Measurement of Higgs boson properties in the diphoton decay channel and a search for di-Higgs production in the $\gamma\gamma b\bar{b}$ final state with the ATLAS detector

Leonor Cerdá Alberich

A dissertation submitted for the degree of
Doctor of Philosophy

Supervised by:
Dr. Luca Fiorini
Dr. Carlos Solans Sánchez

Departament de Física Atòmica, Molecular i Nuclear and IFIC
(Universitat de València - CSIC)
Programa de Doctorat en Física

September 2018

Informe de los directores

Luca Fiorini, investigador doctor de la Universitat de València, y **Carlos Solans Sánchez**, investigador doctor de la Organización Europea para la Investigación Nuclear (CERN),

CERTIFICAN:

Que la presente memoria, ***"Measurement of Higgs boson properties in the diphoton decay channel and a search for di-Higgs production in the $\gamma\gamma b\bar{b}$ final state with the ATLAS detector"***, ha sido realizada bajo nuestra dirección en el *Departament de Física Atòmica, Molecular i Nuclear* de la *Universitat de València* por **Leonor Cerdá Alberich** y constituye su tesis para optar al grado de doctor en Física por la *Universitat de València*.

Y para que conste, en cumplimiento de la legislación vigente, firmamos el presente Certificado en Paterna a 12 de Septiembre de 2018.

Dr. Luca Fiorini

Dr. Carlos Solans Sánchez

Abstract

Measurement of Higgs boson properties in the diphoton decay channel and a search for di-Higgs production in the $\gamma\gamma b\bar{b}$ final state with the ATLAS detector

Leonor Cerdá Alberich

This thesis presents two physics analyses performed with the ATLAS detector at the Large Hadron Collider. Proton-proton collision data was used, corresponding to an integrated luminosity of 36.1 fb^{-1}, obtained at a center-of-mass energy of 13 TeV, during 2015 and 2016. The first analysis is a search for resonant and non-resonant Higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state. The second one is the measurement of the total Higgs boson production-mode cross sections, signal strengths, and simplified template cross sections, as well as the measurement of the fiducial and differential cross sections in the diphoton decay channel. Special emphasis is given to the strategy followed to estimate the uncertainty in the modeling of the parton shower, underlying event and hadronization. The reconstruction and identification techniques of the relevant objects such as photons and jets are covered extensively, and a validation of the calorimeter energy scale is performed by using the ATLAS Tile Calorimeter response to single hadrons with proton-proton collision data obtained at center-of-mass energies of 7 and 8 TeV, during 2010–2012 with the ATLAS detector.

Acknowledgements

Firstly, I would like to express my gratitude to Luca Fiorini and Carlos Solans. Luca has guided and supervised me through my entire PhD. I am grateful that he allowed me to be at CERN for most of it, giving me enough freedom and space to pursue the projects I felt most passionate about. I am particularly thankful to him for helping me become a much better scientist. Carlos was the first person to sit with me and go through my first physics analysis, when I was still a master's student. He trusted me enough to invite me to jointly supervise two CERN summer students during all this time. I will always be thankful for the lessons I learnt from him, not only about physics but about life.

I acknowledge the support from Ministerio de Economía, (Industria) y Competitividad and from Fondo Europeo de Desarrollo Regional (FEDER) de la Unión Europea (UE), through the project Contribuciones al Calorímetro Hadrónico TileCal y al programa de física del experimento ATLAS del Programa Nacional de Física de Partículas (FPA2012-32843, FPA2015-65652-C4-2-R).

I am deeply thankful to Professor Antonio Ferrer. He introduced me to the world of experimental particle physics more than six years ago, when I was just a final year physics student with a big dream. He always cared and believed in me and I strongly think I would not be in this position without him. I would also like to thank Professors Victoria Castillo and Emilio Higón for all the support and advice.

The IFIC TileCal group has been like a family to me during all these years. Some of its members remain, some are gone and others have arrived, but I will never forget the lovely times I had the pleasure to share with all of them. Big thanks to Juan, Alberto, Yesenia, Fernando, Pablo, Luis, Sergi, Pedro, Paco, Adam and Arantxa. I am particularly grateful to Damián for looking after me in every trip, in every school,

in every conference and workshop, in every office we shared, in every evening out, in my very first ATLAS Control Room shift, and basically, in every step I took in the last years. My PhD would not have been the same without him and I could not have asked for a better colleague and friend to go through this experience with.

I am thankful to a long list of $\gamma\gamma$ friends and colleagues, with whom I have shared work and lots of laughs. I thank my enthusiastic conveners, Sandrine Laplace, Dag Gillberg, Elisabeth Petit, Bruno Lenzi, Marco Delmastro, Giovanni Marchiori, Chris Meyer, Ruggero Turra and Haichen Wang. I am grateful for having had the opportunity to present my work in their group, for their advice and for trusting me in numerous occasions to take roles and responsibilities within the group. I am deeply thankful to Bruno for being an endless source of knowledge and encouragement, for his constant patience and availability, and for his friendship. Big thanks to Giovanni for all he has taught me, for motivating and bringing out the best of me. I would also like to thank the $\gamma\gamma b\bar{b}$ team, Jahred Adelman, Magda Chelstowska, Nancy Andari, Tulin Varol, James Robinson, Liza Brost, Marc Escalier and Louis Fayard, from whom I learnt so much that it is impossible to express my gratitude with simple words.

I am extremely grateful to one of the nicest and most hard-working communities in ATLAS, the TileCal collaboration. Thanks to Oleg Solovyanov for trusting me to be Run Coordinator, to Pawel Klimek for the fruitful discussions about performance and data quality, and about life, and to Irene Vichou for being an inspiration to me and to every young woman in the collaboration. I am thankful for her advice, for the always interesting conversations in the office and for giving me the opportunity to present my first ever TileCal poster not long after joining the group. The memory of her will always be in my heart. Big thanks to Jalal, Fabrizio, Henric, Arely, Filipe, Tomas, Stan, Irakli, Jimmy and many others, for their knowledge and kindness.

I would like to thank my South African colleagues and friends who have accom-

panied me in this journey. I am particularly thankful to Robert Reed for being a true friend from day one, for the nice memories in Geneva and the support he has always given to me through the different stages of my life. I am grateful to all my IFIC and CERN colleagues. Special thanks to Javier Jiménez and Marija Kekic, for always being there, for making life easier and brighter and for rescuing me after a tough day at work.

Thanks to my friends in Benicarló, my hometown, for keeping me sane every time that the physics world was getting a bit hard to handle. I am grateful to Cristina, Gemma, Marta, Enzo, Yesmina, Jairo, Carol, Sara, Bea and Sisco, for their sense of humour and their never-ending stories, for never failing to make me laugh, for taking time to visit me in Geneva and for proving that distance makes no difference when friendship is real.

Finally, I would like to thank my family for loving me endlessly and supporting me throughout all my life. Gracias a mi padre, por creer siempre en mí, por dejarme volar y ver mundo por mí misma. Gracias por enseñarme a ser paciente y a conseguir cualquier objetivo de manera limpia y justa. Gracias a mi madre, por quererme más que nada, por inculcarme su carácter luchador e inconformista, por ser la persona más generosa que conozco y por ser una mujer maravillosa. Gracias a ambos por darme el mejor regalo que habría podido desear, una educación y unos valores excelentes, por enseñarme que no hay nada imposible y que uno siempre debe aspirar a ser feliz. Gracias a mi hermano Gabriel por abrirme el camino, por entenderme y protegerme como solo un hermano mayor podría hacerlo. Y junto a María, por ser el mejor ejemplo a seguir que podría tener. No me puedo olvidar del más pequeño de la familia, Gabriel, por esa facilidad que tiene para arrancarme una sonrisa y por hacer que quiera ser mejor persona. I am grateful to Alan for bringing peace, joy, happiness, goals and dreams, for making every day worth living. Esta tesis es vuestra. Gracias.

Contents

Preface

The Large Hadron Collider (LHC) is the biggest and most powerful particle machine in the world. Located underground on the French-Swiss border outside Geneva, protons are circulating around this 27 km electromagnetic track and colliding into one another at the speed of light, recreating conditions of the universe when it was only a trillionth of a second old.

It has been six years since 2012, when the LHC experiments discovered a Higgs boson, the particle that explains why several other elementary particles have mass. That achievement completed the Standard Model (SM) of particle physics, which is in excellent agreement with observations, via Spontaneous Symmetry Breaking and the Higgs mechanism. The SM theory, however, is thought to be far from being complete. In fact it only incorporates three of the four fundamental interactions leaving gravity out of the picture. The search for Physics Beyond the Standard Model has been carried out for many years now in a large variety of possible scenarios. A part of the LHC program in the coming years is to measure the properties of the discovered Higgs boson and compare them to the theoretical predictions. Deviations in cross section measurements, differences in differential cross sections, or new and more precise measurements of the Higgs couplings, can point to new physics.

The "A Toroidal LHC ApparatuS" (ATLAS) experiment is one of the four experiments which records the collisions delivered by the LHC. The first period of successful

data-taking of the ATLAS experiment receives the name of Run 1. ATLAS recorded a 4.5 fb^{-1} dataset of p-p collisions at a centre of mass energy of $\sqrt{s} = 7$ TeV and a 20.3 fb^{-1} dataset of p-p collisions at $\sqrt{s} = 8$ TeV. The second period of data-taking, called Run 2, restarted in 2015 with p-p collisions delivered at $\sqrt{s} = 13$ TeV. During 2015–2017 ATLAS recorded a dataset of 79.8 fb^{-1}. Run 2 is still ongoing and will finish at the end of 2018. The LHC is foreseen to run for many years from now and an upgrade to the High-Luminosity LHC (HL-LHC) is foreseen in 2024, bringing the expected size of the collected p-p collisions dataset by ATLAS to 3000 fb^{-1}. Therefore, the LHC and ATLAS provide an excellent experimental setup to probe the existence of physics beyond the Standard Model in the so far unexplored phase space.

Chapter 1 starts with a review of the Standard Model of particle physics and follows with a description of the phenomenological aspects of the Higgs boson in proton colliders. Special emphasis is made to cover the details about the di-Higgs production and the scenarios which could increase the rate of it, therefore implying new physics.

Chapter 2 is an introduction to the LHC at CERN, followed by an extensive review on the ATLAS experiment with a special emphasis on the description of the Liquid Argon and Tile Calorimeters. The Trigger and Data Acquisition systems are also covered.

Chapter 3 provides a description of the algorithms and methods used in the ATLAS experiment for reconstruction and identification of the particles produced in p-p collisions. The author has contributed to the understanding of the reconstruction of photons by performing photon conversion studies with early Run 2 data.

In Chapter 4 the response of the Tile Calorimeter to single isolated charged hadrons is analysed with Run 1 data collected by the ATLAS detector in 2010–2012. The author has carried out this performance study as part of her service task. It has

been published in [1].

In Chapter 5 a search for di-Higgs production in the $\gamma\gamma b\bar{b}$ final state is presented with Run 2 data collected by the ATLAS detector in 2015–2016. The author has contributed to this search as main analyzer since the beginning of Run 2 [2, 3].

Chapter 6 summarizes the main findings in the measurement of the total Higgs boson production-mode cross sections, signal strengths, and fiducial and differential cross sections in the diphoton decay channel with ATLAS Run 2 data, which is also published in [4]. The author has estimated the uncertainty in the modeling of the parton shower, underlying event and hadronization, which has set a baseline for future progress in this area.

Chapter 1

Theory

This chapter presents the theoretical background necessary to motivate and understand the contents of the thesis. It starts with a description of the Standard Model of Particle Physics, a theory that describes the dynamics of the subatomic world. It follows with an explanation of the electroweak symmetry breaking mechanism that motivates the existence of the Higgs boson, which was first observed at the Large Hadron Collider (LHC) in 2012. Phenomenological aspects of the Higgs boson at the LHC, namely production mechanisms and decay modes are also discussed, with a focus on Higgs boson pair production and its relevance at present and in the upcoming years.

1.1 The Standard Model

The Standard Model (SM) [5–8] constitutes one of the most successful achievements in modern physics. It provides a very elegant theoretical framework, which is able to describe the known experimental facts in particle physics with high precision.

The SM is a non-abelian gauge quantum field theory (QFT), or Yang-Mills the-

ory [9], based on the symmetry group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, where C indicates the color charge, L the weak isospin and Y the hypercharge. The SM describes strong, weak and electromagnetic (EM) interactions, via the exchange of the corresponding spin-1 gauge fields: eight massless gluons and one massless photon, respectively, for the strong and electromagnetic interactions, and three massive bosons, $W^{\pm}$ and Z, for the weak interaction.

The fermionic matter content is given by leptons and quarks, which are organized into three families/generations of increasing mass, each one being composed of two elements, that form a doublet, as shown in Table 1.1. All known stable matter is actually made up of only first generation fermions, in the form of u and d quarks, that constitute protons and neutrons, and electrons. Second and third generation fermions are unstable and decay into lighter particles.

Fermions (Spin = 1/2)						
	Leptons			Quarks		
Generation	Flavor	Charge (e)	Mass	Flavor	Charge (e)	Mass
1st	ν_e	0	<2 eV	u	+2/3	2.2 MeV
	e	-1	0.511 MeV	d	-1/3	4.7 MeV
2nd	ν_μ	0	<0.19 MeV	c	+2/3	1.28 GeV
	μ	-1	105.658 MeV	s	-1/3	96 MeV
3rd	ν_τ	0	<18.2 MeV	t	+2/3	173.1 GeV
	τ	-1	1776.86 MeV	b	-1/3	4.18 GeV

Table 1.1: Fermions in the SM: leptons, quarks and some of their basic properties. Neutrinos are treated as massless in the SM, which is a good approximation, even though current results indicate the masses are very small but non-zero [10].

For leptons, each doublet includes an electrically charged particle (ℓ) and the corresponding neutral particle, the neutrino (ν). A lepton number can be assigned to these particles, and it is 1 for leptons and -1 for antileptons. In the SM, the lepton number is conserved, both overall and within generations. This means that when

particles interact, the number of leptons of the same family remains the same. It also implies, for example, that leptons can only be created in lepton/antilepton pairs of the same generation.

Quarks (q) carry fractional electric charges and the doublets in each generation are formed by a $+2/3$ e-charged up-type quark and a $-1/3$ e-charged down-type quark. The six different types of quarks that exist are commonly referred to as flavors. Quarks have a property (or quantum number) named color, that can be viewled as a charge that is conserved in the SM, just like the electric charge. In fact, each flavor of quark comes in three different colors, red (R), green (G) and blue (B), tripling the actual number of quarks in the SM with respect to what is shown in Table 1.1. Then there are also the antiquarks ($\bar{q}$), which carry anticolors ($\bar{R}$, $\bar{G}$, $\bar{B}$).

Four fundamental forces exist in Nature, with strengths spanning several orders of magnitude. The strength of an interaction is determined by its coupling constant. The gravitational force, not described by the SM, is too weak to play a significant role in elementary particle physics. The strong force is responsible for binding quarks together to form nucleons. The EM force provides attraction between electrons and atomic nuclei. The weak force accounts for the nuclear β-decay. These last three forces are described by gauge theories, so the corresponding exchanging bosons are often called gauge bosons. They are listed in Table 1.2.

Bosons					
Name	Spin	Charge (e)	Mass (GeV)	Force	Relative strength
Gluon, g	1	0	0	Strong	1
Photon, γ	1	0	0	Electromagnetic	10^{-2}
$W^{\pm}$	1	$\pm$ 1	80.385	Weak	10^{-13}
Z	1	0	91.188		
H	0	0	125.09		

Table 1.2: Gauge bosons and fundamental forces in the SM. Particle masses from [10].

The photon is a massless, neutral, spin-1 boson that mediates EM interactions. Photons only couple to electrically-charged particles. Thus, quarks and charged leptons can interact electromagnetically, but neutrinos cannot.

Gluons are also massless, neutral, spin-1 bosons that carry the strong force. They couple to the color charge and, therefore, mediate the interactions between quarks, but not leptons. Gluons themselves have color and can, unlike photons, interact directly with each other. Gluons are bicolored, carrying one unit of color and one unit of a different anticolor. A total of eight gluons exist with different colors.

Charged-current weak interactions are mediated by the $W^{\pm}$ bosons, and are involved in transitions between up-type and down-type quarks, or charged leptons and the corresponding neutrinos. Detailed analysis of the energy and angular distributions of β-decays have shown that only left-handed (right-handed) fermion (antifermion) helicities participate in charged-current interactions, which, therefore, violate parity. Neutral-current weak interactions are mediated by the Z boson and conserve flavor. Similar to the gluons in the strong interaction, the $W^{\pm}$ and Z bosons also couple to each other. Since the $W^{\pm}$ bosons hold electric charge, they can also couple to the photon and interact via the EM force.

The Higgs boson does not mediate any of the fundamental interactions but it is an essential piece of the SM. With a mass of approximately 125 GeV, no intrinsic spin and no charge, the Higgs field delivers mass to other elementary particles through the Spontaneous Symmetry Breaking mechanism (see Section 1.2.2). Consequently, the Higgs boson couples to all particles with mass, including itself.

1.2 Standard Model Lagrangian

The complete Lagrangian of the Standard Model can be factorized in terms that describe the electroweak interactions, the Higgs sector, the Yukawa terms for the

fermion masses and the strong interactions:

$$\mathcal{L} = \mathcal{L}_{EW} + \mathcal{L}_{H} + \mathcal{L}_{strong} + \mathcal{L}_{Yukawa}. \tag{1.1}$$

In this subsection the various terms composing the SM theory are briefly introduced and to help the readability, the usual convention of summing over repeated indices is implicitly used.

1.2.1 Electroweak interaction

The electroweak theory, proposed by Glashow, Salam and Weinberg [11–13] to describe the electromagnetic and weak interactions between quarks and leptons, is based on the gauge symmetry group $SU(2)_L \otimes U(1)_Y$ of weak left-handed isospin and hypercharge.

The electroweak Lagrangian can be written as:

$$\begin{aligned}\mathcal{L}_{EW} = &\sum_{flavors} i(\bar{L}\not{D}L + \bar{Q}\not{D}Q + \bar{\ell}_R\not{D}\ell_R + \bar{u}_R\not{D}u_R + \bar{d}_R\not{D}d_R) \\ &- \frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W^a_{\mu\nu}W^{\mu\nu}_a,\end{aligned} \tag{1.2}$$

where the fields are organized in $SU(2)_L$ doublets (L being the left-handed doublet lepton field, and Q being the left-handed doublet quark field) and singlets (u, d and ℓ being the right-handed singlet quark and lepton fields), respectively. D is the covariant derivative, $\not{D} = D^\mu\gamma_\mu$, where γ are the Dirac matrices and $B_{\mu\nu}$ and $W^a_{\mu\nu}$ are the gauge field kinetic terms defined as:

$$B_{\mu\nu} = \partial_\mu B_\nu + \partial_\nu B_\mu, \tag{1.3}$$

$$W^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu + g\epsilon^{abc}W^b_\mu W^c_\nu, \tag{1.4}$$

where B_μ and W^a_μ are the gauge bosons associated to $U(1)_Y$ and $SU(2)_L$, respectively. The structure constants for $SU(2)$ represented by the component of the three dimensional Levi-Civita tensor ϵ^{abc} with $\epsilon^{abc} = 1$ has been used, and g is the coupling constant for the weak isospin. The covariant derivative in Eq. 1.2 can be written as

$$D_\mu = \partial_\mu - \frac{1}{2} i g' Y B_\mu + \frac{1}{2} i g \tau^a W^a_\mu, \tag{1.5}$$

where τ^a are the Pauli matrices, Y is the hypercharge and g' is the coupling constant for the weak hypercharge. Since the right handed fermions are singlets under $SU(2)$, the term proportional to the weak isospin vanishes. The hypercharge value is fixed by the Gell-Mann-Nishijima relation $Y/2 = Q - T_3$, where Q is the electric charge and T_3 is the eigenvalue of the third component of the weak isospin.

1.2.2 Spontaneous Symmetry Breaking and Yukawa interactions

In a Yang-Mills theory, Dirac or Majorana mass terms for fermions and gauge bosons are not admitted because they would break the gauge invariance. Therefore, all particles in the theory need to be massless in contrast with experimental observations.

The solution to the problems in the weak sector is the mechanism of Electroweak Symmetry Breaking (EWSB), which generates the weak vector boson masses in a way that is minimal and respects the requirements of renormalizability [14] and unitarity. An $SU(2)$ doublet of complex scalar fields ϕ contributing with the Lagrangian terms

$$\mathcal{L}_H = (D_\mu \phi)^2 - \mu^2 \phi^2 - \lambda \phi^4, \tag{1.6}$$

where it is required that $\mu^2 < 0$ and $\lambda > 0$ such that the ϕ field has a non-vanishing

vacuum expectation value given by

$$v = \sqrt{\frac{-\mu^2}{\lambda}}. \tag{1.7}$$

As a consequence, the electroweak $SU(2)_L \otimes U(1)_Y$ symmetry is spontaneously broken to the electromagnetic $U(1)_Q$ symmetry, where Q is the electric charge, giving in this way mass to the weak force carrier bosons and leaving the photon massless, at the cost of the introduction of an extra physical scalar field H.

The generators of $SU(2)_L \otimes U(1)_Y$ (B_μ for $U(1)_Y$ and W^0_μ, W^1_μ, W^2_μ of $SU(2)_L$) mix with each other and with the degrees of freedom of the complex scalar field, resulting in the physical gauge boson fields. The masses of the physical states are given by

$$m^{\pm}_W = \frac{1}{2}vg, \quad m_Z = \frac{m_W}{cos\ \theta_W}, \quad m_\gamma = 0, \tag{1.8}$$

where v is the vacuum expectation value, g is the coupling constant for the weak isospin and θ_W is the mixing angle called the Weinberg angle, defined by the condition $sin\ \theta_W = g'/\sqrt{g'^2 + g^2}$, where g' is the coupling constant for the weak hypercharge.

The massless gauge boson, the photon, corresponds to a superposition of B_μ and W^0_μ and has no mass because the $U(1)$ symmetry of Quantum Electrodynamics (QED) remains unbroken in the process. The two massive vector bosons W^+ and W^- correspond to the charged weak currents and the third massive boson, Z, is the neutral vector boson responsible for the neutral-current weak interactions. The extra physical field which couples to particles in proportion to their mass is called the Higgs field and has a corresponding mass given by $m_H = \sqrt{2\lambda}v$, where λ is the quartic coupling.

The mechanism described above was developed in the 1960s, including the spontaneous symmetry breaking mechanism (the Brout-Englert-Higgs or BEH mecha-

nism [15]) and the unification of the electromagnetic and weak forces. The new theory has all the desirable properties of a quantum field theory: the Lagrangian is Lorentz invariant and the theory is renormalizable.

A mass term for the fermions can also be obtained through the BEH mechanism in a way that preserves gauge invariance adding to the EW Lagrangian the Yukawa terms [16]. They can be written as:

$$\mathcal{L}_{Yukawa} = \sum_{flavors} (-\lambda_\ell \bar{L}\phi\ell_R - \lambda_d \bar{Q}\phi d_R - \lambda_u \epsilon^{ab}\bar{Q}_a\phi_b^\dagger u_R + h.c.), \tag{1.9}$$

where ϵ^{ab} is the two dimensional total anti-symmetric tensor with $\epsilon^{12} = 1$, and λ_e, λ_d and λ_u are arbitrary parameters. After symmetry breaking we get the following mass terms for the fermion fields:

$$m_\ell = \lambda_\ell \frac{v}{\sqrt{2}},\ m_d = \lambda_d \frac{v}{\sqrt{2}},\ m_u = \lambda_u \frac{v}{\sqrt{2}}. \tag{1.10}$$

The predictions of the unified electroweak theory were confirmed in the following years as the massive W and Z bosons were discovered in 1983 at CERN by the UA1 and UA2 experiments using collisions at the Super Proton Synchrotron collider [17,18]. The remaining missing piece was the predicted massive scalar boson, the Higgs boson (see Section 1.4).

1.2.3 Strong interaction

The theory of the strong interactions of coloured quarks and gluons is called Quantum Chromodynamics (QCD) and is based on the $SU(3)_C$ gauge symmetry group. The dynamics of this gauge group lead to two important properties of strong interactions. The first is called *asymptotic freedom*, which states that the strong coupling constant becomes smaller when interactions happen at greater energy scales, making pertur-

bative theory applicable at high energies while at low energies analytical calculations are not possible.

The second important property is called $confinement$ and states that quarks are bound into colourless states called hadrons. There are two types of hadrons: mesons, which are composed by a quark-antiquark pair, and baryons, which are composed by three quarks or three anti-quarks. This important property leads to the fact that when quarks are pulled apart, their interaction becomes so strong that quark-antiquark pairs are created from the vacuum. This is at the origin of the hadronization process that leads to the creation of sprays of hadronic particles, called jets, as products of p-p collisions at the LHC.

The QCD Lagrangian can be written in covariant notation as:

$$\mathcal{L}_{strong} = \sum_{q} \bar{\psi}_{q,i}(i\not{D}_{ij} - m_q\delta_{ij})\psi_{q,j} - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a, \tag{1.11}$$

where $\bar{\psi}_{q,i}$ is the fermion field of flavor q and colour i and the covariant derivative operator D^μ and the gluon fields kinetic terms $G^a_{\mu\nu}$ are given by:

$$D^\mu = \partial_\mu + ig_s t^a G^a_\mu, \tag{1.12}$$

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f^{abc} G^b_\mu G^c_\nu, \tag{1.13}$$

where g_s is the strong running coupling constant and f^{abc} are the $SU(3)$ structure constants given by commutation relations between the group algebra generators $[t^a, t^b]$ $= if^{abc}t^c$.

1.3 Proton-proton collisions at hadron colliders

In order to fully understand the production of particles through proton collisions it is important to give a brief description of the proton's composition. In this section, we will discuss the parton model and some phenomenological aspects of the LHC, namely which tools are used at present to describe the complicated environment at hadron colliders [19, 20].

1.3.1 Phenomenological aspects

Figure 1.1 illustrates a typical high-energy proton-proton collision, such as the ones occurring at the LHC at CERN. A proton is a very complex object composed by valence and sea quarks as well as gluons responsible for the strong interactions. The constituents of the protons are called partons. The valence quarks are, in the case of protons, two up quarks and a down quark and are the ones defining the quantum numbers of the protons. The sea quarks are extremely short living quark/anti-quark pairs created by random vacuum fluctuations, but that can still interact with the partons in the colliding proton.

A collision between two partons, one from each side, with large momentum transfer, gives the hard process of interest. The remaining partons in each proton can continue to travel essentially along the original direction, forming the beam remnants, or also interact in the same p-p collision. These interactions are usually soft, as they involve low momentum transfer between the partons, and constitute the underlying event (UE). Though with a much smaller probability, more than one hard parton interaction can occur in the same p-p collision, and these are referred to as multi-parton interactions (MPI).

A high energy collision involves accelerated color and EM charges, that can radiate gluons and photons (Bremsstrahlung). Emissions associated with the incoming

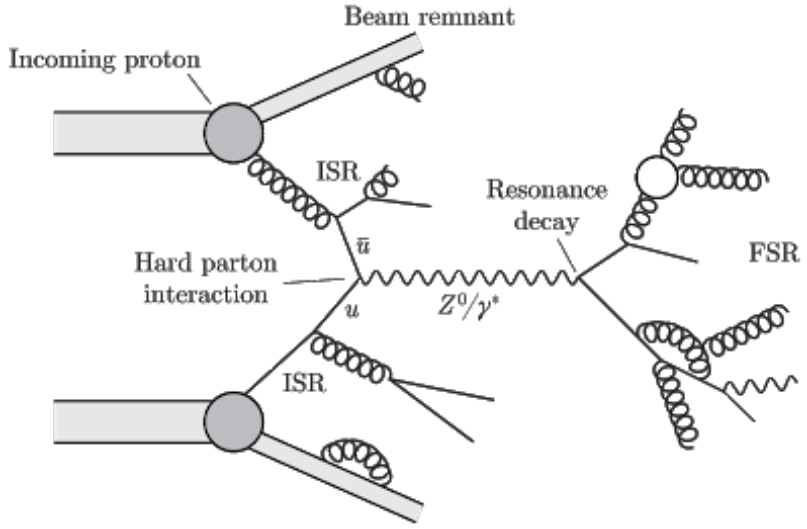


Figure 1.1: Schematic diagram of an event in a p-p collision [21].

colliding partons are referred to as Initial State Radiation (ISR). Emissions made by final state particles are called Final State Radiation (FSR).

As the distance between outgoing partons increases, so does the strong interaction between them, forcing them to be confined in colorless states. At this point the process of hadronization occurs, by which a collimated bunch of hadrons, a jet, is created from the color fields between the partons.

Each of the partons carries a fraction of the proton's momentum x, defined mathematically as the ratio between the magnitude of the parton momentum over the magnitude of the proton momentum. The proton can be described using form factors $f(x, Q^2)$ for each parton inside the proton, called parton distribution functions (PDFs). Here, Q^2 is the momentum transfer. The form factor has a strong dependence on x, and only a small dependence on Q^2. This dependence is provided by the DGLAP equation [22].

