GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Fakultät für
Physik

Bachelor's Thesis

Suche nach Doppelhiggsproduktion im Zerfallskanal $\gamma\gamma WW^*$ in geboosteter Topologie

Search for di-Higgs production in the $\gamma\gamma WW^*$ decay channel in the boosted topology

prepared by

Joshua Falco Beirer

from Schaffhausen

at the II. Physikalisches Institut

Thesis number:	II.Physik-UniGö-BSc-2017/04
Thesis period:	10th April 2017 until 17th July 2017
First referee:	Prof. Dr. Stan Lai
Second referee:	Priv. Doz. Dr. Jörn Große-Knetter

Abstract

In the following thesis a search for non-resonant and resonant di-Higgs boson production in the $hh \to \gamma\gamma W_{had} W^*_{had}$ decay channel is presented. The analysis is based on 36.1 fb^{-1} of collision data recorded by the ATLAS experiment at the LHC in 2015 and 2016 at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$. The channel combines the high resolution of the $h \to \gamma\gamma$ process and the high branching ratio of the $h \to WW^*(\to q\bar{q}q\bar{q})$ decay channel. The analysis was optimised for a resonant mass point $m_H = 750\,\text{GeV}$ and applied to boosted topologies in a mass range $500\,\text{GeV} < m_H < 3\,\text{TeV}$, where the quarks could be resolved by at least one large Anti-$k_t(R = 1.0)$ jet. The optimisation was performed by maximizing the Cowan significance using a cut based approach. In this regard, the discrimination power of kinematic variables as well as substructure variables such as N-subjettiness for boosted topologies is studied. The channel was studied for a tight photon identification as well as for a loose photon ID. The background is simulated using a data-driven method on the photon side in addition to Monte-Carlo simulated Standard Model $h \to \gamma\gamma$ samples.

As no evidence for di-Higgs boson production was observed, upper limits were set on $\sigma(gg \to H) \times \text{BR}(H \to hh)$ and $\sigma(gg \to hh)$ for the resonant and non-resonant Higgs boson pair production respectively. Upper limits for non-resonant production are at $27.16^{+40.11}_{-19.57}\,\text{pb}$ (expected) and $48.37\,\text{pb}$ (observed). The most stringent limits for the resonant samples could be set for $m_H = 1.5\,\text{TeV}$ with $3.69^{+3.06}_{-1.49}\,\text{pb}$ (expected) and $2.07\,\text{pb}$ (observed).

Contents

1. Introduction

Since the middle of the 20$^{\text{th}}$ century, the Standard Model (SM) of particle physics has been developed to be the most successful theory of contemporary physics. Its triumph arguably began in 1961 when Sheldon Glashow discovered a way to integrate electromagnetic and weak interactions into one single electroweak theory [1]. In 1967, Steven Weinberg and Abdus Salam incorporated the Higgs mechanism [2] in Glashow's electroweak theory [3, 4] and shaped the SM as we know it today. Since then, one observation after another confirmed the predictions of the SM. The electroweak theory was confirmed in 1983 with the experimental discovery of the Z [5] and W [6] bosons. With the discovery of the Higgs boson in 2012 [7, 8], the full range of particles predicted by the Standard Model was observed for the first time. Although this was a major breakthrough, many questions still remain open today. For instance, with the amount of data collected today, it is not possible to study Higgs boson self-coupling, an important prediction of the SM. Furthermore, the SM has limitations and is not able to describe all observations in nature. As a consequence, several models try to extend the SM by adding more than one Higgs boson. To validate these models, it is necessary to further investigate the SM as well as extended Higgs sectors.

In this thesis, a search for non-resonant and resonant di-Higgs boson production in the $hh \to \gamma\gamma W_{had}W^*_{had}$ decay channel is presented. As the high precision era at the LHC approaches, the investigation of di-Higgs boson production is likely to gain more importance. It is the most promising possibility to study the Higgs boson self-coupling and is also sensitive to contributions of models that predict resonant Higgs boson pair production.

The thesis is structured as follows: in Chapter 2, a brief description of the Standard Model and its limitations is provided. In addition, the Higgs mechanism is outlined and the prospects of Higgs boson pair production in the SM framework and in extended theories are discussed. The ATLAS detector and its most important components are described in Chapter 3. Chapter 4 presents the experimental methods which were specifically used in this analysis and Chapter 5 outlines the analysis strategy, especially why the $hh \to \gamma\gamma W_{had}W^*_{had}$ channel is chosen and how the background estimation is performed. The optimisation for the signal selection is presented in Chapter 6 and the final results, including limits on di-Higgs boson production, are presented in Chapter 7. Finally, a brief summary of the results as well as an outlook is given in Chapter 8.

2. Theoretical background

In the following section, a brief description of the Standard Model and its limitations is provided. In addition, the Higgs mechanism is outlined and the prospects of Higgs boson pair production in the SM framework and extended theories are discussed.

2.1. The Standard Model

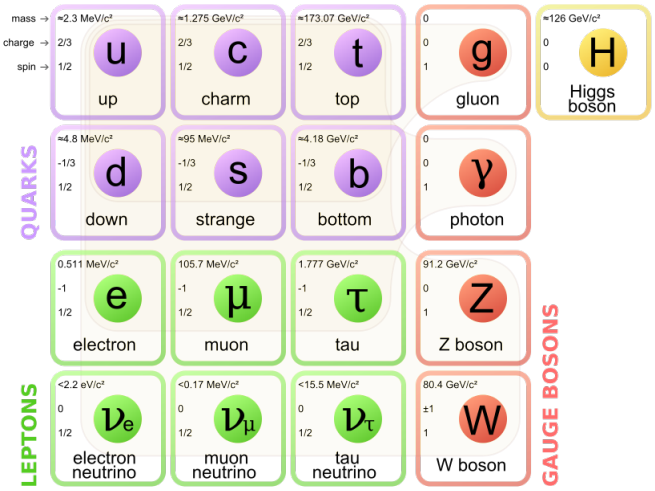


Fig. 1: The fermions and bosons of the Standard Model [9]

The Standard Model aims to describe the constituents of matter at the most fundamental level. To date, it provides the best known description of the properties and interactions of the elementary particles. According to the SM, the universe is made up of 12 spin-$\frac{1}{2}$ particles known as fermions. Fermions can be classified into two groups, namely the leptons and quarks. In addition to a quantum property called colour charge, quarks can either carry $\frac{2}{3}$ or $-\frac{1}{3}$ of the elementary electric charge. In contrast, the electron, muon and tau carry an electric charge of -1 and their respective neutrino partners are charge-less. Furthermore, fermions are grouped into three generations, each comprised of two leptons and two quarks. The first generation includes up and down quarks, the second strange and charm quarks, and the third bottom and top quarks. As particles in higher generations are heavier and less stable, they tend to decay to lower-generation particles. As a consequence, only first generation particles exist commonly in nature. Finally, for each fermion there is a corresponding antifermion. For instance, the antiparticle that corresponds to the electron is the positron. They are identical in terms of mass, lifetime and spin but are oppositely charged. This applies for the electric charge as well as for

the mentioned colour charge. Apart from being categorized in terms of their intrinsic properties like mass, SM particles can be classified by the forces by which they are influenced. In Table 1, the four fundamental forces are listed with their relative strength compared to the strong force. Note that gravity is over 29 orders of magnitude weaker than the weak force and is not included in the Standard Model.

Force	Relative Strength	Boson
Strong	1	Gluon (g)
Electromagnetism	10^{-3}	Photon (γ)
Weak	10^{-8}	$W^{\pm}$
		Z Boson (Z)
Gravity	10^{-37}	Graviton? (G)

Tab. 1: Relative strength of the four fundamental forces and their respective force carriers.

Each force is mediated through a spin-1 gauge boson. The SM predicts four different force carriers: the gluon for the strong force, the photon γ for electromagnetism and the $W^{\pm}$ bosons and Z boson for the weak interaction. In contrast, the Higgs boson is a scalar boson, i.e. a spin-0 particle and plays a vital role in the Standard Model (see section 2.1.4). The particles of the Standard Model are depicted in Figure 1.

2.1.1. Quantum electrodynamics

The electromagnetic interaction is described by quantum electrodynamics (QED). All quantum field theories describe particles as excitations of certain fields. For instance, spin-$\frac{1}{2}$ particles are described as excitations of a spinor field satisfying the Dirac equation. In order to fully describe the interactions of the particles, quantum field theory (QFT) requires the Lagrangian to be invariant under a local phase transformation. The underlying symmetry associated with QED is invariance under U(1) local phase transformations. In order to ensure said invariance, so-called gauge fields are introduced in the Lagrange density. This leads to terms in the Lagrangian that can be interpreted as massless gauge bosons, namely the photon in quantum electrodynamics. As a consequence of the U(1) symmetry, the QED coupling constant α_{EM}, i.e. the strength with which a fermion couples to a photon, is proportional to the elementary charge.

All charged fermions interact with the photon.[1] It is important to remark that the calculation of probability amplitudes for a certain interaction always yields complicated integrals over a large number of variables. However, the use of perturbation theory allows us to represent perturbative contributions to the probability amplitudes in form of Feynman diagrams. The total matrix element is then the sum over all interactions represented by the Feynman diagrams. As $\mathcal{M}_{LO} \propto \alpha^2$ and $\mathcal{M}_{NLO} \propto \alpha^4$ we can neglect higher order terms for low energies, since $\alpha_{EM} \approx \frac{1}{137} \ll 1$. Nonetheless, α_{EM} is a *running* coupling constant and increases for higher momentum transfers. Therefore, the leading order approximation will yield poorer results for a higher centre-of-mass energy.

2.1.2. Quantum chromodynamics

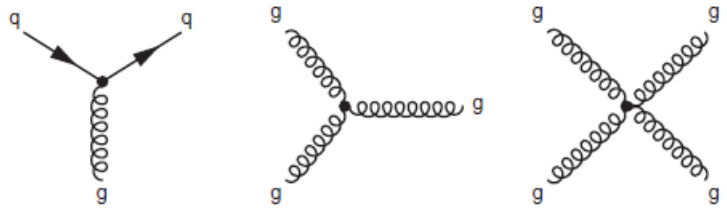


Fig. 2: The predicted QCD interaction vertices arising from the requirement of SU(3) local gauge invariance. [10]

The strong force is the strongest force in the SM and thus dominates interactions whenever possible. Its underlying symmetry is invariance under local SU(3) phase transformations. In order to interact strongly, a particle needs to carry a colour charge, denoted by colour states red (r), green (g) or blue (b). These correspond to orthogonal states of the SU(3) symmetry. As leptons cannot carry a colour charge, only quarks are able to interact through QCD. Furthermore, the gluon itself is a colour carrier and can take on one of eight different colour states. This allows the gluon to self-interact as shown in Figure 2. An important empirical observation is the phenomena of confinement. Neither quarks nor gluons are able to exist as free particles, but always come in bound colourless states. Colourless states can be mesons $|q\bar{q}\rangle$, baryons $|qqq\rangle$ or antibaryons $|\bar{q}\bar{q}\bar{q}\rangle$.[2] When gluons and quarks are separated with a high energy, quark-antiquark pairs are spontaneously created from the vacuum due to gluon self interaction and create an initial spray

[1] The Standard Model does not allow flavour-changing neutral currents (FCNC's) prior to tree level. For instance, a tau lepton cannot decay into an electron through radiation of a photon. However, related scenarios are discussed in BSM physics.

[2] In 2015, LHCb researchers at CERN discovered a new bound pentaquark state ($uudc\bar{c}$) with a statistical significance of 9σ [11].

of particles. Next, the quarks all recombine into hadrons, a process known as hadronisation. Since at $|q| \approx 1\,\text{GeV}$ the coupling constant of QCD is of $\mathcal{O}(1)$ and perturbation theory cannot be used, no straightforward calculations are possible. However, there are different models and methods which are successful in phenomenologically modelling these processes. It is also worth noting that in contrast to QED, the coupling constant in QCD decreases noticeably with rising energy scales. For instance, at $|q| \approx 100\,\text{GeV}$ $\alpha_s \approx 0.1$, which is sufficiently small to apply perturbation theory. This phenomenon is called asymptotic freedom and is used to calculate hard processes prior to hadronisation (see section 4.1).

2.1.3. The weak interaction

The weak interaction is unique compared to the strong and electromagnetic forces in a number of ways. First of all the weak force is carried by three gauge bosons, namely Z and $W^{\pm}$ bosons. Whereas gluons and photons are massless, the Z and $W^{\pm}$ bosons have masses of 91.2 GeV and 80.4 GeV respectively and thus are the heaviest fundamental particles after the top quark and the Higgs boson in the SM. As mentioned before, the requirement of invariance under local phase transformations leads *a priori* to massless gauge bosons. The reason why the Z and $W^{\pm}$ are massive is described by the Higgs mechanism. In addition, the weak force is able to couple to all fermions and is the only interaction capable of changing the flavour of quarks. The $W^{\pm}$ bosons can only couple to particles with a left handed chirality state and antiparticles with a right handed chirality state. The $W^{\pm}$ pairs the electron, muon and tau to their corresponding neutrinos. For the quarks, the charge difference must be one in order to preserve charge conservation. However, not all couplings are equally strong. This observation was first explained with the so-called Cabibbo hypothesis [12], which states that the weak interactions of quarks have the same strength as the leptons, but the weak eigenstates of quarks differ from the mass eigenstates. Their relation can be compactly written in the form

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \Lambda \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{1}$$

where Λ is the so-called Cabibbo-Kobayashi-Maskawa (CKM) matrix [13]. The weak eigenstates d', s' and b' are expressed as linear combination of the observable mass eigenstates. The square of the matrix component $|V_{ij}|$ is then proportional to the coupling strength between the W boson and the two corresponding quarks. The highest coupling strength is found on the diagonal of the matrix, i.e. for the couplings of quarks within

their respective generations. Even though generation mixing is allowed, those processes are much less likely.

One of the most most important goals of particle physics is to unify the individual quantum field theories into one single theory. This has already been accomplished for electromagnetism and the weak interaction, resulting in the theory of electroweak interactions. The unification is accomplished by extending the underlying symmetry, $\mathrm{SU}(2)_L$, from the weak interaction to the gauge group $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$. The associated gauge bosons are the three W bosons of weak isospin from SU(2) (W_1, W_2, and W_3), and the B boson of weak hypercharge from U(1). The weak hypercharge is the charge of the electroweak interaction and given by $Y = 2(Q - I_W^3)$, whereas I_W^3 is the third component of the weak isospin. The SM is therefore a gauge theory based on the gauge symmetry group

$$\underbrace{SU(3)_C}_{strong} \otimes \underbrace{SU(2)_L \otimes U(1)_Y}_{electroweak},$$

where the associated charge for the strong force is colour charge and the associated charges for the electroweak force are the weak isospin and the weak hypercharge.

2.1.4. The Higgs mechanism

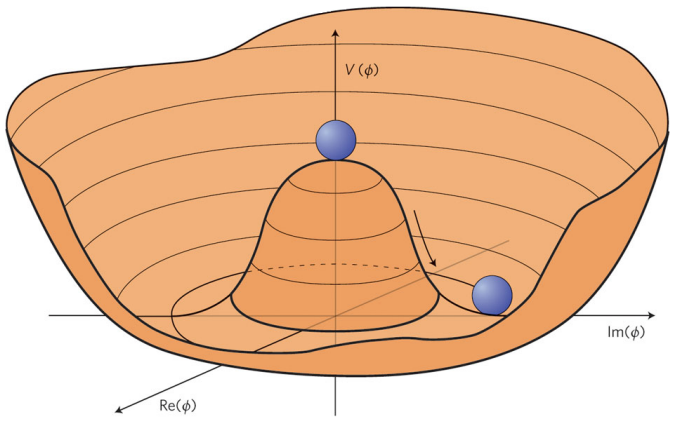


Fig. 3: Form of the Higgs potential for $\mu^2 < 0$ and $\lambda > 0$ [14]

As already stated in section 2.1.3, the underlying symmetries of the interactions require all gauge bosons to be massless. While this is true for the photons and gluons, experiments have shown that the $W^{\pm}$ and Z bosons are among the heaviest fundamental particles in the Standard Model. In fact, even the spin-$\frac{1}{2}$ particles are expected to be

massless to preserve the underlying symmetries. In order to solve this contradiction, a complex doublet of scalar fields is added to the Lagrange density:

$$\phi(\mathbf{x}) = \begin{pmatrix} \phi_1(\mathbf{x}) + i\phi_2(\mathbf{x}) \\ \phi_3(\mathbf{x}) + i\phi_4(\mathbf{x}) \end{pmatrix}, \tag{2}$$

where $\phi_i(\mathbf{x})$ are real scalar fields. The potential of the resulting Lagrangian is known as the Higgs potential and has the form:

$$V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2 \tag{3}$$

where $\mu^2 < 0$ and $\lambda > 0$. The form of the Higgs potential is shown in Figure 3. The lowest energy state does not occur for $\phi(\mathbf{x}) = 0$, as the resulting symmetric potential possesses a global minimum at a distance v from the origin. The potential is thus said to have a non-zero *vacuum expectation value*. When the fields are expressed as an expansion about the new vacuum expectation value, the terms of the masses for the $W^\pm$ and Z bosons naturally emerge. The underlying symmetry of the Lagrangian is conserved but is said to be *hidden*. This phenomena is known as *spontaneous symmetry breaking*. One can simplify the Lagrangian without loss of generality when writing the fields in unitary gauge:

$$\phi(\mathbf{x}) = \begin{pmatrix} 0 \\ v + h(\mathbf{x}) \end{pmatrix}. \tag{4}$$

Here $h(\mathbf{x})$ represents the physical Higgs field. Expanding the Higgs doublet around the electroweak symmetry breaking vacuum expectation value yields terms

$$\mathcal{L} \supset \lambda v^2 h^2 + \lambda v h^3 + \frac{1}{4}\lambda h^4. \tag{5}$$

This process is called Brout-Englert-Higgs mechanism and predicts the existence of a new massive scalar boson, namely the Higgs boson h. The quadratic term can be interpreted as mass term which leads to a Higgs boson mass of

$$m_h = \sqrt{2\lambda}v. \tag{6}$$

The cubic and quartic terms in the Lagrangian predict trilinear and quartic Higgs boson self-coupling. The resulting interaction vertices and their respective coupling strengths are shown in Figure 4.

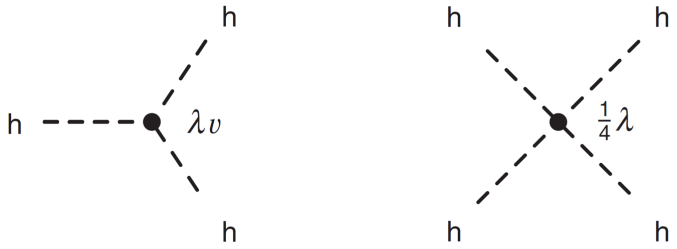


Fig. 4: Trilinear and quartic Higgs-boson self-interaction terms arising from the Higgs mechanism with their corresponding coupling strength.

The mass of the $W^{\pm}$ bosons resulting from spontaneous symmetry breaking can be parametrized as

$$m_W = \frac{1}{2} g_W v, \tag{7}$$

whereas the mass of the Z boson can be expressed as

$$m_Z = \frac{1}{2} \frac{g_W}{\cos\theta_W} v. \tag{8}$$

Here g_W denotes the weak coupling constant and the so-called Weinberg angle is experimentally determined to be $\theta_W \approx 28.74°$ [15]. The Higgs mechanism can also be used to generate the fermion masses via the Yukawa coupling. The fermions interaction with the Higgs field generates the fermion masses

$$m_f = \frac{1}{\sqrt{2}} g_f v, \tag{9}$$

where g_f is known as Yukawa coupling constant.

