation only, the constant term is determined from fits to simulated and measured invariant mass distributions of $Z \to ee$ decays. A Breit-Wigner distribution fixed to the measured Z boson width is convoluted with a Crystal Ball function describing the experimental resolution. Since the Monte Carlo simulation exhibits a slightly better resolution than observed in the experiment, the transverse momentum of the electrons in the Monte Carlo is smeared to model the data distributions correctly.

### 4.2.2 Muons

As minimum ionizing particles (MIP), muons only deposit a very small fraction of their energy in the ATLAS calorimeters. Therefore, the reconstruction and identification of muons is based on tracking information of the inner detector (ID) and the dedicated muon system (MS). Several algorithms to classify muons are available and are described in detail in [66]. These different types of muons can be used for analysis and their main components are displayed in the following.

- **Stand-alone (SA) muons** are reconstructed solely from the tracks measured in the muon spectrometer. By extrapolating the trajectory back to the beam axis, the impact parameters can be determined. The energy loss in the calorimeter is taken into account.
- The reconstruction of **combined (CB) muons** is based on the combination of the SA seeds with the independent track measurement from the inner detector. Two complementary combination schemes are available. The STACO[14] algorithm [67] performs a $\chi^2$-matching of track parameters using the covariance matrix and is also referred to as chain 1. The MUID scheme on the other hand performs a

[14] STAtistical COmbination

global refit of the track using all hits of the two subdetectors and is also called chain 2.

- A **segment-tagged (ST) muon** is an inner detector track extrapolated to the muon spectrometer that can be associated with at least one track segment in the precision muon chambers. This is particularly useful for low $p_T$ muons reaching only the innermost layer of the MS.
- **Calorimeter-tagged (CaloTag) muons** are tracks in the inner detector that can be associated to an energy deposit in the calorimeter as expected from a minimum ionizing particle.

The highest purity is achieved by the CB muons. Therefore, the muons used throughout the presented analysis are STACO muons. The calorimeter-tagged muons have the lowest purity but can be useful to regain acceptance in the uninstrumented regions of the MS.

The reconstruction efficiency of combined muons relies on the ability to form an independent MS track. This is limited especially at $\eta \sim 0$, where the muon detector is only partially equipped because of the needed space for the services of the inner detector and the calorimeter. But also in the transition region between the barrel part and the end caps at $|\eta| \sim 1.2$ the reconstruction is confined because only one muon chamber is passed, which averts the stand-alone momentum measurement. In order to measure the reconstruction efficiency of combined muons, one needs to determine the product of the efficiency in the ID, the efficiency in the MS and the matching efficiency.

This is again performed with the tag-and-probe method using $Z \to \mu\mu$ events. One combined muon serves as tag, while the choice of the probes defines the particular efficiency that is to be measured. The fraction of MS tracks (SA or CB) associated to an ID track within an $\Delta R < 0.05$ gives the efficiency in the inner detector. By matching CaloTag probes to MS tracks within $\Delta R < 0.01$ on the other hand, the MS and matching efficiency can be determined. In this case, the CaloTag probes are favoured over ID probes because of additional background rejection due to the calorimeter information used.

Figure 39 shows the full reconstruction efficiency of STACO muons as a function of $p_T$ and $\eta$. The drops at $\eta \sim 0$ and $|\eta| \sim 1.2$ are clearly visible and illustrate the limited muon chambers coverage, which is the main motivation for the other muon reconstruction types.

Similar to the electrons described in the previous section, the muons searched for in this analysis are expected to be isolated. Isolation criteria help to reject muons emerging from hadron decays as they often appear within so called jets, which are the subject of the following section. Analogously to the electron case, the isolation variables are decoupled from the object reconstruction to maintain the flexibility for the various analyses performed at the ATLAS experiment. The track-based isolation is defined as the scalar sum of the transverse momentum of the tracks in a cone of size $R_{iso}$ without the

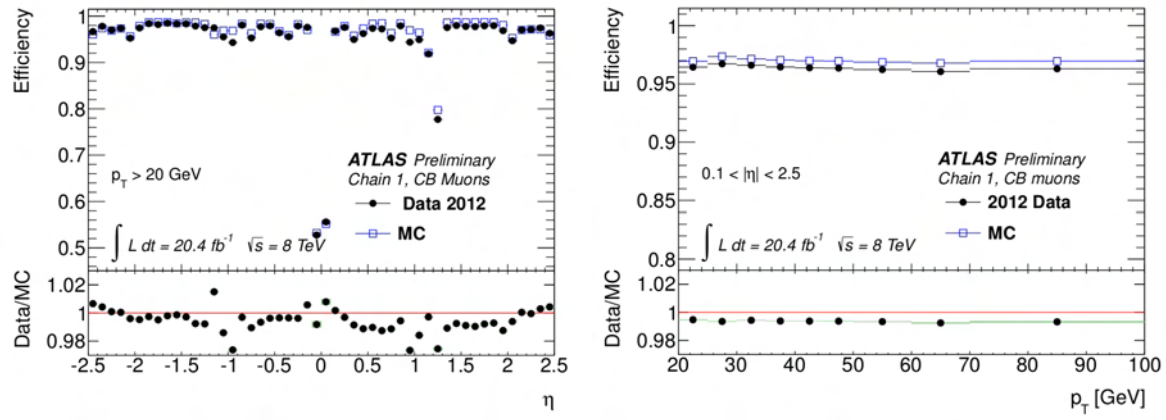



Figure 39: Reconstruction efficiency for STACO muons as a function of $\eta$ for muons with $p_T > 20$ GeV on the left, and on the right as a function of the $p_T$ for muons within $0.1 < |\eta| < 2.5$ [66]. The panel at the bottom shows the ratio between the measured and predicted efficiencies.

considered muon candidate. Accordingly, the calorimeter-based isolation is given by the sum of the transverse energies in the calorimeter cells in a cone of radius $R_{iso}$ excluding the cells of the muon candidate, after correcting for the leakage and pileup contributions. It is convenient to define and measure the relative isolation given as the division by the probe muon's $p_T$ and found to be well modelled by the Monte Carlo simulation. The muons used in this analysis are required to be isolated fullfilling both relative track and calorimeter isolation with a cone size of $R_{iso} < 0.3$.

The momentum scale and the resolution of the muons can be extracted from the width of the di-muon mass distribution in $Z \to \mu\mu$ decays. A good approximation of the fractional momentum resolution is given by:

$$\frac{\sigma(p_T)}{p_T} = a \oplus b \cdot p_T \tag{4.5}$$

where the constant term $a$ describes multiple scattering contributions, while term $b$ describes the intrinsic resolution caused by the spatial resolution and potential misalignment between the two distant detector components. With a template fit technique performed in 16 $\eta$ bins, a $p_T^{corr}$ correction is achieved which also includes the momentum scale correction. The result is illustrated in figure 40 showing the invariant di-muon mass distribution before and after the corrections to the Monte Carlo simulation. For the different detector regions, the resolution ranges from 1.5 to 3 GeV and a corresponding smearing is applied. The measured momentum scale is corrected by $\sim 0.1\%$.

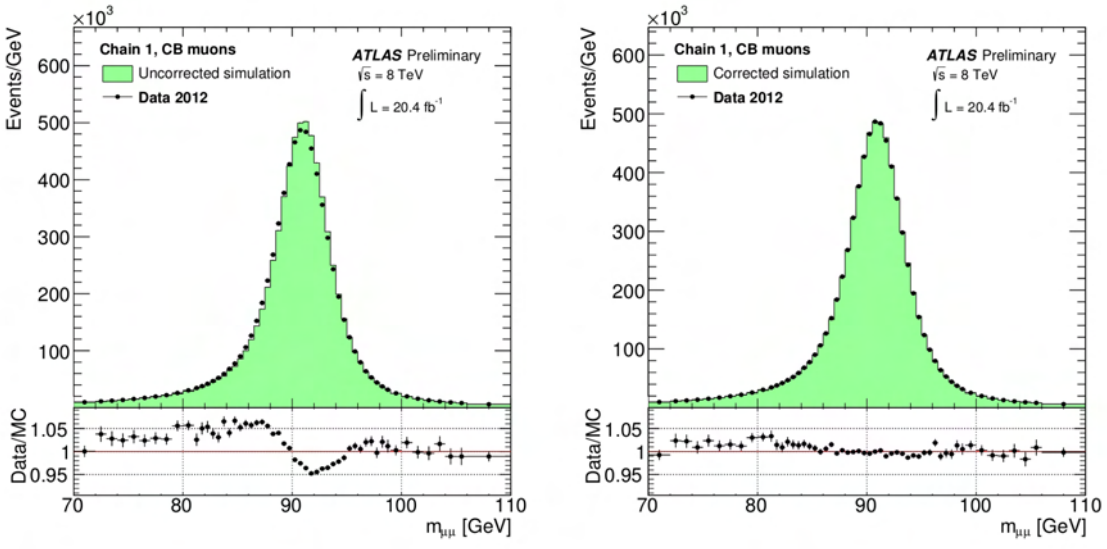



Figure 40: Di-muon invariant mass for STACO Combined muons with $p_T > 25$ GeV [66]. The plot shows the invariant mass for 2012 data and for the simulation of $Z \to \mu\mu$ plus background events. No corrections are applied on the left plot while smearing and scale corrections are applied to the plot on the right. The corrections have been derived from the full 2012 dataset.

### 4.2.3 Jets

Unlike the clearly defined leptons presented so far, jets are localized streams of particles resulting from the fragmentation of hadronized partons. A jet is a general term for a highly active object, usually containing a lot of tracks by the charged particles involved and a calorimeter shower due to multiple interactions and decays. So the jet reconstruction needs to be treated carefully. Since nearly every detector part is involved in the procedure and has to be adjusted to each other, jets typically introduce the dominant source of experimental uncertainty wherever they appear.

The method to reconstruct and calibrate jets used for this thesis starts with a topological cluster algorithm as described in [46]. It associates the energy deposits in the calorimeter to objects with transverse momentum $p_T$ and $\eta$ and $\phi$ coordinates. These clusters are seeded by calorimeter cells with a signal-to-noise (S/N) greater than four. Then their neighbouring cells are added iteratively if they exhibit a lowered threshold of $S/N > 2$. Finally, the nearest neighbours are added without any threshold. They are also called guard cells. If several local maxima are found, the clusters can be split up accordingly. The obtained topological clusters are fed to a sequential jet finder searching for the smallest distance between two clusters $d_{ij}$ or between one cluster and the beam axis $d_{iB}$. Sequential refers to the fact that it goes through the complete list of clusters choosing to either combine them or treat them as separate objects. If $d_{iB}$ is found to be smallest, then cluster $i$ is labeled as individual jet and removed from the list. However, if the

smallest distance is found to be $d_{ij}$, the two clusters $i$ and $j$ are removed from the list and their combination is added to it. This enquiry is repeated until no more clusters remain in the list. The general metric used for this purpose is given by:

$$d_{ij} = min(k_{t,i}^{2p},\ k_{t,j}^{2p}) \cdot \frac{\Delta R_{ij}^2}{R^2} \tag{4.6}$$

$$d_{iB} = k_{t,i}^{2p} \tag{4.7}$$

where there are different variants of concrete algorithms. The quantity $k_t$ is the transverse momentum of the considered cluster, $\Delta R$ the angular distance of the clusters and $R$ a fixed distance parameter. The main difference of the available algorithms lies in the choice of the exponent $p$. Set to zero, it is called Cambridge-Aachen method. If it is chosen to be one, it is simply called $k_t$ algorithm. In case of the presented analysis, the exponent is set to $p = -1$ and referenced as anti-$k_t$ algorithm [68]. Furthermore, the distance parameter is chosen as $R = 0.4$. The advantage of the anti-$k_t$ algorithm is that it prefers to cluster soft activities to large energy deposits rather than combining close-by soft contributions. So high $p_T$ jets tend to accumulate surrounding soft radiation into a conical shape which is insensitive to fluctuations in the soft activity that would be very hard to simulate. Therefore it is infrared and collinear safe. An example of the jet clustering with the anti-$k_t$ algorithm is illustrated in figure 41. It shows a parton level event with a few thousand soft particles.

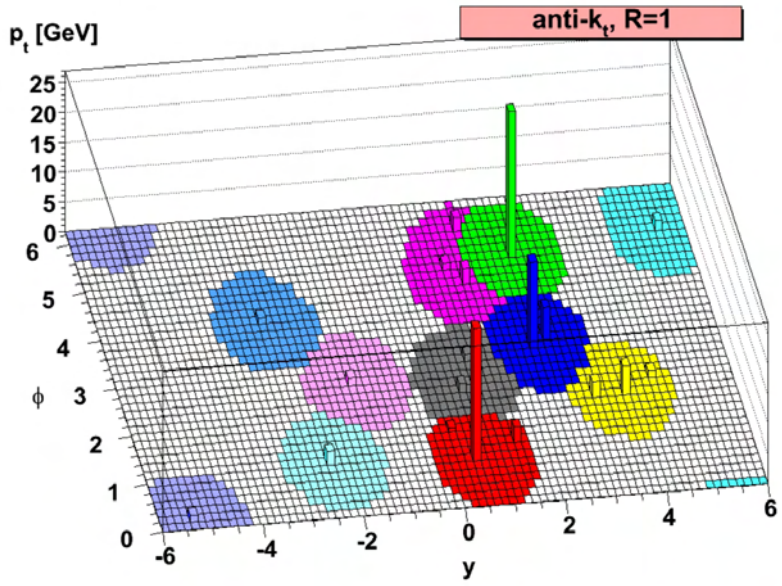



Figure 41: A parton-level event (generated with Herwig[18]), together with many random soft radiations, clustered with the anti-$k_T$ algorithm, illustrating the active catchment areas of the resulting hard jets [68].

A large fraction of jets is expected to originate from pileup interactions rather than from the hard scattering to be analysed. So it is crucial to distinguish hard scatter jets from pileup jets. Since the direction measured in the calorimeters is not precise enough to allow the association to a vertex, inner detector track information is exploited by defining the so called jet vertex fraction (JVF):

$$JVF = \frac{\sum\limits_{tracks\ PV} p_T}{\sum\limits_{all\ tracks} p_T} \tag{4.8}$$

The sum in the denominator runs over all tracks associated to the jet, whereas in the numerator, it runs only over the jet tracks associated to the primary vertex. So the value is close to one, if many tracks emerge from the primary vertex. Or else, pileup jets feature very small values of JVF. But since it relies on the track reconstruction, the jet vertex fraction is only defined within the inner detector coverage of $|\eta| < 2.5$, whereas the jet reconstruction in the calorimeters is highly efficient up to $|\eta| < 4.5$.

A very important point is the jet energy calibration [69]. Initially the calorimeter has been calibrated with electron test beams to give the correct response for electro-magnetic showers. This is called the EM scale. A hadron of the same energy like an electron for example is typically measured with 30% less of its true energy deposit due to the non-compensating nature of the ATLAS calorimeter. This of course has to be adjusted. Starting from the EM scale, a local cluster weighting (LCW) method classifies electro-magnetic and hadronic clusters. The correction is based on single pion Monte Carlo simulation. A dedicated pileup correction subtracts the average additional energy from the measured energy. The correction constants are obtained from in-situ measurements on min-bias data. Also the jet direction is corrected to point to the primary vertex instead of the geometrical center of the ATLAS detector. This recalibration of the four-momentum improves the angular resolution. Further energy and direction corrections of the reconstructed jet are derived from MC truth comparisons. Moreover, some poorly instrumented detector regions tend to reconstruct lower energy jets. This effect is very small on average but not negligible in the transition regions. It is corrected for with the so called jet-$\eta$ calibration. After the application of the full calibration scheme, the resulting objects are referred to as LCW jets.

The jet energy resolution is the result of energy loss in the dead material and fluctuations in the hadronic showers. It is determined from in-situ measurements of the jet response asymmetry in di-jet events [70]. Due to the momentum conservation in the transverse plane, the jets should be balanced. A fitted Gaussian $\sigma_A$ characterizes this asymmetry and the relative jet resolution is given by:

$$\frac{\sigma(p_T)}{p_T} = \sqrt{2} \cdot \sigma_A \tag{4.9}$$

under the balance assumption. For jets within $20 < p_T < 80$ GeV and $|y| < 2.8$, the effect amounts to 14%. The data is in good agreement with the Monte Carlo simulation.

#### Flavour Tagging

In many cases the substructure of jets is not of interest. For some processes however, the content of the jet objects does matter quite a lot, in the context of this thesis particularly for the selection of $t\bar{t}$ events. The heavy flavour process enters the analysis as background that needs to be rejected efficiently. Since the top quarks almost instantly decay into bottom quarks, the so called b-tagging algorithms [71] are used for the identification of jets originating from b-quarks. The efficiency measurement is described in detail in [72].

B-hadrons have a relatively long life time of $\sim 1.5$ ps. Hence, they typically cover a distance of a few mm until they decay further. This results in a displaced secondary vertex and corresponding impact parameters. To identify these, excellent track and vertex reconstruction are crucial.

At ATLAS, there are three different types of algorithms meeting this task. The first is called IP3D and uses the significance of the transverse and longitudinal impact parameters. The second class is reconstructing the displaced vertex by exploiting the track based invariant mass of the vertices and the flight length significance. Depending on the particular criteria cuts, they are labeled SV0 and SV1. Although they provide a low mistag rate, their efficiency is limited. The third algorithm is called Jet Fitter. It tries to reconstruct the full decay chain of b- and c-quarks with a multivariate technique. An artificial neural network utilizes the best performing track and vertex variables to build a likelihood for the jet to be b/c, light flavour or gluon initiated. In order to exploit high rejection and efficiency, all these methods are further combined into another neural network with proper treatment of the input correlations. This approach is labeled MV1. The b-tagging efficiency versus the light jet rejection for the different algorithms is shown in figure 42. Various working points can be picked from this distribution. For the presented analysis, the b-jet efficiency is chosen to be 85%.

### 4.2.4 Missing Transverse Energy

Due to the conservation of momentum in the transverse plane, the vector sum of all reconstructed particles in an event is expected to vanish. But particles as neutrinos for example are leaving the experiment undetected, carrying away momentum. This leads to an imbalance in the summation of the detected momenta and can therefore serve as indirect measurement. So the missing transverse momentum is defined as:

$$\vec{E}_T^{miss} = -\sum \vec{p}_T \tag{4.10}$$

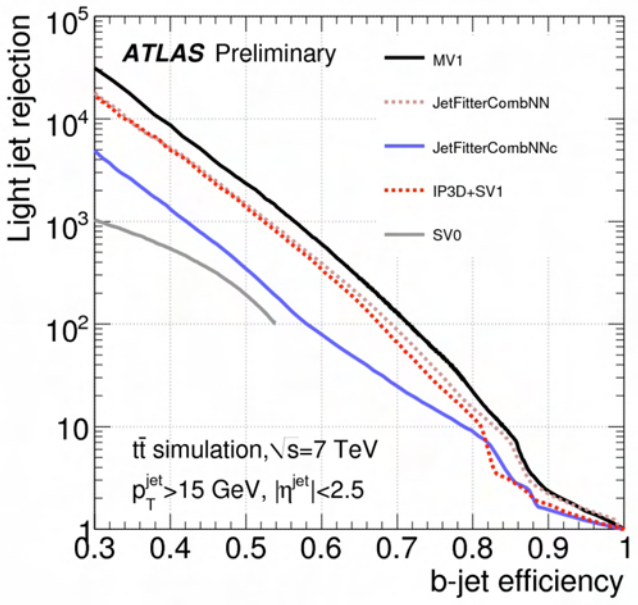



Figure 42: The rejection of light flavour jets as a function of b-jet tagging efficiency in a sample of simulated $t\bar{t}$ events for various b-jet identification algorithms [71].

More common is the phrase missing transverse energy symbolised as $E_T^{miss}$, $\not{E}_T$ or simply MET. However, the synonymous usage is inaccurate, since it is only true for massless particles. The missing transverse momentum relies on the reconstruction of all physics objects, the main ingredients for the determination are the energy deposits in the calorimeter as well as the muon tracks and the inner detector tracks, to account for low $p_T$ objects not well measured in the calorimeter. So it is essential to minimize the effects of limited detector coverage and dead regions, the finite resolution, noise and cosmic ray or beam halo muons. To achieve this, two different concepts can be followed. Either the calculation is based on the calorimeter information or the missing transverse momentum is build from the track parameters. Detailed information on the reconstruction and the efficiency measurement can be found in [73]. The main features are described below.

### Calorimeter based MET

One approach of the missing transverse energy calculation is based on the calorimeter information. Starting from topological clusters, the calorimeter cells are calibrated depending on the object they are associated to. The specific order of the calibration begins with electrons, followed by photons and hadronically decaying tau-leptons. Then the jets are calibrated as described in the previous section before taking charge of the muons. To be considered for this procedure, the leptons have to fulfill certain quality requirements. Only central electrons of $p_T > 10$ GeV passing the medium cut-based identification criteria and the muons of the STACO reconstruction chain with $p_T > 6$ GeV covered by

the inner detector are taken into account. Additionally, the requirement on the impact parameter $z_0 \cdot \sin\Theta < 1$ mm is imposed. The jets are reconstructed with the anti-$k_T$ algorithm and a size of $R = 0.4$. Cells that are not associated to any physics object are summed up to the so called CellOut term. Finally all the separate terms are added up:

$$\not{E}_{x/y} = \not{E}^{e}_{x/y} + \not{E}^{\gamma}_{x/y} + \not{E}^{\tau}_{x/y} + \not{E}^{jets}_{x/y} + \not{E}^{soft\ jets}_{x/y} + \not{E}^{calo\ \mu}_{x/y} + \not{E}^{CellOut}_{x/y} + \not{E}^{track\ \mu}_{x/y} \tag{4.11}$$

The soft jets term covers the low $p_T$ region from 7 GeV to 20 GeV, the jets term everything above 20 GeV. Also the muon term is split into one part accounting for their calorimeter loss and the other for the track $p_T$. In technical jargon, the result of the calibration scheme is also referred to as "METRefFinal". More object-like are the quantities:

$$E_T^{miss} = \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2} \tag{4.12}$$

$$\phi^{miss} = \arctan(E_y^{miss}/E_x^{miss}) \tag{4.13}$$

where equation 4.12 is usually meant when speaking of missing transverse energy.

The performance of the reconstruction of missing transverse energy is quantified in $W \to \ell\nu$ and $Z \to \ell\ell$ events. Due to the neutrinos in the W-decay, these events contain real missing energy, which allows to study the $E_T^{miss}$ scale and uncertainty. The difference in data and simulation are found to be less than 5% and are generally in good agreement. On the other hand, a Z-boson decaying into two leptons should not provide missing energy. Nevertheless it arises from imperfections of the detector and the reconstruction schemes. That way, the width of $E_{x/y}^{miss}$ and therefore the $E_T^{miss}$-resolution can be determined. Figure 43 shows the $E_T^{miss}$ distribution observed in $Z \to \mu\mu$ events.

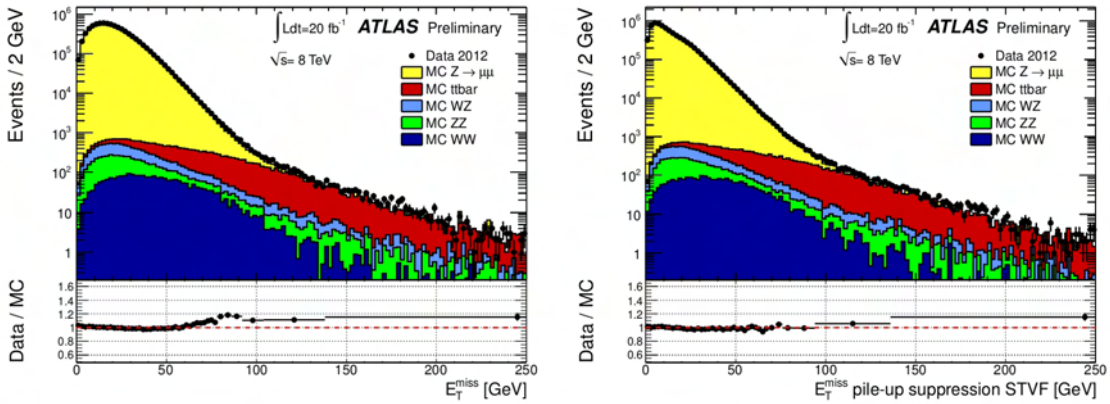



Figure 43: Distribution of $E_T^{miss}$ as measured in a $Z \to \mu\mu$ data sample without pileup correction on the left, and on the right with pileup correction using the STVF method [73].

The resolution is degraded with increasing pileup activity due to the limited calorimeter response. Most affected are the soft jet and the CellOut terms, sometimes simply referred to as soft terms. To reduce the impact on the missing transverse energy reconstruction, the so called soft term vertex fraction (STVF) is introduced:

$$STVF = \frac{\sum_{soft\ term\ tracks,\ PV} p_T}{\sum_{soft\ term\ tracks} p_T} \tag{4.14}$$

similarly to the JVF in the previous section. Here, the summation is performed for all tracks that are not associated to any physics object but included in the soft term, distinguishing whether they belong to the primary vertex or not. This measure can be used for the pileup correction. The effect is visible in figure 44 on the left, where several pileup correction methods are compared, the STVF algorithm performing best. The distribution on the right illustrates the nice agreement of data and Monte Carlo simulation.

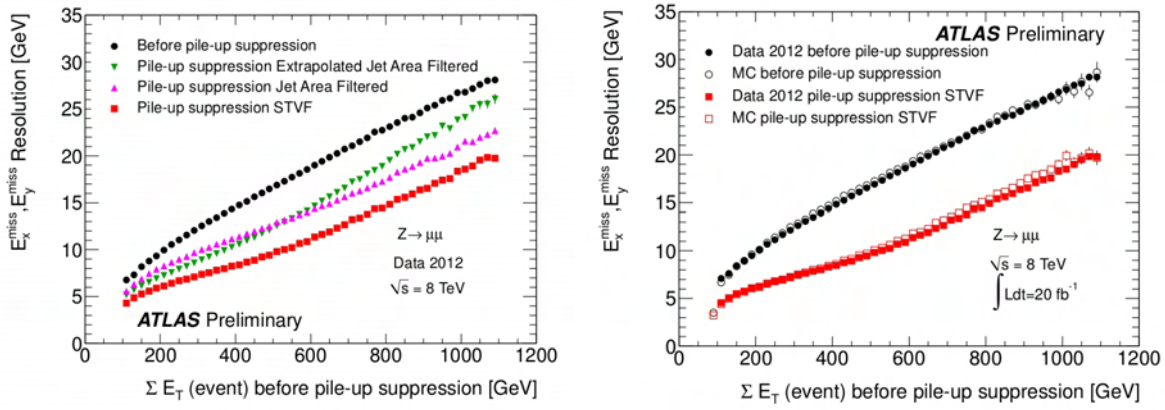



Figure 44: Missing transverse energy resolution as measured in a $Z \to \mu\mu$ data sample [73]. On the left, several pileup suppression methods are compared, while the right hand side shows the data/MC comparison.

The jet term inherits some residual pileup dependence. This can be reduced by scaling the jets with the jet vertex fraction as described in the previous section and leads to a further improvement of the missing transverse energy resolution.

### Track based MET

Another approach to reconstruct the missing transverse momentum is based on the tracking information. The $E_T^{miss,track}$ is defined as the negative sum of the transverse momenta of every track that is is fulfilling a set of quality criteria. This definition is of course restricted to the tracking coverage of $|\eta| < 2.5$. Only tracks with $p_T > 500$ MeV that have been extrapolated to the primary vertex are taken into account. Furthermore, they are required to have at least one hit point in the pixel detector and at least six hits in the SCT. The criteria for the impact parameter with respect to the primary vertex $d_0^{PV}$ and $z_0^{PV} \cdot \sin\Theta$ have to be less than 1.5 mm each. In case the tracks of the leptons fail this set of requirements, but the reconstructed leptons themselves fulfill the set of criteria of the calorimeter based definition above, the tracks are also considered for the track-based calculation. This definition of missing transverse energy reduces the impact of detector effects and mismeasurements and is also less affected by pileup interactions.

# 5

# Search for the Higgs Boson in the $H \to WW \to \ell\nu\ell\nu$ Decay

With the information summarized in section 4 it is now possible to focus on the analysis of the decay channel $H \to WW \to \ell\nu\ell\nu$. Its specific signature forms the basis of the physics objects to search for and therefore opens this chapter. A number of background processes sharing the same or similar decay products has to be taken into account. The second part deals with the collected data sample and the fundamental trigger selection used for the events of interest. Basic quality criteria and cleaning procedures are summarized. The following part then concentrates on the details of the Monte Carlo (MC) simulation which is to be compared to data, in order to analyse the dataset statistically. This includes the ATLAS computing framework and the detector simulation as well as specific pileup simulation and further reweightings to account for effects that are not properly simulated. A summary of the MC generation for all considered signal and background processes is given. The application of basic object criteria to data and MC concludes this chapter and finally builds the bridge from general physics objects to the actual event selection in section 6, where the special signature of the signal process is used to separate it from the background contributions.

## 5.1 Signature of the $H \to WW \to \ell\nu\ell\nu$ Final State

The characteristic signature of this search channel results from the leptonic decay of the two W-bosons leading to two isolated, oppositely charged leptons ($\ell = e$ or $\mu$) and two neutrinos. While the final state leptons enable an efficient differentiation from hadronic backgrounds, the neutrinos escape the ATLAS detector without further interactions and can not be detected. But nevertheless, they leave a trace of large missing transverse energy which is an important handle. The leading order Feynman diagram of a Higgs

boson produced via gluon fusion and decaying through a W-pair into two leptons and neutrinos is shown in figure 45.

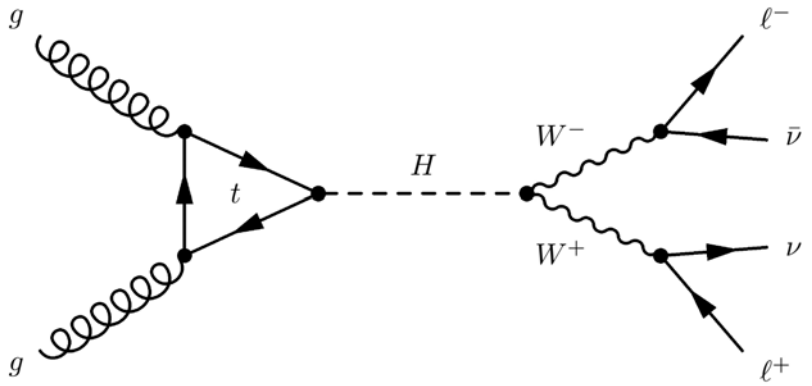



Figure 45: Leading order Feynman diagram of a Higgs boson produced via gluon fusion. The decay into two W-bosons which decay leptonically themselves leads to the special signature of two isolated, oppositely charged leptons and large missing transverse energy in the ATLAS detector.

Depending on the particular W-decays, the final state can be made of either $e^-\bar{\nu}_e e^+\nu_e$ or $\mu^-\bar{\nu}_\mu\mu^+\nu_\mu$ as well as $e^-\bar{\nu}_e\mu^+\nu_\mu$ and $\mu^-\bar{\nu}_\mu e^+\nu_e$ . They are all referred to as dilepton states. But in the following, the short notations $ee$, $e\mu$ and $\mu\mu$ are used since the electrical charge concerns only the lepton pairs and the neutrinos are not detected separately. While the $W \to e\nu$ and $W \to \mu\nu$ decays make up most of the signal, the W-decay via an intermediate $\tau$-lepton into an electron or muon is also implicitly included. There may also be additional jets in the final state, whether originating from the VBF production process or simply from gluon radiation. So the separation into three disjoint search channels H+0 jets, H+1 jet and H+$\geq$2 jets according to their jet multiplicity per event is helpful in order to maximize the sensitivity.

Unfortunately, it is not possible to reconstruct the invariant mass of all decay products to observe the Higgs decay because the neutrinos escape the detector and carry away the information. Therefore, a transverse mass $m_T$ [74] computed from the leptons and the missing transverse momentum is considered to test the presence of a signal. But the resolution is of course affected due to the neutrinos.

Nevertheless, there is one particular characteristic in the dilepton final state related to the presence of the neutrinos in the decay chain. The spin correlation [75] shown in figure 46 leads to a unique feature in the signature of the $H \to WW \to \ell\nu\ell\nu$ channel. The Standard Model Higgs boson is a spin-0 particle. Its decay into two spin-1 W-bosons allows only a combination of (+1, –1). The further leptonic W-decays into two fermions each leave the combinations of (+1/2, +1/2) and (–1/2, –1/2) for the lepton-neutrino pairs respectively. Imposed by the V–A structure of the weak interaction, there exist only

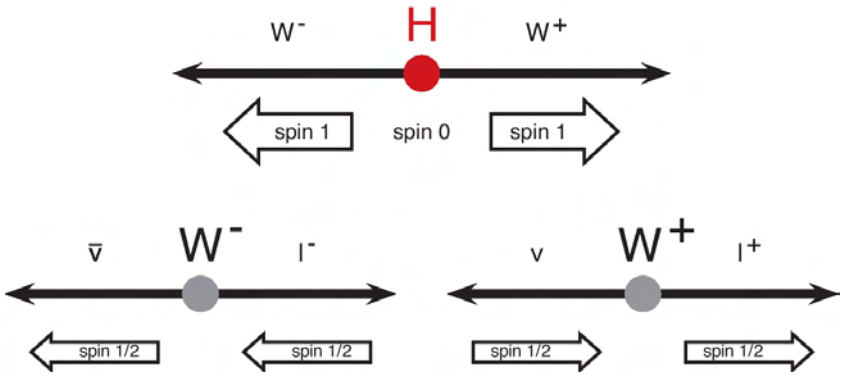



Figure 46: Spin correlation for the SM $H \to WW \to \ell\nu\ell\nu$ decay. While the two leptons are emitted in one direction, the two neutrinos are expected in the opposite direction.

left-handed neutrinos and right-handed anti-neutrinos, meaning that the helicity[15] state forces the neutrinos into a special direction with the lepton partner back-to-back. The possible spin combinations result in the neutrinos preferably travelling in one direction and the lepton pair in the opposite direction. Of course, this argument accurately only holds in the rest frame of the W-bosons. But also in the laboratory frame, the leptons tend to have small opening angles with large missing transverse energy in the opposite direction, since the relatively low mass $m_H = 125$ GeV of the Higgs boson prevents large boosts.

## 5.2 Backgrounds of the $H \to WW \to \ell\nu\ell\nu$ Process

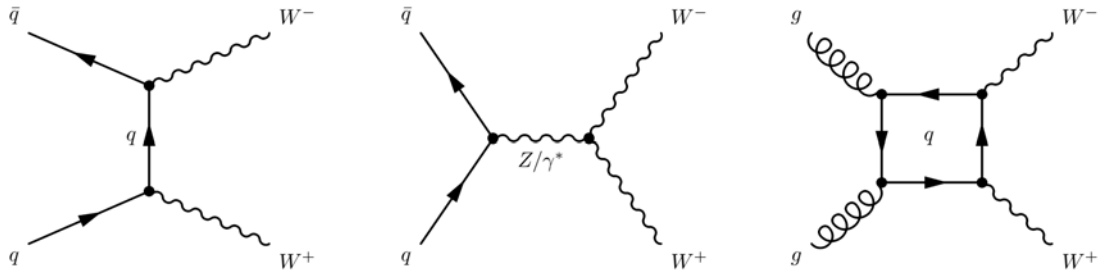
Although the signature of the $H \to WW \to \ell\nu\ell\nu$ final state is quite rare, it is unfortunately not unique. There are various background processes that either share the same final state particles or mimic them due to misidentification, additional pileup remnants or the limited detector coverage. Since the latter of those typically show obvious differences, they are called reducible backgrounds. However, if the final state is exactly the same as the signal, refined analysis techniques need to be developed. This kind of background processes is called irreducible.

The non-resonant Standard Model WW production is such an irreducible process in the presented analysis. For this reason, it is the dominant background source. But also other diboson mechanisms such as WZ, ZZ or W$\gamma$ need to be accounted for. Processes that involve either $t\bar{t}$-pairs or single top-quarks mainly decay via W-bosons and enter the analysis as well. Last but not least, $Z/\gamma^*$ and W production with associated jets make up a large fraction of the irreducible backgrounds. The main backgrounds of the $H \to WW \to \ell\nu\ell\nu$ channel are briefly described in the following sections.

[15] Projection of the spin on the direction of motion

### 5.2.1 Standard Model WW Production

The dominant background contribution to the $H \to WW \to \ell\nu\ell\nu$ arises from the Standard Model $W^+W^-$ production. The leptonic decay of the two W-bosons leads to the exact same final state as the signal and is therefore an irreducible background. Figure 47 shows the leading order Feynman diagrams for the s- and t-channel, as well as the gluon fusion production mechanism. At the LHC it is mainly produced in quark-antiquark annihilations dominated by the t-channel graph. The s-channel contribution involves the triple gauge coupling and makes up $\sim 10\%$ of the total production, while the gluon fusion accounts for only $\sim 3\%$.




(a) (b) (c)

Figure 47: Leading order Feynman diagrams for the Standard Model production of $W^+W^-$ pairs at the LHC: (a) illustrates the quark initiated t-channel, the s-channel is shown in (b). The production via gluon fusion mediated by a quark loop is shown in (c).

The $W^+W^-$ production cross section has been measured in the leptonic decay channels at the LHC with the 2012 dataset of 20.3 fb$^{-1}$ collected at $\sqrt{s} = 8$ TeV with the ATLAS detector [76]. The measured value $\sigma^{tot}_{WW} = 71.4\ ^{+1.2}_{-1.2}$ (stat) $^{+5.0}_{-4.4}$ (syst) $^{+2.2}_{-2.1}$ (lumi) pb. This result is about 22% higher than the theoretical prediction of $58.7\ ^{+3.0}_{-2.7}$ pb calculated using MCFM [77], which provides NLO QCD calculation for the $q\bar{q} \to WW$ and the LO calculation for the $gg \to WW$ process. The observed enhancement of the cross section has a statistical significance of 2.1 $\sigma$, illustrated in figure 48. The contribution from the $gg \to H \to WW$ process is already included to obtain the total Standard Model cross section.

In this thesis, the normalization of the WW process is performed within a so called control region which uses a dedicated selection to enrich the background and deplete the signal process to remain independent of the NLO prediction. Unlike the Higgs boson in the signal process, the quark-antiquark initial state can come in three different spin states –1, 0 and +1. So only one of them exhibits a similar spin correlation as explained in figure 46, while the ±1 states do not possess this feature. When averaging over all three states, considerable differences in the invariant mass and angular separation distributions of the two leptons are expected. Further details follow in section 6.5.1.

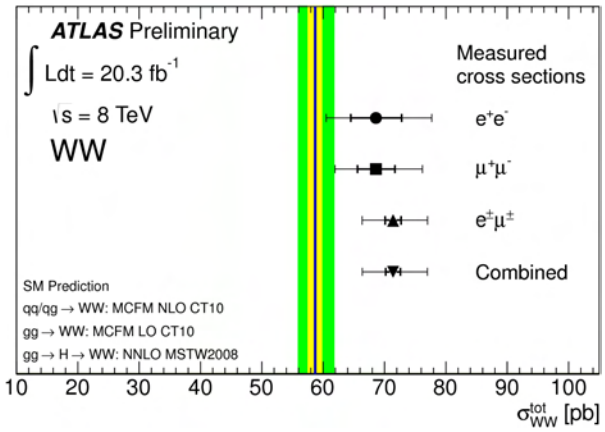



Figure 48: Comparison between predicted and measured WW production cross section [76] in $ee$, $\mu\mu$, $e\mu$ and combined channels. The yellow and green shaded bands represent the PDF and total theoretical uncertainties respectively. The filled symbols show the measured total cross section with the statistical and total uncertainty.

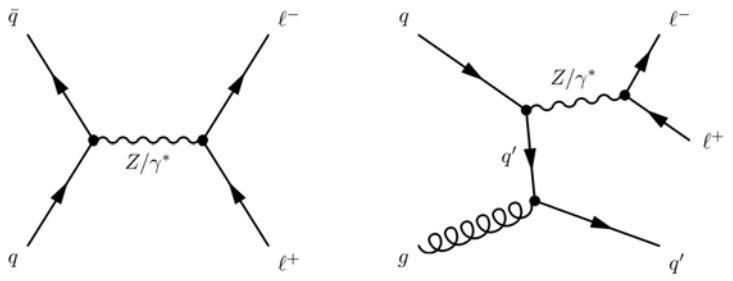
### 5.2.2 $WZ/ZZ/W\gamma$ Production

Another set of background processes in this context is called diboson [78] production, referring to WZ, ZZ and W$\gamma$. Note that the WW production of the previous section is explicitly excluded from this set since it stands out as the dominant background process, although it actually belongs to the same category. But the WZ, ZZ and W$\gamma$ processes significantly differ from the $H \to WW \to \ell\nu\ell\nu$ signal. Although they contain isolated leptons and missing transverse momentum emerging from leptonic W-decays, they hardly mimic the exact signature. A simple veto on a third or more charged leptons already reduces their contribution significantly. Therefore, these processes are expected to be small and are determined purely by Monte Carlo simulation.

### 5.2.3 $Z/\gamma^*$+jets Production

The Drell-Yan process has a large cross section at hadron colliders such as the LHC. With a subsequent leptonic decay of the Z-boson, a background of two isolated, oppositely charged leptons arises. True missing transverse momentum appears only in the case of $Z/\gamma^* \to \tau\tau$ decays, though. But mismeasurement of the leptons and especially of additional associated jets can also result in a wrong reconstruction. And as mentioned before, pileup contributions can lead to a significant degradation of the energy measurement. Figure 49 illustrates the production mechanisms of the $Z/\gamma^*$+jets background process with and without associated jets.

The $Z/\gamma^*$ decays into two same flavour leptons and therefore mostly affects the ee and $\mu\mu$ channel. However, the decay via two $\tau$ leptons leads to a non-negligible contribution




(a) (b)

Figure 49: Feynman diagrams of the Drell-Yan production process in leading order with (b) and without an associated jet (a).

to the e$\mu$ channel as well. In order to reduce the background contribution, same flavour lepton pairs with an invariant mass close to the Z boson mass are removed from the analysis. This is often referred to as Z-veto. With stringent requirements on the missing transverse energy, the remaining fraction of the $Z/\gamma^*$+jets background can be reduced further. The main problem is the large cross section. Compared to the signal process, the total leptonic production cross section of the Drell-Yan background is about four orders of magnitude larger. In [79], the total $Z/\gamma^*$ production cross section times the leptonic branching ratio is measured at $\sqrt{s} = 7$ TeV with the ATLAS detector. The combined electron and muon channels are exploited in the window of the invariant mass of $66 < m_{ll} < 116$ GeV. Within the assigned uncertainties, the measured value $\sigma^{tot}_{Z/\gamma^*} \times BR(Z/\gamma^* \to \ell\ell) = 0.82 \pm 0.06(\text{stat}) \pm 0.05(\text{syst}) \pm 0.09(\text{lumi})$ nb is found to be in good agreement to the prediction of $0.96 \pm 0.05$ nb, which includes NNLO QCD corrections. Similar to the case of the WW-background, the predicted rate of events in this analysis is normalized in a dedicated control region, as explained in detail in section 6.5.4.

### 5.2.4 Production of $t\bar{t}$ and Single Top

A number of background processes to the $H \to WW \to \ell\nu\ell\nu$ analysis are related to top-quarks. In the proton proton collisions at the LHC, top-antitop pairs are produced via gluon-gluon and quark-antiquark annihilation shown in figure 50 (a), (b) and (c). The gluon induced production dominates with its 90% contribution to the total $t\bar{t}$ cross section. Single top-quarks can be produced via three different mechanisms. Either in the decay of a virtual W-boson (s-channel), in the exchange of a W-boson (t-channel) or in association with a W-boson (Wt). The Feynman diagrams for those modes are shown in figure 50 (d), (e) and (f).

Since the top-quarks almost exclusively decay into a bottom-quark and a W-boson, the top-related backgrounds enter the analysis through the leptonic W-decay. While the $t\bar{t}$ and Wt processes lead to signatures containing a $W^+W^-$-pair plus additional

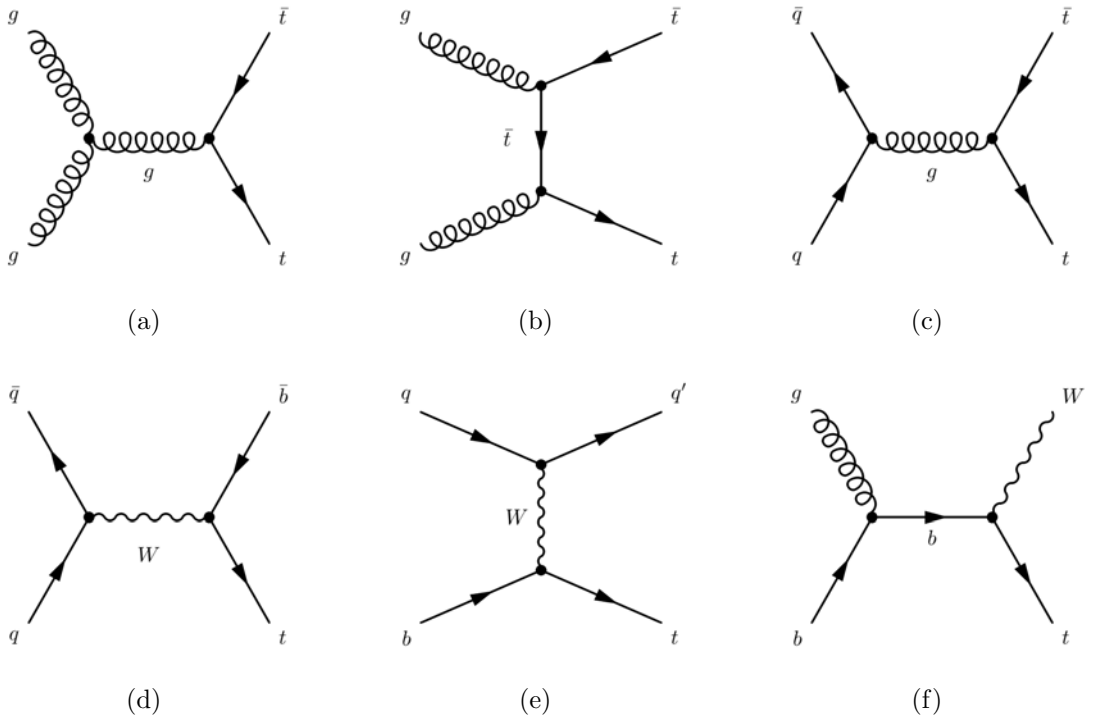



Figure 50: Feynman diagrams of the top-antitop pair production process via gluon annihilation (a), gluon scattering (b) and quark annihilation (c) and the single top-quark production in the s-channel (d) and t-channel (e) contributions. The production in association with a W-boson is shown in (f).

jets, the other single top channels can mimic the signal only due to mismeasurements. In any case, the identification of the top-quark background relies on the associated bottom-quarks. By determining events that contain b-tagged jets, the contribution can be reduced significantly. Therefore, a dedicated control region is used to estimate and normalize the top background described in section 6.5.2.

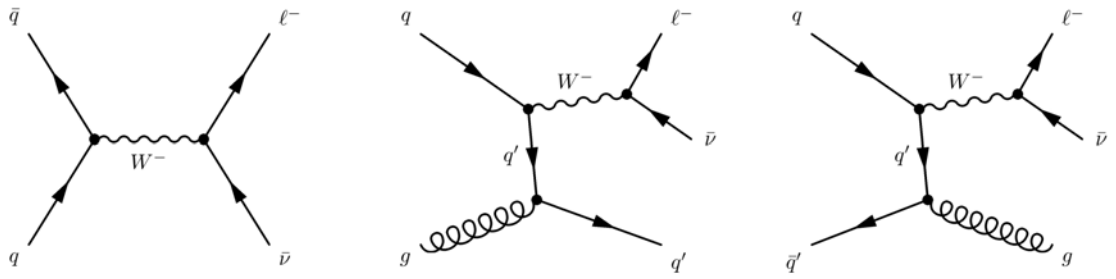
The $t\bar{t}$ cross section has been measured by the ATLAS collaboration in the $\sqrt{s} = 8$ TeV collisions by selecting events that contain two isolated leptons, large missing transverse energy and either exactly one or two b-tagged jets in [80]. The measured value is in good agreement with the theoretical prediction, which is calculated at full NNLO accuracy in the strong coupling constant $\alpha_S$, including the resummation of NNLL soft gluon terms. A similar approach with two leptons, missing transverse energy and exactly one jet in the final state can be used to measure the Wt cross section [81]. With a selection of only one isolated lepton and jets identified to emerge from bottom-quarks, the single top s-channel [82] and t-channel [83] cross sections are measured. Those analyses have been performed on the 2011 dataset at $\sqrt{s} = 7$ TeV. All the results are summarized in table 6 and show agreement between experiment and the predictions, which are calculated for a top-quark mass of $m_t = 172.5$ GeV at NLO with NNLL corrections.

| Process | Measured Cross Section in [pb] | Predicted Cross Section in [pb] |
|---|---|---|
| $t\bar{t}$ | $242 \pm 1.7(\text{stat}) \pm 9.3(\text{syst}) \pm 4.2(\text{lumi})$ | $252.9 \pm 13$ |
| s-chan | $< 26.5$ (at 95% CL) | $4.6 \pm 0.3$ |
| t-chan | $83 \pm 4(\text{stat}) \,^{+20}_{-19}(\text{syst})$ | $64.6 \,^{+3.3}_{-2.6}$ |
| Wt | $16.8 \pm 2.9(\text{stat}) \pm 4.9(\text{syst})$ | $15.7 \,^{+1.3}_{-1.4}$ |

Table 6: Comparison of the measured cross sections of the top-quark pair production and the single top-quark production to the predicted values taken from [80], [81], [82] and [83]. The single top measurements have been performed with $\sqrt{s} = 7$ TeV pp collisions recorded with the ATLAS detector, while the $t\bar{t}$ cross section has been measured for $\sqrt{s} = 8$ TeV.

### 5.2.5 W+jets and QCD Production

The production of a W-boson in association with jets can also contribute to the background of the $H \to WW \to \ell\nu\ell\nu$ channel, if the W-boson decays leptonically and the additional jet gets misidentified as a second lepton. These events contain real missing transverse energy accounting for the neutrino of the W-decay. Figure 51 shows the leading order Feynman diagrams for the W+jets background.




(a) (b) (c)

Figure 51: Feynman diagrams for the production of a W-boson without (a) and with an associated jet (b) and (c).

The likelihood to misidentify a jet as a high-$p_T$ lepton is expected to be quite small, at the order of $10^{-4}$. However, the total production cross section of the W+jets background is so large that it compensates the low misidentification rate. In [79], the product of the production cross section times the leptonic branching ratio is measured with the 2011 data collected from proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. The experimental result of $\sigma_W^{tot} = 9.96 \pm 0.23(\text{stat}) \pm 0.50(\text{syst}) \pm 1.10(\text{lumi})$ nb is compatible with the theoretical NNLO calculations that predict a value of $10.46 \pm 0.52$ nb. Thus, the W+jets background and the signal process are expected to have similar size. Since they also share similar kinematics, the key to suppress the W+jets contribution lies

in the misidentified second lepton. Strong requirements on the track and calorimeter isolation help to distinguish the falsely reconstructed hadronic jets or real non-isolated leptons from the signal process. The Monte Carlo simulation is not expected to model such misidentification problems sufficiently well. That is why in this analysis, both rate and shape of the W+jets background are extracted with a data-driven approach and a separate control region as described in section 6.5.5.

Another contribution to the background processes is QCD dijet production. In this case, both signature leptons are the result of the misidentification of the two jets. Again, the large cross section does compensate the unlikliness to wrongly reconstruct both high-$p_T$ objects in the event. The estimation of the QCD background is explained in section 6.5.6. It is closely related to the techniques developed to extract the W+jets background.

## 5.3 Data Samples

In section 3.1.3, the dataset collected in the $\sqrt{s} = 8$ TeV proton proton collisions in 2012 has already been introduced. Here, the focus lies on the conditions and administrative preparations to be performed in order to set up the analysis. This part starts with the selection of triggers that record the dataset with regard to the signal process. Further considerations on the data quality have to be taken into account before the total integrated luminosity ready for analysis can be determined.

### 5.3.1 Trigger Selection

The trigger selection for $H \to WW \to \ell\nu\ell\nu$ candidate events is based on the signature leptons. A combination of multiple unprescaled single lepton and dilepton triggers has been used to maximize the efficiency of collecting potentially interesting events. Unprescaled means that actually every event fulfilling the trigger requirements has been recorded and none has been condemned in order to reduce the storage space. Thus, the collected dataset corresponds to the measured total integrated luminosity. To prohibit double counting, the combination is performed with a logical or and depends on the lepton channel separation. So an event is taken into account if any of the triggers in the list has fired. The trigger setup of the presented analysis is summarized in table 7.

All triggers evaluated after the event building procedure refer to the third and final step of the trigger chain (see section 3.2.4). The triggered objects are electrons or muons with a certain $p_T$-threshold. For the lowest unprescaled single electron trigger, this threshold is 24 GeV but varies slightly for different $\eta$-regions. It contains a so called hadronic core veto that refuses objects with an energy leakage in the $0.2 \times 0.2$ ($\eta \times \phi$) layer behind the EM calorimeter. A relative track isolation requirement of $\sum_{IDtracks} p_T / p_T^e < 0.1$ within a cone of $\Delta R = 0.2$ is also applied. The medium identification scheme as explained in section 4.2.1 provides further assortment. To recover an efficiency loss at higher $p_T$, a second single electron trigger with a threshold of 60 GeV and medium identification

| Channel | Trigger Setup | Category |
|---|---|---|
| ee | EF_e24vhi_medium1 ‖ EF_e60_medium1 ‖ | single electron |
| | EF_2e12Tvh_loose1 ‖ EF_2e12Tvh_loose1_L2StarB | dielectron |
| $\mu\mu$ | EF_mu24i_tight ‖ EF_mu36_tight ‖ | single muon |
| | EF_mu18_tight_mu8_EFFS | dimuon |
| e$\mu$ | EF_e24vhi_medium1 ‖ EF_e60_medium1 ‖ | single electron |
| | EF_mu24i_tight ‖ EF_mu36_tight ‖ | single muon |
| | EF_e12Tvh_medium1_mu8 | electron-muon |

Table 7: Trigger setup for the different lepton flavour channels. Every channel uses a dedicated combination of single lepton and dilepton triggers.

criteria is combined via an OR-gate. In the dielectron channel, it is possible to expand the search to lower transverse momentum by requiring two electrons with $p_T > 12$ GeV per event. This threshold is again $\eta$-dependent and the hadronic core veto is applied. But both electron objects only need to fulfill the loose identification criteria. A second version of this dielectron trigger has been developed and introduced at a late stage of the 2012 data taking to improve the efficiency in the end cap regions by using a special tracking strategy.

The primary single muon trigger of use has a $p_T$-threshold of 24 GeV as well. It contains a requirement on the relative track isolation of $\sum_{\text{ID tracks}} p_T/p_T^{\mu} < 0.12$ within a cone of $\Delta R = 0.2$ around the muon. Tracks are considered if they exceed a transverse momentum of $p_T > 1$ GeV and fulfill a requirement on the longitudinal impact parameter $\Delta z_0 = |z_0(\mu) - z_0(\text{ID track})| < 6$ mm. The label tight denotes that the triggered muon candidate belongs to the STACO chain (see section 4.2.2). A second single muon trigger with a $p_T$-threshold of 36 GeV secures high efficiencies. In the dimuon channel, a dedicated trigger seeded by candidates above a threshold of $p_T = 18$ GeV performs a full detector scan searching for secondary muons above a threshold of $p_T = 8$ GeV. Thus, the collection of events is expanded to lower transverse momentum. The e$\mu$ channel is set up with a combination of the single lepton triggers mentioned above and a trigger looking for isolated, medium electron candidates above 12 GeV and an additional muon above 8 GeV. The efficiencies for these triggers have been measured with the tag-and-probe method on Z-data. They are designed to deliver 90% (70%) efficiency for electrons and end cap muons (barrel muons) at the mentioned $p_T$-threshold reaching the plateau shortly above. By including the dilepton triggers, the signal efficiency in the different channels improves by 10 – 20% in case of the H+0 jets search. The main contribution comes from the reduction of the transverse lepton momenta.

### 5.3.2 Data Quality

To ensure that the collected data is ready for a reliable physics analysis, the calibration, the alignment and the general condition of all subdetectors have to be confirmed. Only if all relevant components have been working properly under nominal conditions, the events are considered. The ATLAS data quality [84] is monitored online and offline by a dedicated working group. While the states of the different detector components is monitored online, the offline data quality monitoring provides quick feedback on the prompt reconstruction of the data. So any irregularities that undermine the quality and could lead to problems in the later analysis stages are uncovered. The data quality information is condensed into a so called Good Run List[16] (GRL) which filters the problematic data blocks.

The data taking period at $\sqrt{s} = 8$ TeV in 2012 has been summarized in section 3.1.3. After the data quality examination, the dataset collected by the ATLAS detector is reduced from 21.7 fb$^{-1}$ to the final amount of 20 fb$^{-1}$ available for physics analysis. The mean number of inelastic interactions per bunch crossing in 2012 was $\langle\mu\rangle = 20.7$. References [85] and [86] describe the determination and calibration of the luminosity with several different detector devices.

## 5.4 Detector Simulation and Event Reconstruction

This section summarizes the general approach of event simulation of the ATLAS collaboration. To obtain realistic predictions that are directly comparable with the measured events, the Monte Carlo simulation needs to take into account the detector effects. Thus, two different types of ATLAS detector simulation are available and used in this analysis. The following part lists the particular Monte Carlo samples of the signal and background processes considered. After that, some dedicated techniques to further improve the simulation with respect to the data taking conditions are described. The last part of this section focuses on the basic object criteria invoked by the considerations of the signal and background processes introduced at the beginning of this chapter.

### 5.4.1 ATLAS Detector Simulation

The events created by Monte Carlo generators need to be modified to match the output format of the data recorded by the ATLAS detector. At that step, the detector effects are taken into account. The simulation program that performs this transition is integrated into a software framework called Athena [87]. The full simulation infrastructure is described in detail in [88]. The exact distribution of the detector material is simulated with the Geant4 toolkit [89]. It emulates the response to the particles

[16] The presented analysis uses version: data12_8TeV.periodAllYear DetStatus-v61-pro14-02_DQDefects-00-01-00_PHYS_StandardGRL_All_Goo_overlapd.xml

and events propagated through the ATLAS detector, using an accurate physics model of the interactions with the detector material and the simulation of energy measurement as well as the readout electronics. These detailed considerations are expected to lead to precise results. Since the procedure involves the full detector information available, it is referred to as full simulation. But this level of detail comes with a huge amount of computing time. The simulation of an average ATLAS event can take several minutes. So the amount of events and processes run through the full simulation is limited and has to be planned carefully.

About 75% of the time for the full simulation of the ATLAS detector is consumed by the electromagnetic calorimetry. That is why a faster simulation procedure called ATLFAST II has been developed to complement the full simulation studies with a large number of events. A tradeoff between computation speed and sufficient detail has been found in a parameterization of the electromagnetic and hadronic showers in the calorimeters, constructed from the energy profiles of fully simulated single photon and charged pion events. This fast simulation procedure supports the full simulation of the inner detector and the muon system reducing the processing time by a factor of 10 to 20.

### 5.4.2 Monte Carlo Simulated Samples

The signal processes and the relevant background processes are simulated using a variety of different Monte Carlo generators. A list is given in table 8 together with the product of the cross section times the branching ratio. For many processes, separate generators are utilized for the hard scattering, the parton shower and hadronization stages as well as decay and underlying event simulation. The hard scattering for the gluon fusion and vector boson fusion signal processes for example are generated with the POWHEG [90] program whereas showering, hadronization, decay and underlying event are committed to the PYTHIA [17] program while the branching fraction for the Higgs boson decay comes from HDECAY [91]. The POWHEG method is particularly designed to deliver NLO QCD calculations for interfacing parton shower generators. PYTHIA's showering algorithms provide FSR and ISR including coherence effects matched to NLO matrix elements for gluon emission. Hadronization and decay apply the Lund string model. The same combination is also used for some of the major background processes such as $q\bar{q}/g \to WW$, $t\bar{t}$ or single top. Only the single top t-channel uses the ACERMC [92] program for the generation of the hard process instead, which is a tree level generator dedicated for Standard Model processes at LHC collisions. The associated Higgs boson production modes have been generated solely with the PYTHIA package without input from other programs.