The cross section is a measure of the probability of a specific scattering process, under some given set of initial and final conditions. The cross section calculation of

a *p*-*p* process can be factorized into two parts: the perturbative QCD calculation in orders of α_s of the hard-scatter process and the physics of the proton described by the parton model. The scale Q^2 that is assumed to separate these two components is called the factorization scale, μ_F. The strong coupling α_s is small for the hard scattering component of a collision. The hard processes can, therefore, be described using perturbative QCD (pQCD), which relies on an order-by-order expansion of the observables in α_s. The soft part of a collision, on the other hand, involves low momentum transfers and high α_s. It is, thus, essentially non-perturbative, and precise calculations cannot be done.

Let us assume a *p*-*p* interaction between protons A and B with a hard scattering process occurring between partons a and b, which results in the final state X: $ab \to X$. The total cross section can be written as:

$$\sigma_{AB} = \int dx_a dx_b f_{a/A}\left(x_a, \mu_F^2\right) f_{b/B}\left(x_b, \mu_F^2\right) \times \left[\sigma_0 + \hat{\alpha}_S\left(\mu_R^2\right)\hat{\sigma}_1 + ...\right]_{ab\to X}. \quad (1.14)$$

The terms $f_{a/A}$ and $f_{b/B}$ are PDFs that contain the non-perturbative soft component of the collision. They essentially provide the probability that the proton A has a parton a carrying a fraction x_a of its momentum (similar for B). PDFs depend on the non-physical factorization scale μ_F. Another unphysical scale is the renormalization scale μ_R, which corresponds to the value at which α_s is computed. A cross section computed at a fixed order depends on μ_R and μ_F, and variations of these scales provide handles on the theoretical uncertainties. The dependence becomes weaker with increasing orders of calculation, and would vanish if the cross section was computed to all order in perturbation theory. Thus, results obtained only at leading order (LO) suffer from large uncertainties and higher-order corrections (next-to-leading order, NLO, or next-to-next-to-leading order, NNLO) can be significant. It is useful to define the K-factor, as the ratio between the NLO and LO cross sections.

PDFs are universal and independent of the hard process. They are measured experimentally in deep inelastic scattering, Drell-Yan and p-$p \rightarrow$ jets processes [23]. Several collaborations provide PDFs that can be used to make QCD cross section predictions [24–27].

1.3.2 Monte Carlo event generators

The complexity of a high-energy hadron collision is simulated by programs referred to as Monte Carlo (MC) event generators [28]. MC generators predict observables, such as momentum or angular distributions, that can be compared to data. They work sequentially, to implement the different steps of a collision, and are essential to any physics analysis. First, a Matrix Element Monte Carlo, convoluted with the PDFs described in Section 1.3.1, will generate a hard-scatter (HS) $2 \rightarrow n$ process at a fixed-order (LO, NLO, NNLO, etc) where n represents a large number of partons in the final state. In the case of the production of short-lived resonances, the decays to stable particles are also performed, taking into account the partial decay widths and any other properties, such as spin correlations.

Higher-order effects in perturbation theory are added to the HS event, using a Parton Shower (PS) MC, which simulates the multiple emission of soft radiation. PS methods are non-perturbative and, thus, less precise than the ME. Several methods exist and they are approximate, allowing for some tuning to measurements. The combination of a ME with a PS generator (matching) must be treated carefully, to avoid double counting of diagrams or any gaps in the phase-space. PS shower generators are also used to generate FSR and ISR.

The process of hadronization, by which the outgoing partons end up confined in hadrons, is non-perturbative, as already mentioned. Different phenomenological models exist, such as the Lund string model [29] or the cluster model [30], and are

implemented in MC generators. They are derived from experimental input and can usually be tuned to provide better predictions. Finally, the simulation of the pile-up[1] is usually handled by overlaying extra $2 \to 2$ scatterings referred to as minimum bias events. Because these processes are, again, mostly non-perturbative, implementations in MC programs rely heavily on tuning to data.

1.4 The Higgs boson phenomenology at hadron colliders

1.4.1 Production

The main production modes of the Higgs boson at a *p*-*p* collider are presented in Figure 1.2. The cross sections and uncertainties of each production mechanism are provided by the LHC Cross Section Working Group [31–34]. Figure 1.3 shows the cross sections for each production process, as a function of the center-of-mass energy $\sqrt{s}$, for a Higgs boson of mass $m_H = 125$ GeV.

Gluon-gluon fusion (ggH) is the main Higgs boson production channel at the LHC up to masses of 1 TeV, representing 87% of Higgs boson production at $\sqrt{s} = 13$ TeV.[2] The main contribution to the loop arises from the top quark, due to the large Yukawa coupling. The cross section for ggH is currently known at N^{3}LO (next-to-next-to-next-to-leading order) in QCD, with soft gluon contributions resummed up to NNLL (next-to-next-to-leading logarithm), and also includes NLO EW corrections.

Vector boson fusion (VBF) is the second largest contribution to the Higgs boson production at $\sqrt{s} = 13$ TeV (6.8%) and possesses a very distinct signature, with two hard jets produced in the forward regions, originating from the two outgoing quarks. The Higgs boson appears between the jets, in the central region of the detector. The VBF cross section is currently computed at approximate NNLO in QCD and full

[1] Pile-up is defined as the mean number of interactions per bunch crossing.
[2] Higgs boson cross sections are quoted given a mass of 125.09 GeV.

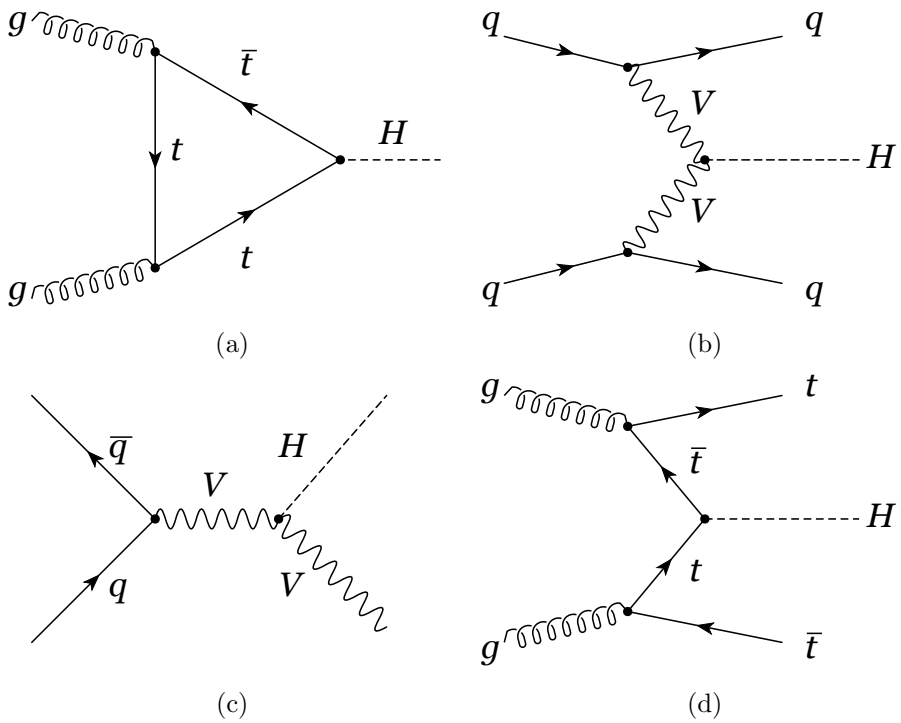


Figure 1.2: Leading-order Feynman diagrams of Higgs boson production modes, through a) gluon- gluon fusion via heavy quark loops, b) vector boson (W or Z) fusion with two forward jets, c) radiation of a Higgs boson (Higgs-Strahlung) from a V (W or Z) boson and d) Higgs production in association with top quarks.

NLO in QCD and EW. The theoretical uncertainties on VBF are much smaller than on ggH because the process is purely electroweak (at leading order) and the QCD corrections have a smaller impact.

Associated production with a vector boson (WH, ZH) is the third most common production mechanism which corresponds to 4.0% of the total Higgs boson production cross section at $\sqrt{s} = 13$ TeV. The process is also called Higgs-Strahlung due to the radiation of a Higgs boson analogous to that of Bremsstrahlung. The cross sections are currently calculated at NNLO in QCD and NLO in EW.

Top associated production ($t\bar{t}H$ and tH) has one of the lowest rates of Higgs boson

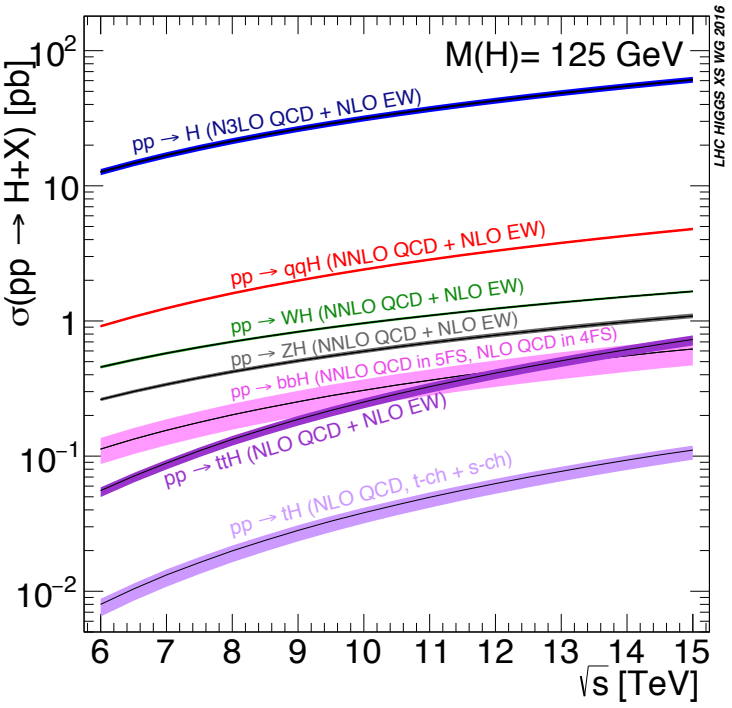


Figure 1.3: Cross sections for the different production mechanisms of the SM Higgs boson at the LHC, as a function of the center-of-mass energy and for a Higgs boson mass of 125 GeV [34].

production at the LHC at $\sqrt{s} = 13$ TeV. Both single and double top production is difficult to measure due to the many final states with b quarks and leptons. The cross section estimation is done at NLO in EW and QCD. This production mode is important to probe as it offers direct access to the top Yukawa coupling.

Finally, bb fusion ($b\bar{b}H$) production has a cross section calculated with two flavor schemes. The four flavor scheme (4FS), computed at NLO, and the five flavor scheme (5FS) computed at NNLO. The 5FS EW corrections are ignored since they are found to be small. The 4FS can be viewed as gluon splitting similar to the $t\bar{t}H$ process while the 5FS has the b quarks added into the parton distribution function (PDF) for the two incoming protons.

1.4.2 Decays

The couplings of the Higgs boson are always proportional to some mass scale, as discussed in Section 1.2.2. The $Hf\bar{f}$ interaction grows linearly with the fermion mass, while the HWW and HZZ vertices are proportional to m_W^2 and m_Z^2, respectively. Therefore, the most probable decay mode of the Higgs boson is the one into the heaviest possible final state. This is clearly illustrated in Figure 1.4.

The Higgs boson branching ratios mentioned in this section are extracted from [34] given $m_H = 125.09$ GeV. The $H \to b\bar{b}$ decay channel is by far the dominant one ($\mathcal{B} \approx 0.581$) as the Higgs boson mass is below the W^+W^- production threshold. Alongside the $H \to \tau^+\tau^-$ decay channel ($\mathcal{B} \approx 0.0626$), these fermionic decay channels are useful in determining the Higgs boson couplings to different quark types and leptons but tend to be difficult to distinguish from their large background processes.

The Higgs boson can also decay into a pair of gauge bosons, $H \to W^+W^-$ ($\mathcal{B} \approx 0.215$) and $H \to ZZ$ ($\mathcal{B} \approx 0.0264$). The $ZZ \to 4\ell$ decay has a very high signal to background ratio and the WW decay is characterized by a poor mass resolution due to neutrinos in the final state. For $m_H > 2m_t$, the $H \to t\bar{t}$ decay width is also sizeable, although smaller than the WW and ZZ ones because of the different dependence of the corresponding Higgs boson coupling with the mass scale (linear instead of quadratic). The design of the LHC detectors has taken into account all these very characteristic properties in order to optimize the search for new particles.

The Higgs boson can also decay into two photons $H \to \gamma\gamma$ through a radiative one-loop transition with a virtual top-antitop-quark pair, but can also be replaced with a W boson loop. Although this decay mode is quite suppressed ($\mathcal{B} \approx 0.00227$), it plays an important role in the LHC Higgs boson search at small masses because of its clean signature, in comparison with the big backgrounds affecting the dominant $H \to b\bar{b}$ channel. The decays of the Higgs boson are shown as Feynman diagrams in

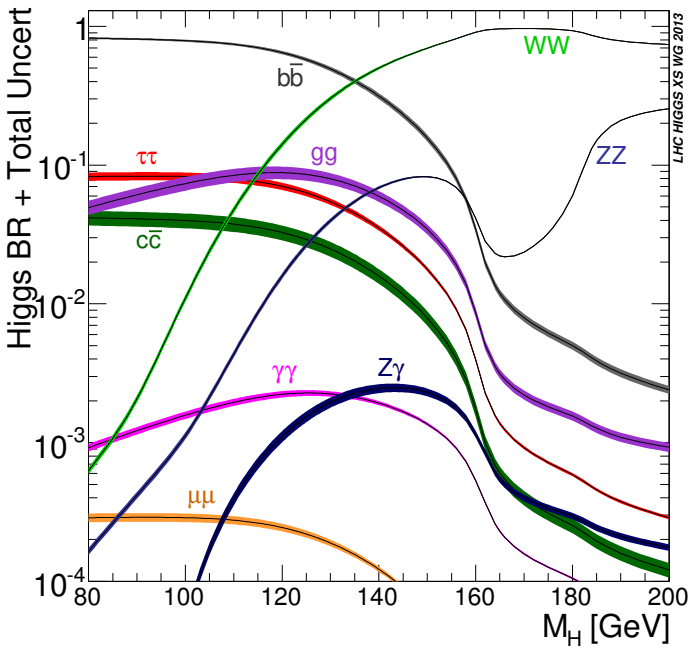


Figure 1.4: Branching ratios of the SM Higgs boson as a function of its mass [33].

Figure 1.5.

1.4.3 Discovery and property measurements of the Higgs boson

A priori, there is no theoretical prediction within the SM for the Higgs boson mass. Combining the data from the four detectors of the Large Electron-Positron collider, a lower bound of 114.4 GeV (95% CL) was set on the Higgs boson mass [35]. Data collected at the Tevatron was used to exclude masses between 100 and 109 GeV, and between 158 and 175 GeV at 95% CL [36]. The discovery of the Higgs boson was eventually made using collisions at the Large Hadron Collider (LHC). On July 4, 2012, the ATLAS and CMS experiments announced the discovery of a particle with Higgs-like properties, with a mass of about 125 GeV [37, 38], with the significance shown in Figure 1.6.

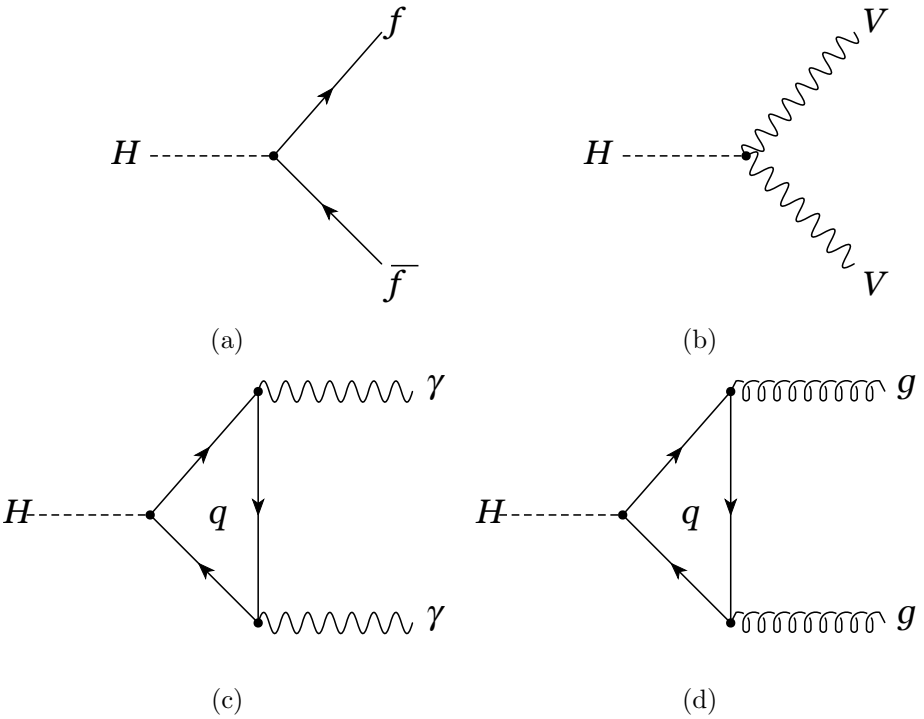


Figure 1.5: Feynman diagrams for the decay modes of the SM Higgs boson into: a) fermion-antifermion pairs, b) WW or ZZ spin-1 boson pairs, c) two photons via virtual quark loop (which can also be replaced with a W boson loop) and d) two gluons.

The discovery of the Higgs boson was made by combining the results from $H \to ZZ$, $H \to \gamma\gamma$, $H \to W^+W^-$, $H \to b\bar{b}$ and $H \to \tau^+\tau^-$. The $H \to \gamma\gamma$ and $H \to ZZ$ decay channels enable the most precise measurement of the Higgs boson mass since the Higgs boson candidate can be fully reconstructed and the momentum and energy of the final state particles (photons, electrons and muons) can be measured with high precision. The combined mass measurement in these decay channels with the ATLAS and CMS experiments using the Run 1 data set was:

$$m_H = 125.09 \ \pm 0.21\,(\text{stat.}) \pm 0.11\,(\text{syst.})\,\text{GeV}. \tag{1.15}$$

The ATLAS and CMS collaborations also combined results for Higgs boson pro-

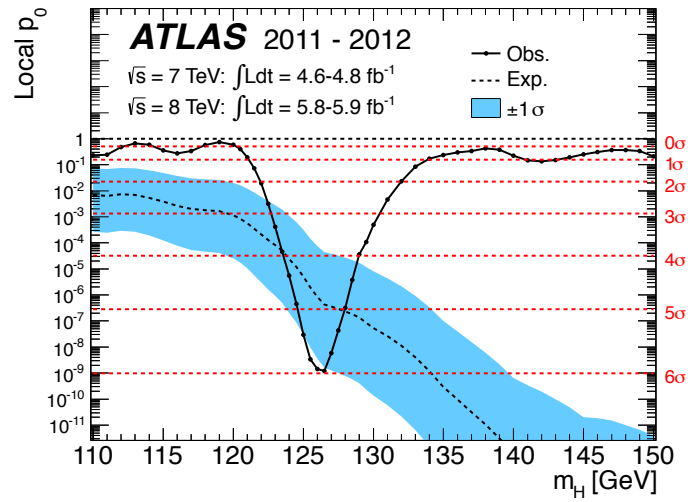
duction and decay modes using the Run 1 data set. These measurements can be presented in terms of signal strength, which is defined for each production and decay mode as:

$$\mu = \frac{\sigma \times \mathcal{B}}{\sigma_{SM} \times \mathcal{B}_{SM}}, \tag{1.16}$$

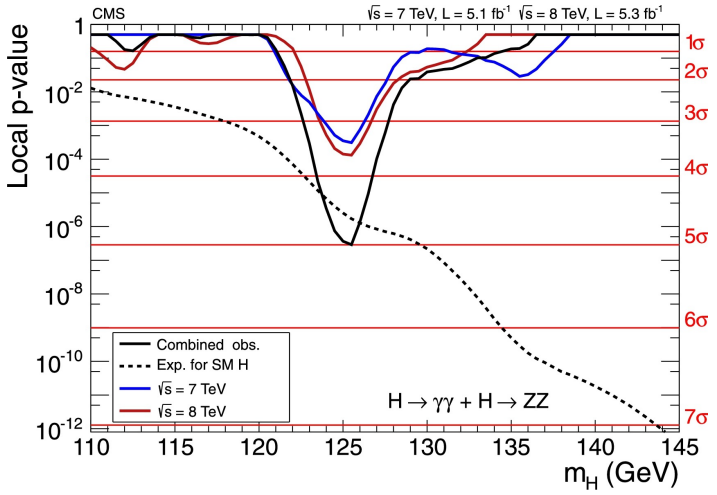
where $\sigma \times \mathcal{B}$ is the measured cross section times branching ratio and $\sigma_{SM} \times \mathcal{B}_{SM}$ is the corresponding SM expectation. The production modes considered were ggH, VBF, WH, ZH and $t\bar{t}H$ while the decay modes considered were ZZ, $\gamma\gamma$, W^+W^-, $\tau^+\tau^-$ and $b\bar{b}$. Without any extra assumptions, only the production mode multiplied by the branching ratio can be measured experimentally. It is possible to measure the individual production modes by assuming the decay modes are equal to their SM expectations. Similarly, it is possible to measure the individual decay modes by assuming the production modes are equal to their SM expectations. The combined results for the production and decay signal strengths with the ATLAS and CMS experiments using the Run 1 data set are shown in Figures 1.7 and 1.8. All results are found to be in good agreement with their SM predictions with the exception of the $t\bar{t}H$ process, where the measured signal strength was $\mu_{ttH} = 2.3^{+0.7}_{-0.6}$.

1.4.4 Current status

In the last decades the Standard Model has revealed itself as a theory capable to describe and predict the behaviour of the subatomic particles with great precision. However, the SM is a theory far from being complete as it does not include the gravitational interaction and there are fundamental issues that lead to the search for Physics beyond the Standard Model (BSM). For example, phenomena such as dark-matter and dark-energy [40], which compose 96% of the Universe, or the observed matter-antimatter asymmetry, remain unexplained by the SM. Moreover, the lightness of the discovered Higgs boson has some theoretical consequences, the most important



(a)



(b)

Figure 1.6: The local probability p_0 measured by a) ATLAS [37] and b) CMS [38] for a background-only experiment to be more signal-like than the observation, for 7 and 8 TeV data and their combination in the low mass range of a) $110 < m_{\gamma\gamma} < 150$ GeV and b) $110 < m_{\gamma\gamma} < 145$ GeV. The full curve gives the observed combined p_0. The dashed curve shows the median expected value under the hypothesis of a SM Higgs boson signal at that mass. The horizontal lines indicate the p_0 corresponding to significances of a) 0σ to 6σ and of b) 1σ to 7σ.

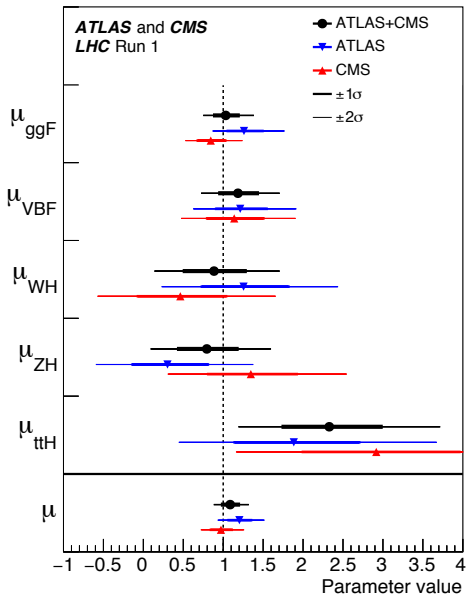


Figure 1.7: Best fit results for the production signal strengths for the combination of ATLAS and CMS data. Also shown are the results from each experiment. The error bars indicate the 1σ (thick lines) and 2σ (thin lines) intervals. The measurements of the global signal strength μ are also shown [39].

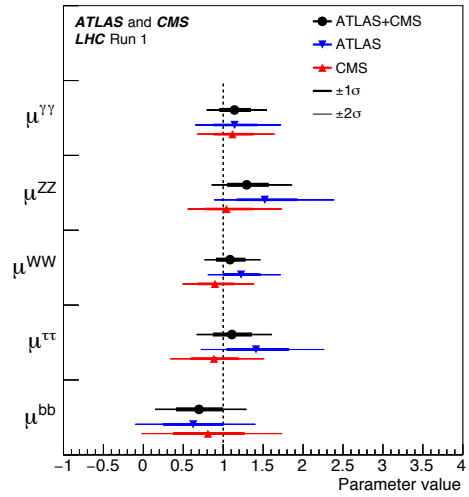


Figure 1.8: Best fit results for the decay signal strengths for the combination of ATLAS and CMS data. Also shown are the results from each experiment. The error bars indicate the 1σ (thick lines) and 2σ (thin lines) intervals [39].

one being the so-called naturalness or hierarchy problem, which is related to the level of the fine tuning of some of the free parameters of the theory. One-loop radiative corrections to the Higgs boson mass are quadratically divergent, and tend to make the Higgs boson heavier. The presence of new particles at the TeV scale would cancel these divergences reducing by far the fine tuning needed.

Part of the LHC program in the coming years is to measure the properties of the discovered Higgs boson with greater precision and compare them to the latest theoretical predictions. Deviations in cross section measurements, differences in differential cross sections, or new and more precise measurements of the Higgs boson couplings, can point to new physics.

1.5 Higgs boson pair production

Higgs boson pair production [41] is expected to play a key role in understanding the structure of the underlying potential that results in the BEH spontaneous symmetry breaking mechanism. Not only it is the simplest production process that is sensitive to the self-coupling λ, but it also provides one with a wealth of possibilities for probing higher-dimensional interactions as well as the existence of heavier states coupled to the Higgs.

Unfortunately, the rates for Higgs boson pair production at the LHC are quite small in the SM, so a measurement of the HH production cross sections necessitates considerable integrated luminosity even at 13–14 TeV centre-of-mass energy. However, new physics can produce sizable enhancements in several scenarios. These enhancements fall into two basic categories: Non-resonant (see Section 1.5.1) and resonant (see Section 1.5.2) production. In any case, precise predictions for rates and distributions are needed in order to be able to extract valuable information on λ or on new physics effects in general.

Analogously to single-Higgs boson production, several channels can lead to a final state involving two Higgs bosons. They entail the Higgs boson coupling to either the top quark (as in the case of gluon-gluon fusion and of $t\bar{t}H$ associated production), or vector bosons (in VBF, and in W and Z associated production), or both (for single-top associated production). The dominant production mechanism is gluon-gluon fusion via virtual top quarks, *i.e.*, box and triangle diagrams (see Figure 1.9), exactly as in the case of single-Higgs boson production. Cross sections corresponding to the other channels are at least one order of magnitude smaller, even though possibly interesting because of different sensitivity to λ or to new physics, and because of the possibility of exploiting a wider range of Higgs boson decay signatures. Figure 1.10

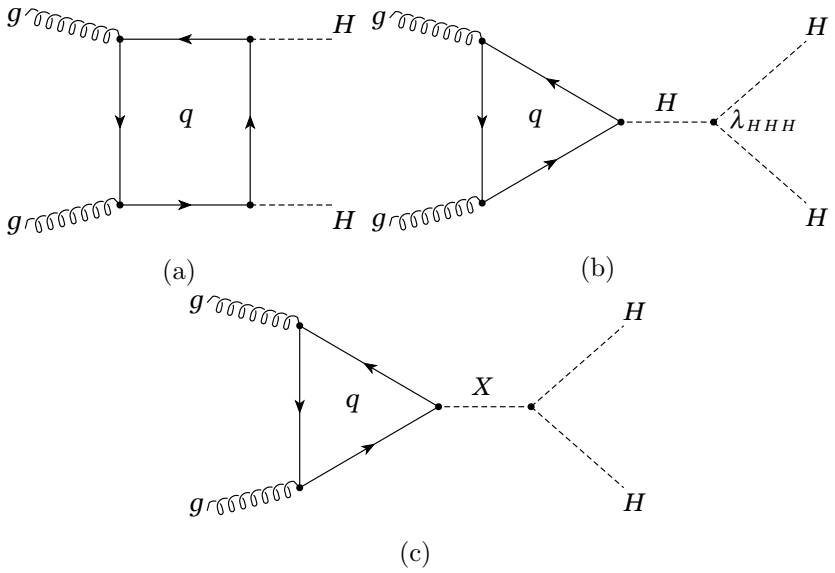


Figure 1.9: Feynman diagrams for the leading order production modes for SM Higgs boson pair production through: a) a heavy quark loop ("box diagram") and b) the Higgs boson self-coupling ("trilinear diagram"). BSM Higgs boson pair production can occur by changing the SM couplings in a) and b) or through c) an intermediate particle, X, called a resonance.

shows the predictions for the total rates at the NLO accuracy matched with parton shower for all the relevant Higgs-pair production channels in the SM at p-p colliders

with up to 100 TeV center-of-mass energy. The thickness of the curves corresponds to the scale and PDF uncertainties added linearly. It is observed that contrary to what happens in single-Higgs boson production, the top-pair associated channel is the third-largest starting at about $\sqrt{s} = 10$ TeV, and becomes the second-largest when center-of-mass energies approach $\sqrt{s} = 100$ TeV. The theoretical uncertainties due to scale variations in the three most important processes (gluon-gluon fusion, VBF, and associated production) are sizably reduced by the inclusion of the NLO corrections.