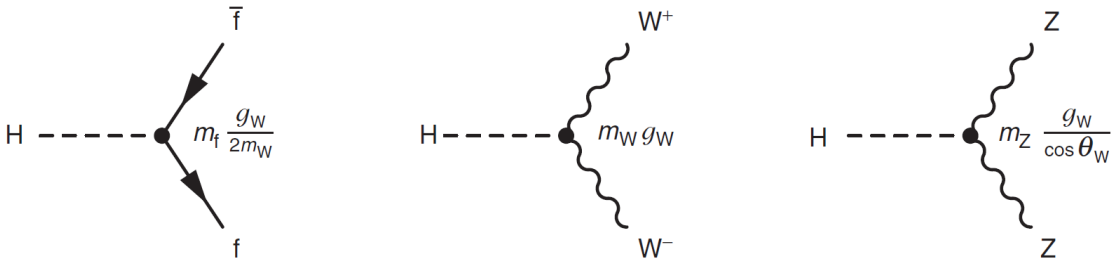


Fig. 5: Primary interaction vertices of the Higgs boson with fermions, $W^{\pm}$ and Z bosons with their corresponding coupling strengths.

Note that the vacuum expectation value v sets the mass scale for all particles. It has been experimentally determined to be $v \approx 246\,\text{GeV}$. Also note that coupling strengths to the Higgs boson are proportional to the particle's mass (see Figure 5).

2.1.5. Higgs boson production

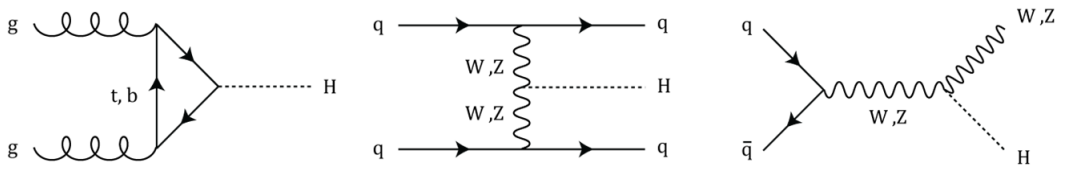


Fig. 6: The three most common singe Higgs production modes at the LHC: gluon-gluon fusion, vector boson fusion and Higgsstrahlung respectively.

The Standard Model allows the Higgs boson to be produced through a number of different processes in pp collisions. The cross sections are heavily dependent on the centre-of-mass energies of the collider (see Figure 7). The cross-sections relevant to this analysis, i.e. for a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$, are shown in Table 2. The most common process at the LHC is known as gluon-gluon fusion (ggF). In this process, the two gluons fuse via a virtual top quark loop to yield a Higgs boson in the final state. The gluons can radiate further gluons which become visible as jets in the detector. The process with the second most highest cross section is the production via the fusion of two vector bosons, namely the $W^{\pm}$ or Z boson. The process is known as vector boson fusion (VBF). The vector bosons are radiated from quarks which are conserved in the process and become visible as highly energetic jets in the hadronic calorimeter. For VBF, the scattered quarks are expected to travel almost parallel to the beam direction, and this topology makes it easier to suppress the background and identify the Higgs signal. The radiation of a Higgs boson from off-shell $W^{\pm}$ or Z bosons, denoted Higgstrahlung (WH and ZH respectively), is also observed. As the gauge bosons are conserved, signatures of decaying W and Z bosons can help to identify this process.

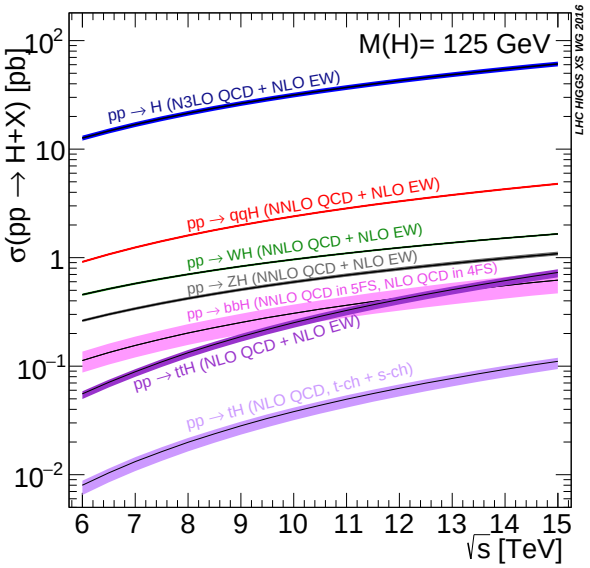


Fig. 7: Predictions for the Standard Model Higgs boson production cross sections as a function of $\sqrt{s}$, assuming a Higgs mass of $m_h = 125\,\text{GeV}$. The lighter colors are the theoretical uncertainties. [16]

The Feynman diagrams of these processes are depicted in Figure 6. The last process considered is the associated production of a Higgs boson with a $t\bar{t}$ pair (ttH). It has the smallest cross section and therefore only plays a minor role in this analysis. While the cross section of the ggF is by far the largest, it is difficult to separate the Higgs signal from the large QCD background that gluon-gluon fusion produces.

Process	Cross section [pb]
ggF	43.92
VBF	3.748
WH	1.380
ZH	0.9753
ttH	0.5085

Tab. 2: Cross sections for the various Higgs boson production processes for $\sqrt{s} = 13\,\text{TeV}$ relevant at the ATLAS experiment. [17]

2.1.6. Higgs boson decay channels

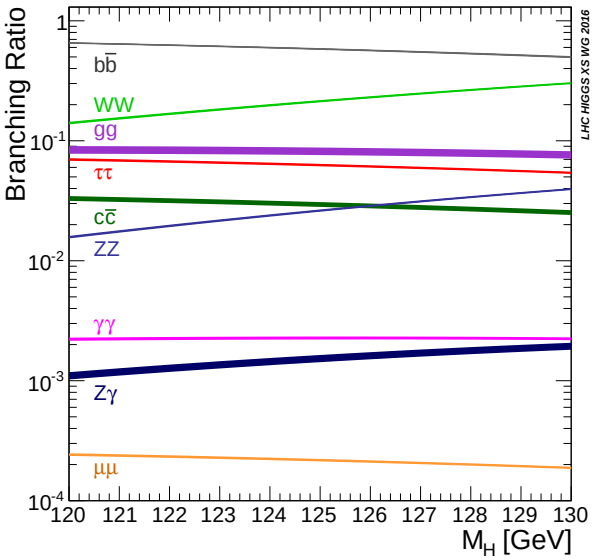


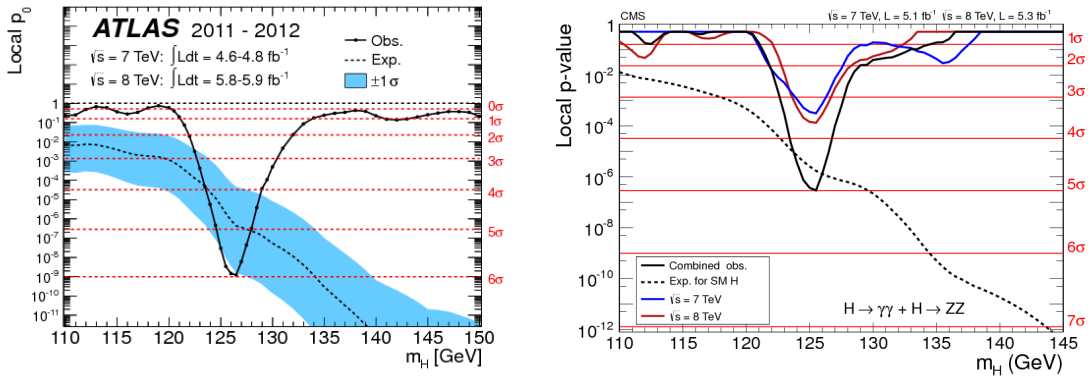
Fig. 8: Predictions for the Standard Model Higgs boson branching ratios as a function of the Higgs boson mass. The lighter colors are the theoretical uncertainties. [16]

With the exception of the heavy top quark, the SM allows the Higgs to potentially decay into all SM particles. As the coupling strength of the Higgs boson is proportional to the mass of the particle (see section 2.1.4), the Higgs boson preferably decays into heavy particles. Therefore, the highest branching ratios for a Higgs decaying into fermions are for the decays into a $b\bar{b}$ pair, a $c\bar{c}$ pair or into $\tau^+\tau^-$. As $m_h < 2m_W$, the Higgs boson is not able to decay into two massive W bosons. However, the second largest branching ratio is given by the decay $h \to WW^*$, where the star indicates a virtual, also known as off-mass-shell, W boson. Even though processes involving off-shell particles are suppressed, the high coupling constant to the massive W boson, $m_W g_W$, compensates for this effect. The Higgs boson can also decay into massless particles via loops of virtual W bosons and top quarks. As the couplings to the heavy particles in the loop are strong, processes like $h \to gg$ and $h \to \gamma\gamma$ play an important role in the decay of the Higgs boson. The branching ratios of the most common decays as a function of the Higgs boson mass are shown in Figure 8. The branching ratios assuming a Higgs boson mass of $m_h = 125$ are shown in Table 3.

Decay mode	Branching ratio
$h \to b\bar{b}$	57.8 %
$h \to WW^*$	21.6 %
$h \to gg$	8.6 %
$h \to \tau^+\tau^-$	6.4 %
$h \to c\bar{c}$	2.9 %
$h \to ZZ^*$	2.7 %
$h \to \gamma\gamma$	0.2 %

Tab. 3: Branching ratios for common Higgs boson decays, assuming a Higgs mass $m_h = 125\,\text{GeV}$. [17]

2.1.7. Higgs boson discovery



(a) Atlas [7]

(b) Cms [8]

Fig. 9: The observed local p-value as a function of m_h in the low mass range. The dashed curve shows the expected local p-value under the hypothesis of a SM Higgs boson signal at that mass. The horizontal dashed lines indicate the p-values corresponding to significances of 1 to 6σ.

As mentioned in the introduction, a new particle was observed 2012 by the Atlas and Cms collaboration with a significance of 5.9 and 5.0σ respectively. The observed local p_0 value and the corresponding significance as a function of m_h is shown in Figure 9. The p-value describes the probability to obtain the observed results if the hypothesis of a SM Higgs boson is wrong. After the discovery there were some doubts whether or not the observed particle was really the Higgs boson. However, further measurements of the particles properties such as spin, parity and coupling strength all strongly pointed to the SM Higgs boson.

Even though the $h \to q\bar{q}$ channels have large branching ratios, they were not of great importance in Higgs searches, as the signatures are subject to large QCD background. The two most important processes which led to the Higgs boson discovery were the $h \to \gamma\gamma$ as well as $h \to ZZ^*(\to 4\ell)$ interactions. Even though they provide low branching ratios (see section 2.1.6), the background processes are well understood and the decay products provide a clear signature in the detector.

The mass of the Higgs boson was experimentally determined to be $126.0 \pm 0.4\,(\text{stat.}) \pm 0.4\,(\text{syst.})$ GeV [7] by the ATLAS collaboration and $125.3 \pm 0.4\,(\text{stat.}) \pm 0.5\,(\text{syst.})$ GeV [8] by the CMS collaboration. The current best value from a combined measurement using Run I data from the ATLAS and CMS experiments is $m_h = 125.09 \pm 0.21\,(\text{stat.}) \pm 0.11\,(\text{syst.})$ GeV [18].

2.2. Limitations of the Standard Model

The SM is not a complete theory of everything (ToE). Gravity is not incorporated in the SM. Until now, every attempt to unite general relativity with the SM has failed, which seems to indicate a fundamental incompatibility between both theories [19]. Additionally, observations suggest that our universe is comprised of much more mass than indirectly measured through cosmic radiation. This has led to the assumption of the existence of an unknown form of matter known as *dark matter* (DM) [20]. In order to explain the lack of observed radiation, DM is expected to consist of weakly interacting massive particles (WIMPs) [21], for which the SM cannot provide a suitable candidate. Finally, the SM predicts a slight asymmetry between matter and antimatter due to CP violation, but there is currently no mechanism that can explain the overwhelming excess of particles to antiparticles in our universe [22]. From a theoretical perspective, there are also some aesthetic issues with the SM. For instance, the Higgs mass receives massive quantum corrections, which are much larger than its actual mass. This means that the *bare mass* must be largely *fine tuned* in order for those corrections to cancel out. Such behaviour is known as *hierarchy problem* [23] and is considered *unnatural* and may indicate the existence of a more fundamental theory.

2.3. Extended models

Alternative theories to the SM have been around since its development. One of the most promising extensions is considered to be supersymmetry (SUSY). SUSY associates a new particle, called superpartner, to every particle in the SM. SUSY does not only resolve

the hierarchy problem, but provides a promising candidate for dark matter, the so-called *neutralino*. Also it allows the unification of all three forces at around 10^{16} GeV, which is considered aesthetic from a theoretical point of view. There exist several SUSY models, including the minimal supersymmetric model (MSSM) [24]. The MSSM only includes the minimal number of new particles and interactions needed for supersymmetry. Some more exotic scenarios dealing with extra dimensions such as the Kaluza–Klein theory [25] and the Randall-Sundrum model [26] are also discussed as possible extensions.

2.3.1. 2HDM models

In section 2.1.4, a complex doublet of scalar fields is added to the Lagrange density as part of the Higgs mechanism and electroweak symmetry breaking, which lead to the prediction of a single spin-0 boson. In fact, the SM Higgs boson is the only known fundamental scalar boson in nature known at the present day. However, extended models propose adding additional scalar multiplets. One of the most promising approaches to study the scalar nature of the electroweak $SU(2)_L \otimes U(1)_Y$ gauge theory is the parameter ρ. Consider n scalar multiplets, ϕ_i, with weak isospin I_i, weak hypercharge Y_i, and vacuum expectation value v_i, then the parameter ρ at tree level becomes [27]:

$$\rho = \frac{\sum_{i=1}^{n} [I_i(I_i+1) - \frac{1}{4}Y_i^2]v_i}{\sum_{i=1}^{n} \frac{1}{2}Y_i^2 v_i} \tag{10}$$

The parameter has been experimentally determined to be very close to one [28] and thus helps to exclude some of the theoretical proposals. The simplest possible extensions of the SM compatible with $\rho \approx 1$ are known as Two-Higgs-Doublet Models (2HDMs) [29]. 2HDM models are considered to be a promising approach as they are capable of solving various SM problems at once. For instance, in supersymmetric models like the MSSM, one Higgs doublet is not enough to generate the masses of up- and down-type quarks simultaneously. Furthermore, the cancellation of gauge anomalies in SUSY models also requires a second doublet [30]. 2HDMs also allow for additional sources of CP violation and might help to explain baryogenesis [31]. The two complex scalar $SU(2)$ doublets that are added in 2HDM models can be compactly written as

$$\Phi_a = \begin{pmatrix} \phi_a^+ \\ (v_a + \rho_a + i\eta_a)\sqrt{2} \end{pmatrix}, \tag{11}$$

where a = 1,2. Three of the eight degrees of freedom are responsible for generating the masses of the $W^{\pm}$ and Z bosons. The other remaining fields lead to five physical Higgs bosons: the known Higgs h from the SM, a heavier neutral Higgs H, a CP odd pseudoscalar A and two charged $H^{\pm}$. Assuming CP conservation in the Higgs sector and the absence of spontaneous CP breaking, the most general potential yields [30]

$$\begin{aligned} V = {} & m_{11}^2 \Phi_1^\dagger \Phi_1 + m_{22}^2 \Phi_2^\dagger \Phi_2 - m_{12}^2 \left(\Phi_1^\dagger \Phi_2 + \Phi_2^\dagger \Phi_1\right) + \frac{\lambda_1}{2}\left(\Phi_1^\dagger \Phi_1\right)^2 + \frac{\lambda_2}{2}\left(\Phi_2^\dagger \Phi_2\right)^2 \\ & + \lambda_3 \Phi_1^\dagger \Phi_1 \Phi_2^\dagger \Phi_2 + \lambda_4 \Phi_1^\dagger \Phi_2 \Phi_2^\dagger \Phi_1 + \frac{\lambda_5}{2}\left[\left(\Phi_1^\dagger \Phi_2\right)^2 + \left(\Phi_2^\dagger \Phi_1\right)^2\right] \end{aligned}, \tag{12}$$

for two doublets Φ_1 and Φ_2 and real parameters m_{ij} and λ_k. 2HDMs are usually categorised into Type-I and Type-II models. In Type-I models, all quarks couple to just one of the Higgs-doublet, Φ_2 by convention. In Type-II models, right-handed up-type quarks couple to Φ_2 and right-handed down-type quarks couple to Φ_1. Note that the MSSM is a Type-II 2HDM.

2.4. Di-Higgs production

Up to the present day, only single Higgs production events have been observed. However, the Standard Model also predicts Higgs boson pair production. In fact, as stated in 2.1.4, the Higgs particle is a consequence of the electroweak symmetry breaking, which predicts Higgs-self coupling and therefore a mechanism for Higgs pair production. In addition, pair production is also allowed for other interactions such as through Higgs-fermion Yukawa couplings as shown in Figure 10.

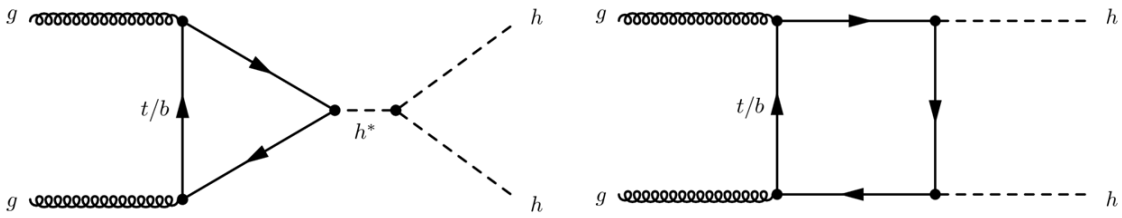


Fig. 10: Leading-order Feynman diagrams of the nonresonant production of Higgs boson pairs in the Standard Model through the Higgs boson self-coupling and the Higgs-fermion Yukawa interactions only. [32]

The main motivation of studying non-resonant di-Higgs production is the measurement of the trilinear Higgs boson self-coupling constant, which is predicted to be

$$\lambda_{hhh}^{SM} = \frac{3m_h^2}{v^2} \tag{13}$$

and only accessible in Higgs boson pair production modes [33]. Note that the predicted SM cross section for di-Higgs boson production at NNLO is approximately 33.45 fb [16] for $\sqrt{s} = 13\,\text{TeV}$. One reason for the small cross section of non-resonant double Higgs boson production is the destructive interference between the Feynman diagrams depicted in Figure 10. However, many BSM models assume a different trilinear coupling constant λ_{hhh}. The cross sections of the di-Higgs boson production modes are directly related to the trilinear coupling constant, assuming that no other unknown processes alter the cross sections (see Figure 11).

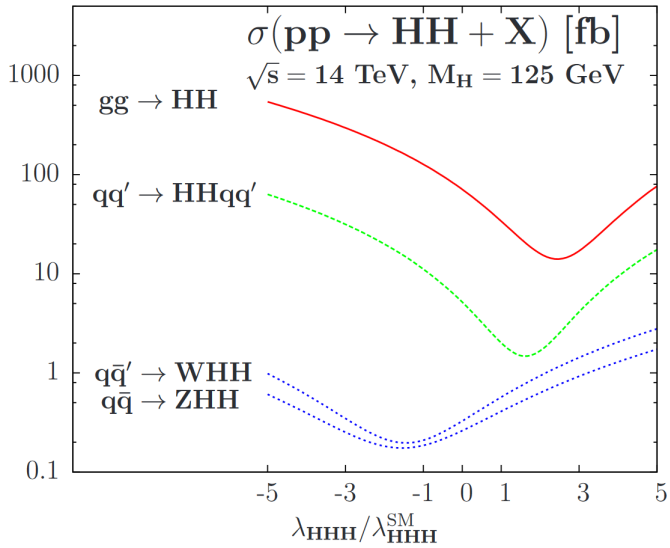


Fig. 11: Cross sections of the various Higgs pair production processes as a function of the relative trilinear Higgs boson self coupling $\lambda_{hhh}/\lambda^{SM}_{hhh}$ at a centre-of-mass energy $\sqrt{s} = 14\,\text{TeV}$, assuming a Higgs boson mass of 125 GeV. Note that H in the plot refers to the SM Higgs boson. [33]

As a consequence, the trilinear Higgs boson self coupling constant can directly be inferred through the measurement of the cross sections. If the observations are consistent with the SM, it is expected that $\frac{\lambda_{hhh}}{\lambda^{SM}_{hhh}} \approx 1$. Deviations of this value would hint to either a different interference than predicted by the SM or to new unknown double Higgs boson production processes. Insufficient data has been recorded by the LHC to be sensitive to the non-resonant cross section which are predicted by the SM. However, if one considers extended Higgs sectors such as those of the MSSM, more possibilities for di-Higgs production arise. For instance, the decay of a heavy 2HDM Higgs boson H into two ligher SM Higgs h is shown in Figure 12.