Another general purpose Monte Carlo event generator is called SHERPA [93]. It uses multi-parton matrix elements with QCD parton cascades and the fragmentation is based on the phenomenological cluster model. The so called CKKW matching scheme [94] is used as factorization prescription for the merging of collinear radiation. A simple multiple interaction model generates the underlying event. The SHERPA generator is

| Process | MC generator | $\sigma \times BR$ in [pb] |
|---|---|---|
| ggF $H \to WW$ | Powheg + Pythia | 0.435 |
| VBF $H \to WW$ | Powheg + Pythia | $36 \cdot 10^{-3}$ |
| WH/ZH $H \to WW$ | Pythia | $25 \cdot 10^{-3}$ |
| $gg \to WW$ | GG2WW + Herwig | 0.20 |
| $q\bar{q}/g \to WW$ | Powheg + Pythia | 5.68 |
| QCD $WW$ + 2 jets | Sherpa | 0.568 |
| EW $WW$ + 2 jets | Sherpa | 0.039 |
| $t\bar{t}$ dileptonic | Powheg + Pythia | 26.6 |
| $tW/tb$ leptonic | Powheg + Pythia | 4.17 |
| $tqb$ leptonic | AcerMC + Pythia | 28.4 |
| inclusive $W$ | Alpgen + Herwig | $37 \cdot 10^{3}$ |
| inclusive $Z/\gamma^*$ | Alpgen + Herwig | $16.5 \cdot 10^{3}$ |
| EW $Z/\gamma^*$ | Sherpa | 5.36 |
| $W(Z/\gamma^*)$ | Powheg + Pythia | 12.7 |
| $W(Z/\gamma^*)$ ($m_{Z/\gamma^*} < 7$ GeV) | Sherpa | 12.2 |
| EW $WZ$ + 2 jets | Sherpa | $13 \cdot 10^{-3}$ |
| $Z^{(*)}Z^{(*)} \to 4\ell$ | Powheg + Pythia | 0.73 |
| $Z^{(*)}Z^{(*)} \to 2\ell 2\nu$ | Powheg + Pythia | 0.50 |
| EW $ZZ$ + 2 jets ($4\ell$) | Sherpa | $73 \cdot 10^{-5}$ |
| EW $ZZ$ + 2 jets ($2\ell 2\nu$) | Sherpa | $12 \cdot 10^{-4}$ |
| $W\gamma$ | Alpgen + Herwig | 369 |
| $Z\gamma$ ($p_T^{\gamma} > 7$ GeV) | Sherpa | 163 |

Table 8: Monte Carlo generators utilized to model the signal and background processes. The product of cross section times branching ratio is given for $\sqrt{s} = 8$ TeV. Only leptonic decays are included and summed over lepton flavours. The quoted numbers for the signal processes correspond to $m_H = 125$ GeV.

used for the electroweak and strong production of W-pairs and other diboson background processes. Alpgen [95] is used for the generation of the hard interaction of the W+jets, $Z/\gamma^*$+jets and $W\gamma$ backgrounds in this analysis. It is a tree level generator specialized in final states with large jet multiplicities with the exact leading order evaluation of partonic matrix elements. For the parton showering and the hadronization it is interfaced with Herwig [18]. The showering uses coherent branching algorithms for ISR and FSR. In contrast to Pythia, the Herwig generator employs the cluster hadronization model based on the colour pre-confinement property of angular ordered parton showers. Here, the so called MLM matching scheme [94] is applied for merging the first order matrix elements with the parton shower. Wherever Herwig comes to use in this analysis, it accesses the multiple scattering model of the program Jimmy [96] for the simulation of the underlying event. The $gg \to WW$ background process also uses the advantages of

the HERWIG package, but in this case the hard interaction comes from the dedicated GG2WW generator [97] because it includes the gluon fusion quark loops which actually belong to the NNLO calculation. Nearly all signal and background processes have been run through the full detector simulation. Only the top-quark related backgrounds and the $Z\gamma$ process Monte Carlo samples rely merely on the fast detector simulation in order to provide enlarged statistics.

### 5.4.3 Simulation Corrections

Although huge effort has been put into the various simulation programs, there are still some corrections to be applied that adjust the Monte Carlo simulation to the given data taking conditions. One example is the simulation of the trigger response and the depending efficiency corrections. Dedicated performance working groups examine the ratio of triggered objects in data and Monte Carlo simulation. By weighting the simulated events with this factor, the Monte Carlo is scaled respectively to describe the data. This factor is simply called scale factor (SF). But in case of the presented analysis, several single and dilepton triggers are used as described above accounting for the presence of two charged leptons that can both fire a trigger. So it is necessary to define an event-based scale factor out of the provided per-lepton scale factors as follows:

$$SF_{event}^{trigger} = \frac{Eff_{data}}{Eff_{MC}} = \frac{1 - \prod_{n=1,2} \epsilon_{data,n}}{1 - \prod_{n=1,2} \epsilon_{MC,n}} = \frac{1 - \prod_{n=1,2} SF_n \cdot \epsilon_{MC,n}}{1 - \prod_{n=1,2} \epsilon_{MC,n}} \tag{5.1}$$

where $n$ is enumerating the two leptons, $\epsilon_{data/MC}$ are the trigger efficiencies determined from data and Monte Carlo simulation respectively and the per-lepton scale factors $SF_n$. Moreover, the usage of dilepton triggers adjusts the calculation to

$$Eff = \epsilon_{single\ lep} + \epsilon_{dilep} \cdot (\epsilon_{dilep} - \epsilon_{single\ lep}) \tag{5.2}$$

where the asymmetry of the $p_T$-thresholds in the dilepton triggers needs to be taken into account. Also the modelling of the pileup contribution needs to be refined. Generally, both in-time and out-of-time pileup are included in the Monte Carlo simulation by overlying simulated minimum bias events. Dedicated Monte Carlo tunes have been developed to replicate an average data event. For every simulated event, the number of interactions per bunch crossing is drawn from a Poisson distribution with a mean value of 23, while the varying data taking conditions in 2012 lead to the measured value of $< \mu >= 20.7$ (see figure 23). Therefore, the simulated pileup distributions are reweighted with a 0.9 $\mu$-rescaling.

## 5.5 Object Criteria

This section describes the basic object requirements deduced from the information about the signal signature and the background composition introduced above. In case of the leptons, the criteria have been optimized to maximize the signal significance. The essential jet selection is important for the segmentation into jet multiplicity bins which defines the construction of the analysis strategy. Furthermore, an overlap removal procedure between leptons and jets has to be defined whenever more than one of the reconstructed objects are close together to avoid double counting of events. Since the emergence and the characteristics of the missing transverse energy vary for the different flavour and the same flavour final states, multiple versions are used to account for the specific properties. The detailed performance studies are collected in [98].

### 5.5.1 Leptons

The final state leptons are the essential handle to separate signal from background. So already the basic selection criteria have been optimized to achieve high background rejection while conserving the signal efficiency as well as possible. The main ingredients are vertexing and isolation. In addition to the standard ATLAS reconstruction and identification described in section 4 and the trigger requirements explained in section 5.3.1, supplementary and tightened criteria for the transverse and longitudinal impact parameters and the calorimeter and track based isolation variables have been developed. They vary with the lepton $p_T$ because of the changing background composition. At low transverse lepton momenta in particular, the W+jets background is dominant and drives the seek for background rejection. But it drops off sharply with increasing lepton $p_T$ and thus, the requirements can be relaxed to restore signal efficiency.

The figure of merit of the optimization procedure [98] is the significance

$$S = \frac{N_{sig}}{\sqrt{N_{sig} + N_{bkgd} + \sigma^2_{bkgd}}} \tag{5.3}$$

with $N_{sig}$ and $N_{bkgd}$ being the number of signal and background events passing the selection and $\sigma_{bkgd}$ the assigned uncertainty. A two dimensional scan of the track-based versus the calorimeter isolation variables is performed, which directly accounts for the correlation among them. Also the impact parameter requirements are probed in a two dimensional map of $d_0/\sigma_{d_0}$ vs. $z_0 \cdot \sin(\Theta)$. The projection of the longitudinal impact parameter accounts for the effect, that tracks in the forward direction have a longer projection on the z-axis which leads to larger uncertainties. While the optimization of the isolation requirements is performed in four different lepton $p_T$ bins of 10 – 15 GeV, 15 – 20 GeV, 20 – 25 GeV and >25 GeV, the one of the impact parameters is not $p_T$ dependent. In the following, the optimization results for electrons and muons are summarized.

### Electrons

The optimization scheme described above has been applied to both the cut-based and the multivariate electron identification schemes as they already provide different measures of signal efficiency and background rejection. Below 25 GeV, the VERY TIGHT LH performs best because of its improved rejection of backgrounds from non-prompt electrons, which particularly reduces the W+jets contamination. For electrons with $E_T > 25$ GeV, the W+jets background plays only a minor role. Therefore, the cut-based medium identification criterion becomes affordable. Its much looser requirements maximize the signal efficiency. The medium working point includes two requirements for the rejection of conversions, namely the conversion bit and a hit in the innermost pixel layer (b-layer). But in the default implementation, they are only applied to central electrons within $\eta < 2.37$. In this analysis, they are extended to the full detector coverage and the identification scheme is labeled CBL on that account.

In section 4.2.1, the relative calorimeter and track isolation are defined. As they vary with the transverse energy of the leptons, the optimization scheme scans the parameter space for the best working points. For low $E_T$ electrons, the isolation requirements are very tight to enforce the background rejection, especially of the W+jets contribution. With increasing transverse energy, this obligation can be loosened to restore the signal efficiency. The result is summarized in table 9. The optimized values for the impact parameter requirements are $d_0/\sigma_{d_0} < 3.0$ and $z_0 \sin\Theta < 0.4$ mm and do not depend on the electron energy since they describe the vertex quality. To account for these particular selection requirements, scale factors for the isolation and impact parameter efficiency are derived using the Z tag-and-probe method as described earlier.

| $E_T$ [GeV] | Calorimeter Isolation | Track Isolation | Impact Parameters |
|---|---|---|---|
| 10-15 | $E_T^{cone,0.3}/E_T < 0.20$ | $p_T^{cone,0.4}/E_T < 0.06$ | $d_0/\sigma_{d_0} < 3.0$, $z_0 \sin\Theta < 0.4$ mm |
| 15-20 | $E_T^{cone,0.3}/E_T < 0.24$ | $p_T^{cone,0.3}/E_T < 0.08$ | |
| 20-25 | $E_T^{cone,0.3}/E_T < 0.28$ | $p_T^{cone,0.3}/E_T < 0.10$ | |
| $> 25$ | | | |

Table 9: Electron selection requirements for different $E_T$ bins. Below 25 GeV, the likelihood identification is used at the VERY TIGHT working point. Above 25 GeV, the cut-based medium identification with extended conversion bit and b-layer requirement is applied.

In case of high proximity of two electrons within $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.1$, the one with larger transverse momentum is kept, while the one with lower $p_T$ is removed from the list of objects. When overlapping with a muon at a distance of $\Delta R < 0.1$, electrons are discarded because it is very likely that they have been reconstructed from the muon's inner detector track and its energy deposit in the calorimeter.

### Muons

The muons are reconstructed from tracks in the inner detector which are combined with the muon detector tracks using the STACO algorithm. Following the recommendation of the muon performance group, the muon candidates are required to have passed at least one pixel sensor and four SCT sensors while they are not allowed to contain more than two so called holes, which are missing hits in crossed sensors. Moreover, if the track is within the TRT acceptance, a successful extension to the TRT has to be found.

The isolation and impact parameter requirements are optimized with the same scheme as for the electrons described above. Table 10 summarizes the optimization results. For low transverse muon momentum, the isolation criteria are very tight and can be loosened towards higher $p_T$. The working point of the significance of the transverse impact parameter is the same as in the electron case. For the longitudinal impact parameter, a value of $z_0 \sin\Theta < 1.0$ mm is found to be optimal. Again, efficiency scale factors for the isolation and the impact parameter are derived with the Z tag-and-probe method.

| $p_T$ [GeV] | Calorimeter Isolation | Track Isolation | Impact Parameters |
|---|---|---|---|
| 10-15 | $E_T^{cone,0.3}/E_T < 0.06$ | $p_T^{cone,0.4}/p_T < 0.06$ | $d_0/\sigma_{d_0} < 3.0$, $z_0 \sin\Theta < 1.0$ mm |
| 15-20 | $E_T^{cone,0.3}/E_T < 0.12$ | $p_T^{cone,0.3}/p_T < 0.08$ | |
| 20-25 | $E_T^{cone,0.3}/E_T < 0.18$ | $p_T^{cone,0.3}/p_T < 0.12$ | |
| $> 25$ | $E_T^{cone,0.3}/E_T < 0.30$ | | |

Table 10: Muon selection requirements as function of $p_T$.

Muons that are overlapping with a reconstructed jet within $\Delta R < 0.3$ are removed from the list of objects. Most likely, those muons belong to the substructure of the jet and emerge from hadronic decays, even if the isolation criteria already suppress such a vicinity in many cases.

### 5.5.2 Jets

The exact jet selection has a non-negligible impact on the analysis performance. Splitting the final state into different jet multiplicities leads to dedicated strategies tailored specifically to the needs of the varying composition of background processes and their associated systematic uncertainties, especially those related to top quarks. That is why the jets used for the analysis binning are sometimes referred to as tagging jets.

The basic jet reconstruction has been described in section 4.2.3. An additional cleaning procedure [99] is used to reject background jets emerging from beam gas interactions, beam halo events, overlapping cosmic ray muons or calorimeter noise by exploiting the characteristic calorimeter pulse shape. This analysis uses the so called looser jet cleaning,

which is designed to ensure an efficiency for jets produced in proton proton collisions above 99.8% for jets with $p_T > 20$ GeV. However, much more important for the jet multiplicity binning is the choice of the $p_T$-threshold for the jets because it can often lead to migration of events with objects close to the threshold. This trading effect can impact the overall sensitivity. Therefore, the transverse momentum of jets and their jet vertex fraction (JVF) have been optimized by exploiting the poisson significance of the full statistical framework, which is discussed in section 7.1. This resulting selection requires more than 25 GeV of transverse momentum for all tagging jets within the coverage of the inner detector of $|\eta| < 2.4$ and additionally a jet vertex fraction of $|JVF| > 0.5$ if the jet $p_T$ is below 50 GeV. In the region $2.4 < \eta < 4.5$, where the jet reconstruction is based only upon the calorimeter information, the $p_T$ threshold is raised to 30 GeV.

In order to suppress the top related backgrounds, jets emerging from bottom quarks follow a different prescription. The flavour identification algorithm evaluates objects above a lowered $p_T$ threshold of only 20 GeV. Since this jet collection is not related to the actual jet multiplicity binning, these jets are referred to as sub-threshold jets. The b-tagging is only available within the tracking coverage of $|\eta| < 2.4$. Any event containing such a b-tagged jet is rejected by the so called b-jet veto requiring $N_{b-jet} = 0$. If a reconstructed jet is as close to an electron as $\Delta R < 0.3$, it is removed from the jet collection while the electron is kept in the event. This is particularly important because every electron is also entering the jet collection through the cluster reconstruction. Thus, the overlap removal avoids double counting of the objects and secures the correct categorization of the event. The tight isolation criteria applied to the electron retain its object quality, while the jet measurement presumably has been affected by the presence of the electron's energy deposit.

### 5.5.3 Missing Transverse Energy

The different methods of the reconstruction of missing transverse energy are outlined in section 4.2.4. However, for the analysis of the $H \to WW \to \ell\nu\ell\nu$ channel some refined considerations have been developed to account for the specific conditions. The signal process with the presence of two neutrinos as well as the majority of background processes evoke true missing transverse energy, especially in the different flavour final states $e\mu$ and $\mu e$. In the same flavour channels $ee$ and $\mu\mu$ on the other hand, one of the most significant backgrounds is the Z/Drell-Yan process, which does not involve neutrinos. The shape $E_T^{miss}$ of this kind of events is dominated by detector effects or mismeasurements. One possibility to reduce this effect is to use the projection of the missing transverse energy onto the axis of the nearest hard object, because the direction of $E_T^{miss}$ is correlated with the direction of the mismeasured object for events without real missing energy. This so called relative missing transverse energy is defined as:

$$E_{T,rel}^{miss} = \begin{cases} E_T^{miss} \cdot \sin\Delta\phi & \text{if } \Delta\phi < \pi/2 \\ E_T^{miss} & \text{if } \Delta\phi \geq \pi/2 \end{cases} \tag{5.4}$$

where $\Delta\phi$ is the angular difference between $E_T^{miss}$ and the closest lepton or jet. Furthermore, a combination of the calorimeter based and the track based reconstruction approach has been developed to enhance their advantages. The so called jet corrected track MET uses the full energy of the (tagging) jets by replacing the transverse momenta of all tracks associated to each jet with its energy deposit in the calorimeter:

$$E_T^{miss,jetCorr} = -\sum_{i\ tracks} \vec{p}_T^{\,i} + \sum_{j\ jets} (\vec{p}_T^{\,j,track} - \vec{p}_T^{\,j,Calo}) \tag{5.5}$$

which improves the resolution on the true missing transverse energy of the event and the robustness against pileup. In case of the 0 jets channel, the definition reduces to the standard track based approach. Since all different definitions rely on the jet selection, the jet cleaning procedure described above affects also the $E_T^{miss}$ calculations. The result is summarized as so called event cleaning, documented in [100].

By requiring the selected events to pass a certain $E_T^{miss}$ threshold, the background contamination gets significantly reduced. But the missing transverse energy also impacts other distributions that are crucial for the categorization of signal and background events. As an estimation of the neutrino momenta, it is included in the reconstruction of the transverse mass of the Higgs boson and similarly in the calculation of the transverse mass of the W-boson which are presented in the following section 6.1. To account for this extensive impact, an optimization study [98] has ascertained the best performing approach by maximizing the significance of the full statistical framework, including a scan of the missing transverse energy threshold and a fit of the transverse mass of the Higgs boson for all the aforementioned definitions. The result is a threshold of 20 GeV on the jet corrected track MET, harmonized for all jet multiplicities of the different flavour (DF) channels. In the same flavour (SF) channels on the other hand, the usage is not as uniform. For 0 jets, the projection onto the closest object of both the calorimeter based and the track based missing transverse energy are required to exceed 40 GeV, while the latter is reduced to 35 GeV for the 1 jet channel. In case of the 2 jet channel, requirements on the calorimeter based MET and its projection are imposed. The complete MET selection is summarized in table 11. However, the impact of the $E_T^{miss}$ definition on the transverse mass of the Higgs boson has been proven to be considerably more significant. The usage of the jet corrected track MET improves the significance up to 8% by separating the signal from the diboson and W+jets background processes and is therefore applied uniformly throughout all analysis channels.

| | Same Flavour ($ee,\ \mu\mu$) | Different Flavour ($e\mu,\ \mu e$) |
|---|---|---|
| 0-jets | $\not\!\!E_{T,rel}^{trk} > 40$ GeV, $\not\!\!E_{T,rel}^{calo} > 40$ GeV | $\not\!\!E_{T}^{jetCorr} > 20$ GeV |
| 1-jet | $\not\!\!E_{T,rel}^{trk} > 35$ GeV, $\not\!\!E_{T,rel}^{calo} > 40$ GeV | $\not\!\!E_{T}^{jetCorr} > 20$ GeV |
| 2-jets | $\not\!\!E_{T}^{calo} > 45$ GeV, $\not\!\!E_{T,rel}^{calo} > 25$ GeV | $\not\!\!E_{T}^{jetCorr} > 20$ GeV,$\not\!\!E_{T,rel}^{calo} > 45$ GeV |

Table 11: Selection cuts for the analysis subchannels using various definitions of missing transverse energy. The different flavour channels use the jet corrected track MET for all jet multiplicities. In the same flavour channels, the projection onto the direction of the closest object is used for the track- and the calorimeter-based missing transverse energy reconstruction method with varying thresholds.

# 6

# Event Selection and Background Estimation

This chapter explains the selection of $H \to WW \to \ell\nu\ell\nu$ candidate events and the determination of the background processes as described in [101]. By cutting on kinematical and topological distributions of the immanent observables, the signature of the signal process is enhanced, while the background contributions are reduced. The phasespace defined by this procedure is called signal region (SR). With the increased purity of these selected candidates, the statistical analysis is performed as explained in the following chapter. Similarly, dedicated background enhanced control regions (CR) can be defined to determine specific processes, examine their modelling in simulation and derive their normalization.

To achieve the desired separation power, not only the basic kinematic distributions of the final state objects are used but also some which defragment these parameters to combined observables. After the definition of the common observables in the first section, the so called preselection is illustrated. It is based on a few kinematic distributions and quantities to separate the most obvious reducible background contributions from the signal process, before splitting the analysis into subchannels depending on the jet multiplicity. Within these subchannels, the anatomy of the topological observables is used to primarily suppress the irreducible background components and purge the signal region. The dominant background sources are determined in their dedicated control regions, which are explained in detail in the following sections. The final remarks of this chapter belong to the theoretical and experimental systematic uncertainties related to the selection, which are a crucial input to the final statistical analysis.

## 6.1 Common Observables

In addition to the basic kinematic measures of the final state objects like the $p_T$ and $\eta$ distribution of the leptons or jets and the missing transverse energy, combined topological observables can be calculated to enhance the signal topology over the background contributions. Some of them have already been mentioned in section 5.1, when introducing the signal signature. The angular difference of the two leptons for example:

$$\Delta\phi_{\ell\ell} = \phi_{\ell_1} - \phi_{\ell_2} \tag{6.1}$$

is expected to be rather small, because of the spin correlation in the Higgs boson decay but tends to large values for background processes with back-to-back decay. Similarly, the angular difference between the dilepton system and the missing transverse energy:

$$\Delta\phi_{\ell,E_T^{miss}} = \phi_{\ell\ell} - \phi_{E_T^{miss}} \tag{6.2}$$

can prove beneficial for the event categorization. For the signal process, those angles should be spread back-to-back while the difference tends to small values when the $E_T^{miss}$ results from mismeasurements. Since the two leptons are the most distinct objects in the $H \to WW \to \ell\nu\ell\nu$ decay channel, another very important common observable is the mass of the dilepton system defined as:

$$m_{\ell\ell} = \sqrt{\left(E^{\ell_1} + E^{\ell_2}\right)^2 - \left(\vec{p}^{\,\ell_1} + \vec{p}^{\,\ell_2}\right)^2} \tag{6.3}$$

Again, for leptons emitted closely together as expected for the signal, the resulting dilepton mass is quite small and rather large for background events with large angles between the leptons. So the observables $m_{\ell\ell}$ and $\Delta\phi_{\ell\ell}$ are deeply correlated. The same holds for the transverse momentum of the dilepton system:

$$p_T^{\ell\ell} = \left|\vec{p}_T^{\,\ell\ell}\right| = \left|\vec{p}_T^{\,\ell_1} + \vec{p}_T^{\,\ell_2}\right| \tag{6.4}$$

which is in principle already part of the $m_{\ell\ell}$ definition. It is expected to feature low values for background but typically large values in case of signal events, for the exact same reasons as stated above.

Related to these quantities is the mass of two tau-leptons emerging from a Z-boson, each decaying further to one electron or muon and two neutrinos. It can be reconstructed with the so called collinear approximation [102], which assumes that the leptons and neutrinos are emitted collinearly with the tau-leptons because of the large mass gap between these particles ($m_{H,Z} \gg m_\tau \gg m_{\ell,\nu}$). It is mathematically defined as:

$$m_{\tau\tau} \overset{coll.approx.}{\approx} \frac{m_{\ell\ell}}{\sqrt{x_{\tau_1} x_{\tau_2}}} \tag{6.5}$$

with $x_{\tau_i} = p_T^{\ell_i}/p_T^{\tau_i} \in [0,1]$, the momentum fraction each lepton receives in the tau decay. This of course can only be determined by associating the missing transverse momentum with the neutrinos.

Another quantity, particularly composed to reduce the $Z/\gamma^* \to \ell\ell$ background contribution in the same flavour channels, is the so called hadronic recoil variable:

$$f_{recoil} = \frac{\left| \sum_{softjets} JVF \cdot \vec{p}_T \right|}{p_T^{\ell\ell}} \tag{6.6}$$

In this event topology without neutrinos and thus, no real missing transverse momentum, some hadronic object must balance the dilepton system. However, especially in absence of high-$p_T$ jets as in the 0-jets channel, only soft jets below the transverse momentum threshold are left to be considered. So the sum in equation 6.6 runs over all soft jets above 10 GeV opposite to the dilepton system $\frac{3}{4}\pi < \Delta\phi_{\ell\ell,softjets} < \frac{5}{4}\pi$. The product with $JVF$ reduces the pileup sensitivity. So finally, $f_{recoil}$ measures the strength of the recoil system relative to the dilepton system. For the 1-jet channel, this construct can be extended by adding the tagging jet-$p_T$ to the denominator.

The most important observable to confirm the signal process is the transverse mass of the Higgs boson. It can be reconstructed from the kinematical quantities of the two leptons and the missing transverse momentum:

$$m_T = \sqrt{\left(E_T^{\ell\ell} + E_T^{miss}\right)^2 - \left|\vec{p}_T^{\,\ell\ell} + \vec{p}_T^{\,miss}\right|^2} \tag{6.7}$$

with $E_T^{\ell\ell} = \sqrt{|\vec{p}_T^{\,\ell\ell}|^2 + m_{\ell\ell}^2}$. As explained in section 5.1, an excess of signal events over the background events is expected at the range of the Higgs boson mass $m_H$. However, the lack of information on the neutrinos prevents a full reconstruction of the Higgs decay.

The invariant mass of one W-boson suffers the same problem and again, only the transverse mass is accessible. It is defined as transverse mass of each lepton with the missing transverse energy as if for a leptonic W-decay:

$$m_T^W = \sqrt{2 p_T^{\ell} E_T^{miss} \left(1 - \cos\Delta\phi_{\ell, E_T^{miss}}\right)} \tag{6.8}$$

Particularly the maximum of this distribution $\max(m_T^W)$ helps to identify processes with at least one real W-boson, because typically one of the leptons leads to a large value of $m_T^W$. So background processes such as $Z/\gamma^* \to \tau\tau$ or QCD can be rejected by a

lower bound on the maximum value. It is also useful for the determination of W+jets background events.

## 6.2 Preselection of Candidate Events

The so called preselection of candidate events uses basic criteria to select the signal from the $H \to WW \to \ell\nu\ell\nu$ decay and reject other final state topologies and background processes before further splitting into jet multiplicities. The first requirement imposed is the presence of exactly two leptons (electron or muon) with opposite electrical charge, defining the four lepton flavour channels. The one with higher transverse momentum is called leading lepton, the softer one is called subleading lepton. Both have to exceed a $p_T$ threshold, which is 10 GeV for the subleading lepton and 22 GeV for the leading one. They are chosen as low as possible corresponding to the trigger selection described in section 5.3.1. Especially the lowered threshold on the subleading lepton has an important impact on the subsequent selection cuts. For the relatively light mass of the Higgs boson, one of the leptons in the decay chain is expected to be rather soft. So a low threshold is very important for the overall signal acceptance. But of course, also a lot of background events can enter the analysis this way, particularly W+jets and QCD processes, which has to be considered in the following strategy.

Figure 52 shows the kinematic distributions of the two leptons after the opposite charge requirement for all flavour channels combined. The signal process is superimposed for better visibility. The majority of the background processes is derived from Monte Carlo simulation. Only the contributions of the W+jets and the QCD background are referring to dedicated data driven estimates, which are described in section 6.5.5 and 6.5.6 later on. The large amount of selected events generally shows good agreement of data with the Monte Carlo simulation. But the background contributions still surmount the signal process by several orders of magnitude, lead by the Z+jets processes.

With the two energetic and isolated leptons in the final state, a low bound on the dilepton mass of $m_{\ell\ell} > 10(12)$ GeV for the DF (SF) channels cleans the selection from Drell-Yan production of the $\Upsilon$-meson resonance and $\gamma^*$ contributions. But the dominant mode of Drell-Yan lepton pair production is the peaking at the Z-boson mass. It leads to an overwhelmingly large background contribution, which is suppressed with the so called Z-veto requiring $|m_{\ell\ell} - M_Z| > 15$ GeV for the same flavour channels $ee$ and $\mu\mu$. The Z-veto rejects nearly 90% of the $Z \to \ell\ell$ background events. Figure 53 shows the $m_{\ell\ell}$ distributions for the same flavour (SF) and different flavour (DF) channels separately, before the application of the low bound and the Z-veto. The histograms illustrate the different background composition within each of the lepton channels and motivate the splitting.

The requirements on the missing transverse energy have been discussed in section 5.5.3. Since the application is so elementary for the reduction of Z+jets, W+jets and QCD events, one part of it is already applied at preselection level. Both distributions are

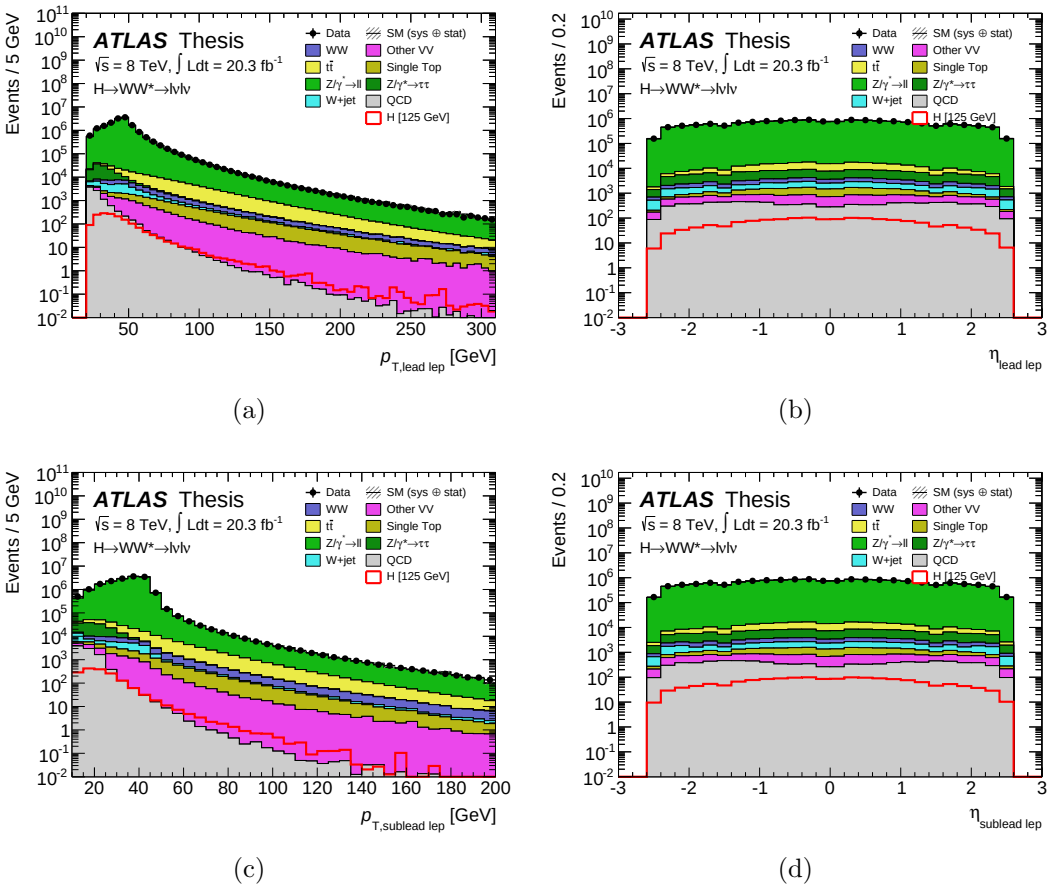
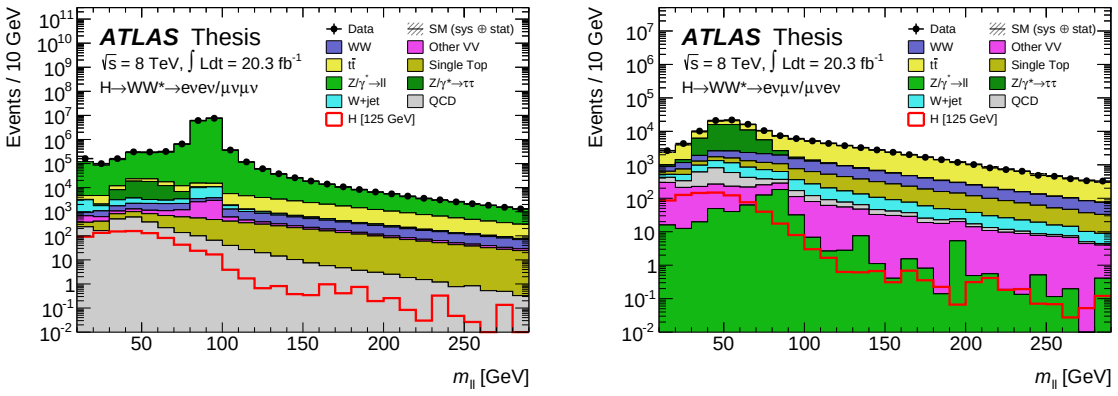



Figure 52: Kinematic distributions of the leading lepton on top and the subleading lepton below. Figures (a) and (c) show the $p_T$ distribution and figures (b) and (d) the $\eta$ distribution at preselection stage, after requiring exactly two oppositely charged leptons. All lepton flavour channels are combined. The signal process is superimposed for better visibility.

shown in figure 54 after the execution of the $m_{\ell\ell}$ cuts. Particularly in the same flavour channels, the threshold is chosen to be very large at 40 GeV for the calorimeter based MET projected onto the axis of the closest object. That removes about 98% of the Z+jets and QCD background processes as well as 90% of the W+jets events from the same flavour collection. For the different flavour channels, the threshold is relaxed to a 20 GeV cut on the jet corrected track MET which results in a rejection of more than 70% of the $Z \to \tau\tau$ and QCD contributions and another 40% of the $t\bar{t}$ and W+jets events.

Finally, at the end of the preselection, 99.2% of the collected background events have been rejected, while keeping the signal efficiency at 50%, which translates into a signal yield of 879 events expected from the Monte Carlo simulation. The analysis is now ready to be categorized into jet multiplicities. Figure 55 shows the $N_{jet}$ distribution of all four




(a) (b)

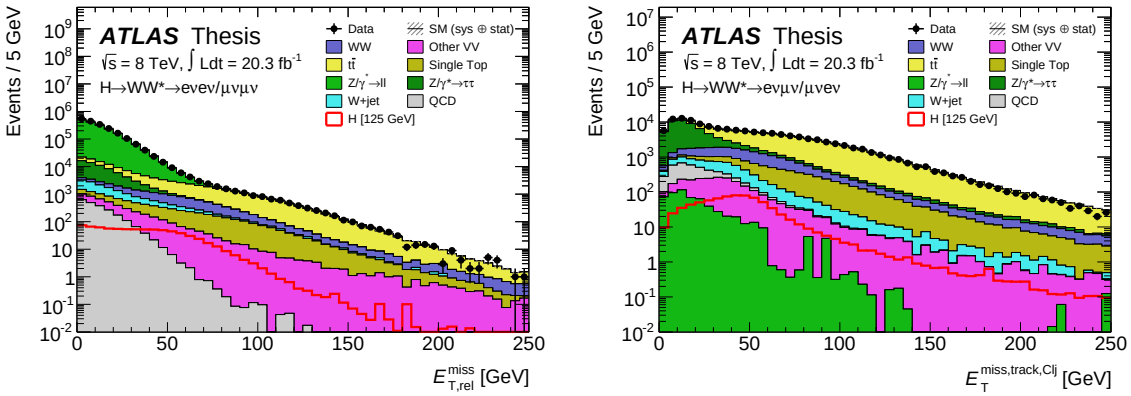
Figure 53: The $m_{\ell\ell}$ distribution for the SF (a) and DF (b) channels at preselection stage, after the lepton cuts. The signal process is superimposed for better visibility.

lepton flavour channels after the missing transverse energy requirements. A normalization factor for the top related backgrounds as explained in section 6.5.2 has already been applied. The dashed band refers to the statistical and the systematical uncertainties. The signal process is superimposed and scaled up by a factor of 20 for better visibility. Only the cases of zero and one additional jet are considered for further analyses of the gluon fusion Higgs boson production in this thesis. To retain orthogonality, the 2-jets channel is reserved for a dedicated vector boson fusion analysis. The same flavour channels are still dominated by the Drell-Yan background, whereas in the different flavour channels, the WW process becomes equally important. With increasing number of jets, also the fraction of top related backgrounds rises. So the focus lies clearly on the different flavour channels of the 0-jets selection, which show the best signal-to-background ratio.

The so called cutflow of the event selection is given in table 12 and lists the number of events after each selection requirement. The data column labels the number of observed events, while the signal and the background processes refer to the Monte Carlo prediction normalized to the integrated luminosity of the data.

| Selection Cuts | Summary | | | Background Composition | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Data | Signal | Bkgd | $WW$ | $t\bar{t}$ | st | diboson | $Z \to \ell\ell$ | $Z \to \tau\tau$ | W+jets | QCD |
| Object selection | 17242569 | 1757 | 16825941 | 23494 | 120769 | 11597 | 14991 | 16502082 | 107699 | 25538 | 19771 |
| lepton $p_{\mathrm{T}}$ | 16713082 | 1688 | 16459085 | 23210 | 119772 | 11509 | 13967 | 16160568 | 93842 | 24771 | 11446 |
| $m_{\ell\ell} > 12, 10$ GeV | 16530227 | 1603 | 16391213 | 23066 | 124912 | 11918 | 9559 | 16097716 | 93542 | 25895 | 4606 |
| $Z$ veto (SF) | 2150203 | 1546 | 2039698 | 20540 | 111751 | 10656 | 4664 | 1784381 | 91740 | 11596 | 4370 |
| $\not{E}_{\mathrm{T,rel}} > 45, 25$ GeV | 139935 | 879 | 135355 | 13198 | 56393 | 6391 | 2102 | 40378 | 13080 | 3168 | 645 |

Table 12: Cutflow for the preselection of $H \to WW \to \ell\nu\ell\nu$ candidate events of the $\sqrt{s} = 8$ TeV dataset. The number of observed events is listed in the first column. The signal processes for $m_H = 125$ GeV and the background processes refer to the MC prediction normalized to the total integrated luminosity.




(a) (b)

Figure 54: Missing transverse energy distributions at preselection stage, after $m_{\ell\ell}$ requirements. Figure (a) shows $\not{E}_{T,rel}^{calo}$ of the SF channels, while figure (b) illustrates $\not{E}_T^{jetCorr}$ of the DF channels. The signal process is superimposed for better visibility.

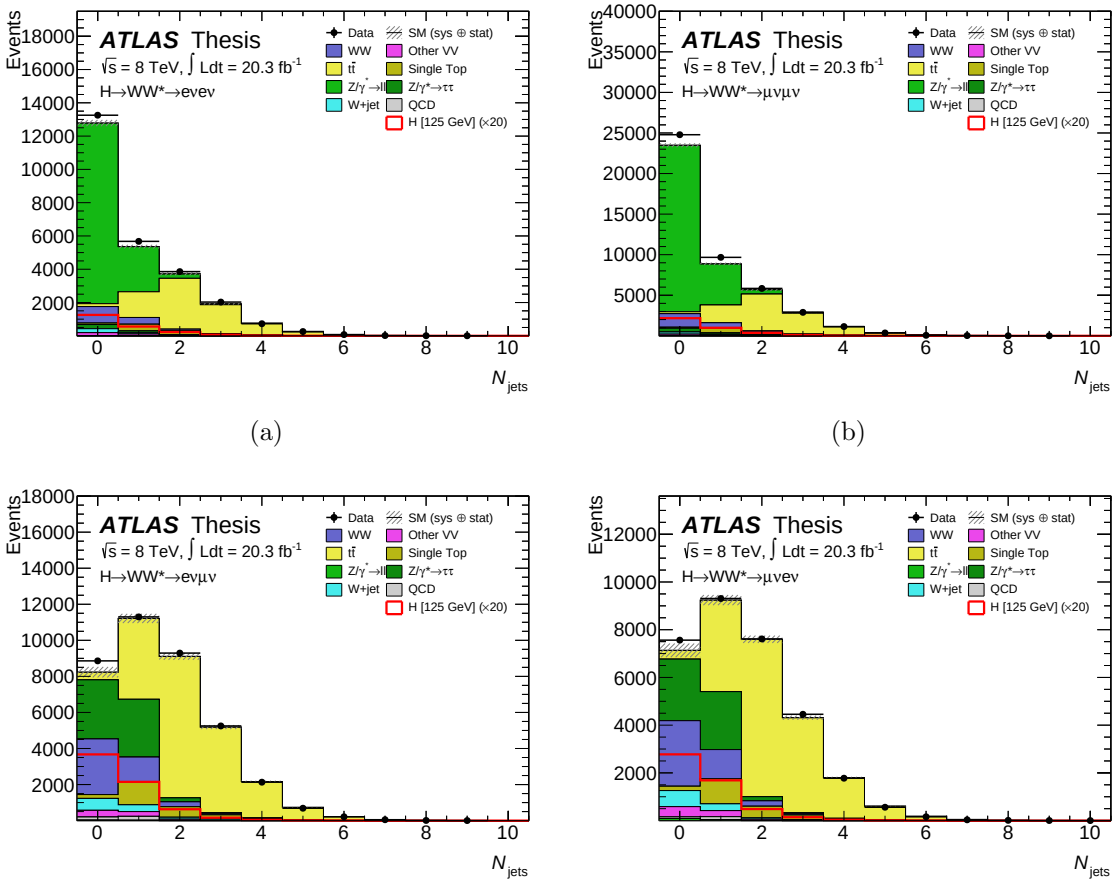



Figure 55: Distributions of the jet multiplicity at preselection stage, after the missing transverse energy requirements. The lepton flavour channels are separated into (a) $ee$, (b) $\mu\mu$, (c) $e\mu$ and (d) $\mu e$. The signal process is superimposed and scaled up by a factor of 20 for better visibility.

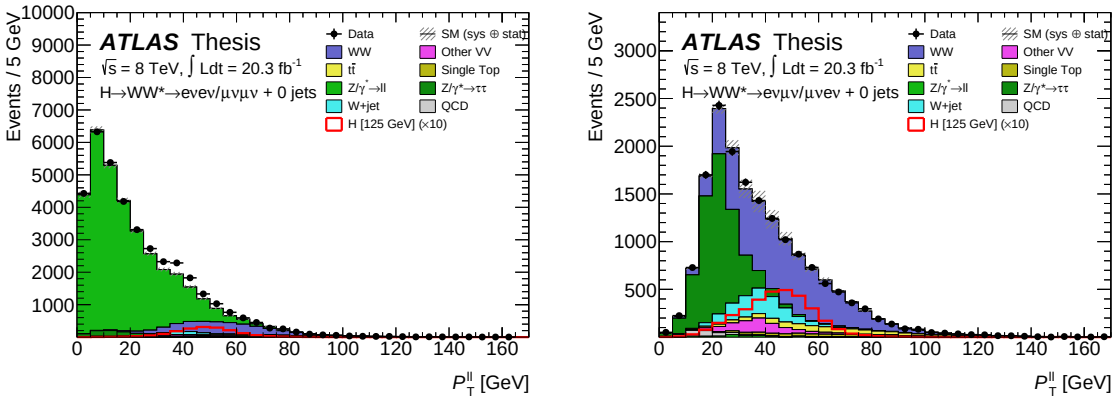
## 6.3 Selection for the 0-jets Analysis

After the basic dilepton preselection, the analysis is split into different jet multiplicity bins. A veto on any tagging jet (see section 5.5.2) identifies the 0-jets channel. It consists of the very first bin of the $N_{jet}$ distributions shown in figure 55. The subsequent selection criteria in the 0-jets analysis focus on the spin correlation of the $H \to WW \to \ell\nu\ell\nu$ decay and its resulting impact on the observables introduced at the beginning of this chapter. They are applied homogeneously to all lepton flavour combinations. But due to the extensive Drell-Yan background contributions, the same flavour channels need additional treatment.

The jet veto excludes nearly 98% of the $t\bar{t}$ background. The remaining contribution makes up only 2% of the total background fraction. In the same flavour channels, the $Z/\gamma^* \to \ell\ell$ contributions dominate with a fraction of 85% of the total background. The WW background rather accounts for about one half of the total background in the different flavour channels, $Z/\gamma^* \to \tau\tau$ still making up more than one third. On the other hand, nearly two thirds of the 494 expected signal events in the 0-jets analysis end up in the different flavour channels due to the softer requirements on the missing transverse energy. In total, the signal to background ratio is smaller than 0.1% at this stage.

The first selection cut requires the opening angle between the dilepton system and the missing transverse energy $\Delta\phi(\ell\ell, E_T^{miss})$ to be larger than $\pi/2$, removing events where these observables point into the same direction. While this criterion is 99.9% efficient for the signal process, it rejects about 7% of the Drell-Yan background, predominantly in the same flavour channels.

Significantly more substantial impact has the requirement on the transverse momentum of the dilepton system $p_T^{\ell\ell}$ to exceed 30 GeV. While it rejects about 76% of the $Z/\gamma^* \to \ell\ell$ contribution to the same flavour selection, the different flavour selection benefits from sorting out nearly 86% of the $Z/\gamma^* \to \tau\tau$ background. These Drell-Yan processes typically feature back-to-back leptons, which leads to low values of the transverse dilepton momentum. As explained in section 6.1, the Higgs decay via two W-bosons favours small opening angles. So overall, the signal efficiency is about 88% for this requirement. The distribution is shown in figure 56 for the same flavour and different flavour channels separately, each after the application of the $\Delta\phi(\ell\ell, E_T^{miss})$ cut. The signal process is superimposed and scaled up by a factor of 10 for better visibility. Most of the background processes are derived from Monte Carlo simulation. But additional normalization factors estimated in dedicated control regions are applied if available at the respective cutstage. The W+jets and the QCD background always refer to a data driven estimate. All these techniques are described in detail in section 6.5. It has to be mentioned that the Monte Carlo description of the data in the same flavour channels is rather poor. The overshoot around $p_T^{\ell\ell} = 40$ GeV is not an indication of a signal. The $p_T^{\ell\ell}$ requirement is derived from the more sensitive different flavour channels, which show a convincing agrement between data and Monte Carlo simulation. Disharmonazations between channels might be motivated but are avoided at this stage.



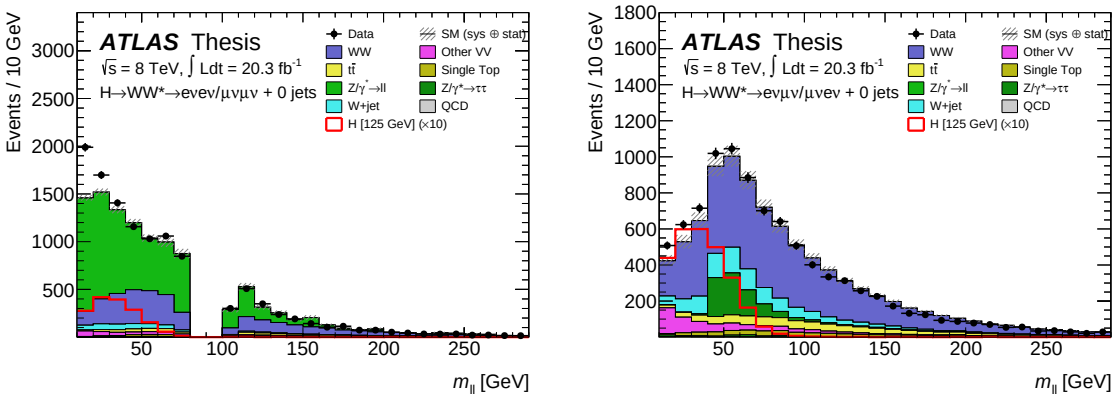

(a) (b)

Figure 56: $p_T^{\ell\ell}$ distributions after the full preselection and the 0-jets requirement. Figure (a) shows the same flavour channels and (b) the different flavour channels. The signal process is superimposed and scaled up by a factor of 10 for better visibility.

Similarly, the dilepton invariant mass $m_{\ell\ell}$ is required to be lower than 55 GeV. This criterion rejects more than half of the total background processes, but keeps another 87% of the signal events. Figure 57 shows the $m_{\ell\ell}$ distribution after the cut on $p_T^{\ell\ell}$, separately for the same flavour and the different flavour channels. The latter illustrates the strong reduction of the WW and $Z/\gamma^* \to \tau\tau$ contributions. On the other hand, the impact on the same flavour channels is not as strong. The residual excess in data observed before, is now located in the low $m_{\ell\ell}$ region, whereas the absence of events around 90 GeV illustrates the application of the Z-veto at preselection level.

Since the same flavour channels still suffer from the large Drell-Yan background contributions, an additional cut on the missing transverse energy addresses this problem. While the calorimeter based definition was used at the preselection level, now the projection of the track based quantity onto the axis of the closest object $\not{E}_{T,rel}^{jetCorr}$ is utilized to exploit the $ee$ and $\mu\mu$ channels further. The right diagram of figure 58 shows the distribution for the same flavour channels after the $m_{\ell\ell}$ cut. By requiring $\not{E}_{T,rel}^{jetCorr} > 40$ GeV, 86% of the $Z/\gamma^* \to \ell\ell$ events can be rejected. Thereby, the fraction of Drell-Yan drops down to only 30% of the total background and leaves the WW process as dominant background source. This impact can be observed by comparing the figures 58 (a) and 59 (a), which show the $\Delta\phi_{\ell\ell}$ distribution for the same flavour channels before and after the cut. The former certainly identifies a mismodelling of the $Z/\gamma^* \to \ell\ell$ Monte Carlo simulation evident in the shift towards larger opening angles. However, the normalization of the Drell-Yan process agrees very well, since it is extracted with a data-driven technique described in detail in section 6.5.4.

The dilepton opening angle $\Delta\phi_{\ell\ell}$ itself is required to be smaller than 1.8 radians to further assess the spin correlation of the $H \to WW \to \ell\nu\ell\nu$ decay. Its distribution is

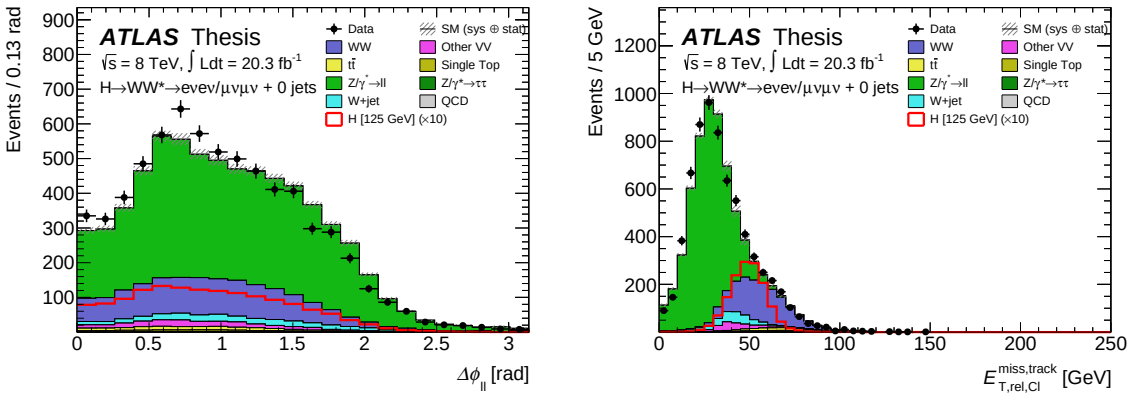



(a) (b)

Figure 57: $m_{\ell\ell}$ distributions for the 0-jets channel after application of the $p_T^{\ell\ell}$ requirements. Figure (a) shows the same flavour channels and (b) the different flavour channels. The signal process is superimposed and scaled up by a factor of 10 for better visibility.

shown in figure 59, for the same flavour channel after the additional cut on the missing transverse energy and for the different flavour channel after the $m_{\ell\ell}$ cut. The signal process is stacked on top of the background. Below the distribution, the ratio between data and the Standard Model expectation (including the signal process) is illustrated for comparison. The yellow band refers to the relative uncertainties. In case of the different flavour selection, the $m_{\ell\ell}$ requirement reduces the remaining contributions of the Z+jets and W+jets backgrounds and further enhances the dominance of the WW process to 64%. The same flavour channels are hardly affected at all. Signal and background efficiency are about 97% homogeneous. In case of the $ee$ and $\mu\mu$ channel, the information of the $\Delta\phi_{\ell\ell}$ distribution has already been exploited by the requirements on $p_T^{\ell\ell}$ and $m_{\ell\ell}$, which are highly correlated with one another.

Figure 60 shows the distributions of the subleading lepton $p_T$, the dilepton invariant mass $m_{\ell\ell}$ and the transverse mass $m_T$ for the different flavour channels after the requirement on $\Delta\phi_{\ell\ell}$. All indicate a good agreement of data with the Monte Carlo simulation. The signal contribution of 209 events is now about 9% compared to the size of the total predicted background, which is dominated by the WW process. The other diboson backgrounds and the W+jets process follow with 14% and 12% respectively. Altogether, they describe the observed data very well. But this is not yet the final signal region used in the statistical fit model. After the $\Delta\phi_{\ell\ell}$ cut, both different flavour channels $e\mu$ and $\mu e$ are treated separately. Each is subdivided depending on the transverse momentum of the subleading lepton and the invariant dilepton mass distribution. Three bins of $10 < p_T^{sub} < 15$ GeV, $15 < p_T^{sub} < 20$ GeV and $p_T^{sub} > 20$ GeV as well as two bins with $10 < m_{\ell\ell} < 30$ GeV and $30 < m_{\ell\ell} < 55$ GeV make up a total of 12 signal regions in the different flavour channels of the 0-jets analysis. The reason for this is visible



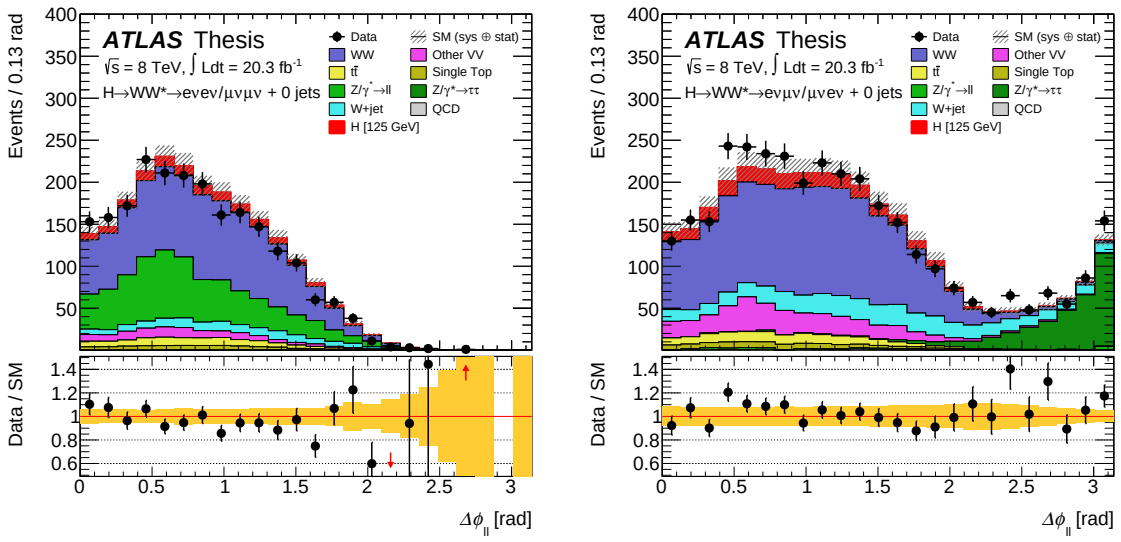

(a) (b)

Figure 58: (a) $\Delta\phi_{\ell\ell}$ and (b) $\not{E}_{T,rel}^{jetCorr}$ distribution for the same flavour channel after the application of the $m_{\ell\ell}$ requirements in the 0-jets selection. The signal process is superimposed and scaled up by a factor of 10 for better visibility.

in figure 60 (a) and (b). All of the bins are characterized by a similar signal-to-background ratio, but the composition of the background differs significantly. While the high $p_T^{sub}$ and high $m_{\ell\ell}$ regions contain mostly WW background, the lower bins show a lot more diboson and W+jets contributions. So the splitting of the signal regions enables the fit model to constrain different contributions for each separate region. A detailed description of this procedure is given in section 7.1.2. The full set of all signal region transverse mass distributions is shown in appendix B with elaborate tables.

Somewhat simpler is the case of the same flavour channels $ee$ and $\mu\mu$. They are considered as one combined selection. But since the $\Delta\phi_{\ell\ell}$ cut does not help to improve the signal to background ratio, an additional requirement on the recoil strength of the dilepton system $f_{recoil}$ is used to suppress the remaining background contributions. It is part of the data-driven Drell-Yan estimate (see section 6.5.4) mentioned above. The distribution is illustrated in figure 61. By selecting only the events that satisfy $f_{recoil} < 0.1$, more than 86% of the Drell-Yan and nearly one half of the total background is rejected. The signal efficiency is kept at 64%, leaving about 75 signal Monte Carlo events for the statistical treatment. The resulting transverse mass distribution is shown in figure 62. This selection defines the signal region of the combined same flavour channel.

Table 13 summarizes the cutflow of the 0-jets event selection. The upper box lists the different flavour channels and the lower box provides the numbers for the same flavour channels. The background processes refer to the MC prediction and are normalized to the total integrated luminosity of the 2012 dataset collected at $\sqrt{s} = 8$ TeV. Some of them include additional normalization factors derived from dedicated background control regions as explained in section 6.5 where applicable. The number of events in the different signal regions of the 0-jets analysis is given in table 28 in appendix A.

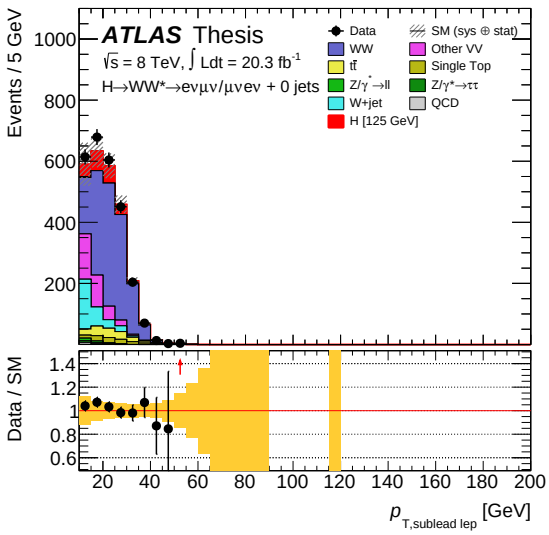



(a) (b)

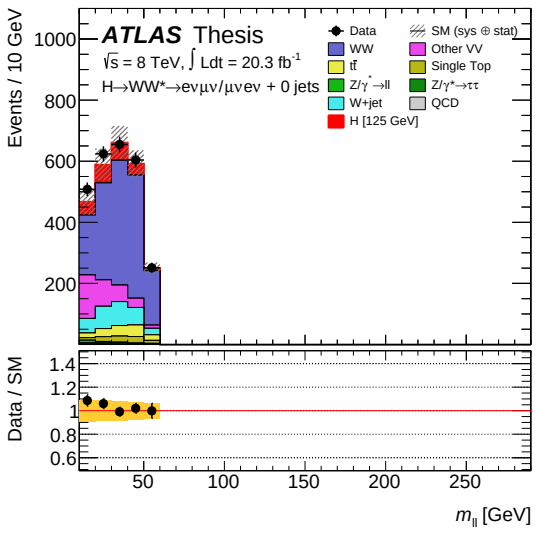
Figure 59: $\Delta\phi_{\ell\ell}$ distributions for the 0-jets selection. Figure (a) shows the same flavour channels after the cut on $\not{E}_{T,rel}^{jetCorr}$ and (b) the different flavour channels after the $m_{\ell\ell}$ cut. The signal process is stacked on the total background. The ratio between data and SM expectation is shown below the distributions, where the yellow band illustrates the relative uncertainties.

| Selection Cuts | Summary | | | Background Composition | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Data | Signal | Bkgd | $WW$ | $t\bar{t}$ | st | diboson | $Z \to \ell\ell$ | $Z \to \tau\tau$ | W+jets | QCD |
| DF ($e\mu$, $\mu e$) | | | | | | | | | | | |
| jet veto | 16423 | 322 | 16333 | 7113 | 820 | 407 | 739 | 115 | 5567 | 1335 | 237 |
| $\Delta\phi_{\ell\ell,MET} > 1.57$ | 16339 | 322 | 16271 | 7108 | 812 | 405 | 736 | 114 | 5532 | 1333 | 230 |
| $p_{T,\ell\ell}$ >30 GeV | 9339 | 273 | 9280 | 5690 | 730 | 363 | 571 | 60 | 783 | 1054 | 28 |
| $m_{\ell\ell} < 55$ GeV | 3411 | 232 | 3059 | 1670 | 141 | 79 | 353 | 27 | 350 | 427 | 12 |
| $\Delta\phi_{\ell\ell} < 1.8$ | 2642 | 209 | 2352 | 1503 | 132 | 75 | 324 | 19 | 12 | 278 | 9.2 |
| SF ($ee$, $\mu\mu$) | | | | | | | | | | | |
| jet veto | 38040 | 171 | 36520 | 3256 | 418 | 211 | 358 | 31059 | 685 | 504 | 29 |
| $\Delta\phi_{\ell\ell,MET} > 1.57$ | 35445 | 171 | 33892 | 3253 | 416 | 211 | 355 | 28518 | 622 | 493 | 26 |
| $p_{T,\ell\ell}$ >30 GeV | 11660 | 161 | 11037 | 3009 | 394 | 201 | 309 | 6704 | 21 | 396 | 2.6 |
| $m_{\ell\ell} < 55$ GeV | 6786 | 147 | 6713 | 1256 | 109 | 64 | 179 | 4843 | 8.7 | 251 | 2.0 |
| $\not{E}_{T,rel}^{jetCorr} > 40$ GeV | 2197 | 121 | 2156 | 1097 | 99 | 59 | 106 | 660 | 0.3 | 133 | 0.5 |
| $\Delta\phi_{\ell\ell} < 1.8$ | 2127 | 117 | 2097 | 1068 | 96 | 57 | 104 | 649 | 0.3 | 122 | 0.5 |
| $f_{\text{recoil}} < 0.1$ (SF) | 1108 | 75 | 1096 | 786 | 41 | 31 | 69 | 91 | 0.1 | 79 | 0.1 |

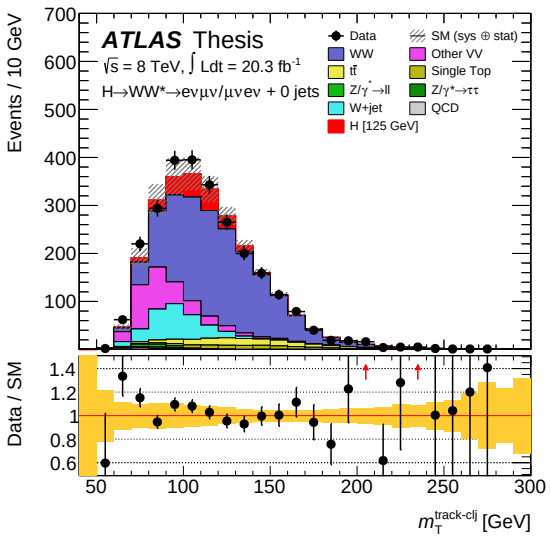
Table 13: Event selection of the 0-jets analysis. The different flavour channels are shown in the upper box, the same flavour channels below. The observed number of events is listed in the second column, signal and background processes refer to the MC prediction. Additional pre-fit normalization factors deduced from dedicated control regions are included where applicable.