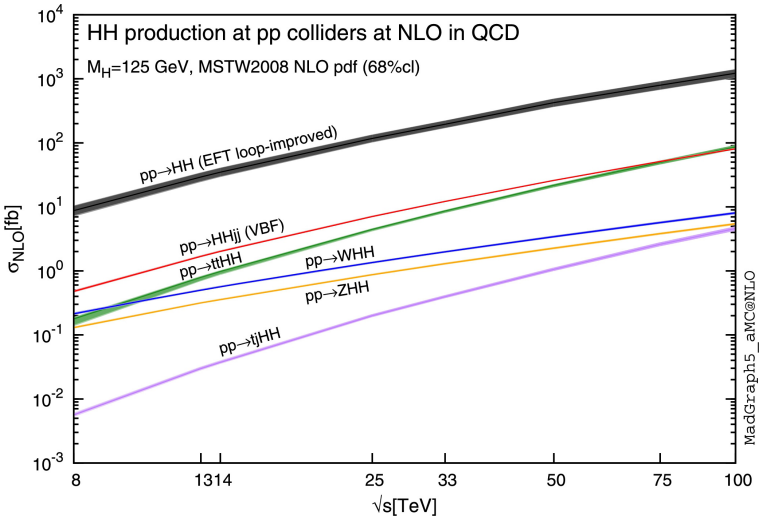


Figure 1.10: Total cross sections at the NLO in QCD for the six largest HH production channels at p-p colliders [41].

The branching ratios of the Higgs boson pair production channels are shown in Figure 1.11 and are extracted from [34] given $m_H = 125.09$ GeV. The search for Higgs boson pair production in the fully hadronic final state ($HH \to b\bar{b}b\bar{b}$) exploits the high $H \to b\bar{b}$ branching ratio, but it is characterized by a large multijet background. It is the analysis with the highest sensitivity at high resonance mass. The $HH \to \gamma\gamma b\bar{b}$ channel is characterized by a very high purity but a low statistic, as a result of exploiting both the $H \to \gamma\gamma$ to reduce QCD background and the $H \to b\bar{b}$ in the

attempt to have the largest possible signal yield. The search for $HH \to b\bar{b}\ell\ell$ includes both the $HH \to b\bar{b}W(\ell\nu)W(\ell\nu)$ and $HH \to b\bar{b}Z(\ell\ell)Z(\nu\nu)$ signals: in both cases, the presence of two neutrinos in the final state degrades the invariant mass resolution on the Higgs boson decaying to leptons. The main background is the top-pair production decaying fully leptonically. The $HH \to b\bar{b}\tau\tau$ represents a good compromise between the branching fraction and the background contamination. It is characterized by a poor $m_{b\bar{b}\tau\tau}$ resolution, degraded by the presence of neutrinos in the final state. The main backgrounds are $t\bar{t}$, Drell-Yan, and the multijet events.

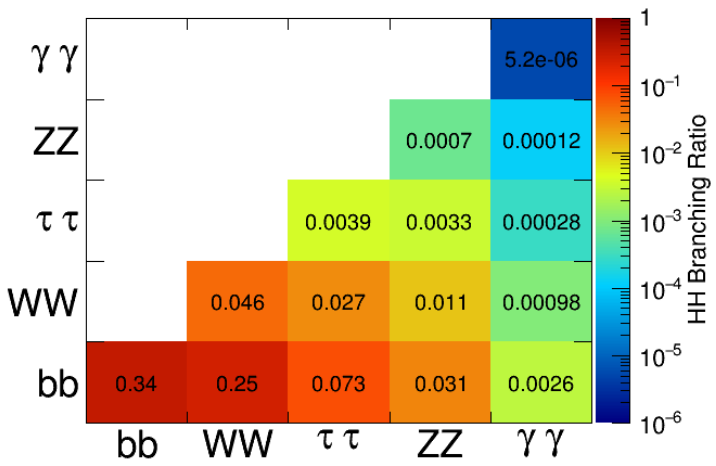


Figure 1.11: Branching ratios of the most relevant HH production channels given a Higgs mass of 125.09 GeV.

1.5.1 Non-resonant production

In theories beyond the SM, there can be contributions to the effective potential from dimension six Higgs boson operators that are induced by integrating out heavy degrees of freedom, or from compositeness. The Higgs boson mass and λ then are independent parameters, and the interactions of the Higgs boson with the electroweak gauge

bosons are modified from their SM values. The gluon-gluon fusion subprocesses that correspond to the box and the trilinear diagrams are the dominant HH production mechanisms. The interference of the two amplitudes is sensitive to the HHH coupling and thereby modifying any of the SM couplings would vary the expected Higgs boson pair production rate. Simply "turning off" the self-coupling λ_{HHH} would result in a doubling of the Higgs boson pair production rate in gluon-gluon fusion, and changing its sign would lead to a quadrupling of the rate [42].

Maximal destructive interference of the real amplitudes occurs at $\lambda^{HHH} \approx 2.45$ $\lambda^{HHH}_{\mathrm{SM}}$. Figure 1.12 shows the total LO and NLO cross sections for the six dominant HH production channels at the LHC with $\sqrt{s} = 14$ TeV, as a function of the self-interaction coupling λ. The dashed (solid) lines and light- (dark-) colour bands correspond to the LO (NLO) results and to the scale and PDF uncertainties added linearly. The SM values of the cross sections are obtained at $\kappa_\lambda = \lambda/\lambda_{SM} = 1$.

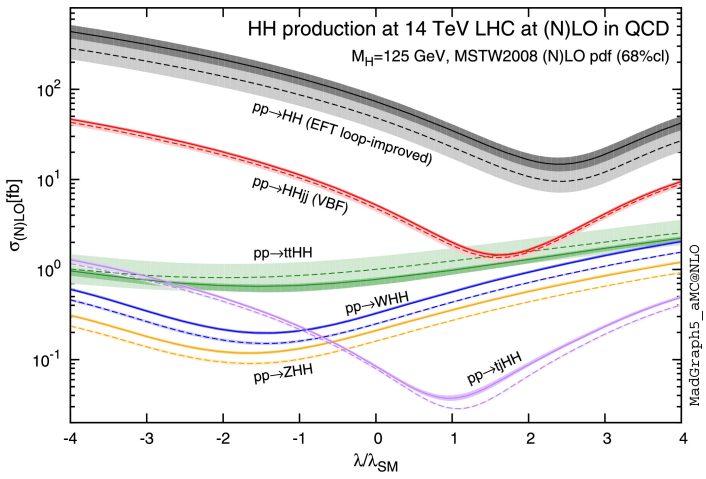


Figure 1.12: Total cross sections at the LO and NLO in QCD for HH production channels, at the $\sqrt{s} = 14$ TeV LHC as a function of the self-interaction coupling λ [41].

Enabling a direct $ttHH$ vertex would likewise lead to a significant enhancement of

the $pp \to HH$ rate [43]. This process would be similar to Figure 1.9b but with a single vertex and without the intermediate Higgs. There is also the option of including light colored scalars [44] among many other possibilities.

1.5.2 Resonant production

Resonant production allows for potentially significant enhancements with striking features that are well defined for a physics analysis. Many beyond-the-Standard-Model (BSM) theories predict the existence of heavy particles that can decay to a pair of Higgs bosons. These could then be identified as a resonance in the Higgs boson pair invariant mass spectrum. They could be produced, for example, through the ggH mode shown in in Figure 1.9c, where "X" is a new intermediate particle.

Models with two Higgs boson doublets [45], such as the minimal supersymmetric extension of the SM (MSSM) [46], twin Higgs boson models [47] and composite Higgs boson models [48, 49] involve the addition of a second complex scalar doublet. This implies the existence of a heavy Higgs boson that could then decay to two of its lighter, SM-like, partners.

Two-Higgs-doublet model

The two-Higgs-doublet model (2HDM) is one of the simplest extensions of the SM. The addition of the second Higgs boson doublet leads to a richer phenomenology as there are five physical scalar states: the CP-even neutral Higgs bosons h and H (where H is heavier than h by convention), the CP-odd pseudoscalar A and two charged Higgs bosons $H^{\pm}$. The current convention is to assume that the Higgs boson discovered at $m_H \sim 125$ GeV is the lightest CP-even neutral Higgs boson proposed by this model, h.

The couplings of the light boson h are SM-like for the alignment limit, $\beta - \alpha =$

$\pi/2$, where $\tan\beta$ is defined as the ratio between the vacuum expectation values (v) and α denotes the mixing angle of the two CP-even neutral Higgs bosons.

Two-Higgs-doublet models can introduce flavor-changing neutral currents which have not been observed so far. The Glashow-Weinberg condition, requiring that each group of fermions (u-type quarks, d-type quarks and charged leptons) couples exactly to one of the two doublets, is sufficient to avoid flavor-changing neutral currents.

Depending on which type of fermions couples to which doublet ϕ, one can divide two-Higgs-doublet models into different classes, stated in Table 1.3.

Type	Description	u-type quarks couple to	d-type quarks couple to	Charged leptons couple to
Type I	Fermiophobic	ϕ_2	ϕ_2	ϕ_2
Type II	MSSM-like	ϕ_2	ϕ_1	ϕ_1
X	Lepton-specific	ϕ_2	ϕ_2	ϕ_1
Y	Flipped	ϕ_2	ϕ_1	ϕ_2

Table 1.3: Classification of Two-Higgs-doublet models which lead to a natural flavor conservation. By convention, ϕ_2 is the doublet to which *up*-type quarks couple [50].

The type II 2HDM, characterized by u- and d-type quarks coupling to separate doublets, is by far the most studied, since it is the structure present in supersymmetric models. The most crucial difference between the general type II 2HDM and the MSSM is that the general type II 2HDM does not have a strict upper bound on the mass of the lightest Higgs boson, which is an important characteristic of the MSSM. In addition, the scalar self-couplings are now arbitrary. Another important difference is that the mixing parameter α, which in the MSSM is given in terms of $\tan\beta$ and the scalar and pseudoscalar masses, is now arbitrary. Finally, in the MSSM the charged-scalar and pseudoscalar masses are so close that the decay of the charged Higgs boson into a pseudoscalar and a real W boson is kinematically forbidden, while it is generally allowed in the type II 2HDM.

A very interesting study in the context of the CP-conserving two-Higgs-doublet model (2HDM) of type II probes that the hhh coupling cannot exceed its SM value,

but can be reduced by a factor of 0.56 at the 2σ level [51].

The role of MSSM heavy Higgs boson production is also studied and results suggest that the $H \to hh$ decay can nontrivially affect the h self-coupling measurement in low $\tan\beta$ regime when the mass of the heavy Higgs boson lies between 250-600 GeV and depending on the parameter space it may be seen as an enhancement of the self-coupling of 125 GeV Higgs boson [52].

Composite Higgs boson models

The composite Higgs boson models are speculative extensions of the SM where the discovered Higgs boson is not considered a fundamental particle but a bound state arising from an unknown interaction that breaks the EW symmetry. These models can solve the naturalness problem of the SM (see Section 1.4.4) since the Higgs boson is not an elementary particle so that a new energy scale exists that can be explained dynamically similarly to the mass of the proton. So far no direct or indirect signs that the Higgs boson or other SM particles are composite have been detected.

Other models

In an alternative model containing the SM Higgs boson doublet, ϕ, and an additional Higgs boson singlet, S, the Higgs boson pair production enhancement can be as large as a factor of $\sim$18 (13) for the mass of the heavy Higgs boson around 270 (420) GeV relative to the SM rate at 14 TeV for parameters corresponding to a global electroweak minimum [53].

The Randall-Sundrum model of warped extra dimensions [54] predicts both spin-0 radions and spin-2 gravitons that could couple to a Higgs boson pair in this way. The model is a braneworld theory developed while trying to solve the hierarchy problem of the SM. It involves a finite five-dimensional bulk that is extremely warped and

contains two branes: the Planckbrane (where gravity is a relatively strong force; also called "Gravitybrane") and the Tevbrane (our home with the SM particles; also called "Weakbrane"). In this model, the two branes are separated in the not-necessarily large fifth dimension by approximately 16 units (the units based on the brane and bulk energies). The Planckbrane has positive brane energy, and the Tevbrane has negative brane energy. These energies are the cause of the extremely warped spacetime.

1.5.3 ATLAS searches for Higgs boson pair production

The ATLAS collaboration has performed searches for resonant and non-resonant Higgs boson pair production in the $b\bar{b}b\bar{b}$, $\gamma\gamma b\bar{b}$, $b\bar{b}\tau\tau$ and $\gamma\gamma WW$ final states using the Run 1 data set at a center-of-mass energy of 8 TeV. The searches in these various final states were combined [55] and no evidence of HH production was found in both the resonant and non-resonant searches.

Figure 1.13 shows the combined observed and expected 95% CL upper limits on the resonant production cross section as a function of the resonance mass, m_X. For the low mass region of 260–500 GeV, both the $\gamma\gamma b\bar{b}$ and the $b\bar{b}\tau\tau$ analyses contribute significantly to the sensitivity of the search, while for masses above 500 GeV, the sensitivity of the analysis is almost entirely driven by the $b\bar{b}b\bar{b}$ final state.

The most significant excess observed in the resonant search is at a mass of 300 GeV, which has a local significance of 2.5σ, largely due to a 3.0σ excess observed in the $\gamma\gamma b\bar{b}$ analysis [56]. Figure 1.14 shows the diphoton invariant mass spectrum in data for the non-resonant $\gamma\gamma b\bar{b}$ analysis and the four body invariant mass distribution in data for the resonant $\gamma\gamma b\bar{b}$ analysis.

The observed upper limit placed on $\sigma\ (g\text{-}g \to X) \times \mathcal{B}\ (X \to HH)$ varies from 2.1 pb at 260 GeV to 0.011 pb at 1000 GeV.

For the non-resonant search, the observed and expected upper limits at 95% CL

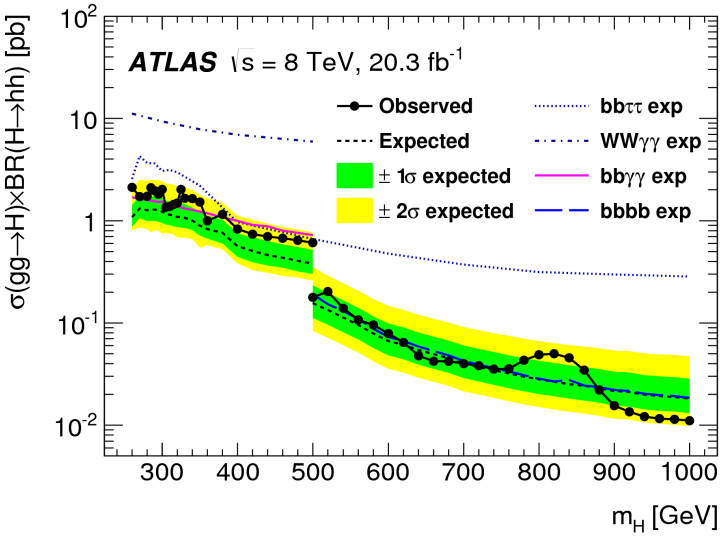
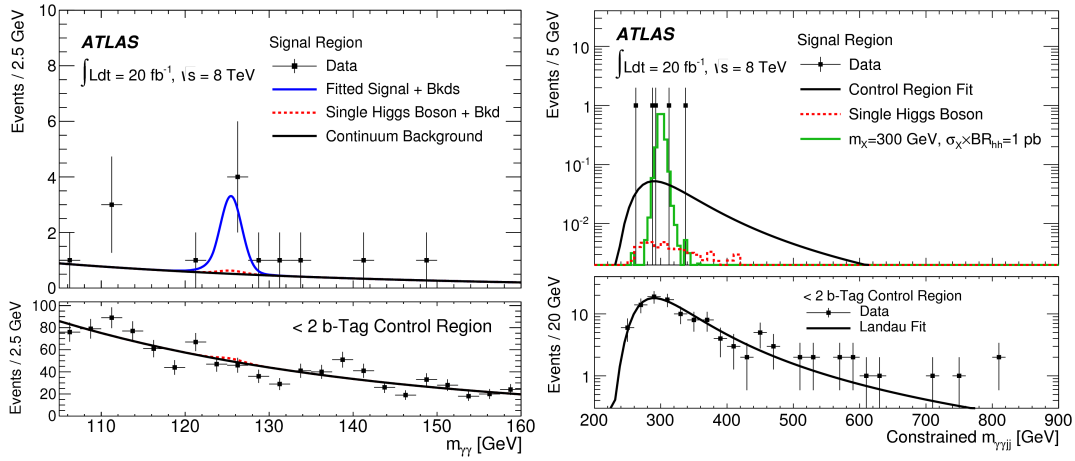


Figure 1.13: The combined observed and expected 95% CL upper limits on the resonant production cross section, $\sigma\ (g\text{-}g \to X) \times \mathcal{B}\ (X \to HH)$ as a function of m_X. The expected limits from the individual analyses are also shown [55].



(a) (b)

Figure 1.14: a) The diphoton invariant mass spectrum for the data and the corresponding fitted signal and background in the non-resonant $\gamma\gamma b\bar{b}$ search. The lower plot shows the diphoton invariant mass spectrum in data for events with fewer than two b-tagged jets. The shape of the fitted function in this control region is used in the upper plot. b) The four body invariant mass for data events in the resonant $\gamma\gamma b\bar{b}$ search with the expected backgrounds shown [56].

on the cross section of SM Higgs boson pair production are shown for each individual analysis and for the combination in Table 1.4. The combined observed (expected) upper limit on σ (g-$g \to HH$) is 0.69 (0.47) pb, corresponding to 70 (48) times the cross section predicted by the SM. The $HH \to b\bar{b}b\bar{b}$ analysis has the best expected sensitivity followed by the $HH \to \gamma\gamma b\bar{b}$ analysis. The observed combined limit is weaker than expected due to the excess in the $HH \to \gamma\gamma b\bar{b}$ which can be seen in Figure 1.14.

Table 1.4: The expected and observed 95% CL upper limits on the cross sections of non-resonant Higgs boson pair production at $\sqrt{s} = 8$ TeV from the individual analyses and their combinations. SM values are assumed for the Higgs boson decay branching ratios. The cross section limits normalized to the SM value are also included.

Analysis	$\gamma\gamma b\bar{b}$	$\gamma\gamma WW$	$bb\tau\tau$	$b\bar{b}b\bar{b}$	Combined
	Upper limit on the cross section [pb]				
Expected	1.0	6.7	1.3	0.62	0.47
Observed	2.2	11	1.6	0.62	0.69
	Upper limit on the cross section relative to the SM prediction				
Expected	100	680	130	63	48
Observed	220	1150	160	63	70

The search for a generic scalar decaying to a pair of Higgs bosons in the $\gamma\gamma b\bar{b}$ final state is one of the aims of this thesis. Details can be found in Chapter 5.

Chapter 2

The LHC and the ATLAS Experiment

This chapter introduces the experimental setup used for the development of the work described in this thesis. It starts with a general description of the LHC machine and its collision conditions. Both Run 1 and Run 2 data-taking periods are covered. It continues with a detailed explanation of the ATLAS detector, including the subsystems that compose it and the data acquisition process, essential to achieve the physics goals set after the discovery of the Higgs boson.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [57] is a two-ring-superconducting-hadron accelerator and collider located at CERN (European Organization for Nuclear Research). Situated underground, it is approximately circular with a circumference of 27 km and it straddles the border between France and Switzerland just outside Geneva. Its

primary goal is to deliver proton-proton collisions[1] at unprecedented conditions of energy and luminosity, with the key objectives of exploring the Standard Model in the TeV energy scale, and searching for the Higgs boson and for potential new physics beyond the Standard Model. In what follows, different aspects of the LHC will be discussed, using the information available in [57–60].

Protons are injected into the LHC using a chain of particle accelerators as shown in Figure 2.1. Hydrogen atoms are ionized and the protons are accelerated to 50 MeV by the Linac, a linear accelerator. The protons continue to the Proton Synchrotron Booster (PSB), followed by the Proton Synchrotron (PS) and Super Proton Synchrotron (SPS); each accelerates the protons to energies of 1.4 GeV, 25 GeV, and 450 GeV respectively, before they are delivered to the LHC. The PS and SPS deliver protons in "bunches" separated by roughly 25 ns each; in total 2808 bunches can be injected into the LHC ring. There are about 10^{11} protons per bunch. Once inside the LHC, the proton bunches are kept circulating using superconducting magnets cooled to 1.9 K with liquid helium. Radiofrequency (RF) cavities with a maximum oscillation of 400 MHz are used to accelerate the particles from 450 GeV to a maximum of 7 TeV.

The LHC has eight arcs and eight straight sections between the arcs. The straight sections are around 528 m long. Each straight section has associated with it surface and underground installations, lifts, and a wide variety of technical infrastructure. These locations are referred to as the LHC points and are illustrated in Figure 2.2. The four main experiments are situated at point 1 (ATLAS [62]), point 2 (ALICE [63]), point 5 (CMS [64]) and point 8 (LHCb [65]). Injection of clockwise (anti-clockwise) beam takes place at point 2 (8). The main collimator installations are at points 3 and 7, the RF system is situated at point 4 and the beam dump system is situated at point 6. Importantly each sector can be cooled and powered independently.

[1]The LHC also collides lead ions, but those will not be discussed in this thesis.

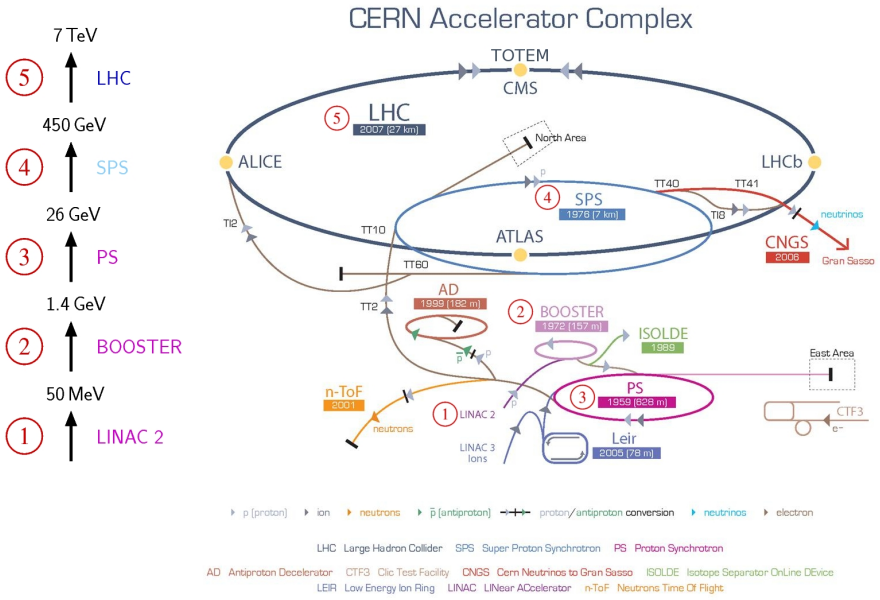


Figure 2.1: The Large Hadron Collider, its injection chain, and the four main experiments, ATLAS, ALICE, CMS, and LHCb, to which it delivers collisions [61].

The energy of the LHC is essentially determined by the bending radius of the ring and the achievable dipole field strength. Instantaneous luminosity is defined as the number of p-p interactions per second and unit surface area. This can be written in terms of the beam parameters in the following way:

$$L = \frac{1}{4\pi} \frac{N^2 \; f_{rev} \; n_{bunch}}{\sigma_x \sigma_y \; t} \sim 10^{34} \; cm^{-2} \; s^{-1}, \tag{2.1}$$

where $N = 10^{11}$ is the number of protons per bunch (squared because a proton could interact with any proton in the oncoming bunch), $f_{rev} = 11245$ Hz is the LHC revolution frequency, $n_{bunch} = 2808$ is the number of bunches and σ_x and σ_y are the transversal dimensions of the bunch, which is of the order of a few μm at the interaction point (IP).

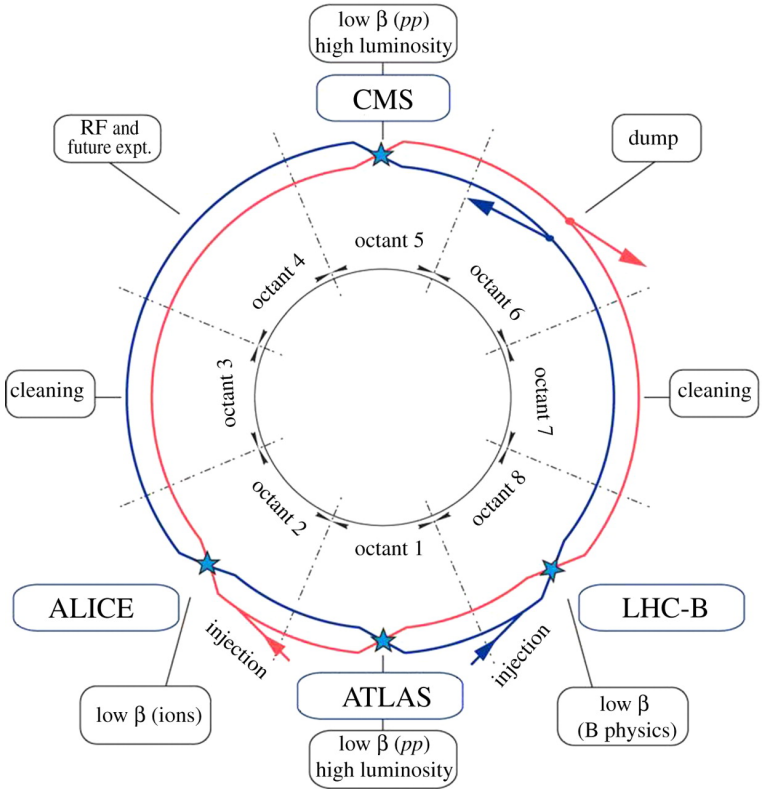
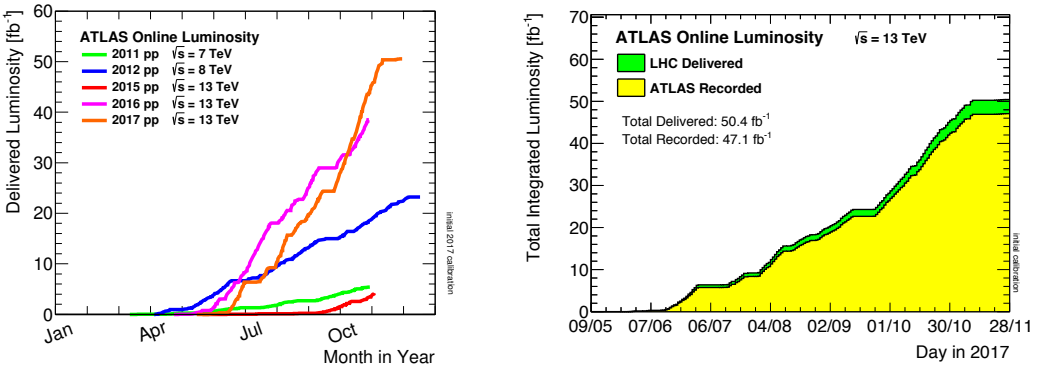


Figure 2.2: Layout of the main LHC ring [66].

High instantaneous luminosity is achieved by a large number of bunches, high bunch population and small beam size at the interaction points. The peak instantaneous luminosity in 2015 was 5.0×10^{33} cm^{-2} s^{-1}; already in 2016 the peak luminosity exceeded 1×10^{34} cm^{-2} s^{-1}, its design luminosity, and in 2017 a peak luminosity of twice the design luminosity was achieved. Figure 2.3a reports the integrated luminosity delivered to ATLAS in each year of data taking from 2011 to 2017, and Figure 2.3b shows the cumulative luminosity delivered to and recorded by ATLAS in the year 2017.

As a consequence of the high instantaneous luminosity delivered by the LHC, the number of inelastic *p*-*p* interactions in a single bunch crossing at the detector IP is



(a) (b)

Figure 2.3: a) Cumulative p-p collision luminosity delivered to the ATLAS detector versus month of the year, separately for years between 2011 and 2017 and b) cumulative luminosity versus time delivered to (green) and recorded by ATLAS (yellow) during stable beams for p-p collisions at 13 TeV centre-of-mass energy in 2017 [67].

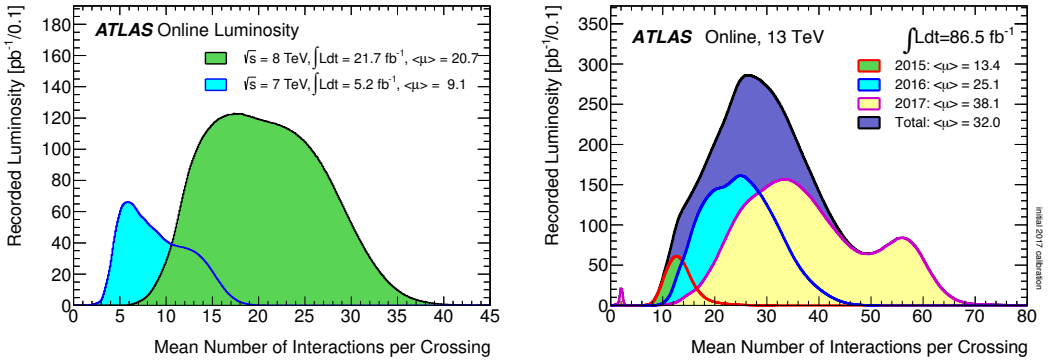
larger than one. This phenomenon is referred to as in-time pile-up. The mean number of interactions per crossing corresponds to the mean of the Poisson distribution on the number of interactions per crossing calculated for each bunch. It is calculated from the instantaneous per bunch luminosity as:

$$\mu = \frac{L_{bunch}\ \sigma_{inel}}{f_{rev}}, \tag{2.2}$$

where L_{bunch} is the per bunch instantaneous luminosity, σ_{inel} is the inelastic cross section which we take to be 71.5 mb for 7 TeV collisions, 73.0 mb for 8 TeV collisions and 80.0 mb for 13 TeV collisions, and f_{rev} is the LHC revolution frequency.