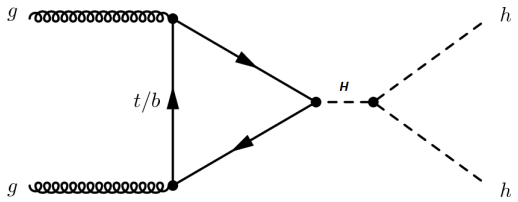


Fig. 12: Feynman diagram depicting the production of a heavy Higgs H like expected in extended 2HDM models such as the MSSM, which then decays into two lighter Higgs bosons h of the SM [32]

3. Experimental setup

The ATLAS experiment is one of the four major experiments at CERN in Geneva. In the following a short overview of the Large Hadron Collider and the ATLAS detector is provided.

3.1. The Large Hadron Collider

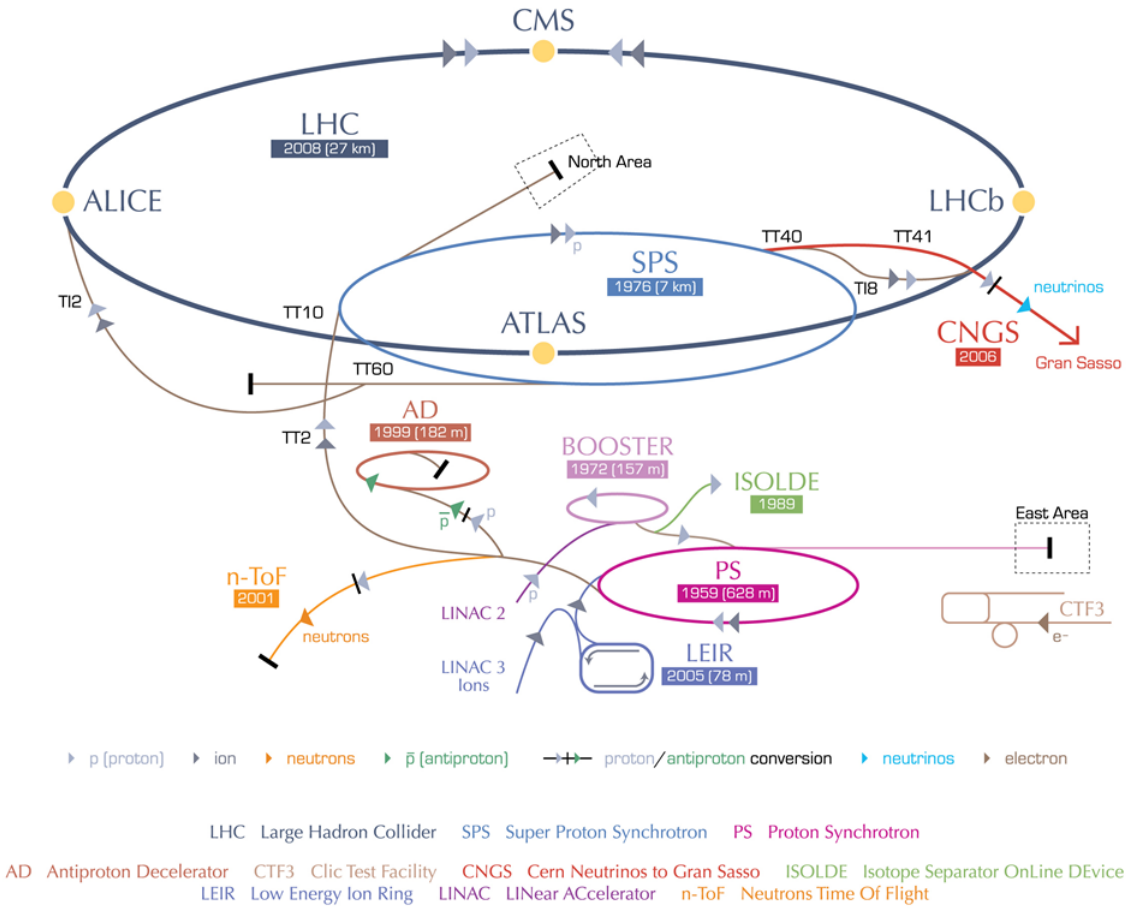


Fig. 13: Sketch of the various pre-accelerators and detectors of the LHC at CERN [34].

In order to study the properties of elementary particles and their fundamental interactions, sophisticated accelerators operating at high centre-of-mass energies are essential. In particular, high luminosity and centre-of-mass energies are needed in order to produce heavy particles and observe extremely rare events such as Higgs boson production. The LHC is a circular proton-proton collider with a circumference of approximately 26.7 km,

located 175 m below the surface in order to shield it from atmospheric radiation. During Run I which took place from 2010 to 2012, the LHC operated at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV producing 5.46 fb^{-1} and 22.8 fb^{-1} of data respectively. The LHC was upgraded and restarted with a centre-of-mass energy of 13 TeV in 2015. In 2016, it delivered a total integrated luminosity of 38.5 fb^{-1}. It is expected that the centre-of-mass energy will be increased to the full design energy of $\sqrt{s} = 14$ TeV in 2018. The protons are not directly injected to the LHC but pass a series of pre-accelerators that successively increase their energy as depicted in Figure 13. They are then accelerated in two adjacent parallel beam pipes in opposite directions within the LHC. These beamlines intersect at four points, namely at the four major experiments ATLAS, CMS, LHCb and ALICE. Dipole magnets are used to keep the beams on their circular path and quadrupole magnets are used to keep the beam focussed. Whereas ATLAS and CMS are multi-purpose experiments, LHCb specializes in the study of b-physics and ALICE in heavy ion collisions.

3.2. The ATLAS detector

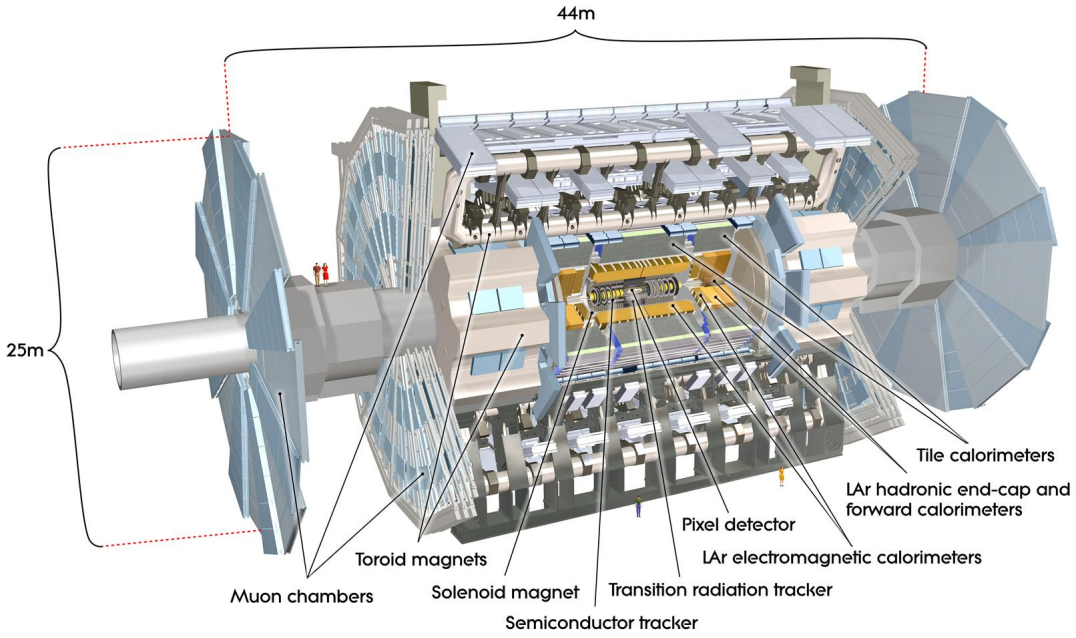


Fig. 14: The ATLAS detector and subsystems. [35]

The ATLAS (A Toroidal LHC ApparatuS) detector consists of end-caps and a series of concentric layers around the collision point which act as various detector types. The

ATLAS detector was built to be nearly hermetic around the interaction point, i.e. to cover the largest possible area around the collision so that sensitivity is ensured for all possible decay products. It can be divided into three major parts: the inner detector, the calorimeters and the muon spectrometer.

3.2.1. Coordinate System

Due to the concentric layout of the detector, ATLAS uses a cylindrical coordinate system. The interaction point, i.e. the point where the beam pipes intersect, is defined as the origin of the coordinate system. The z-axis corresponds to the beam line which traverses the perpendicular $x-y$ plane. Furthermore, θ denotes the polar angle with respect to the z-axis and ϕ the azimuthal angle around the z-axis. An important quantity is the pseudorapidity $\eta \equiv -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]$. The reason why pseudorapidity is preferred over the polar angle θ is because differences in pseudorapidity are Lorentz invariant under boosts along the longitudinal axis in the relativistic limit. As a measure of angular separation between two objects in the detector the quantity $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ is used.

3.2.2. Inner detector

The inner tracking detector begins a few centimetres from the collision point and extends to a length of 6.2 m and a radius of 1.2 m. It is composed of the pixel detector, the Semi-Conductor Tracker (SCT) and the Transition Radiation Tracker (TRT), as shown in Figure 14. It is surrounded by a 2T solenoidal magnetic field, which causes charged particles to bend. By measuring the radius of curvature, the transverse momentum, p_T, can be determined. The silicon pixel detector is composed of approximately 1,740 modules, each containing approximately 74,000 pixels (50 μm $\times$ 400 μm) which can be read out individually. It thus provides extremely precise tracking. The SCT is the middle component of the inner detector and resembles the Pixel detector, except it uses narrow strips instead of pixels. The TRT consists of drift tubes, which are filled with ionizing gas. In order to produce transition radiation, materials with differing refraction indices are incorporated between the tubes. The resolution of the ATLAS tracking detector and all other detector components is shown in Table 4.

3.2.3. Calorimeters

The calorimeters are positioned outside the solenoid magnet. Their function is to measure the deposited energy of incident particles. The calorimeter system consists of a electromagnetic calorimeter (ECAL) and an outer hadronic calorimeter (HCAL). In

ATLAS, both are *sampling calorimeters*, i.e. one material is responsible for developing particle showers and another for detecting the signals. Energies of electrons and photons are measured by the LAr ECAL barrel and end-cap calorimeters. It uses liquid argon as sampling and lead as the absorbing material. The HCAL absorbs energy from particles which interact strongly. In this calorimeter, steel is used as the absorber and plastic scintillators as the active material. Note that there is not a clean split between EM interactions in the ECAL and hadronic interactions in the HCAL. Both types of particles shower in both calorimeters. However, EM showers are shorter and very narrow while hadronic showers are much larger and very wide, so that shower shapes as well as the energy split between the ECAL and the HCAL can be used to distinguish the showers. Also note that the ECAL has a much higher precision in terms of measuring the energy of the particles as well as the location where the energy was deposited (see Table 4).

3.2.4. Muon spectrometer

As muons are approximately 200 times heavier than electrons, they lose only little energy due to bremsstrahlung when traversing matter. The energy loss at GeV scales for electrons and muons is primarily due to inelastic collisions with matter. This has a much larger effect on electrons since they scatter off of other electrons. Muons, however, only lose small amounts of energy per scatter due to their large mass. As a consequence, the muons lose only small energies when travelling through matter and leak out of the detector altogether. The muon chamber is thus furthest away from the interaction point and the largest detector component: it starts at a radius of 4.25 m and extends to the full 11 m radius of the ATLAS detector. The muon spectrometer consists of a series of gas chambers which are filled with a gas mixture mostly containing argon and carbon dioxide. Due to the interactions of the muons with the gas mixture, they produce a track in the detector which is then converted into an electronic signal. It is embedded in a system of toroidal magnets with a magnetic field of 0.5 T.

3.2.5. Trigger system

Another important component of the ATLAS detector is its trigger system. If the raw data were recorded in entirety, it would fill disk space at a rate of approximately 600 Tb/s and thus exceed all current technical limitations in terms of data processing and storage. Furthermore, only around 1 of 10^{13} events is a promising event for finding rare particles such as the Higgs boson or indications for SUSY. To solve this problem efficiently ATLAS

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

Tab. 4: General performance of the ATLAS detector showing the resolutions and coverage angles of the various components. E and p_T are in GeV. [36]

uses a multilevel trigger system, which decides which events to discard and which to save permanently. It consists of a hardware-based Level-1 (L1) and a software-based high level trigger (HLT) [37]. The L1 trigger decides with a latency of 2.5 μs on which events to discard permanently. In general, the L1 trigger scans for the presence of high p_T particles, more specifically, for the presence of high p_T leptons, photons and neutrinos. These might be decay products of heavy particles of interest. In addition, the L1 trigger defines one or more Regions-of-Interest (RoI), where its selection process has identified interesting features. The L1 trigger has an input data stream of 40 MHz and reduces it to $\sim$ 100 kHz . The HLT selections use all the available reconstructed detector data within the RoI and further reduce the event rate to $\sim$1 kHz.

4. Experimental methods

In the following chapter a short description of simulation methods in high energy physics is provided. Furthermore, the production of MC samples for this analysis is described. Jet reconstruction algorithms and further methods relevant to this analysis are outlined.

4.1. Simulation methods

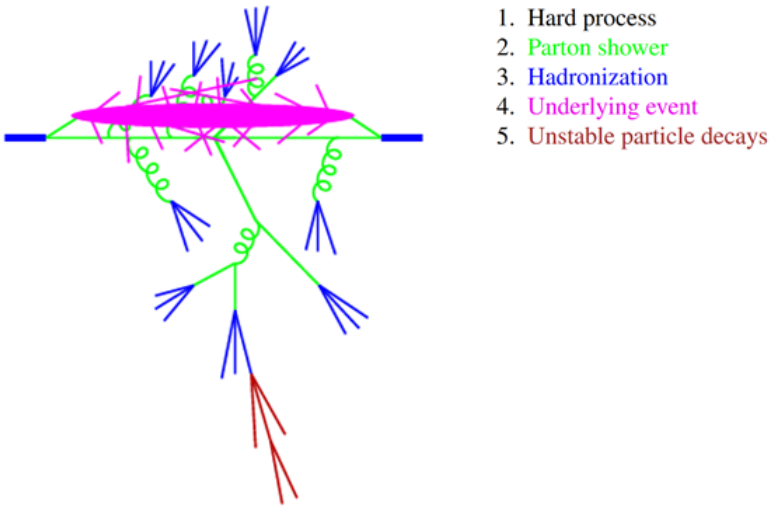


Fig. 15: Diagram showing the structure of a proton-proton collision, where the different colours indicate the different stages involved in event generation. [38]

In particle physics, it is of the utmost importance to model the studied processes based on theoretical predictions of the SM. These simulations are made with so-called Monte Carlo (MC) event generators. The event generation of a proton-proton collision at the LHC can be described by five main different steps as shown in Figure 15. Since most processes that occur in the collision of two protons are not of main interest, not all possible proton interactions are simulated. Instead, MC generators start by simulating hard scatter processes which are related to high momentum transfers and thus likely to generate events containing interesting features. This first step is easily implemented as the probability amplitudes of hard processes can be calculated with perturbation theory. At proton colliders with high centre-of-mass energies like the LHC, the collisions are not between the protons themselves but between its constituents, e.g. gluons and quarks. To take these into account, Parton Distribution Functions (PDFs), which describe the probability distribution of the momentum fraction that a parton carries inside of a proton,

are used to simulate the hard scatter processes. In Figure 16, the parton distribution functions, calculated with HERAPDF1.5 NNLO [39], are shown for a typical momentum transfer region of $Q^2 = 10000\,\text{GeV}$ at hadron colliders such as TEVATRON and the LHC.

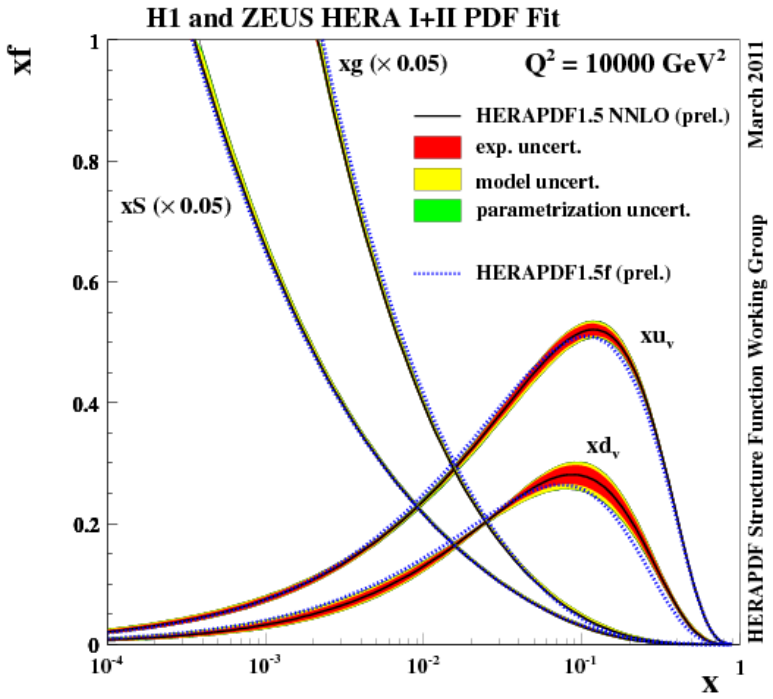


Fig. 16: The parton distribution functions from HERAPDF1.5 NNLO at $Q^2 = 10000\,\text{GeV}$, i.e. a region relevant for the hadron colliders TEVATRON and LHC. The gluon and sea qurk distributions are scaled down by a factor 20. The experimental, model and parametrisation uncertainties are shown separately. [40]

After the hard scattering is calculated, parton showers are considered in a second step (see Figure 15). As particles involved in the hard collisions are gluons and quarks, they all carry colour charge. Similar to electrically charged particles that radiate photons when accelerated (Bremsstrahlung) in QED, colour charged particles can radiate gluons. The main difference is however that the gluons themselves radiate new gluons due to self-coupling in QCD. This leads to aforementioned parton showers, that is, a sequence of gluon radiation with ever decreasing momentum scales. At some point, perturbation theory cannot be used to model this process and phenomenological hadronisation models are required. In a fourth step, interactions apart from main hard scattering are simulated. These include all secondary collisions between proton remnants and is known as

underlying event. Many of the produced hadrons in the hadronisation process as well as in the underlying event are unstable and rapidly decay. As depicted in Figure 15, this is then simulated in a final fifth step. In addition to the pure event generation, a simulation of the ATLAS detector response is needed to make a direct comparison of the measured data with theoretical predictions. The simulation is performed with GEANT4 [41]. The detector response can be simulated with two different methods: either a full simulation or a fast simulation (AFII) can be performed [42]. In the full simulation, every aspect of the detector is modelled and therefore yields an accurate result. However, a considerable amount of computing time is needed. Thus, a fast simulation is often more feasible: it focusses on the most important parts of the detector and provides less sophisticated results but requires considerably less computation time. Whereas the inner detector is modelled identically in AFII and full simulation, approximation methods are used to simulate the electromagnetic and hadronic calorimeter response.