(a)




(b)




(c)

Figure 60: Distributions of (a) $p_T^{sub}$, (b) $m_{\ell\ell}$ and (c) $m_T$ for the different flavour channels of the 0-jets selection after application of the $\Delta\phi_{\ell\ell}$ cut. The signal regions are extracted by further splitting of the lepton transverse momentum and the invariant dilepton mass. The signal process is stacked on the total background.

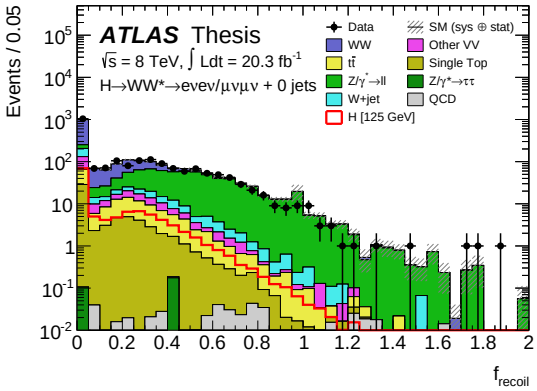



Figure 61: The $f_{recoil}$ distribution of the same flavour channels after the application of the $\Delta\phi_{\ell\ell}$ cut for the 0-jets selection. The signal process is superimposed for better visibility.

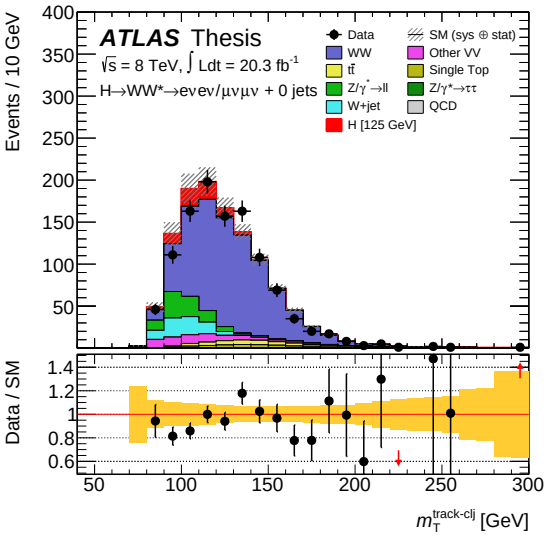
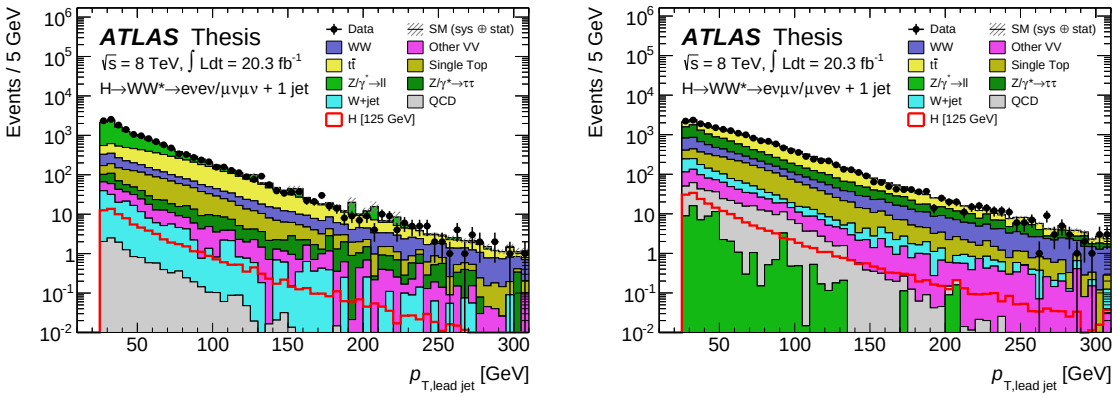



Figure 62: The $m_T$ distribution of the same flavour signal region in the 0-jets selection after cutting on $f_{recoil}$. The signal process is stacked on the total background.

## 6.4 Selection for the 1-jet Analysis

The selection strategy of the 1-jet analysis follows closely the prescriptions above, derived for the 0-jets case. However, due to the differing background composition, a few adjustments and dedicated considerations have to be applied. The starting point is the requirement of exactly one tagging jet in the events that have passed the preselection. Figure 63 shows the distributions of the transverse momentum of the tagging jet for the same flavour and different flavour channels respectively. The signal process is superimposed for better visibility. In contrast to the 0-jets analysis, the top related background processes dominate the selection making up nearly one half of the Monte Carlo simulated events, which describe the observed data very well. 31% of the pre-selected signal events end up in the 1-jet selection, leaving about 296 events for analysis. But this translates into a signal to background ratio which is still below 0.01.

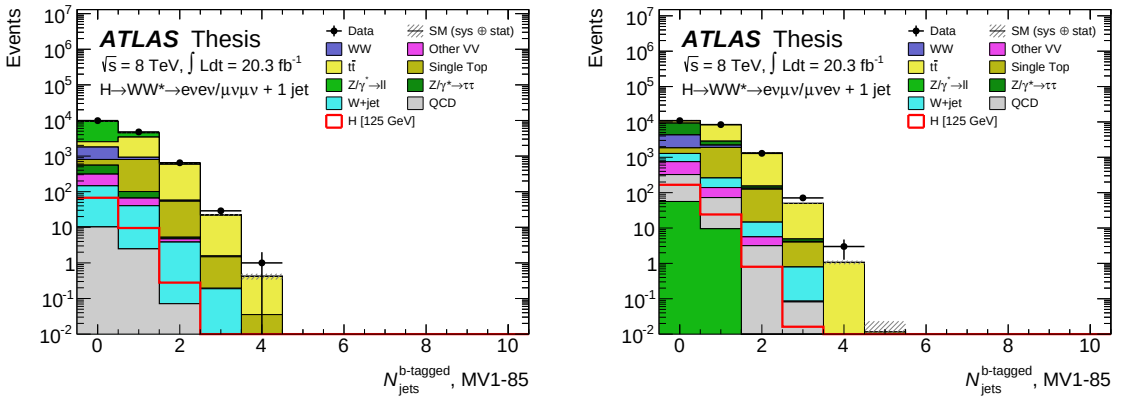



(a) (b)

Figure 63: Transverse momentum of the tagging jet for the same flavour (a) and different flavour (b) channels at the 1-jet selection. The signal process is superimposed for better visibility.

To enhance the signal fraction and in particular to address the top related background processes, the so called b-veto is introduced. Any jet with $p_T^{jet} > 20$ GeV originating from a bottom quark, tagged by the MV1 algorithm at the 85% efficiency working point as described in section 4.2.3 is rejected from the selection. The $N_{jet}^{btag}$ distributions for same and different flavour are shown in Figure 64. Selecting the very first bin $N_{jet}^{btag} = 0$ dismisses more than 80% of the top related processes and reduces its fraction of the total background to only 12%.

After the b-veto, the different flavour channels are dominated by $Z/\gamma^* \to \tau\tau$ background. So two dedicated cuts investigating this topology are imposed on the $e\mu$ and $\mu e$ channel. The first one is a requirement on the $\max(m_T^W)$ distribution, which is shown in figure 65 (a). In case of a $Z/\gamma^* \to \tau\tau$ decay, both final state leptons result in a rather low $m_T^W$ value. The same argument holds for QCD events, where leptons and missing



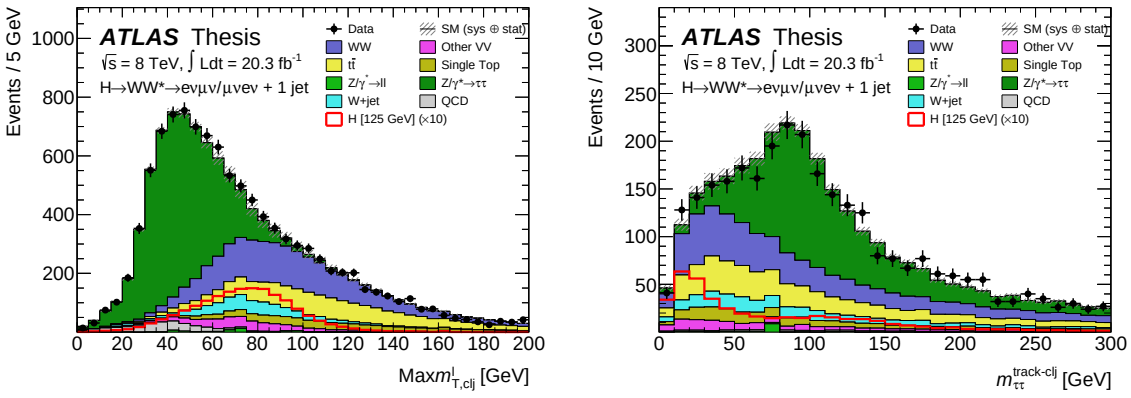

(a) (b)

Figure 64: Distribution of the number of b-tagged jets with $p_T > 20$ GeV identified by the MV1 algorithm at the 85% efficiency operating point for the same flavour (a) and different flavour (b) channels. The signal process is superimposed for better visibility.

transverse energy are a result of mismeasurements and misidentifications. By requiring the reconstructed transverse W boson mass to be $\max(m_T^W) > 50$ GeV, nearly 60% of the Drell-Yan and 77% of the QCD background are sorted out. The second criterion is the so called $Z \to \tau\tau$-veto. It requires the reconstructed mass of the two $\tau$ leptons based on the collinear approximation to be smaller than $(m_Z - 25)$ GeV. The distribution is shown in figure 65 (b) after the application of the $\max(m_T^W)$ cut. For better visibility, the signal process is superimposed and scaled up by a factor of 10. Only 35% of the simulated $Z/\gamma^* \to \tau\tau$ events survive the $m_{\tau\tau} < 66.187$ GeV cut and contribute 15% to the total background afterwards. On the other hand, both requirements each retain about 85% of the signal events.

The next step is similar to the 0-jets analysis. By requiring the invariant dilepton mass to be less than 55 GeV, more than two thirds of the top and WW backgrounds are rejected. The distribution is shown in figure 66, for the same flavour channels after the b-veto and the different flavour channels after the $Z \to \tau\tau$-veto. The impact of the $m_{\ell\ell}$ cut is slightly larger in the different flavour channels. In the same flavour channels, the Z-veto window of the preselection is visible, but the Drell-Yan process is still the dominant background contribution.

Following the same strategy as in the 0-jets same flavour analysis, an additional requirement on the missing transverse energy is utilized to assess this problem for the $ee$ and $\mu\mu$ channel. The jet corrected missing transverse momentum distribution for the combined same flavour channels is illustrated in figure 67 (b). Only 6% of the $Z/\gamma^* \to \ell\ell$ events satisfy the $\not{E}_{T,rel}^{jetCorr} > 35$ GeV requirement. Thus, the fraction of the Drell-Yan process on the total background is cut down to only 23%. The Standard Model WW production now dominates the selection, similarly to the different flavour channels. However,



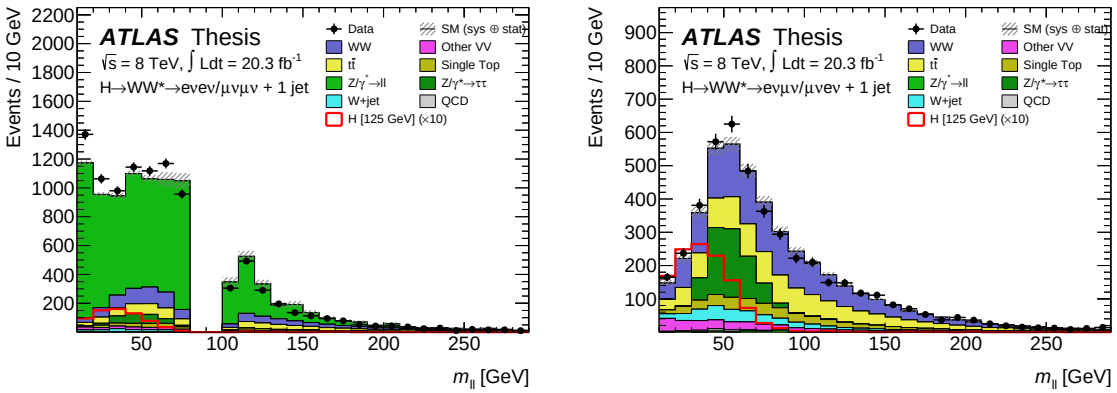

(a) (b)

Figure 65: Figure (a) shows the $\max(m_T^W)$ distribution for the different flavour channels in the 1-jet selection after the b-jet veto. After cutting on this quantity, figure (b) shows the resulting $m_{\tau\tau}$ distribution. The signal process is superimposed and scaled up by a factor of 10 for better visibility.

the description of the data looks rather poor. Although the Drell-Yan process is derived from a data-driven technique (see section 6.5.4), the missing transverse energy of $Z/\gamma^* \to \ell\ell$ events with one addtional jet seems to be underestimated. This can also be observed in figure 68 (a) illustrating the $\Delta\phi_{\ell\ell}$ distribution of the same flavour channels. Certainly, shape and normalization of the Drell-Yan process are insufficiently modelled. The impact of the missing transverse energy cut is demonstrated by the comparison of figure 67 (a) and 68 (a), which show the distribution before and after the application.

To further assess the spin correlation of the $H \to WW \to \ell\nu\ell\nu$ decay, the opening angle between the two leptons $\Delta\phi_{\ell\ell}$ is exploited. Its distribution is shown in figure 68 for the same (a) and different (b) flavour channels, where the signal expectation is stacked on the total background. The requirement of $\Delta\phi_{\ell\ell} < 1.8$ is very beneficial in case of the $e\mu$ and the $\mu e$ channel. It particularly rejects more than 94% of the remaining $Z/\gamma^* \to \tau\tau$ background but only 12% of the signal process. In the same flavour channels, the effect is obviously marginal and the handle on the spin correlation has already been used up.

After the $\Delta\phi_{\ell\ell}$ cut, the different flavour distributions of the subleading lepton $p_T$, the dilepton invariant mass and the transverse mass are shown in figure 69. Decent agreement of data and simulation is evident. The remaining 88 signal Monte Carlo events constitute 8.5% compared to the total number of background events. Dominating sources of background are the Standard Model WW production with a fraction of 40%, followed by the $t\bar{t}$ process with 26%. As explained in the previous section, the final selection of the different flavour channels used as input to the statistical fit is subdivided into 12 exclusive signal regions, depending on the binning in $p_T^{sub}$ and $m_{\ell\ell}$. They are shown in appendix B.

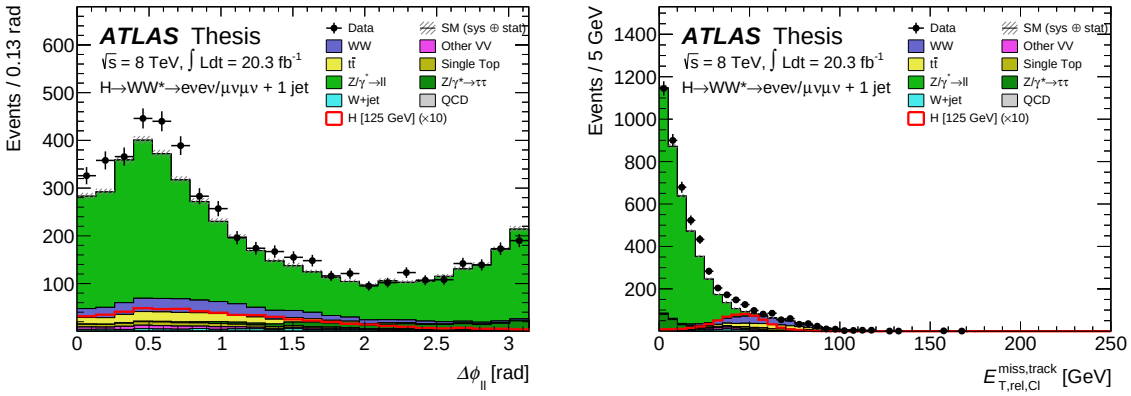



(a) (b)

Figure 66: $m_{\ell\ell}$ distributions for the 1-jet channel after application of the $Z \to \tau\tau$ veto. Figure (a) shows the same flavour channels and (b) the different flavour channels. The signal process is superimposed and scaled up by a factor of 10 for better visibility.

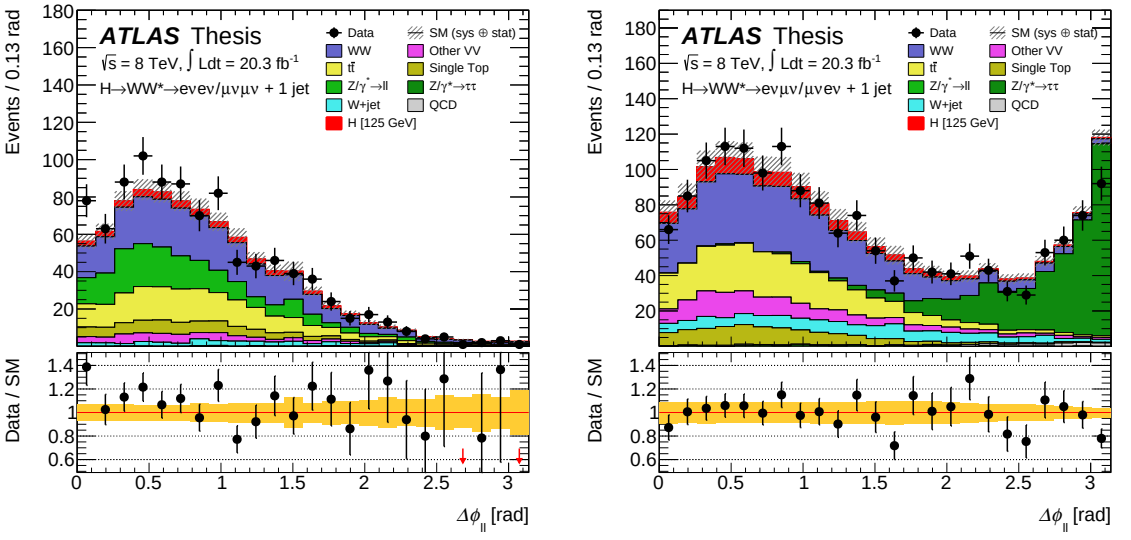
To reach a similarly high signal to background ratio in the same flavour channels as well, another requirement has to be imposed the same way as in the 0-jets analysis. The recoil strength of the dilepton system extended to the 1-jet environment $f_{recoil}^{ext}$ has to be smaller than 0.1. It is shown in figure 70. This reduces very efficiently 86% of the Drell-Yan contributions but also about one half of the remaining signal events. The transverse mass distribution for the combined same flavour channels is illustrated in figure 71 after the $f_{recoil}^{ext}$ cut. It is used as signal region input for the fit model. The Monte Carlo simulation indicates a composition of 46% of Standard Model WW production and 24% $t\bar{t}$ background. The signal expectation of 23 events is similar to the $Z/\gamma^* \to \ell\ell$ yield and contributes 6% compared to the total background estimation.

Table 14 displays the cutflow of the 1-jet event selection. In the upper box, the different flavour channels are listed and the lower box provides the numbers for the same flavour channels. The background processes refer to the MC prediction and are normalized to the total integrated luminosity. These also include additional normalization factors derived from dedicated background control regions where applicable, as explained in section 6.5. The event numbers for the different 1-jet signal regions is given in table 28 in appendix A.



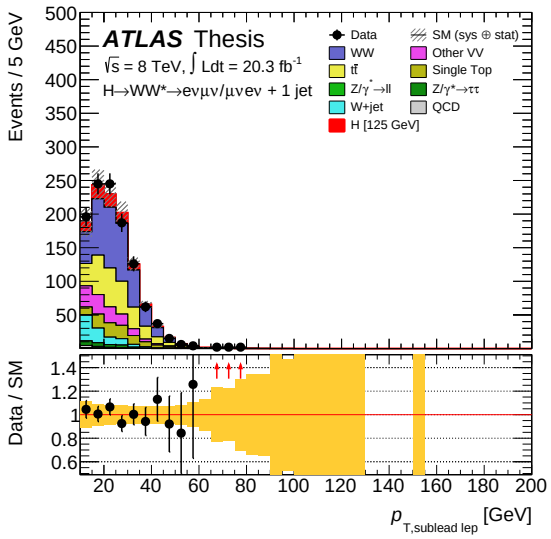

(a) (b)

Figure 67: (a) $\Delta\phi_{\ell\ell}$ and (b) $\not{E}_{T,rel}^{jetCorr}$ distribution for the same flavour channel after the application of the $m_{\ell\ell}$ requirements in the 1-jet selection. The signal process is superimposed and scaled up by a factor of 10 for better visibility.

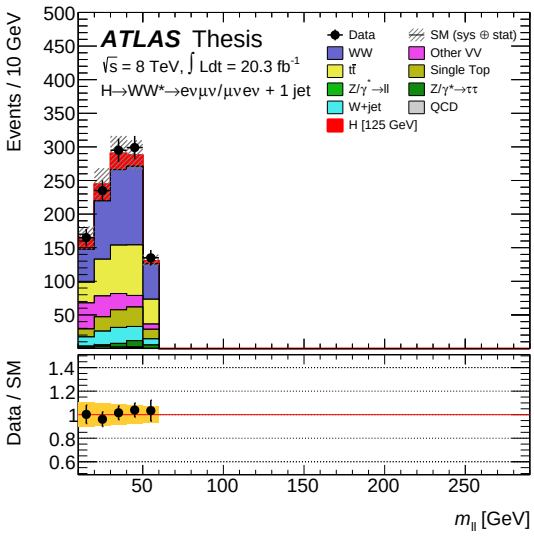



(a) (b)

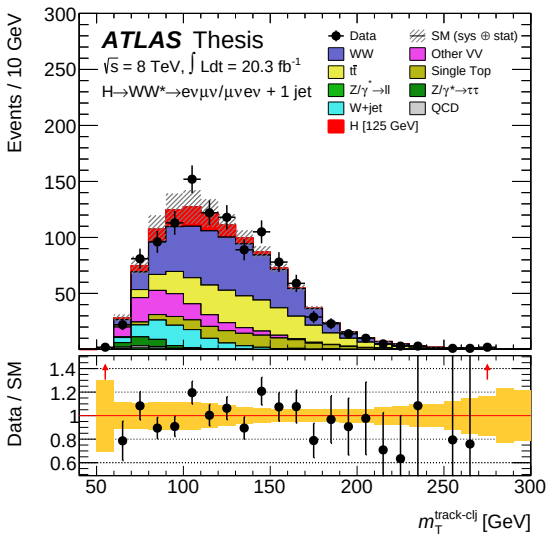
Figure 68: $\Delta\phi_{\ell\ell}$ distributions for the 1-jet selection. Figure (a) shows the same flavour channels after the cut on $\not{E}_{T,rel}^{jetCorr}$ and (b) the different flavour channels after the $m_{\ell\ell}$ cut. The signal process is stacked on the total background.




(a)




(b)




(c)

Figure 69: Distributions of (a) $p_T^{sub}$, (b) $m_{\ell\ell}$ and (c) $m_T$ for the different flavour channels of the 1-jet selection after application of the $\Delta\phi_{\ell\ell}$ cut. The signal regions are extracted by further splitting of the lepton transverse momentum and the invariant dilepton mass. The signal process is stacked on the total background.

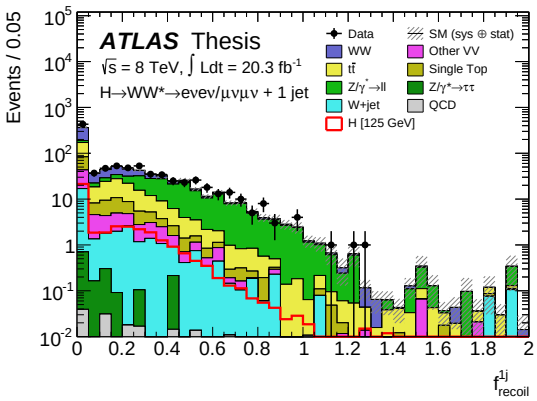



Figure 70: The $f_{recoil}^{ext}$ distribution of the same flavour channels after the application of the $\Delta\phi_{\ell\ell}$ cut for the 1-jet selection. The signal process is superimposed for better visibility.

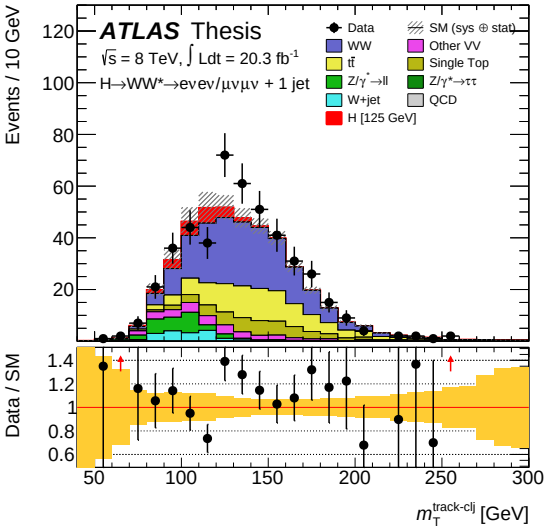



Figure 71: The $m_T$ distribution of the same flavour signal region for the final 1-jet selection after cutting on $f_{recoil}^{ext}$. The signal process is stacked on the total background.

| Selection Cuts | Summary | | | Background Composition | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Data | Signal | Bkgd | $WW$ | $t\bar{t}$ | st | diboson | $Z \to \ell\ell$ | $Z \to \tau\tau$ | W+jets | QCD |
| DF ($e\mu$, $\mu e$) | | | | | | | | | | | |
| one jet | 20607 | 192 | 20671 | 2754 | 8413 | 2314 | 496 | 66 | 5632 | 663 | 334 |
| $b$-jet veto | 10859 | 165 | 10767 | 2407 | 1614 | 554 | 423 | 56 | 4910 | 535 | 268 |
| $Z \to \tau\tau$ veto | 4574 | 119 | 4491 | 1667 | 1106 | 390 | 275 | 21 | 688 | 311 | 32 |
| $m_{\ell\ell} < 55$ GeV | 1656 | 100 | 1569 | 486 | 297 | 111 | 139 | 6.4 | 381 | 129 | 19 |
| $\Delta\phi_{\ell\ell} < 1.8$ | 1129 | 87 | 1030 | 418 | 269 | 102 | 119 | 5.0 | 22 | 88 | 6.1 |
| SF ($ee$, $\mu\mu$) | | | | | | | | | | | |
| one jet | 15344 | 77 | 14642 | 1111 | 3772 | 999 | 192 | 8100 | 279 | 178 | 13 |
| $b$-jet veto | 9897 | 67 | 9138 | 972 | 725 | 245 | 163 | 6645 | 240 | 137 | 10 |
| $Z \to \tau\tau$ veto | 9897 | 67 | 9715 | 972 | 725 | 245 | 163 | 7222 | 240 | 137 | 10 |
| $m_{\ell\ell} < 55$ GeV | 5127 | 58 | 4706 | 351 | 226 | 85 | 79 | 3717 | 167 | 73 | 7.8 |
| $\not{E}_{\mathrm{T,rel}}^{\mathrm{track}} > 35$ GeV | 960 | 43 | 842 | 292 | 193 | 73 | 49 | 194 | 1.7 | 38 | 0.2 |
| $\Delta\phi_{\ell\ell} < 1.8$ | 889 | 39 | 783 | 265 | 179 | 68 | 44 | 194 | 1.5 | 30 | 0.2 |
| $f_{recoil}^{ext} < 0.10$ | 467 | 23 | 404 | 188 | 98 | 44 | 29 | 26 | 0.9 | 17 | 0.1 |

Table 14: Event selection of the 1-jet analysis. The different flavour channels are shown in the upper box, the same flavour channels below. The observed number of events is listed in the second column, signal and background processes refer to the MC prediction. Additional pre-fit normalization factors deduced from dedicated control regions are included where applicable.

## 6.5 Background Estimation

All significant background processes are estimated separately using either a data-driven approach or a dedicated control region (CR). In such a control region, the normalization of the Monte Carlo simulation is determined to improve the description in the signal region (SR). On the other hand, data-driven estimation techniques derive the shape and the normalization of individual processes directly from data and replace the Monte Carlo simulation. In either case, the basic strategy is to define a selection which enhances one source of background to high purity and preferably no signal contamination. In such an environment, the background can be exploited at high accuracy. Extrapolating back the information finally predicts the number of background events in the signal region. To reduce the uncertainties on this extrapolation, the control regions should collect a similar phase space. This is typically achieved by loosening or reversing some of the specific criteria imposed on the signal region. All of the different approaches have been documented in [103].

The Standard Model WW production, the other diboson processes WZ, ZZ and W$\gamma^{(*)}$ or the $Z/\gamma^* \to \tau\tau$ background are estimated in a dedicated control region and the derived Monte Carlo normalization factor is then applied to the signal region. The $Z/\gamma^* \to \ell\ell$ background is investigated with a fully data-driven method comparing $Z/\gamma^*$ enriched and depleted regions. Top related backgrounds are exploited in an inclusive control region to derive the normalization. Additionally, in the 0-jets analysis, a jet veto efficiency estimation in a b-tagged control sample is applied. In case of the 1-jet analysis, the b-veto efficiency is derived from a two jets control region. A different approach is used to determine the number of W+jets and QCD events with a fully data-driven method. Here, the control sample is defined by the exact same cuts as the signal region but using a different lepton selection. The application of the measured rate at which jets are misidentified as leptons performs the extrapolation to the signal region. All these different methods are explained in detail in the following sections.

### 6.5.1 WW Control Region

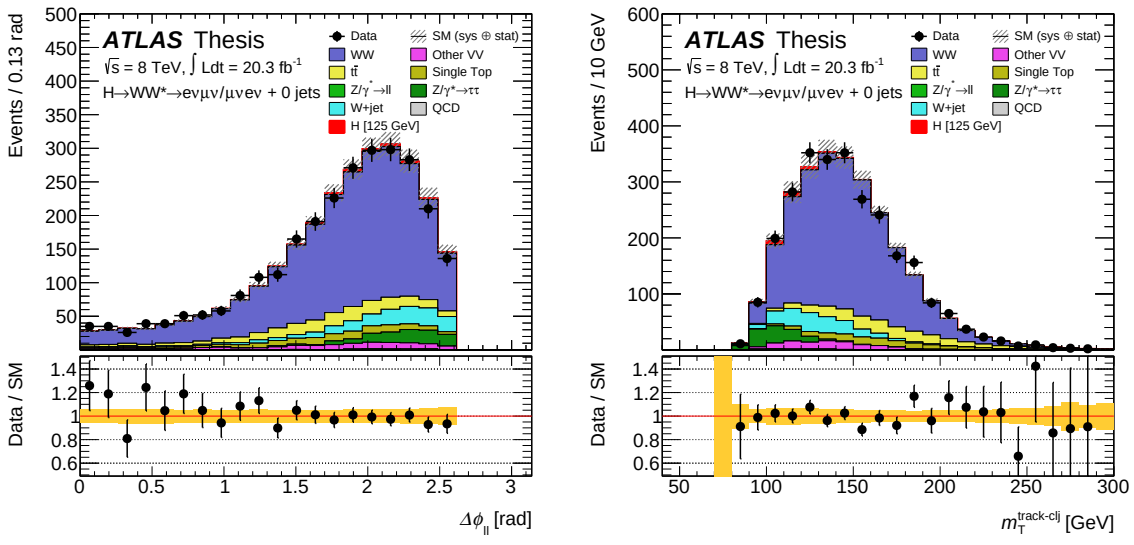
As the dominant source of background to the $H \to WW \to \ell\nu\ell\nu$ analysis, the Standard Model WW production has to be investigated carefully. Dedicated control regions exploit this process separately for the 0-jets and the 1-jet analysis. In case of the 0-jets channel, the control region is selected after the $p_T^{\ell\ell}$ requirement and includes the full preselection with missing transverse energy cuts and jet veto. Some further criteria define the control region and reduce remaining contributions from other backgrounds. The subleading lepton transverse momentum is required to be $p_T^{sub} > 15$ GeV in order to reduce the W+jets contamination. With a cut on $\Delta\phi_{\ell\ell} < 2.6$, a large fraction of the $Z/\gamma^* \to \tau\tau$ events is rejected. Finally, to explicitly separate the WW control region from the signal region, the range of the dilepton mass is chosen to be $55 < m_{\ell\ell} < 110$ GeV. Within this selection, the WW background is estimated by subtracting all other Monte Carlo

simulated background processes from the observed data. The ratio of the remaining data events to the predicted WW yield defines the normalization factor:

$$NF^{WW} = \frac{Data - MC^{non-WW}}{MC^{WW}} \tag{6.9}$$

The application of this factor to the number of WW Monte Carlo events in the signal region gives the resulting WW estimate $N_{WW}^{SR} = NF^{WW} \cdot N_{WW}^{MC,SR}$. Due to the large Drell-Yan contamination in the same flavour channels, the normalization factor is only determined in the different flavour channel $e\mu + \mu e$. There, the purity of WW events is 73% and the resulting normalization factor is $1.22 \pm 0.03(\text{stat}) \pm 0.10(\text{syst})$. It is applied to all lepton flavour combinations.

Figure 72 shows the $\Delta\phi_{\ell\ell}$ and the $m_T$ distributions of the different flavour channel in the WW control region for the 0-jets analysis. The normalization factor on WW is already applied. The contributions of the $Z/\gamma^* \to \tau\tau$, diboson and the top background have been normalized with the results from their own dedicated control regions. The number of W+jets events is estimated from the data-driven method (see section 6.5.5). Both distributions show good agreement between data and simulation. Table 15 shows the number of selected events per process.

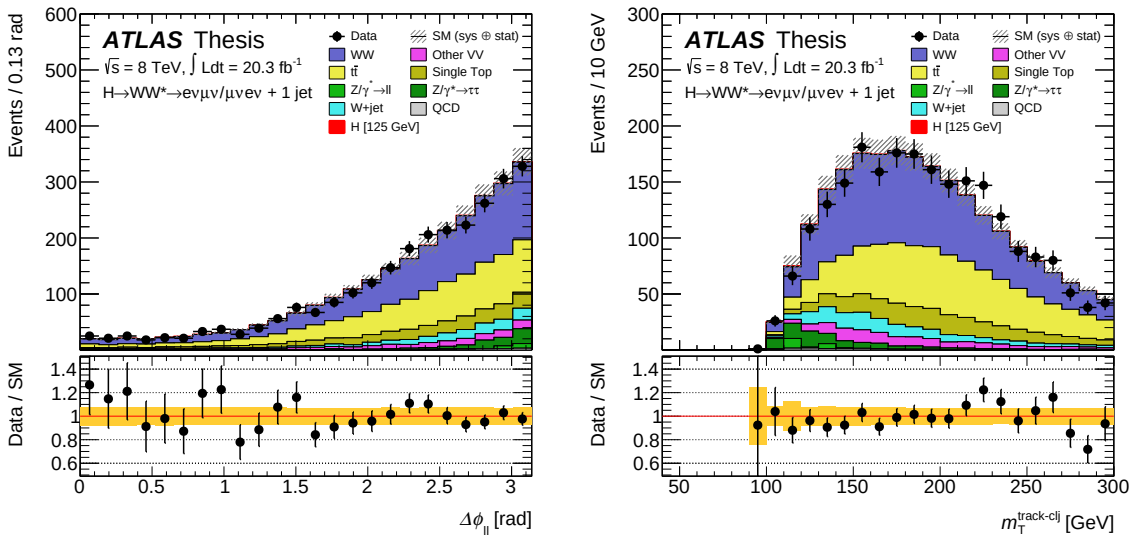



(a) (b)

Figure 72: (a) $\Delta\phi_{\ell\ell}$ and (b) $m_T$ distributions of the different flavour channel in the WW control region for the 0-jets analysis. The normalization factor on the WW simulation is already applied.

The same procedure is also employed in the 1-jet analysis. After the cut on $\max(m_T^W)$ in the 1-jet selection, the WW control region is defined by collecting events that fulfill $|m_{\tau\tau} - m_Z| > 25$ GeV, and $p_T^{sub} > 15$ GeV to reduce other background processes. With a lower bound on the dilepton mass $m_{\ell\ell} > 80$ GeV the signal region events are excluded from the control region. The normalization factor is determined by taking the ratio of the background-subtracted data to the simulated WW events. Again, it is only estimated for the different flavour channels and applied to all lepton combinations. The result is $NF_{WW}^{1j} = 1.05 \pm 0.05(\text{stat}) \pm 0.24(\text{syst})$.

The $\Delta\phi_{\ell\ell}$ and the $m_T$ distribution of the different flavour channels in the 1-jet WW control region are shown in figure 73. The normalization factor is already applied to the WW Monte Carlo simulation. Although dominated by the WW process, the purity in the 1-jet control region is only 43%. Particularly the top backgrounds contribute a large fraction of events. Therefore, they are corrected with the factors determined in the top estimation. Also the diboson and $Z \to \tau\tau$ backgrounds are normalized according to their control regions and the W+jets contribution is estimated using the data-driven approach. The result is a good agreement between data and Monte Carlo simulation. Table 15 lists the number of selected events per process.




(a) (b)

Figure 73: (a) $\Delta\phi_{\ell\ell}$ and (b) $m_T$ distributions of the different flavour channel in the WW control region for the 1-jet analysis. The normalization factor on the WW simulation is already applied.

Due to the limited accuracy of the Monte Carlo prediction in the control region, a theoretical systematic uncertainty is assigned to each of the extrapolation factors described above. Higher order corrections and different modelling parameters are taken into account. The evaluation is described in detail in section 6.6. Depending on the transverse momentum of the subleading lepton, the systematic uncertainty on the WW background

| Control Region | Summary | | | Background Composition | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Data | Signal | Bkgd | $WW$ | $t\bar{t}$ | st | diboson | $Z \to \ell\ell$ | $Z \to \tau\tau$ | W+jets | QCD |
| 0j: $WW$ CR | 2713 | 28 | 2685 | 1954 | 216 | 119 | 97 | 8.7 | 106 | 182 | 2.0 |
| 1j: $WW$ CR | 2647 | 4.2 | 2643 | 1148 | 834 | 270 | 127 | 17 | 81 | 152 | 13 |

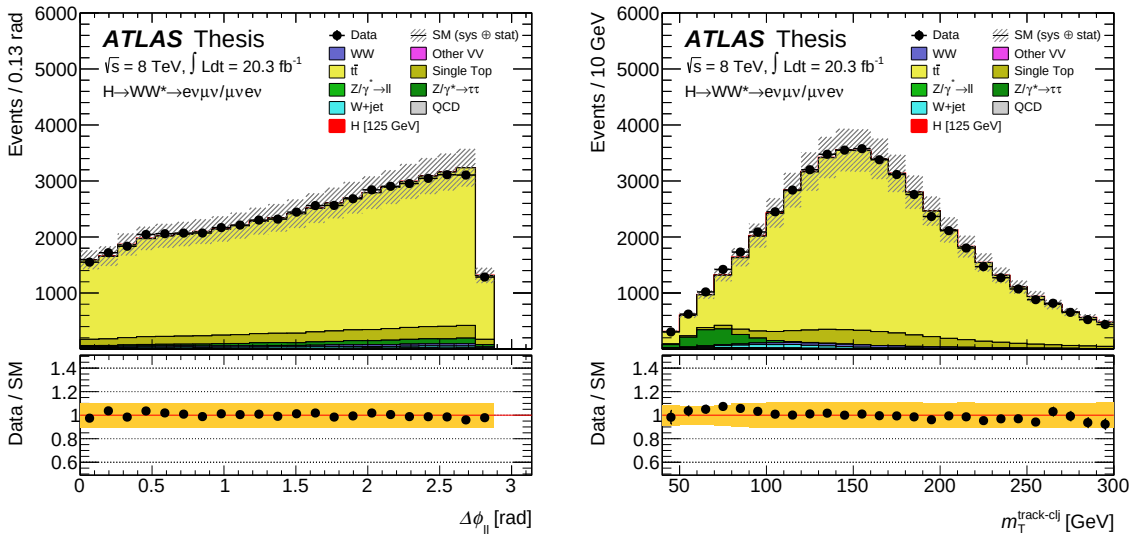
Table 15: Event selection of the WW control region for the different flavour channels in the 0-jets and the 1-jet analyis. The observed number of events is listed in the first column, signal and background processes refer to the MC prediction. Additional normalization factors derived from other background control regions as well as $NF_{WW}$ have already been applied.

in the signal regions ranges from 2.0% to 4.8% in case of the 0-jets selection, and from 3.9% to 7.1% for the 1-jet selection. These numbers are cross-checked by extrapolating the WW control region normalization factors to a so called validation region (VR), which is defined by selecting the high $m_{\ell\ell}$ region above 110 GeV. Considering all systematic uncertainties, in the validation region, the normalized prediction is consistent with the data at a level of 1.1 standard deviations.

### 6.5.2 Top Background Estimation

Since the majority of dilepton events with additional missing transverse energy at the LHC are produced by $t\bar{t}$ or single top processes decaying via W-bosons, the preselection of the $H \to WW \to \ell\nu\ell\nu$ analysis is dominated by top related backgrounds. So the estimation of the top related backgrounds starts with a so called inclusive top control region. After the missing transverse energy requirement, a cut on $\Delta\phi_{\ell\ell} < 2.8$ is imposed to reduce the $Z \to \tau\tau$ contributions. All jet multiplicities are included, but at least one b-tagged jet with $p_T > 20$ GeV identified by the MV1 algorithm at the 85% efficiency working point has to be present. The top normalization factor is derived by taking the ratio of the background subtracted data to the sum of simulated top backgrounds. This is performed in the different flavour channels only, because of the larger statistics and higher purity. The result of $NF_{Top} = 1.05 \pm 0.005$ is then applied to all lepton flavour combinations at the preselection stage.

Figure 74 shows the $\Delta\phi_{\ell\ell}$ and the $m_T$ distributions of the different flavour channel in the inclusive top control region, which illustrate the high purity of 95%. The normalization factor on the $t\bar{t}$ and the single top processes is already applied. The $Z/\gamma^* \to \tau\tau$ contributions are normalized in a dedicated control region and the W+jets events are estimated with the data-driven method. Both distributions show very good agreement between data and simulation. Table 16 shows the number of selected events per process.




(a) (b)

Figure 74: (a) $\Delta\phi_{\ell\ell}$ and (b) $m_T$ distributions of the different flavour channel in the inclusive top control region. The normalization factor on the $t\bar{t}$ and single top simulation is already applied.

### 0-jets: Jet Veto Efficiency

In the 0-jets selection, the estimated top background is based on the jet veto efficiency. Typically, the top events constitute one or more jets which should not pass the jet veto. But due to the very large cross section, a significant number of events enters the 0-jets analysis. The contribution can be estimated with:

$$N^{exp}_{Top,0j} = (N_{Data} - N_{non-top}) \times P_2^{exp} \tag{6.10}$$

where $N_{Data}$ is the number of observed events after the preselection cuts plus the requirement of $\Delta\phi_{\ell\ell} < 2.8$. The contributions from other processes are removed by subtracting $N_{non-top}$ predicted by the Monte Carlo simulation or the available data-driven estimate. $P_2^{exp}$ is the fraction of top events passing the jet veto determined from data. It can be expressed as:

$$P_2^{exp} = P_2^{MC} \times \left(\frac{P_1^{Data}}{P_1^{MC}}\right)^2 \approx P_2^{MC} \times \left(\frac{P_1^{b-tag,Data}}{P_1^{b-tag,MC}}\right)^2 \tag{6.11}$$

with the single jet veto survival probability (JVSP) $P_1^{Data/MC}$ derived from data or Monte Carlo respectively, because the probability of two jets to pass the cut is the same as the probability that each jet does ($P_2 = P_1^2$). In this analysis, the single jet veto survival probability is derived from a control sample requiring at least one b-tagged jet from the MV1 algorithm. With all these measures at hand, the normalization factor for the 0-jet signal region is calculated to be $NF^{SR,0j}_{Top} = N^{exp}_{Top,0j}/N^{MC}_{Top,0j} = 1.08 \pm 0.024(\text{stat})$ leading to the final top background yield.

However, this procedure inherits a number of systematic uncertainties, which need to be taken into account. Theoretical uncertainties arise from the use of Monte Carlo simulation. They are evaluated by varying the different scales and modelling parameters of the generators to check the impact on the different topologies of the control region and the b-tagged region. The effect of the additional cuts after the preselection and experimental uncertainties like those of the jet energy scale and resolution for example, are also taken into account. The total uncertainty on the top background yield in the 0-jets signal region is 8%. More details are given in section 6.6.

### 1-jet: b-tagging Efficiency

The top background estimate for the 1-jet analysis is based on the measurement of the b-tagging efficiency. In this data-driven method, events with two jets are selected and at least one b-tagged jet is required to achieve a $t\bar{t}$ dominated sample. Other background sources are subtracted using their Monte Carlo expectation with available normalization factors deduced from control regions and the data-driven values for the W+jets and QCD

background. The jet b-tagging efficiency in this 2-jets control region $\epsilon_{data}^{CR,2j}$ is defined as the fraction of events in which the second jet is also b-tagged. There are several ways to execute this jet tag-and-probe method. Here, the tag jet is chosen randomly and the other one is defined as probe jet. It has turned out to be the best option, because the kinematic distributions of the probe jets agree much better with those populating the 1-jet selection than using for example a $p_T$-ordered choice. This means that the phase space of the control region is closer to the signal region, which reduces extrapolation uncertainties. The mathematical formulation of the method is given as:

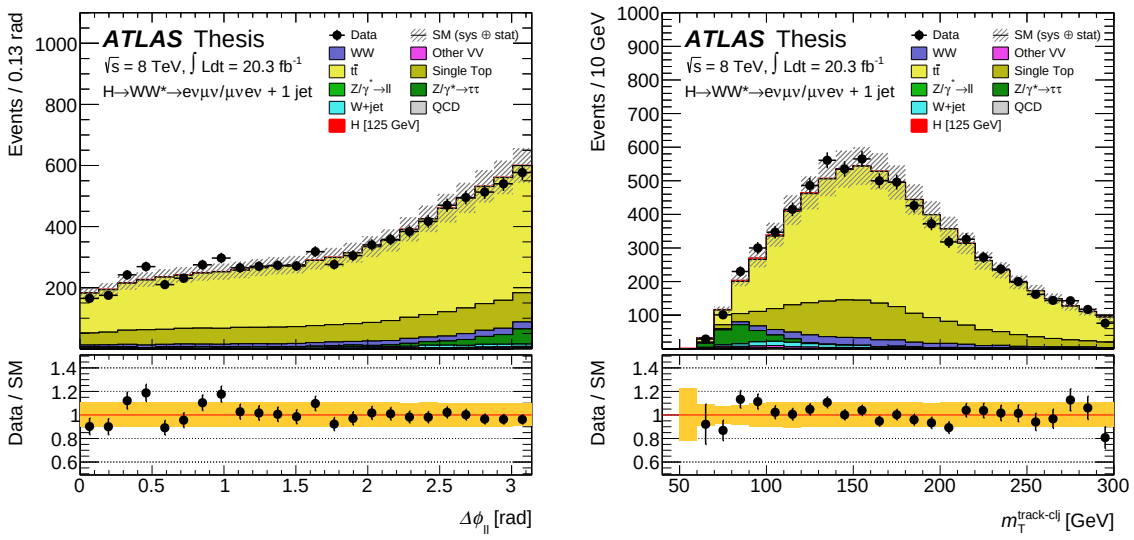
$$\epsilon_{data}^{CR,2j} = \frac{N_{data}^{2tag}}{\frac{1}{2}N_{data}^{1tag} + N_{data}^{2tag}} \tag{6.12}$$

because the number of events with only 1 b-tagged jet $N_{data}^{1tag}$ could also be accounted to the category with 2 b-jets, if the choice of jets is inverted. The measured jet b-tagging efficiency can now be applied to the 1-jet analysis. Since the signal region features a b-veto, the 1-jet control region reverts this criterion to the requirement of at least one b-tagged jet. The top background estimation is thus given as:

$$N_{Top}^{SR,1j} = \frac{N_{data}^{CR,1j}}{\epsilon_{data}^{CR,2j} \cdot f_c} \times (1 - \epsilon_{data}^{2jCR} \cdot f_c) \tag{6.13}$$

with a small Monte Carlo based correction factor $f_c = \epsilon_{MC}^{CR,1j}/\epsilon_{MC}^{CR,2j}$ which addresses the extrapolation from the 2-jets measurement to the 1-jet analysis. Finally, the top background Monte Carlo simulation in the signal region is normalized to match the data-driven estimate. The resulting normalization factor is $NF_{Top}^{SR,1j} = 1.06 \pm 0.025(\text{stat})$. Systematic uncertainties are related to the use of Monte Carlo simulation in the estimate. A theoretical component due to the modelling parameters is taken into account as well as an experimental component, which is dominated by the uncertainty on the b-tagging efficiency. In the 1-jet signal region, they lead to a 5% systematic uncertainty on the top background yield.

The 1-jet top control region is illustrated in figure 75 and the event selection is listed in table 16. The $\Delta\phi_{\ell\ell}$ and the $m_T$ distributions are shown for the different flavour channel. Both show good agreement between data and simulation and a high purity of 70% $t\bar{t}$ and 21% single top events. A normalization factor of $NF_{Top}^{CR,1j} = 1.02 \pm 0.015(\text{stat})$ has been applied to both these Monte Carlo processes to simplify the shape comparison. However, this factor does not enter the signal region, since only the data of the 1-jet control region is used to estimate the top contribution in the analysis.




(a) (b)

Figure 75: (a) $\Delta\phi_{\ell\ell}$ and (b) $m_T$ distributions of the different flavour channel for the 1-jet top control region. A normalization factor on the $t\bar{t}$ and single top simulation has been applied.

| Control Region | Summary | | | Background Composition | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Data | Signal | Bkgd | $WW$ | $t\bar{t}$ | st | diboson | $Z \to \ell\ell$ | $Z \to \tau\tau$ | W+jets | QCD |
| Top CR (incl) | 51856 | 71 | 51856 | 598 | 45598 | 3672 | 150 | 23 | 1156 | 536 | 122 |
| 1j: Top CR | 7941 | 21 | 7908 | 301 | 5755 | 1478 | 59 | 4.1 | 187 | 107 | 16 |

Table 16: Event selection of the top control region for the different flavour channels in the 0-jets and the 1-jet analyis. The observed number of events is listed in the second column, signal and background processes refer to the MC prediction. Additional normalization factors derived from other background control regions as well as $NF_{Top}$ have already been applied.

### 6.5.3 Same Sign Control Region

One of the very basic criteria in the $H \rightarrow WW \rightarrow \ell\nu\ell\nu$ analysis is the presence of two oppositely charged leptons in the final state. Charge conservation requires the electrically neutral Higgs boson to decay into a $W^+W^-$ pair and the subsequent leptonic W-decays cannot change the sign of charge. The same sign control region utilizes the reversion of this criterion to build an event sample almost free from the signal process and backgrounds like WW, top and $Z/\gamma^* \rightarrow \ell\ell$, which enter mainly due to charge flips. Therefore, it rather enhances the collection of the other diboson contributions as well as W+jets and QCD events. This enables the extraction of a normalization factor for the so called non-WW diboson backgrounds $W\gamma$, $W\gamma^*$, $WZ$ and $ZZ$.

The key factor is their symmetry of the opposite sign (OS) and same sign (SS) distributions. For the $W\gamma$, $W\gamma^*$ and $WZ$ it is equally likely to produce a second lepton of either positive or negative electrical charge in addition to the lepton originating from the W-decay. If a real photon converts to an electron for example, the detection efficiency is independent of its charge. So, as the kinematics and normalization are identical, they share the same phase space. In case of the decay of a virtual photon or a Z-boson to a same flavour lepton pair on the other hand, there are several possible scenarios of collecting two leptons and missing another. Thus, only for the different flavour channels $e\mu$ and $\mu e$, the OS-SS symmetry sustains. For the $ZZ^{(*)}$ process, the symmetry is broken if at least one of the Z-bosons decays via a $\tau$-lepton. However, its contribution to the control region is below 1% and even less in the signal region.

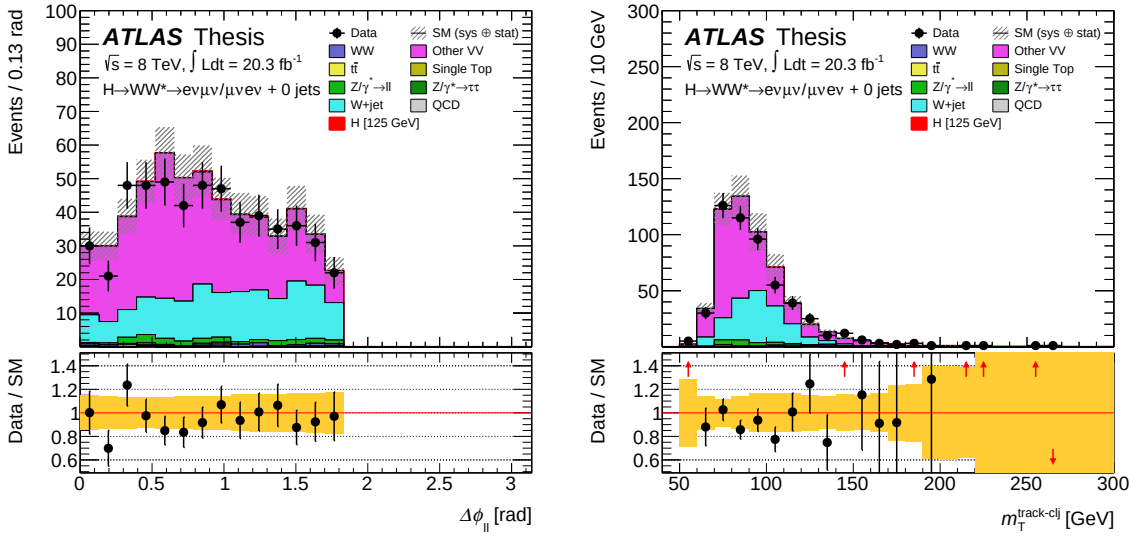
So for the different flavour channel, the same sign control region can be used to determine the normalization of the non-WW background. It is built from all events passing the exact same cuts as the signal region up to the $\Delta\phi_{\ell\ell}$ requirement, only the opposite sign selection is replaced by the same sign criterion. But the collected statistics are to low to further distinguish the different non-WW diboson processes with the current dataset. That is why only a global scale factor for the combined four $W\gamma$, $W\gamma^*$, $WZ$ and $ZZ$ processes is calculated with:

$$NF_{nonWW} = \frac{Data^{SS} - MC^{SS}_{other}}{MC^{SS}_{nonWW}} \tag{6.14}$$

and applied to the signal region. In case of the same flavour channels $ee$ and $\mu\mu$, the same sign control region is only used to validate the Monte Carlo simulation. The symmetry assumption between OS and SS is not applicable and the statistics are significantly lower. A general agreement of data and simulation can be observed.

Figure 76 shows the $\Delta\phi_{\ell\ell}$ and $m_T$ distributions in the same sign control region for the different flavour channel of the 0-jets analysis. The purity of the combined non-WW diboson processes is about 60%. The W+jets background still contributes about one third to the same sign control region. Moreover, it is not symmetric between OS and SS. But since it is determined directly from data, the final extrapolation uncertainty from

same sign to opposite sign is negligibly small, compared to the statistical uncertainties. The derived normalization factor of $NF^{0j}_{nonWW} = 0.92 \pm 0.07(\text{stat})$ is already applied to the non-WW diboson Monte Carlo simulation, which describes the data reasonably well.

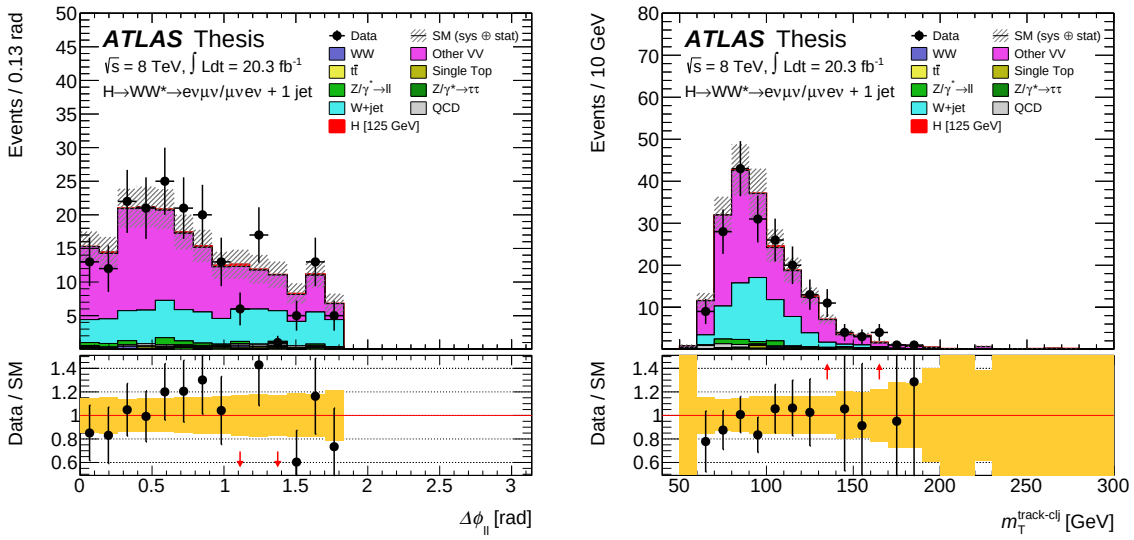



(a) (b)

Figure 76: (a) $\Delta\phi_{\ell\ell}$ and (b) $m_T$ distribution of the different flavour channel shown in the 0-jets same sign control region. The normalization factor on the diboson Monte Carlo simulation has been applied.

The same distributions are shown in figure 77 for the 1-jet analysis. In this case, the normalization factor is found to be $NF^{1j}_{nonWW} = 0.96 \pm 0.12(\text{stat})$. It is already applied to the non-WW simulation, the W+jets contribution is estimated with the data-driven method. The composition of events is very similar to the 0-jets channel with a purity of 61% diboson and 32% W+jets background. However, the number of events in the 1-jet same sign control region is much less. Thus, the statistical uncertainty on the normalization factor is larger.

Both normalization factors described above inherit systematic uncertainties, of course. But since the symmetry between the same charge and the opposite charge selection has been validated, no uncertainty on the extrapolation itself is applied. The only contributions taken into account for the diboson background are the theoretical uncertainties on the cross sections of the different processes and their correlation among the jet bins.




(a) (b)

Figure 77: (a) $\Delta\phi_{\ell\ell}$ and (b) $m_T$ distributions of the different flavour channel shown in the 1-jet same sign control region. The normalization factor on the diboson Monte Carlo simulation has been applied.

### 6.5.4 Drell-Yan Background Estimation

The Drell-Yan background production $pp \to Z/\gamma^* \to \ell\ell$ ($\ell = e, \mu, \tau$) with its large cross section can be reconstructed with a significant amount of missing transverse energy. These events can either contain real missing transverse energy due to neutrinos from leptonic $\tau$-decays or they enter the analysis as mismeasurement because of the degraded resolution in the high pileup environment. Both cases are addressed separately in the $H \to WW \to \ell\nu\ell\nu$ analysis. The following sections describe the normalization strategies for the Drell-Yan background. While the $Z/\gamma^* \to \tau\tau$ contributions are investigated in a dedicated different flavour control region, the $Z/\gamma^* \to ee/\mu\mu$ background is derived with a data-driven technique in the same flavour channel.

#### $Z/\gamma^* \to \ell\ell$ Estimation in the Same Flavour Channel

The same flavour channel is dominated by the Drell-Yan background $Z/\gamma^* \to ee/\mu\mu$. To derive a data-driven estimate for this contribution, a technique called Pacman[17] has been developed. The idea is to exploit the shape of the $f_{recoil}$ distribution, which shows a significant difference between the Drell-Yan (DY) and other processes (nonDY), with a template fit technique. The templates are extracted directly from data in Z-enriched and Z-depleted control regions and their normalizations are used as fit parameters. Figure 78 illustrates the strategy of the Pacman method.