Figure 2.4a shows the distributions of the mean number of interactions per bunch crossing in ATLAS, for the 2011 and 2012 data-taking periods (there was no pile-up in 2010). In 2012, the larger instantaneous luminosity caused a significant increase in pile-up, with respect to 2011. A similar behavior is observed for the 2015, 2016 and 2017 data-taking periods, as illustrated in Figure 2.4b. In 2017 the average number

of interactions per crossing was maximum at a value of $\langle\mu\rangle = 38.1$, with tails up to ~ 70 interactions.



(a) (b)

Figure 2.4: a) The luminosity-weighted distribution of the mean number of interactions per crossing for the 2011 and 2012 p-p collision data at 7 and 8 TeV, respectively, centre-of-mass energy [68] and b) for the 2015, 2016 and 2017 p-p collision data at 13 TeV centre-of-mass energy [67].

2.2 The ATLAS Experiment

The ATLAS (A Toroidal LHC Apparatus) [69] experiment is a general-purpose detector designed to reconstruct events from colliding hadrons at the LHC. ATLAS is about 44 meters long, more than 25 meters high and has an overall weight of approximately 7000 tonnes. The detector has a cylindrical symmetry providing coverage of almost 4π in solid angle.[2]

The ATLAS detector measures the energy and momentum of light, electromagnetically interacting particles (electrons and photons), hadronic jets, and muons. To achieve this, the detector is composed of several nested, cylindrical sub-systems, as shown in Figure 2.5. Closest to the IP, a tracking system is used to measure the

[2]Solid angle is the two-dimensional angle in three-dimensional space. 4π equates to a surface that completely covers a sphere.

momentum of charged particles while absorbing as little of a particle's energy as possible. Next, a system of calorimeters is used to stop electrons, photons and hadronic jets, measuring their energy in the process. An electromagnetic calorimeter is closer to the IP, and absorbs nearly all the energy from electrons and photons. Hadronic jets are stopped by the electromagnetic and hadronic calorimeters.

Muons escape the calorimeter; the outermost layer of the ATLAS detector is a spectrometer composed of tracking detectors and toroidal magnets designed to measure the muons momentum. The sub-systems are designed in ways to make it easier to identify electrons, photons, hadronic jets, taus and muons and discriminate between the different classes in order to understand the fundamental interaction in the p-p collision.

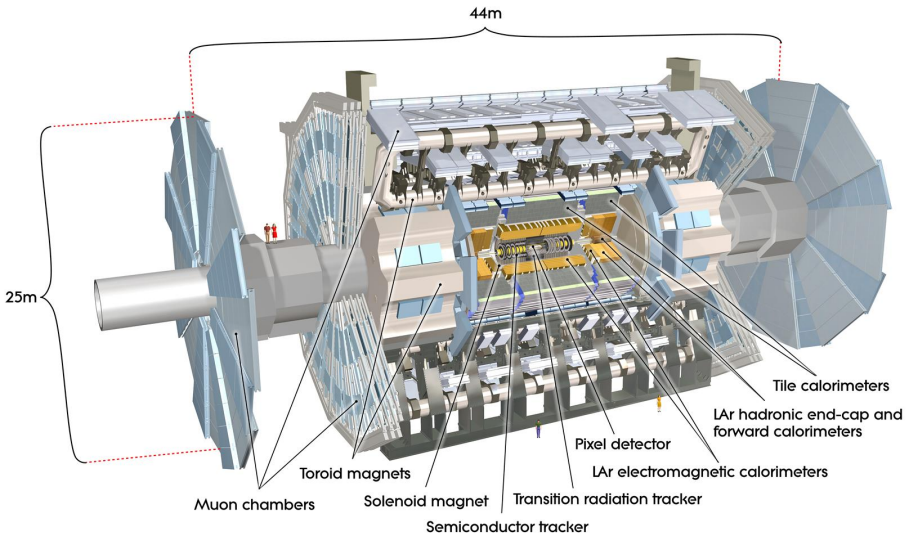


Figure 2.5: Cut-away view of the ATLAS detector [69].

ATLAS uses a right-handed coordinate system, with the origin placed at the IP. The beam direction defines the z-axis. The A-side (C-side) of the detector is defined as that with positive (negative) z. The positive x-axis is defined as pointing

to the center of the LHC ring, and the positive y-axis points upwards. The $x-y$ plane is orthogonal to the beam direction and referred to as the transverse plane. Quantities such as transverse momentum, p_{T}, or transverse energy, E_{T}, are defined in this transverse plane. The azimuthal angle ϕ is measured around the beam axis and the polar angle θ is the angle from the beam axis. In the limit where the particle is travelling close to the speed of light, or equivalently in the approximation that the mass of the particle is negligible, the rapidity, y, and pseudorapidity η are defined as:

$$y = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right) \text{ and } \eta = -\ln\tan\left(\frac{\theta}{2}\right), \tag{2.3}$$

where E denotes the energy and p_z the component of the momentum along the beam direction. In the limit of massless particles, $y = \eta$, and both are invariant under Lorentz boosts along the beam axis. The distance ΔR in the $\eta - \phi$ coordinate space is commonly used and is defined as:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \tag{2.4}$$

2.2.1 Magnet System

ATLAS features a unique hybrid system of a central superconducting solenoid and three outer superconducting toroids. This magnetic system is 22 m in diameter and 26 m in length, with a stored energy of 1.6 GJ. Figure 2.6 shows the general layout with the four superconducting magnets which provide a magnetic field over a volume of approximately 12,000 m^3.

The solenoid provides the Inner Detector with a 2 T axial magnetic field. Crucially, this high field strength is obtained whilst keeping the solenoid thin in order to reduce the material in front of the electromagnetic calorimeter. The solenoid is aligned to the beam axis, it is 5.8 m in length and has an outer diameter of 2.56 m.

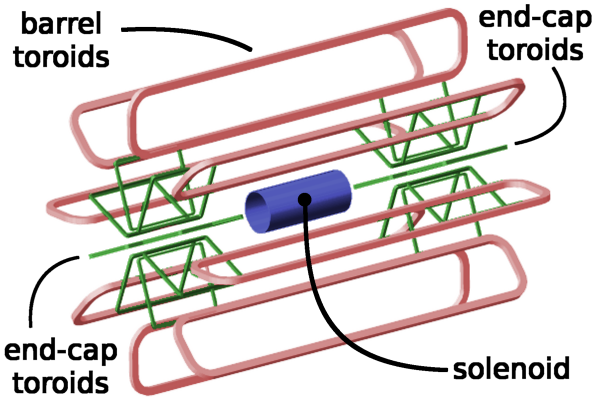


Figure 2.6: Schematic diagram of the ATLAS magnet system [70].

The toroid system is divided into three regions, the barrel and two end-caps. The barrel region is constructed from eight coils and produces a toroidal magnetic field of approximately 0.5 T for the central muon detectors. The end-cap toroids are also constructed from 8 coils each and produce a magnetic field of approximately 1 T for the muon detectors in the end-cap regions.

2.2.2 Inner Detector

The Inner Detector (ID), illustrated in Figure 2.7, is the innermost layer of ATLAS. The ID is composed of three different technologies which decrease in granularity the further away from the interaction point. The Pixel detector is closest to the interaction point and is composed of four layers of silicon pixel detectors. This is followed by the Semi-Conductor Tracker (SCT) which is built from four layers of silicon strip detectors. Finally, straw tubes make up the Transition Radiation Tracker (TRT) which surrounds the Pixel and SCT detectors.

The Pixel and SCT detectors are divided along the beam axis into a barrel region and two end-cap sections. In the barrel region, the detectors are arranged in concentric cylinders around the beam axis while in the end-cap regions they are located on disks

perpendicular to the beam axis. In this way, the Pixel and SCT detectors cover the region $|\eta| < 2.5$.

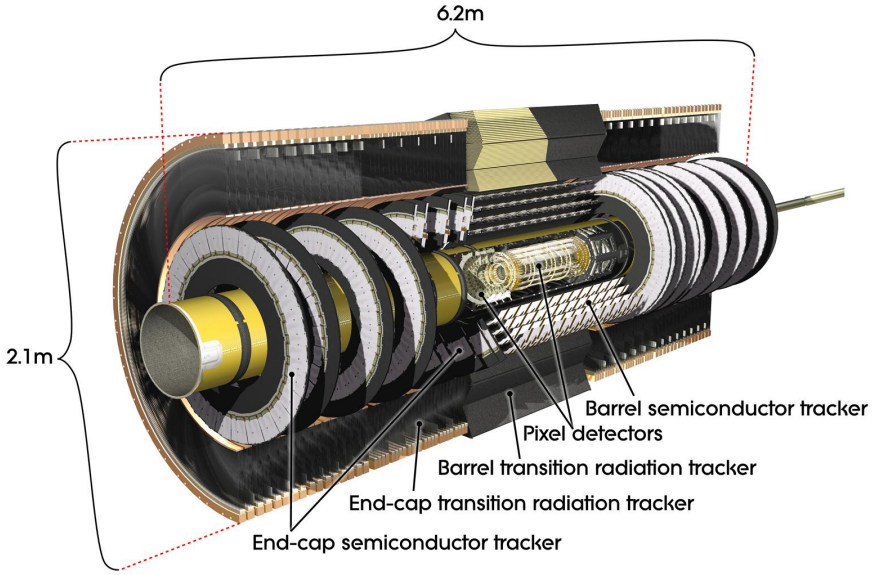
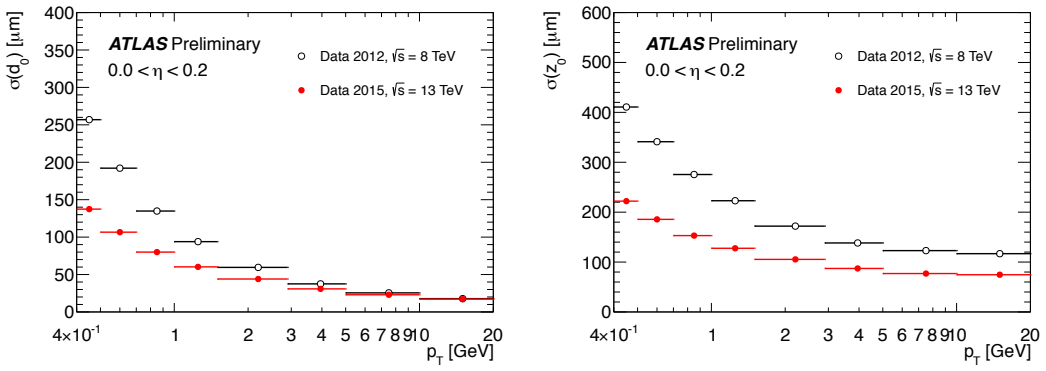


Figure 2.7: Cut-away view of the ATLAS Inner Detector [69].

The Pixel detector is the first detector that a particle will cross from beam interactions. The sensors are finely segmented in $R - \phi$ and z with typically four pixel layers crossed by each particle. In the first layer, the segmentation is 50 μm ($R - \phi$) $\times$ 250 μm (z), while in the other layers it is 50 μm ($R - \phi$) $\times$ 400 μm (z). This high level of segmentation gives precise space point measurements but requires a very large number of readout channels ($>$ 80 million).

The innermost pixel layer in the barrel, the Insertable B-Layer (IBL) was installed between Run 1 and Run 2. The IBL was inserted at a radius of 33 mm from the interaction point. The first layer of the original Pixel detector was located at a radius of 50 mm. In order to have the space to install the IBL, a new beryllium beam pipe was also installed with a reduced diameter (from 59 mm to 47 mm). The original

Pixel detector was designed for a peak luminosity of $1.0 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$, which has been exceeded in Run 2. This upgrade was essential to provide an improvement in tracking, vertex reconstruction and b-tagging performance. Figure 2.8 shows the improvement of resolution on the transverse impact parameter (the distance d_0 in the $x-y$ plane between the track closest point to the z-axis and the z-axis itself) and on the longitudinal impact parameter (the z coordinate z_0 of the closest point of approach to the z-axis) determination as a function of p_{T} due to the additional IBL pixel layer.



(a) (b)

Figure 2.8: a) Transverse impact parameter and b) longitudinal impact parameter resolution measured with data taken in 2015, at $\sqrt{s} = 13$ TeV, with the Inner Detector including the IBL, as a function of p_{T}, for values of $0.0 < |\eta| < 0.2$, compared to that measured with 2012 data, at $\sqrt{s} = 8$ TeV [71].

In the SCT detector, typically eight strip layers (four space points) are crossed by each track. In the barrel region, the detector uses small-angle (40 mrad) stereo strips to measure both coordinates, with one set of strips in each layer parallel to the beam direction, measuring $R - \phi$. They consist of two 6.4 cm long daisy-chained sensors with a strip pitch of 80 μm. In the end-cap region, the detectors have a set of strips running radially and a set of stereo strips at an angle of 40 mrad. The mean pitch of the strips is also approximately 80 μm. The total number of readout channels in the

SCT is approximately 6.3 million.

A large number of hits is provided by the 4 mm diameter Xe-based gas mixture straw tubes of the TRT, which enables tracking up to $|\eta| = 2.0$. The TRT only provides $R - \phi$ information, for which it has an intrinsic accuracy of 130 μm per straw. In the barrel region, the straws are parallel to the beam axis and are 144 cm long, with their gold-plated wires divided into two halves, approximately at $\eta = 0$. In the end-cap region, the 37 cm long straws are arranged radially in wheels. The total number of TRT readout channels is approximately 351,000. Significant modifications of the TRT detector were made for Run 2 mainly to improve response to the expected much higher rate of hits and to mitigate leaks of the Xe-based active gas mixture. Many gas leaks were repaired and the gas system was modified to use a cheaper Ar-based gas mixture in some channels.

The overall ID momentum resolution that was achieved during Run 1 (so before the IBL insertion) can be approximated by the formula:

$$\frac{\sigma_{p_{\mathrm{T}}}}{p_{\mathrm{T}}} = 0.05\% \cdot p_{\mathrm{T}} \oplus 1\% \tag{2.5}$$

2.2.3 Calorimeters

The ATLAS calorimeter system consists of several different electromagnetic and hadronic detectors, with full ϕ symmetry. They are placed around the ID and the solenoid magnet, with a layout shown in Figure 2.9. The innermost layer of the calorimeter system, the so-called Liquid Argon (LAr) calorimeter, is composed of one electromagnetic (EM) calorimeter in the barrel (EMB), one EM calorimeter (EMEC) and one hadronic calorimeter (HEC) at each end-cap; and a forward calorimeter (FCal), that covers the regions closest to the beam. The TileCal is a hadronic calorimeter, that forms the outer layer of the system, and it is composed of

one central barrel (LB) and two extended barrels (EB) on each side.

Calorimeter showers can be characterized according to their longitudinal depth and its lateral width, which vary according to the type of incident particle and the nature of the absorber used in the calorimeters. Electrons and photons passing through

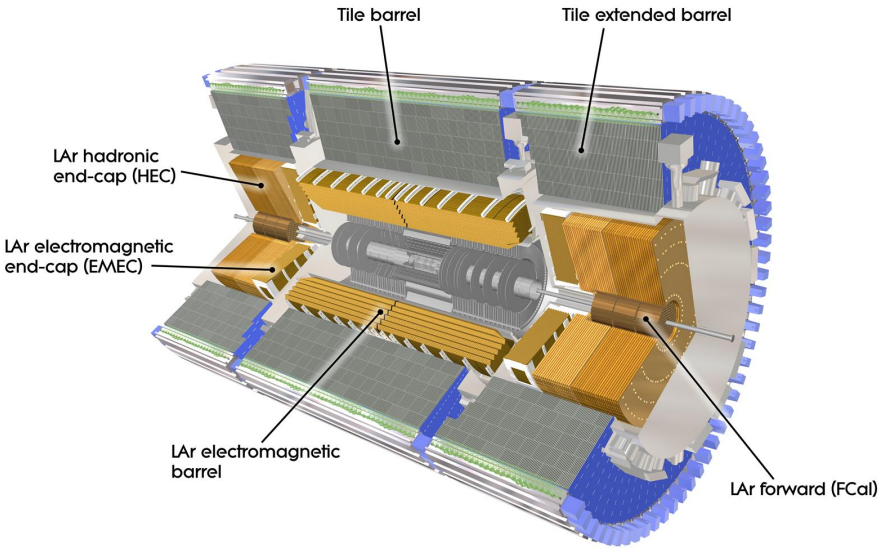


Figure 2.9: Cut-away view of the ATLAS calorimeter system [69].

an absorber will shower electromagnetically: photons produce e^+e^- pairs and electrons emit bremsstrahlung photon radiation. The daughter electrons and photons also interact, resulting in a particle shower. Most of the energy will have been absorbed after traversing about 20 radiation lengths X_0 of absorber (longitudinal depth). The lateral width of the shower in a material is characterized by its Molière radius, the radius of a cone in which 90% of the shower energy is contained.

Hadrons, on the other hand, shower mostly via strong interactions. The shower develops in a similar cascade of decay products, about 1/3 of which are π^0 which decay electromagnetically, initiating an electromagnetic shower component to the

total shower. Hadronic showers are characterized by a core energy deposit, partially due to the electromagnetic showering, and a larger tail of energy deposition.

The LAr and Tile calorimeters are both sampling calorimeters, alternating absorber and active layers. The LAr was specifically designed to aid electron and photon discrimination against hadronic jets. Its relatively small Molière radius (9 cm) results in compact electromagnetic showers from electrons and photons. These local shower deposits can be identified against the more diffuse showers of hadronic jets.

The Liquid Argon Electromagnetic Calorimeter

The EM calorimeter uses liquid argon as the active material and lead plates as the absorber. Liquid argon was chosen for its intrinsic linear behavior, stability of response over time, and radiation-hardness. The lead-LAr layers have an accordion-shaped geometry that provide complete ϕ symmetry without azimuthal cracks.

The EM calorimeter is divided into a barrel part (EMB, $|\eta| < 1.475$) and two end-cap components (EMEC, $1.375 < |\eta| < 3.2$), each housed in their own cryostat held at around 90 K. The EMB calorimeter consists of two identical half-barrels, separated by a small gap of 4 mm at $z = 0$. Each EMEC calorimeter is mechanically divided into two coaxial wheels: an outer wheel covering the region $1.375 < |\eta| < 2.5$, and an inner wheel covering the region $2.5 < |\eta| < 3.2$. The total thickness of the EM calorimeter is > 22 radiation lengths (X_0) in the barrel and $X_0 > 24$ in the end-caps. In the region of $|\eta| < 1.8$, a presampler detector is used to correct for the energy lost by electrons and photons upstream of the calorimeter. The presampler consists of an active LAr layer of thickness 1.1 cm (0.5 cm) in the barrel (end-cap) region.

Figure 2.10 illustrates the geometry of the EM calorimeter in the barrel region, showing it is finely segmented, both laterally and longitudinally, into cells (individual readout elements) of varying sizes. The lead thickness in the absorber plates has been

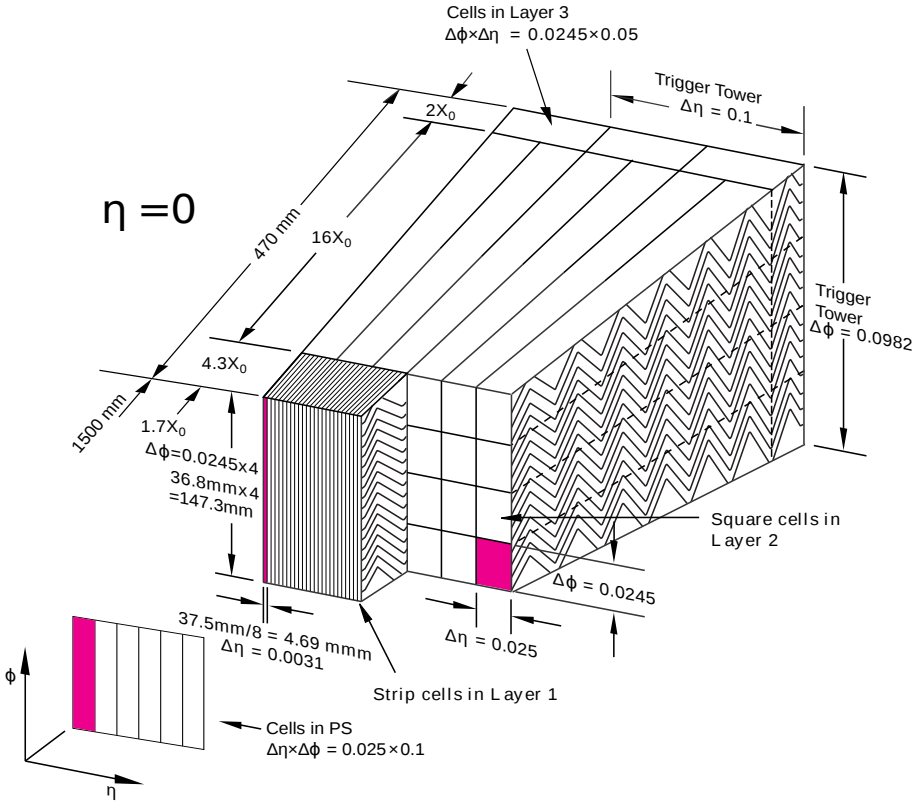


Figure 2.10: Schematic diagram of the cross section of the EM barrel calorimeter, including the presampler (labelled "PS") [72].

optimized as a function of η in terms of EM calorimeter performance in energy resolution. Over the region devoted to precision physics ($|\eta| < 2.5$), the EM calorimeter is segmented in three layers in depth. The first layer is very finely segmented in η, providing accurate position measurement. The second layer has a segmentation of 0.025×0.025 ($\Delta\eta \times \Delta\phi$), and collects the largest fraction of the energy of the EM shower. The third layer only collects the tail of the electromagnetic shower and is, therefore, less segmented in η. For the end-cap inner wheel ($2.5 < |\eta| < 3.2$), the calorimeter is segmented in two sections in depth, with a coarser lateral granularity. The transition region between the LAr barrel and end-caps located at roughly 1.37

$< |\eta| < 1.52$ is dedicated to detector services and is thus not fully instrumented. The EM calorimeter is designed to have an energy resolution of $\sigma/E \approx 10\%/\sqrt{E(\text{GeV})} \oplus 0.7\%$.

The Tile Calorimeter

The Tile Calorimeter (TileCal) is the central section of the hadronic calorimeter of the ATLAS experiment. TileCal covers the $|\eta| < 1.7$ region of the detector. Extending from an inner radius of 2.28 m to an outer radius of 4.25 m, the TileCal has a thickness of 9.7 interaction lengths (λ) at $\eta = 0$, providing containment of the hadronic showers produced in the LHC collisions. It consists of one barrel and two extended barrel sections and surrounds the LAr barrel electromagnetic and end-cap hadronic calorimeters. The Long Barrel (LB) and Extended Barrel (EB) sections roughly correspond to $|\eta| < 1.0$ and $|\eta| > 1.0$, respectively. The crack region between the TileCal LB and EB and the LAr EMB and hadronic end-cap sections are covered with special cells made by scintillators (E-cells).

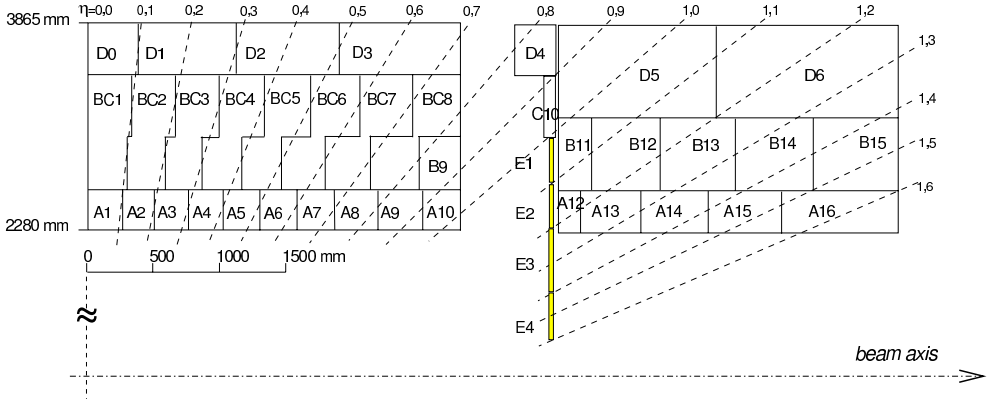


Figure 2.11: The layout of the TileCal cells, denoted by a letter (A to E) plus a number (integer). The A-layer is closest to the beam-line. The E-cells appear in yellow. The naming convention is repeated on each side of $\eta = 0$ [69].

The TileCal provides important information for reconstruction of hadrons, jets,

hadronic decays of tau-leptons and missing transverse energy. This sampling calorimeter uses iron plates as absorber and scintillating tiles as active medium. For precision measurements involving the reconstruction of jets, the TileCal is designed to have an energy resolution for jets of $\sigma/E \approx 50\%/\sqrt{E(\mathrm{GeV})} \oplus 3\%$. The calorimeter readout is segmented into about 5000 cells (longitudinally and transversally), each of them being read by two photomultipliers (PMTs). TileCal cells and electronics are organized into 4 partitions, LBA and LBC for the A-side and C-side of the barrel region, and separate EBA and EBC partitions in the extended barrel region. Each partition is divided into 64 symmetric slices (modules) of $\Delta\phi \sim 0.1$, with 45 instrumented channels in LB modules and 32 channels in EB modules. Both LB and EB have up to three layers: the A-layer being the closest to the beam axis, followed by the B(C) and D-layers. The η and radial structure of the TileCal cells is shown in Figure 2.11. The so-called Intermediate Tile Calorimeter (ITC) cells (D4, C10 and E-cells) are located between the LB and EB, and provide a coverage in the range $0.8 < |\eta| < 1.6$. The E-cells, also called gap- and crack-cells, cover the range $1.0 < |\eta| < 1.6$ and are partly closer to the beam axis than the A-layer cells, which exposes them to high radiation. They are only composed of the scintillator and are exceptionally read out by only one PMT.

The TileCal has three key elements: the optical part (scintillators and fibers), the PMTs and the read-out electronics. Figure 2.12 shows a schematic of the assembly of these components. The light is produced in scintillating tiles, collected and routed by wavelength shifting (WLS) fibers and converted into electric currents by the PMTs. The sampling and the digitization is realized by Analog to Digital Converters (ADCs). The signal of the PMTs is then shaped and amplified with two gains. In order to reduce the total data bandwidth during the normal operation only one of the two gains and only seven consecutive digital samples are read out, taken every 25 ns. Trigger

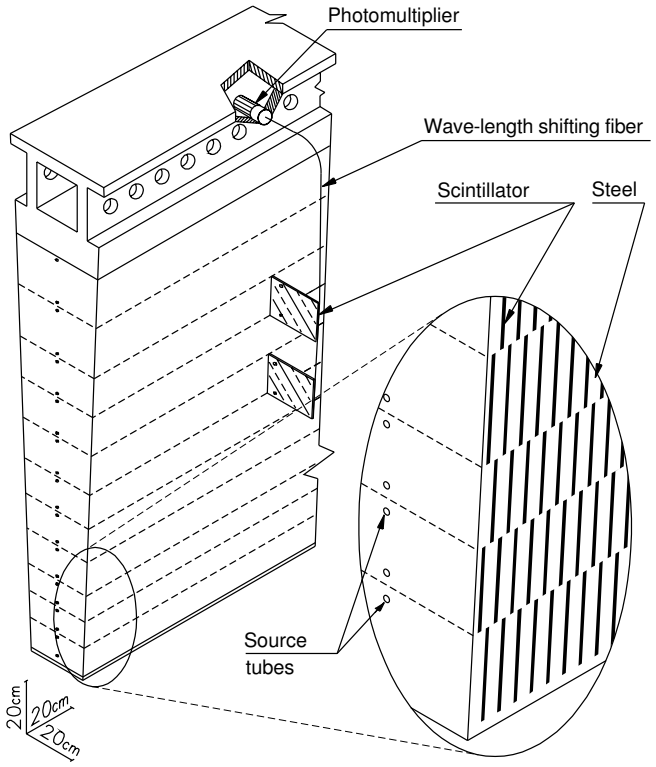


Figure 2.12: A schematic of a Tile Calorimeter wedge-shaped module showing the plastic scintillators sandwiched between steel absorbers. The front-end electronics drawer is located in the outer radius to read out the photomultiplier tubes [69].

signals are formed by an analogue sum of input signals and sent to the calorimeter trigger system, which also considers input from other calorimeters. During collisions, if an event is selected by the trigger system, the digitized signals are collected and processed by a Read-Out Driver (ROD) system in the back-end electronics system. The data flow rate is controlled using a busy feedback signal from the back-end electronics to the Central Trigger Processor (CTP). The busy signal is generated by the RODs modules when the input buffers are full. This signal is transmitted to the Trigger and Busy Module (TBM) and ROD busy module which distributes it to the CTP, informing that it is not possible to accept new events. In parallel to

this, integrators measure the integrated current from the PMTs which is used for calibration with a cesium source and to measure the rate of soft interactions during LHC collisions.