4.1.1. Signal simulation

In terms of the signal, a non-resonant and several samples for resonant di-Higgs production in the $hh \to \gamma\gamma WW^*$ decay channel are simulated. For this, resonant mass points in the range $260\,\text{GeV} \leq m_H \leq 3\,\text{TeV}$ are used. The detector response is simulated using AFII in GEANT4. The hard scattering processes, i.e. the calculation of the matrix elements is performed with the MADGRAPH5_aMC@NLO [43] event generator. For the parton showering, hadronisation and underlying event, HERWIG++[44] is used. As for the parton distribuition functions, the CTEQ6L1 PDF [45] set with the UEEE5 [46] tune was applied.

Mass point m_H [GeV]	Events generated
non-resonant	970000
260	180000
300	199000
400	174000
500	179000
750	184000
1000	200000
1500	195000
2000	179000
2500	197000
3000	199000

Tab. 5: Non-resonant and resonant signal samples produced for several different values of m_H

Table 5 shows the number of events generated for every sample. Assuming a cross section of $\sigma = 1\,\text{pb}$, cross section times branching ratio yields $\sigma \cdot BR(hh \to \gamma\gamma WW^*) = 0.98\,\text{fb}$. A full list of the sample names can be found in the appendix A.1.

4.1.2. Background simulation

To generate the background, only the most common SM background single Higgs boson production $h \to \gamma\gamma$ MC samples were considered. These include the Higgs-boson production modes gluon-gluon fusion (ggF), vector boson fusion (VBF), WH/ZH (Higgsstrahlung) and ttH (associated top pair production) as described in section 2.1.5. The matrix elements for ggF and VBF are calculated using Powheg [47], whereas Pythia8 [48] is used to model hadronisation, showering and the underlying event. The CT10 [49] PDF set with tune AZNLOCTEQ6L1 [50] is used. For ZH and WH, the Pythia8 generator and the NNPDF23LO [51] PDF set with the A14NNPDF23LO [52] tune applied is used. The matrix element calculations for the ttH sample are done with the Madgraph5_aMC@NLO [43] event generator and Herwig++ [44] is used for hadronisation, parton showering and underlying event calculations. In terms of parton distribution functions, the CTEQ6L1 PDF [45] set with the UEEE5 [46] tune is utilised.

Channel	Events generated	$\sigma \cdot BR(h \to \gamma\gamma)$ [fb]
ggF	1930000	110.1
VBF	174000	8.6
WH	246200	3.1
ZH	247800	2.0
ttH	927400	1.2

Tab. 6: Background SM $h \to \gamma\gamma$ samples

Table 6 shows the number of events generated for every background sample and cross section times branching ratio $\sigma \cdot BR(h \to \gamma\gamma)$. A full list of the background sample names can be found in the appendix A.1.

4.1.3. Monte Carlo weighting

It is not possible to simulate every aspect from the collision to the detector response. Therefore, different weights are applied on the MC data to fit the collected collision data of the ATLAS experiment. The following weights are applied:

(i) A weight to account for NNLO effects and beyond is applied by the MC generator itself.

(ii) MC generators are usually designed before the actual data-taking and as a consequence can only make an estimate of the pileup. The actual pileup contributions may look different, wherefore a reweighting is applied such that the pileup profile in MC is adjusted to the data.

(iii) A vertex weight is applied to correct for mismodelling of the z position of the vertex.

(iv) A photon and a lepton scale factor is applied to correct for mismodelling of the reconstructed objects.

(v) A b-tag scale factor corrects mismodelling of the b-tag algorithm, specifically adjusting differences in b-tagging efficiency between data and MC.

4.2. Hadronic jet reconstruction

As discussed in section 2.1.2, unbound partons are not colourless and hadronise as a consequence of confinement. They form bound states which then become visible in the

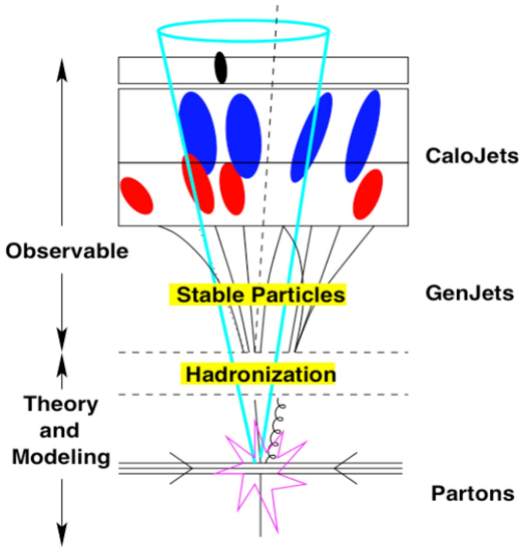


Fig. 17: A cartoon of an event showing the point of collision, the fragmentation and hadronisation of the quarks and gluons and the resulting jet found through the detection of the stable particles. Calojets are those jets created using the calorimeter output whereas Genjets are jets created using stable simulated particles. The dashed line represents the direction of the missing energy. [53]

detector as a collimated spray of hadrons. In order to gain insight into kinematics of the originating parton, it is useful to group all particles coming from the same source to a single object, known as a jet (see Figure 17). Jets form the link between theoretical modelling at the most fundamental level and the observable particles in the detector. It is therefore important to use sophisticated algorithms for jet reconstruction.

4.2.1. Jet reconstruction algorithms

A great variety of jet reconstruction algorithms exist. Notwithstanding that all jet algorithms have advantages and disadvantages, there is a broad consensus on the requirements for such algorithms. Apart from requiring a feasible implementation of the algorithm, i.e. it should be simple to implement in an experimental analysis and in the theoretical calculation, jet algorithms should be infrared and collinear (IRC) safe. Infrared safety means that additional soft particles, should not significantly change the output of the jet algorithm. An infrared unsafe algorithm is therefore sensitive to underlying event and pile-up activity. Collinear safety means that a splitting of a high-p_T particle into two collinear particles should not change the outcome of the algorithm. Both scenarios are illustrated in Figure 18. Another important aspect to consider is the

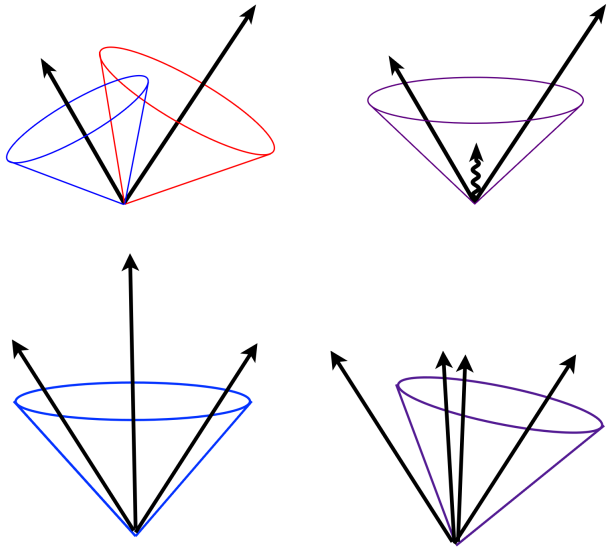


Fig. 18: Illustration of the infrared sensitivity of a cursory designed jet algorithm (top). Illustration of the product of a collinear unsafe jet algorithm. A collinear splitting changes the number of jets (bottom). [54]

size of the jet. The size of a jet can parametrized through the jet radius R, a parameter of the algorithm that determines its angular reach. Note that the jet radius does not correspond to its geometrical radius, as the algorithms do not necessarily lead to perfect cones. A large jet radius helps to collect all hadrons and thus yields more precise calculations of the jets mass and energy. Nevertheless, a smaller radius helps to reduce the inclusion of pile-up and underlying event contributions in the calculations and therefore prevents an overestimation of the jets energy and mass. At the present day, two main classes of jet algorithms are in use. The first class of jet algorithms are known as cone algorithms and assume that all hadronised particles originating from the initial parton will show up in a conical region in the detector. Cone algorithms are generally IRC unsafe. The second class of algorithms are known as sequential jet clustering algorithms and include the Cambridge-Aachen [55] algorithm as well as the k_t [56] and Anti-k_t [57] algorithm. Together they are known as the generalized k_t algorithm. Cone algorithms were preferred by experimentalists in the past, as they are simple to implement and fast in terms of processing time. However, since the introduction of the FastJet [58] program, sequential jet clustering algorithms gained considerable computational performance and are now used extensively as they provides the benefit of being infrared and collinear safe. In the following the generalized k_t algorithm is discussed in detail. In

general, the k_t algorithm family can use various sources of information such as charged particle tracks from the tracking detectors, energy clusters from the calorimeters or truth hadrons. In this analysis only information based on the calorimeters was considered as input (LCTopo clusters). The algorithm begins by creating a list of proto-jets and calculating the distance

$$d_{ij} = \min\left(p_{T,i}^{2p}, p_{T,j}^{2p}\right) \frac{\Delta R_{ij}^2}{R^2} \tag{14}$$

between every pair of proto-jets i and j and the distance

$$d_{i,B} = p_{T,i}^{2p} \tag{15}$$

between each proto-jet i and the Beam B. $p_{T,i}$ is the transverse momentum of the cluster i. The quantity $\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ is a measure of the angular separation between the two proto-jets and $y = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right)$ is the rapidity. The parameter p defines the type of algorithm. A value $p = +1$ corresponds to the k_t algorithm, $p = 0$ to the Cambride-Aachen algorithm and $p = -1$ to the Anti-k_t algorithm, which is used in this analysis. The algorithms then finds the minimum of the pairs of proto-jets. If the minimal distance is between two proto-jets, both objects are merged and the individual clusters removed from the list. If the minimal value is $d_{i,B}$, the proto-jet i is removed from the list and called a jet. The procedure is then repeated until no proto-jets are left. Even though all methods are valid, their results differ substantially.
A comparison between the k_t algorithms and the cone algorithm SISCone [59] is shown in Figure 19. From equation 14, it can be inferred that the k_t algorithm first combines soft particles, whereas the Anti-k_t algorithm initially clusters the highest energy depositions. As can be seen from Figure 19, the Anti-k_t yields the best jet shape. As it starts reconstruction from high-p_T particles it is also less sensitive to pile-up and underlying event contributions.

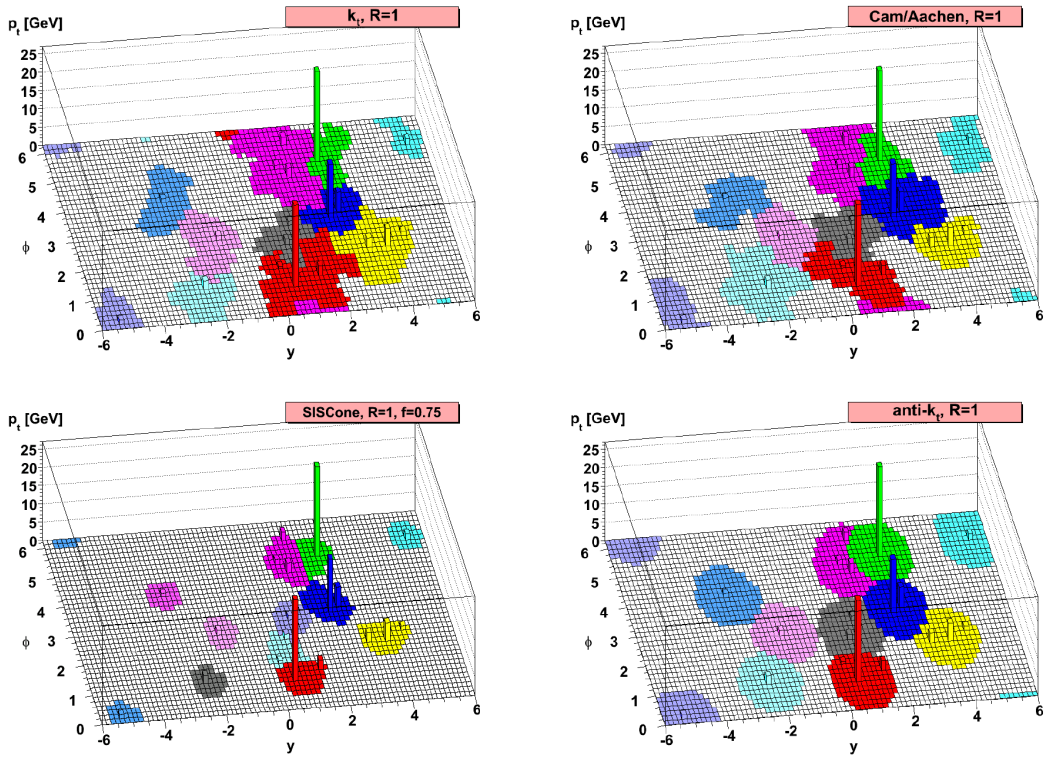


Fig. 19: Jet reconstruction using the k_t, Anti-k_t, Cambridge-Aachen and the cone algorithm SISCone. Engergy depositions are depicted in the y-ϕ plane and the respective transverse momentum p_T is shown on the z-axis. The same Monte-Carlo event is used. [54]

4.2.2. Large radius jets

The angular separation between two objects in the lab frame resulting from a two body decay can be approximated to be

$$\Delta R \approx \frac{2m}{p_T}, \tag{16}$$

for negligible masses of the decay products. Here, m is the mass and p_T the transverse momentum of the mother particle. For instance, a hadronically decaying W boson originating from a SM Higgs boson will lead to an angular separation $\Delta R \approx 0.4$ of the two quarks, assuming a transverse momentum $p_T = 400\,\text{GeV}$ of the W boson. At ATLAS, usually a jet radius of $R = 0.4$ for Anti-k_t jets is used. This approach aims to resolve individual jets resulting from single partons originating from high-p_T particles which are likely produced from relatively heavy particles like the SM Higgs boson. However,

particles that originate from even heavier resonances will have even higher transverse momentum and therefore will be closer to each other in the lab frame, so the that highly boosted topologies cannot be resolved using one Anti-$k_t(R = 0.4)$ jet for each parton.

4.2.3. Calibration and grooming procedures

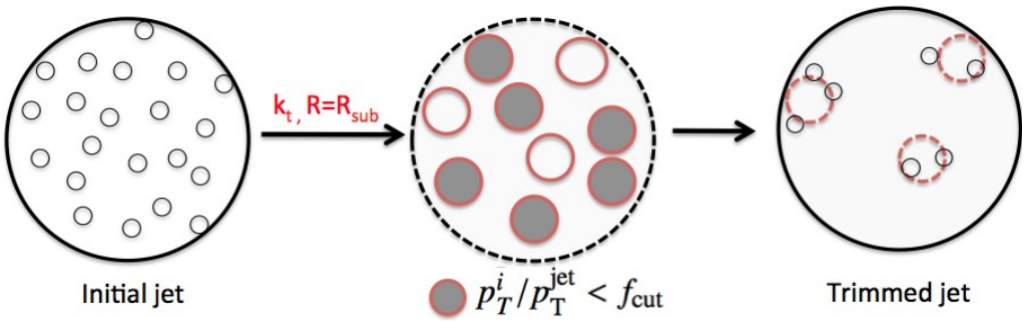


Fig. 20: Schematic depiction of the grooming method known as trimming algorithm. Subjets within a seed jet are discarded if the ratio between the transverse momentum p_T^i of the subjet i and the transverse momentum p_T^{jet} of the associated seed jet is smaller than a fixed cut value f_{cut}. The remaining jet is known as trimmed jet. [60]

Jet reconstruction algorithms are generally not experiment-specific and therefore do not account for several effects which are unique to a certain experiment. For instance, the ATLAS detector mostly uses sampling calorimeters (see section 3.2.3) and thus not all of the energy is deposited in the active medium. This will lead to an underestimation of the deposited energy. Also, the finite detector size as well as further effects regarding the detector design also alter the jet related measurements. To correct for these effects, an energy calibration is necessary. To achieve this, the reconstructed jet energy is compared to Monte-Carlo data and then scaled correspondingly. This is mostly done by studying γ + jets events, as the photon side can be measured with high precision and the information on the jet side extracted via momentum conservation. The jet mass is calibrated using well known mass distributions such as for top quarks and W bosons [61].

As mentioned in section 4.2.1, large radius jets are susceptible to considerably more amount of pile-up and underlying event contributions compared to small radius jets. To reduce this effect, several so-called jet grooming [62] techniques are available. All grooming methods seek to remove soft objects from pile-up or underlying event contributions and considerably improve the jet reconstruction. Similar to the jet reconstruction

algorithms, there are a few algorithms available with their own advantages and disadvantages, the most importants being the mass-drop filtering [63], trimming [60] and pruning [64] algorithms. Throughout this analysis trimming is used. A schematic depiction of the trimming algorithm is shown in Figure 20. In a first step, jets are clustered as usual, in this case with the Anti-k$_t$ algorithm. The resulting jets are known as seed jets. Then the constituents within each seed jet are re-clustered using a smaller parameter $R = R_{sub}$. A subjet of the corresponding seed jet is then discarded if the ratio between the transverse momentum p_T^i of the subjet i and the transverse momentum p_T^{jet} of the associated seed jet is smaller than a fixed cut value f_{cut}. The remaining subjets are then merged and the output is known as trimmed jet. In this analysis the values $f_{cut} = 0.05$ and $R_{sub} = 0.2$ are used. For small radius jets, no grooming is applied and a different approach known as JetVertexTagger [65] is used to suppress jets that are largely made from particles from pileup.

4.2.4. Jet substructure

As the highly boosted jets cannot be resolved individually, it is necessary to find alternative methods to extract information about the original parton. One possibility is to consider so-called jet substructure variables which provide insight in the jet shape and energy distributions. Substructure variables are designed to allow for discrimination of signal and background processes, each having their own advantages and disadvantages. A powerful variable is denoted N-subjettiness τ_N [66]. First, N subjets are generated within the jet of interest. N-subjettiness is then defined as

$$\tau_N = \frac{\sum\limits_{i \in J} p_{T,i} \min\left(\Delta R_{k,i} | k = 1,...,N\right)}{\sum\limits_{i \in J} p_{T,i}\, R_0}, \tag{17}$$

where the summation is over all jet constituents i of the jet J. τ_N is a measure of how well a jet J can be considered as containing N subjets. Note that τ_N calculates a p_T weighted average distance between each constituent and the nearest of the N subjets. Therefore, low values τ_N indicate good compatibility with an N-subjet hypothesis. As soon as an incompatible constituent is grouped to a subjet, the N-subjettiness rises sharply. As a consequence, τ_{N-1} should be much smaller than τ_N if the jet consists of N subjets. This is also the reason why mostly ratios $\tau_{N,N-1} = \frac{\tau_N}{\tau_{N-1}}$ are studied when N-subjettiness is used. Example distributions of τ_{21} and τ_{32} for Anti-$k_t(R = 1.0)$ jets with $300 < p_T < 400\,\text{GeV}$ are shown in Figure 21.