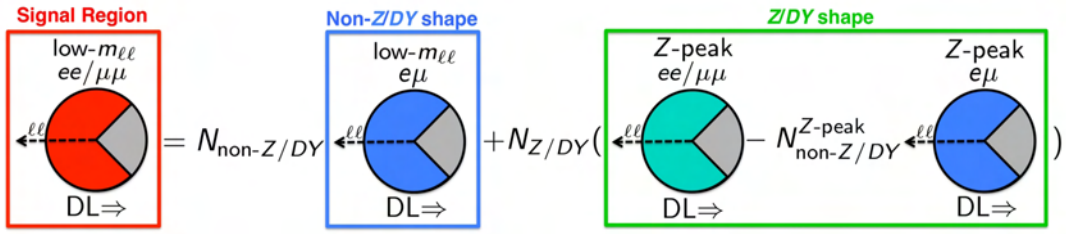



Figure 78: Scheme of the Pacman procedure, illustrating the data-driven $Z/DY$ background estimation [104].

Because the $f_{recoil}$ cut is used to define the same flavour signal region, the implementation of the Pacman method used in the $H \to WW \to \ell\nu\ell\nu$ analysis is rather simple. Instead of exploiting the full shape of the $f_{recoil}$ distribution, only two bins are used: passing the cut or rejected by the cut. This way, the templates are translated to cut efficiencies and the fit is reduced to an analytical equation.

The non-Drell-Yan contributions are extracted from the different flavour channel. They are nearly 100% pure in nonDY, but very close to the same flavour signal region. For the most of these processes, their relative fraction is identical in same and different flavour,

[17] The name is borrowed from the character of the US arcade game Pac-Man, because of its similarity with the scheme in figure 78.

which is sometimes referred to as flavour universality. The cut-efficiency of $f_{recoil}$ has to be measured for the signal region (SR) characterized by its low-$m_{\ell\ell}$ selection and a Z-enriched region (Z-peak) with the requirement of $|m_{\ell\ell} - m_Z| < 15$ GeV:

$$\epsilon_{nonDY} = \frac{NO_{DF}^{SR}}{N_{DF}^{SR}} \quad \text{and} \quad \epsilon_{nonDY}^{Z-peak} = \frac{NO_{DF}^{Z-peak}}{N_{DF}^{Z-peak}} \tag{6.15}$$

where $N_{DF}^{SR/Z-peak}$ counts all data events before the cut, and $NO_{DF}^{SR/Z-peak}$ denotes the number of observed data events passing the cut in the different flavour channel. They describe the blue contributions in the Pacman scheme in figure 78.

The Drell-Yan cut-efficiency $\epsilon_{DY}$ is extracted from the Z-peak control region. It is symbolized by the green box in the Pacman scheme. But due to the various cuts applied to the same flavour channel deselecting the $Z/\gamma^* \to \ell\ell$ background like those operating on the missing transverse energy, the purity in the Z-window is only about 50%. Therefore, the non-Drell-Yan contributions need to be subtracted using their Monte Carlo expectation $N(\text{nonDY})_{SF}^{Z-peak}$. Then, the efficiency of the $f_{recoil}$ cut on the Drell-Yan background can be measured from the same flavour data in the Z-peak control region:

$$\epsilon_{DY} = \frac{NO_{SF}^{Z-peak} - \epsilon_{nonDY}^{Z-peak} \times N(\text{nonDY})_{SF}^{Z-peak}}{N_{SF}^{Z-peak} - N(\text{nonDY})_{SF}^{Z-peak}} \tag{6.16}$$

So now, with the efficiencies measured in data and the observed number of data events before and after the cut on $f_{recoil}$, the following system of two equations can be solved analytically to derive the Drell-Yan yield in the same flavour signal region. Only $\epsilon_{DY}$ depends on the non-Drell-Yan Monte Carlo simulation.

$$N = N_{DY} + N_{nonDY} \tag{6.17}$$

$$NO = \epsilon_{DY} \times N_{DY} + \epsilon_{nonDY} \times N_{nonDY} \tag{6.18}$$

The resulting estimate of Drell-Yan events in the same flavour signal region $N_{DY}^{SR}$ is:

$$N_{DY}^{SR} = \epsilon_{DY} \times \frac{NO_{SF}^{SR} - \epsilon_{nonDY} \times N_{SF}^{SR}}{\epsilon_{DY} - \epsilon_{nonDY}} \tag{6.19}$$

Thus, a normalization factor for the Monte Carlo simulation is derived to match the estimate of the Pacman method $NF_{DY} = N_{DY}^{SR}/N_{DY,MC}^{SR}$. For the 0-jets analysis it is found to be $NF_{DY}^{0j} = 2.15 \pm 1.05(\text{tot.})$. In the 1-jet analysis the factor is derived from the $f_{recoil}^{ext}$ distribution. All other technicalities are exactly the same. The result of the measurement is $NF_{DY}^{1j} = 1.6 \pm 0.7(\text{tot.})$. In both cases, the normalization is applied to the same flavour channels.

Systematic uncertainties related to the $f_{recoil}$ selection efficiencies are taken into account. For the non-DY components, they are dominated by the statistical uncertainties and the extrapolation from the different flavour to the same flavour selection, which is evaluated by comparing different Monte Carlo samples. They sum up to a maximum of 4.5% uncertainty reached in the 1-jet channel. The same Monte Carlo generator comparison is performed to evaluate the uncertainties on $\epsilon_{DY}$. Here, they account for the extrapolation from the Z-peak region to the signal region and the related statistics. These components add up to an uncertainty of more than 30% in both jet multiplicities. In total, the systematic uncertainty on the Drell-Yan background yield is 49% (45%) in the 0-jets (1-jet) selection.

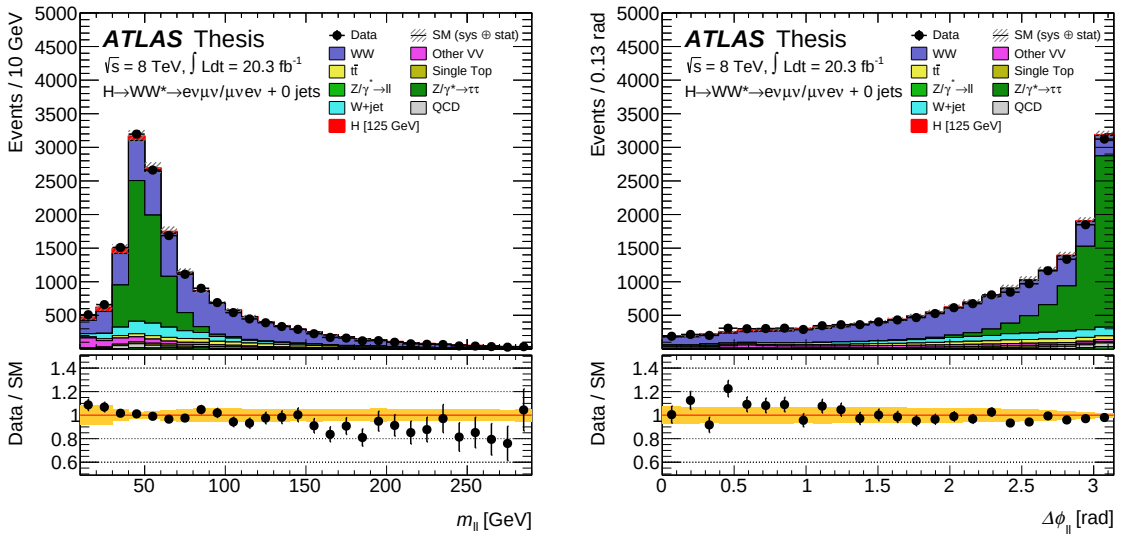
### $Z/\gamma^* \to \tau\tau$ Control Region in the Different Flavour Channel

The $Z/\gamma^* \to \tau\tau$ background contribution is estimated in a dedicated control region. In the same flavour channels, it is nearly impossible to separate it from the other Drell-Yan processes. So normalization factors are derived in the different flavour channels $e\mu$ and $\mu e$ only. Subsequently, they are applied to all lepton flavour combinations.

After the full preselection and the jet veto, the 0-jets analysis is dominated by WW background. Two additional requirements are imposed to enhance the $Z/\gamma^* \to \tau\tau$ contributions. Both exploit the back-to-back nature of the $Z/\gamma^*$ decay. The first one is the selection of $m_{\ell\ell} < 80$ GeV and the second requirement sets a high boundary on $\Delta\phi_{\ell\ell} > 2.8$ radians. Figure 79 shows both distributions at the jet veto stage before the definition of the $Z/\gamma^* \to \tau\tau$ control region.

Applying the cuts on $m_{\ell\ell}$ and $\Delta\phi_{\ell\ell}$ reduces the dominant WW background to only 3%. The $Z/\gamma^* \to \tau\tau$ purity in the $e\mu + \mu e$ channel is above 90% with negligible contributions from $Z/\gamma^* \to \ell\ell$. Table 17 lists the number of selected events per process. Figure 80 presents both distributions after the cuts in the $Z/\gamma^* \to \tau\tau$ control region. The Monte Carlo simulation agrees well with the observed data. The top, WW and other diboson background contributions have been normalized with the scale factors derived in their corresponding control regions, while the W+jets process is estimated with the data-driven technique. Subtracting all these processes from data and comparing the result to the $Z/\gamma^* \to \tau\tau$ simulation leads to the normalization factor of $NF^{0j}_{Z\tau\tau} = 0.998 \pm 0.018(\text{stat})$. This is imposed on all lepton flavour combinations in the 0-jets analysis.

The 1-jet analysis uses a similar concept. After the cut on $\max(m_T^W)$ in the one jet selection, two requirements are used to build the $Z/\gamma^* \to \tau\tau$ control region. First, an $m_{\ell\ell} < 80$ GeV cut is imposed to enhance the back-to-back feature of the $\tau$-leptons (see section 6.1), identical to the 0-jets case. The second requirement exploits the distribution of $m_{\tau\tau}$ reconstructed with the collinear approximation. In the 1-jet analysis, this method leads to a reasonably well reconstructed mass resolution, because the $\tau\tau$ system is likely to be boosted. So a low bound on $m_{\tau\tau} > 66$ GeV helps to suppress the dominant top and




(a) (b)

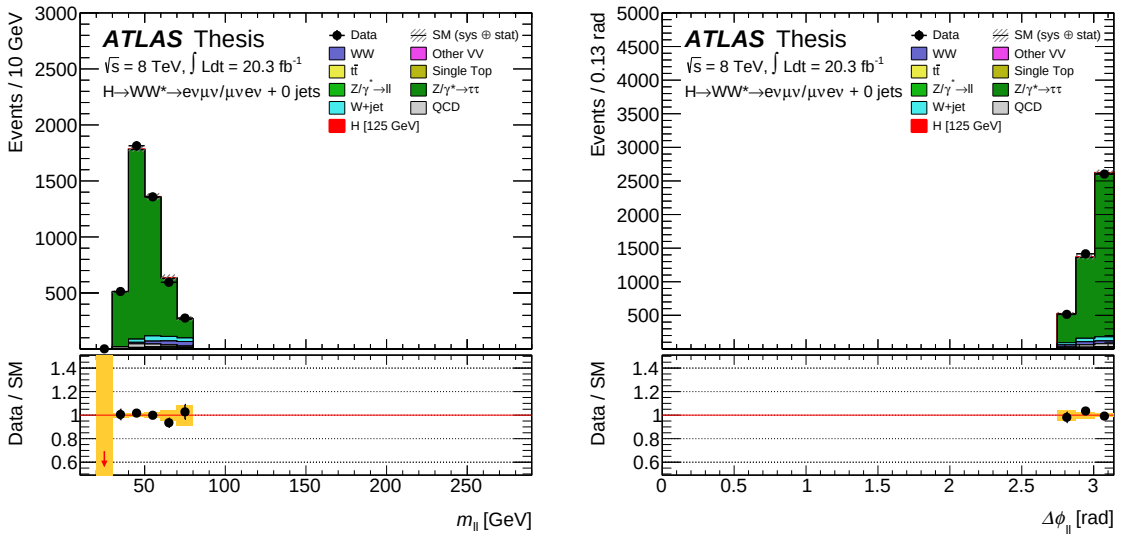
Figure 79: Distributions of (a) $m_{\ell\ell}$ and (b) $\Delta\phi_{\ell\ell}$ in the different flavour channels of the 0-jets analysis at the jet veto stage. Normalization factors on WW, other diboson and top background have already been applied.

WW backgrounds. Figure 81 shows both distributions after the $\max(m_T^W)$ cut. Again, the contributions of top, WW and other diboson backgrounds have been scaled with the factors derived in their control regions.

In the 1-jet $Z/\gamma^* \to \tau\tau$ control region, the purity is 80% with only 0.4% contributions from $Z/\gamma^* \to \ell\ell$. The number of events per process is listed in table 17. The $m_{\ell\ell}$ and $m_{\tau\tau}$ distributions after the application of the cuts are shown in figure 82. In both cases, the Monte Carlo simulation describes the data very well. The derived normalization factor of $NF^{1j}_{Z\tau\tau} = 1.05 \pm 0.036(\text{stat})$ is already applied, as are those of the WW, diboson and top background processes. It is used for all lepton flavour combinations in the 1-jet signal region.

| Control Region | Summary | | | Background Composition | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Data | Signal | Bkgd | $WW$ | $t\bar{t}$ | st | diboson | $Z \to \ell\ell$ | $Z \to \tau\tau$ | W+jets | QCD |
| 0j: $Z\tau\tau$ CR | 4557 | 22 | 4535 | 117 | 11 | 5.0 | 33 | 28 | 4102 | 146 | 93 |
| 1j: $Z\tau\tau$ CR | 1540 | 18 | 1516 | 99 | 56 | 19 | 27 | 6.6 | 1225 | 64 | 20 |

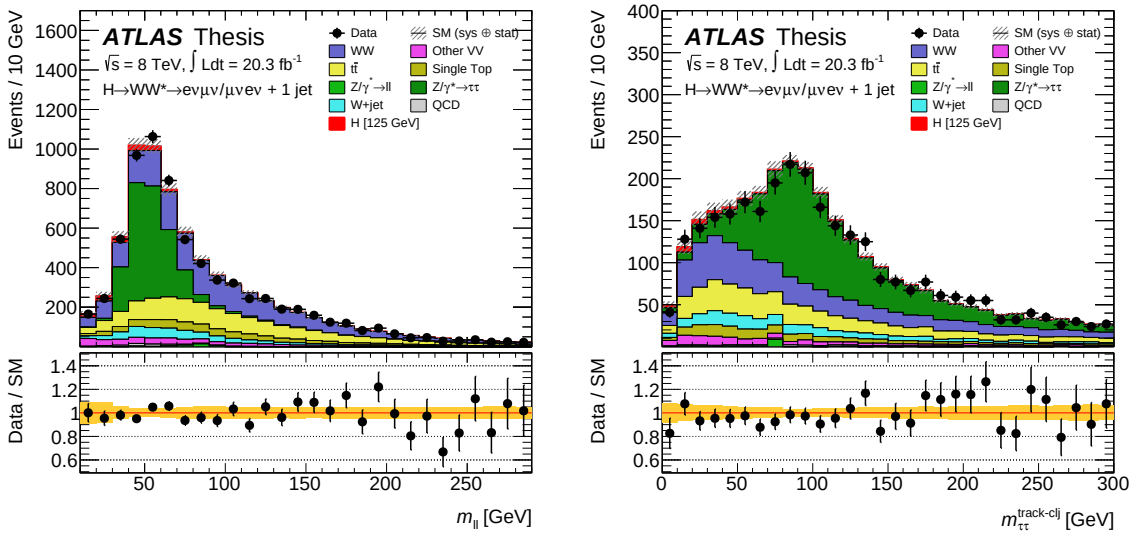
Table 17: Event selection of the $Z \to \tau\tau$ control region for the different flavour channels in the 0-jets and the 1-jet analyis. The observed number of events is listed in the second column, signal and background processes refer to the MC prediction. Additional normalization factors derived from other background control regions as well as $NF_{Z\tau\tau}$ have already been applied.




(a) (b)

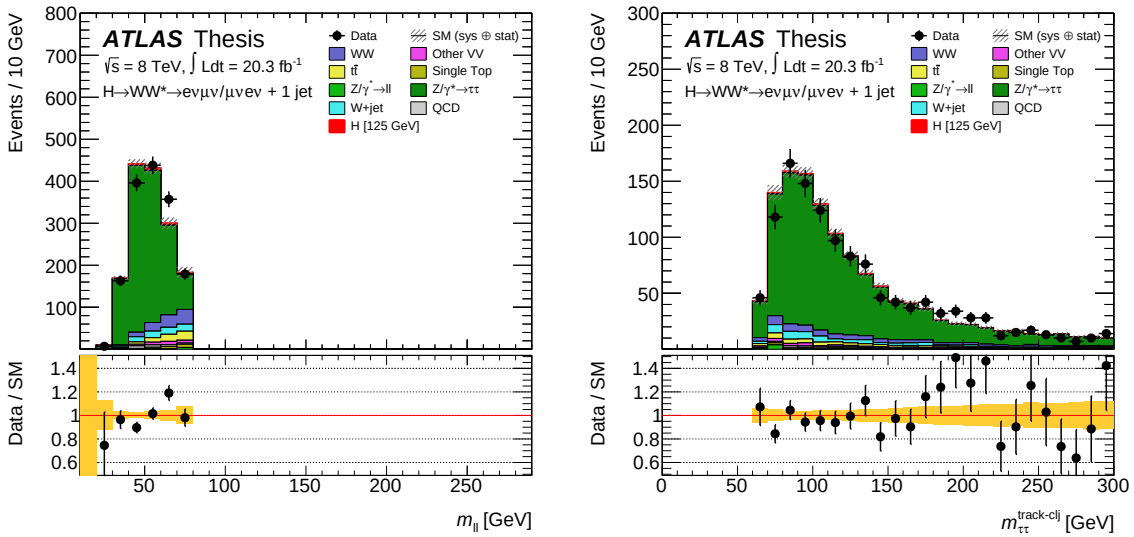
Figure 80: Distributions of (a) $m_{\ell\ell}$ and (b) $\Delta\phi_{\ell\ell}$ in the different flavour channels of the 0-jets analysis in the $Z/\gamma^* \to \tau\tau$ control region. Normalization factors on $Z/\gamma^* \to \tau\tau$, WW, other diboson and top background have already been applied.

The systematic uncertainties on the extrapolation from the $Z/\gamma^* \to \tau\tau$ control region to the signal region consider variations of the QCD scale and the PDF sets. Effects of the generator modelling are evaluated by comparing ALPGEN+HERWIG and ALPGEN+PYTHIA samples. In case of the 0-jets channel, an additional uncertainty on the $p_T^Z$ reweighting is applied. This component dominates the systematic uncertainties on the background yield in the signal regions with contribution of 19%.




(a) (b)

Figure 81: Distributions of (a) $m_{\ell\ell}$ and (b) $m_{\tau\tau}$ in the different flavour channels after the cut on $\max(m_T^W)$ in the 1-jet selection. Normalization factors on WW, other diboson and top background have already been applied.




(a) (b)

Figure 82: Distributions of (a) $m_{\ell\ell}$ and (b) $m_{\tau\tau}$ in the different flavour channels of the 1-jet analysis in the $Z/\gamma^* \to \tau\tau$ control region. Normalization factors on $Z/\gamma^* \to \tau\tau$, WW, other diboson and top background have already been applied.

### 6.5.5 W+jets Estimation

The production of W-bosons in association with jets leads to a significant background contribution in the $H \to WW \to \ell\nu\ell\nu$ analysis. One isolated lepton and momentum imbalance due to a neutrino result from the leptonic W-decay. Together with a second, non-prompt lepton emerging either from the decay of hadrons that contain heavy quarks or from the fragmentation of quarks or gluons, these objects mimic the expected Higgs signal final state. Those leptons, that are reconstructed from jets are referred to as fake leptons, or simply fakes. Although the probability of such mis-identifications is pretty low, the large production cross section of the W+jets background yields a substantial contribution. The same holds for the QCD background described in section 6.5.6, which also arises from non-prompt leptons reconstructed from mis-identified jets. So the estimates of these contributions are closely related to one another. Thus, the same methods apply for both. The processes leading to the mis-identification of a jet as a lepton are hard to model in simulation. Therefore, the W+jets and the QCD estimate are derived directly from data and do not rely on Monte Carlo predictions.

#### The Fake Factor Method

The data-driven approach of estimating the W+jets background features two key aspects that have already been introduced before: a control region with enhanced W+jets contribution and an extrapolation factor connecting it to the signal region of the Higgs analysis. But here, they follow a special strategy, focusing on the selection of the non-prompt and fake leptons. So both depend on the definition of the objects that enter the analysis through mis-identification and fulfill only a loosened set of identification and isolation criteria. At first, the control region for the W+jets background needs to be specified. The event selection criteria stay the same as for the signal region, but the different object definition collects a completely orthogonal event sample. The W+jets control region is built from one isolated lepton as defined for the signal region (described in sections 4 and 5.5.1) and one lepton emerging from jet mis-identification. The first category is referred to as ID lepton, the latter is called anti-ID lepton since it is requested to fail the full identification scheme. This way, the other dilepton backgrounds contributing to the signal region and the signal itself are reduced significantly and leave a very pure environment. But to make use of this control region and estimate the W+jets background in the signal region of the Higgs analysis, an extrapolation factor addressing the distinction of the different lepton definitions is needed. This factor is derived from data of an independent jet-enriched sample. It is simply called fake factor, defined as:

$$f_\ell = \frac{N_{ID}}{N_{anti-ID}} \tag{6.20}$$

which is evaluated for electrons and muons separately. $N_{ID}$ denotes the number of ID leptons found in the jet-enriched sample and $N_{anti-ID}$ counts the number of objects

fulfilling the looser criteria of the anti-ID selection[18]. Therefore, the identified objects are sometimes simply referred to as numerators and the anti-ID candidates as denominators.

The anti-ID lepton definition aims to select non-prompt and fake leptons produced by hadronic activity over so called true leptons from W- or Z-decays. Although the selection is exclusive, it has to remain very close to the full identification chain imposed on the signal region. Otherwise, the extrapolation would be too large and only apply for unfeasible objects. For anti-ID electrons, the basic criteria like the $p_T > 10$ GeV and $\eta < 2.47$ range as well as the minimum number of hits in the SCT and the pixel detector and the criteria on the longitudinal and transverse impact parameter are exactly the same as for the ID electrons. But the calorimeter isolation is loosened to $E_T^{cone,30}/E_T < 0.30$ along with the track isolation, which is loosened to $p_T^{cone,30}/p_T < 0.16$ independently of the objects transverse momentum. Furthermore, the requirement of the conversion flag and a hit in the innermost layer of the pixel detector (b-layer) are dropped and the medium identification chain needs to be failed explicitly. But the latter argument removes the possibility of an anti-ID electron to fire the single electron trigger. So to include the case of a fake electron triggering the event instead of an identified lepton, a second version of the anti-ID definition is approached and outlined as triggered anti-ID electron. The requirements on the calorimeter and track isolation are removed completely and the criterion on the longitudinal impact parameter is loosened to $|z_0 \times \sin\Theta| < 1.2$ mm, the one on the transverse impact parameter is relaxed to $|d_0/\sigma_{d_0}| < 9$. But the medium identification chain needs to be passed, which includes requirements on the conversion flag and b-layer hits and finally enables the single electron trigger used for the analysis. Table 18 summarizes the important aspects of the anti-ID and the triggered anti-ID definition and displays the differences compared to the full identification scheme.

| Electron Parameter | Anti-ID | Triggered Anti-ID | ID |
|---|---|---|---|
| $\lvert z_0 \times \sin\Theta\rvert$ | $< 0.4$ mm | $< 1.2$ mm | $< 0.4$ mm |
| $\lvert d_0/\sigma_{d_0}\rvert$ | $< 3$ | $< 9$ | $< 3$ |
| $E_T^{cone,30}/E_T$ | $< 0.30$ | removed | $(0.20 - 0.28)$ |
| $p_T^{cone,30}/p_T$ | $< 0.16$ | removed | $(0.06 - 0.10)$ |
| conversion flag + b-layer hit | removed | removed | required |
| medium identification scheme | fail | pass | included in LH |
| ID electron selection | fail | fail | pass |

Table 18: Requirements on the anti-ID and triggered anti-ID electron objects compared to the full signal selection identification scheme. Numbers in parentheses refer to variation for different ranges of lepton $p_T$.

[18]Since the lepton definitions are not a subset of each other but are exclusive, the ratio $f$ is not a fake rate. It is rather a ratio of fake rates and thus does not explicitly describe a probability of mis-identification.

A similar prescription holds for the case of the muons. Anti-ID muons need to be reconstructed with the STACO algorithm, which combines the independent track measurements of the muon system and the inner detector. The standard requirements on $p_T > 10$ GeV and $\eta < 2.5$ are also the same as for the ID muons. While the criterion on the longitudinal impact parameter is preserved, the one on the transverse impact parameter is removed. Furthermore, the calorimeter isolation is relaxed for the different transverse momentum regions and the track isolation is dropped completely. But the full identification chain imposed on the signal region muons needs to be failed. To include the cases of fake muons firing the trigger, a second definition of triggered anti-ID muons is developed. It only applies to objects above $p_T > 25$ GeV. The calorimeter isolation requirements are removed completely, but a loose track isolation of $p_T^{cone,30}/p_T < 0.12$ is imposed. Both categories are summarized in table 19 and compared to the ID muon requirements.

| Muon Parameter | Anti-ID | Triggered Anti-ID | ID |
|---|---|---|---|
| $p_T$ threshold | $> 10$ GeV | $> 25$ GeV | $> 10$ GeV |
| $\lvert d_0/\sigma_{d_0}\rvert$ | removed | removed | $< 3$ |
| $E_T^{cone,30}/E_T$ at [10,15] | $< 0.15$ | removed | $< 0.06$ |
| $E_T^{cone,30}/E_T$ at [15,20] | $< 0.25$ | removed | $< 0.12$ |
| $E_T^{cone,30}/E_T$ at [20+] | $< 0.30$ | removed | $(0.18 - 0.30)$ |
| $p_T^{cone,30}/p_T$ | removed | $< 0.12$ | $(0.06 - 0.12)$ |
| ID muon selection | fail | fail | pass |

Table 19: Reqirements on the anti-ID and triggered anti-ID muon objects compared to the full signal selection identification scheme. Numbers in parentheses refer to variation for different ranges of lepton $p_T$.

The introduced anti-ID objects are only very loosely isolated by construction, if at all. So the overlap removal procedure between all the different objects used in the analysis needs to be revisited, especially for the jets. It basically follows the rules already introduced for the ID leptons introduced in section 5.5.1. The distance between lepton candidates has to fulfill $\Delta R > 0.1$. Muons are always favoured versus electrons, independently of their affiliation to the ID or anti-ID category. In case of two overlapping electron objects, the one with higher transverse energy is kept. This applies also to anti-ID electrons. Even if the one with lower $E_T$ is fully identified, it is nevertheless removed from the object list. The separation of leptons and jets is a crucial point for the W+jets and QCD background estimation, since it addresses exactly the mis-identifiaction of one object with the other. If a jet overlaps with an ID or anti-ID electron within $\Delta R < 0.3$, the jet is always removed. But for muons, the consistency of this procedure with respect to the signal region is not retained. Fully identified muons are removed from the object list in favour of the overlapping jet. But the anti-ID muons are preferred versus jets, although they most often emerge from heavy flavour decays. However, they need to be retained to enlarge the small population in the jet-enriched sample, which is used to derive the muon extrapolation fake factor.

### Z+jets Fake Factor

Both lepton fake factors for the W+jets estimation are measured from jets in events with a Z-boson candidate of a dedicated Z+jets sample. It is collected by requiring an opposite charge electron or muon pair with a dilepton invariant mass of $81 < m_{\ell\ell} < 107$ GeV using the same triggers as the signal region selection. To reduce the real contamination of leptons from electroweak processes, several objections are imposed. A ZZ-veto for example rejects events with a second lepton pair in a window of $76 < m_{\ell\ell} < 107$ GeV. Contributions of WZ events are suppressed with a veto on a third lepton with $m_T^W > 30$ GeV. But these criteria are imposed on any lepton candidate with transverse momentum greater than 7 GeV without further identification or isolation requirements. This way, both the ID and anti-ID definitions are included. About 80% of these background contributions are suppressed by those kinematic criteria. Remaining contributions are subtracted from data using the Monte Carlo simulation prediction. Figure 83 shows the $p_T$ distributions of the muon and electron categories in the Z+jets sample before Monte Carlo subtraction. The systematic errors on the background simulation are dominated by the cross section uncertainties of the respective processes.

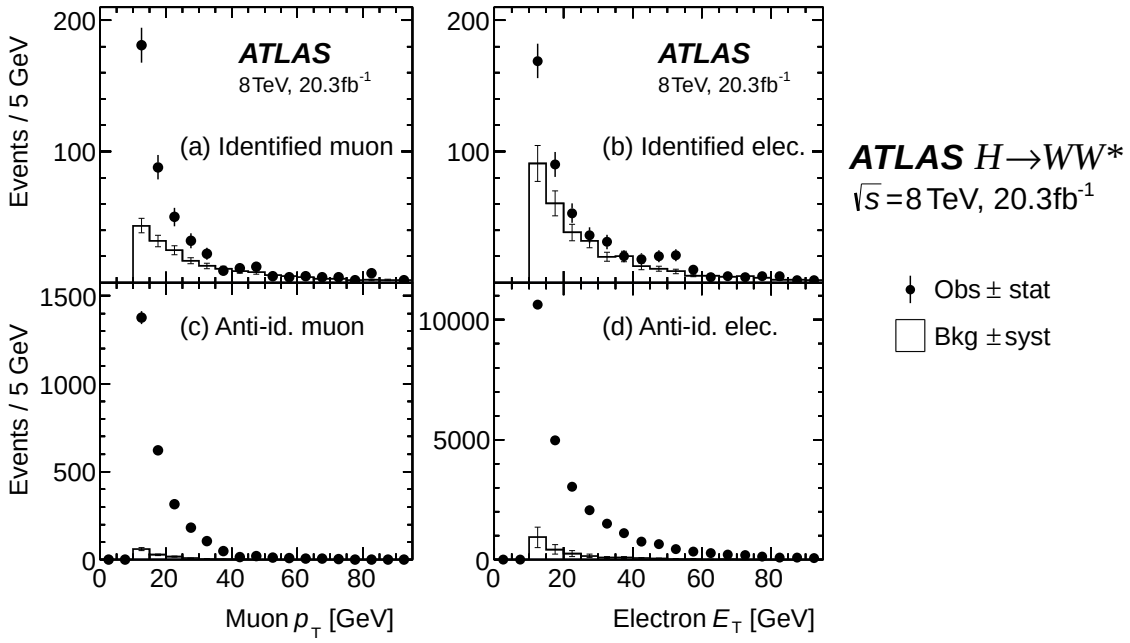



Figure 83: Misidentified lepton $p_T$ distributions in the Z+jets sample: (a) ID muon, (b) ID electron, (c) anti-ID muon, and (d) anti-ID electron [101]. The data points are compared to the sum of the electroweak processes other than the Z-boson production in association with jets, predicted by the Monte Carlo simulation.

The ratio of the identified and the anti-ID value per bin gives the $p_T$ dependent fake factor. It is shown in figure 84 for the muons on the left and the electrons on the right. The grey bands symbolize the systematic uncertainties associated to the measurement.

The statistical component refers to the limited number of objects found in the Z+jets data sample. Another contribution results from the uncertainties on the Monte Carlo background subtraction. The expected difference in sample composition between Z+jets and W+jets events is split into an opposite charge (OC) component, reflecting the signal region selection and a same charge (SC) component, derived for the same sign control region (see section 6.5.3).

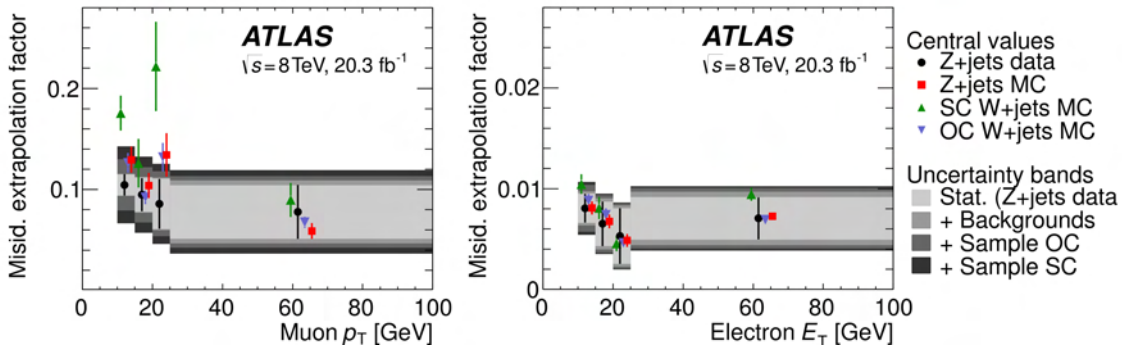



Figure 84: Misidentified lepton extrapolation factors for anti-ID muons on the left and electrons on the right [101]. The data is compared to the Monte Carlo simulation of Z+jets, opposite-charge (OC) W+jets, and same-charge (SC) W+jets which are used to derive additional MC correction factors. The different bands represent the related uncertainties.

An additional correction factor is calculated to account for this so called sample dependence by taking the ratio of the fake factors derived from W+jets and the Z+jets Monte Carlo simulation. To obtain the systematic uncertainty of this correction, it is evaluated with three different Monte Carlo generators. The baseline generator Alpgen+Pythia6 is shown for Z+jets (red) and W+jets (blue and green) in figure 84. It is compared to the different matrix element and parton shower simulations of Alpgen+Herwig and Powheg+Pythia8. But while some of the contributing processes are charge symmetric, especially the W production in association with a charm quark shows a significant charge asymmetry (see the Feynman graphs in figure 51 (b) and (c)). The asymmetry is investigated in [105] and [106]. The higher population in the opposite charge selection leads to smaller fake factors compared to the same charge case. A detailed study of the flavour composition is performed in [103]. The resulting Monte Carlo correction factors in the opposite charge selection are $c_e^{OC} = f_e^{W+jets}/f_e^{Z+jets} = 0.99 \pm 0.20$ for the electrons and $c_\mu^{OC} = 1.00 \pm 0.22$ for the muons. So the flavour composition of the W+jets and the Z+jets Monte Carlo simulation is very similar in the opposite charge case. It is illustrated by the red and blue points in figure 84 and shows different behaviour for the same charge selection shown in green. There, the correction factors are much larger with $c_e^{SC} = 1.25 \pm 0.31$ for the electrons and $c_\mu^{SC} = 1.40 \pm 0.49$ for the muons. The uncertainties resemble the difference of the three generators which feature a large statistical component. Since the jet composition does not show strong variation with the fake lepton transverse momentum, the correction factors are averaged over $p_T$. A complete

list of the contributions to the systematic uncertainties on the fake factor measurement is given in table 20. The total systematic uncertainty sums up in quadrature.

| Electron $E_T$ | Stat. | MC subtr. | Sample corr. | | Total | |
|---|---|---|---|---|---|---|
| | | | OC | SC | OC | SC |
| $10-15$ GeV | 18 | 11 | 20 | 25 | 29 | 32 |
| $15-20$ GeV | 34 | 19 | 20 | 25 | 44 | 46 |
| $20-25$ GeV | 52 | 25 | 20 | 25 | 61 | 63 |
| $>25$ GeV | 30 | 23 | 20 | 25 | 43 | 45 |
| Muon $p_T$ | Stat. | MC subtr. | Sample corr. | | Total | |
| | | | OC | SC | OC | SC |
| $10-15$ GeV | 10 | 3 | 22 | 35 | 25 | 37 |
| $15-20$ GeV | 18 | 5 | 22 | 35 | 37 | 46 |
| $20-25$ GeV | 29 | 8 | 22 | 35 | 37 | 46 |
| $>25$ GeV | 34 | 21 | 22 | 35 | 46 | 53 |

Table 20: Uncertainties (in %) on the extrapolation factor $f_\ell$ for the determination of the W+jets background, for electrons in the upper box and for muons in the lower one. The different contributions of the statistics in the Z+jets control sample, the MC background subtraction and the correction factor for the opposite charge (OS) and same charge (SC) W+jets control region are given as well as the total sum of systematic uncertainties.

### W+jets Control Region

The W+jets control region uses the same anti-ID lepton definitions. By selecting exactly one ID and one anti-ID lepton, the control region is constructed orthogonal to the signal region requiring exactly two ID leptons. All other criteria of the event selection run in parallel, so that no additional selection efficiencies need to be taken into account. In the signal region, the different background components can be written as:

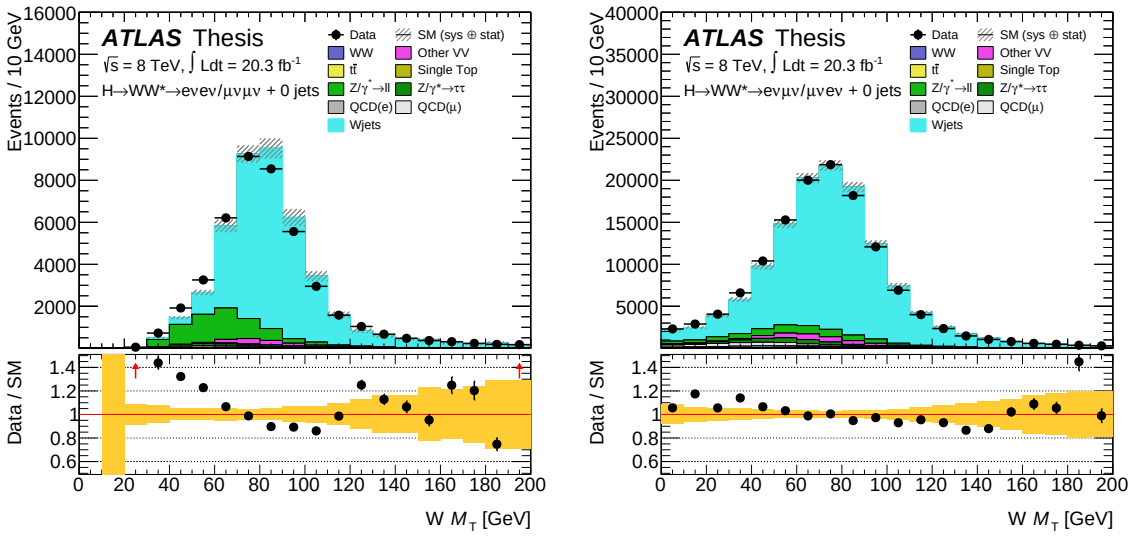
$$N_{ID+ID}^{Data} = N_{ID+ID}^{Signal} + N_{ID+ID}^{EW} + N_{ID+ID}^{QCD} + N_{ID+ID}^{W+jets} \tag{6.21}$$

The signal contribution is taken from the Monte Carlo simulation as well as the electroweak background components with the normalizations described in the previous sections. The QCD and W+jets components are derived from data. With the extrapolation factor at hand, the number of W+jets events in the signal region can be estimated via:

$$N_{ID+ID}^{W+jets} = f^{W+jets} \times (N_{ID+anti-ID}^{Data} - N_{ID+anti-ID}^{EW} - N_{ID+anti-ID}^{QCD}) \tag{6.22}$$

with the selected number of events in the W+jets control region $N_{ID+anti-ID}$. Here, $f^{W+jets}$ labels the corrected fake factor measured in the Z+jets control sample, depending on the flavour of the anti-ID lepton. Remaining contributions from other processes than W+jets are subtracted using either the Monte Carlo simulation as for the case of the electroweak backgrounds or the data-driven QCD estimate explained in detail in the following section.

Figure 85 shows the $\max(m_T^W)$ distribution for the same flavour channels on the left and the different flavour channels on the right at the jet veto cut stage of the W+jets control region. So neither the background subtraction nor the extrapolation factor have been applied. The light blue contributions illustrate the W+jets Monte Carlo simulation, which is normalized to the data for comparison. Only statistical errors are included in the histograms. The control region shows a high purity of 84% after the jet veto.



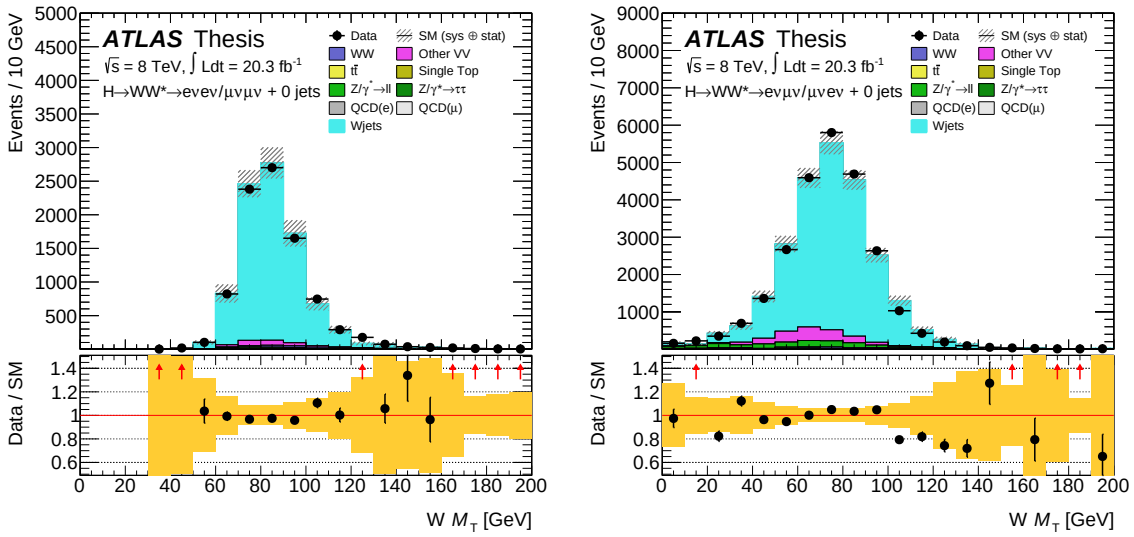

(a) (b)

Figure 85: The $\max(m_T^W)$ distributions for the W+jets control region at the jet veto cut stage for the (a) same flavour and (b) different flavour channels. The QCD background is given by the data driven estimate, while the other processes are purly Monte Carlo simulation. The W+jets Monte Carlo is normalized to data for comparison. Only statistical errors are included.

Generally, the W+jets Monte Carlo overestimates the number of events in the control region and thus the number of fake leptons except for the $\mu\mu$ channel compared to data. This normalization information is one benefit of the presented method. But the data driven estimate also investigates the shape of the distributions, which could not be realized with a simple Monte Carlo scale factor. This can be observed from the ratio histogram at the bottom of each figure comparing data to the sum of the simulated events. Both distributions in figure 85 show significant discrepancies. In this case, the W+jets Monte Carlo is shifted towards larger $\max(m_T^W)$ and systematically

overestimates the value of the observable, revealing the mis-modelling of fake leptons in simulation. Thus, the derivation of the W+jets background directly from data relieves the analysis from the weak confidence in the simulation of this complex and important process.

This is particularly of interest in the final signal region selection entering the statistical fit. Figure 86 shows the $m_T^W$ distribution for the same flavour and different flavour channels after application of the $\Delta\phi_{\ell\ell}$ requirement in the W+jets control region. The purity of 89% is very high at this stage. The biggest contribution comes from the $\mu e$ channel, because in most cases the subleading electron has been misidentified. In general, different flavour population has larger statistics, the $\mu\mu$ channel adds the lowest share.

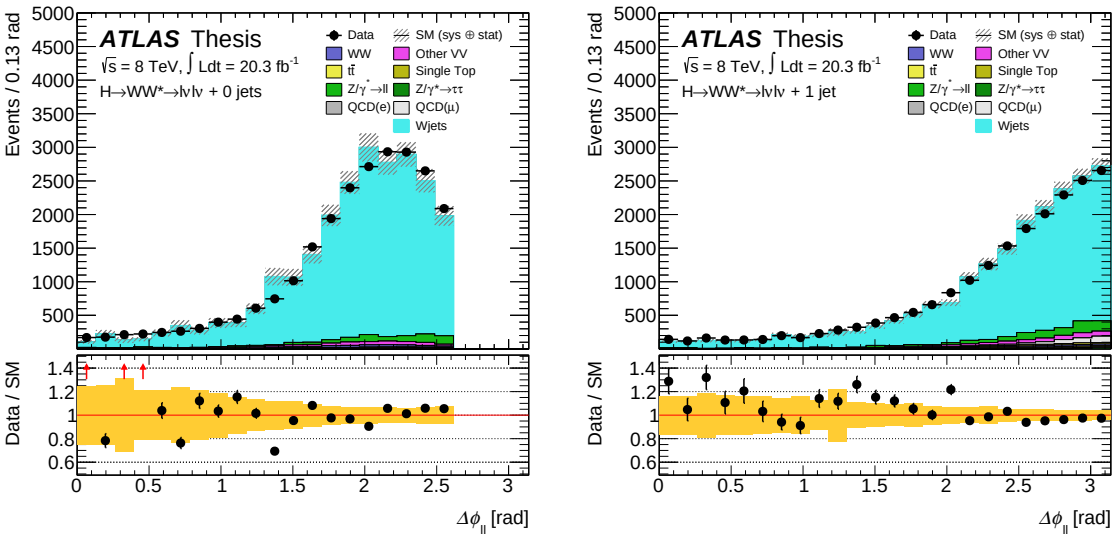



(a) (b)

Figure 86: The max($m_T^W$) distributions for the W+jets control region after the $\Delta\phi_{\ell\ell}$ requirement in the 0-jets analysis for the (a) same flavour and (b) different flavour channels. The W+jets Monte Carlo is normalized to data for comparison. Only statistical errors are included.

The data driven W+jets estimate also serves as input for other background control regions as described in the previous sections. It is extracted in exactly the same way as the signal region contribution. The requirements imposed to select the dedicated background are also applied to the sample containing one ID and one anti-ID lepton.

One example is illustrated in figure 87. It shows the $\Delta\phi_{\ell\ell}$ distributions with the criteria of the WW control region described in section 6.5.1 for the 0-jets and the 1-jet analysis but applied to the W+jets control sample. Again, these selections are very pure with 92% in the 0-jets and 87% in the 1-jet analysis. The appliance of equation 6.22 then leads to the estimated W+jets contribution in the WW control region. This works for all




(a) (b)

Figure 87: $\Delta\phi_{\ell\ell}$ distributions with the WW control region requirements applied to the W+jets control sample for the (a) 0-jets and (b) 1-jet analysis. The W+jets Monte Carlo is normalized to data for comparison. Only statistical errors are included.

of the different background estimations mentioned before. In case of the other diboson processes $W\gamma$, $W\gamma^*$, $WZ$ and $ZZ$ derived from the same sign control region, the same charge extrapolation factor $f^{SC}$ is used.

### 6.5.6 QCD Estimation

The QCD estimate is also derived from data with the fake factor method already introduced for the W+jets process in the previous section. It focuses on the description of dijet events, where both leptons in the signal region selection are faked by misidentified jets. Such kind of events can only enter the analysis when at least one of these anti-ID leptons fires a trigger. Therefore, their acceptance is predominantly realized by the dilepton triggers which feature low transverse momentum thresholds and very loose isolation requirements. The procedure to estimate the QCD contributions follows the same steps as before. A QCD control region is collecting two anti-ID leptons with the same definitions as summarized in tables 18 and 19 in the W+jets section. The event selection cuts are imposed parallel to the signal region. A fake factor is derived to extrapolate from control to signal region. But in this case, the extrapolation factor is measured in a dijet sample rather than the Z+jets sample, because its jet composition is supposed to resemble more the expected hadronic activity.

#### Dijet Fake Factor

The lepton misidentification extrapolation factor measured in the dijet sample $f_\ell^{dijet}$ is simply called dijet fake factor. The sample is selected with two special sets of triggers for the electrons and the muons in a jet enriched environment. In case of the electrons, the anti-ID objects are included with two triggers searching for an EM cluster above 5 and 24 GeV respectively, without further requirements. To avoid a bias due to the thresholds the first is used for the measurement below 20 GeV, the second above. Identified electrons are triggered with the medium identification scheme at a 5 GeV threshold. Both the ID and the anti-ID muon population on the other hand, are collected with single muon triggers that do not impose any requirements on the isolation or the impact parameters. One with a threshold of 6 GeV covers the measurement in the $p_T$ range below 15 GeV, a second one everything above. A list of all the so called supporting triggers in use is given in table 21.

| | |
|---|---|
| Electron trigger | EF_e5_medium1<br>EF_e5_etcut<br>EF_g24_etcut |
| Muon trigger | EF_mu6<br>EF_mu15 |

Table 21: Trigger selection for the dijet sample of the QCD estimate

But such loose trigger criteria lead to very high rates at the LHC. Thus, all of the triggers are heavily prescaled, which has to be accounted for. The presented combination of triggers aims to increase the number of ID and anti-ID objects and thus reduces the

statistical uncertainties. In the end, the collected dijet dataset corresponds to roughly 2 $\text{pb}^{-1}$ for the electrons and about 23 $\text{pb}^{-1}$ for the muons.

To reduce the electroweak contamination of W- or Z-decays, the events for the fake factor measurement require a reconstructed, so called away-side jet with $p_T > 15$ GeV and $\Delta\phi(\ell, jet) > 0.7$. The application of a Z-veto $|m_{\ell\ell} - m_Z| > 13$ GeV and the requirement $m_T^W < 30$ GeV suppress these processes further. Remaining contributions are subtracted from the sample using their Monte Carlo expectation, which includes also the W+jets simulation in this case. The subtraction is validated in the Z-peak region, reverting the veto. Finally, the dijet fake factor can be measured as $f^{dijet} = N_{ID}/N_{anti-ID}$, separately for the electrons and the muons. Figure 88 shows the fake factor binned in $E_T$ and $\eta$ for electrons at the top and muons at the bottom. The electroweak Monte Carlo subtraction is illustrated along with its variations to estimate the incidental systematic uncertainty. But it contributes only a relative uncertainty of 2 – 5% which is comparable to the statistical component.

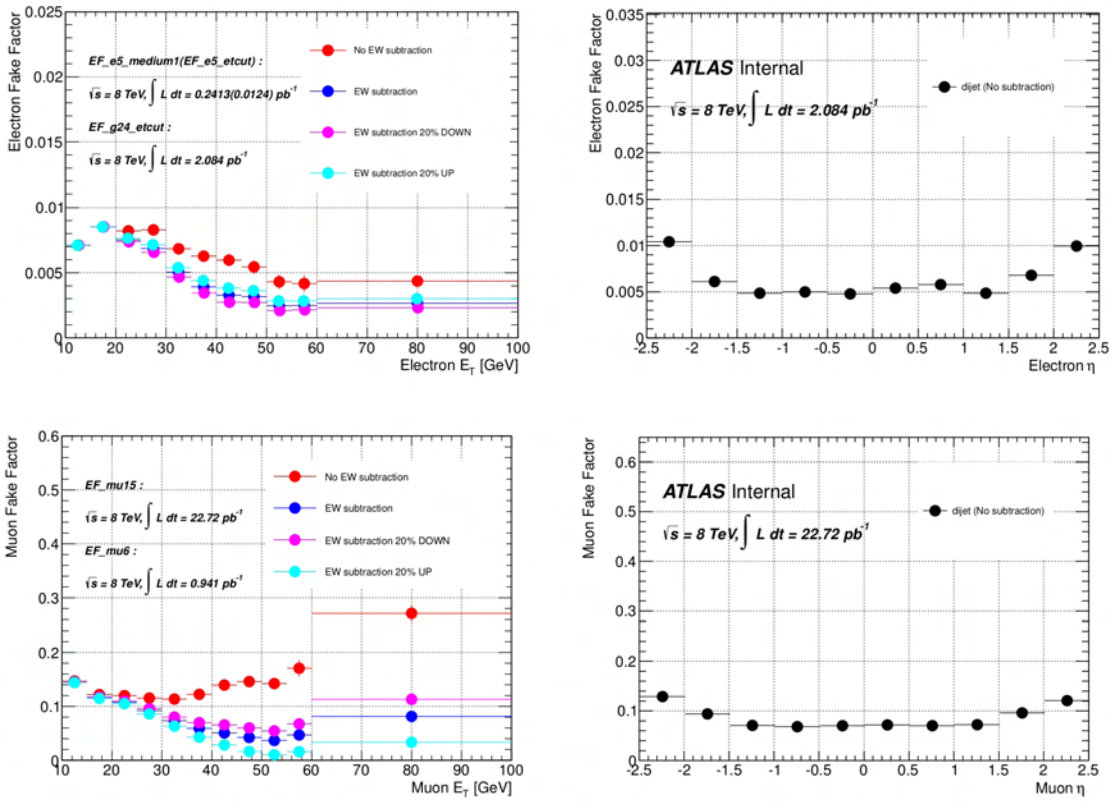



Figure 88: Dijet fake factor as a function of $p_T$ and $\eta$ for electrons on top and muons below [103]. On the left, the red points show the $p_T$ dependence before and the blue points after the subtraction of the EW contamination. A ±20% variation is also illustrated. The $\eta$ dependence on the right is shown before the subtraction.

Splitting the dijet sample into a low and a high pileup set leads to variations in the jet composition and thus the fake factor measurement. The associated uncertainty is below 10%. But the largest systematic uncertainty is assigned to the difference between the dijet sample and the QCD control sample which consists of events with two anti-ID leptons (either without or with one additional jet). To account for this sample dependence, correction factors are derived from simulation:

$$c_\ell = \frac{1}{f_\ell^{dijet}} \times \frac{N_{ID}}{N_{anti-ID}} \tag{6.23}$$

An ansatz is made for the jet composition involving the three categories light flavour (LF), c-jets and b-jets, which can be written as $N = N_{LF}\xi_{LF} + N_c\xi_c + N_b\xi_b$. Each component itself is corrected by taking the ratio of the number of jets $J$ of category $X$ with and without the bias of a particular away-side object $\xi_X = \frac{J_X^{away}/J^{away}}{J_X/J}$. This leads to a large number of correction factors for every combination of jets, ID lepton and anti-ID lepton. But the low statistics in simulation demand simplifications to this procedure which is described in detail in [103]. The resulting correction factors range typically between 1 and 3. Depending on the lepton flavour, the systematic uncertainty on the correction factors ranges from 30% to 50%. It includes a variation of the jet $p_T$ as well as the variation of the b- and c-jet fractions.

Of course, the dijet fake factor describes the same properties as the Z+jets fake factor and thus could also be used for the W+jets estimate. It has in fact been utilized for the previous iteration of the W+jets estimation [107], which did not imply a dedicated QCD estimate yet. But the events collected in the different control samples populate different phase space regions, which leads to large extrapolation uncertainties reducing the sensitivity of the analysis. In order to retain small extrapolation factors, the control sample needs to be chosen as close as possible to the phase space of the process it constrains. That is why the dijet fake factor is used for the estimation of the QCD background, while the W+jets process is described by the fake factors derived from the Z+jets control sample. This refined procedure has been developed within the presented thesis and is a major contribution to the $H \to WW \to \ell\nu\ell\nu$ analysis.

### QCD Control Region and Extrapolation

The procedure for the QCD control region works analogously to the W+jets control region. It selects two anti-ID leptons and uses the same event selection requirements as the signal region apart from that. Residual electroweak background contributions are subtracted using their Monte Carlo expectation. This includes also the W+jets simulation in this case. Figure 89 shows the $p_T$ distribution of the subleading lepton for the 0-jets and the 1-jet selection in the QCD control region before the application of the $m_{\ell\ell}$ requirement. The purity is very high with 74% for 0-jets and 99% for the 1-jet case.

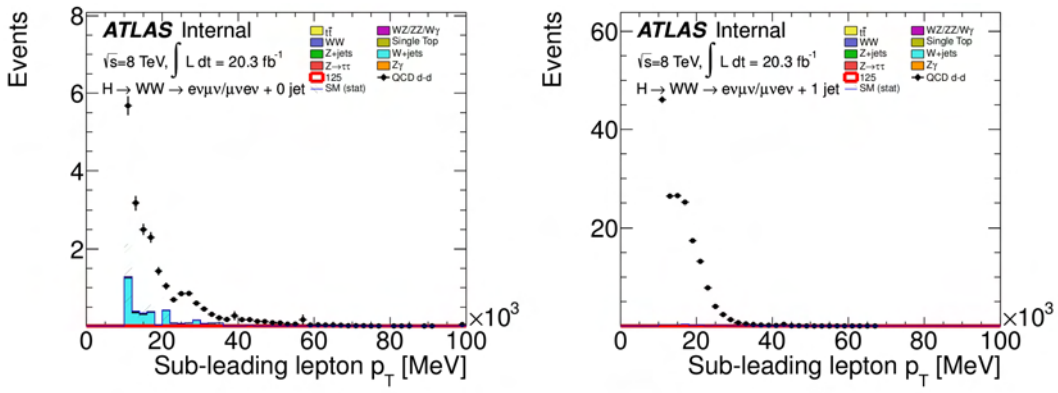



Figure 89: Subleading lepton $p_T^{sub}$ in the QCD control region before $m_{\ell\ell}$ requirement [103]. The 0-jets selection is shown on the left and the 1-jet selection on the right. The EW background contamination is small and has not yet been subtracted.

With the control region and the extrapolation factor at hand, the QCD background contribution to the signal region can now be calculated via:

$$N_{ID+ID}^{QCD} = f'_{dijet} \times f''_{dijet} \times (N_{anti+anti}^{Data} - N_{anti+anti}^{W+jets\ MC} - N_{anti+anti}^{EW\ MC}) \tag{6.24}$$

where the fake factor $f' = c'f$ is corrected for an away-side ID lepton and $f'' = c''f$ refers to the correction in presence of an away-side anti-ID lepton. They both need to be applied because two anti-ID leptons need to be extrapolated to the signal region with respect to each other. By multiplying $f''$, one of the two is extrapolated and afterwards appears as away-side ID lepton with respect to the second, which is accounted for with $f'$. The resulting QCD event yield of $9.20 \pm 1.65$ makes up only 0.3% of the total background and contributes almost exclusively to the different flavour channels.

### QCD Contributions to the W+jets CR

The QCD process also contributes to the W+jets estimate and needs to be subtracted. It enters the W+jets control region through the misidentification of one jet as a lepton. However, this case can be estimated with the already presented measurements. It technically means the extrapolation from the QCD control region with two anti-ID objects to the W+jets control region with one identified lepton and one anti-ID object. This can be calculated by:

$$N_{ID+antiID}^{QCD} = 2f''_{dijet} \times (N_{anti+anti}^{Data} - N_{anti+anti}^{W+jets\ MC} - N_{anti+anti}^{EW\ MC}) \tag{6.25}$$

Here, the away-side object has to be an anti-ID lepton by definition and therefore, $f''$ is applied. Since either one could be picked, an additional factor of 2 has to be taken

into account. This procedure is included in the W+jets estimate described in section 6.5.5 and works generally very well. Nevertheless, some conspicuous discrepancies can be observed that seem to be related to the QCD contribution. Figure 90 shows the calorimeter based missing transverse energy projected onto the axis of the closest object in the W+jets control region at several different stages of the event selection cuts of the $e\mu$ channel. The data-driven QCD estimate is split up into contributions of anti-ID electrons and muons. In the low tail of the distributions, a significant deviation is visible, exactly where the QCD background is expected. This is a strong hint for a systematic underestimation of the QCD contribution derived from data. The same holds for the 1-jet analysis. It has to be noted, that the W+jets Monte Carlo simulation is already normalized to data, which distorts the visibility slightly. The discrepancies are less prominent in the $\mu e$ channel, because of the smaller relative QCD contributions.

A quick study aims to quantify the discrepancy by deriving a normalization factor for the QCD contributions in the low missing transverse energy region below 40 GeV, which is performed separately from the W+jets Monte Carlo simulation normalization. The resulting scale factors calculated for QCD do not distinguish between electron or muon fakes retaining the estimated proportions. They range consistently between 2 and 7 for the different flavour channels of the 0-jets and the 1-jet analysis, depending on the cutstage. The distributions after the QCD normalization are shown in figure 91 for the example of the 0-jets $e\mu$ channel.

Generally, a better agreement of data and the expectation compared to the former case can be observed. But with the proceeding event selection, the QCD background is subsequently reduced, as is the power of this procedure. Anyhow, the impact of this observation on the analysis is very small. Even if the data-driven QCD estimate might be underestimated, it is compensated by the W+jets yield. Because less background is subtracted from data in the W+jets control region, exactly those events enter the signal region hidden in the W+jets estimate. The result is a migration from one process to the other. But at the final steps of the event selection, the fraction of such events is negligibly small. Therefore, more precise investigations have not been initiated.