2.2.4 Muon Spectrometer

The ATLAS Muon Spectrometer (MS), whose layout is shown in Figure 2.13, occupies the outermost part of the ATLAS detector. It consists of monitored drift tubes (MDTs) for precision tracking in the spectrometer bending plane, Resistive Plate Chambers (RPCs) and Thin Gap Chambers (TGCs) for triggering in barrel and end-cap, respectively, and Cathode Strip Chambers (CSCs) for precision measurements in the high-rate end-cap inner layer where MDTs would have occupancy problems.

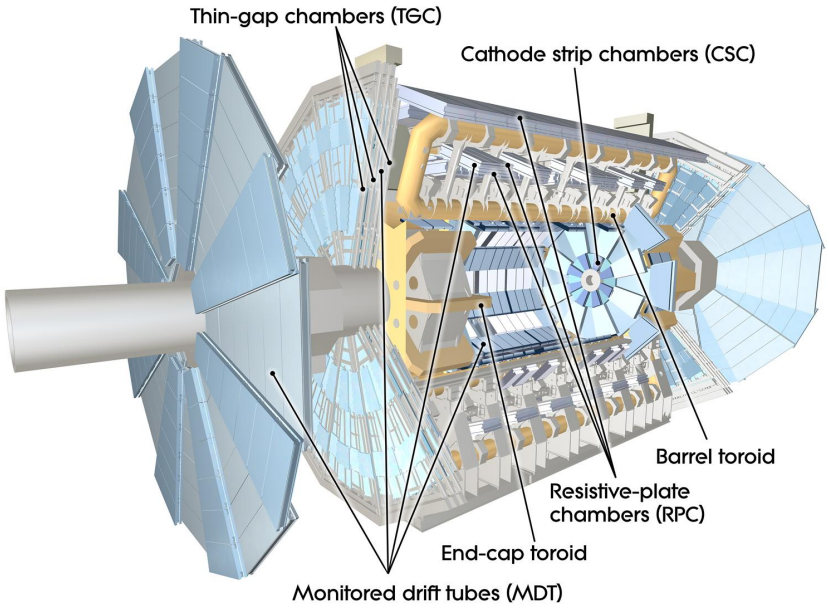


Figure 2.13: Cut-away view of the ATLAS muon system [69].

The ATLAS toroid magnets (see Section 2.2.1) provide a magnetic field for the muon momentum measurements, that is mostly orthogonal to the muon trajectories.

In the $|\eta| < 1.4$ range, magnetic bending of the muons is performed by the large barrel toroid, whereas the toroid end-caps bend the muon trajectories in the region with $1.6 < |\eta| < 2.7$. In the transition region with $1.4 < |\eta| < 1.6$, the magnetic deflection is provided by a combination of the barrel and end-cap toroid fields. For track reconstruction, the field is mapped using computer models of the field which are normalized to measurements from 1850 Hall sensors mounted on spectrometer chambers.

Alignment measurements of the spectrometer are also critical for momentum determination and are accomplished with an optical alignment system of 12,000 sensors. Measurements from these sensors allow a 3 dimensional reconstruction of chamber positions accurate to better than 50 μm. In addition, the optical alignment system is complemented by alignment done with tracks.

The spectrometer is designed so that muons cross three layers of MDT chambers for the sagitta measurement. The track coordinate in the bending plane of the spectrometer is measured by the precision chambers with a resolution of 60-70 μm. In comparison, the sagitta of a 1 TeV muon will be about 500 μm. The trigger chambers are placed on opposite sides of the middle MDT layer and provide a trigger based on muon momentum in addition to identifying the bunch crossing time of the muon. They also provide the second coordinate measurement (non-bending plane) accurate to 5-10 cm.

2.2.5 Forward Detectors

Three smaller detector systems cover the ATLAS forward region. The main function of the first two systems is to determine the luminosity delivered to ATLAS. At ± 17 m from the interaction point lies LUCID (LUminosity measurement using Cerenkov Integrating Detector). It detects inelastic p-p scattering in the forward direction, and

is the main online relative-luminosity monitor for ATLAS. The second detector is ALFA (Absolute Luminosity For ATLAS). Located at ±240 m, it consists of scintillating fibre trackers located inside Roman pots which are designed to approach as close as 1 mm to the beam. The third system is the Zero-Degree Calorimeter (ZDC), which plays a key role in determining the centrality of heavy-ion collisions and is located at ±140 m from the interaction point. The ZDC modules consist of layers of alternating quartz rods and tungsten plates which will measure neutral particles at pseudorapidities $|\eta| \geq 8.2$.

2.3 The ATLAS Trigger, Data Acquisition and Detector Control Systems

The Trigger and Data Acquisition (collectively TDAQ) systems [73], the timing- and trigger-control logic, and the Detector Control System (DCS) are partitioned into sub-systems, typically associated with sub-detectors, which have the same logical components and building blocks. The trigger system has three distinct levels. Each of them refines the decisions made at the previous level and, where necessary, applies additional selection criteria. The data acquisition system receives and buffers the event data from the detector-specific readout electronics over 1600 point-to-point readout links. The ATLAS TDAQ system is illustrated in Figure 2.14.

2.3.1 Trigger System

The beams of the LHC consist of trains of particle bunches. The minimum time interval between passage of successive bunches within a train is 25 ns. Thus collisions can take place every 25 ns within a time interval determined by the lengths of the bunches. At an instantaneous luminosity of 10^{34} cm^{-2} s^{-1} and bunch spacing of 25 ns the aver-

age number of interactions is about 35 per bunch-crossing, corresponding to about 1.4 $\times 10^9$ interactions per second.[3] Selective triggering is therefore required. Association of a unique bunch-crossing with each event is necessary to avoid background from collisions corresponding to other bunch-crossings.

Furthermore, to avoid excessive dead time, the trigger should be able to analyze event data at a rate of 40 MHz. ATLAS employs three levels of trigger to meet these requirements. The first level (L1) is built from custom hardware and can analyze event data at the required rate of 40 MHz. This is achieved by making use of analog sums of calorimeter signals formed on the detector and of signals of dedicated muon trigger chambers (RPCs and TGCs). Consequently event selection is only possible on the basis of energy depositions in the calorimeters and of muon track segments. The L1 trigger is located in the USA15 underground area, as close to the detector cavern as possible, to minimize the lengths of the cables used for forwarding the analog sums to the trigger and to minimize the time needed for sending the trigger accepts to the on-detector readout electronics.

By choosing appropriate thresholds the L1 trigger has been operated during Run 1 with a maximum accept rate of 60-65 kHz, somewhat lower than the maximum design rate of 100 kHz, to prevent excessive dead time. The readout of the detector was upgraded during the long shutdown of 2013 and 2014 to allow for 100 kHz accept rate during Run 2. The L1 trigger can handle an input rate equal to the maximum bunch-crossing rate of 40 MHz. Its maximum latency is about 2.5 μs, i.e. smaller than the maximum of about 3 μs imposed by the depth of the on-detector buffer memories.

This latency includes the transit times of signals between detectors and trigger system and the time required for sending the trigger accepts to the on-detector readout

[3]Except for a few test runs the bunch spacing was 50 ns for Run 1, which resulted in an average of 23 interactions per bunch-crossing.

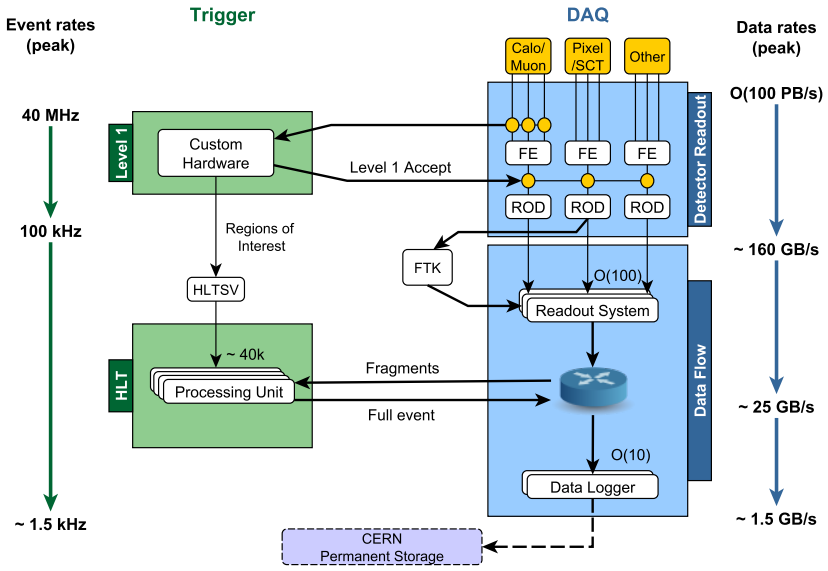


Figure 2.14: Schematic diagram of the TDAQ system of ATLAS [74].

electronics. Data corresponding to events accepted by L1 are further analyzed by software running in computer farms to provide two further levels of triggering. The second level (L2) makes use of a fraction of the full precision detector data and reduces the rate further. The original design aimed for 3.5 kHz and a latency of 40 ms, although a maximum rate of about 5-6 kHz was allowed during Run 1. The design value of the output rate of the last trigger level, the Event Filter (EF), is about 200-300 Hz with an average processing time of about 4 s and with an event size of approximately 1.3 MB, and during Run 1 the maximum output rate was about twice as high. The two levels of the software trigger are collectively known as the High Level Trigger (HLT). During Run 2 the maximum output rate of the HLT is about 1 kHz.

2.3.2 Readout architecture and Data Acquisition

The Readout Drivers (RODs) are detector-specific functional elements of the back-end systems, which achieve a higher level of data concentration and multiplexing by gathering information from several front-end data streams. After an event is accepted by the L1 trigger, the data from the pipelines are transferred off the detector to the RODs. Digitized signals are formatted as raw data prior to being transferred to the data acquisition system (DAQ).

The function of the DAQ system is to efficiently buffer, transport, and record the events that were selected by the trigger system. The first stage of the DAQ, the Readout System (ROS), receives and temporarily stores the data in local buffers. Since Run 2, the HLT farm receives the ROIs[4], requests data from the ROS, and if it passes the HLT trigger, then it requests the rest of the data and runs the rest of the triggers in the same node. Events selected by the HLT are moved to permanent storage at the CERN computer center. In addition to the movement of data, the data acquisition also provides for the configuration, control and monitoring of the hardware and software components which together provide the data-taking functionality.

2.3.3 Detector Control System

The Detector Control System (DCS) permits the coherent and safe operation of the ATLAS detector hardware, and serves as a homogeneous interface to all sub-detectors and to the technical infrastructure of the experiment. It controls, continuously monitors and archives the operational parameters, signals any abnormal behaviour to the operator, and allows automatic or manual corrective actions to be taken. Typical examples are high- and low-voltage systems for detector and electronics, gas and cooling systems, magnetic field, temperatures, and humidity. The DCS also enables

[4]The geographical coordinates in η and ϕ, of those regions within the detector where its selection process has identified interesting features.

bi-directional communication with the DAQ system in order to synchronize the state of the detector with data-taking and handles the communication between the sub-detectors and other systems which are controlled independently, such as the LHC accelerator, the CERN technical services, the ATLAS magnets, and the detector safety system.

2.4 Computing and data management

The field of computing is a crucial part of the CERN past and present history. The communication needs between scientists led to the invention of the World Wide Web (www) by Tim Berners-Lee in 1989 [75], one of the biggest technological milestones in history. The current need for storage and management of a huge amount of data, the Big Data problem, led to the development of the Worldwide LHC Computing Grid (WLCG), usually referred simply as GRID. The GRID is a huge network of interconnected computer centers around the world, which allows the storage and access of the data recorded by the LHC experiments, as well as providing distributed computing services for the analysis of these data.

The WLCG is composed by three levels, or Tiers, numbered 0, 1, and 2 respectively, whose hierarchy is shown in Figure 2.15. Each Tier is made up of several computer centers and provides a specific set of services. Tier-0 is the CERN Data Center, the heart of the network. All of the data from the LHC passes through this central hub, but it provides less than 20% of the GRID's total computing capacity. It also distributes the raw data and the reconstructed output to Tier-1, and reprocesses data when the LHC is not running. Tier-1 consists of 13 computer centers large enough to store LHC data. They provide round-the-clock support for the GRID and are responsible for storing a proportional share of raw and reconstructed data, as well as performing large-scale reprocessing and storing the corresponding output. They

also distribute data to Tier-2 and store a share of the simulated data that the Tier-2 produces. Tier-2 are typically universities and other scientific institutes (around 155 sites around the world) that can store sufficient data and provide adequate computing power for specific analysis tasks. They handle a proportional share of the production and reconstruction of simulated events.

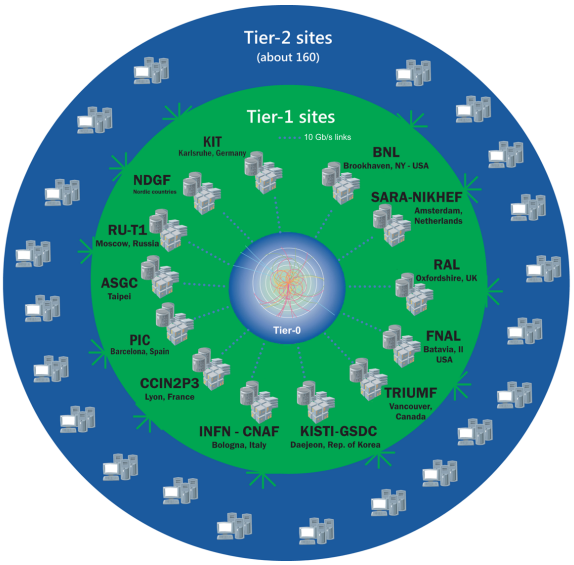


Figure 2.15: Diagram of the Tier hierarchy of the GRID, showing the Tier-0 (CERN Computer Center) in the inner layer, the 13 Tier-1 sites in the green layer and the Tier-2 sites in the blue outer layer [76].

Within ATLAS, there are a few main data formats in use, which facilitate easier management. The RAW and RDO formats are essentially restricted to Tier-0 and Tier-1 sites, with very little additional access available due to the very large computing resources and very complicated processing interface required to reconstruct events from raw events. The next level is the Event Summary Data (ESD), which contains all of the reconstructed data, and thus is a very large data format. These are typically not stored, but in some very specialized cases they are required, and thus the functionality exists for users to access and work with ESDs.

An additional processing step by which the majority of the individual cell information is discarded and other space saving steps are taken forms the Analysis Object Data (AOD). This is the intended format for primary use within ATLAS, where all of the necessary information for reconstruction of interesting objects remains.

Users who work with ESDs or AODs typically also apply some additional analysis selection to reduce the size, thus producing Derived Event Summary Datas (DESDs) or Derived Analysis Object Datas (DAODs). However, the most common data format in use in Run 1 applied a third level of filtering, earning it the name of the Derived^3 Physics Data (D3PD), in addition to flattening the data format. This means that rather than storing the data in objects, every variable was stored as a simple type, such as an integer, vector (list) of integers, or so on. This format and its Run 2 counterpart, which receives the name of Mini-AODs (MAODs), are typically created on a per-analysis or per-working group level, and apply specific selection cuts, triggers, and so on to further reduce the size of the files. Many groups find this necessary in order to reduce datasets to a manageable level. Even with this amount of reduction, a typical analysis is performed on many TeraBytes (TBs) of D3PDs or MAODs.

Chapter 3

Event reconstruction

ATLAS identifies electrons, photons, muons, jets and taus by converting raw detector data to fundamental physics objects using a dedicated set of algorithms, collectively referred to as event reconstruction. The basic building blocks for constructing these objects are the particle trajectories (tracks) reconstructed from hits in the inner detector and muon spectrometer (in the case of muons), providing an estimate of a particle's momentum, and topological clusters constructed from energy deposits in the calorimeters. A second clustering algorithm for determining local energy deposits in the EM calorimeter, a sliding window algorithm, is used in the reconstruction of electrons and photons.

In this chapter, a brief description will be given of the reconstruction algorithms and techniques used to identify the objects used in the Tile Calorimeter performance studies shown in Chapter 4, in the $HH \to \gamma\gamma b\bar{b}$ search, detailed in Chapter 5, and in the measurements of Higgs boson properties in the $H \to \gamma\gamma$ channel, described in Chapter 6.

3.1 Tracks and Primary Vertices

3.1.1 Track reconstruction

Due to the presence of the axial magnetic field along the z direction, charged particles originating from the proton-proton collisions follow helicoidal trajectories that can be parameterized by the following five parameters:

- ϕ_0, the azimuthal angle of the track.
- θ, the polar angle of the track.
- d_0, the transverse impact parameter, i.e. the distance from the point of closest approach to the reference point in the transverse plane.
- z_0, the longitudinal impact parameter, i.e. the z coordinate of the track at the point of closest approach mentioned in the previous bullet point. Strictly speaking, the longitudinal impact parameter is $|z_0| \sin\theta$.
- the ratio q/p, which defines the orientation and the curvature of the helix trajectory, where q denotes the electric charge and p the momentum.

The charged particles trajectories are identified in the inner detector using a set of local and global pattern recognition algorithms, referred to as New Tracking [77], aiming to reconstruct tracks originating from the proton-proton hard scatter (primary tracks), from the decays of long-lived particles (secondary tracks) and from the interaction of particles with the material (conversion tracks).

The inner detector track reconstruction consists of several sequences with different strategies as described in [77, 78]. Primary tracks are required to have a $p_{\mathrm{T}} > 400$ MeV and are reconstructed in the region $|\eta| < 2.5$ using the "inside-out" track finding sequence, which consists of the following components that are executed in the following order:

- Data preparation and space point formation. The initial step of the inner detector reconstruction consists of the creation of clusters and drift circles and the transformation of clusters in the silicon detectors into 3D space points. Clusters are formed by finding connected cells in the pixel and strip detectors.

- Space point seeded track finding. Track finding starts with the formation of space point triplets (seeds). Seeds can be built from space points in the pixel detector only, the strip detector only or any mixed setup. To reduce the number of potential seeds, initial cuts are applied and dedicated care is taken not to extensively use space points in multiple seeds. Seeds that pass the initial requirements are then input to a track finding algorithm that uses a Kalman filter technique [79] and aims to complete the track candidates within the silicon detector.

- Ambiguity solving. Track candidates are then further processed in an ambiguity solving module that aims to eliminate track candidates from random hit combinations (often referred to as fakes) or track duplicates, which can be identified by measurements that are shared with other track candidates. The ambiguity solving relies on a scoring function applying positive scores for unique measurements and good fit quality, while penalizing missing measurements where they would be expected (also called holes) or shared measurements with other track candidates.

- TRT extension. Tracks that successfully pass the ambiguity solving stage and are within the coverage of the TRT detector are then extended into the TRT and completed for measurements in the outermost tracking detector.

In the complementary "outside-in" sequence, the track search starts from segments reconstructed in the TRT, and extends them inwards by adding silicon hits,

in a process referred to as back-tracking. Back-tracking is designed to reconstruct secondary and conversion tracks. Tracks with a TRT segment, but no extension into the silicon detector, are referred to as TRT-standalone tracks.

The increasing detector occupancy with pile-up results in a degradation of the track reconstruction, as it increases the combinatorial fake tracks, i.e. reconstructed tracks which cannot be matched to either a primary vertex or a secondary particle. Robust requirements on the hits (measurement points assigned to a track), holes (non-existing but expected measurement points), and outliers (hits that reduce the quality of the track fit) of the silicon detectors are used to minimize this impact, and improve the quality of the reconstructed track. Moreover, cuts on the d_0 and z_0 impact parameters (or their errors, denoted by σ) are also generally used in analyses, to ensure the track originates from the primary hard-scattering vertex. In particular, the addition of the IBL in Run 2 leads to a significant improvement in the impact parameters resolution.

3.1.2 Primary Vertex reconstruction

Charged particles tracks passing the robust requirements are combined in order to reconstruct the spatial position of p-p interaction vertices, using an iterative procedure, which is divided in two stages: vertex finding and vertex fitting.

First, a vertex seed is found by looking for a global maximum in the distribution of z coordinates of the pre-selected tracks, here computed with respect to the center of the beamspot. Then, the vertex position is determined using the adaptive vertex fitting algorithm [80], constrained by the beamspot position. It is a robust χ^2-based fitting, that takes the input seed and the tracks around it. Each track carries a weight, which is a measure of its compatibility with the fitted vertex. Tracks incompatible with the vertex by more than 7σ are used to seed a new vertex, and the procedure is

repeated until no tracks are left.

The resolution on the vertex position improves as a function of the number of tracks associated to it, N_{trk}, and with the square sum of their transverse momenta $\sum_i^{N_{trk}} p_{\mathrm{T},i}^2$. The primary vertex is the one with highest $\sum_i^{N_{trk}} p_{\mathrm{T},i}^2$ while the other reconstructed vertices are labelled as pile-up vertices.

3.2 Photons

3.2.1 Photon reconstruction

A detailed explanation of the photon reconstruction process can be found in [81]. An energetic photon's interaction with the EM calorimeter is at the origin of the electromagnetic shower, which deposits a significant amount of energy in a small number of neighbouring calorimeter cells. The reconstruction of photons proceeds in the following way:

- Seed cluster building. The reconstruction of photon candidates in the region $|\eta| < 2.5$ starts by building seed clusters out of the energy deposits in the EM calorimeter. Cells in the EM calorimeter from all three layers are grouped into $\Delta\eta \times \Delta\phi$ towers of 0.025×0.025 by a sliding-window algorithm [82] with windows of size 3×5 in $\eta - \phi$ space and transverse momentum above 2.5 GeV, followed by duplicate removal. Based on Monte Carlo simulations, the cluster building efficiency is estimated to be greater than 99% for photons with $E_{\mathrm{T}} > 20$ GeV.

- Track reconstruction. Once seed clusters are reconstructed, a search is performed for inner detector tracks that are loosely matched to the clusters, in order to identify and reconstruct photon conversions. These initial clusters are

used to create Regions of Interest (ROIs).[1] Within these ROIs, standard track pattern reconstruction [77] is first performed everywhere. If the pattern recognition fails for a silicon track seed that is within a ROI, a modified pattern reconstruction algorithm is performed based again on a Kalman filter [79], allowing for up to 30% energy loss at each material intersection. Track candidates are then fitted with the ATLAS global χ^2 fitter [83], allowing for additional energy loss in cases that the standard track fit fails. Tracks with silicon hits loosely matched to EM clusters are refitted using a Gaussian Sum Filter (GSF) fitter [84], a non-linear generalization of the Kalman filter, for improved track parameter estimation.

- Track conversion. Tracks consistent with originating from a photon conversion are used to create conversion vertex candidates.

- Track-cluster matching. Conversion vertex candidates are matched to seed clusters.

- Final cluster creation. A final algorithm decides whether a seed cluster corresponds to an unconverted photon, a converted photon or a single electron based on the matching to conversion vertices or tracks and on the cluster and track(s) four-momenta.

Photon conversion

Conversion-finding is run on the loosely-matched tracks by using both tracks with silicon hits and tracks reconstructed only in the TRT. Converted photons are classified as single-track or double-track. Double-track conversion vertex candidates are reconstructed from pairs of oppositely charged tracks in the inner detector that are likely

[1] For each seed EM cluster passing loose shower shape requirements of $R_\eta > 0.65$ and $R_{had} < 0.1$ a region of interest with a cone-size of $\Delta R = 0.3$ around the seed cluster barycenter is defined.

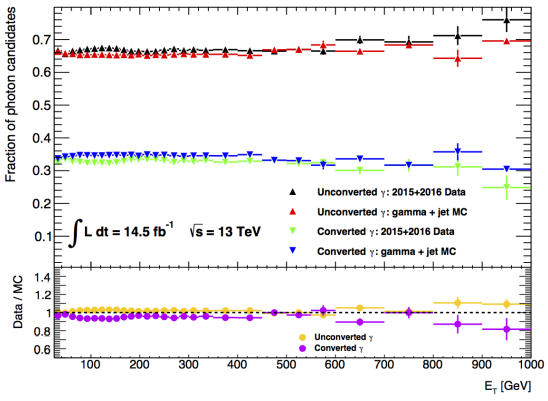
to be electrons. Single-track conversion vertex candidates are essentially tracks that do not have hits in the innermost sensitive layers. To increase the converted photon purity, the tracks used to build conversions must generally have a high probability to be electron tracks as determined by the TRT, especially when building single-track conversions or if using standalone-TRT tracks.

The matching between tracks from the conversion vertices and clusters relies on an extrapolation of the conversion candidates to the second sampling layer of the calorimeter, and the comparison of the extrapolated η and ϕ coordinates to the η and ϕ coordinates of the cluster center (ΔR). The details of the extrapolation depend on the type of the conversion vertex candidate. Quality cuts are applied on the conversion vertices for better pile-up tolerance. Because dead pixels can affect conversion building, the dead pixel map is also used to determine the conversion quality.

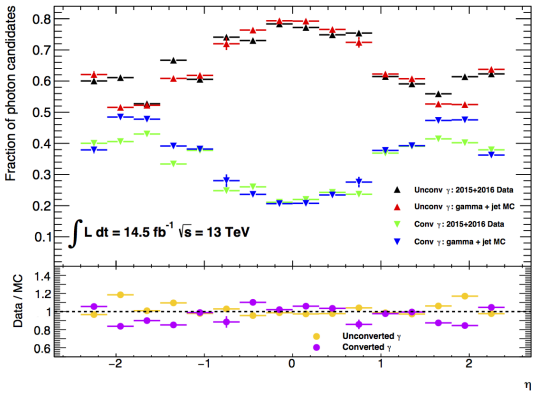
If there are multiple conversion vertices matched to a cluster, double-track conversions with two silicon tracks are preferred over other double-track conversions, followed by single-track conversions. Within each category, the vertex with the smallest conversion radius is preferred. Conversion vertex matching is performed twice, first to the potential seed clusters, and then to the complete photon superclusters. An arbitration relying on the properties of the tracks and conversion vertices matched to a given EM cluster is performed to determine whether an object is reconstructed as an electron, a photon, or both. The fraction of converted/unconverted photons can be seen in Figure 3.1.

3.2.2 Photon identification

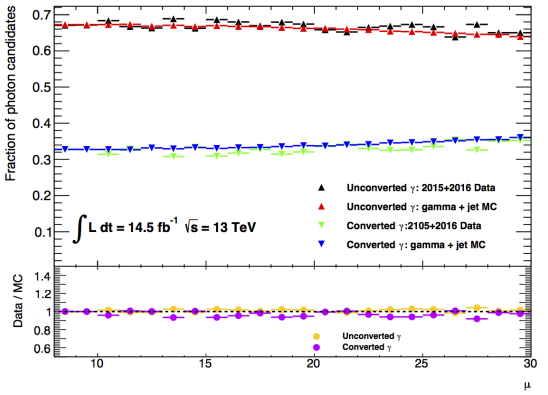
Photon candidates are required to satisfy a set of identification (ID) criteria to reduce contamination from background, primarily associated with neutral pions in jets decaying into photon pairs, based on the lateral and longitudinal shape of the elec-



(a)



(b)



(c)

Figure 3.1: Fraction of converted and unconverted photons as a function of a) transverse energy E_{T}, b) detector pseudorapidity η and c) number of interactions per crossing μ, with early Run 2 data. Comparison with MC simulated γ+jet events is also shown.

tromagnetic shower in the calorimeter [81, 85]. Photon candidates are required to deposit only a small fraction of their energy in the hadronic calorimeter, and to have a lateral shower shape consistent with that expected from a single electromagnetic shower. There are two levels of ID:

- The loose ID exploits the discriminating variables (DV), such as the ratio of E_{T} in the hadronic calorimeter to E_{T} of the EM cluster or the lateral width of the shower, only in the hadronic and in the EM calorimeter second sampling layer, providing a highly efficient selection with fair background rejection, typically used for the trigger and background studies.

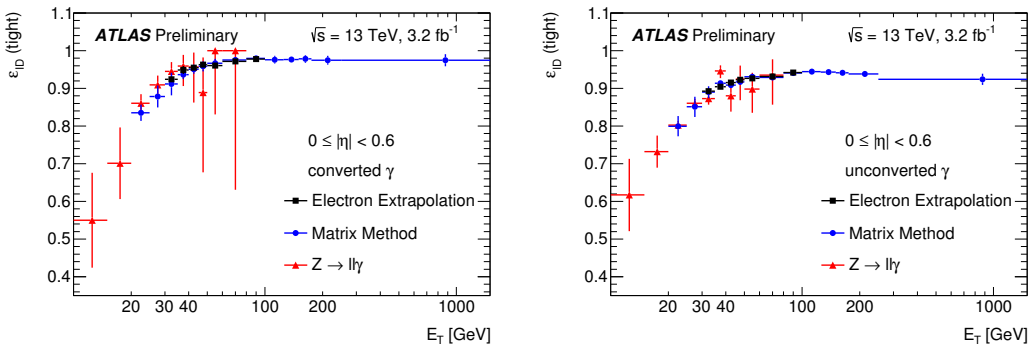
- The tight ID level exploits the full granularity of the EM calorimeter and applies tighter requirements also on the DVs used by the loose ID. The efficiency of the tight identification ranges from 67% (60%) to 90% (95%) for unconverted (converted) isolated photons from p_{T} of 15 GeV to 50 GeV and larger. The average jet background rejection factor is equal to 5000 for this selection.