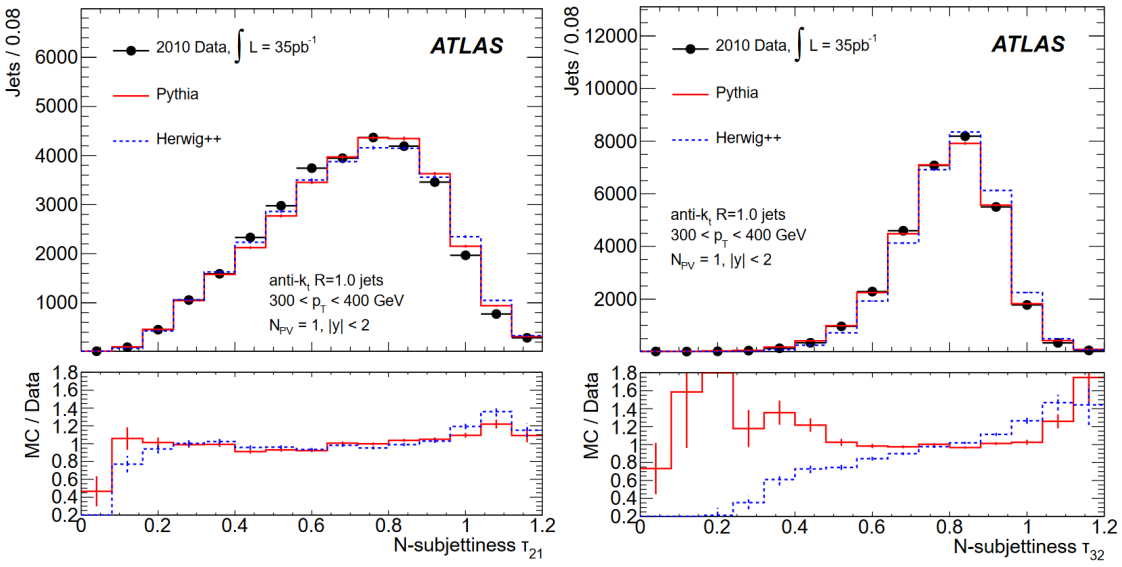


Fig. 21: Distributions of τ_{21} and τ_{32} for Anti-$k_t(R = 1.0)$ jets with $300 < p_T < 400\,\text{GeV}$. [67]

Another set of interesting variables are the so-called energy correlation functions [68]

$$\text{ECF}\,(N,\beta) = \sum_{i_1<i_2<\ldots<i_N\in J} \left(\prod_{a=1}^{N} p_{T_{i_a}}\right) \left(\prod_{b=1}^{N-1}\prod_{c=b+1}^{N} R_{i_b i_c}\right)^{\beta}, \tag{18}$$

where the sum runs over all constituents of a jet J and $R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ is the angular separation as usual. ECFs only use information about the energies and angles of particles within a jet and do not require subjet finding procedures unlike N-subjettiness. However, they yield similar discrimination power as methods involving subjet creation. The parameter β is an arbitrary positive constant and tuned to provide the highest discriminating power. One defines the dimensionless ratio

$$r_N^{(\beta)} = \frac{\text{ECF}(N+1,\beta)}{\text{ECF}(N,\beta)}, \tag{19}$$

as it behaves similarly to τ_N: in the case that the jet indeed consists of N jets, $\text{ECF}(N+1,\beta)$ should be much larger than $\text{ECF}(N,\beta)$. In analogy to $\tau_{N,N-1}$, studies have shown that

$$C_N^{(\beta)} = \frac{r_N^{(\beta)}}{r_{N-1}^{(\beta)}} = \frac{\text{ECF}_{N+1} \times \text{ECF}_{N-1}}{\text{ECF}_N^2} \tag{20}$$

as well as

$$D_N^{(\beta)} = \frac{\mathrm{ECF}_{N+1} \times \mathrm{ECF}_{N-1} \times \mathrm{ECF}_1^N}{\mathrm{ECF}_N^3} \tag{21}$$

are potentially highly discriminating variables.

4.2.5. Track-assisted jet mass

The mass of a jet J reconstructed in a calorimeter is defined as

$$m^{cal} = \sqrt{\left(\sum_{i \in J} E_i\right)^2 - \left(\sum_{i \in J} \vec{p_i}\right)^2}, \tag{22}$$

where the sum goes over all cluster constituent i of the Jet. However, jets solely based on the calorimeters, only provide small information on the origin of the particles. Therefore, an association between calorimeter clusters and particle tracks is performed. In this thesis a procedure known as ghost-matching [69] is applied. The p_T of the tracks is first scaled down by a factor in the order of $\mathcal{O}(10^{-100})$. The tracks are then added to the input list of the corresponding jet algorithm. The scaling ensures that the reconstruction of the jet is not distorted by the addition of tracks. After the procedure, an association of tracks with jet clusters is possible. The so-called track-assisted jet mass is given by

$$m^{TA} = \frac{p_T^{cal}}{p_T^{track}} \cdot m^{track}, \tag{23}$$

where p_T^{track} is the total transverse momentum and m_{track} the invariant mass of the total four-vector of the associated tracks. One further motivation why the track-assisted mass is preferred over m^{cal} is because m^{TA} is less sensitive to the type of simulation that is performed, i.e. AFII or FullSim.

5. Analysis Strategy

In this section, the analysis strategy is presented. The $hh \to \gamma\gamma WW^*$ decay channel is motivated and the different topologies of the channel outlined. Furthermore, the object definition, the event selection and an overlap procedure are described.

5.1. The $hh \to \gamma\gamma WW^*$ decay channel and boosted event topology

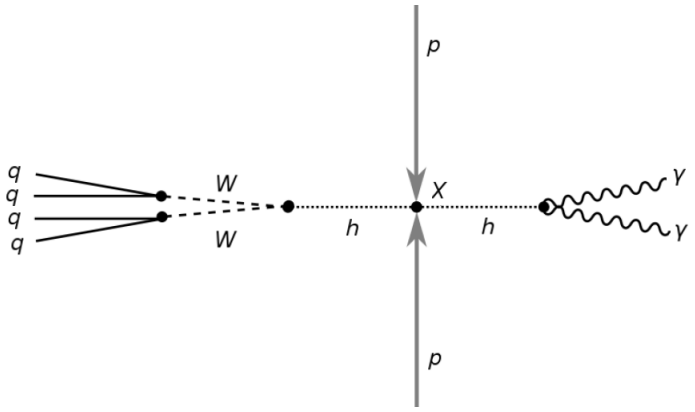


Fig. 22: Di-Higgs production in a proton-proton collision. One Higgs boson decays into two photons and the other one into a pair of W bosons, which then decay hadronically.

In the $hh \to \gamma\gamma WW^*$ decay channel, one Higgs boson decays into two photons and the other one to a pair of W bosons. Only those cases where the W boson decays hadronically as shown in Figure 22 are considered. The channel provides a high resolution and clean signal of the $h \to \gamma\gamma$ decay and on the other hand a high branching ratio for the $h \to WW^*$ process. The branching ratio of the hadronically decaying W boson is around 67 % and thus more common than the leptonic decays. A promising approach for finding new physics such as the $H \to hh$ decay, is the study of boosted topologies. Heavier resonances result in higher Higgs boson p_T, such that the decay products, in this case the SM Higgs bosons, are Lorentz boosted. As a consequence, the decay products will be closer to each other in the lab frame. Often, the resulting jets cannot be resolved using common jet algorithms. In order to study those boosted topologies two different approaches can be taken: either all four quarks can be reconstructed in one large radius jet, known as *fully boosted topology*, or each pair of partons can be reconstructed as one large jet, known as *partially boosted topology*. In addition, it is also possible to use a combination of both topologies.

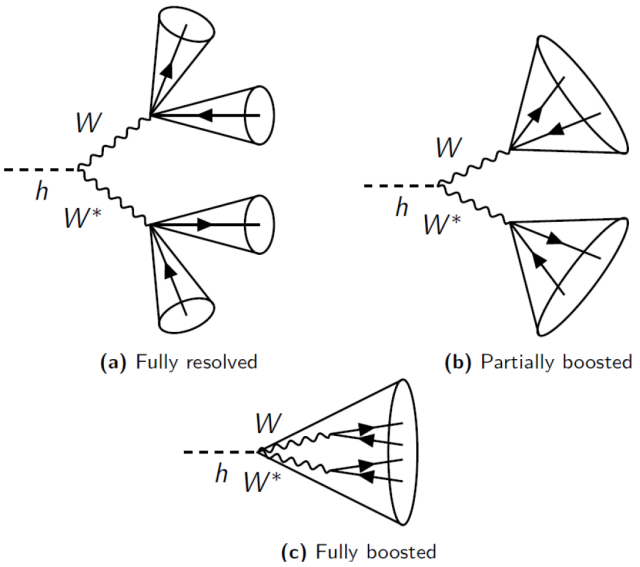


Fig. 23: Sketch of the different topologies of the process $hh \to W_{had}W^*_{had}$. (a) All partons are resolved with small radius jets. (b) The parton pairs surging from the decay of one W boson are reconstructed with one jet each. (c) All partons from the initial Higgs-boson are reconstructed in one large radius jet.

A sketch of the different topologies is shown in Figure 23. Note that even though a fully resolved event is very unlikely with Anti-$k_t(R = 1.0)$ jets, the partially boosted case is not negligible. As a consequence, this analysis was chosen to be inclusive by requiring at least one large radius jet with $R = 1.0$.

5.2. Object definition

The identification (ID) of observable particles using the output signal from the detector is not always trivial. A number of requirements are set to ensure that the considered objects correspond to the actual physical particles. Stringent requirements lead to less background but also discard objects which correspond to the physical particle of interest. Looser requirements are more likely to misidentify objects, but can provide more statistics for the analysis. In the following, each object is defined and how they are reconstructed is described.

5.2.1. Photons

Two possibilities for photon reconstruction exist. The photon is known as converted if it decays into an electron-positron pair $\gamma \to e^+e^-$ before reaching the electromagnetic

calorimeter. In this case the tracks associated to the electron-positron pair production in the inner detector can be matched to the energy depositions in the EM calorimeter surging from the electron and positron. Unconverted photons traverse the inner detector without decaying and directly shower in the EM calorimeter, leaving no tracks in the inner detector. In this analysis the optimisation was performed using 'Loose' and 'Tight' [70] photon identification criteria. Whereas 'Loose' provides an expected efficiency of ∼98 %, 'Tight' is expected to deliver an efficiency of ∼89 %. In combination with the 'Tight' criteria, the isolation working point 'FixedCutLoose' [70] is used. Only photons with $|\eta| < 2.37$ and $|\eta| \notin \{1.37,1.52\}$ are considered as other regions do not provide enough or no sensitivity for photon reconstruction. Note that the exclusion of $1.37 < \eta < 1.52$ is due to the transition between the barrel and end-cap calorimeters. An additional $p_T > 25\,\text{GeV}$ threshold for photons is in place due to poor reconstruction efficiencies of low-p_T photons.

5.2.2. Leptons

Electrons are also reconstructed by matching tracks in the inner detector to energy depositions in the electromagnetic calorimeter. In this analysis, the electron ID 'Medium' [71] with an expected efficiency of ∼90 % and the isolation working point 'Loose' [71] are used. Only electrons with $|\eta| < 2.47$ and $|\eta| \notin \{1.37,1.52\}$ are considered, as tracks can only be reconstructed in a region $|\eta| < 2.47$. An additional $p_T > 10\,\text{GeV}$ threshold for electrons is in place due to poor reconstruction efficiencies of low-p_T electrons.

Muon are first reconstructed independently in the muon calorimeter and in the inner detector. The tracks are then merged together to complete the reconstruction. The muons are identified using the 'Medium' particle ID and the isolation criterion 'GradientLoose' [72]. An efficiency of over 95% is reached. Only muons with $|\eta| < 2.7$ and $p_T > 10\,\text{GeV}$ are considered.

5.2.3. Jets

For the scope of this analysis, Anti-$k_t(R = 1.0)$ jets are used. As only the boosted topology is considered, additional thresholds $|\eta| < 2$ and $p_T > 200\,\text{GeV}$ are in place. As mentioned in section 4.2.3, a jet grooming technique known as trimming is applied with the parameters $f_{cut} = 0.005$ and $R_{sub} = 0.2$. Anti-k_t(R=0.4) jets are used in the overlap removal described in section 5.3. No grooming procedure is applied for small radius jets, but the JetVertexTagger with $JVT < 0.59$ is used. Further requirements

for small radius jets are $p_T > 25\,\text{GeV}$ and $|\eta| < 4.4$. Jets resulting from b quarks are identified using b-tagging. For this scope the tagger MV2C10 [73] with an approximate tagging efficiency of 77 % is used.

5.3. Overlap removal

As the reconstruction algorithms for photons, leptons and jets are independent of each other at the ATLAS experiment, it is possible that one object, i.e. a set of tracks and energy clusters, is associated with more than one physical particle. To account for this effect, a so-called overlap removal is applied. The procedure aims to discard overlapping objects in order to only keep objects that are unambiguously associated with the physical particles. The following procedure was applied:

(i) Electrons are discarded if $\Delta R_{e,\gamma} < 0.4$ for any photon, aiming to remove electrons which are faked by photons.

(ii) Anti-$k_t(R = 0.4)$ jets are discarded if $\Delta R_{J,\gamma} < 0.4$ for any photon, aiming to remove small radius jets faked by photons.

(iii) Anti-$k_t(R = 0.4)$ jets are discarded if $\Delta R_{J,e} < 0.2$ for any electron, aiming to remove small radius jets faked by electrons.

(iv) Leptons are discarded if $\Delta R_{J,l} < \min\left(0.4, 0.04 + \frac{10\,\text{GeV}}{p_{T,l}}\right)$ for any small radius jet, aiming to remove soft leptons that arise from the decay of hadrons in hadronic showers. Only boosted leptons are kept.

(v) Muons are discarded if $\Delta R_{\mu,\gamma} < 0.4$ for any photon, aiming to remove muons which are faked by photons.

(vi) Anti-$k_t(R = 1.0)$ jets are discarded if $\Delta R_{j,\gamma} < 1.0$ for any photon, aiming to remove large radius jets faked by photons.

(vii) Anti-$k_t(R = 1.0)$ jets are discarded if $m_J^{TA} < 15\,\text{GeV}$, aiming to remove large radius jets faked by electrons.

Note that even though this analysis focuses on the topology of hadronically decaying W bosons (see section 5.1), boosted leptons are kept in (iv) to ensure compatibility with possible future semileptonic studies.

5.4. Event selection

Only events satisfying the following selection criteria, hereafter referred to as preselection, are considered in the analysis:

(i) Only events passing a series of data quality checks, namely fulfilling a so-called Good Runs List criterion (GRL) are considered. The GRL ensures that the data is taken in stable beam conditions, with well functioning magnets and detectors.

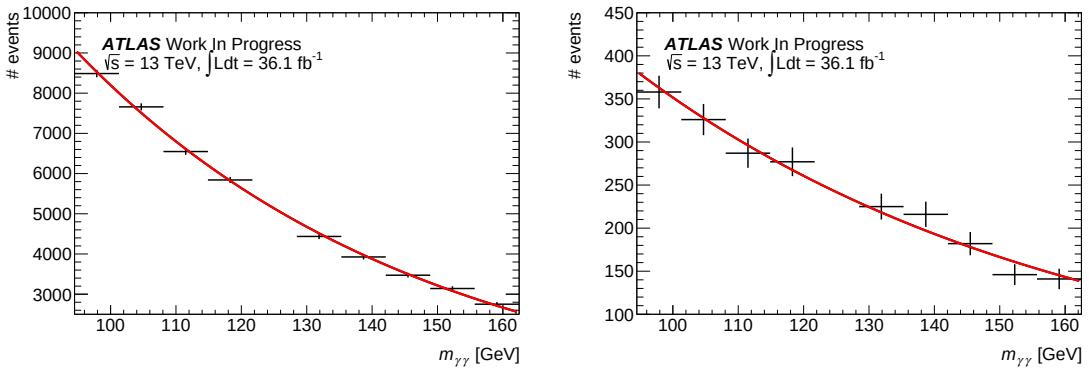
(ii) Only events containing exactly two photons are passed. Furthermore, the photons must fulfill a relative p_T cut, i.e $\frac{p_{T,\gamma_1}}{m_{\gamma\gamma}} > 0.35$ and $\frac{p_{T,\gamma_2}}{m_{\gamma\gamma}} > 0.25$ for $p_{T,\gamma_1} > p_{T,\gamma_2}$. A trigger match criterion ensures that the two photons were registered by the trigger and helps to prevent misidentification. For the loose preselection, no photon isolation is required and the identification criterion 'Loose' discussed in section 5.2 is used. In terms of the tight preselection, the isolation working point 'FixedCutLoose' and the photon ID 'Tight' is in place.

(iii) Only events in the region $94.49\,\text{GeV} < m_{\gamma\gamma} < 162.49\,\text{GeV}$ of the invariant mass spectrum of the diphoton system are considered. This mass region is in the following referred to as fit range (see section 5.5). The interval $125.05 \pm 3.4\,\text{GeV}$ is defined as Higgs mass window and corresponds to $m_h \pm 2\sigma_\gamma$, where σ_γ is the $m_{\gamma\gamma}$ resolution. Events are required to be inside the Higgs mass window or outside, hereafter referred to as sideband or control region.

(iv) Only events containing at least one Anti-$k_t(R = 1.0)$ jet with $|\eta| < 2$ and $p_T > 200\,\text{GeV}$ pass the preselection. As higher jet multiplicities are allowed, an inclusive analysis is performed.

(v) Events containing a lepton, i.e. an electron or muon, are discarded. Furthermore, a vetoe on b-jets is in place.

5.5. Data-driven background estimation



(a) Loose preselection

(b) Tight preselection

Fig. 24: Invariant mass spectrum $m_{\gamma\gamma}$ of the diphoton system, where the signal region is blindend. An exponential fit is performed through the sidebands.

For the optimisation the signal yield s and the background yield b are required (see section 6). As it is not feasible to simulate all processes that contribute to the background, only the most common SM single-Higgs boson production background processes were simulated using MC samples. Therefore, a data-driven method is additionally used in this thesis for background estimation. For this, the decay $h \to \gamma\gamma$ was chosen, as it provides the highest resolution. As the signal is expected to appear in the diphoton mass spectrum around $m_{\gamma\gamma} \approx m_h$, the Higgs mass window introduced in section 5.4 is defined to be the signal region. To estimate the continuous background, the signal region was blinded and a fit is performed through the sideband regions. Because the diphoton mass spectrum is expected to yield a smooth, continuous spectrum, a first order exponential fit function

$$f(m_{\gamma\gamma}) = \exp(a + b \cdot m_{\gamma\gamma}) \tag{24}$$

is used. The integral of the fit function within the signal region then corresponds to the estimated continuous background yield $N_{\gamma\gamma}$. The invariant mass spectrum with the sideband fit for the loose and tight preselection is shown in Figure 24. Note that it is assumed that the exponential form of the $m_{\gamma\gamma}$ spectrum is largely independent of the applied cuts in the optimisation in section 6.

If $[h_+, h_-]$ denotes the Higgs mass window, then the continuous background is estimated with

$$N_{\gamma\gamma} = \int_{h_-}^{h+} f(m_{\gamma\gamma})\, \mathrm{d}m_{\gamma\gamma}. \tag{25}$$

The error can be calculated using the error propagation formula and yields

$$\sigma_{N_{\gamma\gamma}} = \sqrt{N_{\gamma\gamma}^2 \cdot \sigma_a^2 + \frac{1}{b}\left[F\left(h_+, h_-\right) - N_{\gamma\gamma}\right] \cdot \sigma_b^2 + \frac{2N_{\gamma\gamma}}{b}\left[F\left(h_+, h_-\right) - N_{\gamma\gamma}\right] \cdot \sigma_{ab}^2}, \quad (26)$$

where $F\left(h_+, h_-\right) \equiv h_+ f(h_+) - h_- f(h_-)$ and σ_{ab} denotes the uncertainty on the covariance between a and b.

6. Optimisation

The optimisation in this thesis was performed by tuning the selection criteria to maximize the Cowan significance [74]

$$\Sigma = \sqrt{2 \cdot \left((s+b) \ln\left(1 + \frac{s}{b}\right) - s \right)} \xrightarrow{\frac{s}{b} \ll 1} \Sigma = \frac{s}{\sqrt{s+b}} \tag{27}$$

Here, s denotes the signal yield from the MC signal and b the total background yield, which is constituted by the sum of the MC background and the continuous background estimation. The corresponding error is given by

$$\sigma_\Sigma = \sqrt{\frac{1}{\Sigma} \cdot \left[\left[\ln\left(1 + \frac{s}{b}\right) \cdot \sigma_s \right]^2 + \left[\left(\ln\left(1 + \frac{s}{b}\right) - \frac{s}{b} \right) \cdot \sigma_b \right]^2 \right]}, \tag{28}$$

where σ_s is the error on the signal yield and $\sigma_b = \sqrt{\sigma^2_{N_{ggF}} + \sigma^2_{N_{vbf}} + \sigma^2_{N_{tth}} + \sigma^2_{N_{wh}} + \sigma^2_{N_{zh}} + \sigma^2_{N_{\gamma\gamma}}}$ the uncertainty on the total background yield. To maximize the significance, the separation power of potentially high discriminating variables was studied. The following variables were considered:

(a) Kinematic variables

  (i) Transverse momentum $p_{T,\gamma\gamma}$ of the diphoton system, $p_{T,J}$ of the jet and the ratio $\frac{p_{T,\gamma\gamma}}{p_{T,J}}$.