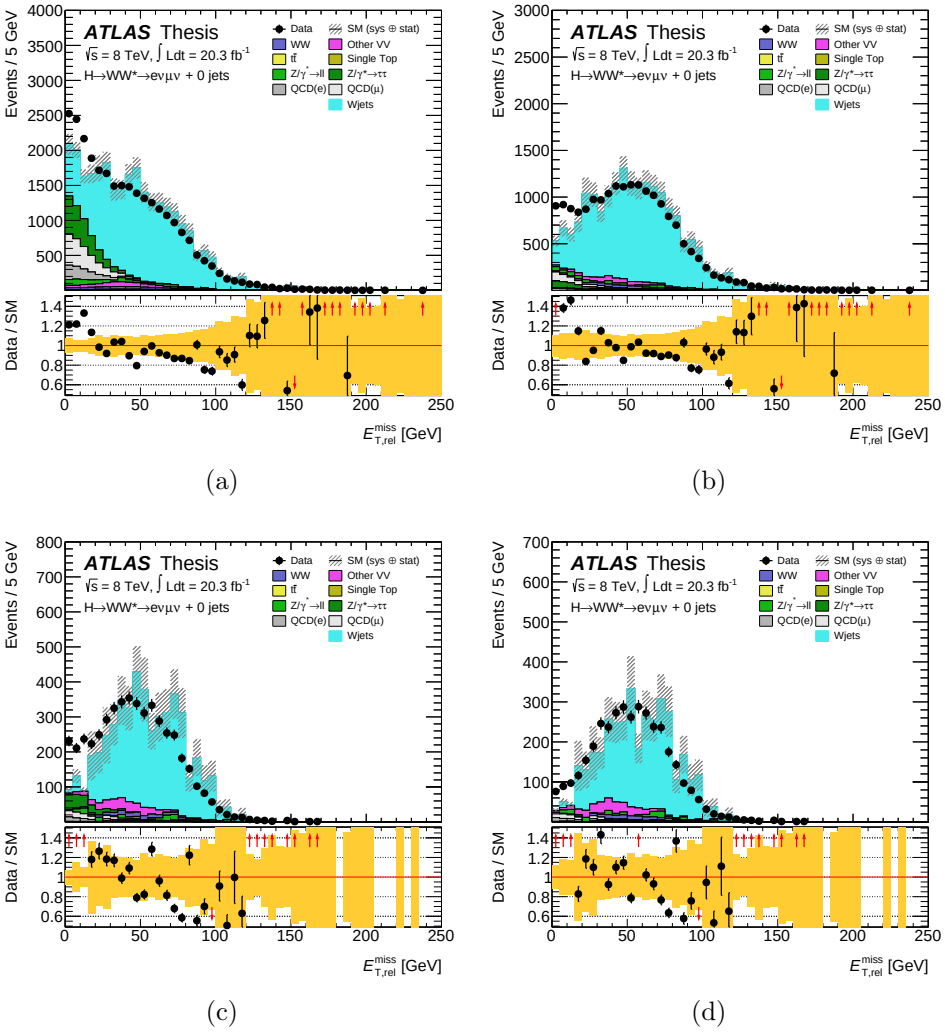



Figure 90: $\not{E}_{T,rel}^{calo}$ distribution in the W+jets control region at (a) jet veto, (b) $p_T^{\ell\ell}$, (c) $m_{\ell\ell}$ and (d) $\Delta\phi_{\ell\ell}$ cut stage for the 0-jets $e\mu$ channel. The W+jets Monte Carlo simulation is normalized to data for comparison. The QCD contributions are data driven. Only statistical errors are included.

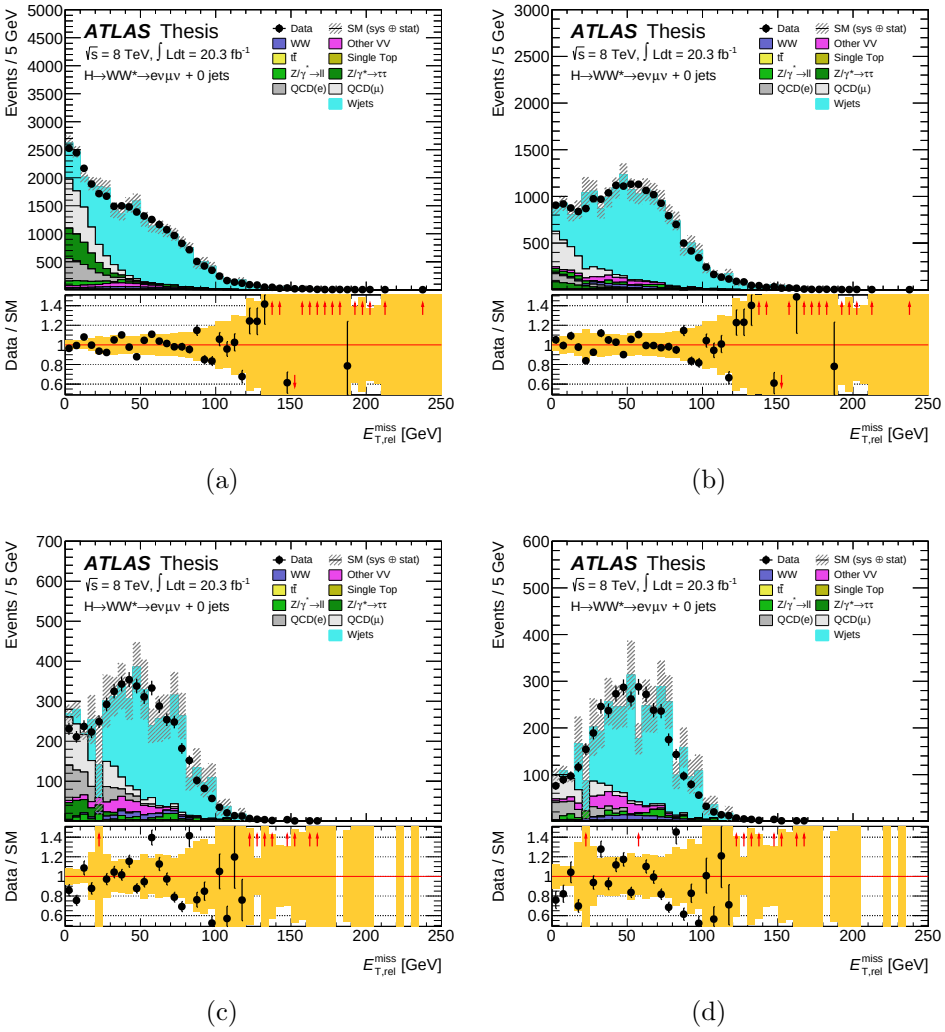



Figure 91: $\not{E}_{T,rel}^{calo}$ distribution in the W+jets control region at (a) jet veto, (b) $p_T^{\ell\ell}$, (c) $m_{\ell\ell}$ and (d) $\Delta\phi_{\ell\ell}$ cut stage for the 0-jets $e\mu$ channel. The QCD contributions are the data driven estimate and are normalized to data below 40 GeV. The W+jets Monte Carlo simulation is normalized to data for comparison. Only statistical errors are included.

## 6.6 Systematic Uncertainties

This section describes the main sources of systematic uncertainties that need to be considered in the $H \to WW \to \ell\nu\ell\nu$ analysis. They serve as input for the statistical analysis of the selected signal region events. A few of them have already been mentioned above. Here, the dominant sources are decomposed to their basics in order to explain their impact on the analysis. Since the different systematics often affect various objects or processes, it is very important to properly treat their correlations. Therefore, the uncertainties are categorized into two groups. Theoretical systematic uncertainties are due to limited accuracy or knowledge of the Standard Model, typically higher order corrections to the gauge theory as they have been mentioned in section 2.3. They are explained in great detail in [108] and the dominant sources are described in the first part of this section. The second category are experimental systematic uncertainties which arise from the limited accuracy of the detector components and their measurements. They range from the luminosity determination to object identification, reconstruction and resolution. An extensive investigation is given in [98]. Those uncertainties that are due to the limited statistics of Monte Carlo Simulation are also counted to the experimental category. Whenever the label statistical uncertainty is used, it only accounts for the limited amount of available data, for example in the control regions described in the previous sections. The full list of all uncertainties considered in the statistical fit is given in table 31 in appendix C and their implementation is explained in chapter 7.1.3. The relative post fit uncertainties on the signal yield and the cumulative background yield are shown in table 23 and a breakdown of the different processes is given by table 34 in appendix C.

### 6.6.1 Theoretical Uncertainties

The theoretical uncertainties of the signal and background processes consider QCD scale variations, differences of the parton shower and underlying event model (PSUE) as well as PDF model variations. For the Higgs boson production, the cross section uncertainties are needed to predict the expected signal region yields. To account for the migration from one signal region to another, proper correlations have to be implemented. But since most of the background processes are normalized in separate data control regions or are derived directly from data, their theoretical cross sections are not required. Only very few cases as the diboson background in the same flavour signal region do actually rely on the theoretical prediction.

#### Theoretical Uncertainties on the Signal Production

The Higgs boson production at the LHC has been introduced in section 2.4.2. The dominant gluon fusion production cross section is calculated to NNLO including initial state

radiation. Resummation of soft QCD radiation has been performed to NNLL and electroweak corrections are applied to NLO. The main contributions to the uncertainty are the QCD scale variations and the parton distribution functions of about 7%. This leads to a 10% systematic uncertainty of the total cross section. But another important parameter are the cross sections for the exclusive jet multiplicities. The POWHEG+PYTHIA8 event generator shows agreement with the best available calculations. A reweighting scheme for the transverse momentum of the Higgs boson $p_T^H$ [109] is applied to reproduce the NNLO+NLL prediction. The jet veto efficiency adds another uncertainty of 11% to the 0-jets analysis. In the 1-jet analysis, the jet binning inherits larger systematic uncertainties of 25%. Relatively small contributions of a few percent are related to the signal acceptance in each category. They include scale and PDF uncertainties as well as differences in the matrix element matching and parton shower models of different Monte Carlo generators.

The total cross section of the vector boson fusion process is obtained with an approximate QCD NNLO computation. Its dominant uncertainty results from the PDF modelling, which is found to be only 2.7%. Both production modes also share an additional systematic uncertainty of 4.2% on the $H \to WW$ branching fraction [110].

#### Theoretical Uncertainties on the WW Production

The WW background process is normalized in the WW control region. Therefore, the theoretical cross section is not used in the predicted yield. But the extrapolation from control region to signal region has systematic uncertainties due to the limited accuracy of the Monte Carlo simulation. By varying the renormalization and factorization scales independently by factors of one half and two, the uncertainties of higher order QCD corrections are evaluated. Electroweak corrections are determined by reweighting the Monte Carlo simulation to NLO calculation [111]. PDF uncertainties cover the largest difference between the nominal CT10 set and other PDF sets. Similar to the gluon fusion Higgs production, generator uncertainties and differences in parton showering and underlying event modelling are evaluated by comparing POWHEG+PYTHIA to combinations including HERWIG and aMC@NLO. All these components range at the percent level, with the generator uncertainties as largest contribution. The total systematic uncertainties on the WW extrapolation factors range from 2% to 4.8% in the 0-jets analysis and from 3.9% to 7.1% in the 1-jet analysis, depending on the transverse momentum of the subleading lepton.

The $gg \to WW$ process inherits larger uncertainties. Renormalization and factorization scale variations to the leading order calculations evaluate the systematic uncertainty to 26% for the 0-jets channel and 33% for the 1-jet channel [112]. The impact on the extrapolation factor can be determined by considering the ratio of $gg \to WW$ to $q\bar{q} \to WW$ in both signal and control region. But increasing the cross section of the gluon induced process shows only minor effects [113] because the contribution to the total WW background is around 6% in the signal region and even less in the WW control region.

Uncertainties that come along with double parton interactions (DPI) are also considered. ATLAS measurements of the $W(\rightarrow \ell\nu)jj$ production [114] are combined with the computed NNLO $W^{\pm}$ cross section to estimate the DPI yield. The uncertainty assigned to this procedure is 60% [115]. But since the multi parton scattering contribution to the WW production is as small as 0.4%, such effects are rather negligible.

Finally, the impact of uncertainties on the shape of the $m_T$ distribution of the WW background in the signal regions is evaluated in addition to the extrapolation uncertainties described above. Higher order matrix element corrections, parton showering and the matching of one to the other are varied in simulation. The resulting uncertainties are very small in the low range of $m_T < 130$ GeV but can increase to 10% in the high tails.

#### Theoretical Uncertainties on the Top Background Production

The theoretical uncertainties on the top background extrapolation arises from the use of Monte Carlo simulation in determining the MC correction factors and b-tagging efficiencies described in section 6.5.2. In order to evaluate them, procedures similar to the case of the WW background are used. The variations of the factorization and renormalization scale, parton showering and underlying event modelling are considered as well as different PDF sets and matrix element matching. The largest contribution comes from the generator comparison and different parton shower models adding a 4% uncertainty to the other components that are at the percent level. The relative fraction of $Wt$ and $t\bar{t}$ production is also taken into account. A $\pm 20\%$ variation of the single top cross section estimates the uncertainty of this interference effect and leads to an uncertainty on the extrapolation factor below 1%. In total, these components add up to a theoretical systematic uncertainty of 8% on the 0-jets top background yield and 5% in case of the 1-jet analysis.

### 6.6.2 Experimental Uncertainties

The experimental systematic uncertainties account for the limited accuracy of the detectors which affect the identification and reconstruction of the objects relevant for this analysis. These are for example the determination of the lepton resolution and identification or the trigger efficiencies. But the dominant source of experimental uncertainties are the jet related quantities such as the jet energy scale, the jet energy resolution or the b-tagging efficiency. These also highly affect the missing transverse momentum measurement. Pileup related uncertainties on the jets or systematics due to the application of the jet vertex fraction on the other hand are found to be negligible.

One major aspect of the high energy scattering experiment is the luminosity determination. In 2012, the uncertainty on the integrated luminosity is derived from a preliminary calibration from beam separation scans by various luminosity sensitive detectors [86].

The estimated 2.8% uncertainty for the $\sqrt{s} = 8$ TeV dataset is slightly larger than it has been for the 7 TeV data collected in 2011 and affects all signal and background yields the same way.

### Lepton Efficiency, Scale and Resolution

The lepton uncertainties before additional cuts on quantities such as isolation or impact parameters are determined by the e/$\gamma$ and muon performance working groups. They evaluate the systematic uncertainties for the reconstruction, identification, isolation and trigger efficiencies from $Z \to ee, \mu\mu$ or $J/\Psi \to ee, \mu\mu$ and $W \to \ell\nu$ decays respectively. Their results are discussed in detail in [62] for the electrons and in [116] for the muons. Typically, the estimated experimental uncertainties are smaller than 1%. Solely the uncertainty of the electron identification efficiency ranges from 0.2% up to 2.7% depending on the $p_T$ and $\eta$ of the object, with the largest contributions at $p_T < 15$ GeV.

The additional selection cuts are another source of experimental uncertainties, particularly those on the isolation variables. The cut efficiencies are computed to correct for mis-modelling in the Monte Carlo simulation and their systematics are determined with the tag-and-probe method on Z-decays. In case of the electrons, the uncertainty on the isolation efficiency reaches up to 1.6% depending on $p_T$ and $\eta$. For the muons, the largest uncertainty contribution is 2.7%.

### Jet Energy Scale and Resolution

The jet energy scale and its uncertainty is determined from in-situ measurements with corrections derived from Monte Carlo simulation for anti-$k_T$ jets from 20 GeV to 1 TeV. The calorimeter response of single hadrons in pp collisions and test-beam data serves as additional input and extends the measurement above 1 TeV. To derive the sytematic uncertainty, it is split up into several independent components. Some of them account for the in-time and out-of-time pileup, Monte Carlo uncertainties and in-situ corrections. Others describe the extrapolation of the jet calibration from the central detectors, subdivided into modelling and statistics. Furthermore, the calibration of light quarks or gluons are investigated as well as the uncertainties on the high $p_T$ behaviour and the b-jet energy scale. Each category is further decomposed to its physical origin. All these independent sources are described in detail in [117]. For the jets used in the presented analysis, the jet energy scale uncertainties range from 1% to 7% depending on the transverse momentum and the $\eta$ direction. The largest contributions are observed at the $p_T$ threshold of 25 GeV.

The jet energy resolution is also determined from in-situ measurements. The resolution itself is found to range from 5% to 20%, while the relative systematic uncertainty varies from 2% to as much as 40%, depending on the jet-$p_T$ and $\eta$. This effect is most prominent at the momentum threshold, again.

### b-tagging Systematics

Another jet related experimental source of uncertainty is the b-jet tagging efficiency as it contributes to the jet multiplicity binning of the analysis and especially to the estimation of the top background. It is measured from a data sample that is dominated by $t\bar{t}$ pairs decaying into two leptons described in [118] based on a likelihood fit procedure. The systematic uncertainties contain a large number of correlated sources like the hadronization and parton shower modelling of the $t\bar{t}$ and other present processes, their normalization, pileup corrections and the largest components resulting from the jet energy scale and resolution. In the end, they are decomposed into six uncorrelated uncertainties with an eigenvector method. The resulting systematics range from below 1% to 7.8% depending on the transverse momentum of the jets, that are b-tagged with the MV1 algorithm at the 85% efficiency working point.

Furthermore, the misidentification rates of light jets and c-jets are estimated in [119] and their systematic uncertainties are evaluated. For light jets that are misidentified as b-jets, the uncertainty ranges from 9% to 19% depending on the jet $p_T$ and $\eta$. Similarly, the systematics on c-jets that are wrongly reconstructed as b-jets vary between 6% and 14% depending on the transverse momentum of the jets.

### Missing Transverse Momentum

The missing transverse momentum is composed from all reconstructed particles in an event. In case of the $H \rightarrow WW \rightarrow \ell\nu\ell\nu$ analysis, these are the leptons and jets. Therefore, the systematic uncertainties of the hard objects like those on the jet energy scale and the lepton momenta are directly propagated to the $\not{E}_T$ measurement [120]. That means, the uncertainties of the missing transverse energy is fully correlated to the input sources.

Nevertheless, the remaining so called soft terms inherit additional systematic uncertainties emerging from the modelling of low energy particle measurements. For the calorimeter based missing transverse energy $\not{E}_T^{calo}$, these soft particles appear as calorimeter clusters not associated with reconstructed objects. The uncertainty is examined from the balance to the sum of the reconstructed high energy objects in $Z \rightarrow \mu\mu$ data and Monte Carlo simulation. Smearing and rescaling the soft term components perpendicular and longitudinal with respect to $\sum \vec{p_T}^{hard}$, leads to the associated systematic uncertainties, which are evaluated in bins of the average number of interactions per bunch crossing. The mean of the soft term varies around 0.3 GeV, which translates to an uncertainty below 20%. The systematic uncertainty on the resolution ranges between 2.0% and 4.5%.

The same method is performed for the track based jet corrected missing transverse momentum $\not{E}_T^{jetCorr}$. In this case, the track based soft term is defined via tracks that are not associated with charged leptons or reconstructed jets. The balance of the track

based soft term to the total transverse momentum of the hard objects is investigated to assess the related systematic uncertainty. In the $Z \to \mu\mu$ data and Monte Carlo simulation, the hard objects are the two identified muons. Four different combinations of MC generators and parton shower implementations are compared to cover modelling effects. The result is an uncertainty on the mean of the track based soft term ranging from 0.3 GeV to 1.4 GeV depending on $\sum \vec{p_T}^{hard}$, which translates to a value below 30%. The resolution of the perpendicular and longitudinal components adds up to an uncertainty between 1.5% and 3.3%.

# 7

# Cross Section Measurement

This chapter discusses the main results of the $H \to WW \to \ell\nu\ell\nu$ analysis leading to the measurement of the Higgs boson production cross section. The first part introduces the basic concepts of the statistical analysis and their application to the case of this measurement, which is based upon the event selection and background estimation presented in the previous chapter. Special emphasis is placed on the treatment of the systematic uncertainties and their impact on the observations. The second part deals with the outcome of the statistical fit procedure and the relevant observables. These are the significance of a Higgs boson signal observed in the $H \to WW \to \ell\nu\ell\nu$ search channels, the measured signal strength and the concluding calculation of the production cross section. A complementary examination of exclusion limits is also performed.

## 7.1 Statistical Treatment

The statistical analysis of the observed data is the key ingredient with regard to the interpretations and conclusions to be drawn. Elaborate descriptions of the use of statistical methods in physics and statistical data analysis can be found in [121] and [122], leading the way from fundamental probability to significance. In this thesis, only the cornerstones are summarized and the actual application to the Higgs boson measurement is introduced step by step.

The goal of a statistical test is to quantify the level of agreement between the measured data and the predicted probabilities given by the so called null hypothesis. The null hypothesis serves as the base model, which is well established and has been measured precisely. In order to explore the terrain of the Higgs boson, the null hypothesis is given by all Standard Model processes without the Higgs contributions. Thus, it describes only the background contributions of the presented analysis and is therefore

called background-only hypothesis (B-only). To establish the Higgs boson signal process, the validity of an alternative hypothesis including signal and background (S+B) is compared to the null hypothesis and the measured data. In the end, the decision taking, whether to accept or reject a hypothesis, is tested with the sample of the observed data.

To formulate the test hypothesis, the probability to produce a specific experimental outcome needs to be identified. The ATLAS collaboration prefers the frequentist definition of probability for their analyses. This approach interprets the limit value $N \to \infty$ of the frequency of a certain outcome $E$ of a repeated experiment as its probability $p(E)$:

$$p(E) = \lim_{N\to\infty} \frac{n}{N} \tag{7.1}$$

where $n$ is the number of events $E$ out of the number of repetitions $N$. This simple and intuitive concept works well for the high energy counting experiments at the LHC with a large number of repeated proton collisions under identical conditions. However, the convergence is a delicate property. Under strict mathematical consideration, the limit does not exist, because there is no deterministic rule connecting the outcome $n$ with $n+1$. So the probability is not defined before the outcome of the experiment is known. Therefore, it is called objective[19] posterior probability. As a result, the probability of an event $E$ is always a joint property of the experiment and the ensemble of all $N$ performances. But when dealing with continuous variables, it is more feasible to describe their distribution via a probability density. The probability $p$ of a variable $x$ to lie in the interval $[x_1, x_2]$ is defined as the integral over the probability density function $f(x)$:

$$p(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} f(x')dx' \tag{7.2}$$

Typically, the hypotheses are more complex and inherit one or more undetermined parameters, as in the case of the Higgs boson analysis. For such a so called composite hypothesis, the functional form of the probability density function is known, while the value of at least one parameter is not. Nevertheless, it is possible to define a confidence interval $[x_1(\mu), x_2(\mu)]$ which covers the unknown parameter $\mu$ with a certain probability according to the definition above:

$$p(x_1(\mu) \leq x \leq x_2(\mu) \mid \mu) = \int_{x_1}^{x_2} f(x'|\mu)dx' = 1 - \alpha \tag{7.3}$$

The conditional probability $p(x|\mu) = 1 - \alpha$ is called confidence level (CL). Commonly quoted is the 95% CL which refers to a value of $\alpha = 0.05$. Note that the integration limit needs to be specified properly, such that the confidence interval is either one-sided or two-sided. For the decision to accept or reject a hypothesis, the true value needs to

[19] Opposite to the subjective probability definition used in Bayesian statistics

be estimated from the measured data sample. The presented analysis uses the so called maximum likelihood (ML) method following in the next section.

### 7.1.1 Maximum Likelihood Method

An estimator for the free parameter of a composite hypothesis is the so called likelihood function. It is constructed as the product of the individual probability density functions of the repeated independent measurements. Thus, it describes the probability to measure the data $x_i$ for a given value of the free parameter $\mu$:

$$L(\mu) = f(x_1|\mu) \cdot f(x_2|\mu) \cdot \ldots \cdot f(x_n|\mu) = \prod_i f(x_i|\mu) \overset{max}{\rightarrow} \prod_i f(x_i|\hat{\mu}) \tag{7.4}$$

However, the likelihood function $L(\mu)$ is not a probability density in $\mu$. The maximum likelihood principle says, that the best estimate $\hat{\mu}$ is the one that maximizes the likelihood function as indicated in equation 7.4. To this effect, $\hat{\mu}$ is not the most likely value of the parameter $\mu$, but the one with the largest probability to obtain the results $\{x_1, \ldots, x_n\}$. This feature is the key characteristic for the interpretation of the statistical analysis of the Higgs boson measurement presented here.

To develop the likelihood formalism towards its actual application, a closer look at the conceptual formulation of the problem is needed. The starting point of the statistical analysis is given as a histogram $\boldsymbol{N} = (N_1, \ldots, N_j)$ with the counted number of events $N_i$ per bin $i$ as for example the distribution of the transverse mass of the Higgs boson $m_T$. The expectation value for each bin can be parametrized as $E(N_i) = \mu \cdot S_i + B_i$ as a function of signal and background events. The undetermined factor $\mu$ is called signal strength, as it scales the signal contribution. In case of the $H \rightarrow WW \rightarrow \ell\nu\ell\nu$ analysis, it expresses the ratio of the measured Higgs production cross section versus its Standard Model prediction $\mu = \sigma^H_{obs}/\sigma^H_{SM}$ and is therefore often referenced as rate measurement. Hence, it is the major parameter of interest for the following considerations. For a large number of independent trials, the probability density is given by a Poisson distribution. This leads to a likelihood function of the form:

$$L(\mu) = \prod f_P(N \mid \mu S + B) = \prod_i \frac{(\mu S_i + B_i)^{N_i}}{N_i!} \times e^{-(\mu S_i + B_i)} \tag{7.5}$$

as product over each independent bin $i$. Maximization of the likelihood function leads to the estimate $\hat{\mu}$, which has the largest probability to produce the outcome histogram $\boldsymbol{N}$. In practice, it is more convenient to use the logarithm of the likelihood function because the execution of the product turns into a simple sum. The monotony of the logarithm does not change the location of the maximum in the parameter space. Also the sign of the function is modified, since many mathematical computer programs available offer minimization algorithms rather than maximization. So the final problem translates to

the so called negative log-likelihood (nLLH) $-2\ln L(\mu) \to min$. The auxiliary factor of 2 is commonly introduced, because this kind of normalization directly relates the problem with the method of least squares ($\chi^2$ minimization) for the case of Gaussian random variables. These aspects are taken on in section 7.1.4.

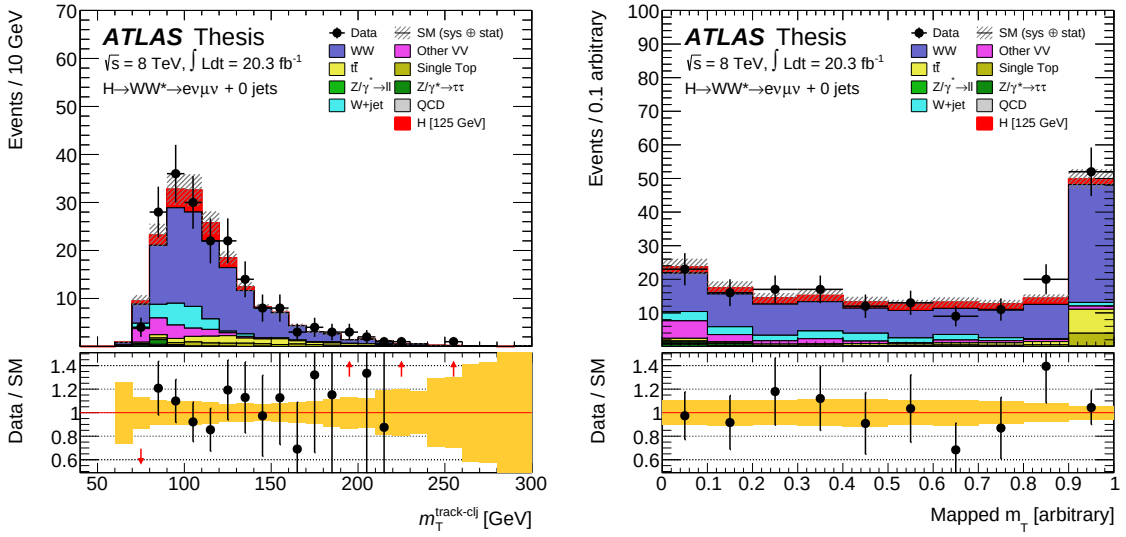
### 7.1.2 Fit Model

The likelihood maximization represents a fit of the distribution of the observable under investigation. It fits the signal yield to the data sample. The parameter of interest is the signal strength $\mu = \sigma^H_{obs}/\sigma^H_{SM}$ binned in the transverse mass distribution of the Higgs boson $m_T$. Several orthogonal signal regions are defined to improve the procedure by selecting different signal and background compositions, each with individual constraints.

While the same flavour signal regions for the 0-jets and 1-jet analysis are defined as combination of the channels $ee$ and $\mu\mu$ after the selection cut on $f_{recoil}$, the different flavour channels $e\mu$ and $\mu e$ are investigated separately and are further subdivided by two additional criteria after the cut on $\Delta\phi_{\ell\ell}$. The first of these splitting criteria is imposed on the dilepton invariant mass distribution leading to a low-$m_{\ell\ell}$ selection between 10 and 30 GeV and a high-$m_{\ell\ell}$ selection from 30 to 55 GeV. The lower and the upper bound have been applied earlier in the selection as explained in chapter 6. The low-$m_{ell}$ selection contains a relatively large number of the W+jets background events and the non-WW diboson processes $W\gamma$, $W\gamma^*$, $WZ$ and $ZZ$, whereas the high mass region is dominated by the WW and top background. The relevant distributions are shown in figures 60 and 69. Similarly, as a second criterion, the distribution of the transverse momentum of the subleading lepton is subdivided into three distinct intervals. The first covers the $p_T^{sub}$ range from 10 to 15 GeV, the second ranges from 15 to 20 GeV and the third selects everything above 20 GeV. All of them are characterized by different fractions of the background processes with varying signal contributions. The W+jets and non-WW diboson backgrounds increase, the lower the transverse momentum gets. This way, the fit of the $m_T$ distribution of each dedicated signal region is executed with a different background composition, which empowers the variety of possible constraints. Finally, the analysis is split into $e\mu/\mu e \times m_{\ell\ell} \times p_T^{sub} = 2 \times 2 \times 3 = 12$ different flavour plus one same flavour signal region per jet multiplicity. So in total, there are 26 orthogonal signal regions incorporated into the fit.

This method is extended by a remapping scheme. In order to maximize the signal significance and stabilize the fit procedure against statistical fluctuations of the background contributions, the binning of each resulting signal region $m_T$ distribution is adjusted, such that every bin shares approximately equal signal yields. In case of the 0-jets analysis, the remapping leads to ten bins of about 5 GeV width in the range of $m_T \in [80, 130]$ GeV including over- and underflow. The 1-jet analysis is rebinned into six bins of approximately 10 GeV width in the main search region. The exact bin boundaries for every individual signal region are listed in tables 29 and 30.

One example of the remapping scheme is illustrated in figure 92. It shows the transverse mass distribution of the Higgs boson for the case of the $e\mu$ channel of the 0-jets analysis with $m_{\ell\ell} > 30$ GeV and $15 < p_T^{sub} < 20$ GeV. The original appearance is given on the left, the rebinned version on the right. Every bin shares approximately equal signal contributions while the background composition obviously differs. This increases the power of the fit procedure. In the lower bins, W+jets and non-WW diboson backgrounds can be constrained very effectively, while the higher bins are predominantly populated by WW production with increasing fractions of the top backgrounds. The latter is a very useful feature which tackles an important problem. Because the WW control region has a large contamination of top related background events, the determination of the processes is deeply correlated.




(a) (b)

Figure 92: Transverse mass distribution of the 0-jets signal region selection in the $e\mu$ channel with $m_{\ell\ell} > 30$ GeV and $15 < p_T^{sub} < 20$ GeV. The original binning (a) is shown on the left and the remapped distribution (b) on the right. While most bins focus on the region $m_T \in [80, 130]$ GeV, the very first and last bin sum up the remaining low and high mass range.

The details of the remapping scheme applied to the signal region $m_T$ distributions are collected in appendix B. The original histograms are illustrated in figures 103 to 107. Their remapped versions can be found from figure 108 to 112. A breakdown of the composition of the signal and all considered background processes for every signal region is given in table 28.

But the individual signal regions are not the only fit regions considered in the fit model. As explained in chapter 6.5, the majority of background processes are normalized in dedicated control regions. The measured normalization factors inherit systematic uncertainties and are often correlated among the different processes. These scale factors can

be interpreted and handled the same way as the signal strength parameter. Therefore, the fit model of the $H \to WW \to \ell\nu\ell\nu$ analysis is extended to also fit the normalization of individual background processes. This kind of in-situ normalization is called profiling (see section 7.1.4). The choice which background process is profiled, depends on its general impact on the analysis and the level of contamination in the dedicated control region. The profiled control regions included in the fit are the WW control region, the same sign control region for the $W\gamma$, $W\gamma^*$, $WZ$ and $ZZ$ processes and the $Z \to \tau\tau$ control region. Furthermore, the control regions used in the measurement of the DY and non-DY efficiencies for the determination of the $Z \to \ell\ell$ background are taken into account. In case of the 1-jet analysis, also the top control region is profiled. So all of the related normalization factors and efficiencies are treated as floating parameters in the fit. But none of them contains shape information, as they are all incorporated as a single bin. The technical implementation into the likelihood function is discussed in the next section.

### 7.1.3 Systematic Uncertainties and Nuisance Parameters

Another set of auxiliary parameters is the full list of systematic uncertainties considered. Although their size has been estimated thoroughly, they are typically quoted conservatively and thus are sometimes overestimated. But they can also be determined and constrained by the fit model, if they are properly incorporated into the likelihood function. The modified version of equation 7.5 represents the technical implementation of these so called nuisance parameters $\boldsymbol{\theta} = \{\theta_1, \theta_2, \ldots\}$ and can be written as:

$$L(\mu, \boldsymbol{\theta}) = \prod_{i,b}^{\text{SR, bin}} P\Big(N_{ib} \;\Big|\; \mu \cdot S_{ib} \cdot \prod_{r}^{\substack{\text{signal}\\ \text{systs}}} \nu_{br}(\theta_r) + \sum_{k}^{\substack{\text{bkgd}\\ \text{procs}}} \beta_k \cdot B_{kib} \cdot \prod_{s}^{\substack{\text{bkgd}\\ \text{systs}}} \nu_{bs}(\theta_s)\Big) \tag{7.6}$$

$$\times \prod_{l}^{\substack{\text{profiled}\\ \text{CR}}} P\Big(N_l \;\Big|\; \sum_{k}^{\substack{\text{bkgd}\\ \text{procs}}} \beta_k \cdot B_{kl}\Big) \tag{7.7}$$

$$\times \prod_{t}^{\substack{\text{non-profiled}\\ \text{systs}}} G\left(\vartheta_t \mid \theta_t\right) \tag{7.8}$$

$$\times \prod_{k}^{\substack{\text{bkgd}\\ \text{procs}}} P\left(\xi_{k,b} \mid \zeta_{k,b} \cdot \theta_k\right) \tag{7.9}$$

where the functions $P()$ and $G()$ represent Poissonian and Gaussian probability density functions respectively. In order to disentangle this rather complex construct, the four individual terms are discussed in detail one by one.

The first term (7.6) of the likelihood function is the elaborate version of equation 7.5 and describes the $m_T$ fit. The Poisson probability density has the form $P(N|\alpha) = \alpha^N e^{-\alpha}/N!$ evaluated for every single bin $b$ of the remapped $m_T$ distribution in each individual signal region $i$. Parameter of interest is the signal strength $\mu$, which is applied to the signal yield $S_{ib}$ per bin of the particular signal region selection. This yield is further scaled by the response function $\nu(\theta)$, parametrizing the impact of the systematic uncertainties $\theta$ which varies for every analysis bin. The same holds for the yield $B_k$ of each background process $k$ and its set of corresponding systematic uncertainties, with the background normalization factors $\beta$ determined from the second term. The exact form of the response functions depends on the specific type of uncertainty and is described in the discussion of the third and fourth term. However, statistical uncertainties associated with the random error of a predicted value are explicitly considered by term one.

The second term (7.7) implies the profiled control regions $l$ which are the WW, the same sign, the Drell-Yan control regions and the 1-jet top control regions. They are also realized by Poisson distributions expressing the probability to observe $N$ events resulting from the sum of the background yields $B_k$. Here, the background normalization factors $\beta_k$ are determined one by one. The maximization of the likelihood then picks the best combination of estimates and serves as in-situ normalization, accounting for their correlations by selecting the Poisson terms with the largest probability to produce the observed number of events. Since the normalization factors are applied to the signal regions, they are the same as those appearing in term one. Although they represent the background strength similar to the parameter of interest $\mu$, conceptually they belong to the set of nuisance parameters $\boldsymbol{\theta}$ and are not referred explicitly as arguments of the likelihood function.

The third term (7.8) introduces Gaussian probability density functions that are used to model systematic uncertainties. So their general form is $G(\vartheta|\theta) = e^{-(\vartheta-\theta)^2/2}/\sqrt{2\pi}$. The central value of the measurement is represented by $\vartheta$ with its associated uncertainty given as nuisance parameter $\theta$. Although this assumption is reasonable in most of the cases, it is a renunciation of the frequentist view on probability. Strictly speaking, this implementation is referred to as semi-frequentist method, because such a choice is based on a certain degree of belief. Nevertheless, the constraints of the Gaussian modeled systematic uncertainties affect the predicted yields through the response functions stated in term one. There are two categories of systematics that are treated differently. Normalization uncertainties that do not vary between analysis bins are called flat systematics. They are parametrized by an exponential response function $\nu^{\text{flat}}(\theta) = (1-\epsilon)^\theta$, where the uncertainty $\theta$ has the value $\epsilon$. That means, a relative uncertainty with a certain value quoted at one standard deviation corresponds to $\theta = \pm 1$. The response function $\nu^{\text{flat}}(\theta)$ is distributed log-normally. To stabilize the fit procedure, normalization uncertainties that are less than 0.1% are excluded. The second category of uncertainties is called shape systematics. They affect every analysis bin differently and thus alter the shape of a given distribution. Any global normalization change is separated as flat component and treated as described above. The pure shape component is parametrized as linear response function for each bin $\nu_b^{\text{shape}}(\theta) = 1 + \epsilon_b \cdot \theta$, normally distributed around unity

with the width $\epsilon_b$ determined by measuring $\nu^{\text{shape}}(\theta = \pm 1)$. It is truncated to be positive in order to avoid unphysical responses. Shape systematics with a size below 1% in every bin are excluded from the fit. In general, the $H \to WW \to \ell\nu\ell\nu$ analysis is more sensitive to normalization uncertainties, because they have a larger effect on the $m_T$ shape of the total background, as the different processes are not evenly distributed. Shape variations of individual background processes commonly tend to be compensated by one another.

The fourth term (7.9) represents the systematic uncertainties emerging from the finite number of generated Monte Carlo events for each background process. This so called sample error is described in [123]. It is realized by a Poisson term with the central value $\xi$ of the background estimate $\zeta \cdot \theta$ and constrains the background yield $B$ with the statistical uncertainty $\delta$ via $\zeta = (B/\delta)^2$. Since this systematic is specific to every bin content, it impacts the predicted yields through the linear response function $\nu_b^{\text{stat}}(\theta) = \theta$, similar to the case of shape systematics. The full list of all uncertainties considered in the statistical fit is given in table 31 in appendix C. The background normalization factors and cut efficiencies are listed in table 32.

### 7.1.4 Test Statistic

Now that the likelihood function is properly set up, the final test statistic can be defined with respect to the anticipated interpretations. A very useful paper on likelihood based test statistics, asymptotic formulae and the use of so called Asimov datasets is presented in [124]. The main items applied to the Higgs analysis are summarized below.

The focus lies on the signal strength $\mu$ as parameter of interest, but estimated in presence of a set of nuisance parameters $\boldsymbol{\theta}$. The so called profile likelihood ratio can be used to test a hypothesized value of $\mu$:

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})} \tag{7.10}$$

In the numerator, the nuisance parameters $\hat{\hat{\boldsymbol{\theta}}}$ denote the values maximizing the likelihood function for a specific signal strength. Therefore, it represents a conditional maximum likelihood estimator of $\boldsymbol{\theta}$, which itself depends on $\mu$. The estimation of $\boldsymbol{\theta}(\mu)$ is called profiling, because it determines and fixes the values of the nuisance parameters for the scanned values of $\mu$. On the other hand, the denominator constitutes the unconditionally maximized likelihood function with the maximum likelihood estimators $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$. The test statistic for the discovery of a positive signal aims at the rejection of the background-only hypothesis given by $\mu = 0$. It is labelled $q_0$ and represents a negative log-likelihood ratio (nLLR):

$$q_0 = \begin{cases} -2\ln\lambda(0) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \tag{7.11}$$

As stated above, this is a mathematically convenient definition. The logarithm of the ratio turns into a simple subtraction of the likelihood functions. Although the case of $\hat{\mu} < 0$ may indicate a discrepancy of the experiment and the null hypothesis, the test statistic is set to zero. In case of the $H \to WW \to \ell\nu\ell\nu$ process, the Higgs boson can only supply additional data events, which cannot be deduced from a negative estimator. The level of disagreement between data and the background-only hypothesis can rather be quantified by computing the p-value, in this case called local $p_0$:

$$p_0 = \int_{q_0,\mathrm{obs}}^{\infty} f(q_0|0)dq_0 \tag{7.12}$$

where $f(q_0|0)$ represents the probability density function[20] of the test statistic $q_0$ with the assumption $\mu = 0$. With the observed value as integration limit, the local $p_0$ describes the probability to find data of equal or even greater incompatibility with the predictions of the null hypothesis. This can also be expressed in terms of a Gaussian significance $Z$. That means, a Gaussian distributed variable with an (upper) tail probability of $p_0$ is found to lie $Z = \Phi^{-1}(1 - p_0)$ standard deviations above its mean value as illustrated in figure 93, where $\Phi^{-1}$ denotes the quantile of the standard Gaussian. To claim a discovery, the $5\sigma$ level ($Z = 5$) has been naturalized in the particle physics community, which corresponds to the p-value of $p_0 = 2.87 \times 10^{-7}$. In other words, a fluctuation of the background-only hypothesis resulting in the observed dataset is expected to happen less than once every $\sim 3.5$ million times. For the exclusion of a signal hypothesis with a 95% confidence level on the other hand, the threshold value $p = 0.05$ corresponds to $Z = 1.64$.

Once the background-only hypothesis is rejected to a significant extent, the validity of the signal hypothesis can be investigated. To further quantify the agreement of the data with the signal hypothesis, an alternative test statistic $\tilde{q}_\mu$ for upper limits is defined. Although the signal strength of the Higgs boson $\mu \geq 0$ is bound to be positive, its effective estimator $\hat{\mu}$ is not. Therefore, the following cases need to be accounted:

$$\tilde{q}_\mu = \begin{cases} -2\ln\tilde{\lambda}(\mu) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} = \begin{cases} -2\ln\frac{L(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(0,\hat{\hat{\boldsymbol{\theta}}}(0))} & \hat{\mu} < 0 \\ -2\ln\frac{L(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(\hat{\mu},\hat{\boldsymbol{\theta}})} & 0 \leq \hat{\mu} < \mu \\ 0 & \hat{\mu} > \mu \end{cases} \tag{7.13}$$

[20] To actually calculate the p-values for this thesis, the asymptotic approximations for the distributions $f(q|\mu)$ developed in [124] are used.

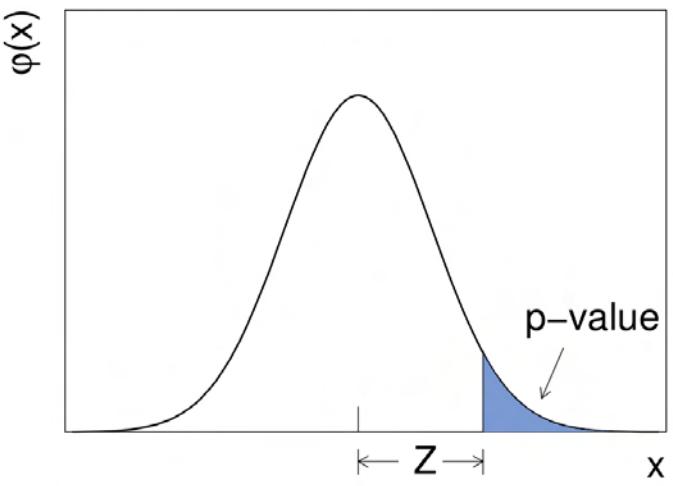



Figure 93: Relation between the significance $Z$ and the p-value [124]. The Gaussian distributed variable $\varphi(x) = e^{-x^2/2}/\sqrt{2\pi}$ found $Z$ standard deviations above its mean has an upper tail probability equal to $p$.

representing the same profile likelihood ratio as before. For negative values of $\hat{\mu}$, the denominator is given by the conditional maximum likelihood estimator of the null hypothesis. On the other hand, if it exceeds the value given by the signal hypothesis $\hat{\mu} > \mu$, the test statistic is set to zero to leave it out of the rejection region. So an observation of even more data events than expected is not regarded as incompatibility but rather as adjustment of the signal hypothesis. Again, the p-values for both hypotheses can be calculated via their sampling distributions:

$$p_{S+B} = \int_{\tilde{q}_{\mu,\mathrm{obs}}}^{\infty} f(\tilde{q}_\mu|\mu, \hat{\theta}_\mu) d\tilde{q}_\mu \tag{7.14}$$

$$p_B = \int_{-\infty}^{\tilde{q}_{\mu,\mathrm{obs}}} f(\tilde{q}_\mu|0, \hat{\theta}_0) d\tilde{q}_\mu \tag{7.15}$$

Figure 94 illustrates the procedure for the example of a likelihood ratio $q = -2 \ln L_{s+b}/L_b$. The histograms represent the sampling distributions of the probability density functions of the test statistic taken from Monte Carlo simulation. The solid curves are the asymptotic formulae for the null hypothesis $f(q|b)$ with $\mu = 0$ and the signal+background hypothesis $f(q|s+b)$ with $\mu = 1$. The p-values can be obtained from the integrals defined by the integration limit $q_{obs}$, which represents the value of the test statistic for the observed data.

In this particular example, the limit is chosen to match the median expectation of the background-only hypothesis. This is an important case to obtain the expected sensitivity of an experiment for an assumed signal strength. The test statistic serves as ordering rule

and separates the two hypotheses. In the example of the negative log-likelihood ratio, signal-like estimates tend to negative values, while background-like estimates are shifted to the positive range as indicated in figure 94. The more the distributions overlap, the harder it is to distinguish between the two hypotheses and the lower is the sensitivity. Systematic uncertainties that enter as nuisance parameters broaden the likelihood ratio, reflecting the loss of information on $\mu$. The coverage of the two distributions gets larger and the sensitivity is degraded. In reverse, the median expected significance provides useful information about the interplay of tested models and the run conditions of the experiment.

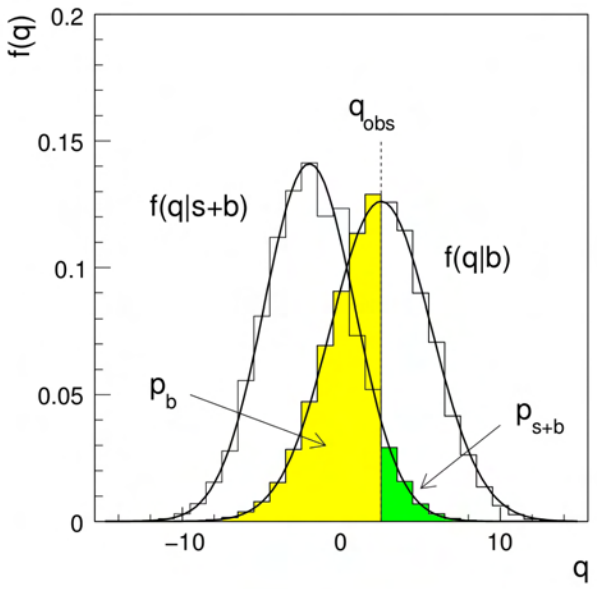



Figure 94: Illustration of the distribution of a test statistic $q = -2 \ln L_{s+b}/L_b$ [124]. The histograms represent the Monte Carlo simulation, while the solid curves are the predictions of the asymptotic formulae $f(q|s+b)$ for $\mu = 1$ and $f(q|b)$ for $\mu = 0$ respectively. The p-values corresponding to the observed value $q_{obs}$ are indicated by the coloured integrals.

Moreover, the p-values described above can be used to determine confidence intervals [125] for the assumed signal hypothesis. They are typically stated at a certain confidence level and test equation 7.3 for the corresponding threshold. But this is predominantly of interest in the absence of a signal. In this case, exclusion limits can be derived, given by the upper limits on the signal strength $\mu$ that can be excluded at a certain confidence level. The so called $CL_S$ technique [126] has been proven to be appropriate for the search of new physics phenomena:

$$CL_S = \frac{CL_{S+B}}{CL_B} = \frac{p_{S+B}}{1 - p_B} \tag{7.16}$$

Both hypotheses need to be tested to ensure the required sensitivity and to prohibit spurious exclusions. The signal hypothesis is rejected, when meeting the condition $CL_S(\mu_{\text{limit}}) < \alpha$. This is obviously fulfilled if the p-value $p_{S+B}$ is low, meaning that signal+background hypothesis has a low probability to result in the observed data. But if the distributions have a large overlap and cannot be distinguished properly because the sensitivity is simply too low, the denominator gets small and prevents the unreasonable exclusion.

## 7.2 Results and Conclusions

This section describes the results of the statistical fit and the conclusions that can be drawn from the determined fit values. The procedure outlined in section 7.1 is performed with the ROOT Data Analysis Framework [127] and the additional ROOFIT [128] package. The probability density functions are incorporated with the tools provided by the ROOSTATS [129] framework. To obtain the estimators of the profile likelihood ratio $\lambda(\mu)$, the analysis framework is interfaced with the minimization package MINUIT [130]. Starting with the closer view on the nuisance parameters representing the systematic uncertainties and auxiliary measurements, the post fit signal and background yields are discussed for the masspoint $m_H = 125$ GeV. The following sections focus on the signal significance $q_0$, the signal strength $\hat{\mu}$ and also exclusion limits for the investigated mass range. With these information at hand, the gluon fusion Higgs production cross section can be calculated and compared to its theoretical expectations in the final section of this chapter.

To begin with, the results for the obtained estimators $\hat{\mu}$, $\hat{\boldsymbol{\theta}}$ and $\hat{\hat{\boldsymbol{\theta}}}$ need to be examined in detail. They are extracted simultaneously, adjusting the pre-fit value of every nuisance parameter $\theta = 0$ and its uncertainty $\Delta\theta = \pm 1$, initially defined by the unit Gaussian construction. The effect of the statistical fit on a single nuisance parameter is often expressed by its so called pull:

$$\delta\hat{\theta} = \frac{\hat{\theta} - \theta}{\Delta\theta} \tag{7.17}$$

comparing the pre- and post fit value, normalized to the uncertainty. A complete list of the pull values for all nuisance parameters is given in table 31 of appendix C for the different fit scenarios. They are also illustrated on the lower scale in figures 95 and 118. The vast majority of the fit values ranges well below 0.01. Only a few cases show larger deviations up to a pull of 0.5. None of the fitted values exceeds the variation range imposed by the expected uncertainties. So there are no indications of either unreasonable initial estimates or suspicious constraints on the nuisance parameters. The fitted background strength parameters $\beta$ are listed separately in table 32 of appendix C. Their initial values $\beta = 1$ refer to the Standard Model expectation, while the post fit values resemble the normalization factors determined from the control regions evaluated simultaneously instead of a sequential determination as described in section 6.5.

The most prominent fit result is the increase of the WW background, which is 21% (11%) higher than predicted for the 0-jets (1-jet) analysis, very close to the observations in section 6.5.1. This is of course a delicate issue. The correct normalization of the dominant and irreducible background process is crucial, because it highly affects the sensitivity of the analysis. Indeed, the ATLAS measurement of the Standard Model WW process [76] described in section 5.2.1 observes a cross section that is about 20% larger than the expected value. Hence, the obtained fit result is in good agreement.

Apart from that, the effects on the other processes are rather subtle. Most of them tend to post fit values above one that increase the expected yields. The more signal strength is injected into the fit scenario, the smaller the variations are.

Very important to know is the impact of systematic uncertainties on the signal strength. A measure to quantify the impact of a single nuisance parameter is:

$$\Delta\hat{\mu} = \hat{\mu}(\hat{\theta} \pm \Delta\hat{\theta}) - \hat{\mu}(\hat{\theta}) \tag{7.18}$$

where $\hat{\mu}(\hat{\theta}\pm\Delta\hat{\theta})$ represents the result of a fit with one parameter varied by the associated uncertainty around its post-fit value, while all others are floating. Of course, this can also be formulated for the pre-fit scenario by substituting $\hat{\theta} \to \theta = 0$ and $\Delta\hat{\theta} \to \Delta\theta = \pm 1$. The post fit uncertainties of the nuisance parameters $\Delta\hat{\theta}$ still need to be determined. This is achieved by a scan of $\chi^2 = -2\ln(L(\hat{\theta} \pm \Delta\hat{\theta})/L(\hat{\theta})) = 1$, where the maximum likelihood ratio follows a $\chi^2$ distribution, because of the construction via unit Gaussian distributed nuisance parameters. The desired reduction of $\Delta\hat{\mu} < 1$ arises from the constraints on the data fit and thus leads to actually reduced systematic uncertainties of the analysis and decreases the impact on the observed signal strength.

Figure 95 illustrates the fit results for the nuisance parameters and their impact on the observed signal strength for a hypothesized Higgs mass of $m_H = 125$ GeV. The pulls of the nuisances parameters $\delta\hat{\theta}$ are visualized as black points plotted versus the lower scale, where the red error bars mark the $\pm 1$ standard deviation range and the black error bars show the post fit uncertainties $\Delta\hat{\theta}$ illustrating the fit constraints. The pre-fit values are indicated by the dotted black lines. The listed nuisance parameters are ranked from top to bottom by their post fit impact $\Delta\hat{\mu}$ on the observed signal strength illustrated by the dashed rectangles, compared to the pre-fit impact shown as filled rectangles. For better visibility, only the thirty highest-ranked parameters are shown. The same illustration for the full set of 134 nuisance parameters is given in figure 118 in appendix C.

The ranking confirms the expectation, that the dominant uncertainties on the signal process and the major background processes possess the largest impact on the determination of the observed signal strength. The first two entries in figure 95 describe the uncertainties on the gluon fusion Higss boson production cross section due to the variation of the renormalization and factorization scale and due to the variation of the parton distribution functions. Unfortunately, the statistical fit does not seem sensitive enough to constrain these systematics from the experimental data as can be observed from the very small pull values and the post fit uncertainties of these nuisance parameters. Therefore, they retain the same impact $\Delta\hat{\mu}$ in the pre- and post fit scenario. The third entry belongs to the WW generator modelling and represents the component introduced by the $m_T$ shape uncertainty, which particularly affects the high tails of the distribution. However, this regime of the signal regions is characterized by rather little signal contributions and a high purity of the Standard Model WW background process. Therefore, the fit has valuable handles to constrain this systematic uncertainty leading to a 30% improvement given by $\Delta\hat{\theta} = 0.7$. This is the largest effect for all 134 considered

nuisance parameters and is also reflected in the reduced impact on the observed signal strength. While the $m_T$ shape matching uncertainty of the WW background surmounts all other systematics with $\Delta\hat{\mu} = ^{+0.12}_{-0.11}$ in the pre-fit scenario, it loses the top position and drops down to a post-fit value of $\Delta\hat{\mu} = ^{+0.08}_{-0.07}$. Another set of highly ranked nuisance parameters belongs to the sample dependence in the fake factor determination used to derive the W+jets background yield, namely the opposite charge Monte Carlo correction factors $c^{OC}_{e,\mu}$ for electrons and muons. They are fitted in the low $p_T^{sub}$ signal regions, since the W+jets process does not belong to the set of profiled backgrounds. For both, the result is a considerable pull with constrained uncertainties, leading to a reduced impact on the analysis. The systematic uncertainty on the integrated luminosity also belongs to the top flight. But it only applies to the signal process and some minor background processes, since the dominant background sources are all normalized to data in dedicated control regions. Thus, the capability of the fit to constrain this parameter is rather small.

The resulting relative post fit uncertainties on signal and background are merged in table 23, for the two jet multiplicities respectively. The top box summarizes the signal systematics according to their physical origin. They are dominated by the uncertainties on the gluon fusion Higgs boson production cross section of about 10% and the similar sized uncertainties on the jet veto efficiencies. Most of the corresponding nuisance parameters have been found to have a large impact on the measurement. The acceptance model of the gluon fusion process and the branching ratio to WW follow in the 4 – 5% range, whereas the vector boson fusion related systematics are negligible due to the very low VBF rate. Uncertainties related to the jet energy scale and resolution are also rather large, compared to those on other objects. In the lower box, the systematic uncertainties on the total background are summarized. The largest values are those of the WW theoretical model with about 1.5% and the W+jets estimate in the range of $\sim 1\%$. They have also been identified to decisively affect the fit procedure. All other sources of systematics on the total background yield are below one percent. For example, the post fit uncertainty on the integrated luminosity is below 0.1% for the background, because most processes are normalized to data. However, the signal expectation is originally scaled with the luminosity applying the full 2.8% systematic, and the remaining fit to data measures the signal strength parameter. A breakdown of the various categories of the post fit systematic uncertainties per process is given in table 34 in appendix C. The absolute values can be drawn from table 22 or 33, which show the post fit signal and background yields for the different channels and all separate signal regions.

Another way to visualize the results of the fit to the 2012 ATLAS data is given in figure 96. It shows the post fit transverse mass distribution for the combination of all 26 signal regions of the different analysis channels, namely the $e\mu$, the $\mu e$ and the same flavour ($ee + \mu\mu$) selections for the two jet multiplicities. Those are shown in figures 103 to 107 for the pre-fit scenario. In the summation illustrated in the upper plot of figure 96, the $m_H = 125$ GeV signal contribution has been scaled by the observed signal strength $\hat{\mu}$ and the post fit normalization factors $\beta_k$ have been applied to the background processes where available. So in total, the post fit distribution exhibits 428 signal events and 4909

background events that describe the 5346 observed data events (see table 22) very well. The gluon fusion process makes up 94% of the signal contribution, the remaining events emerge from vector boson fusion and the production in association with a vector boson. The shape of the excess in data is filled smoothly by the Monte Carlo signal expectation and supports the convincibility of the Higgs boson signal. To further illustrate this peak in the transverse mass distribution, the lower plot of figure 96 shows the background subtracted version of the above. This means, the residuals of the data with respect to the total background contribution are compared to the signal events. The resulting distribution is rather flat across the mass range with a broad clear peak in the region between 80 GeV and 150 GeV and a maximum around 110 GeV, which is in good agreement with the signal expectation of the $m_H = 125$ GeV Standard Model Higgs boson. The complete set of post-fit results of the remapped transverse mass distribution is shown from figure 113 to 117 in appendix C.

The post fit event yields are listed in table 22 for the different channels and their combination, which resembles the contributions shown in figure 96. The quoted uncertainties include all theoretical, experimental and statistical sources and account for the fit constraints and correlations between channels and background processes. A breakdown with the numbers for all signal regions is presented in table 33 in appendix C. They can be compared to the pre-fit selections from table 28 in appendix B.

| Channel | Summary | | | Background Composition | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Data | Bkgd | Signal | $WW$ | Top | diboson | Z/DY | W+jets | QCD |
| 0-jets | 3750 | 3431 ± 126 | 317 ± 50 | 2254 ± 95 | 302 ± 22 | 425 ± 42 | 70 ± 16 | 365 ± 62 | 14 ± 2 |
| 0j: DF | 2642 | 2399 ± 88 | 237 ± 37 | 1520 ± 64 | 200 ± 15 | 353 ± 35 | 31 ± 4 | 283 ± 48 | 12 ± 2 |
| 0j: SF | 1108 | 1032 ± 37 | 80 ± 13 | 735 ± 31 | 101 ± 7 | 72 ± 7 | 39 ± 12 | 82 ± 14 | 2.0 ± 0.3 |
| 1-jet | 1596 | 1478 ± 63 | 111 ± 21 | 630 ± 48 | 535 ± 27 | 144 ± 20 | 53 ± 14 | 110 ± 20 | 7.2 ± 1 |
| 1j: DF | 1129 | 1048 ± 44 | 88 ± 17 | 446 ± 34 | 367 ± 18 | 113 ± 16 | 25 ± 3 | 90 ± 16 | 7.0 ± 1 |
| 1j: SF | 467 | 431 ± 20 | 23 ± 5 | 184 ± 14 | 169 ± 8 | 30 ± 4 | 28 ± 11 | 20 ± 4 | 0.2 ± 0.1 |
| Comb. | 5346 | 4909 ± 189 | 428 ± 71 | 2884 ± 143 | 837 ± 49 | 569 ± 62 | 123 ± 29 | 475 ± 82 | 21 ± 4 |

Table 22: Post fit yields of the different analysis channels and their combination estimated from the 2012 ATLAS dataset. The expected signal yield for a Standard Model Higgs boson of the mass $m_H = 125$ GeV is scaled by the observed signal strength while the background processes use the post fit normalization factors. The quoted uncertainties include theoretical, experimental and statistical sources and account for the fit constraints and correlations between channels and background processes.