Figure 3.2 shows several methods that are used to measure with data the efficiency of the photon identification requirements, to cover a broad energy spectrum:

- At low energy, radiative photons from $Z \to \ell\ell\gamma$ decays are selected by placing kinematic requirements on the dilepton pair, on the invariant mass of the three particles in the final state and on the quality of the two leptons.

- In the medium energy range, similarities between electrons and photon showers are exploited using $Z \to ee$ decays and photon ID efficiencies are obtained using a tag-and-probe method.

- At high energy, inclusive photon final states are used by applying the matrix method, which classifies the photons between prompt and background photon candidates, passing or failing tight ID criteria.



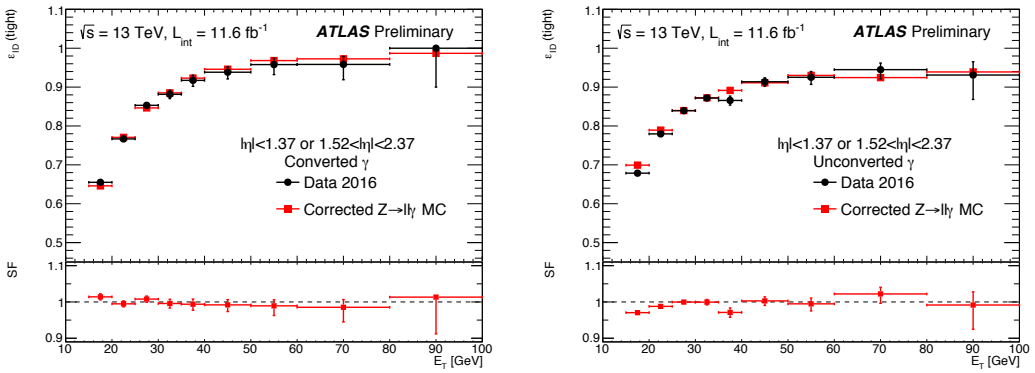
(a) (b)

Figure 3.2: Comparison of the data-driven measurements of the identification efficiency for a) converted photons and b) unconverted photons [86].

Figure 3.3 shows the difference between the simulation and the data-driven measurements, which is taken into account by computing data-to-MC efficiency ratios, also referred to as scale factors (SF). The SF obtained for photons are closer to unity than the ones obtained for electrons (see Section 3.3.2) because of the correction of the shower shape variables in simulation applied on photons to account for the average data-MC difference.

3.2.3 Photon isolation

To reject the hadronic jet background, photon candidates are required to be isolated from any other activity in the calorimeter and the tracking detectors. The calorimeter isolation is computed as the sum of the transverse energies of positive-energy topological clusters [88] in the calorimeter within a cone of $\Delta R = 0.2$ centered around the photon shower barycentre. The transverse energy of the photon candidate is removed and the contributions of the underlying event and pile-up are subtracted according to the method suggested in [89]. The track isolation for a cone size of $\Delta R = 0.2$ is used and for converted photons the tracks associated with the conversion are removed. The



(a) (b)

Figure 3.3: Comparison of the radiative Z boson data-driven efficiency measurements of a) converted and b) unconverted photons to the $Z \to \ell\ell\gamma$ simulation as a function of the E_{T}, in the region $15 \geq E_{\mathrm{T}} \geq 100$ GeV. The lower panel shows the ratio between the data and the simulation [87].

calorimeter (track) isolation is required to be less than 6.5% (5%) of the photon E_{T}. The efficiency of the isolation requirement for photons satisfying the tight identification criteria ranges from approximately 60% for E_{T} of 15 GeV to more than 90% for E_{T} of 40 GeV and larger. The average jet background rejection factor for isolation requirements with respect to the tight identification criteria is approximately 1.5.

3.2.4 Photon calibration

Precise measurements of photon energy are of vital importance in $H \to \gamma\gamma$ analyses. There are a number of effects which impact the energy resolution and bias the measured value of the photon energy, which are summarised below:

- The energy loss due to the amount of material that sits in front of the EM calorimeter.
- The energy loss due to dead material inside the EM calorimeter.
- The energy loss laterally outside the reconstructed cluster.

- The energy loss longitudinally behind the EM calorimeter.

Without a calibration procedure [90], the sum of the energy in the cells will be less than the true energy. A multivariate regression algorithm based on Monte Carlo simulation calibrates the energy of the photons to account for the energy losses described above. After this step, an in-situ measurement using $Z \to ee$ events is performed to determine the energy scale and resolution in both data and simulation. This is then used to correct the energy scale in data and to smear the energy resolution in simulation.

MVA calibration

A multivariate (MVA) technique is used to optimize the energy calibration using a single particle Monte Carlo sample. The regression MVA technique chosen is a a boosted decision tree with gradient boosting that uses the ratio of the true photon energy to the raw energy measured in the calorimeter as its target.

The variables used in the training of the MVA are the total raw cluster energy measured in the calorimeter (E_{acc}), the ratio of the pre-sampler layer energy to the calorimeter energy (E_0/E_{acc}), the ratio between the energy measured in the first two layers of the calorimeter (E_1/E_2), the cluster barycenter η in the ATLAS coordinate system and the cluster barycenter η and ϕ in the EM calorimeter coordinate system.

The algorithm is optimized in different regions of the phase space with the Monte Carlo sample divided into bins in pseudorapidity ($|\eta|$), transverse energy measured in the calorimeter (E_{acc}) and whether the photon is converted or unconverted. The binning in $|\eta|$ is chosen to follow the known detector geometry variations.

In-situ corrections

After applying the MVA based calibration, there may be some disagreement between energy scale and resolution in data. In order to account for this potential discrepancy, a correction is evaluated with an in-situ measurement using $Z \to ee$ events. The energy miscalibration is defined as the difference in response between data and simulation, and is parameterized as:

$$E_i^{\text{data}} = E_i^{\text{MC}}(1 + \alpha_i), \tag{3.1}$$

where E_i^{data} is the electron energy in data and E_i^{MC} is the photon energy in simulation. The α_i represents the deviation from the perfect calibration in a given η bin, which is labeled as i. The difference in energy resolution between data and simulation, which does not depend at first order on the energy, can be modeled by an additional term (c_i') for a given η bin as:

$$\left(\frac{\sigma(E)}{E}\right)_i^{\text{data}} = \left(\frac{\sigma(E)}{E}\right)_i^{\text{MC}} \oplus c_i'. \tag{3.2}$$

The measured values of α_i range between -2% and +2% depending on the η bin. The c_i' values measured range between 0.005 and 0.03 depending on η.

Despite the Z boson decaying to electrons, the corrections are assumed to be valid for photons. This assumption is tested by using photons from radiative Z boson decays in the electron and muon decay channels, where the results are found to be compatible.

3.3 Electrons

3.3.1 Electron reconstruction

Electron reconstruction in the central region of the ATLAS detector ($|\eta| < 2.47$) proceeds in several steps:

- Seed-cluster reconstruction. Electrons are formed by matching tracks reconstructed in the Inner Detector with electromagnetic clusters found using the sliding window algorithm. The efficiency of the cluster reconstruction for a true e/γ ranges from 95% at $E_{\mathrm{T}} = 7$ GeV to more than 99% above $E_{\mathrm{T}} = 15$ GeV.

- Track reconstruction. Track reconstruction proceeds in two steps: pattern recognition and track fit. The pattern recognition uses the pion hypothesis for energy loss in the interactions with the detector material, complemented with a modified pattern recognition algorithm that takes into account energy loss for possible bremsstrahlung. If a track seed with a transverse momentum larger than 1 GeV can not be successfully extended to a full track and it falls within one of the EM RoI, a second attempt is performed using another pattern recognition with the electron hypothesis that allows for larger energy loss. Track candidates are then fitted using the ATLAS global χ^2 track fitter either with the pion hypothesis or the electron hypothesis.

- Electron specific track fit. The obtained tracks are loosely matched to EM clusters using the distance in η and ϕ between the position of the track, after extrapolation, in the calorimeter middle layer and the cluster barycentre. The matching conditions account for energy-loss due to bremsstrahlung and the number of precision hits in the silicon detector. Tracks that have significant number of precision hits (≥ 4) and are loosely associated to electron clusters are refit using an optimized GSF, which takes into account the non-linear bremsstrahlung

effects.

- Electron candidate reconstruction. The matching of the track candidate to the cluster seed completes the electron reconstruction procedure. A similar matching as the one described above is repeated for the refit track with stricter conditions. If several tracks fulfil the matching condition, one track is chosen as primary track. The choice is based on an algorithm using the cluster-track distance R calculated using different momentum hypotheses, the number of pixel hits and the presence of a hit in the first silicon layer [91]. Electron candidates without any associated precision hit tracks are removed and considered to be photons. The electron cluster is then re-formed and its energy is calibrated to the original electron energy using multivariate techniques based on simulated MC samples. Data-driven scale corrections are derived from $Z \to ee$ events by following the same procedure described for photons in Section 3.2.4.

The four-momentum of the electrons is computed using information from both the final calibrated energy cluster and the best track matched to the original seed cluster. The energy is given by the final calibrated cluster, while the ϕ and η directions are taken from the track.

In Run 2 analyses, the electron measurements are performed by requiring the track associated with the electron to be compatible with the primary interaction vertex of the hard collision, in order to reduce the background from conversions and secondary particles. To assess the compatibility with the primary vertex, the Δz_0 between the track and the primary vertex is employed. The efficiency of the electron reconstruction process goes from 97% to 99% for electrons with $p_{\mathrm{T}} > 100$ GeV.

3.3.2 Electron identification

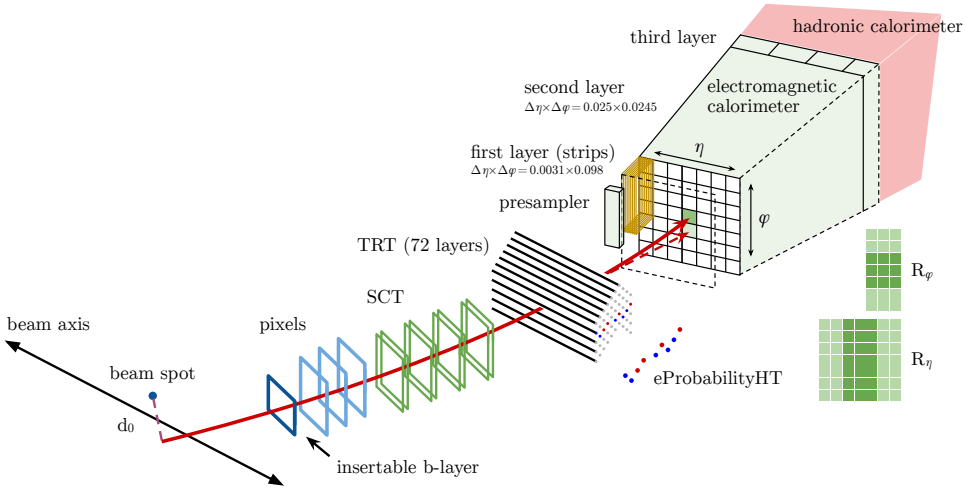
To determine whether the reconstructed electron candidates are signal-like objects or background-like objects such as hadronic jets or converted photons, algorithms for electron identification (ID) are applied. The ID algorithms use quantities related to the electron cluster and track measurements including calorimeter shower shapes, information from the transition radiation tracker, track-cluster matching related quantities, track properties, and variables measuring bremsstrahlung effects for distinguishing signal from background.

Figure 3.4 shows a schematic view of the electron reconstruction and ID. For Run 2, several changes to the input variables used for electron identification have been introduced [92]:

- Taking advantage of the new innermost pixel layer, the IBL, the number of hits in this layer is used for discriminating between electrons and converted photons.
- The change in the TRT gas led to modifications in the detector response and prompted the introduction of a new discriminating variable in the electron ID algorithms. In Run 1, only the fraction of high-threshold hits was used from the TRT as a signature of transition radiation to distinguish electrons from hadrons. In Run 2, a likelihood method based on the TRT high-threshold hits is introduced to compensate for the lower transition radiation absorption probability of the argon.

The measurement of the electron identification efficiencies and isolation cuts are performed with the data using tag-and-probe techniques with a large statistics sample of $Z \to ee$ and $J/\psi \to ee$ decays.

- This method consists on applying strict selection criteria on one of the electron candidates (also known as tag) together with the requirements of the di-electron



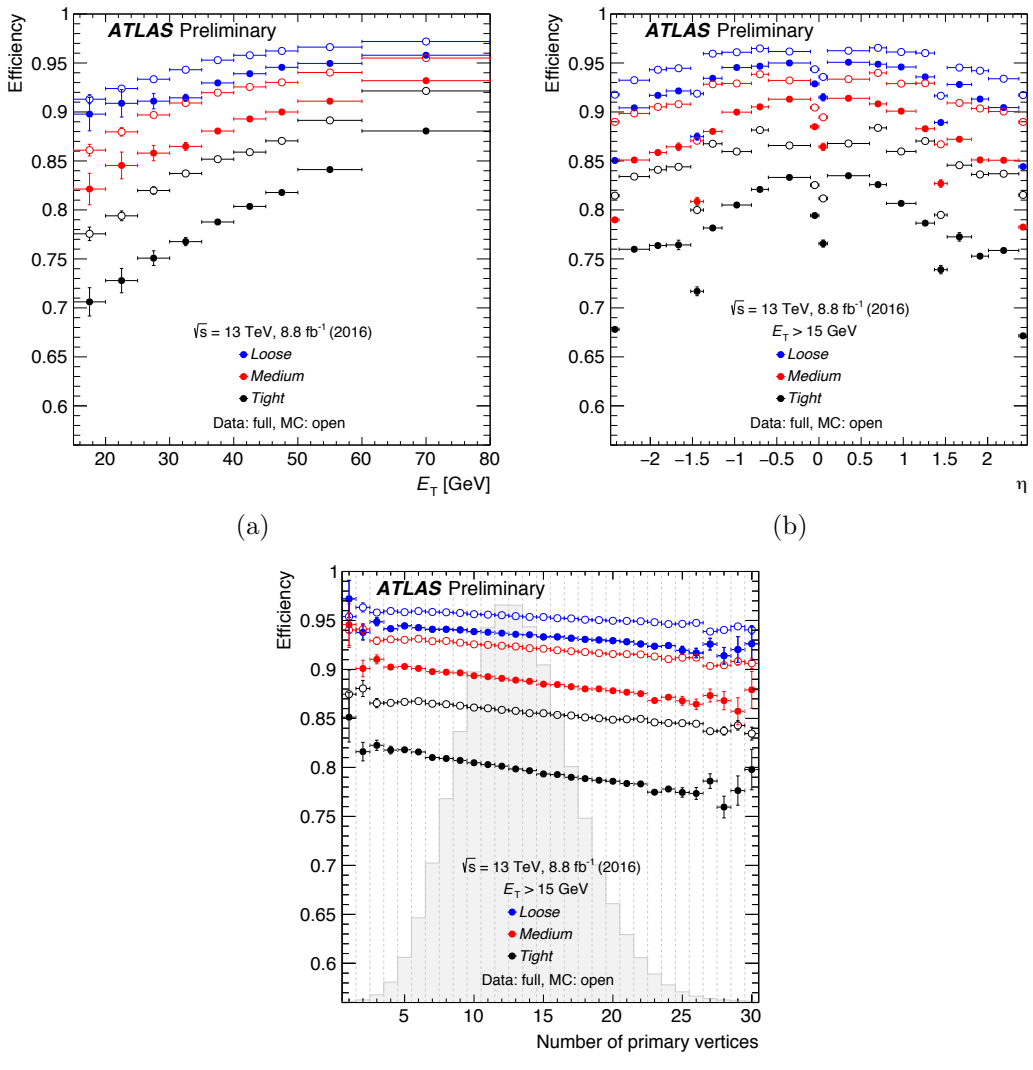
(a)

Figure 3.4: Schematic view of the electron reconstruction and identification [92].

invariant mass, which allows for a loose pre-identification of the other electron candidate (also known as probe).

- The low E_{T} range (from 7 to 20 GeV) is covered by $J/\psi \to ee$ and suffers from a significant background contamination, while $Z \to ee$ events are used for measurements above 15 GeV, as shown in Figure 3.5.

Electrons are identified by different sets of likelihood-based (LH) identification criteria which are chosen to be 95%, 90% and 80% efficient for electrons with $E_{\mathrm{T}} \approx$ 40 GeV, and referred to as loose, medium and tight operating points respectively. The electron ID efficiencies are found to be robust with respect to the number of primary vertices, in the range probed by the available data.



(a) (b)

(c)

Figure 3.5: Electron ID efficiencies in $Z \to ee$ events as a function of a) transverse energy E_{T}, b) pseudorapidity η and c) the number of reconstructed primary vertices. The distribution of the number of reconstructed primary vertices in the selected data events is overlaid in grey [93].

3.4 Muons

The ATLAS experiment uses the information from the muon spectrometer, inner detector and EM calorimeter to identify and precisely reconstruct muons produced in the p-p collisions. Muon identification is performed according to several reconstruction criteria [94, 95], leading to the following muon types:

- Stand-Alone (SA) muons: MS tracks found outside the ID acceptance ($2.5 < |\eta| < 2.7$) and momentum taken from the MS track.
- Combined (CB) muons: a MS track is matched to a reconstructed track in the ID, and the measurements of the momenta are combined. This is the type of reconstructed muons most commonly used by physics analyses.
- Segment-tagged (ST) muons: a partial MS track is matched to an ID track, and the muon momentum is taken from the ID measurement.
- Calorimeter-tagged (CaloTag) muons: in the non-active "crack" of the MS at $|\eta| < 0.1$, muon candidates are reconstructed from a track in the ID that has an EM calorimeter energy deposit compatible with a minimum ionizing particle.

The reconstruction of the muons used in physics analyses has been performed using two independent reconstruction software packages, implementing different strategies named chains, which combine the information from the ID tracks and the MS track segments. The first chain, called Staco or Chain 1, statistically combines the ID and MS track candidates using the corresponding track parameters and covariance matrices. Chain 2 (or Muid), on the other hand, performs a global χ^2 fit of the muon tracks, using both ID and MS measurements.

Chain 3, an improved algorithm that combines the best features of the other two chains, has been used to reconstruct muons in Run 2. The improvements included a

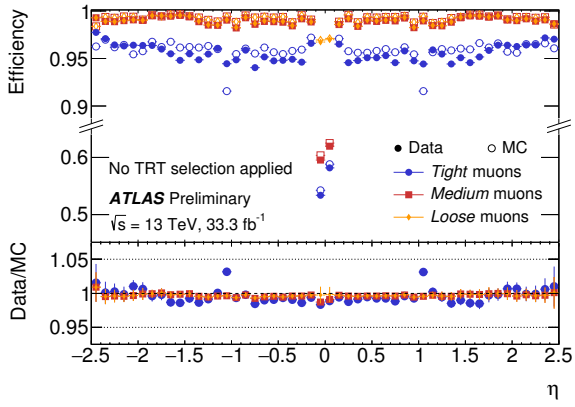
better calculation of the energy loss in the calorimeter, the use of a Hough transformation [96] to increase track-finding efficiency and a higher background rejection.

Muon identification is performed by applying quality requirements that suppress background, mainly from pion and kaon decays, while selecting prompt muons with high efficiency and/or guaranteeing a robust momentum measurement. Several variables offering good discrimination between prompt muons and background muon candidates are studied in simulated $t\bar{t}$ events. Muons from W decays are categorized as signal muons while muon candidates from light-hadron decays are categorized as background. A number of quality requirements can be imposed on the reconstructed muons, including minimum requirements on the number of hits in each of the ID subdetectors and the MS, where applicable. Four muon identification selections (Loose, Medium, Tight, and High-p_{T}) are provided to address the specific needs of different physics analyses. Loose, Medium, and Tight are inclusive categories in that muons identified with tighter requirements are also included in the looser categories.

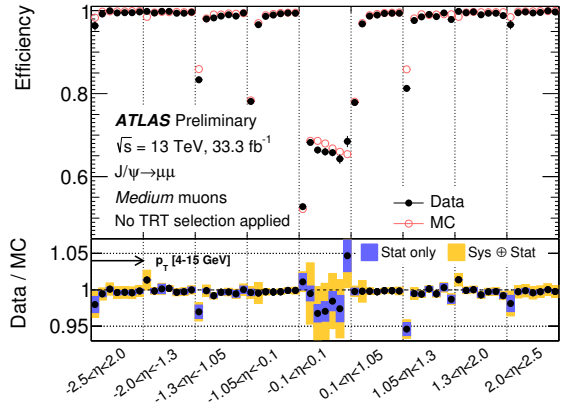
The availability of two independent detectors to reconstruct the muons (the ID and the MS) enables a precise determination of the muon reconstruction efficiency in the region $|\eta| < 2.5$. This is obtained with the so called tag-and-probe method using $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ events that cover a broad muon p_{T} spectrum. In this method one leg of the decay (tag) is required to be identified as a Medium muon that fires the trigger and the second leg (probe) is required to be reconstructed by a system independent of the one being studied. The differences in the measured reconstruction efficiency in Data and MC are taken into account by multiplicative Scale Factors that differ from unity within a few percent.

Figure 3.6a shows the muon reconstruction efficiency as a function of η as measured from $Z \to \mu\mu$ events. The efficiency at low p_{T} is reported in Figure 3.6b as measured from $J/\psi \to \mu\mu$ events as a function of p_{T} in different η regions. The efficiencies of

the Loose and Medium selections are very similar throughout the detector with the exception of the region $|\eta| < 0.1$, where the Loose selection fills the MS acceptance gap using the calorimeter and segment-tagged muons contributions.



(a)



(b)

Figure 3.6: Muon reconstruction efficiency a) as a function of η measured in $Z \to \mu\mu$ events for muons with $p_{\mathrm{T}} > 10$ GeV for Loose, Medium and Tight muon selections, and b) in different η regions measured in $J/\psi \to \mu\mu$ events for Medium muon selection. The error bars on the efficiencies indicate the statistical uncertainty. The panel at the bottom shows the ratio of the measured to predicted efficiencies, with statistical and systematic uncertainties [97].

3.5 Jets

Jets are collimated sprays of hadrons. They are abundantly produced in *p*-*p* collisions at the LHC. The jet calibration procedure should correctly determine the jet energy scale and additionally the best possible energy and angular resolution should be achieved. Good jet reconstruction and calibration facilitates the identification of known resonances that decay to hadronic jets, as well as the search for new particles.

The main jet identification algorithm used by the ATLAS collaboration is the anti-k_{T} algorithm [98] with a distance parameter $R = 0.4$. Various objects can be used as inputs to this algorithm: calorimeter energy deposits, inner detector tracks [99] or a combination of both [100]. Jets reconstructed from tracks, also referred to as the track jets, have low dependence on the pile-up activity since only tracks originating from the primary vertex are used for the jet finding. However, the reconstruction of track jets is limited by the ATLAS tracker acceptance to $|\eta| < 2.5$. Therefore, the majority of ATLAS analyses uses a jet reconstruction based on calorimeter deposits, so called calorimeter jets.

The inputs for calorimeter jets reconstruction are the topologically clustered calorimeter cells, so-called topoclusters [88]. The topological clustering algorithm groups cells with the significant energy deposits, aiming for effective noise suppression. The total noise in calorimeter cells, σ_{Noise}, is calculated as the quadratic sum of the measured electronics and so-called pile-up noise.[2] The clustering algorithm starts from the seed cells with the energy deposits above $4\sigma_{Noise}$. All neighbouring cells are iteratively added to the topocluster if their energy is above $2\sigma_{Noise}$. This is followed by the addition of all adjacent cells. As the final step, the cluster splitting algorithm separates produced topoclusters based on local energy maxima to avoid

[2]Pile-up noise: average energy deposited by concurrent *p*-*p* interactions. Its value increases with the number of interactions per bunch crossing.

overlap. Topoclusters are considered to be massless and only those with positive energy are used for jet reconstruction.

The energy of the calorimeter cells is measured at EM scale, established using electrons at test beams. The local cell weighting (LCW) calibration [101] can be applied to topoclusters classified as hadronic to account for the difference in the detector response to electromagnetic and hadronic particles. The LCW calibration is derived using MC simulation of single pion events.

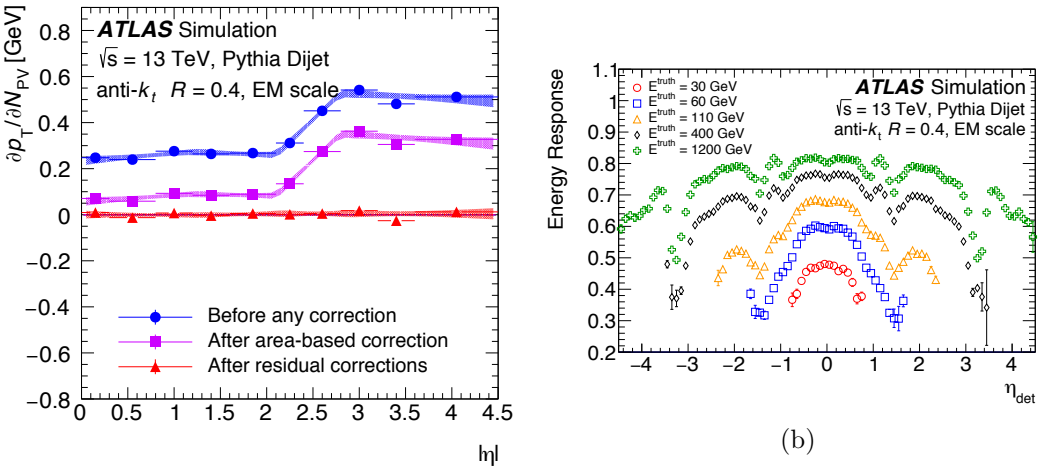
The jet energy scale (JES) calibration [102] restores the energy scale of reconstructed jets to that of simulated truth jets. Different sets of correction factors are developed for jets reconstructed using the EM and LCW topo-clusters. The JES calibration includes origin correction, pile-up correction, absolute correction of the detector response based on MC simulation, global sequential correction and residual in-situ calibration.

- The origin correction forces the four-momentum of the jet to point to the hard-scatter primary vertex rather than to the center of the detector while keeping the jet energy unchanged.

- A two step procedure is used to subtract the pile-up contribution. First, the effect of pile-up is removed by exploiting the average energy density and the area of the jet. Second, a residual correction removes the remaining dependence of the jet on the number of reconstructed primary vertices (N_{PV}) and the expected average number of interactions per bunch crossing ($\langle\mu\rangle$). The performance of the pile-up correction is shown in Figure 3.7a.

- The jet energy scale and η calibration correct the reconstructed jet to the particle-level energy scale to account for the difference of the calorimeter energy response using MC simulation. In addition, it corrects any bias of the

reconstructed jet η caused by the transition between different calorimeter regions and the difference in calorimeter granularity. The energy response as a function of detector η for jets of different simulated truth energy is shown in Figure 3.7b.



(a) (b)

Figure 3.7: a) Dependence of the jet p_{T} with the number of primary vertices, for simulated events, before pile-up corrections (blue), after the area-based correction (violet), and after the residual correction (red). b) The average jet energy response as a function of the detector pseudorapidity for different truth energies [103].

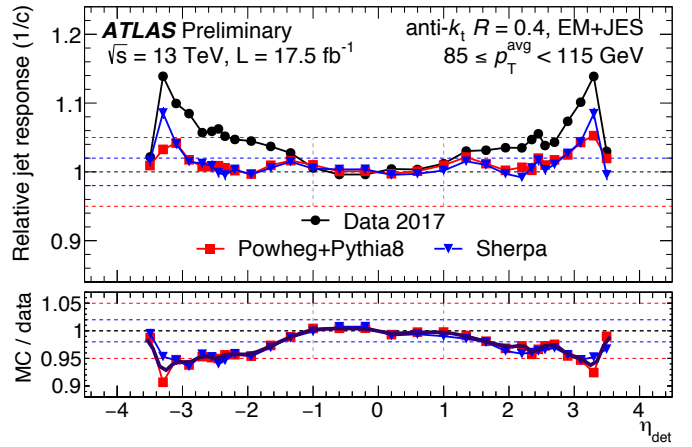
- The global sequential correction (GSC) [104] is designed to reduce the jet response dependence on the flavor of the initiated-jet parton. The correction uses global properties of the jets such as the portion of the jet energy measured in the first layer of the hadronic calorimeter, the portion of the jet energy measured in the third layer of the electromagnetic calorimeter, the average p_{T}-weighted transverse distance in the $\eta - \phi$ space between the jet axis and all tracks associated to the jets, the number of tracks associated to the jet and the number of muon track segments associated to the jet (accounting for punch-through

correction). The correction removes the jet response dependence on the listed observables.