  (ii) The angular separation $\Delta R_{\gamma\gamma}$ of both photons as well as variables related to the angular separation of the jet and the diphoton system, i.e. $\Delta\phi_{\gamma,J}$, $\Delta\eta_{\gamma,J}$ and $\Delta R_{\gamma,J}$.

  (iii) The track assisted mass of the jet, in the following simply denoted as m_J.

(b) Substructure variables

  (i) The N-subjettiness variables τ_{21}, τ_{31}, τ_{32}, τ_{41}, τ_{42} and τ_{43}.

  (ii) The energy correlation function related variables C_1, C_2, C_3, C_4 and D_1, D_2, D_3, D_4.

The optimisation is performed using the $m_H = 750\,\text{GeV}$ sample and later applied to all other resonant mass points as well as to the non-resonant sample. Typically, resolved analysis are relevant up to masses of 500 GeV. By optimising on the 750 GeV sample, it is aimed to ensure that the cuts which are then applied to higher resonance masses are looser and do not cut away too much signal. For the non-resonant sample, the predicted

SM cross section was assumed. For the resonant samples, a cross section of 1 pb was assumed, as no theoretical predictions exist and the cross-section has not been excluded in Run I. To maximize the significance, four hard cuts in the loose and three hard cuts on the tight selection were applied. In this context, hard cuts are cuts on values which yield significances very near to the maximum. As the variables are not independent of each other, previous optimal cuts can change after cutting on further variables. To account for this, the cuts were re-optimised after all iterations, i.e. the first variable is re-optimised with all other cuts set, then the second variable is re-optimised with the new value of the first variable and the other cuts unchanged and so forth. This ensures that the calculated maximum significance is near to the true maximum value. Note that only the gluon-gluon fusion background process from the single $h \to \gamma\gamma$ SM background is depicted in the following distribution plots. As it is the main single Higgs boson SM background contribution, the other processes would not significantly distort the distributions. For the distribution of the continuous background, it is assumed that the sidebands can be used to interpolate into the signal region, i.e. the shapes in the sidebands are largely independent of $m_{\gamma\gamma}$. The validity of this assumption has been confirmed in previous studies [75]. Note that the significance is not calculated if $N_{sig} = 0$ or $N_H = 0$ and $N_{\gamma\gamma} = 0$ or $\sigma_{N_{\gamma\gamma}}/N_{\gamma\gamma} > 0.5$ as a cut becomes unreasonable.

6.1. Loose photon selection

Iteration 1

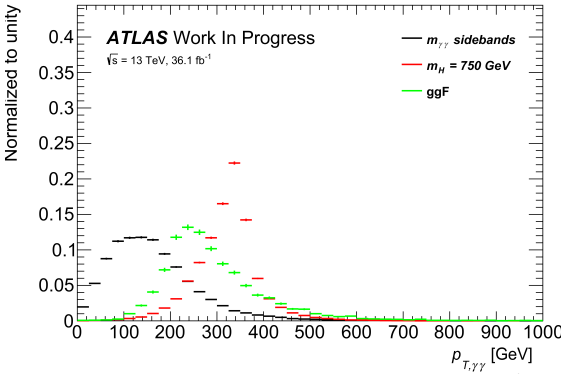


Fig. 25: Normalized distribution of the transverse momentum $p_{T,\gamma\gamma}$ of the diphoton system after the loose preselection.

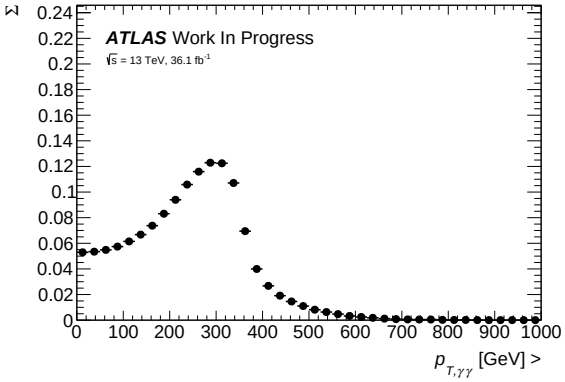


Fig. 26: Significance as a function of a cut $p_{T,\gamma\gamma} > X$ after the loose preselection.

Not surprisingly, the transverse momentum of the diphoton system $p_{T,\gamma\gamma}$ provides the best separation power in the first iteration. As the resonant mass point results in higher photon p_T compared to single Higgs production modes and continuous background, a good background suppression can be reached (see Figure 25). As can be seen from Figure 26, the best significance was achieved for a cut $p_{T,\gamma\gamma} > 275\,\text{GeV}$. The SM background was reduced by approximately 57% from 39.20 to 17.36 events, whereas the expected continuous background was reduced by 86% from 5129.76 to 584.59 events and around 21% of the signal events were lost. The significance increased from 0.05 to 0.12.

Iteration 2

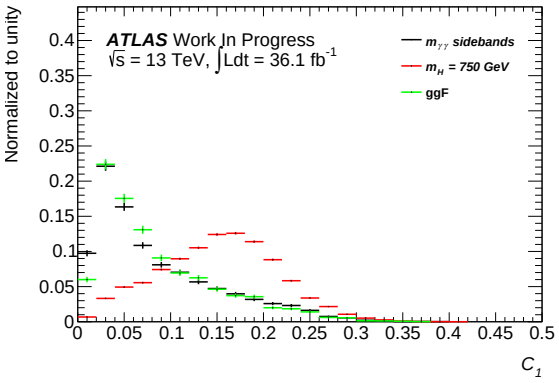


Fig. 27: Normalized distribution of C_1 after the loose preselection and $p_{T,\gamma\gamma} > 275\,\text{GeV}$.

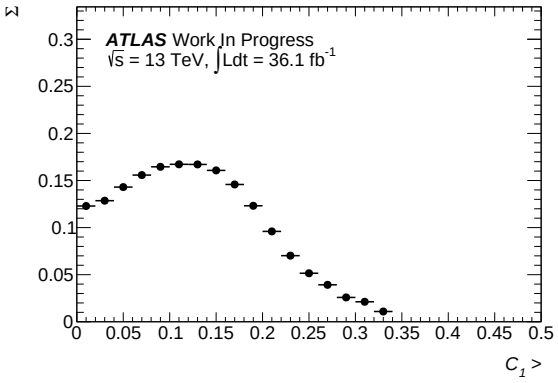


Fig. 28: Significance as a function of a cut $C_1 > X$ after the loose preselection and $p_{T,\gamma\gamma} > 275\,\text{GeV}$.

In the second iteration, C_1 was the best candidate. Whereas the signal sample distribution is shifted to higher values, peaking around $C_1 \approx 0.17$, the background samples tended to peak for lower values around 0.03 (see Figure 27). The highest increase in significance was observed for $C_1 > 0.1$ by approximately 41% to 0.17 (see Figure 28). The SM Higgs contributions was reduced from 17.36 events to 5.18 events by 70%. The continuous background was reduced by 67% and the signal was decreased from 3.02 to 2.35 events by approximately 22%.

Iteration 3

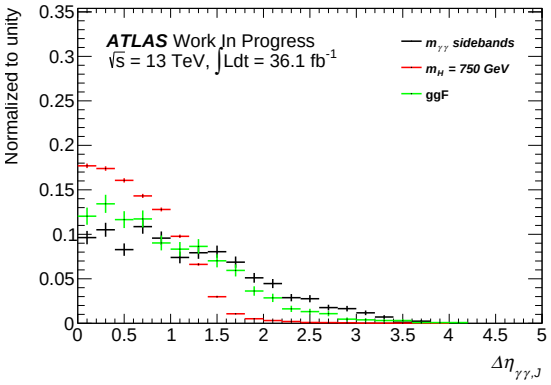


Fig. 29: Normalized distribution of $\Delta\eta_{\gamma\gamma,J}$ after the loose preselection, $p_{T,\gamma\gamma} > 275\,\text{GeV}$ and $C_1 > 0.1$.

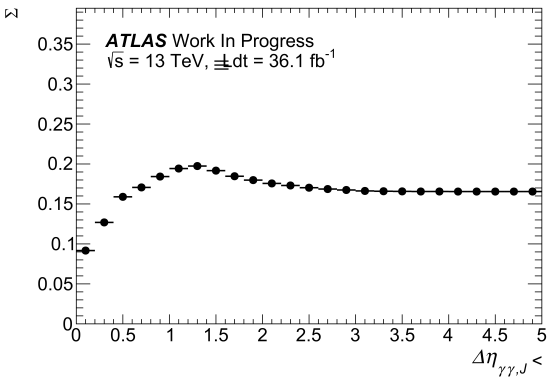


Fig. 30: Significance as a function of a cut $\Delta\eta_{\gamma\gamma,J} < X$ after the loose preselection, $p_{T,\gamma\gamma} > 275\,\text{GeV}$ and $C_1 > 0.1$.

In the third iteration, $\Delta\eta_{\gamma\gamma,J}$ was the most promising cut variable. Even though both distributions peak at $\Delta\eta_{\gamma,\gamma,J} = 0$ and overlap considerably, the background distributions seem to be flatter, whereas the signal sample shows a steeper distribution (see Figure 29). This allows a good separation with the highest significance for a cut $\Delta\eta_{\gamma\gamma,J} < 1.4$ (see Figure 30). Only approximately 5.5% of the signal was discarded, whereas 36% of the continuous and 13% of the SM Higgs background was rejected. The significance was increased by 15% to 0.20.

Iteration 4

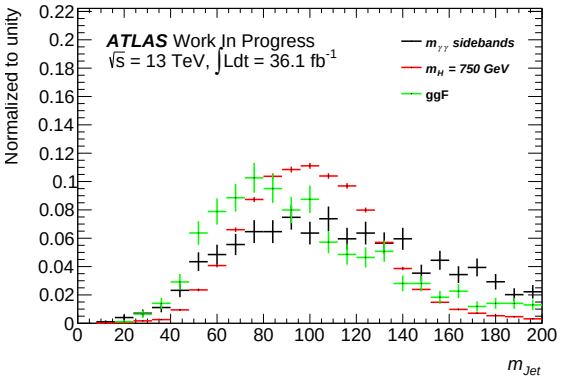


Fig. 31: Normalized distribution of m_J after the loose preselection, $p_{T,\gamma\gamma} > 275\,\text{GeV}$, $C_1 > 0.1$ and $\Delta\eta_{\gamma\gamma,J} < 1.4$.

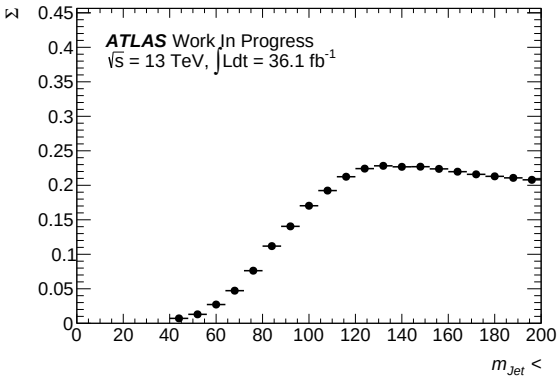


Fig. 32: Significance as a function of a cut $m_J < X$ after the loose preselection, $p_{T,\gamma\gamma} > 275\,\text{GeV}$, $C_1 > 0.1$ and $\Delta\eta_{\gamma\gamma,J} < 1.4$.

In the last iteration the track assisted mass of the jet could provide the best separation power. As can be seen from Figure 31, the sideband events are more evenly distributed and surprisingly there are significantly more background events in the high mass region than signal events. A maximum significance was achieved for $m_J < 150\,\text{GeV}$ (see Figure 32), reducing the signal yield from 2.22 to 2.09 by approximately 6%, the continuous background by 36% and the SM Higgs background by 13%.

Further cuts could only marginally improve the significance and would lead to non-continuous significance curves. As this is a sign that statistical fluctuations affect the significance calculations, no additional cuts were implemented. As mentioned before, a iterative re-optimisation was performed and the highest significance was achieved for $p_{T,\gamma\gamma} > 300\,\text{GeV}$, $C_1 > 0.14$, $\Delta\eta_{\gamma\gamma,J} < 1.2$ and $m_J < 136\,\text{GeV}$. As a consequence, all cuts were tightened and a maximal significance of 0.25 achieved. Through the re-optimisation additional 78% of the continuous background and further 72% of the SM Higgs background where rejected, while losing approximately 44% of the signal events.

Iteration	Cut	$N_{\gamma\gamma}$	N_H	N_{sig}	Σ_{max}
0	none	5129.76 ± 24.31	39.20 ± 8.14	3.81 ± 0.03	$0.05 \pm 8.9 \cdot 10^{-5}$
1	$p_{T,\gamma\gamma} > 275\,\text{GeV}$	584.59 ± 8.14	17.36 ± 0.21	3.02 ± 0.02	$0.12 \pm 4.4 \cdot 10^{-4}$
2	$C_1 > 0.1$	191.50 ± 4.66	5.18 ± 0.11	2.35 ± 0.02	$0.17 \pm 1.0 \cdot 10^{-3}$
3	$\Delta\eta_{\gamma\gamma,J} < 1.4$	122.12 ± 3.72	3.66 ± 0.09	2.22 ± 0.02	$0.20 \pm 1.5 \cdot 10^{-3}$
4	$m_J < 150\,\text{GeV}$	89.72 ± 3.19	3.17 ± 0.08	2.09 ± 0.02	$0.22 \pm 1.9 \cdot 10^{-3}$
Final		20.14 ± 1.51	0.89 ± 0.05	1.17 ± 0.01	$0.25 \pm 4.7 \cdot 10^{-3}$

Tab. 7: Summary of the relevant values after each iteration for the loose preselection.

Summarising, the continuous background was reduced from 5129.76 to 20.14 events by 99.6%, whereas the SM Higgs background was reduced from 39.20 to 0.89 events by 97.7%. Therefore, a total background rejection of 99.59% was accomplished. From 3.81 initial signal events, 1.17 events remain after the optimisation, corresponding to a signal loss of 69.29%. A maximal significance of 0.252 ± 0.005 in the loose selection was achieved. A summary of the relevant values after each iteration and of the final selection is shown in Table 7.

6.2. Tight photon selection

Iteration 1

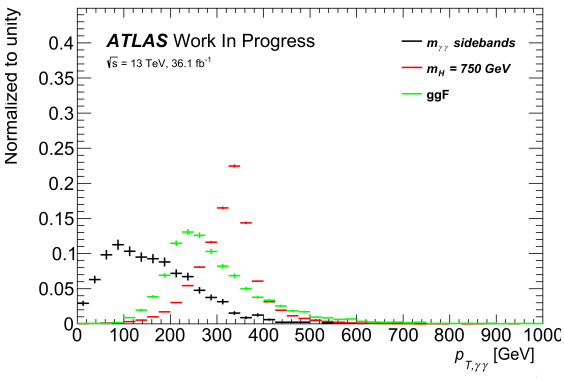


Fig. 33: Normalized distribution of $p_{T,\gamma\gamma}$ after the tight preselection.

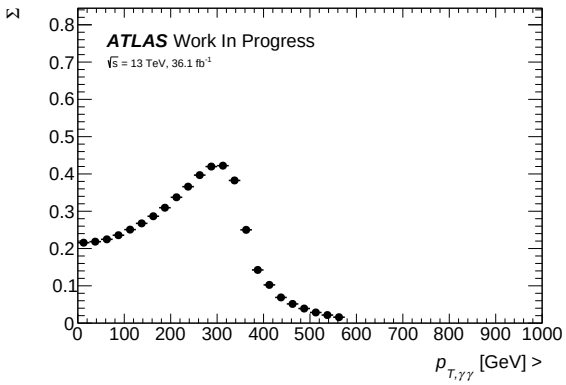


Fig. 34: Significance as a function of a cut $p_{T,\gamma\gamma} > X$ after the tight preselection.

As for the loose selection, $p_{T,\gamma\gamma}$ provided the best separation power in the first iteration (see Figure 33). The best cut value is $p_{T,\gamma\gamma} > 300\,\text{GeV}$ and rejects 91% of the continuous and 64% of single Higgs boson SM background by discarding 31% of the signal events.

The significance was increased from an initial value of 0.22 to 0.42 by almost 91% (see Figure 34).

Iteration 2

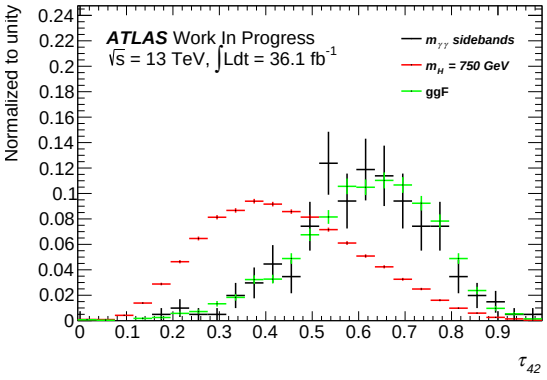


Fig. 35: Normalized distribution of τ_{42} after the tight preselection and $p_{T,\gamma\gamma} > 300\,\text{GeV}$.

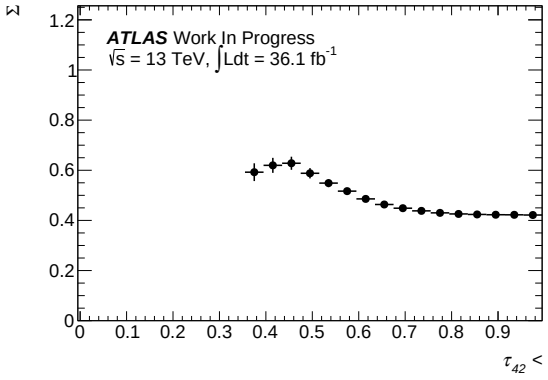


Fig. 36: Significance as a function of a cut $\tau_{42} < X$ after the tight preselection and $p_{T,\gamma\gamma} > 300\,\text{GeV}$

In the second iteration, the N-subjettiness variable τ_{42} could provide the best separation power. As one can see from Figure 35, the distributions for the $m_{\gamma\gamma}$ sidebands and the ggF sample are shifted to the right. Whereas the background distributions are peaking around 0.65, the signal sample has its maximum for lower values around 0.38. The significance was maximized by 50% to a value of 0.63 for a cut $\tau_{42} < 0.475$ (see Figure 36). The continuous and single Higgs boson background were rejected by approximately 87% and 81% respectively, while also discarding 0.98 signal events corresponding to a approximate 40% percentage decrease.

Iteration 3

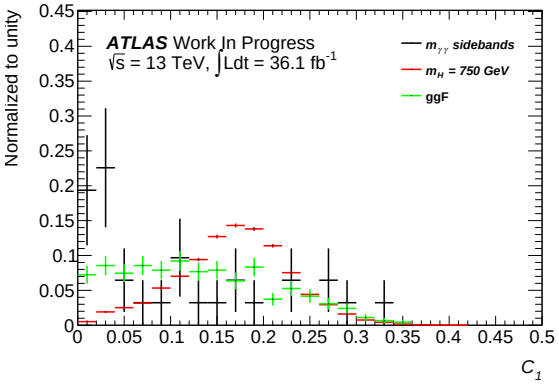


Fig. 37: Normalized distribution of C_1 after the tight preselection, $p_{T,\gamma\gamma} > 300\,\text{GeV}$ and $\tau_{42} < 0.475$.