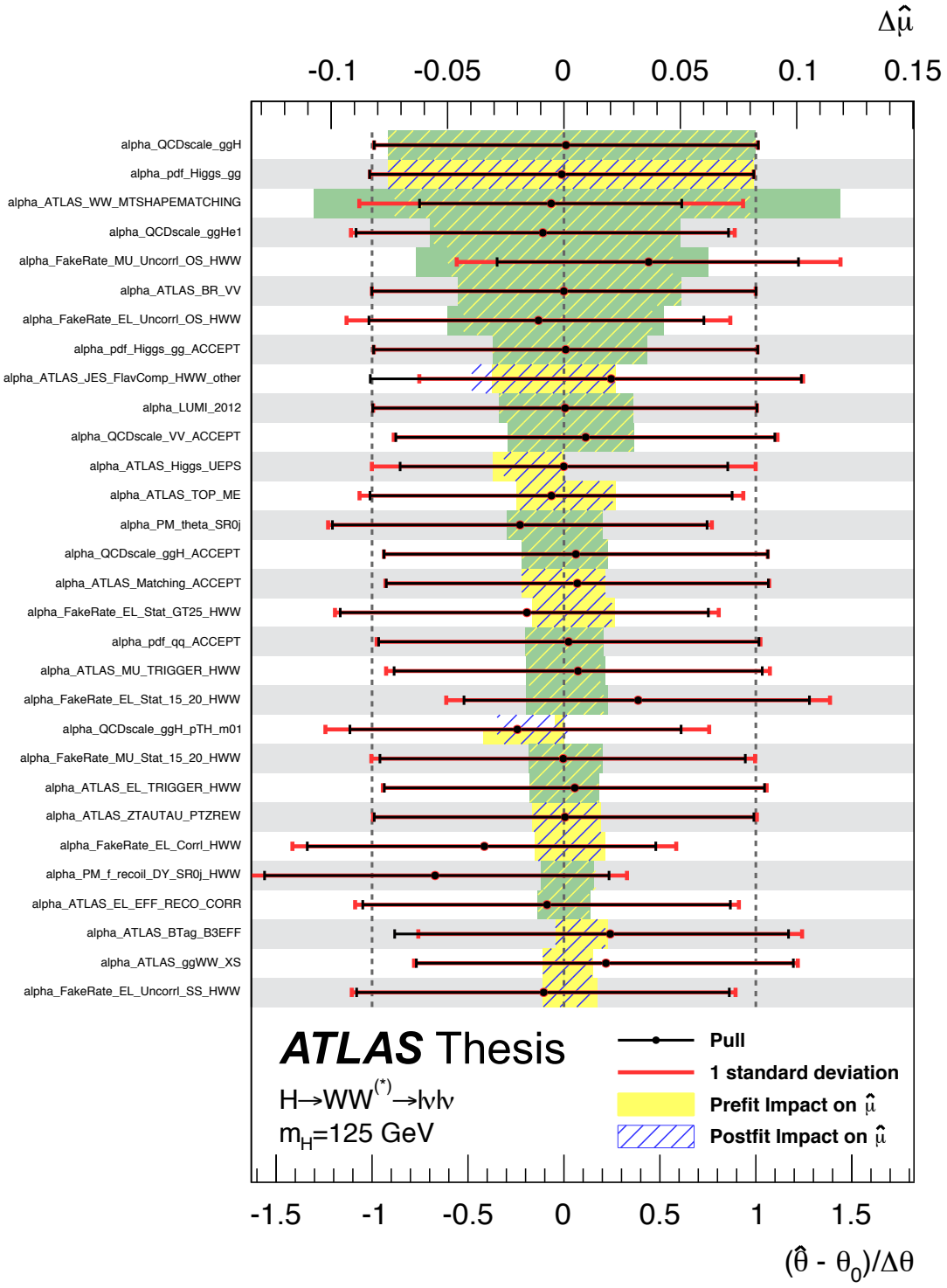



Figure 95: Fit results of the thirty most influential nuisance parameters ranked by their post fit impact on the observed signal strength $\Delta\hat{\mu}$, which is plotted versus the upper scale and illustrated by the dashed rectangles. The black points referring to the lower scale symbolize the pulls $\delta\hat{\theta}$ of the fitted parameters with the corresponding pre- and post fit uncertainties $\Delta\theta$ and $\Delta\hat{\theta}$ respectively.

| Uncertainties on the Signal Yield [%] | 0-jets | 1-jet |
|---|---|---|
| ggF→H total cross section | 10 | 9.1 |
| ggF→H acceptance model | 4.8 | 4.5 |
| ggF→H jet veto efficiency $\epsilon_0$ | 8.1 | 14 |
| ggF→H jet veto efficiency $\epsilon_1$ | – | 12 |
| VBF→H total cross section | – | 0.4 |
| VBF→H acceptance model | – | 0.3 |
| H→WW branching ratio | 4.3 | 4.3 |
| Integrated luminosity | 2.8 | 2.8 |
| Trigger efficiency | 0.8 | 0.7 |
| $f_{recoil}$ efficiency | 2.5 | 2.1 |
| Electron ID, isolation, reco. efficiency | 1.4 | 1.6 |
| Muon ID, isolation, reco. efficiency | 1.1 | 1.6 |
| Jet energy scale and resolution | 5.1 | 2.3 |
| $\not\!\!E_T$ scale and resolution | 0.6 | 1.4 |
| Pileup model | 1.2 | 0.8 |
| **Uncertainties on the background yield [%]** | **0-jets** | **1-jet** |
| WW theoretical model | 1.4 | 1.6 |
| Top theoretical model | – | 1.2 |
| VV theoretical model | – | 0.4 |
| W+jets estimate | 1.0 | 0.8 |
| QCD estimate | 0.1 | 0.1 |
| $Z/\gamma^* \to \tau\tau$ estimate | 0.6 | 0.3 |
| Integrated luminosity | – | – |
| Trigger efficiency | 0.3 | 0.3 |
| $f_{recoil}$ efficiency | 0.5 | 0.5 |
| Electron ID, isolation, reco. efficiency | 0.3 | 0.3 |
| Muon ID, isolation, reco. efficiency | 0.2 | 0.2 |
| Jet energy scale and resolution | 0.4 | 0.7 |
| b-tagging efficiency | – | 0.2 |
| Light- and c-jet mistag | – | 0.2 |
| $\not\!\!E_T$ scale and resolution | 0.1 | 0.3 |
| Pileup model | 0.4 | 0.5 |

Table 23: Relative post fit systematic uncertainties on the predicted signal yield on top and the cumulative background yield at the bottom, given for the 0-jets and the 1-jet analysis in %. Entries that are marked with a dash either do not apply or refer to values below 0.1%.

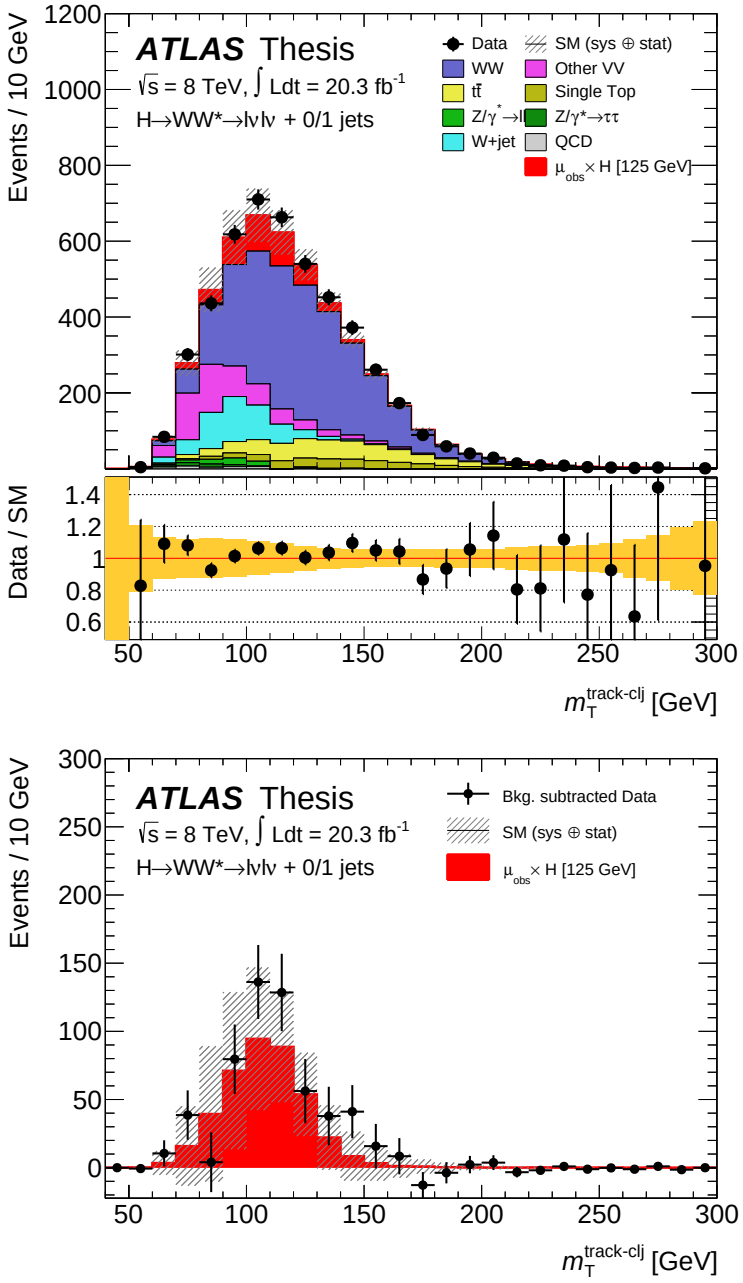



Figure 96: Post fit transverse mass distribution for the combination of all lepton flavours and both jet multiplicities reconstructed from the 2012 ATLAS dataset. The signal process for a Standard Model Higgs boson of the mass $m_H = 125$ GeV is scaled by the observed signal strength while the background processes use the post fit normalization factors. The lower figure illustrates the signal excess by showing the residuals of the data with respect to the total estimated background compared to the signal contribution.

### 7.2.1 Signal Significance

To quantify the excess observed in data, the signal significance derived in equation 7.12 is calculated. It describes the probability of the background-only hypothesis to lead to the observed experiment's outcome. Thus, it evaluates the level of disagreement between data and the background expectation without the presence of an additional signal. The distribution of the local probability $p_0$ versus $m_H$ is shown in figure 97. A scan is performed in 5 GeV steps in the range from 100 GeV to 200 GeV. A broadly distributed drop in the Higgs boson mass range between 100 GeV and 150 GeV illustrates the rather low resolution due to the loss of information about the neutrinos, which is only partially recovered by the use of missing transverse momentum. The most interesting point is the absolute minimum at the anticipated Higgs boson mass of $m_H = 125$ GeV with an observed value of $p_0^{obs} = 1.38 \times 10^{-6}$, while the expected value is $p_0^{exp} = 1.25 \times 10^{-5}$. It can also be expressed as local signal significance $Z_0$, plotted on the right axis of figure 97, with a maximum of $Z_0^{obs} = 4.69\ \sigma$ and $Z_0^{exp} = 4.31\ \sigma$ at the $m_H = 125$ GeV mass hypothesis. Although this provisional result falls marginally short of the 5 $\sigma$ discovery limit, it strongly supports the interpretation of the presence of a Standard Model Higgs boson with a mass of 125 GeV, which serves as reference point from now on.

There are two different expectation scenarios shown by the dashed lines in figure 97 to compare to. The blue one illustrates the expectation of the respective Higgs boson mass over the full mass range, whereas the black one represents the expectation for the fixed $m_H = 125$ GeV hypothesis and its resulting distribution due to the degraded resolution. Hence, they intersect at 125 GeV. The minimum of the blue curve lies around $m_H = 160$ GeV resembling the mass of two real W-bosons, which has the highest branching fraction of the analysis channel. But the observed signal significance in this region is very low. In fact, the observed distribution convincingly follows the expectation for the $m_H = 125$ GeV hypothesis. It is in general slightly larger than the expectation indicating a more signal-like behaviour, but lies well within the $1\sigma$ band across the whole mass range. Overall, the observation nicely fits in the picture of the anticipated Higgs boson mass settled by the measurements of the $H \to \gamma\gamma$, $H \to ZZ^{(*)} \to 4\ell$ and previous versions of the $H \to WW \to \ell\nu\ell\nu$ analyses. Since the whole event selection has been optimized for the mass point $m_H = 125$ GeV, the result confirms and verifies the methods and procedures presented above. However, the observations made in this analysis are not to be confused with an actual mass measurement. The statistical fit evaluates only the transverse mass of the decay products leading to the displayed resolution. For example, the 130 GeV mass point has a signal significance close to the maximum. But even with a finer grid of Monte Carlo generated signal samples the exact Higgs boson mass determination is beyond the scope of the analysis.

To enhance the investigations, table 24 shows the observed and expected signal significance for the mass point $m_H = 125$ GeV decomposed into the different analysis categories depending on the jet multiplicity and the lepton flavour. In general, the different flavour channels ($e\mu + \mu e$) show a larger significance than the same flavour channels ($ee + \mu\mu$).

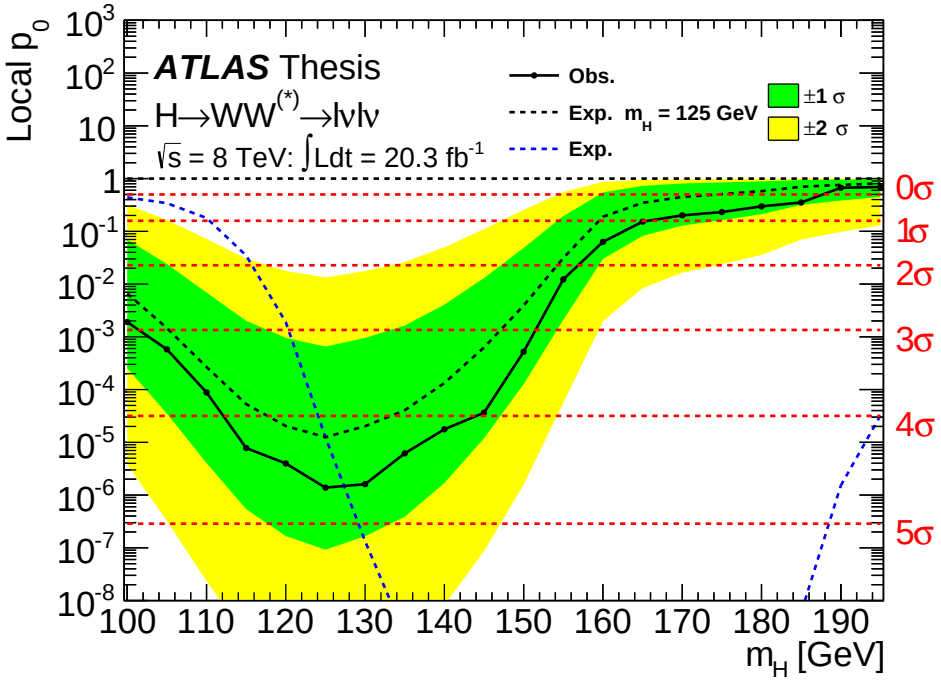



Figure 97: Local $p_0$ as a function of the Higgs boson mass. The observed values for the tested mass points are illustrated by the black dots. The median expected distribution for a Standard Model Higgs boson with $m_H = 125$ GeV is shown as black dashed line with the green (yellow) band representing the $\pm 1(2)\sigma$ range. The expectation without an imposed mass constraint is shown in blue. On the right axis, the corresponding significance $Z_0$ is marked in red.

The observed data even exceed the expectations, such that the different flavour combination of the 0-jets and the 1-jet analysis leads to the highest value of $Z_0^{obs} = 4.94$. But on the other hand, the same flavour channel numbers are considerably lower than expected. Thus, the combination of the different flavour and the same flavour analyses leads to a reduced observed significance, while the expected significance is increased. This behaviour holds for the 0-jets and the 1-jet selection as well as for their combination. The final measurement for the full combination is $Z_0^{obs} = 4.69\ \sigma$ with an expected significance of $Z_0^{exp} = 4.31\ \sigma$, as stated in the bottom right box. While it falls just below the discovery level, it nevertheless represents a substantial improvement to the earlier publicated $H \to WW \to \ell\nu\ell\nu + 0/1$ jets analyses of the ATLAS collaboration [107] with the full dataset. The expected significance has been found to be only 3.5 standard deviations. So the sensitivity increased by $\sim 20\%$. Still, the obvious deviations between the different flavour and the same flavour channels need to be followed up in the next section, which focuses on the measurement of the signal strength $\mu$.

| Channel | DF | | SF | | DF+SF | |
|---|---|---|---|---|---|---|
| | obs | exp | obs | exp | obs | exp |
| 0-jets | 4.37 | 3.41 | 0.52 | 1.36 | 4.16 | 3.54 |
| 1-jet | 2.75 | 2.35 | 0.17 | 0.96 | 2.41 | 2.49 |
| $\leq 1$ jet | 4.94 | 4.08 | 0.57 | 1.68 | 4.69 | 4.31 |

Table 24: Observed and expected local significance $Z_0$ of the different analysis channels and their combination, evaluated for a Standard Model Higgs boson mass of $m_H = 125$ GeV.

### 7.2.2 Measureing the Signal Strength

The determination of the signal significance in the previous section sets the focus on a Standard Model Higgs boson with a mass of $m_H = 125$ GeV. By measuring the actual signal strength $\hat{\mu} = \sigma^H_{obs}/\sigma^H_{SM}$ with the tools of the statistical fit, the signal model is investigated in more detail. Figure 98 shows the observed signal strength versus the hypothesized mass and compares it to the distribution of the expected signal strength of a $m_H = 125$ GeV Higgs boson across the interesting mass range with high signal significance. Therefore, at the 125 GeV mass point, the expected signal strength crosses the horizontal 1-line representing the theoretical Standard Model signal strength. The illustrated bands represent the $\pm 1\sigma$ uncertainty range and show that the observed values all agree very well with the expectation.

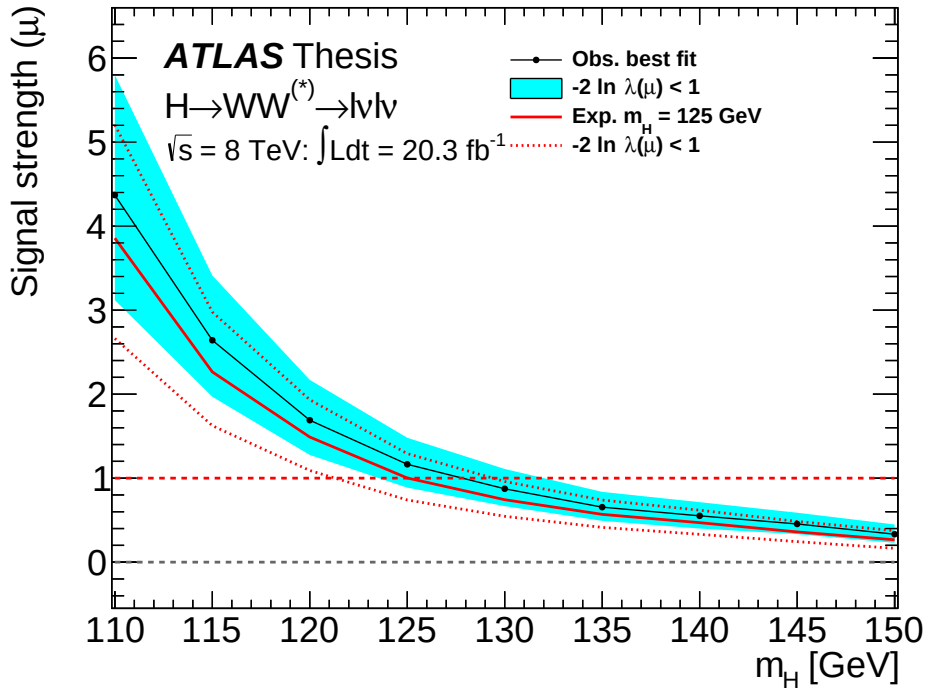



Figure 98: Observed and expected signal strength $\mu$ as a function of the Standard Model Higgs boson mass. The black dots illustrate the observed best fit values for the tested mass points and the blue band marks the $\pm 1\sigma$ range. The red lines refer to the expectation of an injected signal with fixed mass of $m_H = 125$ GeV.

The observed signal strength for a Higgs boson with a mass of $m_H = 125$ GeV has been measured to be:

$$\hat{\mu} = 1.16\ ^{+0.20}_{-0.20}\ (\text{stat})\ ^{+0.24}_{-0.20}\ (\text{syst}) = 1.16\ ^{+0.31}_{-0.28} \tag{7.19}$$

and lies slightly above the Standard Model expectation, with an uncertainty at the $\Delta\mu/\mu = 25\%$ level. Nevertheless, the measurement is compatible to $\mu_{SM} = 1$ within the uncertainties, where the contributions of the statistical and the systematical uncertainties have nearly the same magnitude. Compared to the previous version of the analysis in [107], this is a 15% improvement. The statistics have been enlarged by the refined triggering and the optimized lepton selection, while important sources of systematic uncertainties could be reduced, such as those related to the W+jets estimate. So the advanced methods have been affirmed to be very effective, as the measurement is now on the verge to become statistically limited.

For smaller masses, the observed and expected signal strengths are much higher. Due to the lower branching fraction (see figure 15), a larger cross section rate is determined from the statistical fit to describe the excess in data. For higher masses on the other hand, the $H \to WW$ branching fraction increases rapidly and the signal contributions need to be scaled down. This leads to decreasing signal strength values approaching zero for the mass range above 160 GeV. Yet again, these considerations do not choose one Higgs boson mass point over the other. The observed signal strength extracted from the statistical fit represents the value that makes the data most probable under the assumption of the implemented Higgs boson mass. So the different $\mu$-values themselves possess no information about a comparable likeliness. Their performance can only be viewed with respect to the Standard Model expectation. For example, the measured signal strength $\hat{\mu}_{130} = 0.87\ ^{+0.24}_{-0.21}$ for the 130 GeV mass point is even closer to 1 than the 125 GeV reference point. But a conclusive comparison has to consider the signal significance presented in the previous chapter, which is found to be maximal for a Higgs boson mass of 125 GeV. Conversely, the high mass range above 150 GeV is characterized by a low signal significance and is therefore omitted in figure 98.

A breakdown of the observed signal strength for the 125 GeV mass point is given in table 25, decomposing the measurement into the different anlysis channels. So the entries stand for separate statistical fits of the different selections. The different flavour channels show relatively large signal strength values, whereas the same flavour channels yield considerably lower. Generally, the 0-jets selection exhibits larger signal strength values and smaller uncertainties than the 1-jet selection. So the largest contribution comes from the $e\mu + \mu e$ channel combination without additional jets and accounts for an observed signal strength of $\hat{\mu}_{DF}^{0-jets} = 1.40\ ^{+0.45}_{-0.36}$, whereas the same flavour selection with one additional jet on the other hand determines the lowest signal strength value with only $\hat{\mu}_{SF}^{1-jet} = 0.19\ ^{+1.20}_{-1.13}$. But all observed and expected results are compatible within the uncertainties of the measurement. The final result of the $H \to WW \to \ell\nu\ell\nu + 0/1$ jets analysis is $\hat{\mu} = 1.16\ ^{+0.31}_{-0.28}$ as given in the bottom right box. This means that the observed signal cross section is 1.16 times higher than expected in the Standard Model. Although it obviously averages the mentioned effects, it does not refer to a combination of single entries from the table presented here but rather to a simultaneous fit of all channels and the estimation of the global signal strength parameter from the profile likelihood ratio.

| Channel | DF | | SF | | DF+SF | |
|---|---|---|---|---|---|---|
| | obs | exp | obs | exp | obs | exp |
| 0-jets | $1.40^{+0.45}_{-0.36}$ | $1^{+0.37}_{-0.32}$ | $0.39^{+0.75}_{-0.75}$ | $1^{+0.80}_{-0.74}$ | $1.28^{+0.41}_{-0.35}$ | $1^{+0.36}_{-0.31}$ |
| 1-jet | $1.19^{+0.59}_{-0.47}$ | $1^{+0.54}_{-0.45}$ | $0.19^{+1.20}_{-1.13}$ | $1^{+1.18}_{-1.04}$ | $1.01^{+0.53}_{-0.44}$ | $1^{+0.52}_{-0.43}$ |
| $\leq 1$ jet | $1.31^{+0.37}_{-0.30}$ | $1^{+0.30}_{-0.27}$ | $0.34^{+0.62}_{-0.60}$ | $1^{+0.64}_{-0.60}$ | $1.16^{+0.31}_{-0.28}$ | $1^{+0.29}_{-0.26}$ |

Table 25: Observed and expected signal strength of the different analysis channels and their combination for a Standard Model Higgs boson mass of $m_H = 125$ GeV. The quoted uncertainties include statistical and systematical effects.

This picture follows the trends that have already been contemplated in table 24, showing the breakdown of the local significance. The 0-jets analysis has a generally larger local significance than the 1-jet analysis, where the different flavour channels contribute the most and are noticeably above the expectations, while the values for the same flavour channels are much smaller and fall behind their expectations. But here, every entry is actually found to be compatible with unity, which strengthens the belief in a Standard Model Higgs boson at $m_H = 125$ GeV.

### 7.2.3 Exclusion Limits

The determination of the signal significance and the signal strength described in the previous sections support the existence of a Standard Model Higgs boson with a mass of $m_H = 125$ GeV. Complementary, the focus of this section lies on the possible exclusion of the other tested mass points with the $CL_S$ technique as defined in equation 7.16. The resulting 95% confidence level limits on the cross section ratio $\mu = \sigma^H/\sigma^H_{SM}$ are illustrated in figure 99, depending on the hypothesized Higgs boson mass. The black dots refer to the limits derived from the observed data. The dashed line illustrates the median expectation, which means that the integration limit is chosen to be the median of the probability density function of the test statistic for the background-only hypothesis. Similarly, to determine the green and yellow band that symbolize the uncertainty, the integration limit is set to $\pm 1(2)$ standard deviation around the median.

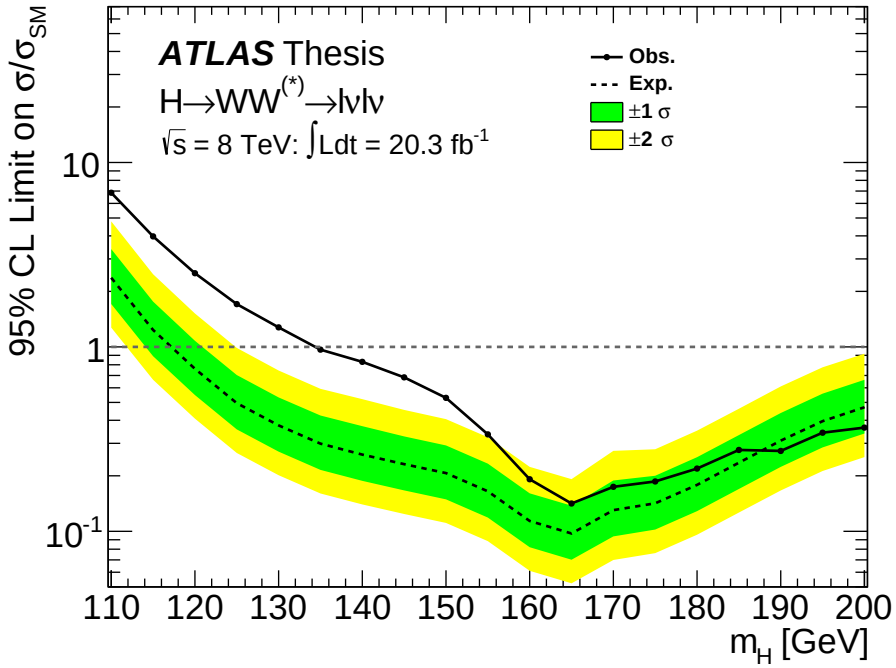



Figure 99: Observed and expected $CL_S$ exclusion limits on the signal strength $\mu = \sigma/\sigma_{SM}$ as a function of the Standard Model Higgs boson mass, stated at the 95% confidence level. The black dots illustrate the observed values of the tested mass points while the dashed line shows the median expectation and the green (yellow) band marks its $\pm 1(2)\sigma$ range.

In the low mass region of figure 99, a considerable deviation is visible. Between a Higgs boson mass of 100 GeV and 150 GeV, the observed limit lies well above the expectation and even wide outside the yellow band. That means that only immoderate high signal cross sections can be excluded at the 95% confidence level. Therefore, the

excess can be interpreted to be signal-like. In the high mass range above $m_H = 150$ GeV, the observed limit converges to the expectation and fluctuates within the $\pm 1\sigma$ band. The 1-line represents signal strength imposed by the Standard Model. Wherever the observed limit falls below one, the Standard Model cross section for the particular Higgs boson mass point is excluded. For the 2012 ATLAS data of the $H \to WW \to \ell\nu\ell\nu$ analysis, this is the case for $m_H = 135$ GeV and all the tested mass points above. So in the complete region from 135 GeV to 200 GeV, a Higgs boson with Standard Model cross section can be excluded at a 95% confidence level. The expected limit crosses the 1-line already below 120 GeV, which quantifies the general sensitivity of the analysis. At the most sensitve point of $m_H = 165$ GeV, the cross section can even be restricted to be less than 0.142 times the Standard Model expectation.

These results correspond very well to those that have already been concluded from the signal significance shown in figure 97. The broad signal-like excess in the low mass region is related to the resolution of the signal significance, whereas the complete exclusion of the high mass hypotheses affirms the absence of a significant signal. Of course, the results presented here would benefit from a finer grid of tested mass points with the respective Monte Carlo simulation samples. But the exclusion limits are only of secondary interest for this thesis. Therefore, their caclulation is reduced to a representative sample of mass points with 5 GeV steps. A complete list of the determined limits is given in table 26.

| Mass $m_H$ | $CL_S$ obs | $CL_S$ exp | $-1\sigma$ | $+1\sigma$ | $-2\sigma$ | $+2\sigma$ |
|---|---|---|---|---|---|---|
| 100 | 33.23 | 13.699 | 9.871 | 19.651 | 7.353 | 28.163 |
| 105 | 13.61 | 5.257 | 3.788 | 7.517 | 2.822 | 10.675 |
| 110 | 6.854 | 2.371 | 1.708 | 3.391 | 1.272 | 4.792 |
| 115 | 3.974 | 1.236 | 0.890 | 1.761 | 0.663 | 2.478 |
| 120 | 2.509 | 0.761 | 0.548 | 1.079 | 0.408 | 1.518 |
| 125 | 1.706 | 0.496 | 0.357 | 0.704 | 0.266 | 0.988 |
| 130 | 1.276 | 0.375 | 0.270 | 0.532 | 0.201 | 0.745 |
| 135 | 0.966 | 0.299 | 0.215 | 0.425 | 0.160 | 0.592 |
| 140 | 0.829 | 0.261 | 0.188 | 0.372 | 0.139 | 0.521 |
| 145 | 0.684 | 0.231 | 0.167 | 0.327 | 0.124 | 0.455 |
| 150 | 0.529 | 0.207 | 0.149 | 0.292 | 0.111 | 0.406 |
| 155 | 0.335 | 0.165 | 0.119 | 0.233 | 0.088 | 0.324 |
| 160 | 0.192 | 0.114 | 0.082 | 0.161 | 0.061 | 0.224 |
| 165 | 0.142 | 0.097 | 0.070 | 0.138 | 0.052 | 0.191 |
| 170 | 0.174 | 0.130 | 0.094 | 0.189 | 0.069 | 0.273 |
| 175 | 0.186 | 0.142 | 0.102 | 0.201 | 0.076 | 0.279 |
| 180 | 0.219 | 0.179 | 0.129 | 0.252 | 0.096 | 0.351 |
| 185 | 0.276 | 0.236 | 0.169 | 0.333 | 0.126 | 0.462 |
| 190 | 0.273 | 0.310 | 0.224 | 0.438 | 0.167 | 0.610 |
| 195 | 0.343 | 0.396 | 0.285 | 0.558 | 0.212 | 0.775 |
| 200 | 0.365 | 0.471 | 0.339 | 0.663 | 0.252 | 0.918 |

Table 26: $CL_S$ exclusion limits on the signal strength for the tested Standard Model Higgs boson mass points, stated at the 95% confidence level. The observed values are compared to the median expectation and its $\pm 1(2)\sigma$ range.

### 7.2.4 Inclusive Cross Section

With the observed signal strength determined from the statistical fit of the 2012 ATLAS dataset recorded at $\sqrt{s} = 8$ TeV, the total inclusive Higgs production cross section $\sigma_{pp\to H\to WW}$ can be measured. It is given by the product of the signal strength and the expected inclusive Standard Model cross section:

$$\sigma^{obs}_{pp\to H\to WW} = \hat{\mu} \cdot (\sigma_{pp\to H} \cdot BR_{H\to WW})_{SM} \tag{7.20}$$

and depends on the Higgs boson mass as can be seen in figures 14 and 15. The production cross section of a $m_H = 125$ GeV Standard Model Higgs boson is $\sigma_{SM} = 21,996\ ^{+2.28}_{-2.29}$ pb at the center of mass energy of $\sqrt{s} = 8$ TeV. This value includes the dominating gluon fusion process, the vector boson fusion process and the production in association with a vector boson that are all considered in this thesis. The expected branching fraction of the Higgs boson into two W-bosons is $BR_{H\to WW} = 0.215 \pm 0.009$. Those numbers can be found in [110] and the updated online resources [33]. The resulting product is the inclusive cross section of the Standard Model $(\sigma_{pp\to H} \cdot BR_{H\to WW})_{SM} = 4,73 \pm 0.53$ pb, that needs to be multiplied with the observed signal strength.

But for a proper treatment of the systematic uncertainties, a few modifications to the measured signal strength need to be applied. The theoretical uncertainties related to the total signal production yield are not part of the measurement described by equation 7.20. They are already accounted for in the term of the Standard Model expectation. The respective contributions are the PDF and the QCD scale uncertainties on the total cross section, as well as the uncertainty on the $H \to WW$ branching fraction. So these need to be extracted from the observed signal strength. In figure 95 and table 23, they have been identified as significant contributions with a substantial impact on $\Delta\hat{\mu}$. The correct itemization of systematic uncertainties is $\hat{\mu} = 1.16 \pm 0.2(\text{stat})\ ^{+0.20}_{-0.16}(\text{syst})\ ^{+0.14}_{-0.10}(\text{sig})$, where the last term labelled (sig) is dropped for the measurement. Finally, the resulting total inclusive cross section is obtained to be:

$$\begin{aligned} \sigma^{obs}_{pp\to H\to WW} &= 5.49 \pm 1.04\ (\text{stat})\ ^{+1.03}_{-0.90}\ (\text{syst})\ \text{pb} \\ &= 5.49\ ^{+1.46}_{-1.37}\ \text{pb} \end{aligned} \tag{7.21}$$

for a Standard Model Higgs boson with a mass of $m_H = 125$ GeV at a center-of-mass energy $\sqrt{s} = 8$ TeV, including the gluon fusion, vector boson fusion and the associated WH and ZH production processes. The uncertainty is at the $\sim 25\%$ level with nearly equal proportions of statistical and systematical uncertainties. Compared to the previous version of the $H \to WW \to \ell\nu\ell\nu$ analysis of the full $\sqrt{s} = 8$ TeV ATLAS dataset [107] published in 2013, the result constitutes a 15% improvement which has

already been stated for the observed signal strength. It can be accredited to the enlarged statistics due to the inclusion of the dilepton triggers on the one hand, and to reduced systematic uncertainties on the refined methods of the background determination such as the W+jets estimate on the other hand.

# 8

# Combination of Search Channels

The analysis methods and procedures presented in this thesis have been incorporated into the paper [101]. However, it also contains further investigations and measurements out of the scope of this thesis. Those additional results are summarized below to provide a thorough picture of the ATLAS $H \to WW \to \ell\nu\ell\nu$ channel.

The major supplement is a dedicated investigation of the vector boson fusion production mode dominating the event selection with two or more jets. The special VBF topology is exploited by a multivariate technique. A so called boosted decision tree (BDT) categorizes signal and background events by a multi-dimensional cut selection on the VBF related common observables. The gluon fusion Higgs boson production is treated as (profiled) background in this case. Top quark related backgrounds dominate the different flavour channels, while the same flavour channels share large contributions from Drell-Yan processes. In contrast to the presented thesis, the Drell-Yan background in the VBF analysis is estimated with the ABCD method. The statistical fit procedure is basically the same. But the fit variable is the distribution of the BDT discriminant exploiting only two different signal regions, given by the lepton flavour. The resulting post-fit distributions are shown in figure 100. The local significance of the VBF signal has been observed to be 3.8 standard deviations and establishes the evidence of the vector boson fusion Higgs production.

A complementary $n_j \geq 2$ gluon fusion enriched signal region is also examined for the $\sqrt{s} = 8$ TeV dataset. It is basically selected by vetoing the VBF topology, where only the different lepton flavour channels are used. All other methods and procedures follow a similar prescription as the $n_j \leq 1$ analyses presented in this thesis. Figure 101 shows the post-fit transverse mass distribution of the $n_j \geq 2$ ggF-enriched signal region. The observed local signal significance of this category alone is 1.4 standard deviations.

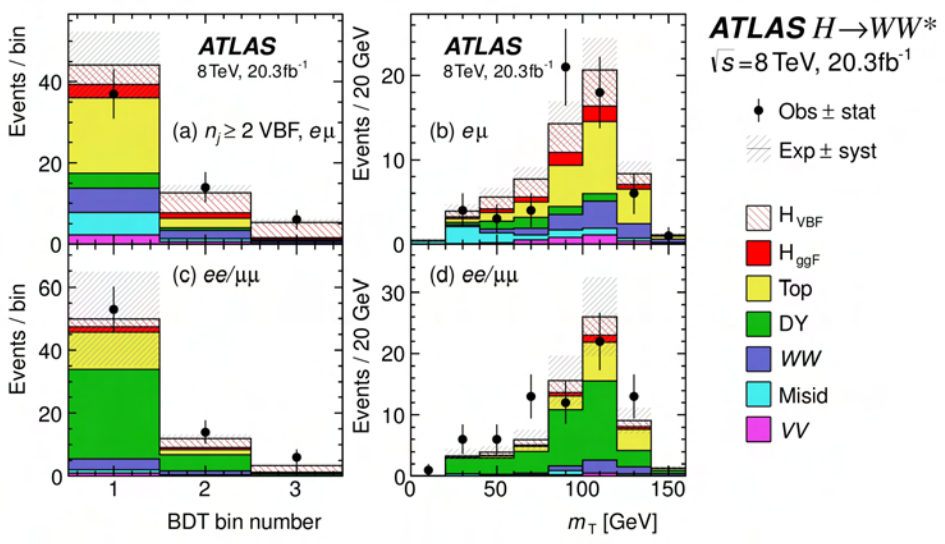



Figure 100: Post-fit BDT and transverse mass distributions in the $n_j \geq 2$ VBF-enriched category in the 8 TeV data analysis [101]. Figure (a) shows the BDT output in $e\mu$, (b) $m_T$ in $e\mu$, (c) BDT output in $ee/\mu\mu$, and (d) $m_T$ in $ee/\mu\mu$. For (b) and (d), the three BDT bins are combined.

Moreover, the $n_j \leq 1$ ggF and the $n_j \geq 2$ VBF analyses have also been performed on the $\sqrt{s} = 7$ TeV ATLAS dataset from 2011 with slight adjustments according to the run conditions. All these contributions are summed up to separately measure the gluon fusion and the vector-boson-fusion signal strength. The results are:

$$\begin{aligned} \mu_{ggF} &= 1.02 \pm 0.19\ (\text{stat})\ ^{+0.22}_{-0.18}\ (\text{syst}) = 1.02\ ^{+0.29}_{-0.26} \\ \mu_{VBF} &= 1.27\ \ ^{+0.44}_{-0.40}\ (\text{stat})\ ^{+0.30}_{-0.21}\ (\text{syst}) = 1.27\ ^{+0.53}_{-0.45} \end{aligned} \tag{8.1}$$

quoted at $m_H = 125.36$ GeV. These two results can be used to measure the Higgs boson couplings to fermions and vector bosons and to test their compatibility with the Standard Model predictions via the ratio $\mu_{VBF}/\mu_{ggF}$. The likelihood scan of this ratio is shown in figure 102. The individual coupling strength values are consistent with the Standard Model expectation.

The full combination of all $H \to WW \to \ell\nu\ell\nu$ search channels with the 2011 and 2012 ATLAS dataset leads to the observed signal strength of:

$$\mu = 1.09\ ^{+0.16}_{-0.15}\ (\text{stat})\ ^{+0.17}_{-0.14}\ (\text{syst}) = 1.09\ ^{+0.23}_{-0.21} \tag{8.2}$$

including all production modes. The local signal significance is found to be $Z_0^{obs} = 6.1$ standard deviations while the expected value is 5.8 $\sigma$ and establishes a discovery-level Standard Model Higgs boson signal at $m_H = 125.36$ GeV.

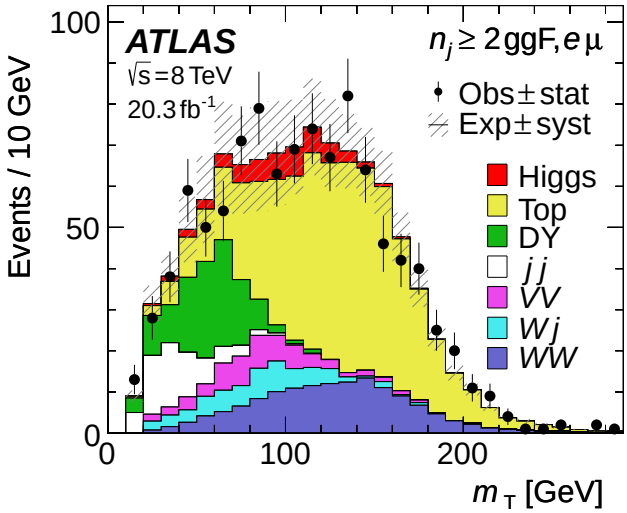



Figure 101: Post-fit transverse mass distribution in the $n_j \geq 2$ ggF-enriched category in the 8 TeV analysis [101].

The measurement of the inclusive cross section is performed for various combinations of channels. The most prominent results are:

$$\begin{aligned} \sigma_{ggF}^{8TeV} \cdot BR_{H\to WW} &= 4.6\ ^{+1.2}_{-1.1}\ \text{pb} \\ \sigma_{VBF}^{8TeV} \cdot BR_{H\to WW} &= 0.51\ ^{+0.22}_{-0.17}\ \text{pb} \end{aligned} \tag{8.3}$$

for the high statistics 2012 dataset collected at $\sqrt{s} = 8$ TeV, where $4.2 \pm 0.5$ pb and $0.35 \pm 0.02$ pb are the values predicted by the Standard Model respectively. Finally, also the fiducial cross sections for the gluon fusion production in the 0-jets and the 1-jet channel are calculated.

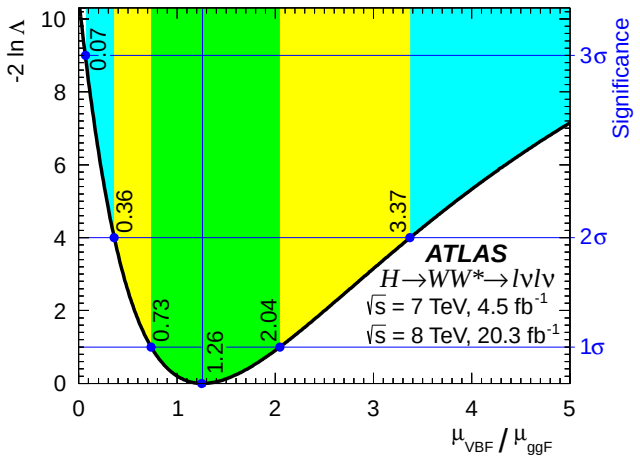



Figure 102: Likelihood scan as a function of $\mu_{VBF}/\mu_{ggF}$ for $m_H = 125.36$ GeV [101]. The value of the likelihood at $\mu_{VBF}/\mu_{ggF} = 0$ gives the significance of the VBF signal at 3.2 standard deviations. The inner (middle) [outer] band shaded in green (yellow) [blue] represents the one (two) [three] standard deviation uncertainty around the central value represented by the vertical line.

# 9

# Summary and Outlook

The discovery of the new elementary particle in 2012 is surely one of the biggest scientific achievements of the recent years. The high expectations on the major project of the Large Hadron Collider and its sophisticated detectors have been fulfilled so far and at the same time set the focus on the accurate measurement of the properties of the Higgs boson.

The thesis at hand describes the thorough analysis of a Standard Model Higgs boson in the $H \to WW \to \ell\nu\ell\nu$ decay channel with the ATLAS detector at the Large Hadron Collider. The dataset used for this examination comprises 20.3 fb$^{-1}$ and has been recorded at a center-of-mass energy of $\sqrt{s} = 8$ TeV. The introduction of the special characteristics of the decay channel has been used to develop a general search strategy. The focus has been set on the gluon fusion process, which is the dominant production mode of the Higgs boson at the Large Hadron Collider. Consequently, the analysis is restricted to events with none or exactly one associated jet. The essential element of the analysis is the particular event selection and the involved methods for the determination of the background processes. Very important in this context is the data-driven estimation of the W+jets background, which enters the event selection through mis-identified leptons. In order to optimize the evaluation of the selected events, a simultaneous fit of different signal and background regions is applied in consideration of the systematic uncertainties. For that purpose, the maximum likelihood method has been introduced and a profile likelihood ratio has been motivated as test statistic.

A signal significance of $Z_0^{obs} = 4.69\ \sigma$ for a Standard Model Higgs boson of the mass $m_H = 125$ GeV has been observed and remains only slightly beneath the discovery level. The expected value is $Z_0^{exp} = 4.31\ \sigma$. Moreover, the existence of a Standard Model Higgs boson in the range of 135–200 GeV has been excluded at a 95% confidence level. The ratio of the observed and the predicted Standard Model production cross section is called signal strength. Its measurement at the mass point $m_H = 125$ GeV yields the

value of $\hat{\mu} = 1.16\ ^{+0.31}_{-0.28}$ and is therefore in agreement with the Standard Model. From this result, the inclusive production cross section of $\sigma^{obs}_{pp\to H\to WW} = 5.49\ ^{+1.46}_{-1.37}$ pb can be calculated. This translates to a 15% improvement compared to the previous version of the measurement [107], which has utilized the same dataset. Most of the insights gained within the compilation of this thesis are documented in the publication [101].

The improved results can be attributed to the increased acceptance due to the use of dilepton triggers on the one hand and the reduction of systematic uncertainties on the other hand. The lowered thresholds of the transverse momentum of the leptons open up the selection for signal events in the considered mass region, but of course also for background processes. This leads to increased statistics in the signal and background control regions reducing the statistical uncertainties and thus enhancing the potential of the global fit. However, the probably most relevant contribution to the ATLAS analysis constituted in the development of this thesis are the refined methods for the W+jets background estimation. Without the precise description of this process, which is typically occuring at low transverse momenta and is hard to model, the extension of the phase space is not accessible in order to achieve an increase in signal significance. Hence, it has been essential to considerably improve the data-driven technique of the determination of jets mis-identified as leptons. For this purpose, the fake factor method has been advanced to use a Z+jets control sample. This procedure reduces the associated systematic uncertainties significantly compared to the previously established treatment. In this course, also the lepton selection has been optimised. Particularly the use of the likelihood based electron identification effectuated an improved background reduction, which influences the purity of the signal and control regions. Furthermore, the approach has been extended to a data-driven estimate of the QCD background.

After a nearly two-year shutdown of revision and upgrade, the Large Hadron Collider starts a new period of data-taking in 2015. The so called Run-2 will provide proton collisions at a center-of-mass energy of 13 TeV and is planned to last until 2018. The bunch-spacing is reduced to its design value of 25 ns and further increases the luminosity. A dataset of about 100 $\text{fb}^{-1}$ is aspired. The Higgs boson production cross section is approximately doubled. With such high event statistics, the era of precision measurement of the Higgs boson is initiated. Characteristic properties such as the differential cross section, the couplings or the spin of the newly discovered elementary particle advance into focus. But the increased amount of pileup events per bunch-crossing requires refined reconstruction methods for the future analyses. Despite the considerably larger instantaneous luminosity, the trigger thresholds of the leptons are attempted to be maintained on the level of 2012. Thus, the general strategy of the Higgs boson analyses can be pursued. Additionally, the higher collision energy enlarges the accessible mass range. This promotes the direct and indirect search for new phenomena, as for example predicted by supersymmetric theories. The data-taking period of Run-2 will therefore approach the precision measurement of the heaviest fundamental particles of the Standard Model as well as explore new physics beyond the Standard Model to possibly find evidence for dark matter or likewise. In any case, high energy physics at the Large Hadron Collider promises further exciting insights into the fundamentals of nature.

# List of Figures

# List of Tables

# References


[1] F. Halzen and A. Martin. *Quarks and Leptons*. John Wiley & Sons, 1984.

[2] M. Peskin and D. V. Schroeder. *An Introduction to Quantum Field Theory*. Westview Press, 1995.

[3] F. Scheck. *Electroweak and Strong Interactions*. Springer, 1996.

[4] G. Altarelli. The Standard Model of Particle Physics. arXiv:hep-ph/0510281.

[5] S. Weinberg. The Making of the Standard Model. arXiv:hep-ph/0401010.

[6] P. W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508, 1964.

[7] R. Brout and F. Englert. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett.*, 13:321, 1964.

[8] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global Conservation Laws and Massless Particles. *Phys. Rev. Lett.*, 13:585, 1964.

[9] M. Kobayashi and T. Maskawa. CP-Violation in the Renormalizable Theory of Weak Interaction. *Progress of Theoretical Physics*, 49:652–657, 1973.

[10] J. M. Campbell, J. W. Huston, and W. J. Stirling. Hard Interactions of Quarks and Gluons: a Primer for LHC Physics. arXiv:physics/0611148.

[11] J. Collins and D. Soper. The Theorems of perturbative QCD, 1987.

[12] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. arXiv:0901.0002 [hep-ph].

[13] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne. Uncertainties of predictions from parton distributions. I: Experimental errors. arXiv:hep-ph/0211080.

[14] M. A. Dobbs, S. Frixione, E. Laenen, and K et al. Tollefson. Les Houches Guidebook to Monte Carlo Generators for Hadron Collider Physics. arXiv:hep-

ph/0403045. Compiled for the Workshop 'Physics at TeV Colliders', Les Houches, France, May 2003.

[15] M. H. Seymour and M. Marx. Monte Carlo Event Generators. arXiv:1304.6677 [hep-ph].

[16] M. L. Mangano and T. J. Stelzer. Tools for the Simulation of Hard Hadronic Collisions, 2005.

[17] T Sjostrand, S. Mrenna, and P. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 0605:026, 2006. arXiv:hep-ph/0603175.

[18] G. Corcella et al. HERWIG 6.5: An event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes). *JHEP*, 0101:010, 2001. arXiv:hep-ph/0011363.

[19] S. Frixione and B. R. Webber. Matching NLO QCD computations and parton shower simulations. *JHEP*, 0206:029, 2002. arXiv:hep-ph/0204244.

[20] S. Alioli, P. Nason, C. Oleari, and E. Re. NLO Higgs boson production via gluon fusion matched with shower in POWHEG. *JHEP*, 0904:002, 2009. arXiv:0812.0578 [hep-ph].

[21] A. Djouadi. The Anatomy of Electro-Weak Symmetry Breaking I. *Phys. Rept.*, 457:1–216, 2008. arXiv:hep-ph/0503172.

[22] T. Hambye and K. Riesselmann. SM Higgs mass bounds from theory. arXiv:9708416 [hep-ph].

[23] J. Ellis, G. Ridolfi, and F. Zwirner. Higgs Boson Properties in the Standard Model and its Supersymmetric Extensions. *Competes Rendus Physique*, 8:999–1012, 2007. arXiv:hep-ph/0702114.

[24] The LEP Working Group for Higgs Boson Searches et al. Search for the Standard Model Higgs Boson at LEP. arXiv:hep-ex/0306033.

[25] The TEVNPH Working Group. Combined CDF and DØ Upper Limits on Standard Model Higgs-Boson Produktion with 2.1 - 4.2 $\text{fb}^{-1}$ of Data. arXiv:0911.3939 [hep-ex].

[26] The ALEPH, CDF, D0, DELPHI, L3, OPAL, and SLD Collaborations. Precision Electroweak Measurements and Constraints on the Standard Model. arXiv:1012.2367 [hep-ex].

[27] The LEP Electroweak Working Group. `http://lepewwg.web.cern.ch/LEPEWWG/`. online resource.

[28] M. Spira, A. Djouadi, D. Graudenz, and P. M. Zerwas. Higgs boson production at the LHC. *Nucl. Phys.*, B453:17–82, 1995. arXiv:hep-ph/9504378.

[29] R. V. Harlander and W. B. Kilgore. Next-to-Next-to-Leading Order Higgs Production at Hadron Colliders. *Phys. Rev. Lett.*, 88:201801, 2002. arXiv:hep-ph/0201206.

[30] S. Catani, D. de Florian, M. Grazzini, and P. Nason. Soft-gluon resummation for Higgs boson production at hadron colliders. *JHEP*, 0307:028, 2003. arXiv:hep-ph/0306211.

[31] G. Degrassi and F. Maltoni. Two-loop electroweak corrections to Higgs production at hadron colliders. *Phys. Lett.*, B600:255–260, 2004. arXiv:hep-ph/0407249.

[32] The LHC Higgs Cross Section Working Group. Handbook of LHC Higgs cross sections: 1. Inclusive Observables. arXiv:1101.0593 [hep-ph].

[33] The LHC Higgs Cross Section Working Group. `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections`. online resource.

[34] The ATLAS Collaboration. `https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/`. online resource.

[35] The ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys.Lett.*, B716:1–29, 2012. arXiv:1207.7214 [hep-ph].

[36] The CMS Collaboration. Observation of a new a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys.Lett.*, B716:30–61, 2012. arXiv:1207.7235 [hep-ph].

[37] 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. *Phys. Rev. Lett.*, 114:191803, 2015. arXiv:1503.07589 [hep-ex].

[38] The Gfitter Group. The global electroweak fit at NNLO and prospects for the LHC and ILC. *Eur. Phys. J.*, C74:3046, 2014. arXiv:1407.3792 [hep-ph].

[39] W Herr and B. Muratori. Concept of luminosity. In *CERN Accelerator School, Zeuthen 2003*, pages 361–378, 2006.

[40] J. Rossbach and P. Schmüser. Basic course on accelerator optics. In *CERN Accelerator School, Jyväskylä 1992*, pages 17–88, 1994.

[41] L. Evans and P. Bryant. LHC machine. *JINST*, 3:S08001, 2008.

[42] J. Haffner. The CERN accelerator complex. OPEN-PHO-ACCEL-2013-056, Oct 2013. `http://cds.cern.ch/record/1621894`.

[43] M. A. Furman. The Møller Luminosity Factor. LBNL-53553, CBP Note-543.

[44] S. Maettig. Luminosity Measurements with the ATLAS Detector. CERN-THESIS-2012-249, Aug 2012.

[45] H. Burkhardt and P. Grafstrom. Absolute Luminosity from Machine Parameters. LHC Project Report 1019, 2007.

[46] The ATLAS Collaboration. Expected Performance of the ATLAS Experiment: Detector, Trigger and Physics. arXiv:0901.0512 [hep-ex]. CERN-OPEN-2008-020.

[47] The ATLAS Collaboration. Technical Proposal for a General-Purpose pp Experiment at the Large Hadron Collider at CERN. CERN-LHCC-9443.

[48] The ATLAS Collaboration. The ATLAS experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.

[49] ATLAS Outreach. ATLAS Fact Sheet : To raise awareness of the ATLAS detector and collaboration on the LHC. `http://cds.cern.ch/record/1457044`, 2010.

[50] K. Kleinknecht. *Detektoren für Teilchenstrahlung*. Teubner, 1992.

[51] W. Demtroeder. *Kern-, Teilchen- und Astrophysik*. Springer, 2004.

[52] The ATLAS Collaboration. Alignment of the ATLAS Inner Detector and its Performance in 2012. ATLAS-CONF-2014-047.

[53] J. Pequenao. Computer generated image of the ATLAS inner detector. `http://cds.cern.ch/record/1095926`, Mar 2008.

[54] The ATLAS Collaboration. Particle Identification Performance of the ATLAS Transition Radiation Tracker. ATLAS-CONF-2011-128.

[55] J. Pequenao. Computer Generated image of the ATLAS calorimeter. `http://cds.cern.ch/record/1095927`, Mar 2008.

[56] R. Achenbach et. al. The ATLAS Level-1 Calorimeter Trigger. *JINST*, 3:03001, 2008.

[57] T. Cornelissen, M. Elsing, I. Gavrilenko, W. Liebig, E. Moyse, and A. Salzburger. The new ATLAS Track Reconstruction (NEWT). *J. Phys.:Conf. Ser.*, 119:032014, 2008. International Conference on Computing in High Energy and Nuclear Physics (CHEP 07).

[58] The ATLAS Collaboration. Performance of the ATLAS Silicon Pattern Recognition Algorithm in Data and Simulation at $\sqrt{s} = 7$ TeV. ATLAS-CONF-2010-072.

[59] The ATLAS Collaboration. Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s} = 7$ TeV in the ATLAS experiment. ATLAS-CONF-2010-069.

[60] W. Waltenberger, R. Frühwirt, and P. Vanlaer. Adaptive vertex fitting. *J. Phys.*, G34:N343–N356, 2007.

[61] The ATLAS Collaboration. Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data. *Eur. Phys. J.*, C74:2941, 2014. arXiv:1404.2240 [hep-ex].

[62] The ATLAS Collaboration. Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data. ATLAS-CONF-2014-032.

[63] The ATLAS Collaboration. Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung. ATLAS-CONF-2012-047.

[64] J. Alison, K. Brendlinger, S. Heim, and C. M. Lester. Description and Performance of the Electron Likelihood Tool at ATLAS using 2012 LHC Data. ATL-COM-PHYS-2013-378.

[65] M. Hance, D. Olivito, and H. H. Williams. Performance Studies for e/gamma Calorimeter Isolation. ATL-COM-PHYS-2011-1186.

[66] The ATLAS Collaboration. Preliminary results on the muon reconstruction efficiency, momentum resolution, and momentum scale in ATLAS 2012 pp collision data. ATLAS-CONF-2013-088.

[67] The ATLAS Collaboration. A muon identification and combined reconstruction procedure for the ATLAS detector at the LHC at CERN. *IEEE Trans. Nucl. Sci.*, 51:3030–3033, 2004.

[68] M. Cacciari, G. P. Salam, and G. Soyez. The anti-$k_t$ jet clustering algorithm. *JHEP*, 04:063, 2008. arXiv:0802.1189 [hep-ph].

[69] The ATLAS Collaboration. Jet energy measurement with the atlas detector in proton-proton collisions at sqrt(s) = 7 tev. *Eur. Phys. J.*, C73:2304, 2013. arXiv:1112.6426 [hep-ph].

[70] The ATLAS Collaboration. Jet energy resolution and selection efficiency relative to track jets from in-situ techniques with the atlas detector using proton-proton collisions at a center of mass energy $\sqrt{s} = 7$ tev. ATLAS-CONF-2010-054.

[71] The ATLAS Collaboration. Performance of b-jet identification in atlas. ATL-PHYS-PROC-2014-040.

[72] The ATLAS Collaboration. Measuring the b-tag efficiency in a top-pair sample with 4.7 $fb^{-1}$ of data from the atlas detector. ATLAS-CONF-2012-097.

[73] The ATLAS Collaboration. Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at 8 TeV. ATLAS-CONF-2013-082.

[74] A. J. Barr, B. Gripaios, and C. G. Lester. Measuring the Higgs boson mass in dileptonic W-boson decays at hadron colliders. *JHEP*, 0907:072, 2009. arXiv:0902.4864 [hep-ph].

[75] C. A. Nelson. Correlation between decay planes in Higgs-boson decays into a W pair (into a Z pair). *Phys. Rev.*, D37:1220, 1988.

[76] The ATLAS Collaboration. Measurement of the $W^+W^-$ production cross section in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. ATLAS-CONF-2014-033.

[77] J. M. Campbell, R. K. Ellis, and C. Williams. Radiative corrections to $Zb\bar{b}$ production. *Phys. Rev.*, D62:114012, 2000. arXiv:hep-ph/0006304.

[78] J. M. Campbell, R. K. Ellis, and C. Williams. Vector boson pair production at the LHC. *JHEP*, 1107:018, 2011. arXiv:1105.0020 [hep-ph].

[79] The ATLAS Collaboration. Measurement of the $W \to \ell\nu$ and $Z/\gamma^* \to \ell\ell$ production cross sections in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *JHEP*, 1012:060, 2010. arXiv:1010.2130 [hep-ex].

[80] The ATLAS Collaboration. Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with $b$-tagged jets in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS detector. *Eur. Phys. J.*, C74:3109, 2014. arXiv:1406.5375 [hep-ex].

[81] The ATLAS Collaboration. Evidence for the associated production of a W boson and a top quark in ATLAS at $\sqrt{s} = 7$ TeV. *Phys. Lett.*, B716:142–159, 2012. arXiv:1205.5764 [hep-ex].

[82] The ATLAS Collaboration. Search for s-Channel Single Top-Quark Production in pp Collisions at $\sqrt{s} = 7$ TeV. ATLAS-CONF-2011-118.

[83] The ATLAS Collaboration. Measurement of the t-channel single top-quark production cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Phys. Lett.*, B717:330–350, 2012. arXiv:1205.3130 [hep-ex].

[84] J. Adelman et al. ATLAS Offline Data Quality Monitoring. *J. Phys.:Conf. Ser.*, 219:042018, 2010. 17th International Conference on Computing in High Energy and Nuclear Physics (CHEP09).

[85] The ATLAS Collaboration. Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector at the LHC. *Eur. Phys. J.*, C71:1630, 2011. arXiv:1101.2185 [hep-ex].

[86] The ATLAS Collaboration. Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC. *Eur. Phys. J.*, C73:2518, 2013. arXiv:1302.4393 [hep-ex].

[87] The ATLAS Collaboration. ATLAS Computing - Technical Design Report. CERN-LHCC-2005-002. ATLAS TDR 017.

[88] The ATLAS Collaboration. The ATLAS Simulation Infrastructure. *Eur. Phys. J.*, C70:823–874, 2010. arXiv:1005.4568 [physics.ins-det].

[89] S. Agnostelli et al. Geant4 – a simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.

[90] S. Frixione, P. Nason, and C. Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. arXiv:0709.2092 [hep-ph].

[91] A. Djouadi, J. Kalinowski, and M. Spira. HDECAY: a Program for Higgs Boson Decays in the Standard Model and its Supersymmetric Extension. *Comput. Phys. Commun.*, 108:56–74, 1998. arXiv:hep-ph/9704448.

[92] B. P. Kersevan and E. Richter-Was. The Monte Carlo Event Generator AcerMC 2.0 with Interfaces to PYTHIA 6.2 and HERWIG 6.5. arXiv:hep-ph/0405247.