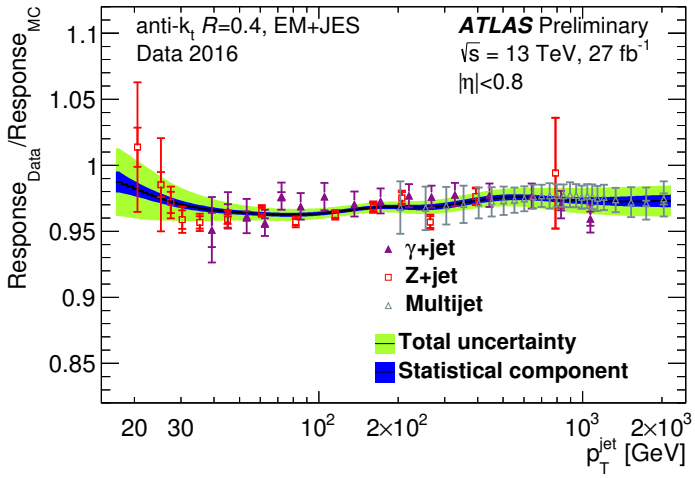
- The in-situ JES calibration is applied to jets measured in data. The correction is calculated as the jet response difference between data and MC simulation using the transverse momentum balance of a jet and a well-measured reference object. Some of the in-situ methods are, for example: relative η-intercalibration, where the jet response in a particular region is corrected relative to the jet response in the central region; direct transverse momentum balance between a photon or a Z boson and a jet, which allows for a correction of the jet response in the central region; and balance between a high p_{T} jet and a recoil system of low p_{T} jets (Multijet Balance), which have been well calibrated by the previous techniques. Figure 3.8a shows the relative jet response as a function of detector η in a single jet p_{T} bin. The lower panel depicts the η-intercalibration correction. The combined in-situ corrections are shown in Figure 3.8b as a function of the jet p_{T} measured using Z+jet, γ+jet and multijet events.

3.5.1 Jet energy scale uncertainty

The steeply falling jet p_{T} spectrum causes jets to easily shift in and out of analyses selections. For this reason, the uncertainty on the JES is frequently the dominant systematic uncertainty in many physics analyses, and arises from a variety of sources. Figure 3.9 shows the total uncertainty on the JES as a function of p_{T} and η of the jet. The total uncertainty at low jet p_{T} reaching up to 5% is driven by the in-situ methods uncertainties, and can be assessed by varying the event selection criteria, as well as from the selection, calibration and modeling of the objects used as reference. Additional JES uncertainties due to specific event topologies, such as selection of samples with different flavor compositions, are also accounted for. The



(a)



(b)

Figure 3.8: a) Relative jet response as a function of detector pseudorapidity in a single jet p_{T} bin. The lower panel depicts the η-intercalibration correction, and b) Data-to-MC ratio of the jet response, as a function of the transverse momentum of the jet, for three in situ techniques combined, to determine the in-situ energy scale correction [105, 106].

in-situ methods uncertainties of about 1% solely dominate at high jet transverse momentum ($p_{\mathrm{T}} > 200$ GeV). The uncertainties originating from the pile-up correction are evaluated in-situ, using track-jets and the p_{T} balance between the jet and a Z boson.

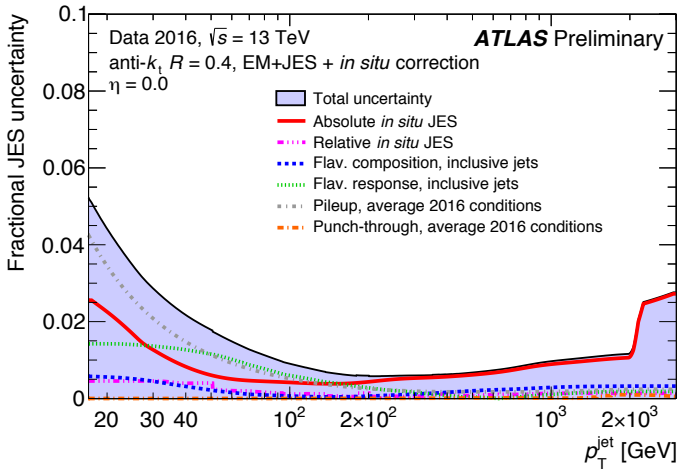
Jets with transverse momenta above 2 TeV are not covered by Multijet Balance calibration and a larger uncertainty is taken from the single particle response measurement. Each uncertainty is independent from the others and fully correlated across p_{T} and η. A reduced set of systematic uncertainties is available for physics analyses while minimising the loss of correlations.

3.6 *b*-jet tagging

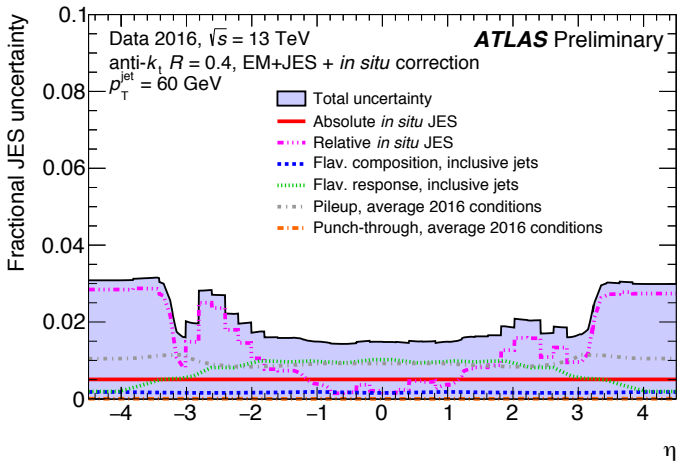
The search for Higgs boson pair production presented in this thesis is characterized by the presence of jets originating from b-quarks in the final state. Therefore their precise identification is of key importance for both selecting the HH signal and reducing the contribution of some of the expected backgrounds.

The rate at which a true b-jet is identified defines the b-tagging efficiency for a particular b-tagging algorithm. The true flavor of a jet is defined in simulated data using a spatial ΔR matching between stable hadrons and reconstructed jets. A hierarchical matching is performed, first checking whether a b-hadron can be matched, followed by a charm hadron and followed by a τ-lepton. This matching procedure results in a jet being classified respectively as either a b-jet, c-jet, τ-jet or a light-jet (if no match is found).

In ATLAS, various algorithms [107], based on the information from the reconstructed tracks and displaced secondary vertices in the ID, are used in order to identify



(a)



(b)

Figure 3.9: Fractional jet energy scale systematic uncertainty components, as a function of a) jet transverse momentum at $\eta = 0$ and b) η for jets with $p_{\mathrm{T}} > 60$ GeV [106].

b-jets, taking advantage of the relatively long lifetime of b-hadrons.[3]

There are three main categories of b-tagging algorithms commonly used in ATLAS. There are impact parameter based algorithms (IP2D, IP3D); inclusive secondary vertex reconstruction algorithms (SV); and decay chain reconstruction algorithms (JetFitter). These algorithms contribute complementary information and can be combined using a multivariate function to create a single b-tagging discriminant (MV).

The impact parameter algorithms use the signed impact parameter significances, d_0/σ_{d_0} and z_0/σ_{z_0}, of the tracks which are matched to a jet. A positive or negative sign is assigned to the impact parameter, based on the assumption that the decay point of the b-hadron lies along its flight path. Thus, tracks from b-hadron decays tend to have positive-signed impact parameters, indicating they originated in front of the PV, whereas tracks originating from the primary vertex will have the same probability for a negative or positive sign, due to resolution effects. Furthermore, the division of the impact parameter by its uncertainty (σ), ensures that more weight is given to tracks which are measured more precisely.

The secondary vertex algorithms attempt to reconstruct an inclusive displaced secondary vertex which is formed by selected tracks within a jet. All possible two-track vertices are identified from the track candidates. Vertices which are likely from long-lived particles (K_s,Λ), photon conversions or hadronic interactions with the detector material are removed. A single inclusive vertex is reconstructed by iteratively removing outlier tracks until a good vertex candidate is identified. Once reconstructed, the kinematic properties of the tracks can be used to offer discrimination between the b-jet and light-jet hypotheses.

The multi-vertex decay chain algorithm attempts to reconstruct the weak decay chain of a b-hadron produced at the primary vertex which subsequently decays to a

[3]The lifetime of B hadrons is of the order of 1.6 ps. Considering the mass of a B hadron of 5 GeV and its energy of 30 GeV, the typical flight path is approximately ~ 3 mm.

charm hadron (producing a tertiary vertex). A Kalman filter is used to identify the b-hadron flight path using the assumption that the tertiary vertex lies along the same flight path. The tracks which are associated to the reconstructed vertices can also provide discriminating information, in a similar way to the SV algorithm.

The three types of algorithms provide input to a multivariate classifier (MV). In Run 2 this classifier is a boosted decision tree (BDT) trained to discriminate b-jets from light-jets. The background samples of light-jets can contain a mixture of charm and light-jets in order to improve the charm-jet rejection. The discrimination performance is shown in Figure 3.10a for the MV2 algorithm with a 10% charm-jet admixture (MV2c10). The performance between the Run 1 MV algorithm (MV1) and the Run 2 MV algorithm (MV2) has seen an improvement of 10%, primarily driven by the inclusion of the IBL. Between 2015 and 2016, retraining of the b-tagging classifier has further improved the charm-jet rejection by around 40% at the 77% working point. The charm-jet rejection rate as a function of b-tagging efficiency is shown in Figure 3.10b. Jets which have been selected using a top-pair event selection are shown in Figure 3.10c comparing the performance of the b-tagging weight (MV2c10) in simulation and data, where a good agreement can be observed.

Calibrations of b-tagging algorithms derive a correction factor applied to simulated data to match the b-tagging efficiency measured in data. This correction factor is dependent on the true flavor of the simulated jet and the jet kinematics. A correction factor is derived both for true b-jet which are correctly identified as well as for light and charm-jets which are incorrectly b-tagged. In order to derive these scale factors, the performance of the b-tagging algorithm needs to be measured using events in data which are enriched in a particular jet flavor whilst correctly modelling the backgrounds.

The b-tagging efficiency is measured by using dileptonically decaying top-pair

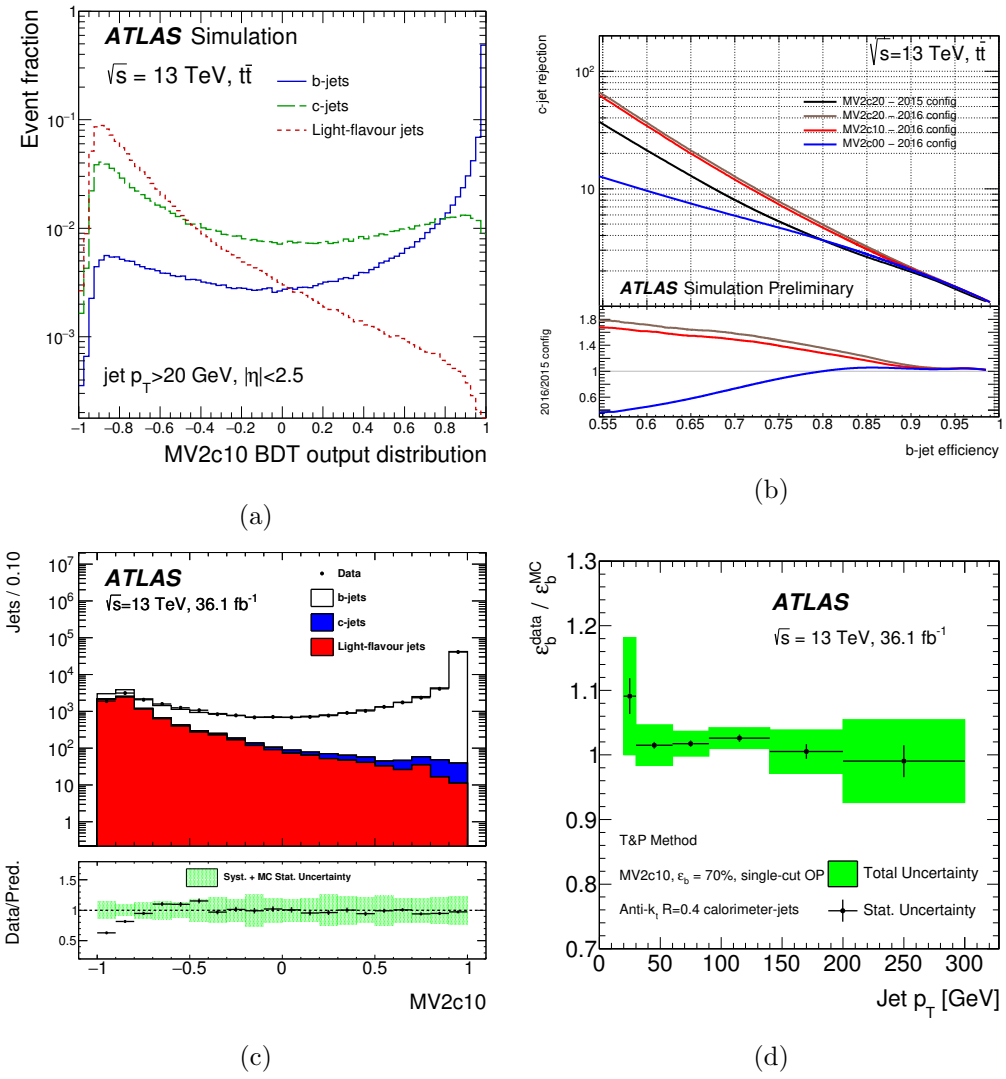


Figure 3.10: a) Normalized MV2c10 distribution for light-flavor, c- and b-quark initiated jets in simulated $t\bar{t}$ events [108]. b) c-jet rejection versus b-jet efficiency for the previous (2015) and the current configuration (2016) of the MV2 b-tagging algorithm evaluated on $t\bar{t}$ events [109]. c) The distribution of the MV2c10 multivariate discriminant applied to jets in a $t\bar{t}$-dominated sample of events selected by requiring an opposite-sign $e\mu$ pair and at least two jets with 2015+2016 data [108]. d) Data-to-simulation scale factors for the MV2c10 algorithm at the 70% working point as a function of the p_{T} of the probe jet obtained selecting $t\bar{t}$ dilepton events [108].

events as they are pure in b-jets. Two methods are employed to extract this efficiency, a combinatorial likelihood approach [110] and a tag-and-probe method. The current b-tagging efficiency scale factors calculated with the latter method as a function of jet p_{T} are shown in Figure 3.10d, where the scale factors agree with unity within errors. Jets in the intermediate p_{T} range have a total uncertainty of 2–4%, and are dominated by the effect of top-pair modelling.

3.7 Missing transverse energy

Neutrinos or other weakly-interacting particles produced by proton collisions will not interact with the ATLAS detector, and thus escape undetected. Indirect information on the presence of these particles can be inferred imposing the momentum conservation in the plane transverse to the beam direction. Since the initial state has zero momentum component in the transverse plane, an imbalance in the total measured transverse momentum in the final state would indicate the presence of an invisible particle being produced.

The missing transverse momentum ($\overrightarrow{E_{\mathrm{T}}^{\mathrm{miss}}}$) is reconstructed as the negative vector sum of transverse momenta ($\overrightarrow{p_{\mathrm{T}}}$) of reconstructed physics objects. The magnitude of the missing transverse energy is denoted by $E_{\mathrm{T}}^{\mathrm{miss}}$. The physics objects considered in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation are electrons, photons, muons, τ-leptons and jets (hard terms). The reconstructed momentum not associated to any of the hard terms is referred as the soft term and is also considered in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. Several algorithms can be used to reconstruct the $E_{\mathrm{T}}^{\mathrm{miss}}$ soft term using calorimeter energy deposits or tracks [111]. The main algorithm for the soft term reconstruction used by ATLAS during Run 2 fully relies on tracks and it is called Track Soft Term (TST). The algorithm is very robust against varying pile-up condition, but it misses the contribution from neutral particles.

The removal of pile-up jets is essential for $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution. This is done with the jet-vertex-tagger (JVT) technique which extracts the pile-up jets using track-to-vertex association method [112]. In addition, a novel forward pileup tagging technique (fJVT) that exploits the correlation between central and forward jets originating from pileup interactions was developed for Run 2 [113]. The fJVT improves the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution in high pile-up conditions as shown in Figure 3.11. The performance of

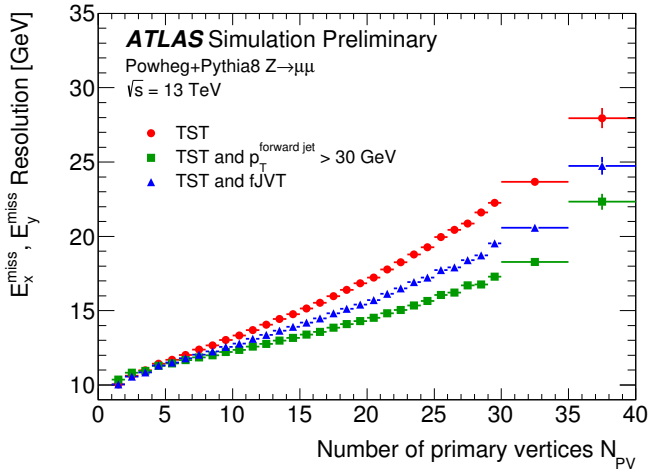
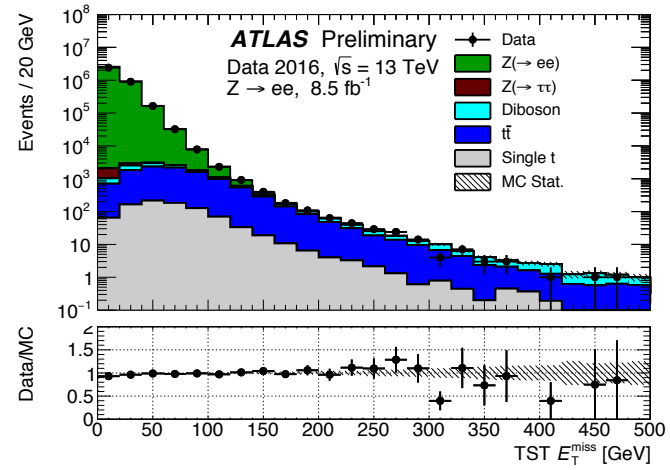


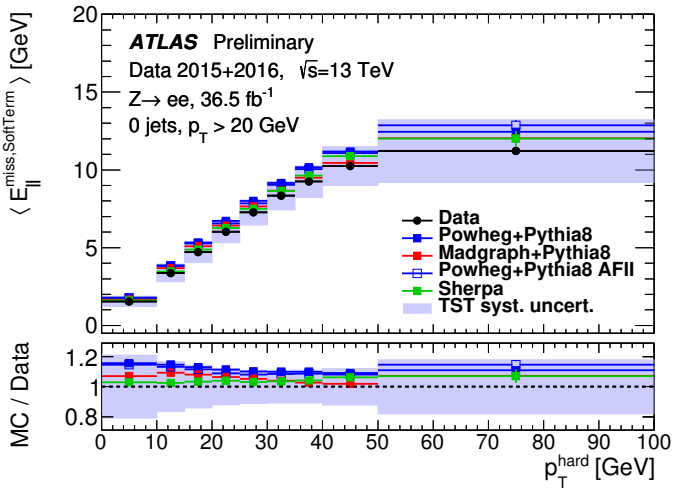
Figure 3.11: The Track Soft Term $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution as a function of N_{PV} measured in MC simulated $Z \to \mu\mu$ events using different strategies for pile-up suppression [114].

TST $E_{\mathrm{T}}^{\mathrm{miss}}$ is validated using $Z \to \ell\ell$, $W \to \ell\nu$ and $t\bar{t}$ events. Good agreement of TST $E_{\mathrm{T}}^{\mathrm{miss}}$ measured using 8.5 fb^{-1} Run 2 ATLAS data and MC simulation is observed as shown in Figure 3.12a with $Z \to ee$ events.

The TST systematic uncertainties are evaluated exploiting the differences between data and Monte Carlo using the balance of a soft term and a calibrated physics objects. The systematic uncertainties of each hard term are propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$. The mean of the TST distribution as a function of the hard term p_{T} measured using 36.5 fb^{-1} Run 2 ATLAS data agrees with MC simulation within the systematic uncertainty as shown in Figure 3.12b.



(a)



(b)

Figure 3.12: a) TST $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for a selection of Z boson decays to a pair of electrons at Run 2 ATLAS data. The expectation is superimposed by POWHEG+PYTHIA8 MC simulated events for the relevant signal physics processes including some background processes while diboson backgrounds use SHERPA [115]. b) The mean of the TST distribution projected in the direction longitudinal to the hard term p_{T} for $Z \to ee$ events measured using Run 2 ATLAS data and MC simulation [114].

Chapter 4

TileCal Response to Single Isolated Charged Particles

Contributions to the data-taking, the data acquisition system and the performance of the Tile Calorimeter (see Section 2.2.3) were carried out in the scope of this thesis. The performance of the TileCal is probed in-situ by using the calorimeter response to single hadrons. These objects, along with jets, deposit more energy in the hadronic calorimeter compared to muons and therefore the response to higher energies can be probed. They are selected from p-p collision data, and only isolated particles are considered.

The response is characterized by the ratio of the energy of a charged hadron (E) as measured by the TileCal, to that hadron's momentum (p), as measured by the ATLAS inner detector system (see Section 2.2.2). The hadron energy in both data and MC events are calibrated to the electromagnetic energy scale. The double ratio of the mean value ($\langle E/p \rangle$) between data and MC simulation should be approximately one, with deviations from unity possibly due to poor electromagnetic scale calibration in the data or differences in the MC description due to a relatively complex hadron

shower development.

This chapter is organized as follows: data and simulated samples used in the analysis are listed in Section 4.1, event and particle selection is described in Section 4.2, measurement results are discussed in Section 4.3, systematics are detailed in Section 4.4, and finally Section 4.5 contains the conclusions of this study.

4.1 Data and Simulated Samples

The data used in this analysis spans over the three years of data-taking of LHC Run 1. The 2012 data set corresponds to an integrated luminosity of 20.3 fb^{-1} before prescaling, collected between April and November of 2012. Events triggered by a fixed rate random trigger (EF_rd0_filled_NoAlg) [116] correspond to an integrated luminosity of 129 nb^{-1}. In 2012 the spacing between proton bunches was either 25 or 50 ns and the average number of interactions per crossing $\langle\mu\rangle$ was between 3 and 30. The 25 ns bunch spacing data corresponds to one single run (216399) with a total luminosity of 0.419 nb^{-1}. Both 25 and 50 ns bunch spacing runs have been used in this analysis. The simulated and experimental samples from the 2012 data set are listed in Table 4.1. The number of events quoted corresponds to the number of events from the DESD containers that passes the trigger selection.

The 2011 data set corresponds to an integrated luminosity of 4.7 fb^{-1}, collected between February and October of 2011. The events were triggered using the same fixed rate random trigger found in the Minimum Bias stream. This corresponds to an integrated luminosity of 15.5 nb^{-1} when trigger prescale is taken into account. In 2011 the spacing between proton bunches was either 50 or 75 ns and $\langle\mu\rangle$ was between 10 and 18. The corresponding samples from the 2011 data set are listed in Table 4.2.

The 2010 data set collected between March and September of 2010 corresponds to an integrated luminosity of 35 pb^{-1} and comprises a data set of 92 nb^{-1} after the

Period	Name	#events
A	data12_8TeV physics_MinBias repro14_v01	3061147
B	data12_8TeV physics_MinBias repro14_v01	2241203
C	data12_8TeV physics_MinBias repro14_v01	5612455
D	data12_8TeV physics_MinBias t0pro13_v01	1450355
E	data12_8TeV physics_MinBias repro14_v01	878221
G	data12_8TeV physics_MinBias repro14_v01	430147
H	data12_8TeV physics_MinBias repro14_v01	267602
I	data12_8TeV physics_MinBias t0pro14_v01	584590
J	data12_8TeV physics_MinBias t0pro14_v01	3717415
L	data12_8TeV physics_MinBias t0pro14_v01	371514
M	data12_8TeV physics_MinBias t0pro14_v01	984688
MC	mc12_8TeV 119997 Pythia8_A2MSTW2008LO minbias_ND recon e1119_s1479_s1470_r3843	13385253

Table 4.1: The simulated and experimental samples for the 2012 data set.

Period	Name	#events
B	data11_7TeV physics_MinBias pro10_v01	1151146
D	data11_7TeV physics_MinBias pro10_v01	1044827
E	data11_7TeV physics_MinBias pro10_v01	249682
F	data11_7TeV physics_MinBias pro10_v01	1221351
G	data11_7TeV physics_MinBias pro10_v01	1504734
H	data11_7TeV physics_MinBias pro10_v01	819663
I	data11_7TeV physics_MinBias pro10_v01	995167
J	data11_7TeV physics_MinBias pro10_v01	464306
K	data11_7TeV physics_MinBias pro10_v01	1287780
L	data11_7TeV physics_MinBias pro10_v01	1224198
M	data11_7TeV physics_MinBias pro10_v01	920106
MC	mc11_7TeV 105001 pythia_minbias recon e815_s1310_s1300_r3043	4998695

Table 4.2: The simulated and experimental samples for the 2011 data set.

requirement of the Minimum Bias Trigger Scintillators (MBTS) [117] is applied. The bunch crossing spacing was 150 ns, and $\langle\mu\rangle$ was much lower than in the case of 2011

and 2012. The 2010 data set samples are listed in Table 4.3.

The simulated samples correspond to minimum bias non-diffractive events generated with Pythia 6 [118] in 2010 and 2011 periods, and Pythia 8 for 2012. The detector is simulated using Geant4 9.4 [119] with the QGSP_BERT [120–122] physics lists. The reconstruction and analysis was done using Athena release 17 for both data and simulated samples.

Period	Name	#events
A	data10_7TeV physics_MinBias repro05_v02	6527176
B	data10_7TeV physics_MinBias repro05_v02	13235025
C	data10_7TeV physics_MinBias repro05_v02	1982540
D	data10_7TeV physics_MinBias repro05_v02	3177312
E	data10_7TeV physics_MinBias repro05_v02	2170958
F	data10_7TeV physics_MinBias repro05_v02	1014949
G	data10_7TeV physics_MinBias repro05_v02	1284495
H	data10_7TeV physics_MinBias repro05_v02	839477
I	data10_7TeV physics_MinBias repro05_v02	931879
MC	mc10_7TeV 105001 pythia_minbias recon e577_s1024_s946_r1926	4999689

Table 4.3: The simulated and experimental samples for the 2010 data set.

4.2 Event Selection

The samples reported in Section 4.1 are skimmed and slimmed into the form of D3PDs using the software package TileD3PDMaker [123]. At this stage, a pre-selection is applied on the data and the MC simulation requiring the two following conditions:

- The tracks are required to be isolated in p_{T}: the sum of the tracks p_{T} in a cone $\Delta R = 0.4$ around the track is less than 15% of that of the p_{T} of the track.
- The tracks are required to have a momentum higher than 2 GeV in order to reach

the TileCal, and that the extrapolated track, using the ATLAS Extrapolator Tool [124], falls within the limits of the calorimeter, $|\eta| < 2$.

These two selection requirements reduce the number of events approximately by a factor 2 in all data sets.

Event selection is made on the processed D3PDs using the following selection criteria:

- It is required that the event run number and lumiblock is marked as good in the Good Run Lists (GRL) [125] and that there is at least one track with $p_{\mathrm{T}} > 2$ GeV, that does not extrapolate into a dead module for which the TileTripReader (TTR) [126] tool is used.

- Tracks are selected if they have at least one hit in the pixel, one hit in the TRT and six hits in the SCT detectors. This cut eliminates outliers from the distribution of the number of hits. The loose requirement on the number of TRT hits associated to the track is applied in order to include particles that may have undergone a hadronic interaction earlier in the inner detector [127].

- It is required that the tracks are inside the TileCal geometrical acceptance ($|\eta| < 1.7$) when extrapolated to each of the calorimeter layers.

- The impact parameters (d_0 and z_0) have to $|d_0| < 1.5$ mm and $|z_0 \sin(\theta)| < 1.5$ mm, respectively.

- It is required that $\langle\mu\rangle$ is between 3 and 25 for the 2012 data set, to have sufficient event statistics for the data to MC comparisons both at the high end and at the low end of the μ distribution. This requirement is not present for the 2010 or 2011 data sets.

- The hadron candidate should be associated with an energy deposit in the calorimeter. The energy of the hadron is defined as the sum of the energy of the cells calibrated to the EM scale that belong to all the topoclusters that have their center within $\Delta R = 0.2$ of the track extrapolated into the Tile Calorimeter. This selection requirement is based on a previous MC study [128]. An alternative determination of the energy deposited by the hadron using the total energy of the clusters within $\Delta R = 0.2$ of the extrapolated track was also considered but it was found to be sub-optimal for this study.

- The energy deposited by the hadron in the LAr is required to be compatible to that of a minimum ionizing particle, thus depositing less than 1 GeV along the path without a lower bound energy cut allowing for clusters of negative energy to be considered.

- The amount of energy deposited by the hadron in the TileCal (R) is required to be at least 75% of its total energy in order to increase the purity of hadrons in the sample.

The effect of this selection criteria is discussed in Section 4.4. Tables 4.4, 4.5 and 4.6 show the total and relative efficiency reduction per track of each selection criteria for the 2012, 2011 and 2010 data sets, respectively. The requirement on the inner detector hits is observed to eliminate a higher percentage of the tracks in the MC simulation than in the data in all the three years of Run 1 data-taking. This is explained by the data / MC discrepancies in the distributions of the number of Pixel, TRT and SCT hits, where the MC presents a much larger relative peak in zero than the data in all of them.

	Data		MC	
Selection criteria	Total (%)	Relative (%)	Total (%)	Relative (%)
Track $p_{\mathrm{T}} > 2$ GeV & GRL & TTR	51.1	51.1	48.6	48.6
One hit in Pixel and TRT, six hits in SCT	25.4	49.7	10.2	20.9
Extrapolated track $\lvert\eta\rvert < 1.7$	21.6	84.8	8.5	83.4
$\lvert d_0\rvert < 1.5$ mm and $\lvert z_0 \sin\theta\rvert < 1.5$ mm	20.9	97.2	8.3	97.7
$3 < < \mu > < 25$	16.3	77.9	7.0	84.2
ΔE(LAr) < 1 GeV	9.4	57.8	3.8	54.8
ΔE(Tile)/ΔE(LAr+Tile) > 0.75	2.2	23.1	0.9	23.6

Table 4.4: Total and relative efficiencies (in %) of each consecutive selection criteria for the 2012 data and simulation.