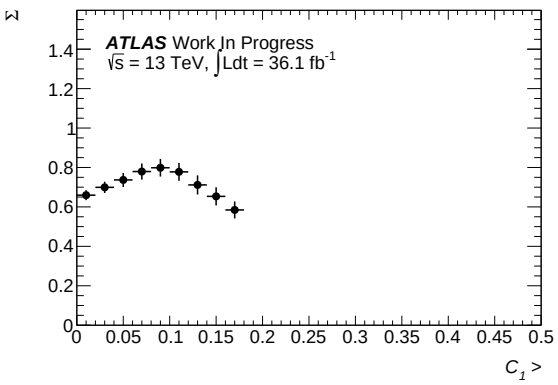


Fig. 38: Significance as a function of a cut $C_1 > X$ after the tight preselection, $p_{T,\gamma\gamma} > 300\,\text{GeV}$ and $\tau_{42} < 0.475$.

For the last iteration in the tight selection, the substructure variable C_1 again led to the maximal increase in significance. One can see from Figure 37, that the uncertainties on the distributions increase significantly due to low statistics. However, the significance curve in Figure 38 remains smooth, wherefore it was opted for a further loose $C_1 > 0.08$ cut. Approximately 44% of the continuous and 40% of the single Higgs boson SM background are rejected, while discarding only 8% of the signal events. Consequently, the significance increases by 17% to 0.74.

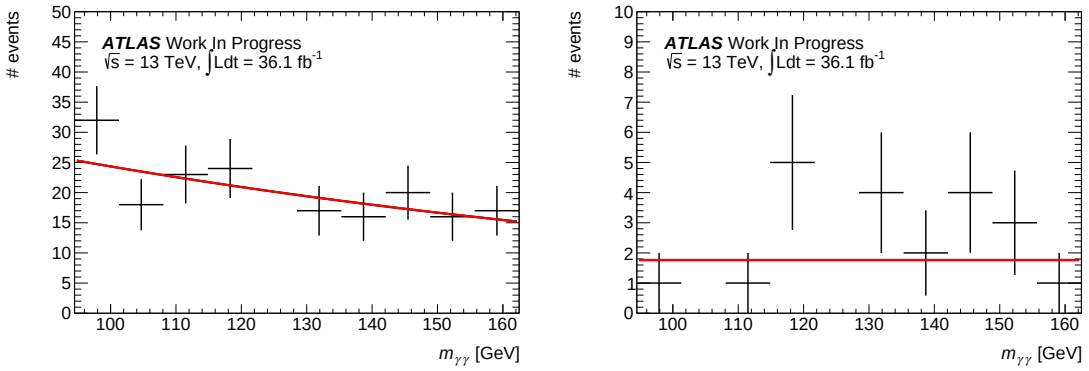
The final re-optimisation led to a looser $p_{T,\gamma\gamma} > 275\,\text{GeV}$ cut, while the other cuts remained identical. The continuous background increased from 1.61 events to 1.76 by 9.3%, the SM background increased by 27.7% from 1.26 events to 1.61. However, the signal events also increased by 17% from 1.34 to 1.57 events, leading to a final significance of 0.8. Due to increasing statistical fluctuations, further cuts were not feasible. Also, the fit quality for the continuous background estimation would decline significantly.

Iteration	Cut	$N_{\gamma\gamma}$	N_H	N_{sig}	Σ_{max}
0	none	241.69 ± 5.24	32.95 ± 0.29	3.58 ± 0.03	$0.22 \pm 1.2 \cdot 10^{-3}$
1	$p_{T,\gamma\gamma} > 300\,\text{GeV}$	20.90 ± 1.52	11.86 ± 0.18	2.44 ± 0.02	$0.42 \pm 6.7 \cdot 10^{-3}$
2	$\tau_{42} < 0.475$	2.67 ± 0.55	2.24 ± 0.07	1.46 ± 0.02	$0.63 \pm 2.6 \cdot 10^{-2}$
3	$C_1 > 0.08$	1.61 ± 0.57	1.26 ± 0.05	1.34 ± 0.02	$0.74 \pm 5.6 \cdot 10^{-2}$
Final		1.76 ± 0.47	1.61 ± 0.06	1.57 ± 0.02	$0.80 \pm 4.5 \cdot 10^{-2}$

Tab. 8: Summary of the relevant values after each iteration for the tight preselection.

Summarising, the continuous background was reduced from 241.69 to 1.76 events by 99.3%, whereas the single Higgs boson background was reduced from 32.95 to 1.61 events by 95.1%. Therefore, a total background rejection of 98.8% was accomplished. From 3.58 initial signal events, 1.57 events remain after the optimisation, corresponding to a signal loss of 56.1%. A maximal significance of 0.799 ± 0.045 in the tight selection was achieved. A summary of the relevant values after each iteration and of the final selection is shown in Table 8.

6.3. Continuous background estimation



(a) Loose preselection (b) Tight preselection

Fig. 39: Sideband fit for the invariant mass spectrum $m_{\gamma\gamma}$ of the diphoton system with all corresponding cuts described in section 6 applied.

The final exponential fits for the loose and tight preselection are shown in Figure 39. To investigate the quality of the fit, it is useful to calculate

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}, \tag{29}$$

where O_i is the observed number of events in bin i and E_i the expected number of events in bin i based on the fit. The goodness of the fit can be estimated by studying the reduced χ^2 ratio χ^2/NDf, where NDf denotes the number of degrees of freedom. A reduced χ^2 value around one indicates a good fit quality [76]. In Table 9 and 10 the parameters of the fits and the reduced χ^2 are listed after each optimisation iteration for the loose and tight selection respectively.

Iteration	Cut	a	b	red. χ^2
0	none	10.88 ± 0.03	-0.019 ± 0.001	1.09
1	$p_{T,\gamma\gamma} > 275\,\text{GeV}$	7.89 ± 0.09	-0.012 ± 0.001	0.40
2	$C_1 > 0.1$	6.80 ± 0.15	-0.012 ± 0.001	0.69
3	$\Delta\eta_{\gamma\gamma,J} < 1.4$	6.44 ± 0.20	-0.013 ± 0.002	1.28
4	$m_J < 150\,\text{GeV}$	5.95 ± 0.23	-0.012 ± 0.002	1.12
Final		3.95 ± 0.49	-0.008 ± 0.004	0.68

Tab. 9: Development of the fit parameters and the reduced χ^2 after each iteration and the final fit for the loose selection.

For the loose selection, the final fit parameters were determined to be $a = 3.95 \pm 0.49$ and $b = -0.008 \pm 0.004$ with a reduced χ^2 value of approximately 0.7. As can be seen from Table 9, the reduced χ^2 parameter seems to fluctuate around one and thus indicates a relatively good fit quality throughout the cuts.

Iteration	Cut	a	b	red. χ^2
0	none	7.69 ± 0.29	-0.017 ± 0.002	0.62
1	$p_{T,\gamma\gamma} > 300\,\text{GeV}$	4.81 ± 1.01	-0.011 ± 0.007	0.56
2	$\tau_{42} < 0.475$	1.65 ± 2.93	-0.003 ± 0.039	0.21
3	$C_1 > 0.08$	4.81 ± 4.07	-0.030 ± 0.030	0.53
Final		0.57 ± 0.27	0	0.46

Tab. 10: Development of the fit parameters and the reduced χ^2 after each iteration and the final fit for the tight selection.

For the tight criteria, a worse fit quality was expected due to lower statistics. In fact, the original fit of the $m_{\gamma\gamma}$ spectrum in Figure 39 (b) resulted in a positive parameter b and thus in a rising diphoton mass spectrum. This seems to result from the fact that the two first bins only contain one event, even though they are expected to yield more events than any other bins. Whether this effect is a systematic error or a simple statistical fluctuation should be investigated in the future. In order to still calculate a sensible background estimation, the exponential function was constrained to give a decreasing diphoton spectrum, i.e. the parameter b was constrained to be zero or negative. According to expectation, this led to $b \approx 0$ and thus an almost constant diphoton mass distribution. The final fit parameters were determined to be $a = 0.57 \pm 0.27$ and $b = -1.6 \cdot 10^{-9} \pm 9.8 \cdot 10^{-3}$ with a reduced χ^2 of 0.46. The parameter b can as a consequence considered to be 0 and the resulting fit a simple constant. Note that as this

being a low statistics analysis, the fits were already expected to yield high uncertainties. However, the uncertainties on the expected continuous background yield $\sigma_{N_{\gamma\gamma}}$ remain reasonable (see Table 7 and 8).

7. Results

In the following chapter the signal yields and the final background estimation are presented. The signal region is unblinded and limits are set on $\sigma(gg \to H) \times \mathrm{BR}(H \to hh)$ and $\sigma(gg \to hh)$ for resonant and non-resonant Higgs boson pair production respectively.

7.1. Cutflow and event yields

Sample	Event yield
Cont. bkg.	20.138 ± 1.509
ggF	0.468 ± 0.037
VBF	0.202 ± 0.024
WH	0.087 ± 0.007
ZH	0.080 ± 0.004
$t\bar{t}H$	0.055 ± 0.004
Σ_{bkg}^{loose}	21.030 ± 1.510
0.5 TeV	0.009 ± 0.001
0.75 TeV	1.166 ± 0.015
1 TeV	0.810 ± 0.012
1.5 TeV	0.091 ± 0.005
2 TeV	0.033 ± 0.003
2.5 TeV	0.018 ± 0.001
3 TeV	0.010 ± 0.001

Tab. 11: Event yields for the loose selection

Sample	Event yield
Cont. bkg.	1.765 ± 0.470
ggF	0.783 ± 0.051
VBF	0.460 ± 0.032
WH	0.141 ± 0.009
ZH	0.099 ± 0.007
$t\bar{t}H$	0.126 ± 0.006
Σ_{bkg}^{tight}	3.376 ± 0.474
0.5 TeV	0.021 ± 0.002
0.75 TeV	1.570 ± 0.017
1 TeV	2.343 ± 0.020
1.5 TeV	2.482 ± 0.020
2 TeV	1.664 ± 0.018
2.5 TeV	0.922 ± 0.012
3 TeV	0.545 ± 0.009

Tab. 12: Event yields for the tight selection

The event yields for the signal samples as well as background contributions are shown in Table 11 and Table 12 for the final loose and tight selection respectively. Note that the signal samples $m_H < 500\,\mathrm{GeV}$ are not included as basically all events are lost after applying the optimisation of the 750 GeV mass point. For the loose selection, the total background estimation was determined to be

$$\Sigma_{bkg}^{loose} = 21.030 \pm 1.510\,(\text{stat-only}). \tag{30}$$

For the tight selection, the background yield was calculated to

$$\Sigma_{bkg}^{tight} = 3.376 \pm 0.474\,(\text{stat-only}). \tag{31}$$

Note, however, that only statistical uncertainties were considered. The detailed cutflow tables for the final selection are shown in Table 13 and 14 for the loose and tight selection respectively.

Cut	$m_{\gamma\gamma}$ sidebands	MC-bkg	non-reso.	750 GeV	unbl. data
N_{DxAOD}	111112680	3485.7076	0.3844	12.4359	111112680
DQ	97814823	3346.1636	0.3670	12.1078	97814823
2 loose photons	29866531	2772.3157	0.2957	9.8296	29866531
trigger match	29794061	2766.7179	0.2945	9.7793	29794061
rel. $p_{T,\gamma}$ cuts	23805404	2511.8504	0.2655	9.2516	23805404
fit range	8288643	2507.1722	0.2640	9.2026	8288643
$m_{\gamma\gamma} = 125.09$	7568261	2227.9247	0.2230	8.1099	720382
> 1 large-R jet	49501	44.50144	0.0361	4.5901	5517
b-veto	46463	39.8963	0.0351	4.4764	5175
lepton veto	46262	39.2037	0.0300	3.8094	5153
$p_{T,\gamma\gamma} > 300$ GeV	3792	13.7192	0.0099	2.5742	412
$C_1 > 0.14$	740	2.2017	0.0049	1.5081	80
$\Delta\eta_{\gamma\gamma,J} < 1.2$	423	1.4381	0.0042	1.3913	50
$m_J < 136$ GeV	183	0.8922	0.0033	1.1661	25

Tab. 13: Cutflow for the final selection in the loose topology for the $m_{\gamma\gamma}$ sidebands, Monte-Carlo background, non-resonant sample, resonant $m_H = 750$ GeV sample and the data in the unblinded signal region.

The detailed cutflows for the other resonant samples can be found in the appendix A.2. After the event and background yields were fixed, the signal region was unblinded. For the loose selection, 25 events were observed in the signal regions after all corresponding cuts were applied (see Table 13). Around four additional events are observed than expected in the background-hypothesis only. However, the number is consistent within uncertainties. Note that the final optimal selection for the $m_H = 750$ GeV mass point leads to a very high signal loss for the other resonant samples (see appendix A.2). The reason for this, seems to be the last cut on m_J. Whereas only approximately 16% of the signal is lost for the 750 GeV sample, 75% of the events are lost for the 1.5 TeV sample, with an increasing signal loss for higher resonances. It therefore might be sensible to remove this cut when studying boosted resonances. The background yields excluding the m_J cut in the loose selection are shown in appendix A.3. The total background yield excluding the m_J cut was determined to 47.945 ± 2.302. A total of 50 events were observed in the signal region, being consistent with the background-only hypothesis.

Cut	$m_{\gamma\gamma}$ sidebands	MC-bkg	non-reso.	750 GeV	unbl. data
N_{DxAOD}	111112680	3485.7076	0.3844	12.4359	111112680
DQ	97814823	3346.1636	0.3670	12.1078	97814823
2 loose photons	29866531	2772.3157	0.2957	9.8296	29866531
trigger match	29794061	2766.7179	0.2945	9.7793	29794061
tight ID	4422423	2403.5774	0.2945	9.7793	4422423
isolation	1635695	2156.1191	0.2616	8.9808	1635695
rel. $p_{T,\gamma}$ cuts	1427304	1987.0086	0.2389	8.5893	1427304
fit range	490962	1986.2892	0.2385	8.5792	490962
$m_{\gamma\gamma} = 125.09$	445828	1800.4735	0.2040	7.6050	45134
> 1 large-R jet	2331	37.1145	0.0336	4.3145	284
b-veto	2176	33.4883	0.0327	4.2076	264
lepton veto	2158	32.9486	0.0274	3.5819	260
$p_{T,\gamma\gamma} > 275\,\text{GeV}$	283	14.9583	0.0123	2.8562	41
$\tau_{42} < 0.475$	42	2.8261	0.0071	1.7091	14
$C_1 > 0.08$	21	1.6112	0.0063	1.5701	7

Tab. 14: Cutflow for the final selection in the tight topology for the $m_{\gamma\gamma}$ sidebands, Monte-Carlo background, non-resonant sample and resonant $m_H = 750\,\text{GeV}$ sample and the data in the unblinded signal region.

For the tight selection, a small excess of 7 events was observed in the signal region compared to 3.376 ± 0.474 background events after all corresponding cuts were applied (see Table 14). The result is consistent within a (stat-only) 2σ confidence level.

7.2. Limits

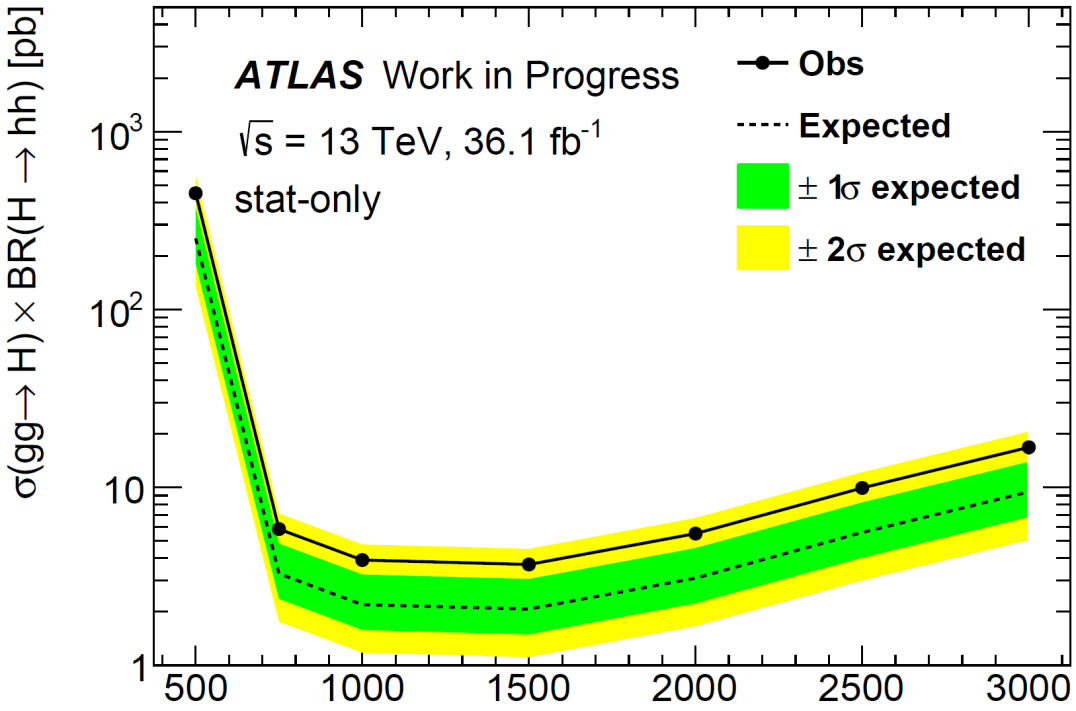


Fig. 40: Observed and expected upper limits for various values of m_H. Only statistical uncertainties are considered.

As no evidence for di-Higgs boson production has been observed, limits on $\sigma(gg \to H) \times \mathrm{BR}(H \to hh)$ and $\sigma(gg \to hh)$ were set with the HistFactory [77] package for resonant and non-resonant production respectively. Because the tight channel offered a more promising significance, only limits for the tight selection were calculated. Figure 40 and Table 15 show the observed and expected limits for non-resonant production as well as for the resonant mass points in a range $0.5\,\mathrm{TeV} < m_H < 3\,\mathrm{TeV}$.

	Non-res.	0.5 TeV	0.75 TeV	1 TeV	1.5 TeV	2 TeV	2.5 TeV	3 TeV
Expected	27.16	251.22	3.27	2.19	2.07	3.09	5.58	9.44
Observed	48.37	450.25	5.83	3.90	3.69	5.50	9.93	16.80
$+2\sigma$	59.24	566.71	7.14	4.78	4.51	6.73	12.16	20.58
$+1\sigma$	40.11	375.23	4.83	3.24	3.06	4.56	8.23	13.94
-1σ	19.57	181.02	2.36	1.58	1.49	2.23	4.02	6.80
-2σ	14.58	134.84	1.76	1.18	1.11	1.66	2.99	5.07

Tab. 15: Observed and expected upper limits in units of picobarns for resonant and non-resonant production (stat-only)

Note that only statistical uncertainties were considered. All observed limits are within the 2σ region of the expected limits. Upper limits for non-resonant production are at $27.16^{+40.11}_{-19.57}$ pb (expected) and 48.37 (observed). The most stringent limits for the resonant samples could be set for $m_H = 1.5\,\text{TeV}$ with $3.69^{+3.06}_{-1.49}$ pb (expected) and 2.07 pb (observed). Note that no sensible limit calculation was possible for $m_H < 500\,\text{GeV}$, as below this point, too few signal events pass the cut selection. This is due to the fact, that a resolution with one large R-jet becomes infeasible for lower mass points, especially because the Anti-$k_t(R = 1.0)$ jets are required to have a minimum transversal momentum of 200 GeV. This also explains why upper limits decrease for the 1 TeV and 1.5 TeV samples: as the partons are more boosted, they are closer to each other and a large R-jet reconstruction is easier. Because the photons are also boosted, they are equally close to each other in the lab frame. Due to the limited resolution of the detector, this will lead to a poorer identification of the photons and could therefore explain the rise of the upper limits for even higher mass samples.