[93] T. Gleisberg et al. Event generation with SHERPA 1.1. *JHEP*, 0902:007, 2009. arXiv:0811.4622 [hep-ph].

[94] J. Alwall et al. Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions. *Eur. Phys. J.*, C53:473–500, 2008. arXiv:0706.2569 [hep-ph].

[95] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 0307:001, 2003. arXiv:hep-ph/0206293.

[96] J.M. Butterworth, J.R. Forshaw, and M.H. Seymour. Multiparton Interactions in Photoproduction at HERA. arXiv:hep-ph/9601371.

[97] T. Binoth, M. Ciccolini, N. Kauer, and M. Krämer. Gluon-induced W-boson pair production at the LHC. *JHEP*, 0612:046, 2006. arXiv:hep-ph/0611170.

[98] The ATLAS Collaboration. Object selections in the $H \to WW^{(*)} \to \ell\nu\ell\nu$ analysis with 20.3 fb$^{-1}$ of data collected with the ATLAS detector at $\sqrt{s} = 8$ TeV. ATLAS-CONF-2013-1504.

[99] The ATLAS Collaboration. Selection of jets produced in proton-proton collisions with the ATLAS detector using 2011 data. ATLAS-CONF-2012-020.

[100] The ATLAS Collaboration. Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of $\sqrt{s} = 7$ TeV. ATLAS-CONF-2010-038.

[101] The ATLAS Collaboration. Observation and measurement of Higgs boson decays to $WW^*$ with the ATLAS detector. *Phys. Rev. D*, 92:012006, Jul 2015.

[102] R. K. Ellis, I. Hinchliffe, M. Soldate, and J. J. Van der Bij. HIGGS DECAY TO $\tau^+\tau^-$ – A possible signature of intermediate mass Higgs bosons at high energy hadron colliders. *Nucl. Phys.*, B297:221–243, 1988.

[103] The ATLAS Collaboration. Background estimation in the $H \to WW^{(*)} \to \ell\nu\ell\nu$ analysis with 20.7 fb$^{-1}$ of data collected with the ATLAS detector at $\sqrt{s} = 8$ TeV. ATLAS-CONF-2013-1630.

[104] The ATLAS Collaboration. Analysis of $H \to WW \to \ell\nu\ell\nu$ ggF and VBF production modes with 20 fb$^{-1}$ and 5 fb$^{-1}$ of data collected with the ATLAS detector at $\sqrt{s} = 8$ and 7 TeV. ATLAS-CONF-2014-466.

[105] W. J. Stirling and E. Vryonidou. Charm production in association with an electroweak gauge boson at the LHC. *Phys. Rev. Lett.*, 109:082002, 2012. arXiv:1203.6781 [hep-ph].

[106] The ATLAS Collaboration. Measurement of the production of a W boson in association with a charm quark in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *JHEP*, 05:068, 2014.

[107] The ATLAS Collaboration. Measurements of the properties of the Higgs-like boson in the $WW^{(*)} \to \ell\nu\ell\nu$ decay channel with the ATLAS detector using 25 fb$^{-1}$ of proton-proton collision data. In *48th Rencontres de Moriond on QCD and High Energy Interactions*, 2013. ATLAS-CONF-2013-030.

[108] The ATLAS Collaboration. Theoretical studies for the $H \to WW$ search and measurement. ATL-COM-PHYS-2013-1541.

[109] The LHC Higgs Cross Section Working Group. Handbook of LHC Higgs cross sections: 2. Differential Distributions. CERN-2012-002. arXiv:1201.3084 [hep-ph].

[110] The LHC Higgs Cross Section Working Group. Handbook of LHC Higgs cross sections: 3. Higgs Properties. CERN-2013-004. arXiv:1307.1347 [hep-ph].

[111] A. Bierweiler, T. Kasprzik, and J. H. Kühn. Vector-boson pair production at the LHC to $\mathcal{O}(\alpha^3)$ accuracy. arXiv:1305.5402 [hep-ph].

[112] T. Melia, K. Melnikov, R. Rontsch, M. Schulze, and G. Zanderighi. Gluon fusion contribution to W+W- + jet production. arXiv:1205.6987 [hep-ph].

[113] M. Bonvini, F. Caola, S. Forte, K. Melnikov, and G. Ridolfi. Signal-background interference effects for $gg \to H \to W^+W^-$ beyond leading order. *Phys. Rev.*, D88:034032, 2013. arXiv:1304.3053 [hep-ph].

[114] The ATLAS Collaboration. Measurement of hard double-parton interactions in $W(\to \ell\nu) + 2$ jet events at $\sqrt{s} = 7$ TeV with the ATLAS detector. *New J. Phys.*, 15:0330038, 2013. arXiv:1301.6872 [hep-ex].

[115] B. Blok, Y. Dokshitzer, L. Frankfurt, and M. Strikman. Perturbative QCD correlations in multi-parton collisions. arXiv:1306.3763 [hep-ph].

[116] The ATLAS Collaboration. Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data. *Eur.Phys.J.*, C74:3130, 2014. arXiv:1407.3935 [hep-ex].

[117] The ATLAS Collaboration. Jet energy measurement and its systematic uncertainty in proton–proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. CERN-PH-EP-2013-222. arXiv:1406.0076 [hep-ex].

[118] The ATLAS Collaboration. Calibration of $b$-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment. ATLAS-CONF-2014-004.

[119] The ATLAS Collaboration. Calibration of the performance of $b$-tagging for $c$ and light-flavour jets in the 2012 ATLAS data. ATLAS-CONF-2014-046.

[120] The ATLAS Collaboration. Performance of missing transverse momentum reconstruction in proton-proton collisions at 7 TeV with ATLAS. *Eur.Phys.J.*, C72:1844, 2012. arXiv:1108.5602 [hep-ex].

[121] R. J. Barlow. *Statistics - A Guide to the Use of Statistical Methods in the Physical Sciences*. John Wiley & Sons, 1999.

[122] G. Cowan. *Statistical Data Analysis*. Oxford University Press, 1998.

[123] R. Barlow and C. Beeston. Fitting using finite Monte Carlo samples. *Comput. Phys. Commun.*, 77:219–282, 1993.

[124] G. Cowan, K. Cranmer, E. Gross, and O. Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur.Phys.J.*, C71:1554, 2011. arXiv:1007.1727 [physics.data-an].

[125] G. J. Feldman and R. D. Cousins. A Unified Approach to the Classical Statistical Analysis of Small Signals. *Phys. Rev.*, D57:3873–3889, 1998. arXiv:physics/9711021 [physics.data-an].

[126] A. L. Read. Presentation of search results: The $CL_S$ technique. *J. Phys.*, G28:2693, 2002.

[127] R. Brun and F. Rademakers. ROOT: An object oriented data analysis framework. *Nucl. Instrum. Meth.*, A389:81–86, 1997.

[128] W. Verkerke and D. Kirkby. The RooFit toolkit for data modeling. In *CHEP03 Conference Proceedings*, 2003. arXiv:physics/0306116 [physics.data-an].

[129] L. Moneta, K. Belasco, and K. Cranmer et al. The RooStats Project. In *ACAT2010 Conference Proceedings*, 2010. arXiv:1009.1003 [physics.data-an].

[130] F. James and M. Roos. Minuit - a system for function minimization and analysis of the parameter errors and correlations. *Comput. Phys. Commun.*, 10:343–367, 1975.

# A

# Selection Cutflow

| | Signal [125 GeV] | $WW$ | Other VV | $t\bar{t}$ | Single Top | $Z \to \ell\ell + \gamma$/jets | $Z \to \tau\tau + \gamma$/jets | $W$+jets | QCD | Total Bkg. | Observed | Data/MC |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| blinding | 1758.10 ± 2.40 | 23493.91 ± 20.96 | 14991.35 ± 32.27 | 120768.97 ± 40.90 | 11596.56 ± 10.31 | 16502082.15 ± 8146.00 | 107699.32 ± 138.07 | 25537.99 ± 169.65 | 19771.20 ± 15.52 | 16825941.46 ± 8149.15 | 17242569 | 1.02 ± 0.00 |
| lepton $p_{\mathrm{T}}$ | 1689.98 ± 2.37 | 23209.82 ± 20.84 | 13966.81 ± 30.58 | 119772.37 ± 40.74 | 11509.18 ± 10.24 | 16160568.40 ± 8102.70 | 93841.55 ± 128.97 | 24771.30 ± 157.11 | 11445.91 ± 11.75 | 16459085.34 ± 8105.45 | 16713082 | 1.02 ± 0.00 |
| Scale factors | | | | NF = 1.05 | NF = 1.05 | | | | | | | |
| OS leptons | 1633.72 ± 2.20 | 23141.36 ± 20.81 | 9851.72 ± 23.44 | 125239.61 ± 42.66 | 11944.51 ± 10.17 | 16127141.47 ± 8094.87 | 93632.69 ± 128.83 | 22541.61 ± 151.84 | 9130.17 ± 10.47 | 16422623.15 ± 8097.50 | 16671419 | 1.02 ± 0.00 |
| $m_{\ell\ell} > 10, 12$ GeV | 1603.85 ± 2.19 | 23065.53 ± 20.78 | 9558.73 ± 22.70 | 124911.38 ± 42.61 | 11917.69 ± 10.15 | 16097715.51 ± 8094.57 | 93542.25 ± 128.78 | 25895.49 ± 147.96 | 4605.50 ± 7.95 | 16391212.06 ± 8097.13 | 16530227 | 1.01 ± 0.00 |
| $Z$ veto (SF) | 1546.81 ± 1.96 | 20539.87 ± 19.59 | 4664.05 ± 18.63 | 111750.67 ± 40.30 | 10656.14 ± 9.57 | 1784380.63 ± 1975.58 | 91739.91 ± 127.50 | 11596.31 ± 79.33 | 4370.30 ± 7.70 | 2039697.89 ± 1981.92 | 2150203 | 1.05 ± 0.00 |
| Scale factors | | | | NF = 1.05 | NF = 1.05 | NFs applied | NFs applied | | | | | |
| $E_{\mathrm{T,[rel],(clj)}}^{\mathrm{miss,(track)}} > [40], (20)$ | 879.39 ± 1.42 | 13197.71 ± 15.70 | 2102.45 ± 13.73 | 56392.92 ± 28.63 | 6391.24 ± 7.40 | 40377.76 ± 258.86 | 13079.52 ± 46.86 | 3167.74 ± 23.75 | 645.02 ± 3.62 | 135354.36 ± 266.63 | 139935 | 1.03 ± 0.00 |
| $Z$ VR (incl) | 837.38 ± 1.71 | 13848.21 ± 16.06 | 6911.70 ± 19.01 | 75171.83 ± 33.08 | 7172.48 ± 7.89 | 14313907.38 ± 7849.86 | 46408.15 ± 90.56 | 18025.41 ± 128.54 | 2411.61 ± 6.41 | 14483856.78 ± 7851.55 | 14515758 | 1.00 ± 0.00 |
| Top VR (incl) | 97.40 ± 0.64 | 882.37 ± 4.19 | 220.75 ± 4.25 | 64043.28 ± 30.61 | 5295.92 ± 6.57 | 2834.17 ± 67.03 | 1238.99 ± 14.89 | 732.09 ± 12.46 | 128.30 ± 1.40 | 75375.88 ± 76.73 | 75898 | 1.01 ± 0.00 |
| Scale factors | | NF = 1.22 | | NF = 1.08 | NF = 1.08 | NFs applied | | | | | | |
| 0j: jet veto | 494.16 ± 0.93 | 10367.68 ± 15.40 | 1097.81 ± 9.84 | 1238.53 ± 4.31 | 618.03 ± 2.45 | 31174.10 ± 230.68 | 6251.58 ± 33.42 | 1838.74 ± 16.83 | 266.04 ± 2.74 | 52852.52 ± 234.48 | 54463 | 1.03 ± 0.01 |
| 0j: $\Delta\phi_{\ell\ell,MET} > 1.57$ | 493.57 ± 0.93 | 10359.01 ± 15.39 | 1092.36 ± 9.83 | 1227.97 ± 4.29 | 615.70 ± 2.45 | 28632.22 ± 218.23 | 6153.15 ± 33.19 | 1825.84 ± 16.56 | 256.09 ± 2.72 | 50162.36 ± 222.18 | 51784 | 1.03 ± 0.01 |
| 0j: $p_{\mathrm{T},\ell\ell}$ >30 GeV | 434.03 ± 0.87 | 8697.42 ± 14.12 | 880.68 ± 8.83 | 1124.82 ± 4.11 | 564.48 ± 2.30 | 6763.39 ± 98.73 | 804.34 ± 11.85 | 1450.30 ± 11.01 | 30.46 ± 1.85 | 20315.90 ± 101.55 | 20999 | 1.03 ± 0.01 |
| 0j: $m_{\ell\ell} < 55$ GeV | 379.31 ± 0.76 | 2926.00 ± 8.17 | 531.98 ± 7.23 | 249.94 ± 1.93 | 143.06 ± 1.14 | 4870.59 ± 31.28 | 358.91 ± 7.89 | 677.10 ± 8.11 | 14.37 ± 1.73 | 9771.96 ± 35.12 | 10197 | 1.04 ± 0.01 |
| 0j: $E_{\mathrm{T,rel}}^{\mathrm{miss,track}} > 40$ GeV (SF) | 353.35 ± 0.74 | 2766.75 ± 7.95 | 459.63 ± 6.72 | 239.82 ± 1.89 | 137.79 ± 1.11 | 687.62 ± 15.32 | 350.49 ± 7.81 | 560.02 ± 7.00 | 12.92 ± 1.73 | 5215.03 ± 21.47 | 5608 | 1.08 ± 0.02 |
| 0j: $\Delta\phi_{\ell\ell} < 1.8$ | 325.74 ± 0.71 | 2570.18 ± 7.67 | 428.42 ± 6.51 | 228.44 ± 1.84 | 132.22 ± 1.06 | 667.62 ± 14.79 | 12.60 ± 1.61 | 399.88 ± 5.93 | 9.63 ± 1.65 | 4449.00 ± 19.10 | 4769 | 1.07 ± 0.02 |
| 0j: $f_{\mathrm{recoil}} < 0.1$ (SF) | 283.40 ± 0.66 | 2288.37 ± 7.24 | 393.53 ± 6.30 | 173.01 ± 1.60 | 105.86 ± 0.95 | 109.32 ± 5.57 | 12.43 ± 1.60 | 356.15 ± 5.59 | 9.20 ± 1.65 | 3447.88 ± 12.77 | 3750 | 1.09 ± 0.02 |
| 0j: low $p_{\mathrm{T}}$ | 56.10 ± 0.28 | 257.18 ± 2.41 | 167.96 ± 4.27 | 26.26 ± 0.62 | 14.25 ± 0.35 | 30.97 ± 2.87 | 6.11 ± 1.10 | 196.89 ± 4.47 | 5.53 ± 1.57 | 705.16 ± 7.51 | 773 | 1.10 ± 0.04 |
| 0j: high $p_{\mathrm{T}}$ | 227.30 ± 0.60 | 2031.19 ± 6.83 | 225.58 ± 4.63 | 146.74 ± 1.48 | 91.61 ± 0.89 | 78.35 ± 4.77 | 6.32 ± 1.16 | 159.26 ± 3.36 | 3.67 ± 0.51 | 2742.72 ± 10.33 | 2977 | 1.09 ± 0.02 |
| 0j: SSCR | 2.35 ± 0.17 | 3.41 ± 0.25 | 382.93 ± 6.32 | 0.78 ± 0.10 | 0.79 ± 0.13 | 18.94 ± 1.90 | 2.74 ± 0.82 | 215.27 ± 4.66 | 5.63 ± 0.67 | 630.49 ± 8.15 | 622 | 0.99 ± 0.04 |
| 0j: $Z$ VR | 407.03 ± 0.89 | 8841.30 ± 12.87 | 3423.67 ± 12.96 | 1041.78 ± 3.89 | 524.68 ± 2.30 | 11046360.82 ± 6871.64 | 34516.62 ± 80.21 | 12400.06 ± 104.84 | 1339.63 ± 5.02 | 11108448.55 ± 6872.94 | 11073081 | 1.00 ± 0.00 |
| Scale factors | | NF = 1.22 | | NF = 1.08 | NF = 1.08 | NFs applied | | | | | | |
| 0j: $WW$ CR | 35.71 ± 0.32 | 2539.40 ± 7.64 | 139.78 ± 3.15 | 284.82 ± 2.07 | 158.95 ± 1.28 | 127.74 ± 20.92 | 106.34 ± 4.22 | 216.97 ± 3.53 | 2.05 ± 0.48 | 3576.04 ± 23.29 | 3577 | 1.00 ± 0.02 |
| 0j: $WW$ VR | 1.14 ± 0.15 | 1822.93 ± 6.49 | 90.45 ± 2.44 | 318.38 ± 2.19 | 144.48 ± 1.05 | 71.18 ± 19.15 | 0.73 ± 0.37 | 76.34 ± 1.73 | 1.01 ± 0.12 | 2525.50 ± 20.58 | 2382 | 0.94 ± 0.02 |
| Scale factors | | NF = 1.05 | | NF = 1.06 | NF = 1.06 | NFs applied | NFs applied | | | | | |
| 1j: one jet | 268.76 ± 0.84 | 3860.98 ± 8.68 | 687.75 ± 7.41 | 12186.05 ± 13.38 | 3313.75 ± 5.37 | 8165.51 ± 110.03 | 5910.53 ± 29.15 | 840.07 ± 13.75 | 347.02 ± 2.25 | 35311.67 ± 116.14 | 35951 | 1.02 ± 0.01 |
| 1j: $b$-jet veto | 231.85 ± 0.75 | 3375.62 ± 8.06 | 586.07 ± 6.86 | 2338.93 ± 5.82 | 799.56 ± 2.90 | 6700.36 ± 97.62 | 5150.12 ± 27.00 | 672.58 ± 12.06 | 278.40 ± 2.02 | 19901.64 ± 102.78 | 20756 | 1.04 ± 0.01 |
| 1j: Max lepton-MT (DF) | 206.84 ± 0.71 | 3231.96 ± 7.89 | 529.09 ± 6.49 | 2267.24 ± 5.73 | 775.32 ± 2.84 | 7264.96 ± 105.84 | 2222.68 ± 18.08 | 614.09 ± 9.22 | 72.58 ± 1.03 | 16977.92 ± 108.45 | 17265 | 1.02 ± 0.01 |
| 1j: $Z \to \tau\tau$ veto (DF) | 185.81 ± 0.66 | 2636.59 ± 7.13 | 438.71 ± 5.93 | 1831.37 ± 5.15 | 635.16 ± 2.59 | 7243.26 ± 105.39 | 928.32 ± 12.00 | 448.57 ± 7.99 | 41.94 ± 0.84 | 14203.92 ± 106.94 | 14471 | 1.02 ± 0.01 |
| 1j: $m_{\ell\ell} < 55$ GeV | 157.15 ± 0.57 | 836.69 ± 4.01 | 218.08 ± 4.46 | 523.46 ± 2.75 | 196.05 ± 1.54 | 3425.99 ± 21.44 | 548.58 ± 9.11 | 201.95 ± 5.89 | 26.40 ± 0.65 | 5977.20 ± 24.97 | 6783 | 1.13 ± 0.01 |
| 1j: $E_{\mathrm{T,rel}}^{\mathrm{miss,track}} > 35$ GeV (SF) | 142.10 ± 0.55 | 777.99 ± 3.87 | 188.51 ± 4.18 | 490.48 ± 2.66 | 184.14 ± 1.42 | 200.77 ± 8.03 | 383.06 ± 7.84 | 166.30 ± 4.47 | 18.81 ± 0.59 | 2410.06 ± 13.70 | 2616 | 1.09 ± 0.02 |
| 1j: $\Delta\phi_{\ell\ell} < 1.8$ | 125.79 ± 0.52 | 682.44 ± 3.63 | 163.67 ± 3.88 | 448.63 ± 2.55 | 170.70 ± 1.32 | 199.04 ± 8.38 | 23.77 ± 1.99 | 118.45 ± 3.39 | 6.23 ± 0.39 | 1812.93 ± 11.06 | 2018 | 1.11 ± 0.03 |
| 1j: $f_{recoil} < 0.10$ (SF) | 110.69 ± 0.48 | 605.34 ± 3.42 | 148.73 ± 3.75 | 367.51 ± 2.30 | 146.37 ± 1.21 | 31.42 ± 3.11 | 23.12 ± 1.96 | 105.54 ± 3.17 | 6.09 ± 0.39 | 1434.12 ± 7.50 | 1596 | 1.11 ± 0.03 |
| 1j: low $p_{\mathrm{T}}$ | 18.28 ± 0.19 | 69.88 ± 1.15 | 38.36 ± 1.99 | 47.58 ± 0.82 | 17.94 ± 0.48 | 9.44 ± 1.36 | 6.98 ± 1.10 | 45.02 ± 2.23 | 3.08 ± 0.27 | 238.26 ± 3.77 | 260 | 1.09 ± 0.07 |
| 1j: high $p_{\mathrm{T}}$ | 92.41 ± 0.44 | 535.45 ± 3.22 | 110.37 ± 3.18 | 319.94 ± 2.16 | 128.43 ± 1.12 | 21.98 ± 2.80 | 16.15 ± 1.62 | 60.51 ± 2.26 | 3.01 ± 0.28 | 1195.85 ± 6.48 | 1336 | 1.12 ± 0.03 |
| 1j: SSCR | 2.19 ± 0.19 | 1.22 ± 0.15 | 139.73 ± 3.73 | 1.71 ± 0.15 | 2.11 ± 0.50 | 4.73 ± 0.91 | 0.85 ± 0.37 | 75.95 ± 2.88 | 2.91 ± 0.25 | 229.21 ± 4.86 | 228 | 0.99 ± 0.07 |
| 1j: $Z$ VR | 243.14 ± 0.95 | 3615.53 ± 8.21 | 2186.24 ± 10.84 | 10805.46 ± 12.54 | 3030.53 ± 5.09 | 2410699.80 ± 3256.31 | 8487.04 ± 33.85 | 3699.34 ± 61.83 | 737.47 ± 3.31 | 2443261.41 ± 3257.13 | 2504811 | 1.03 ± 0.00 |
| Scale factors | | NF = 1.05 | | NF = 1.06 | NF = 1.06 | NFs applied | NFs applied | | | | | |
| 1j: $WW$ control region | 5.60 ± 0.23 | 1414.12 ± 5.23 | 156.77 ± 3.35 | 1079.49 ± 3.96 | 345.42 ± 1.73 | 69.47 ± 19.40 | 81.32 ± 3.52 | 166.43 ± 3.18 | 12.84 ± 0.43 | 3325.85 ± 21.36 | 3298 | 0.99 ± 0.02 |
| 1j: $WW$ control region (SF $f_{\mathrm{recoil}}$) | 5.42 ± 0.22 | 1330.77 ± 5.07 | 146.26 ± 3.25 | 964.95 ± 3.75 | 315.27 ± 1.65 | 25.50 ± 12.13 | 81.32 ± 3.52 | 161.52 ± 3.13 | 12.83 ± 0.43 | 3038.42 ± 14.91 | 3018 | 0.99 ± 0.02 |
| 1j: Top CR | 22.06 ± 0.29 | 324.68 ± 2.58 | 63.44 ± 2.23 | 6395.78 ± 9.52 | 1950.15 ± 3.86 | 52.89 ± 10.20 | 204.20 ± 5.62 | 106.73 ± 4.34 | 12.32 ± 0.40 | 9110.19 ± 16.48 | 9180 | 1.01 ± 0.01 |

Table 27: Event selection for the combined lepton flavour channels of the 0-jets and the 1-jet analysis with relevant control and validation regions. The uncertainties include the statistical errors and the pre-fit systematics. The normalization factors derived from the background control regions are applied to the Monte Carlo simulation and the W+jets and the QCD background are given by the data-driven estimates.

| Signal Region Selection | | | Summary | | | Background Composition | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Data | Bkgd | Signal 125 | $WW$ | $t\bar{t}$ | st | diboson | $Z \to \ell\ell$ | $Z \to \tau\tau$ | W+jets | QCD |
| 0j: $e\mu$ | $m_{\ell\ell} < 30$ GeV | $10 < p_T^{sub} < 15$ GeV | 209 | 171 | 15 | 54 | 5 | 2.3 | 59 | 2.8 | 2 | 42.5 | 4.3 |
| | | $15 < p_T^{sub} < 20$ GeV | 177 | 139 | 20 | 80 | 5.9 | 3.4 | 35 | 2.2 | 0.4 | 12.1 | 0.8 |
| | | $p_T^{sub} > 20$ GeV | 214 | 191 | 25 | 148 | 13 | 7.7 | 16 | 1 | 0.1 | 5.7 | 0.2 |
| | $m_{\ell\ell} > 30$ GeV | $10 < p_T^{sub} < 15$ GeV | 137 | 117 | 11 | 54 | 7 | 3.7 | 10 | 0.7 | 1.7 | 39.4 | 0.8 |
| | | $15 < p_T^{sub} < 20$ GeV | 190 | 168 | 18 | 116 | 13 | 6.8 | 14 | 0.6 | 1 | 16.7 | 0.8 |
| | | $p_T^{sub} > 20$ GeV | 503 | 464 | 31 | 372 | 31 | 18 | 21 | 1 | 2.4 | 18.7 | 0.4 |
| 0j:$\mu e$ | $m_{\ell\ell} < 30$ GeV | $10 < p_T^{sub} < 15$ GeV | 161 | 152 | 10 | 39 | 3.3 | 1.6 | 63 | 4.3 | 0.5 | 40 | 0.6 |
| | | $15 < p_T^{sub} < 20$ GeV | 153 | 127 | 14 | 62 | 4.6 | 2.5 | 41 | 2.3 | 0.9 | 13 | 0.8 |
| | | $p_T^{sub} > 20$ GeV | 218 | 173 | 20 | 132 | 11 | 6.4 | 15 | 1.1 | 0.2 | 7.5 | 0.3 |
| | $m_{\ell\ell} > 30$ GeV | $10 < p_T^{sub} < 15$ GeV | 107 | 107 | 6.9 | 39 | 4.7 | 2.3 | 17 | 1.2 | 1.9 | 41 | 0 |
| | | $15 < p_T^{sub} < 20$ GeV | 159 | 134 | 13 | 85 | 8.6 | 4.6 | 14 | 0.6 | 0.6 | 21 | 0.6 |
| | | $p_T^{sub} > 20$ GeV | 414 | 408 | 26 | 325 | 26 | 16 | 19 | 0.9 | 0.7 | 20 | 0.2 |
| 0j: $ee + \mu\mu$ | $f_{recoil} < 0.1$ | | 1108 | 1096 | 75 | 786 | 41 | 31 | 69 | 91 | 0 | 79 | 0 |
| 1j: $e\mu$ | $m_{\ell\ell} < 30$ GeV | $10 < p_T^{sub} < 15$ GeV | 45 | 48 | 4.9 | 12 | 8.1 | 3.1 | 11 | 0.4 | 1.8 | 9.4 | 1.7 |
| | | $15 < p_T^{sub} < 20$ GeV | 67 | 51 | 6.1 | 18 | 12 | 4.4 | 10 | 0.3 | 0.2 | 5.5 | 0.5 |
| | | $p_T^{sub} > 20$ GeV | 115 | 100 | 12 | 43 | 28 | 11 | 13 | 0.2 | 0.1 | 4.1 | 0.2 |
| | $m_{\ell\ell} > 30$ GeV | $10 < p_T^{sub} < 15$ GeV | 71 | 51 | 3.5 | 15 | 12 | 4.5 | 4.1 | 0.2 | 3.1 | 12 | 1.1 |
| | | $15 < p_T^{sub} < 20$ GeV | 73 | 76 | 6.1 | 30 | 22 | 7.9 | 5.7 | 0.2 | 3.5 | 6.3 | 0.5 |
| | | $p_T^{sub} > 20$ GeV | 250 | 233 | 16 | 109 | 67 | 27 | 14 | 0.6 | 5.5 | 11 | 0.5 |
| 1j: $\mu e$ | $m_{\ell\ell} < 30$ GeV | $10 < p_T^{sub} < 15$ GeV | 37 | 37 | 3.1 | 8.8 | 5.3 | 1.7 | 12 | 0.7 | 0.6 | 7.6 | 0.2 |
| | | $15 < p_T^{sub} < 20$ GeV | 47 | 39 | 5 | 14 | 9.3 | 3.2 | 8.3 | 0.4 | 0 | 3.6 | 0.4 |
| | | $p_T^{sub} > 20$ GeV | 89 | 93 | 11 | 40 | 23 | 9.4 | 16 | 1 | 0.1 | 3.9 | 0.2 |
| | $m_{\ell\ell} > 30$ GeV | $10 < p_T^{sub} < 15$ GeV | 43 | 38 | 2.4 | 11 | 8.2 | 2.8 | 4.8 | 0.2 | 1.4 | 10 | 0.1 |
| | | $15 < p_T^{sub} < 20$ GeV | 58 | 57 | 4.2 | 22 | 15 | 5.3 | 5.5 | 0.3 | 1.8 | 6.4 | 0.6 |
| | | $p_T^{sub} > 20$ GeV | 234 | 208 | 14 | 95 | 60 | 22 | 15 | 0.6 | 4.3 | 9.3 | 0.3 |
| 1j: $ee + \mu\mu$ | $f_{recoil} < 0.1$ | | 467 | 404 | 23 | 188 | 98 | 44 | 29 | 26 | 0.9 | 17 | 0 |

Table 28: Pre-fit signal region event selection of the 0-jets and the 1-jet analysis serving as input for the statistical fit. Background normalization factors derived from control regions have already been applied to the respective Monte Carlo simulation. The W+jets and the QCD background are given by the data-driven estimate.

# B

# Transverse Mass Remapping

| 0j: Signal Region Selection | | | Bin Boundaries on $m_T$ [GeV] | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $e\mu$ | $m_{\ell\ell} < 30$ | $10 < p_T^{sub} < 15$ | 74.45 | 80.65 | 85.55 | 90.05 | 94.55 | 98.95 | 103.95 | 109.65 | 117.95 |
| | | $15 < p_T^{sub} < 20$ | 81.55 | 87.85 | 92.55 | 96.65 | 100.65 | 104.75 | 109.25 | 114.55 | 122.15 |
| | | $p_T^{sub} > 20$ | 93.65 | 99.85 | 104.45 | 108.45 | 112.25 | 116.05 | 120.35 | 125.45 | 133.25 |
| | $m_{\ell\ell} > 30$ | $10 < p_T^{sub} < 15$ | 84.15 | 90.55 | 95.15 | 99.45 | 103.45 | 107.85 | 112.15 | 117.25 | 124.75 |
| | | $15 < p_T^{sub} < 20$ | 86.35 | 92.35 | 97.05 | 101.45 | 105.45 | 109.45 | 113.65 | 118.75 | 125.85 |
| | | $p_T^{sub} > 20$ | 93.25 | 100.25 | 105.05 | 109.25 | 113.05 | 116.95 | 121.15 | 126.45 | 135.25 |
| $\mu e$ | $m_{\ell\ell} < 30$ | $10 < p_T^{sub} < 15$ | 76.75 | 82.55 | 87.05 | 91.25 | 95.25 | 99.75 | 104.45 | 110.15 | 118.05 |
| | | $15 < p_T^{sub} < 20$ | 80.85 | 86.75 | 91.55 | 95.95 | 99.85 | 103.95 | 108.55 | 113.85 | 121.55 |
| | | $p_T^{sub} > 20$ | 93.15 | 99.85 | 104.75 | 108.85 | 112.65 | 116.55 | 120.75 | 125.95 | 134.35 |
| | $m_{\ell\ell} > 30$ | $10 < p_T^{sub} < 15$ | 84.95 | 90.95 | 95.65 | 99.85 | 103.75 | 107.75 | 112.05 | 117.35 | 125.15 |
| | | $15 < p_T^{sub} < 20$ | 85.15 | 91.75 | 96.55 | 100.65 | 104.65 | 108.55 | 112.75 | 117.95 | 125.75 |
| | | $p_T^{sub} > 20$ | 93.55 | 100.35 | 105.35 | 109.55 | 113.35 | 117.15 | 121.45 | 126.95 | 135.85 |
| SF | $f_{recoil} < 0.1$ | | 95.15 | 100.15 | 104.05 | 107.55 | 110.85 | 114.25 | 117.85 | 122.15 | 128.75 |

Table 29: Bin boundaries of the $m_T$ remapping scheme for the different signal regions of the 0-jets analysis. The binning maximizes the significance of the Monte Carlo simulated signal events for the mass point $m_H = 125$ GeV, which leads to an equal signal distribution across the full transverse mass range with a bin width of approximately 5 GeV in the essential region.

| 1j: Signal Region Selection | | | Bin Boundaries on $m_T$ [GeV] | | | | |
|---|---|---|---|---|---|---|---|
| $e\mu$ | $m_{\ell\ell} < 30$ GeV | $10 < p_T^{sub} < 15$ GeV | 78.55 | 89.05 | 97.55 | 106.35 | 117.45 |
| | | $15 < p_T^{sub} < 20$ GeV | 81.55 | 92.05 | 101.65 | 110.65 | 120.05 |
| | | $p_T^{sub} > 20$ GeV | 86.65 | 97.85 | 106.85 | 116.05 | 127.85 |
| | $m_{\ell\ell} > 30$ GeV | $10 < p_T^{sub} < 15$ GeV | 88.05 | 98.05 | 106.05 | 113.45 | 123.55 |
| | | $15 < p_T^{sub} < 20$ GeV | 88.15 | 97.95 | 105.85 | 113.75 | 124.15 |
| | | $p_T^{sub} > 20$ GeV | 92.15 | 101.55 | 109.85 | 118.65 | 130.25 |
| $\mu e$ | $m_{\ell\ell} < 30$ GeV | $10 < p_T^{sub} < 15$ GeV | 79.65 | 88.75 | 97.95 | 106.25 | 116.05 |
| | | $15 < p_T^{sub} < 20$ GeV | 82.25 | 92.35 | 101.65 | 110.25 | 120.65 |
| | | $p_T^{sub} > 20$ GeV | 87.45 | 98.55 | 107.55 | 116.85 | 128.05 |
| | $m_{\ell\ell} > 30$ GeV | $10 < p_T^{sub} < 15$ GeV | 87.05 | 96.45 | 105.25 | 112.25 | 121.75 |
| | | $15 < p_T^{sub} < 20$ GeV | 87.55 | 97.05 | 104.95 | 113.55 | 123.65 |
| | | $p_T^{sub} > 20$ GeV | 91.75 | 101.65 | 109.95 | 118.15 | 129.65 |
| SF | $f_{recoil} < 0.1$ | | 96.85 | 105.15 | 111.65 | 118.25 | 127.15 |

Table 30: Bin boundaries of the $m_T$ remapping scheme for the different signal regions of the 1-jet analysis. The binning maximizes the significance of the Monte Carlo simulated signal events for the mass point $m_H = 125$ GeV, which leads to an equal signal distribution across the full transverse mass range with a bin width of approximately 10 GeV in the essential region.

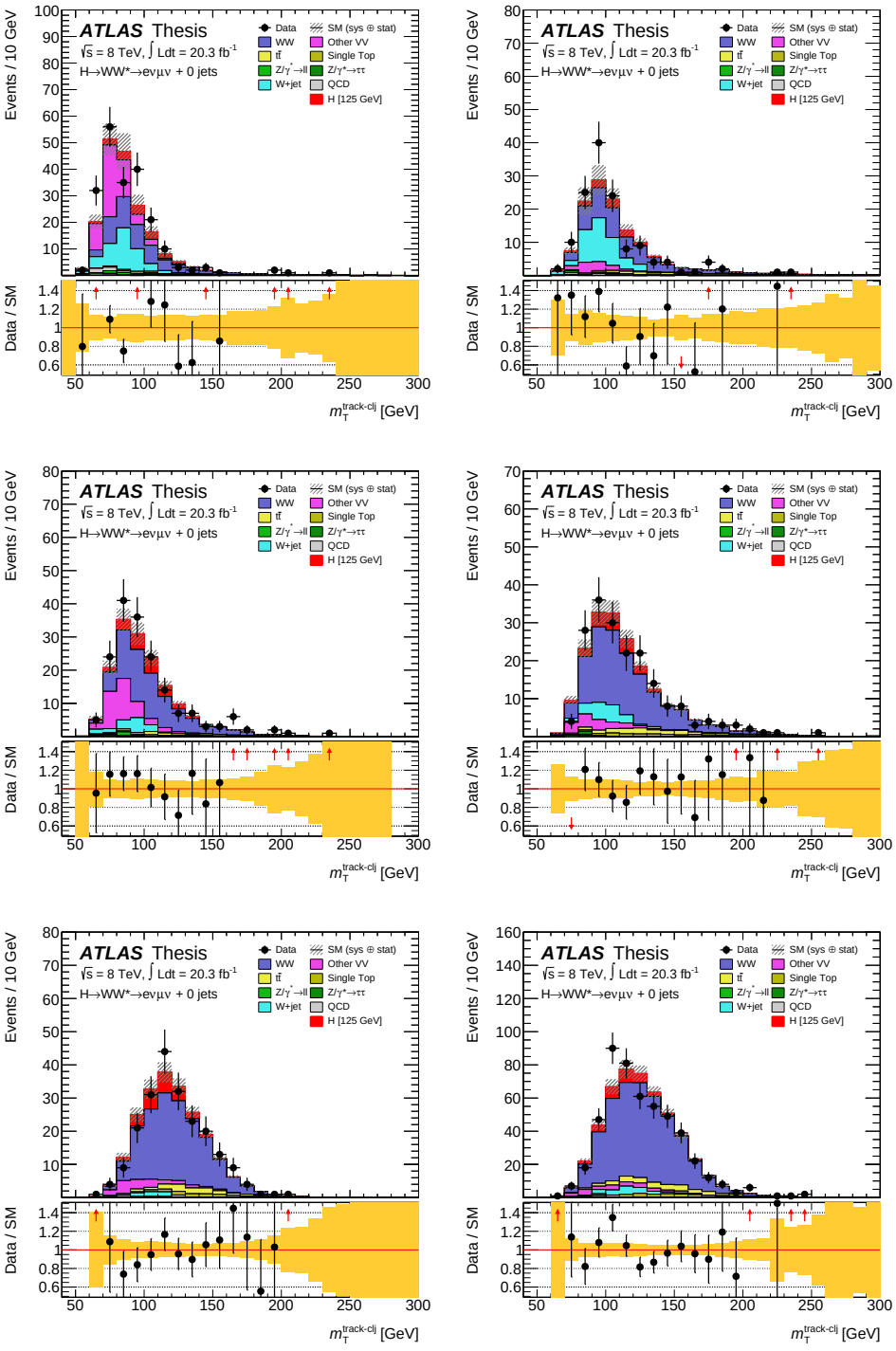



Figure 103: Transverse mass distribution of the $e\mu$ signal regions of the 0-jets analysis. The plots on the left show the $m_{\ell\ell}$ selection below 30 GeV, those on the right above. The $p_T^{sub}$ range increases from $[10, 15]$ GeV on top to $[15, 20]$ GeV in the middle and $[20, \infty]$ GeV on the bottom.

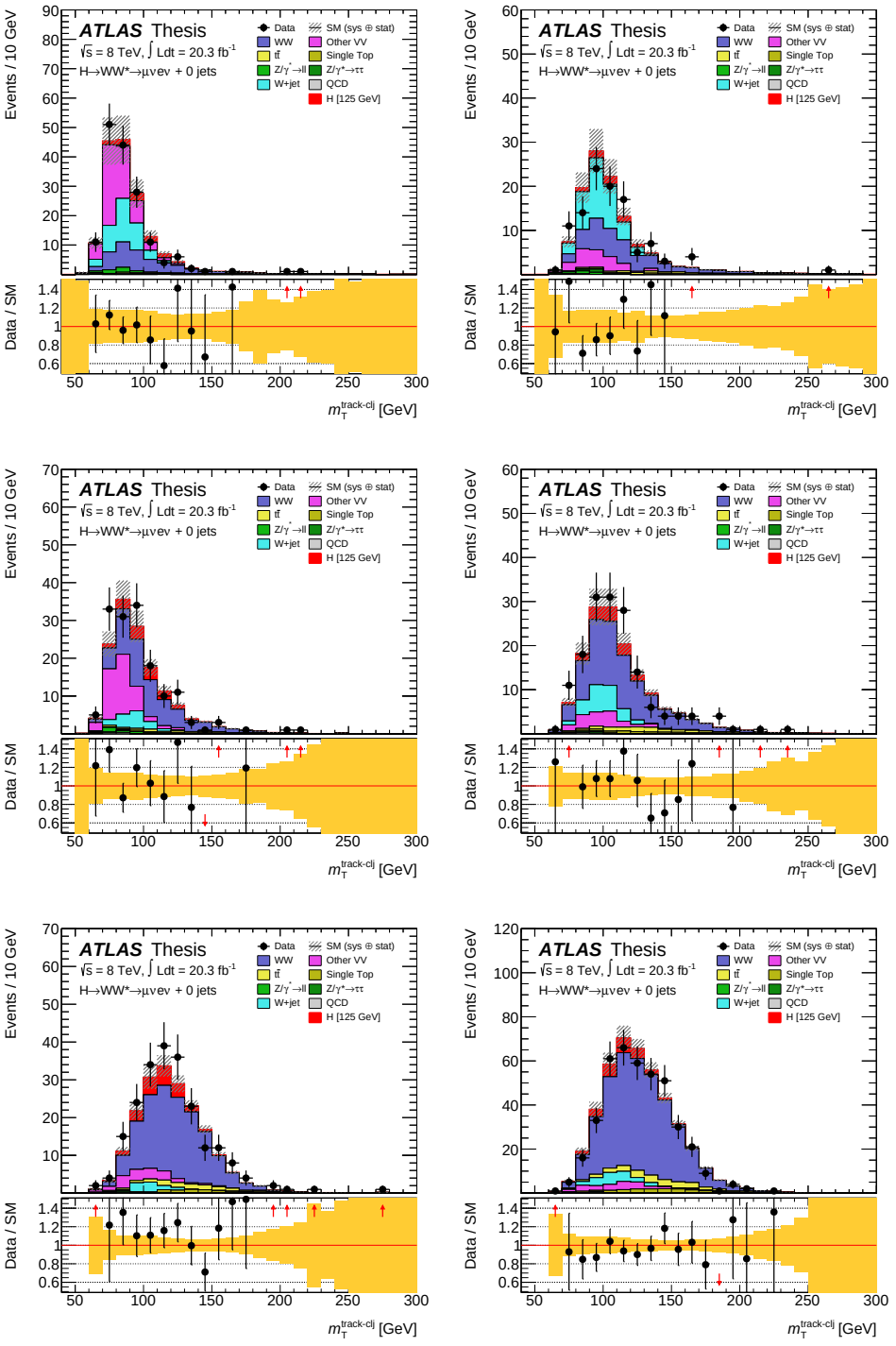



Figure 104: Transverse mass distribution of the $\mu e$ signal regions of the 0-jets analysis. The plots on the left show the $m_{\ell\ell}$ selection below 30 GeV, those on the right above. The $p_T^{sub}$ range increases from $[10, 15]$ GeV on top to $[15, 20]$ GeV in the middle and $[20, \infty]$ GeV on the bottom.

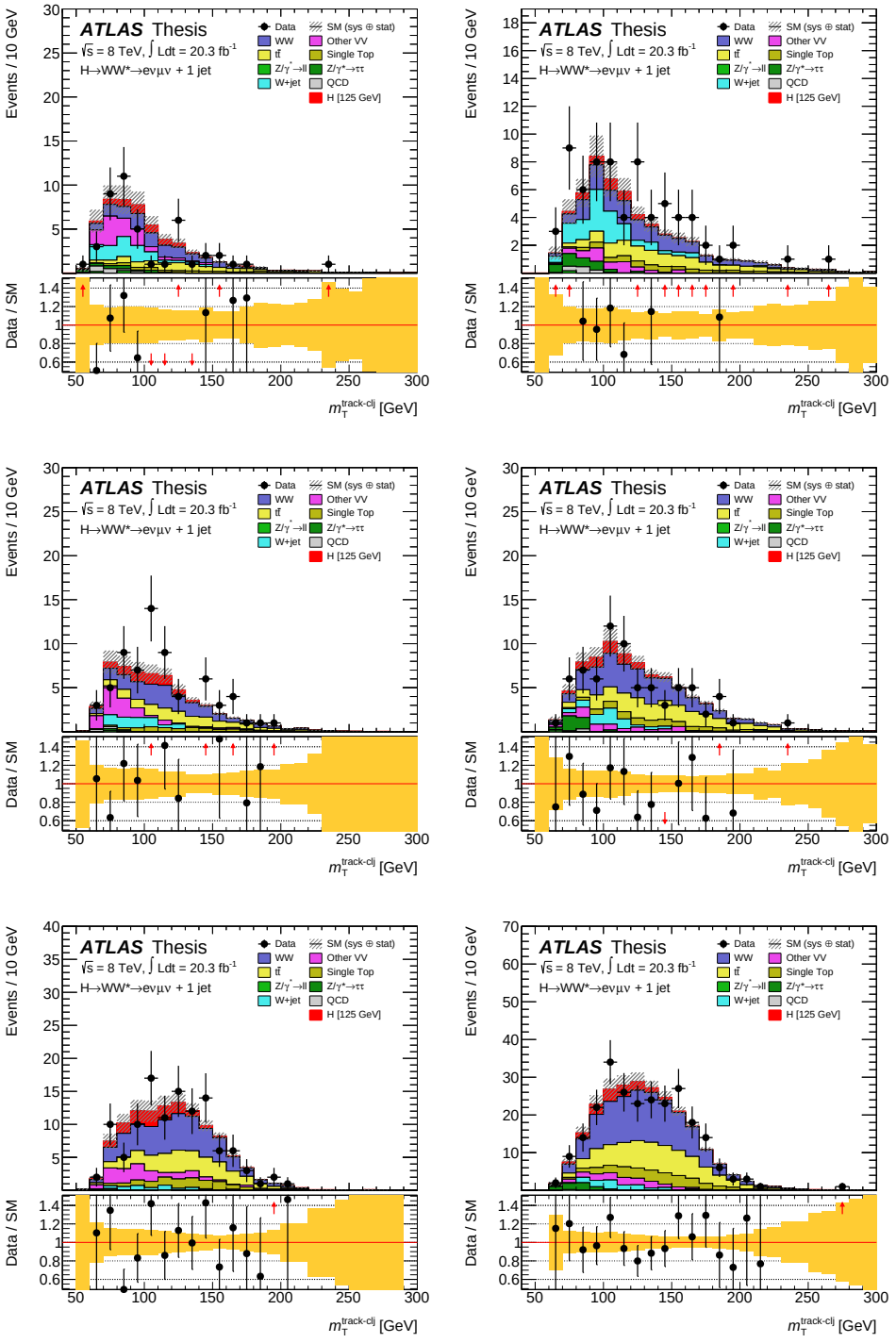



Figure 105: Transverse mass distribution of the $e\mu$ signal regions of the 1-jet analysis. The plots on the left show the $m_{\ell\ell}$ selection below 30 GeV, those on the right above. The $p_T^{sub}$ range increases from $[10, 15]$ GeV on top to $[15, 20]$ GeV in the middle and $[20, \infty]$ GeV on the bottom.

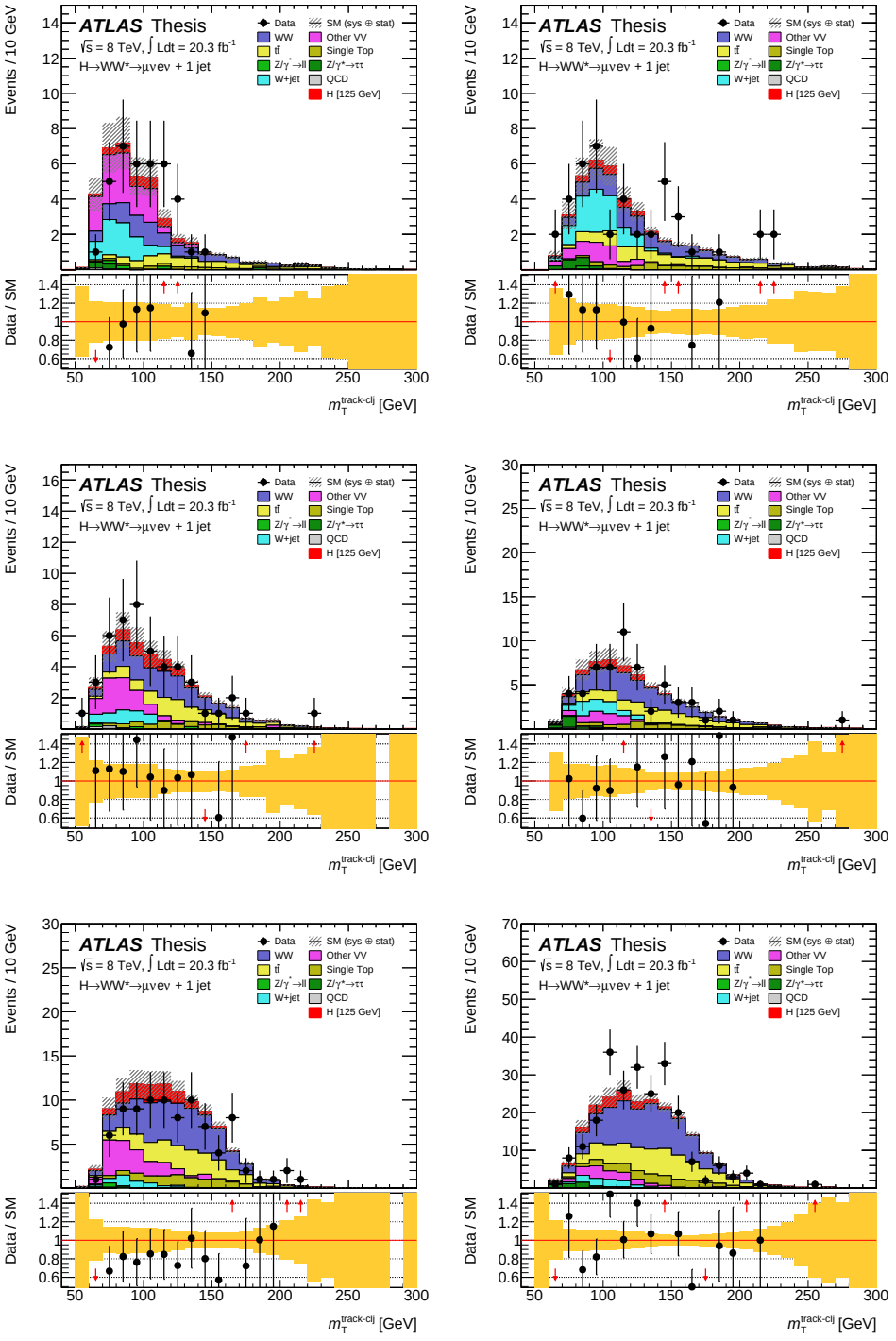



Figure 106: Transverse mass distribution of the $\mu e$ signal regions of the 1-jet analysis. The plots on the left show the $m_{\ell\ell}$ selection below 30 GeV, those on the right above. The $p_T^{sub}$ range increases from $[10, 15]$ GeV on top to $[15, 20]$ GeV in the middle and $[20, \infty]$ GeV on the bottom.

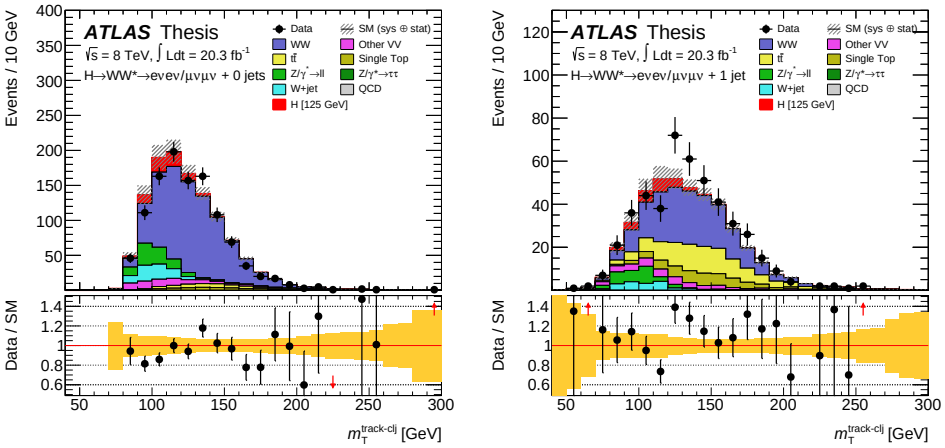



Figure 107: Transverse mass distribution of the same flavour ($ee + \mu\mu$) signal regions. The 0-jets channel is shown on the left and the 1-jet channel on the right.

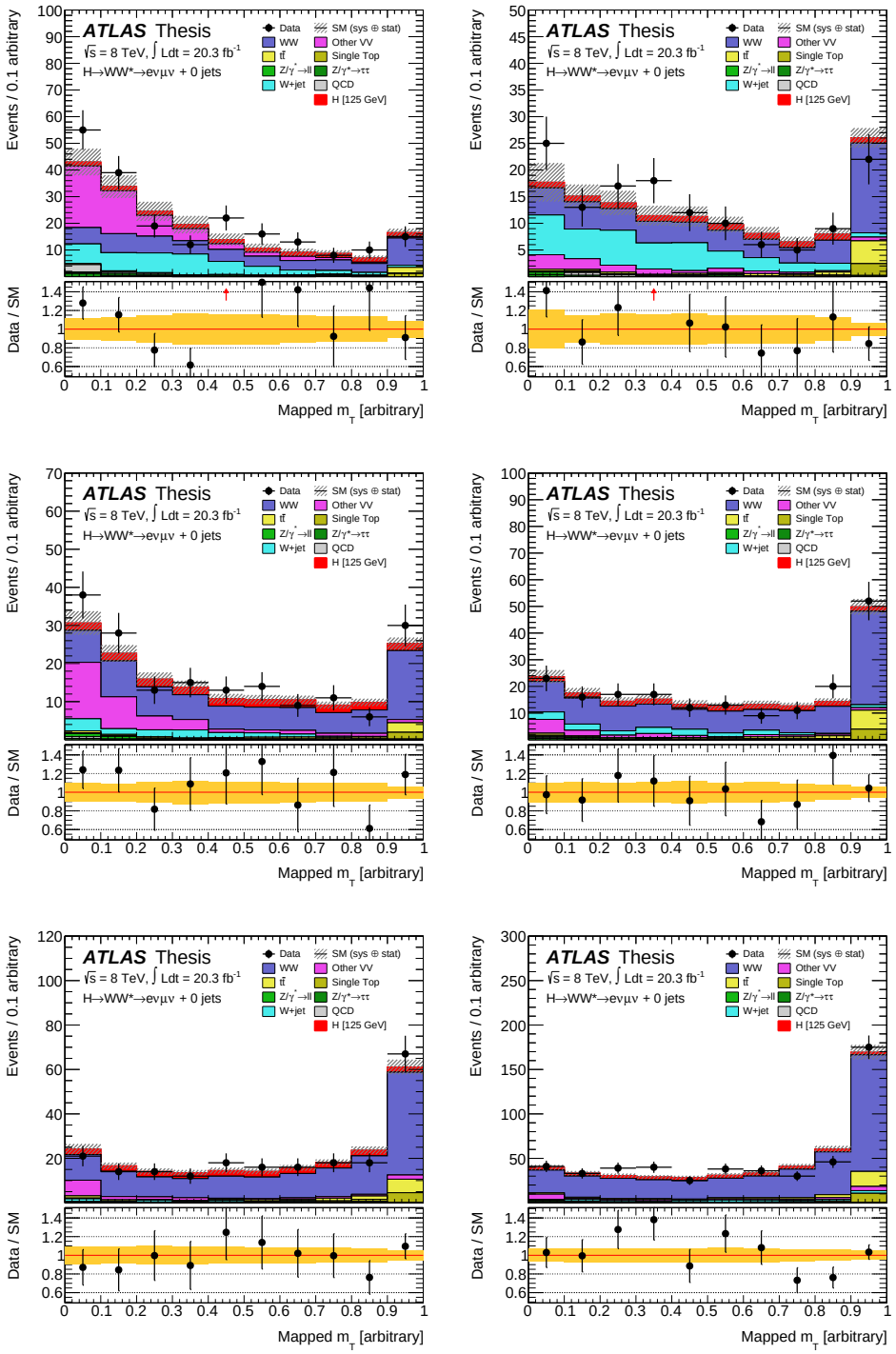



Figure 108: Remapped transverse mass distribution of the $e\mu$ signal regions of the 0-jets analysis. The plots on the left show the $m_{\ell\ell}$ selection below 30 GeV, those on the right above. The $p_T^{sub}$ range increases from $[10, 15]$ GeV on top to $[15, 20]$ GeV in the middle and $[20, \infty]$ GeV on the bottom.

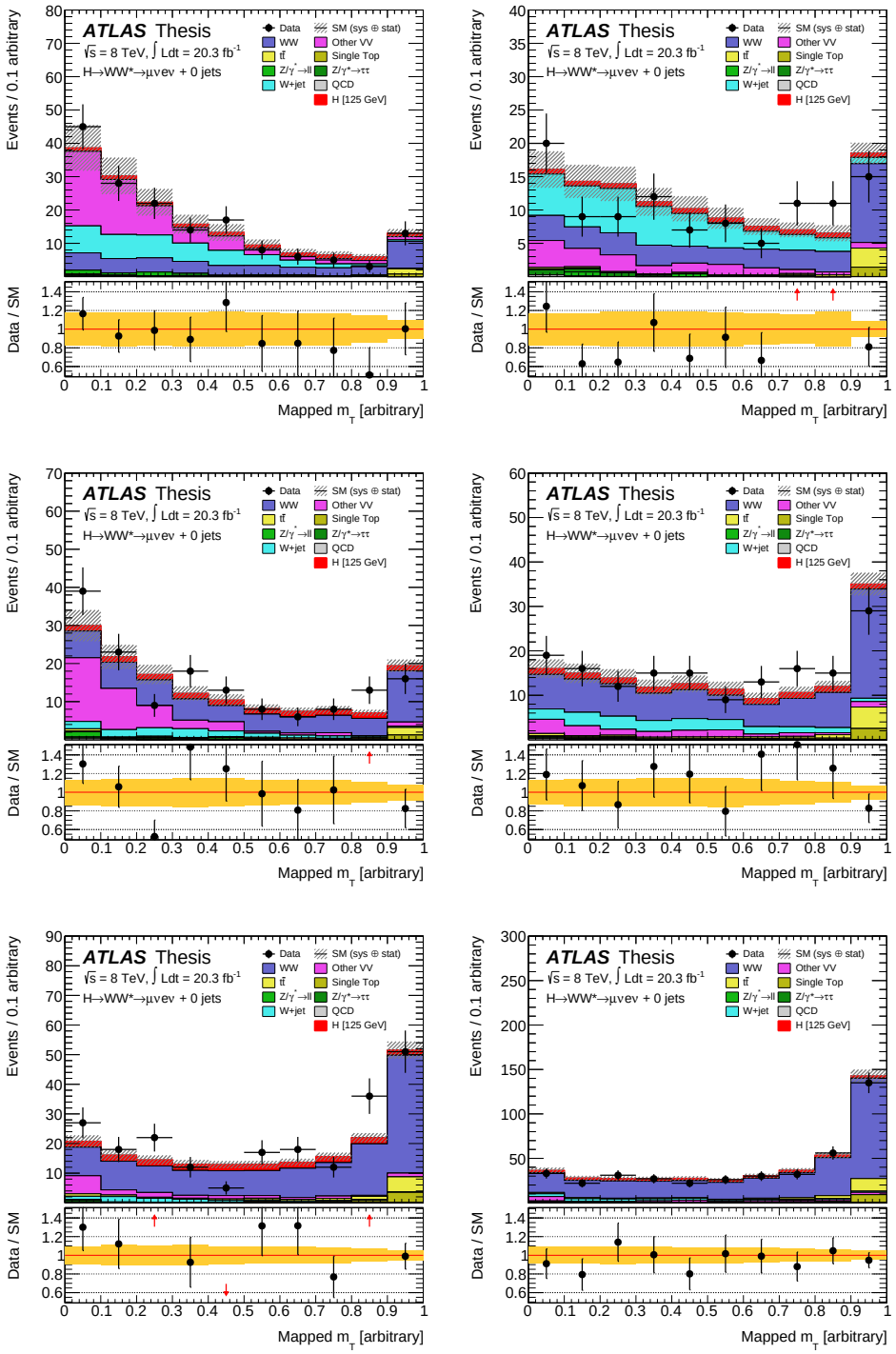



Figure 109: Remapped transverse mass distribution of the $\mu e$ signal regions of the 0-jets analysis. The plots on the left show the $m_{\ell\ell}$ selection below 30 GeV, those on the right above. The $p_T^{sub}$ range increases from $[10, 15]$ GeV on top to $[15, 20]$ GeV in the middle and $[20, \infty]$ GeV on the bottom.

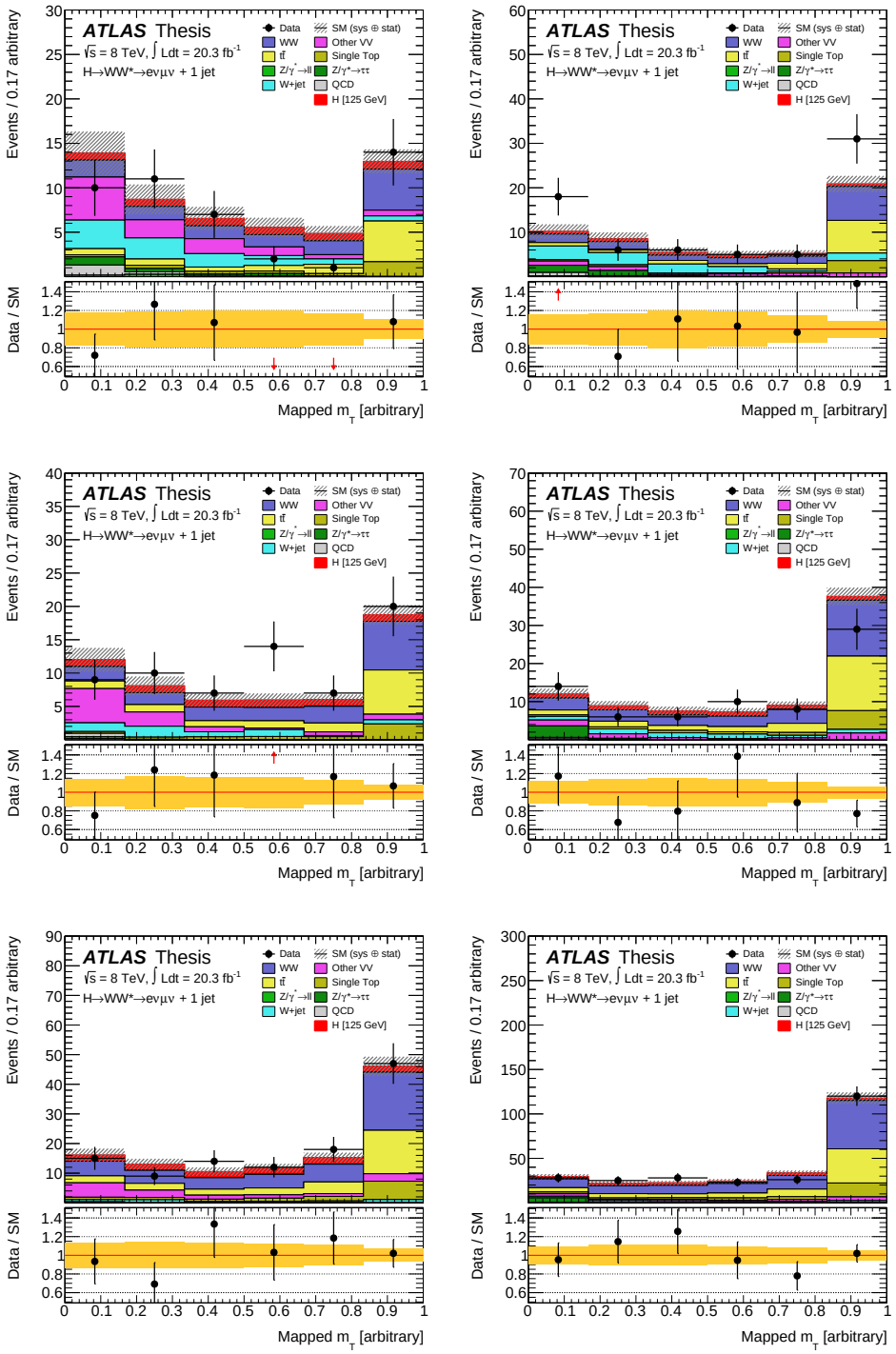



Figure 110: Remapped transverse mass distribution of the $e\mu$ signal regions of the 1-jet analysis. The plots on the left show the $m_{\ell\ell}$ selection below 30 GeV, those on the right above. The $p_T^{sub}$ range increases from $[10, 15]$ GeV on top to $[15, 20]$ GeV in the middle and $[20, \infty]$ GeV on the bottom.

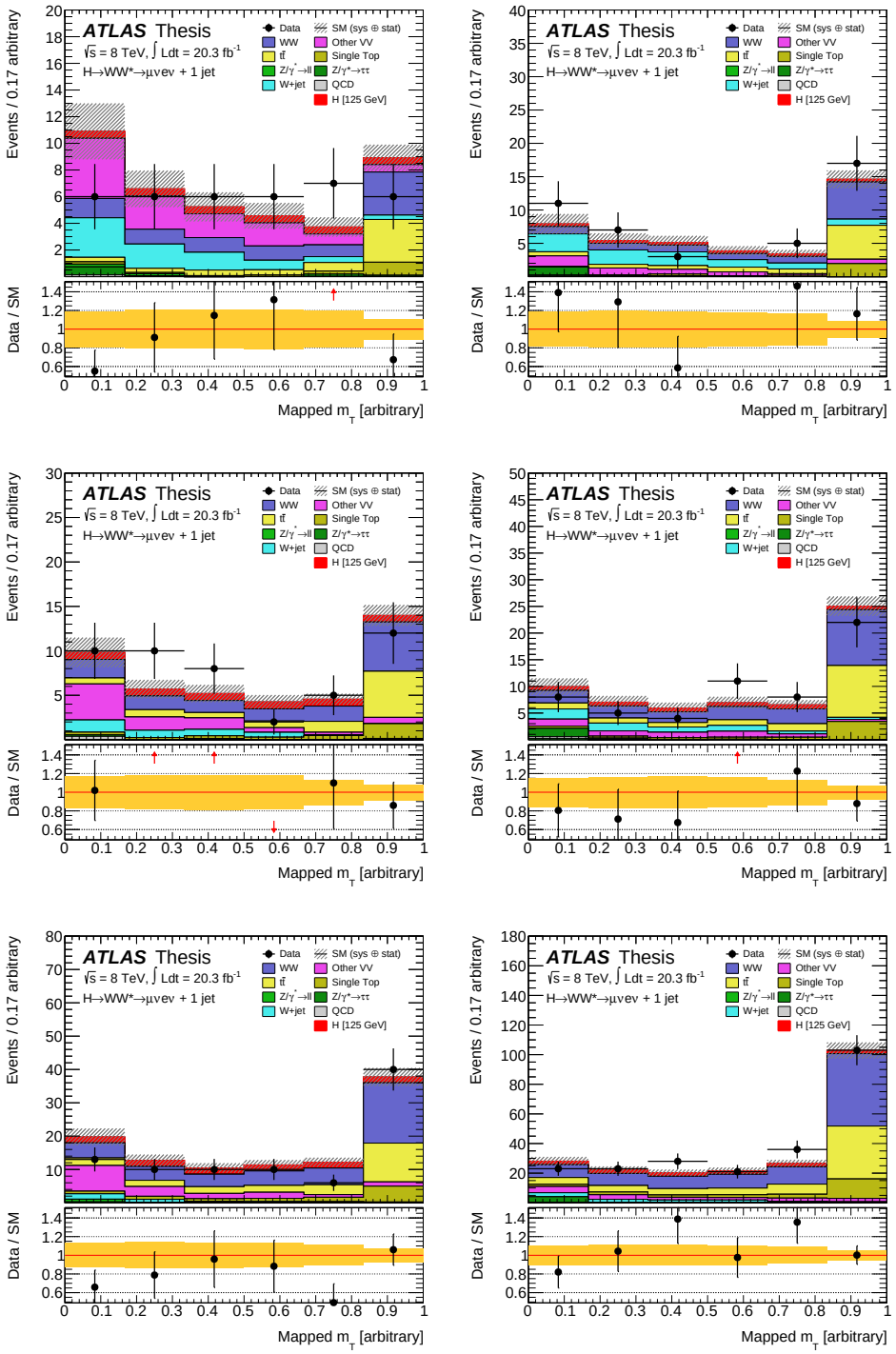



Figure 111: Remapped transverse mass distribution of the $\mu e$ signal regions of the 1-jet analysis. The plots on the left show the $m_{\ell\ell}$ selection below 30 GeV, those on the right above. The $p_T^{sub}$ range increases from $[10, 15]$ GeV on top to $[15, 20]$ GeV in the middle and $[20, \infty]$ GeV on the bottom.

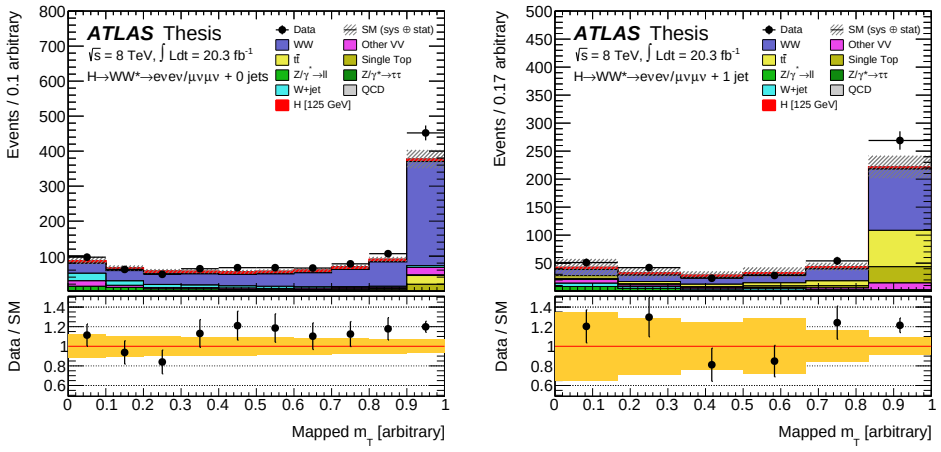



Figure 112: Remapped transverse mass distribution of the same flavour ($ee + \mu\mu$) signal regions. The 0-jets channel is shown on the left and the 1-jet channel on the right.

# C

# Post-Fit Results

Table 31: Fit values of the nuisance parameters symbolizing the systematic uncertainties of the 0-jets and the 1-jet analysis depending on the signal strength parameter of the different fit scenarios: $\mu = 0$ for the background-only hypothesis, $\mu = 1$ for the SM signal hypothesis and $\hat{\mu} = 1.16$ for the observed signal strength.

| Parameter | Prefit | $\mu = 0$ | $\mu = 1$ | $\hat{\mu} = 1.16$ |
|---|---|---|---|---|
| alpha_ATLAS_BR_TAUTAU | 0 | -1.47e-07 | -0.000118 | 9.14e-05 |
| alpha_ATLAS_BR_VV | 0 | -5.94e-07 | 0.0913 | 0.000372 |
| alpha_ATLAS_BTag_B1EFF | 0 | -0.0247 | -0.0238 | -0.0202 |
| alpha_ATLAS_BTag_B2EFF | 0 | 0.00347 | 0.00111 | 0.000593 |
| alpha_ATLAS_BTag_B3EFF | 0 | -0.00282 | 0.207 | 0.236 |
| alpha_ATLAS_BTag_B4EFF | 0 | 0.0848 | 0.0919 | 0.103 |
| alpha_ATLAS_BTag_B5EFF | 0 | -0.101 | -0.0953 | -0.101 |
| alpha_ATLAS_BTag_B6EFF | 0 | 0.405 | 0.437 | 0.406 |
| alpha_ATLAS_BTag_CEFF | 0 | 0.00878 | 0.00699 | 0.00718 |
| alpha_ATLAS_BTag_Herwig_LEFF | 0 | -0.0247 | -0.0185 | -0.018 |
| alpha_ATLAS_BTag_LEFF | 0 | 0.0841 | 0.052 | 0.074 |
| alpha_ATLAS_BTag_Pythia6_LEFF | 0 | 0.0646 | 0.0458 | 0.0558 |
| alpha_ATLAS_BTag_Sherpa_LEFF | 0 | 0.00816 | 0.00734 | 0.00712 |
| alpha_ATLAS_DIL_TRIGGER_HWW | 0 | 0.0146 | 0.00794 | 0.00719 |
| alpha_ATLAS_DPI_XS | 0 | -0.0538 | -0.0249 | -0.0201 |
| alpha_ATLAS_EL_EFF_ID_CORRLOW | 0 | 0.342 | 0.163 | 0.133 |
| alpha_ATLAS_EL_EFF_ID_HIGHPT | 0 | 0.153 | 0.0501 | 0.0118 |
| alpha_ATLAS_EL_EFF_RECOID80010 | 0 | -0.0258 | -0.0749 | -0.0838 |
| alpha_ATLAS_EL_EFF_RECOID80015 | 0 | 0.29 | 0.159 | 0.136 |
| alpha_ATLAS_EL_EFF_RECO_CORR | 0 | 0.106 | -0.0541 | -0.083 |
| alpha_ATLAS_EL_EFF_RECO_CORRLOW | 0 | 0.0319 | 0.0178 | 0.0158 |
| alpha_ATLAS_EL_ESCALE | 0 | -0.124 | -0.00937 | -0.00148 |
| alpha_ATLAS_EL_ISO | 0 | 0.0396 | -0.00146 | -0.0231 |
| alpha_ATLAS_EL_RES | 0 | 0.0589 | 0.0612 | 0.0579 |
| alpha_ATLAS_EL_TRIGGER_HWW | 0 | 0.269 | 0.087 | 0.0558 |
| alpha_ATLAS_Higgs_UEPS | 0 | -2.02e-06 | -5.37e-09 | 0.00129 |
| alpha_ATLAS_JER | 0 | -0.41 | -0.249 | -0.236 |
| alpha_ATLAS_JES_2012_Detector1 | 0 | 0.196 | 0.202 | 0.206 |
| alpha_ATLAS_JES_2012_Eta_StatMethod | 0 | 0.293 | 0.228 | 0.226 |
| alpha_ATLAS_JES_2012_Modelling1 | 0 | 0.0616 | 0.0436 | 0.0382 |
| alpha_ATLAS_JES_2012_PilePt | 0 | 0.0296 | 0.0295 | 0.0276 |
| alpha_ATLAS_JES_2012_PileRho_HWW | 0 | 0.0573 | -0.0295 | -0.0351 |
| alpha_ATLAS_JES_BJET | 0 | 0.159 | 0.139 | 0.146 |
| alpha_ATLAS_JES_Eta_Modelling | 0 | -0.0403 | 0.0176 | 0.0426 |
| alpha_ATLAS_JES_FlavComp_HWW_WW | 0 | 0.00883 | -0.0057 | -0.00786 |
| alpha_ATLAS_JES_FlavComp_HWW_other | 0 | -0.135 | 0.161 | 0.234 |
| alpha_ATLAS_JES_FlavComp_HWW_tt | 0 | 0.00733 | 0.0164 | 0.0166 |
| alpha_ATLAS_JES_FlavResp | 0 | -0.177 | -0.222 | -0.204 |
| alpha_ATLAS_JES_MU | 0 | 0.159 | 0.094 | 0.0874 |
| alpha_ATLAS_JES_NPV | 0 | -0.0184 | -0.0159 | -0.00866 |
| alpha_ATLAS_JES_NonClosure_AFII | 0 | 0.000778 | -0.0129 | -0.0403 |
| alpha_ATLAS_MET_RESOSOFT | 0 | -0.0351 | -0.0286 | -0.0291 |
| alpha_ATLAS_MET_SCALESOFT | 0 | -0.00415 | -0.0022 | -0.00207 |
| alpha_ATLAS_MU_EFF | 0 | 0.118 | 0.0339 | 0.00933 |
| alpha_ATLAS_MU_ESCALE | 0 | -0.0334 | -0.0184 | -0.015 |
| alpha_ATLAS_MU_ID_RES | 0 | 0.0124 | -0.00209 | -0.00348 |
| alpha_ATLAS_MU_ISO | 0 | 0.452 | 0.272 | 0.234 |
| alpha_ATLAS_MU_MS_RES | 0 | -0.144 | -0.0781 | -0.0704 |
| alpha_ATLAS_MU_RESCALE_lvlv_2012 | 0 | -0.176 | -0.0717 | -0.0598 |
| alpha_ATLAS_MU_TRIGGER_HWW | 0 | 0.2 | 0.0894 | 0.0648 |
| alpha_ATLAS_Matching_ACCEPT | 0 | -0.000124 | 0.0343 | 0.0758 |
| alpha_ATLAS_PTllRewSyst | 0 | 3.64e-07 | -3.27e-09 | -8.32e-05 |
| alpha_ATLAS_TOP_ME | 0 | -0.476 | -0.105 | -0.07 |
| alpha_ATLAS_TOP_PDF | 0 | -0.0203 | -0.00398 | -0.00241 |
| alpha_ATLAS_TOP_PS | 0 | -0.219 | -0.138 | -0.141 |
| alpha_ATLAS_TOP_SCALEF_NONTOP_0j_HWW | 0 | -0.134 | -0.0292 | -0.0174 |

*Continued on next page*

Table 31 – *Continued from previous page*

| Parameter | Prefit | $\mu = 0$ | $\mu = 1$ | $\hat{\mu} = 1.16$ |
|---|---|---|---|---|
| alpha_ATLAS_TOP_SCALEF_STATS_0j_HWW | 0 | -0.114 | -0.025 | -0.0148 |
| alpha_ATLAS_TOP_SCALEF_THEO_0j_HWW | 0 | -0.168 | -0.0606 | -0.0486 |
| alpha_ATLAS_TOP_Scale | 0 | -0.162 | -0.0279 | -0.0148 |
| alpha_ATLAS_TRACKMET_RESOPARASOFT | 0 | 0.115 | 0.21 | 0.221 |
| alpha_ATLAS_TRACKMET_RESOPERPSOFT | 0 | -0.0953 | -0.0574 | -0.0532 |
| alpha_ATLAS_TRACKMET_SCALESOFT | 0 | 0.0838 | 0.102 | 0.102 |
| alpha_ATLAS_VGammaShapeLepPt | 0 | 0.22 | 0.155 | 0.152 |
| alpha_ATLAS_WW_EWCorr_HWW | 0 | -0.206 | -0.0564 | -0.0421 |
| alpha_ATLAS_WW_MTSHAPEMATCHING | 0 | 0.882 | 0.0326 | -0.0697 |
| alpha_ATLAS_WW_MTSHAPEPSUE | 0 | 0.0405 | 0.0393 | 0.0452 |
| alpha_ATLAS_WW_MTSHAPESCALE | 0 | 0.135 | 0.174 | 0.178 |
| alpha_ATLAS_WgsJetBin0 | 0 | 0.0323 | 0.0334 | 0.0343 |
| alpha_ATLAS_WgsJetBin1 | 0 | -0.163 | -0.203 | -0.202 |
| alpha_ATLAS_WgsJetBin2 | 0 | 0.0315 | 0.0267 | 0.0256 |
| alpha_ATLAS_WgsMTscale | 0 | -0.00753 | -0.00806 | -0.00804 |
| alpha_ATLAS_ZTAUTAU_MODELING | 0 | 0.128 | 0.00307 | -0.0136 |
| alpha_ATLAS_ZTAUTAU_PDF | 0 | -0.0507 | -0.0183 | -0.0151 |
| alpha_ATLAS_ZTAUTAU_PTZREW | 0 | -0.316 | -0.0261 | 0.0066 |
| alpha_ATLAS_ZTAUTAU_PYTHIAMCSTAT_SR_0j | 0 | 0.00524 | -0.00283 | -0.00442 |
| alpha_ATLAS_ZTAUTAU_PYTHIAMCSTAT_SR_1j | 0 | -0.00576 | -0.0273 | -0.029 |
| alpha_ATLAS_ZTAUTAU_PYTHIAMCSTAT_WWCR_0j | 0 | -0.0683 | -0.00493 | 0.0027 |
| alpha_ATLAS_ZTAUTAU_PYTHIAMCSTAT_WWCR_1j | 0 | -0.0481 | -0.0168 | -0.0146 |
| alpha_ATLAS_ZTAUTAU_SCALE | 0 | 0.207 | 0.0247 | 0.00677 |
| alpha_ATLAS_btag21j_extrap | 0 | 0.0222 | 0.0565 | 0.0617 |
| alpha_ATLAS_ggWW_XS | 0 | -0.124 | 0.155 | 0.186 |
| alpha_ATLAS_ggfMTPSUE | 0 | 3.93e-05 | -0.0302 | -0.0244 |
| alpha_ATLAS_ggfMTmatching | 0 | 5.42e-05 | -0.0382 | -0.0335 |
| alpha_ATLAS_ggfMTscale | 0 | 1.25e-05 | -0.00377 | -0.00777 |
| alpha_FakeRateCorr_QCD_HWW | 0 | 0.0446 | 0.12 | 0.126 |
| alpha_FakeRateOther_QCD_HWW | 0 | 0.0121 | 0.0515 | 0.0555 |
| alpha_FakeRateStat_QCD_HWW | 0 | 0.0368 | 0.0552 | 0.0573 |
| alpha_FakeRate_EL_Corrl_HWW | 0 | -0.569 | -0.426 | -0.415 |
| alpha_FakeRate_EL_Other_HWW | 0 | -0.209 | -0.248 | -0.249 |
| alpha_FakeRate_EL_Stat_10_15_HWW | 0 | -0.506 | -0.508 | -0.508 |
| alpha_FakeRate_EL_Stat_15_20_HWW | 0 | 0.822 | 0.449 | 0.406 |
| alpha_FakeRate_EL_Stat_20_25_HWW | 0 | -0.05 | -0.162 | -0.176 |
| alpha_FakeRate_EL_Stat_GT25_HWW | 0 | -0.541 | -0.219 | -0.191 |
| alpha_FakeRate_EL_Uncorrl_OS_HWW | 0 | 0.594 | -0.0607 | -0.131 |
| alpha_FakeRate_EL_Uncorrl_SS_HWW | 0 | -0.272 | -0.133 | -0.107 |
| alpha_FakeRate_MU_Corrl_HWW | 0 | 0.194 | 0.14 | 0.144 |
| alpha_FakeRate_MU_Other_HWW | 0 | 0.0809 | 0.0107 | -0.000134 |
| alpha_FakeRate_MU_Stat_10_15_HWW | 0 | 0.254 | 0.225 | 0.213 |
| alpha_FakeRate_MU_Stat_15_20_HWW | 0 | 0.357 | 0.0396 | 0.00506 |
| alpha_FakeRate_MU_Stat_20_25_HWW | 0 | 0.0618 | -0.0221 | -0.0307 |
| alpha_FakeRate_MU_Stat_GT25_HWW | 0 | -0.383 | -0.325 | -0.318 |
| alpha_FakeRate_MU_Uncorrl_OS_HWW | 0 | 1.17 | 0.494 | 0.419 |
| alpha_FakeRate_MU_Uncorrl_SS_HWW | 0 | -0.182 | -0.0499 | -0.0277 |
| alpha_LUMI_2012 | 0 | -0.00144 | 0.0635 | 0.00468 |
| alpha_PM_f_recoil_DY_SR0j_HWW | 0 | -0.48 | -0.635 | -0.647 |
| alpha_PM_f_recoil_DY_SR1j_HWW | 0 | 0.166 | 0.0657 | 0.0658 |
| alpha_PM_f_recoil_NDY_SR0j_HWW | 0 | -0.0997 | -0.104 | -0.104 |
| alpha_PM_f_recoil_NDY_SR1j_HWW | 0 | 0.234 | 0.16 | 0.154 |
| alpha_PM_f_recoil_NDY_ZP0j_HWW | 0 | 0.0957 | 0.118 | 0.124 |
| alpha_PM_f_recoil_NDY_ZP1j_HWW | 0 | -0.0617 | -0.0308 | -0.0238 |
| alpha_PM_theta_SR0j | 0 | 0.000441 | -0.196 | -0.235 |
| alpha_PM_theta_SR1j | 0 | 8.99e-05 | -0.0423 | -0.0554 |
| alpha_QCDscale_V | 0 | 0.0103 | 0.00129 | 0.000271 |
| alpha_QCDscale_VH | 0 | -9.17e-07 | 0.000663 | 0.000569 |
| alpha_QCDscale_VV | 0 | 0.0711 | 0.0728 | 0.0822 |
| alpha_QCDscale_VV_ACCEPT | 0 | 0.694 | 0.157 | 0.0981 |
| alpha_QCDscale_Wg_ACCEPT0j_HWW | 0 | -0.0709 | -0.0914 | -0.0939 |
| alpha_QCDscale_Wg_ACCEPT1j_HWW | 0 | 0.392 | 0.205 | 0.199 |
| alpha_QCDscale_Wg_ACCEPT2j_HWW | 0 | -0.000983 | -0.00656 | -0.0065 |
| alpha_QCDscale_ggH | 0 | -1.78e-05 | 0.151 | 0.0112 |

*Continued on next page*

Table 31 – *Continued from previous page*

| Parameter | Prefit | $\mu = 0$ | $\mu = 1$ | $\hat{\mu} = 1.16$ |
|---|---|---|---|---|
| alpha_QCDscale_ggH_ACCEPT | 0 | -9.71e-05 | 0.0994 | 0.0597 |
| alpha_QCDscale_ggH_pTH_m01 | 0 | 0.000431 | -0.261 | -0.235 |
| alpha_QCDscale_ggHe1 | 0 | 0.000176 | -0.0168 | -0.107 |
| alpha_QCDscale_qqH | 0 | 5.05e-07 | -0.000131 | -0.000313 |
| alpha_QCDscale_qqH_ACCEPT | 0 | 8.93e-06 | -0.00231 | -0.00554 |
| alpha_pdf_Higgs_gg | 0 | 1.73e-05 | -0.15 | -0.0109 |
| alpha_pdf_Higgs_gg_ACCEPT | 0 | -1.32e-05 | 0.0731 | 0.00821 |
| alpha_pdf_Higgs_qq | 0 | 6.56e-06 | -0.00106 | -0.00407 |
| alpha_pdf_Wg_ACCEPT_HWW | 0 | 0.00414 | -0.0162 | -0.0172 |
| alpha_pdf_Wgs_ACCEPT_HWW | 0 | -0.00332 | -0.00759 | -0.0074 |
| alpha_pdf_gg | 0 | -0.0229 | 0.0232 | 0.0279 |
| alpha_pdf_gg_ACCEPT | 0 | 0.0144 | 0.00782 | 0.00704 |
| alpha_pdf_qq | 0 | 0.097 | 0.0704 | 0.0786 |
| alpha_pdf_qq_ACCEPT | 0 | 0.423 | 0.0491 | 0.00969 |

| Parameter | Prefit | $\mu = 0$ | $\mu = 1$ | $\hat{\mu} = 1.16$ |
|---|---|---|---|---|
| ATLAS_norm_Diboson0j | 1 | 0.98 | 0.982 | 0.982 |
| ATLAS_norm_Diboson1j | 1 | 0.894 | 0.879 | 0.873 |
| ATLAS_norm_SF_MUSR_DY0j | 1 | 1.17 | 1.12 | 1.12 |
| ATLAS_norm_SF_MUSR_DY1j | 1 | 1.65 | 1.51 | 1.49 |
| ATLAS_norm_SF_MU_DY0j | 1 | 1.05 | 1.06 | 1.06 |
| ATLAS_norm_SF_MU_DY1j | 1 | 1.44 | 1.43 | 1.42 |
| ATLAS_norm_Top1j | 1 | 1.03 | 1.03 | 1.03 |
| ATLAS_norm_TopPF2j | 1 | 1.02 | 1.02 | 1.02 |
| ATLAS_norm_WW0j | 1 | 1.26 | 1.22 | 1.21 |
| ATLAS_norm_WW1j | 1 | 1.14 | 1.11 | 1.11 |
| ATLAS_norm_Ztautau0j | 1 | 0.977 | 0.986 | 0.987 |
| ATLAS_norm_Ztautau1j | 1 | 1.04 | 1.05 | 1.05 |
| ATLAS_norm_btag | 1 | 0.984 | 0.984 | 0.984 |
| PM_EFF_f_recoil_DY0j | 0 | 0.106 | 0.107 | 0.107 |
| PM_EFF_f_recoil_DY1j | 0 | 0.136 | 0.125 | 0.121 |
| PM_EFF_f_recoil_NDY_SR0j | 0 | 0.685 | 0.685 | 0.684 |
| PM_EFF_f_recoil_NDY_SR1j | 0 | 0.651 | 0.647 | 0.647 |
| PM_EFF_f_recoil_NDY_ZP0j | 0 | 0.682 | 0.683 | 0.683 |
| PM_EFF_f_recoil_NDY_ZP1j | 0 | 0.651 | 0.653 | 0.653 |

Table 32: Post-fit background normalization factors and Drell-Yan efficiencies for the 0-jets and the 1-jet analysis depending on the signal strength parameter of the different fit scenarios: $\mu = 0$ for the background-only hypothesis, $\mu = 1$ for the SM signal hypothesis and $\hat{\mu} = 1.16$ for the observed signal strength.

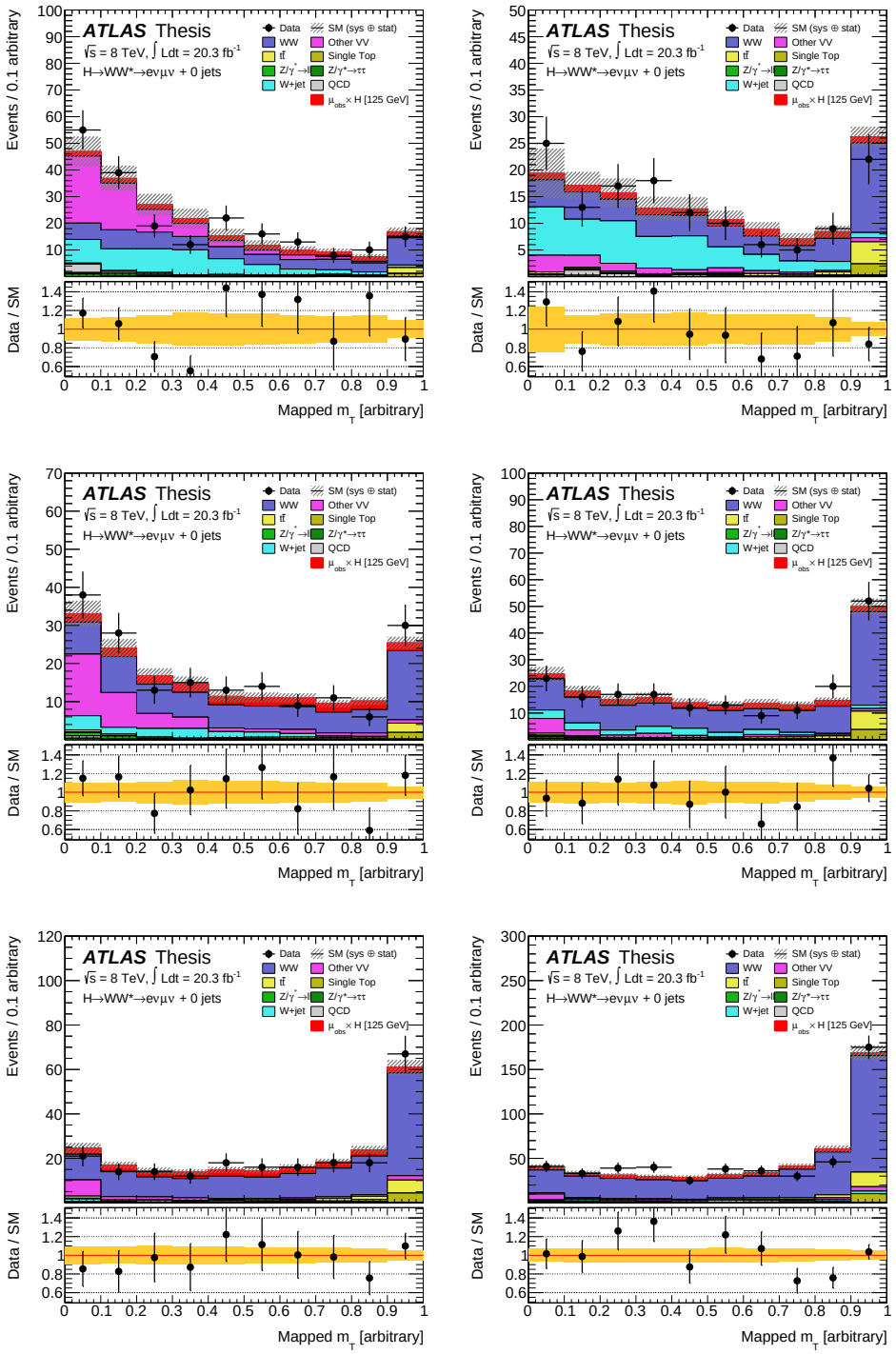



Figure 113: Post-fit remapped transverse mass distribution of the $e\mu$ signal regions of the 0-jets analysis. The plots on the left show the $m_{\ell\ell}$ selection below 30 GeV, those on the right above. The $p_T^{sub}$ range increases from $[10, 15]$ GeV on top to $[15, 20]$ GeV in the middle and $[20, \infty]$ GeV on the bottom.

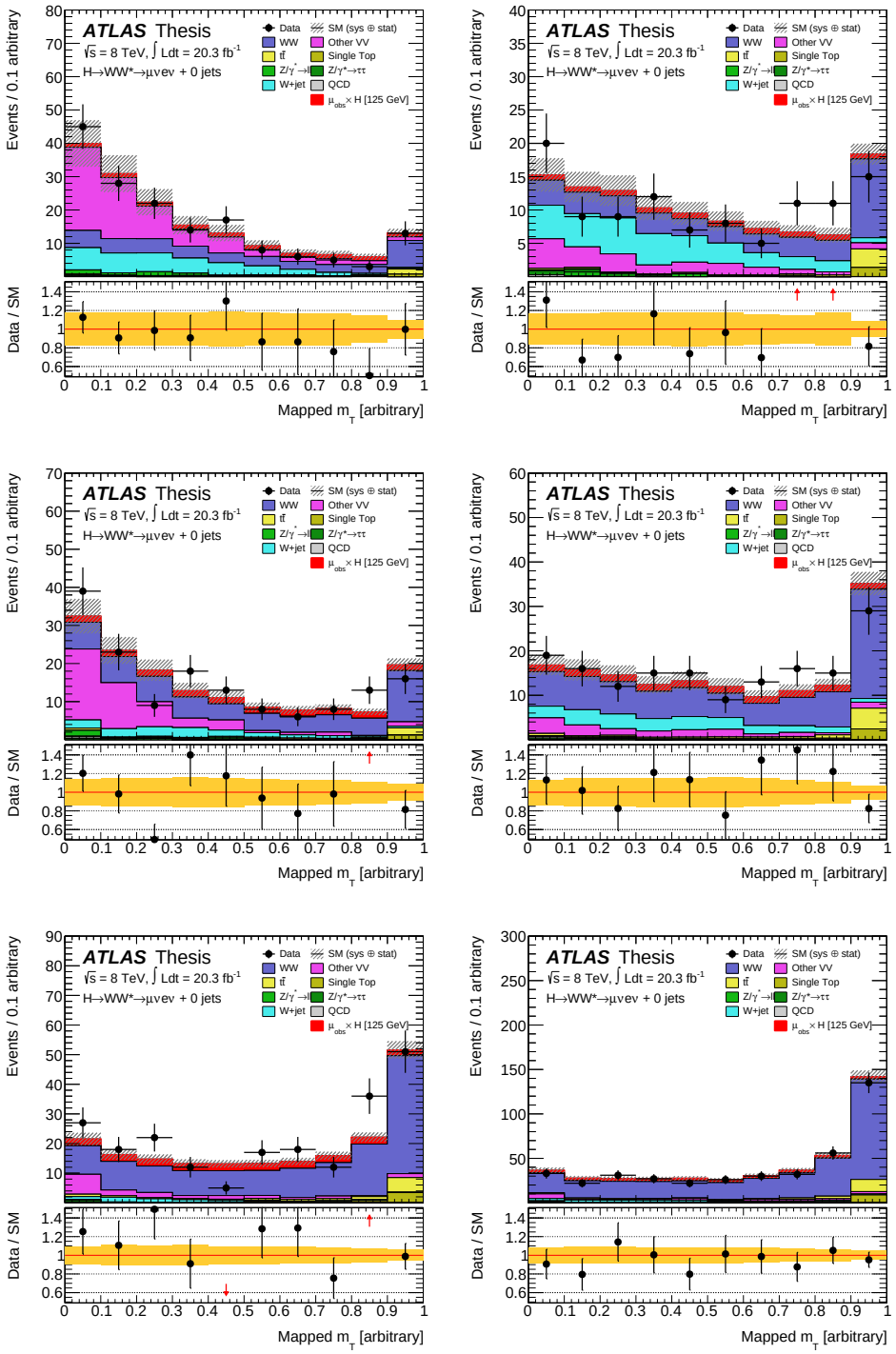



Figure 114: Post-fit remapped transverse mass distribution of the $\mu e$ signal regions of the 0-jets analysis. The plots on the left show the $m_{\ell\ell}$ selection below 30 GeV, those on the right above. The $p_T^{sub}$ range increases from $[10, 15]$ GeV on top to $[15, 20]$ GeV in the middle and $[20, \infty]$ GeV on the bottom.

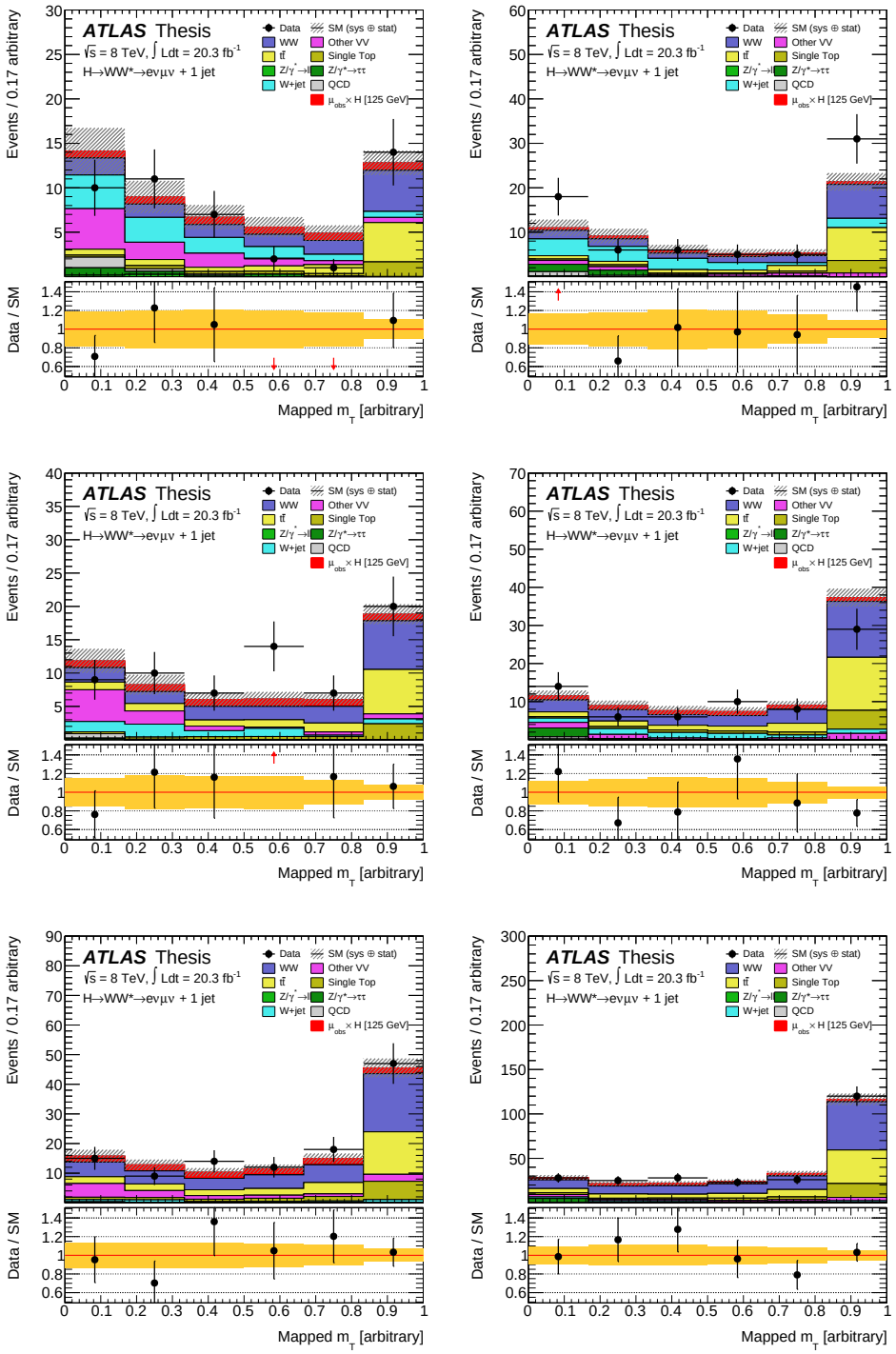



Figure 115: Post-fit remapped transverse mass distribution of the $e\mu$ signal regions of the 1-jet analysis. The plots on the left show the $m_{\ell\ell}$ selection below 30 GeV, those on the right above. The $p_T^{sub}$ range increases from $[10, 15]$ GeV on top to $[15, 20]$ GeV in the middle and $[20, \infty]$ GeV on the bottom.

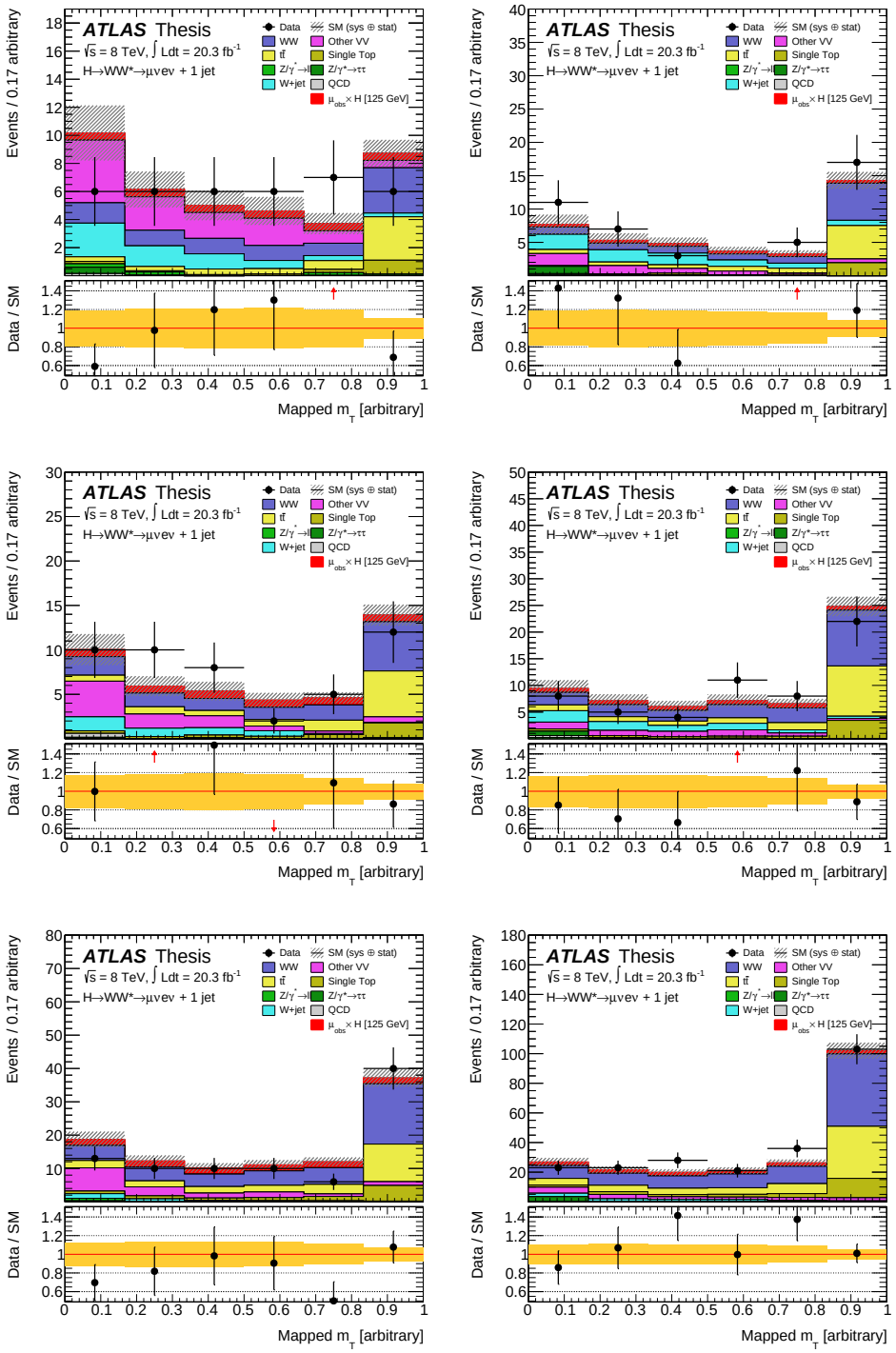



Figure 116: Post-fit remapped transverse mass distribution of the $\mu e$ signal regions of the 1-jet analysis. The plots on the left show the $m_{\ell\ell}$ selection below 30 GeV, those on the right above. The $p_T^{sub}$ range increases from $[10, 15]$ GeV on top to $[15, 20]$ GeV in the middle and $[20, \infty]$ GeV on the bottom.

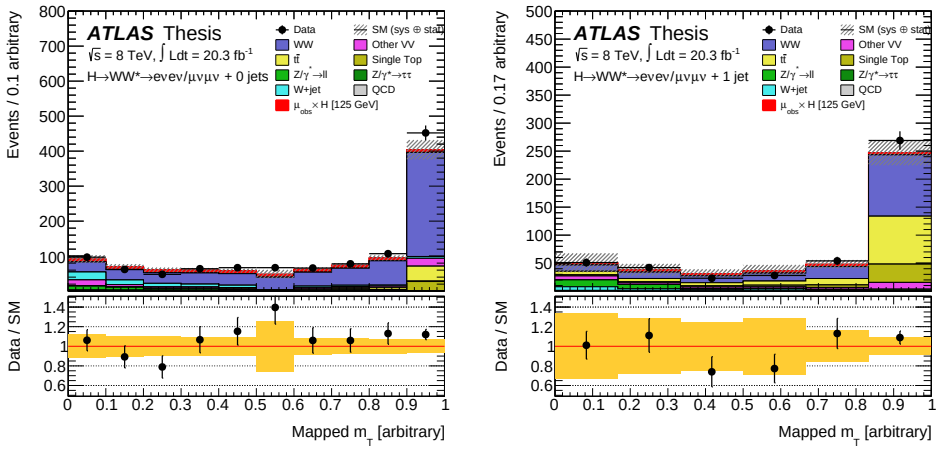



Figure 117: Post-fit remapped transverse mass distribution of the same flavour ($ee+\mu\mu$) signal regions. The 0-jets channel is shown on the left and the 1-jet channel on the right.

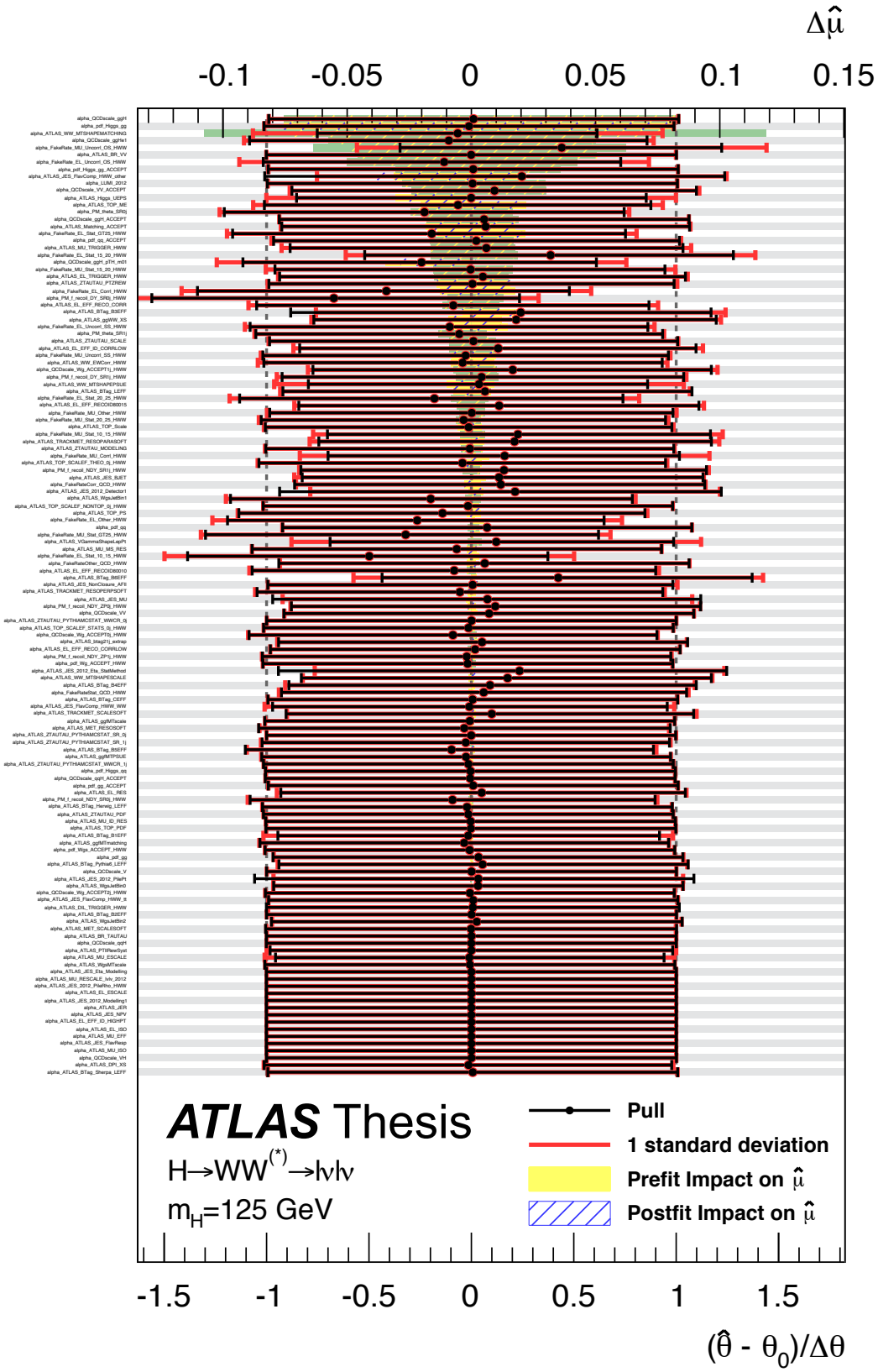



Figure 118: Fit results of all nuisance parameters ranked by their post fit impact on the observed signal strength $\Delta\hat{\mu}$ plotted versus the upper scale and the corresponding pulls $\delta\hat{\theta}$ plotted versus the lower scale.

| Signal region selection | | | Summary | | | Background composition | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Data | Bkgd | Signal 125 | $WW$ | $t\bar{t}$ | st | diboson | $Z \to \ell\ell$ | $Z \to \tau\tau$ | W+jets | QCD |
| 0j: $e\mu$ | $m_{\ell\ell} < 30$ GeV | $10 < p_T^{sub} < 15$ GeV | 209 | 186 ± 11 | 17 ± 3 | 55 ± 2 | 4.9± 0.4 | 2.2 ± 0.2 | 64 ± 6 | 3.0 ± 0.4 | 1.7 ± 0.1 | 51 ± 9 | 4.8 ± 1 |
| | | $15 < p_T^{sub} < 20$ GeV | 177 | 146 ± 6 | 22 ± 4 | 81 ± 3 | 5.7 ± 0.4 | 3.3 ± 0.2 | 38 ± 4 | 2.3 ± 0.4 | 0.4 ± 0.1 | 14 ± 2 | 1.0 ± 0.2 |
| | | $p_T^{sub} > 20$ GeV | 214 | 191 ± 7 | 28 ± 4 | 149 ± 6 | 12 ± 1 | 7.5 ± 1 | 16 ± 2 | 1.0 ± 0.2 | 0.1 ± 0.1 | 5.4 ± 1 | 0.2 ± 0.1 |
| | $m_{\ell\ell} > 30$ GeV | $10 < p_T^{sub} < 15$ GeV | 137 | 128 ± 9 | 12 ± 2 | 55 ± 2 | 6.8 ± 1 | 3.6 ± 0.3 | 11 ± 1 | 0.8 ± 0.1 | 1.5 ± 0.1 | 48 ± 8 | 1.4 ± 0.2 |
| | | $15 < p_T^{sub} < 20$ GeV | 190 | 172 ± 6 | 21 ± 3 | 117 ± 5 | 12 ± 1 | 6.6 ± 0.5 | 15 ± 1 | 0.6 ± 0.1 | 0.9 ± 0.1 | 19 ± 3 | 0.9 ± 0.1 |
| | | $p_T^{sub} > 20$ GeV | 503 | 466 ± 16 | 35 ± 6 | 375 ± 16 | 30 ± 2 | 17 ± 1 | 22 ± 2 | 1.0 ± 0.2 | 2.1 ± 0.2 | 18 ± 3 | 0.5 ± 0.1 |
| 0j: $\mu e$ | $m_{\ell\ell} < 30$ GeV | $10 < p_T^{sub} < 15$ GeV | 161 | 152 ± 9 | 12 ± 2 | 39 ± 2 | 3.2 ± 0.2 | 1.5 ± 0.1 | 70 ± 7 | 4.6 ± 1 | 0.4 ± 0.1 | 33 ± 6 | 0.7 ± 0.1 |
| | | $15 < p_T^{sub} < 20$ GeV | 153 | 134 ± 6 | 16 ± 3 | 63 ± 3 | 4.4 ± 0.3 | 2.4 ± 0.2 | 46 ± 5 | 2.6 ± 0.4 | 0.8 ± 0.1 | 15 ± 2 | 0.9 ± 0.2 |
| | | $p_T^{sub} > 20$ GeV | 218 | 176 ± 6 | 23 ± 4 | 134 ± 6 | 10 ± 1 | 6.2 ± 0.5 | 17 ± 2 | 1.1 ± 0.2 | 0.2 ± 0.0 | 6.7 ± 1 | 0.2 ± 0.1 |
| | $m_{\ell\ell} > 30$ GeV | $10 < p_T^{sub} < 15$ GeV | 107 | 101 ± 6 | 7.8 ± 1 | 39 ± 2 | 4.5 ± 0.3 | 2.2 ± 0.2 | 18 ± 2 | 1.2 ± 0.2 | 1.6 ± 0.1 | 34 ± 6 | 0.5 ± 0.1 |
| | | $15 < p_T^{sub} < 20$ GeV | 159 | 140 ± 6 | 15 ± 2 | 86 ± 4 | 8.3 ± 1 | 4.5 ± 0.3 | 16 ± 2 | 0.6 ± 0.1 | 0.6 ± 0.1 | 23 ± 4 | 0.7 ± 0.1 |
| | | $p_T^{sub} > 20$ GeV | 414 | 406 ± 14 | 29 ± 4 | 326 ± 14 | 25 ± 2 | 15 ± 1 | 20 ± 2 | 0.9 ± 0.1 | 0.7 ± 0.1 | 17 ± 3 | 0.5 ± 0.1 |
| 0j: $ee + \mu\mu$ | $f_{recoil} < 0.1$ | | 1108 | 1032 ± 37 | 80 ± 13 | 735 ± 31 | 63 ± 5 | 38± 3 | 72 ± 7 | 39 ± 12 | 0.2 ± 0.1 | 82 ± 14 | 2.0 ± 0.3 |
| 1j: $e\mu$ | $m_{\ell\ell} < 30$ GeV | $10 < p_T^{sub} < 15$ GeV | 45 | 49 ± 3 | 4.9 ± 1 | 13 ± 1 | 7.9± 0.4 | 3.1 ± 0.2 | 10 ± 1 | 0.4 ± 0.1 | 1.7 ± 0.1 | 11 ± 2 | 1.8 ± 0.3 |
| | | $15 < p_T^{sub} < 20$ GeV | 67 | 52 ± 2 | 6.2 ± 1 | 19 ± 1 | 12 ± 1 | 4.5 ± 0.2 | 8.9 ± 1 | 0.1 ± 0.0 | 0.2 ± 0.1 | 6.6 ± 1 | 0.6 ± 0.1 |
| | | $p_T^{sub} > 20$ GeV | 115 | 101 ± 4 | 12 ± 2 | 46 ± 4 | 27 ± 1 | 11 ± 1 | 12 ± 2 | 0.2 ± 0.1 | 0.1 ± 0.1 | 3.9 ± 1 | 0.2 ± 0.1 |
| | $m_{\ell\ell} > 30$ GeV | $10 < p_T^{sub} < 15$ GeV | 71 | 56 ± 3 | 3.6 ± 1 | 17 ± 1 | 12 ± 1 | 4.6 ± 0.2 | 3.9 ± 1 | 0.2 ± 0.1 | 3.3 ± 0.2 | 14 ± 3 | 1.3 ± 0.2 |
| | | $15 < p_T^{sub} < 20$ GeV | 73 | 77 ± 3 | 6.2 ± 1 | 32 ± 2 | 22 ± 1 | 7.9 ± 0.4 | 5.2 ± 1 | 0.2 ± 0.1 | 2.8 ± 0.2 | 7.4 ± 1 | 0.6 ± 0.1 |
| | | $p_T^{sub} > 20$ GeV | 250 | 237 ± 10 | 16 ± 3 | 116 ± 9 | 65 ± 3 | 26 ± 1 | 13 ± 2 | 0.6 ± 0.2 | 5.0 ± 0.4 | 10 ± 2 | 0.5 ± 0.1 |
| 1j: $\mu e$ | $m_{\ell\ell} < 30$ GeV | $10 < p_T^{sub} < 15$ GeV | 37 | 36 ± 2 | 3.1 ± 1 | 9.4 ± 1 | 5.1 ± 0.3 | 1.7 ± 0.1 | 12 ± 2 | 0.8 ± 0.2 | 0.5 ± 0.1 | 6.2 ± 1 | 0.2 ± 0.1 |
| | | $15 < p_T^{sub} < 20$ GeV | 47 | 41 ± 2 | 4.7 ± 1 | 15 ± 1 | 9.2 ± 0.5 | 3.3 ± 0.2 | 8.4 ± 1 | 0.4 ± 0.1 | 0.0 ± 0.1 | 4.1 ± 1 | 0.5 ± 0.1 |
| | | $p_T^{sub} > 20$ GeV | 89 | 93 ± 4 | 11 ± 2 | 42 ± 3 | 22 ± 1 | 9.3 ± 0.5 | 15 ± 2 | 0.9 ± 0.3 | 0.1 ± 0.1 | 3.4 ± 1 | 0.2 ± 0.1 |
| | $m_{\ell\ell} > 30$ GeV | $10 < p_T^{sub} < 15$ GeV | 43 | 37 ± 2 | 2.4 ± 0.4 | 11 ± 1 | 8.1 ± 0.4 | 2.8 ± 0.1 | 5.1 ± 1 | 0.2 ± 0.1 | 1.3 ± 0.1 | 8.4 ± 2 | 0.2 ± 0.1 |
| | | $15 < p_T^{sub} < 20$ GeV | 58 | 58 ± 2 | 4.3 ± 1 | 24 ± 2 | 15 ± 1 | 5.4 ± 0.3 | 5.4 ± 1 | 0.3 ± 0.1 | 1.0 ± 0.1 | 7.2 ± 1 | 0.6 ± 0.1 |
| | | $p_T^{sub} > 20$ GeV | 234 | 211 ± 9 | 14 ± 3 | 102 ± 8 | 59 ± 3 | 22 ± 1 | 15 ± 2 | 0.6 ± 0.2 | 3.8 ± 0.3 | 7.9 ± 1 | 0.3 ± 0.1 |
| 1j: $ee + \mu\mu$ | $f_{recoil} < 0.1$ | | 467 | 431 ± 20 | 23 ± 5 | 184 ± 14 | 122 ± 6 | 47 ± 2 | 30 ± 4 | 27 ± 11 | 1.0 ± 0.3 | 20 ± 4 | 0.2 ± 0.1 |

Table 33: Post-fit signal region event selection of the 0-jets and the 1-jet analysis. Signal and background Monte Carlo expectation have been scaled by the respective parameters determined by the statistical fit. The post-fit uncertainties include all statistical and systematical components.

| Sample Yield | | Stat. Error | Theo. Syst. | Expt. Syst. | Total Syst. |
|---|---|---|---|---|---|
| 0j | Signal $m_H = 125$ GeV | – | 6.7 | 15 | 16 |
| | Total Background | 1.5 | 1.2 | 1.7 | 2.5 |
| | WW | 2.4 | 2.3 | 2.6 | 4.2 |
| | Top | 2.3 | 4.2 | 5.6 | 7.4 |
| | W+jets/QCD | – | 9.9 | 14 | 17 |
| | VV | 4.8 | 4.6 | 7.4 | 9.9 |
| | $Z/\gamma^* \to \tau\tau$ | 1.7 | 33 | 7.2 | 34 |
| | $Z/\gamma^* \to \ell\ell$ | 14 | 26 | 5.5 | 30 |
| 1j | Signal $m_H = 125$ GeV | – | 5.3 | 22 | 22 |
| | Total Background | 1.7 | 1.4 | 2.1 | 3 |
| | WW | 5.5 | 2.7 | 4.6 | 7.7 |
| | Top | 3.4 | 2.9 | 2.3 | 5 |
| | W+jets/QCD | – | 11 | 14 | 18 |
| | VV | 8.9 | 6.1 | 8.5 | 14 |
| | $Z/\gamma^* \to \tau\tau$ | 3.3 | 26 | 6.3 | 27 |
| | $Z/\gamma^* \to \ell\ell$ | 27 | 26 | 7.4 | 39 |

Table 34: Breakdown of the relative post fit uncertainties on the signal yield, the total background yield as well as the individual background processes in %, each one decomposed into their statistical (Stat.), theoretical (Theo.) and experimental (Expt.) components. The upper box contains the values of the 0-jets analysis, while the lower one represents the 1-jet analysis. Entries that are marked with a dash either do not apply or refer to values below 0.1%.