8. Conclusion and Outlook

A search for non-resonant and resonant Higgs boson pair production in the $hh \to \gamma\gamma W_{had}W^*_{had}$ decay channel based on 36.1 fb^{-1} of collision data recorded by the ATLAS experiment at the LHC in 2015 and 2016 at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ was presented. No significant excess of di-Higgs boson production with respect to the SM background-only hypothesis is observed. Upper limits were set on $\sigma(gg \to H) \times \text{BR}(H \to hh)$ and $\sigma(gg \to hh)$ for the resonant and non-resonant Higgs boson pair production respectively. Upper limits for non-resonant production are at $27.16^{+40.11}_{-19.57}\,\text{pb}$ (expected) and 48.37 pb (observed). The most stringent limits for the resonant samples could be set for $m_H = 1.5\,\text{TeV}$ with $3.69^{+3.06}_{-1.49}\,\text{pb}$ (expected) and 2.07 pb (observed).

Note that this analysis only takes into account statistical uncertainties. A full analysis also requires considering systematic uncertainties, which are expected to contribute to a non-negligible amount to the total errors. The observation that the $m_{\gamma\gamma}$ spectrum in the continuous background estimation seems to indicate an unexpected downward fluctuation in the first two bins, is to be investigated further. Evidently, the uncertainty on the continuous background estimation is the major contributor to the final background yield error. It is therefore essential to ensure a good fit quality. As this being a low statistics analysis, more data from the LHC in the coming years will enhance the background estimation. Also, low statistics prevent the implementation of more cuts, as at some point, the exponential fit becomes infeasible. More data may allow for additional cuts or more sophisticated methods such as boosted decision trees which could further enhance the background subtraction.

In addition to using more data, the combination of different decay channels for analysis with low statistics is crucial. For instance, the semileptonic channel $WW^* \to qq\ell\nu$, seems to be a promising candidate. Even though it offers a slightly lower branching ratio, the lepton will considerably improve the event identification.

The LHC is expected to provide approximately $3000\,\text{fb}^{-1}$ of integrated luminosity by 2030. With this amount of data, combined analysis should become sensitive to non-resonant di-Higgs production as well as to heavy resonances. It can therefore expected that studies on Higgs boson self-coupling as well as boosted mass points will become increasingly important in a near future.

A. Appendix

A.1. Monte Carlo samples

Mass point m_H [GeV]	Signal sample name
non-resonant	mc15_13TeV.342621.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1 _CT10ME_hh_yyWW.merge.AOD.e4419_a766_a821_r7676
260	mc15_13TeV.343756.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1 _CT10ME_Xhh_m260_yyWW.merge.AOD.e5153_a766_a821_r7676
300	mc15_13TeV.343758.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1 _CT10ME_Xhh_m300_yyWW.merge.AOD.e5153_a766_a821_r7676
400	mc15_13TeV.343761.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1 _CT10ME_Xhh_m400_yyWW.merge.AOD.e5153_a766_a821_r7676
500	mc15_13TeV.343763.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1 _CT10ME_Xhh_m500_yyWW.merge.AOD.e5153_a766_a821_r7676
750	mc15_13TeV.343818.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1 _CT10ME_Xhh_m750_yyWW.merge.AOD.e5153_a766_a821_r7676
1000	mc15_13TeV.343819.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1 _CT10ME_Xhh_m1000_yyWW.merge.AOD.e5153_a766_a821_r7676
1500	mc15_13TeV.343820.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1 _CT10ME_Xhh_m1500_yyWW.merge.AOD.e5153_a766_a821_r7676
2000	mc15_13TeV.343821.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1 _CT10ME_Xhh_m2000_yyWW.merge.AOD.e5153_a766_a821_r7676
2500	mc15_13TeV.343822.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1 _CT10ME_Xhh_m2500_yyWW.merge.AOD.e5153_a766_a821_r7676
3000	mc15_13TeV.343823.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1 _CT10ME_Xhh_m3000_yyWW.merge.AOD.e5153_a766_a821_r7676

Tab. 16: Signal sample names for non-resonant production and resonant production for several Higgs-boson mass points m_H

Channel	Background sample name
ggH	mc15_13TeV.341000.PowhegPythia8EvtGen_CT10_AZNLOCTEQ6L1 _ggH125_gamgam.merge.AOD.e3806_s2984_r8585_r7676
VBF	mc15_13TeV.341001.PowhegPythia8EvtGen_CT10_AZNLOCTEQ6L1 _VBFH125_gamgam.merge.AOD.e3806_s2608_r7772_r7676
ZH	mc15_13TeV.341068.Pythia8EvtGen_A14NNPDF23LO _ZH125_gamgam.merge.AOD.e3796_s2608_s2183_r7772_r7676
WH	mc15_13TeV.341067.Pythia8EvtGen_A14NNPDF23LO _WH125_gamgam.merge.AOD.e3796_s2608_s2183_r7772_r7676
ttH	mc15_13TeV:mc15_13TeV.341081.aMcAtNloHerwigppEvtGen_UEEE5 _CTEQ6L1_CT10ME_ttH125_gamgam.merge.AOD.e4277 _s2608_s2183_r7772_r7676

Tab. 17: Background SM $h \to \gamma\gamma$ samples

A.2. Cutflows

Cut	0.5 TeV	0.75 TeV	1 TeV	1.5 TeV	2 TeV	2.5 TeV	3 TeV
N_{DxAOD}	11.4577	12.4359	12.8440	13.3675	13.4486	11.2754	8.7729
DQ	10.9383	12.1078	12.6019	13.2095	13.3293	11.1762	8.6796
2 loose photons	8.7841	9.8296	10.3809	11.0882	11.3000	9.3894	7.0723
trigger match	8.7495	9.7793	10.3173	10.9978	11.1979	9.3076	7.0104
rel. $p_{T,\gamma}$ cuts	7.9847	9.2516	9.9048	10.7025	10.9709	9.1130	6.8397
fit range	7.9302	9.2026	9.8519	10.659	10.9450	9.0909	6.8160
$m_{\gamma\gamma} = 125.09$	6.8281	8.1099	8.7614	9.5683	9.8354	8.1396	6.0425
> 1 large-R jet	0.9999	4.5901	6.2810	7.9579	8.5840	7.2006	5.3276
b-veto	0.9733	4.4764	6.1263	7.7796	8.4067	7.0549	5.2047
lepton veto	0.8284	3.8094	5.1444	6.7608	7.5344	6.3793	4.6899
$p_{T,\gamma\gamma} > 300\,\text{GeV}$	0.0490	2.5742	4.7187	6.6457	7.4919	6.3611	4.6817
$C_1 > 0.14$	0.0178	1.5081	1.7608	0.8713	0.5799	0.3739	0.2525
$\Delta\eta_{\gamma\gamma,J} < 1.2$	0.0104	1.3913	1.1649	0.3670	0.2497	0.1458	0.0836
$m_J < 136\,\text{GeV}$	0.0086	1.1661	0.8097	0.0906	0.0326	0.0179	0.0096

Tab. 18: Cutflow for the final selection in the loose topology for the resonant mass samples. As the analysis is not sensitive to $m_H < 500\,\text{GeV}$, the mass points 260 GeV, 300 GeV and 400 GeV are not shown.

Cut	0.5 TeV	0.75 TeV	1 TeV	1.5 TeV	2 TeV	2.5 TeV	3 TeV
N_{DxAOD}	11.4577	12.4359	12.8441	13.3674	13.4486	11.2754	8.7729
DQ	10.9383	12.1078	12.6019	13.2095	13.3293	11.1762	8.6796
2 loose photons	8.7841	9.8296	10.3809	11.0882	11.3000	9.3894	7.0723
trigger match	8.7495	9.7793	10.3172	10.9978	11.1979	9.3076	7.0104
tight ID	8.7495	9.7793	10.3172	10.9978	11.1979	9.3076	7.0104
isolation	7.8642	8.9808	9.5797	10.3341	10.5803	8.6891	6.1195
rel. $p_{T,\gamma}$ cuts	7.2580	8.5894	9.2877	10.1291	10.4315	8.5648	6.0194
fit range	7.2426	8.5793	9.2710	10.1227	10.4284	8.5621	6.0179
$m_{\gamma\gamma} = 125.09$	6.2984	7.6051	8.2975	9.1202	9.3983	7.6844	5.3401
> 1 large-R jet	0.9141	4.3147	5.9557	7.5901	8.2087	6.8000	4.7515
b-veto	0.8895	4.2076	5.8096	7.4212	8.0404	6.6630	4.6395
lepton veto	0.7552	3.5819	4.8812	6.4482	7.2093	6.0457	4.2337
$p_{T,\gamma\gamma} > 275\,\text{GeV}$	0.0673	2.8561	4.6275	6.3728	7.1847	6.0346	4.2292
$\tau_{42} < 0.475$	0.0264	1.7091	2.7390	3.7604	4.2665	3.5934	2.4922
$C_1 > 0.08$	0.0208	1.5701	2.3434	2.4816	1.6642	0.9217	0.5448

Tab. 19: Cutflow for the final selection in the tight topology for the resonant mass samples. As the analysis is not sensitive to $m_H < 500\,\text{GeV}$, the mass points 260 GeV, 300 GeV and 400 GeV are not shown.

A.3. Yields for the loose selection, excluding the m_J cut

Sample	Event yield
Cont. bkg.	46.507 ± 2.302
ggF	0.736 ± 0.044
VBF	0.367 ± 0.035
WH	0.121 ± 0.008
ZH	0.100 ± 0.007
$t\bar{t}H$	0.114 ± 0.005
Σ_{bkg}^{loose}	47.945 ± 2.302
0.5 TeV	0.010 ± 0.001
0.75 TeV	1.391 ± 0.016
1 TeV	1.165 ± 0.014
1.5 TeV	0.367 ± 0.008
2 TeV	0.250 ± 0.007
2.5 TeV	0.146 ± 0.005
3 TeV	0.084 ± 0.004

Tab. 20: Background yields for the loose selection, excluding the m_J cut. The yields correspond to a significance of 0.200 ± 0.002.

B. References

[1] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961)

[2] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964)

[3] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967)

[4] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. **C680519**, 367 (1968)

[5] UA1 Collaboration, *Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2 at the CERN SPS Collider*, Phys. Lett. **126B** (1983)

[6] UA1 Collaboration, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at s**(1/2) = 540-GeV*, Phys. Lett. **122B** (1983)

[7] 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 (2012)

[8] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012)

[9] D. Fehling, *The Standard Model of Particle Physics: A Lunchbox's Guide*, Website, `http://www.pha.jhu.edu/~dfehling/`; retrieved on 20.03.2017

[10] M. Thomson, *Modern particle physics*, Cambridge University Press (2013)

[11] R. Aaij, et al. (LHCb), *Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays*, Phys. Rev. Lett. **115**, 072001 (2015)

[12] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Physical Review Letters **10**, 531 (1963)

[13] M. Kobayashi, T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Progress of Theoretical Physics **49**, 652 (1973)

[14] CMS Collaboration, *The Discovery of the Higgs Boson with the CMS Detector and its Implications for Supersymmetry and Cosmology*, in *Time and Matter 2013 (TAM2013) Venice, Italy* (2013)

[15] C. Arroyo, et al. (CCFR), *A Precise measurement of the weak mixing angle in neutrino nucleon scattering*, Phys. Rev. Lett. **72** (1994)

[16] D. de Florian, et al. (LHC Higgs Cross Section Working Group), *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector* (2016)

[17] J. R. Andersen, et al. (LHC Higgs Cross Section Working Group), *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties* (2013)

[18] ATLAS Collaboration, *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)

[19] J. F. Donoghue, *The effective field theory treatment of quantum gravity*, AIP Conf. Proc. **1483** (2012)

[20] H. Murayama, *Physics Beyond the Standard Model and Dark Matter*, in *Les Houches Summer School - Session 86: Particle Physics and Cosmology: The Fabric of Spacetime Les Houches, France, July 31-August 25, 2006* (2007)

[21] M. Kamionkowski, *WIMP and axion dark matter*, in *High-energy physics and cosmology. Proceedings, Summer School, Trieste, Italy, June 2-July 4, 1997* (1997)

[22] A. D. Sakharov, *Violation of CP Invariance, c Asymmetry, and Baryon Asymmetry of the Universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967)

[23] C. H. Llewellyn Smith, G. G. Ross, *The Real Gauge Hierarchy Problem*, Phys. Lett. **105B** (1981)

[24] C. Csaki, *The Minimal supersymmetric standard model (MSSM)*, Mod. Phys. Lett. **A11**, 599 (1996)

[25] D. Bailin, A. Love, *KALUZA-KLEIN THEORIES*, Rept. Prog. Phys. **50**, 1087 (1987)

[26] J. L. Hewett, *Warped Extra Dimensions and the Randall-Sundrum Model*, in *Workshop on CP Studies and Non-Standard Higgs Physics* (2006)

[27] P. Langacker, *Grand Unified Theories and Proton Decay*, Phys. Rept. **72** (1981)

[28] K. Nakamura, P. D. Group, *Review of Particle Physics*, Journal of Physics G: Nuclear and Particle Physics **37(7A)** (2010)

[29] T. D. Lee, *A Theory of Spontaneous T Violation*, Phys. Rev. **D8** (1973)

[30] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, M. Sher, J. P. Silva, *Theory and phenomenology of two-Higgs-doublet models*, Phys. Rept. **516** (2012)

[31] J. M. Cline, K. Kainulainen, A. P. Vischer, *Dynamics of two Higgs doublet CP violation and baryogenesis at the electroweak phase transition*, Phys. Rev. **D54** (1996)

[32] ATLAS Collaboration, *Searches for Higgs boson pair production in the hh $\rightarrow$ $bb\tau\tau, \gamma\gamma WW^*, \gamma\gamma bb, bbbb$ channels with the ATLAS detector*, Phys. Rev. **D92**, 092004 (2015)

[33] J. Baglio, A. Djouadi, R. Gröber, M. M. Mühlleitner, J. Quevillon, M. Spira, *The measurement of the Higgs self-coupling at the LHC: theoretical status*, JHEP **04** (2013)

[34] C. Lefèvre, et al., *The CERN accelerator complex. Complexe des accélérateurs du CERN*, Technical report (2008)

[35] ATLAS Collaboration, *Studies of the performance of the ATLAS detector using cosmic-ray muons*, Eur. Phys. J. **C71** (2011)

[36] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3**, S08003 (2008)

[37] *ATLAS Phase-II Upgrade Scoping Document*, Technical Report CERN-LHCC-2015-020. LHCC-G-166, CERN, Geneva (2015)

[38] M. H. Seymour, M. Marx, *Monte Carlo Event Generators*, in *Proceedings, 69th Scottish Universities Summer School in Physics : LHC Phenomenology (SUSSP69): St.Andrews, Scotland, August 19-September 1, 2012* (2013)

[39] F. D. Aaron, et al. (ZEUS, H1), *Combined Measurement and QCD Analysis of the Inclusive e+- p Scattering Cross Sections at HERA*, JHEP **01** (2010)

[40] R. Placakyte, *Parton Distribution Functions*, in *Proceedings, 31st International Conference on Physics in collisions (PIC 2011): Vancouver, Canada, August 28-September 1, 2011* (2011)

[41] Agostinelli, et al., *GEANT4—a simulation toolkit*, Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment **506(3)** (2003)

[42] W. Lukas, *Fast simulation for ATLAS: Atlfast-II and ISF*, in *Journal of Physics: Conference Series*, volume 396, IOP Publishing (2012)

[43] J. Alwall, et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014)

[44] M. Bahr, et al., *Herwig++ Physics and Manual*, Eur. Phys. J. **C58**, 639 (2008)

[45] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. M. Nadolsky, W. K. Tung, *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07**, 012 (2002)

[46] S. Gieseke, C. Rohr, A. Siodmok, *Colour reconnections in Herwig++*, Eur. Phys. J. **C72** (2012)

[47] S. Alioli, et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010)

[48] T. Sjostrand, S. Mrenna, P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008)

[49] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, C. P. Yuan, *New parton distributions for collider physics*, Phys. Rev. **D82** (2010)

[50] ATLAS Collaboration, *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP **09** (2014)

[51] R. D. Ball, et al., *Parton distributions with LHC data*, Nucl. Phys. **B867** (2013)

[52] *ATLAS Run 1 Pythia8 tunes*, Technical Report ATL-PHYS-PUB-2014-021, CERN, Geneva (2014)

[53] R. Atkin, *Review of jet reconstruction algorithms*, J. Phys. Conf. Ser. **645(1)** (2015)

[54] B. Isildak, *Measurement of the differential dijet production cross section in proton-proton collisions at $\sqrt{s} = 7$ TeV*, Ph.D. thesis, Bogazici U. (2011)

[55] CMS Collaboration, *A Cambridge-Aachen (C-A) based Jet Algorithm for boosted top-jet tagging* (2009)

[56] S. Catani, Y. L. Dokshitzer, M. H. Seymour, B. R. Webber, *Longitudinally invariant K_t clustering algorithms for hadron hadron collisions*, Nucl. Phys. **B406** (1993)

[57] M. Cacciari, G. P. Salam, G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **04** (2008)

[58] M. Cacciari, G. P. Salam, G. Soyez, *FastJet User Manual*, Eur. Phys. J. **C72** (2012)

[59] G. P. Salam, G. Soyez, *A Practical Seedless Infrared-Safe Cone jet algorithm*, JHEP **05** (2007)

[60] D. Krohn, J. Thaler, L.-T. Wang, *Jet Trimming*, JHEP **02** (2010)

[61] ATLAS Collaboration, *Measurement of the jet mass scale and resolution uncertainty for large radius jets at* $\sqrt{s} = 8$ *TeV using the ATLAS detector* (2016)

[62] ATLAS Collaboration, *Performance of jet substructure techniques for large-R jets in proton-proton collisions at* $\sqrt{s} = 7$ *TeV using the ATLAS detector*, JHEP **09** (2013)

[63] J. M. Butterworth, A. R. Davison, M. Rubin, G. P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, Phys. Rev. Lett. **100** (2008)

[64] S. D. Ellis, C. K. Vermilion, J. R. Walsh, *Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches*, Phys. Rev. **D81** (2010)

[65] ATLAS Collaboration (ATLAS), *Tagging and suppression of pileup jets* (2014)

[66] J. Thaler, K. Van Tilburg, *Identifying Boosted Objects with N-subjettiness*, JHEP **03**, 015 (2011), 1011.2268

[67] ATLAS Collaboration, *Jet mass and substructure of inclusive jets in* $\sqrt{s} = 7$ *TeV pp collisions with the ATLAS experiment*, JHEP **05** (2012)

[68] A. J. Larkoski, G. P. Salam, J. Thaler, *Energy Correlation Functions for Jet Substructure*, JHEP **06** (2013)

[69] M. Cacciari, G. P. Salam, *Pileup subtraction using jet areas*, Phys. Lett. **B659** (2008)

[70] ATLAS Collaboration, *Expected photon performance in the ATLAS experiment* (2011)

[71] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data* (2016)

[72] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV*, Eur. Phys. J. **C76(5)** (2016)

[73] ATLAS Collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, Technical Report ATL-PHYS-PUB-2016-012, CERN, Geneva (2016)

[74] G. Cowan, K. Cranmer, E. Gross, O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C71** (2011)

[75] ATLAS Collaboration, *Search for Higgs boson pair production in the final state of $\gamma\gamma WW^*(\rightarrow l\nu jj)$ using 13.3 fb^{-1} of pp collision data recorded at $\sqrt{s}$ = 13 TeV with the ATLAS detector* (2016)

[76] P. R. Bevington, D. K. Robinson, *Data reduction and error analysis for the physical sciences; 3rd ed.*, McGraw-Hill, New York, NY (2003)

[77] K. Cranmer, G. Lewis, L. Moneta, A. Shibata, W. Verkerke, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats* (2012)