	Data		MC	
Selection criteria	Total (%)	Relative (%)	Total (%)	Relative (%)
Track $p_{\mathrm{T}} > 2$ GeV & GRL & TTR	42.5	42.5	41.2	41.2
One hit in Pixel and TRT, six hits in SCT	13.8	32.6	7.5	18.3
Extrapolated track $\lvert\eta\rvert < 1.7$	11.6	83.7	6.3	83.7
$\lvert d_0\rvert < 1.5$ mm and $\lvert z_0 \sin\theta\rvert < 1.5$ mm	11.4	98.3	6.1	96.9
$3 < < \mu > < 25$	-	-	-	-
ΔE(LAr) < 1 GeV	5.1	52.9	2.9	48.2
ΔE(Tile)/ΔE(LAr+Tile) > 0.75	1.3	25.5	0.8	26.1

Table 4.5: Total and relative efficiencies (in %) of each consecutive selection criteria for the 2011 data and simulation.

	Data		MC	
Selection criteria	Total (%)	Relative (%)	Total (%)	Relative (%)
Track $p_{\mathrm{T}} > 2$ GeV & GRL & TTR	48.7	48.7	44.0	44.0
One hit in Pixel and TRT, six hits in SCT	11.7	24.1	7.5	17.1
Extrapolated track $\lvert\eta\rvert < 1.7$	10.2	87.0	6.4	85.2
$\lvert d_0\rvert < 1.5$ mm and $\lvert z_0 \sin\theta\rvert < 1.5$ mm	10.1	98.8	6.3	98.2
$3 < < \mu > < 25$	-	-	-	-
ΔE(LAr) < 1 GeV	6.1	60.8	3.7	58.6
ΔE(Tile)/ΔE(LAr+Tile) > 0.75	1.5	24.0	1.0	26.0

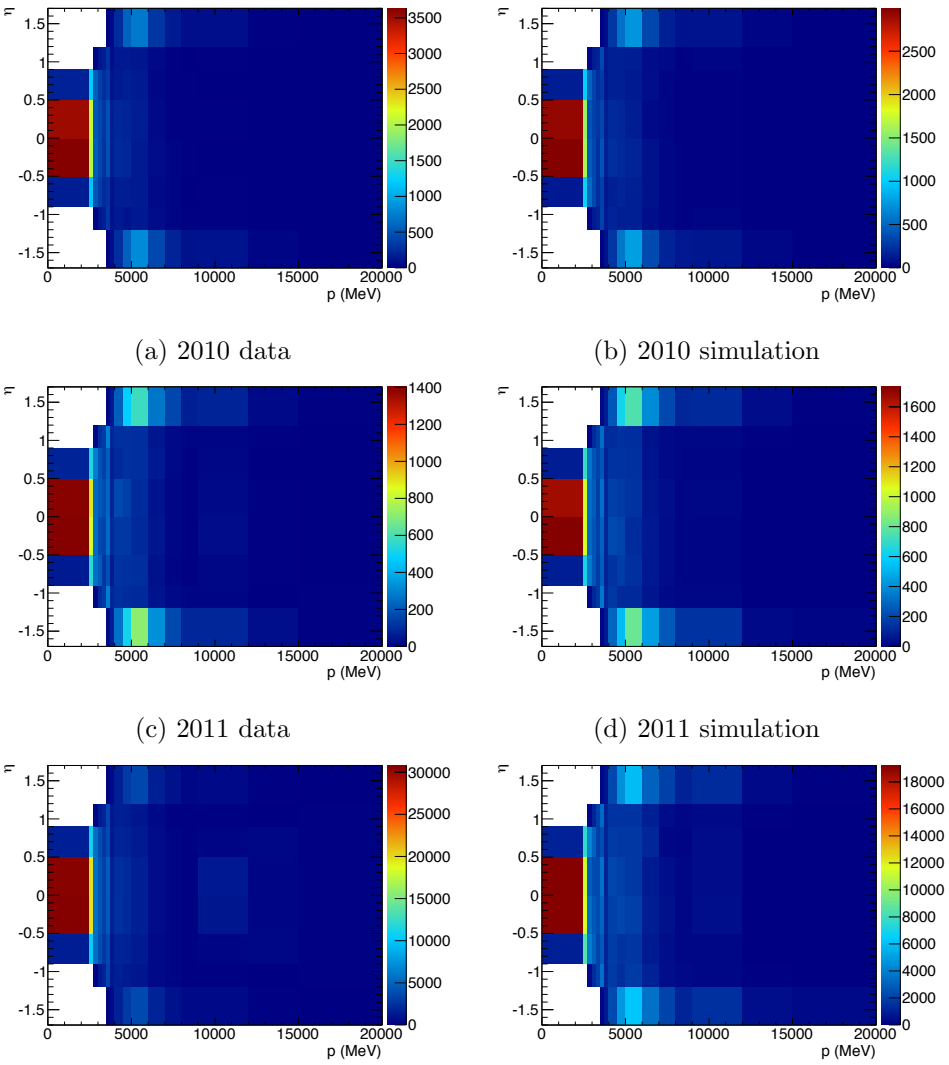
Table 4.6: Total and relative efficiencies (in %) of each consecutive selection criteria for the 2010 data and simulation.

4.2.1 Track Reweighting

Simulated samples are reweighted so that the distribution of the average number of interactions per bunch crossing matches that of the data sample. For this the pile-up reweighting tool [129] is used to compute the weight that is given to every event in the MC.

There is also another reweighting applied to the MC due to the fact that the tracks momentum and η spectra in MC is different to the one in data. It receives the name of Track Reweighting (TRW). The simulation is weighted by the ratio of the number of tracks in data to simulation in 8 bins of pseudorapidity and 16 bins in momentum. The bins are chosen so that there are roughly the same number of tracks per bin as a function of these two variables. Figure 4.1 shows the number of tracks as a function of η and momentum in data and simulated samples for the 2010, 2011 and 2012 data-taking periods. The reach in momentum for this analysis is limited to $\sim$ 10 GeV due to a lack of statistics.

Comparisons between the non-reweighted and reweighted MC and data track η and p distributions are shown in Figures 4.2, 4.3, 4.4 and 4.5, respectively. The agreement data to MC shows an uniform improvement through all periods with a larger impact on the 2012 data set, in both η and p distributions. It is applied to all data sets for consistency.



(a) 2010 data

(b) 2010 simulation

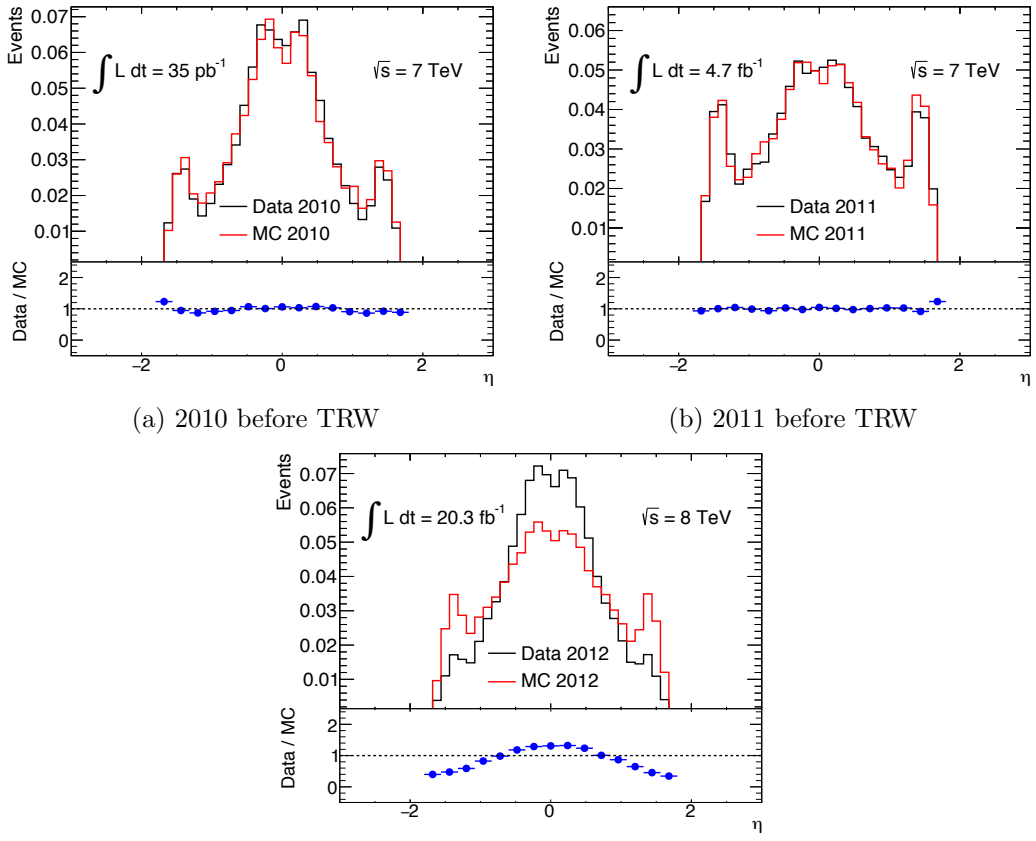
(c) 2011 data

(d) 2011 simulation

(e) 2012 data

(f) 2012 simulation

Figure 4.1: Number of tracks as a function of momentum and η, with $|\eta| < 1.7$ and $p_{\mathrm{T}} > 2$ GeV after full event selection for a) 2010 data, b) 2010 MC, c) 2011 data, d) 2011 MC, e) 2012 data and f) 2012 MC.

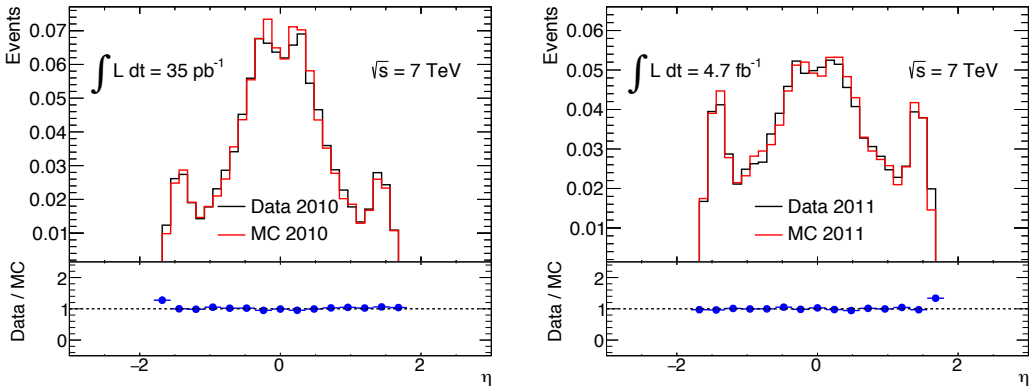


(a) 2010 before TRW

(b) 2011 before TRW

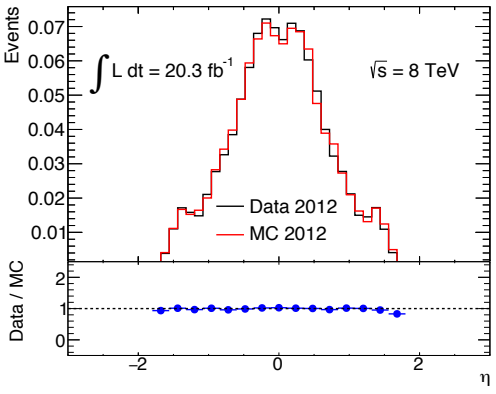
(c) 2012 before TRW

Figure 4.2: Distribution of η of the tracks after full event selection for a) 2010, b) 2011 and c) 2012 before applying the track reweighting.



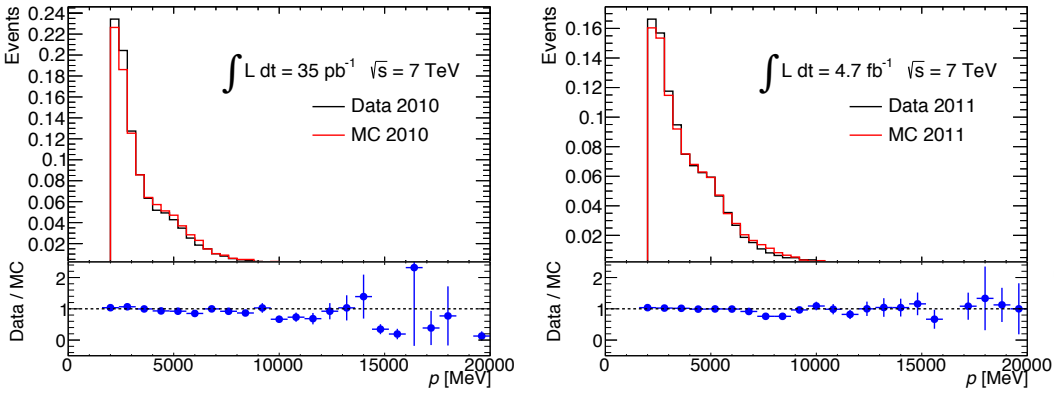
(a) 2010 after TRW

(b) 2011 after TRW



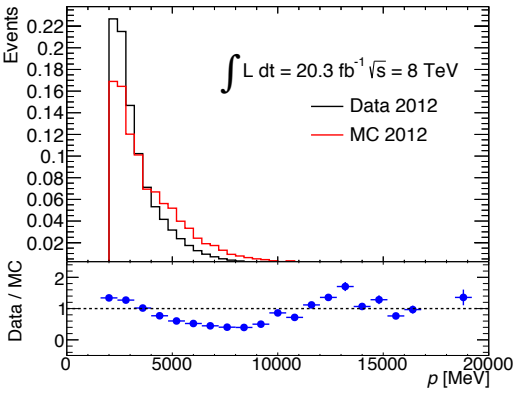
(c) 2012 after TRW

Figure 4.3: Distribution of η of the tracks after full event selection for a) 2010, b) 2011 and c) 2012 after applying the track reweighting.



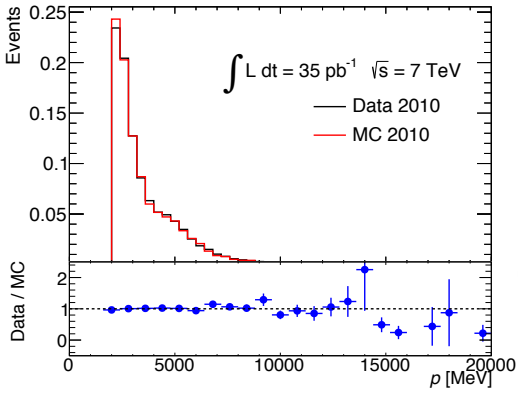
(a) 2010 before TRW

(b) 2011 before TRW

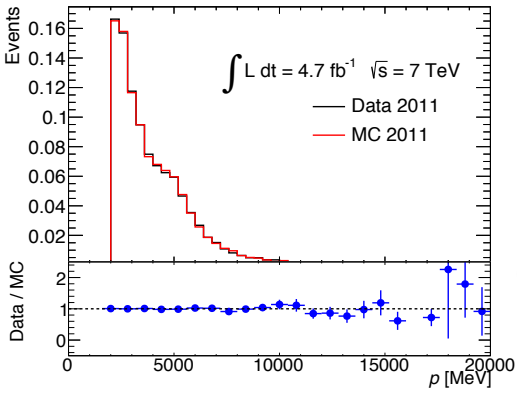


(c) 2012 before TRW

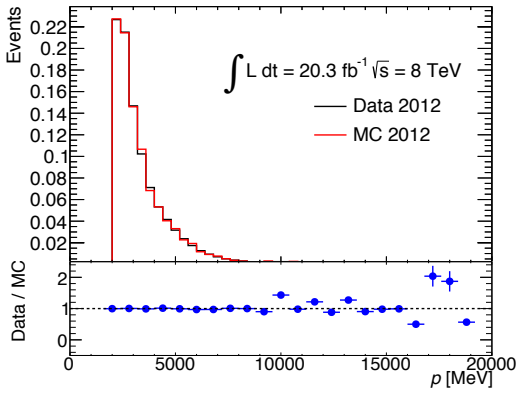
Figure 4.4: Distribution of momentum of the tracks after full event selection for a) 2010, b) 2011 and c) 2012 before applying the track reweighting.



(a) 2010 after TRW



(b) 2011 after TRW



(c) 2012 after TRW

Figure 4.5: Distribution of momentum of the tracks after full event selection for a) 2010, b) 2011 and c) 2012 after applying the track reweighting.

4.3 Results

Figure 4.6 shows the E/p distribution for particles passing the full event selection detailed in Section 4.2. The mean value of the E/p distribution, $\langle E/p \rangle$, integrated over momentum, η and ϕ, is in good agreement throughout the three years of Run 1 data-taking. The large fraction of tracks with $E/p = 0$ observed in all data-taking periods corresponds to tracks with momentum larger than 2 GeV but with no topoclusters within $\Delta R = 0.2$ of the track extrapolated into the Tile Calorimeter. This can happen if either a particle interacts hadronically before reaching the calorimeter or no single energy deposit is large enough to seed a topological cluster. This feature is included in the following figures but not in the analysis. The negative tail of this distribution is caused by noise in the calorimeter, which for data-taking conditions with low $\langle \mu \rangle$ consists mostly of electronics noise, while the long positive tail is caused by the background of neutral particles, since these particles add to the measured E but not to p.

Table 4.7 shows the $\langle E/p \rangle$ for data and MC for the three years of Run 1 data-taking, showing consistent predicted measurements across the years. The difference between data and MC is largest in the lower part of the distribution where the tails of the electronics noise are dominating. This discrepancy is found to be randomly distributed across momentum and η, which suggests that the modeling of hadronic interactions, rather than of geometry, is primarily responsible for it.

	2010	2011	2012
$\langle E/p \rangle^{\text{data}}$	0.507 ± 0.002	0.448 ± 0.002	0.502 ± 0.001
$\langle E/p \rangle^{\text{MC}}$	0.505 ± 0.001	0.479 ± 0.002	0.505 ± 0.001

Table 4.7: $\langle E/p \rangle$ for data and simulated samples for data-taking years from 2010 to 2012. Note that only statistical uncertainties are shown. The peak at zero is not included in this measurement.

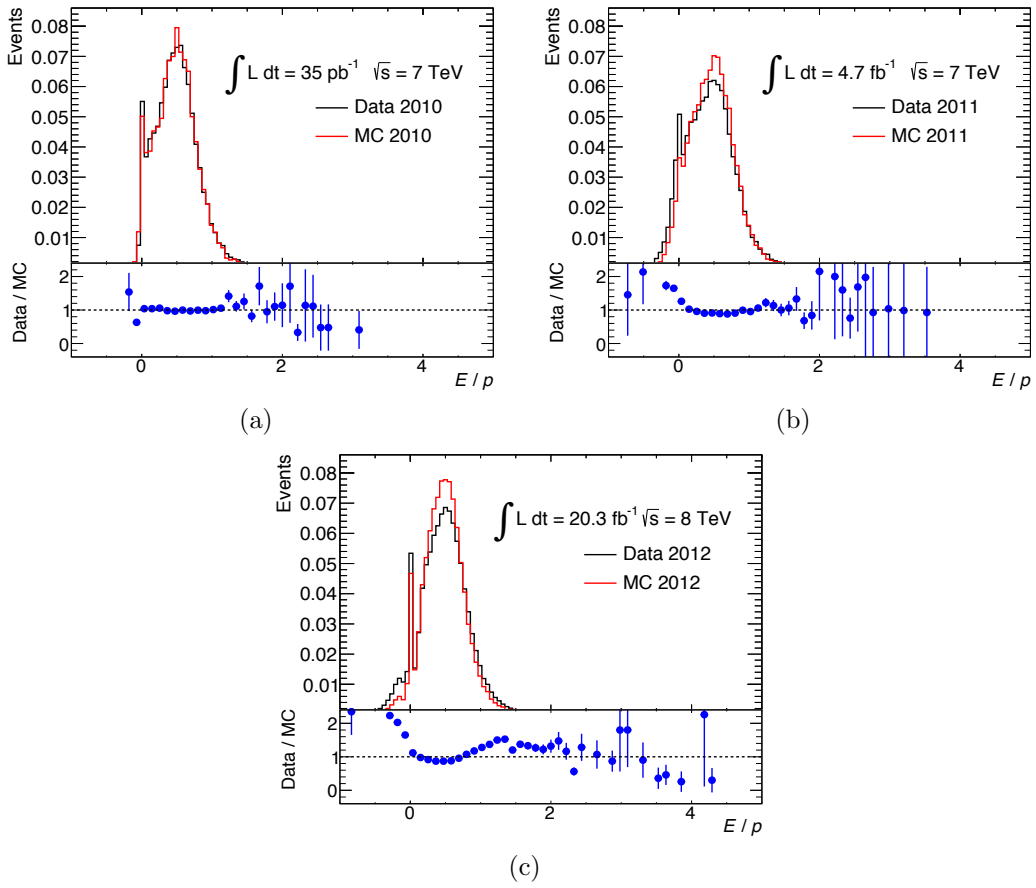


Figure 4.6: E/p distribution response to charged hadrons after full event selection (see Section 4.2) for a) 2010, b) 2011 and c) 2012 data sets.

Double ratios of the experimental to simulated data of $\langle E/p \rangle$ are used to validate the agreement between the data and the MC. The results for all three years are listed in Table 4.8. The double ratio in 2011, which shows the largest deviation from unity, agrees with this result within 1.5σ.

$\langle E/p \rangle$ is evaluated as a function of η, ϕ, momentum and average number of interactions for the 2010, 2011 and 2012 data-taking years, as shown in Figures 4.7, 4.8, 4.9, and 4.10 respectively. The agreement between data and MC as a function of η

	2010	2011	2012
Double ratio of $\langle E/p \rangle$	1.003 ± 0.004	0.936 ± 0.007	0.994 ± 0.004

Table 4.8: Double ratios of the experimental to simulated data of the $\langle E/p \rangle$ for data-taking years from 2010 to 2012. Note that only statistical uncertainties are shown.

is not very good in the transition region between the Long Barrel and the Extended Barrel ($0.9 < |\eta| < 1.3$), with a disagreement that reaches up to 15% in 2010, 25% in 2011 and 15% in 2012 and is explained by the increased number of forward events in 2011 as shown in Figure 4.3b. This large discrepancy is not so clear as a function of ϕ, that yields a value close to 0.5 for all the years.

The increase of the value of the $\langle E/p \rangle$ as a function of momentum in the range between 2 and 5 GeV is observed in the 2010 and 2012 data sets as a result of the trigger configuration used. An increase in $\langle E/p \rangle$ is also observed in 2012 above 10 GeV. These effects are well reproduced in the MC. As a function of the average number of interactions, $\langle E/p \rangle$ stays around 0.5 with very small variation. The maximum disagreement is 10% in 2011 systematically for all values of $\langle \mu \rangle$ and with data being consistently lower than MC. However, it is less than 5% in 2012 with no particular outlier.

Table 4.9 shows the mean value of the double E/p ratio ($\langle E/p \rangle^{\text{data}} / \langle E/p \rangle^{\text{MC}}$) for the different years as a function of η, ϕ, momentum and number of collisions per bunch crossing.

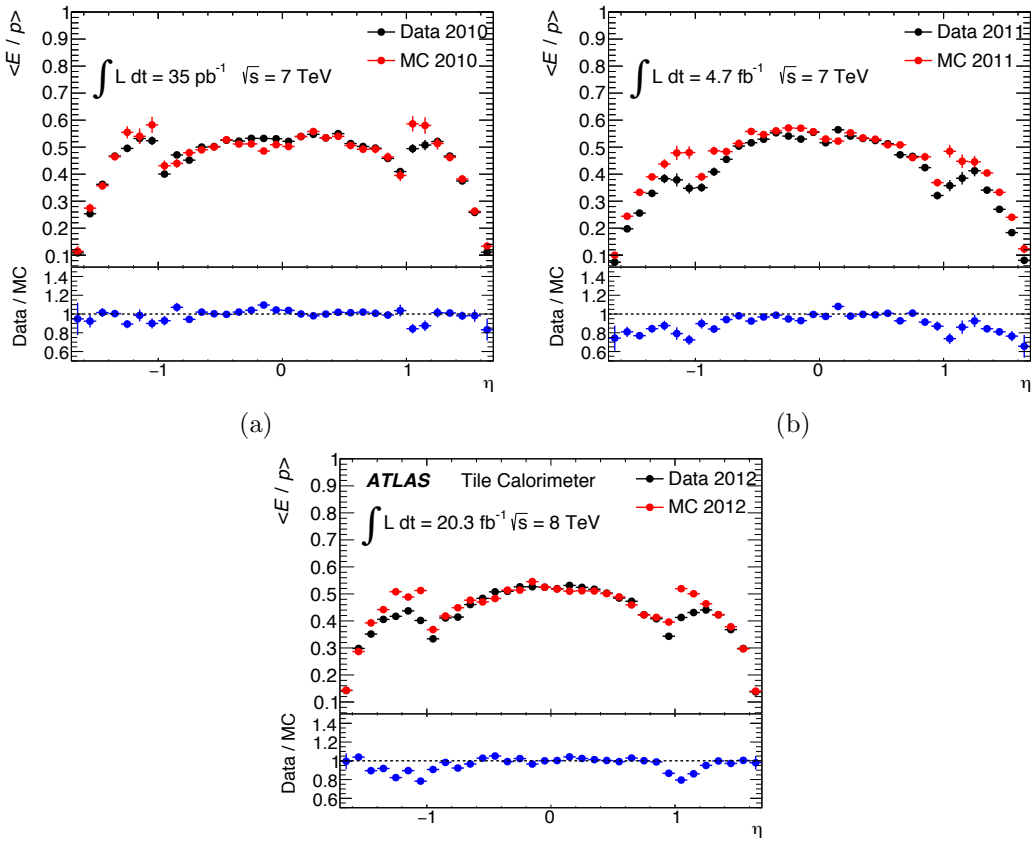
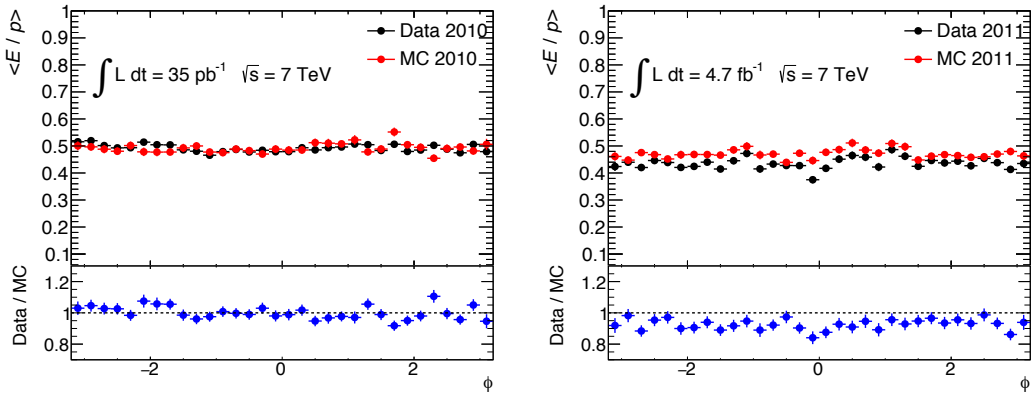


Figure 4.7: Distributions of $\langle E/p \rangle$ for single charged hadrons as a function of η after full event selection (see Section 4.2) for a) 2010, b) 2011 and c) 2012 data sets. Note that only statistical uncertainties are shown.

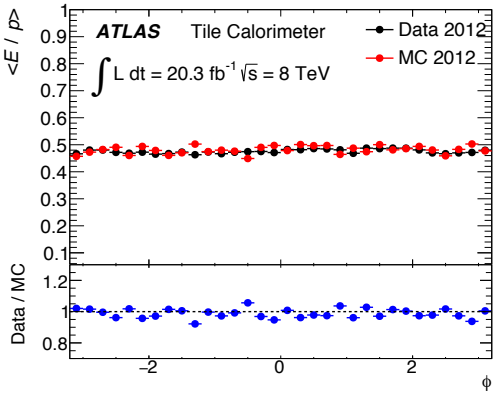
	2010	2011	2012
Double ratio of $\langle E/p \rangle$ vs η	1.003 ± 0.004	0.936 ± 0.006	0.994 ± 0.004
Double ratio of $\langle E/p \rangle$ vs ϕ	1.003 ± 0.004	0.936 ± 0.007	0.994 ± 0.004
Double ratio of $\langle E/p \rangle$ vs p	1.003 ± 0.004	0.936 ± 0.006	0.993 ± 0.005
Double ratio of $\langle E/p \rangle$ vs $\langle \mu \rangle$	-	0.936 ± 0.005	0.994 ± 0.005

Table 4.9: Double ratios of the experimental to simulated data of the $\langle E/p \rangle$ as a function of η, ϕ, momentum and average number of interactions per crossing for data-taking years from 2010 to 2012. Note that only statistical uncertainties are shown.



(a)

(b)



(c)

Figure 4.8: Distributions of $\langle E/p \rangle$ for single charged hadrons as a function of ϕ after full event selection (see Section 4.2) for a) 2010, b) 2011 and c) 2012 data sets. Note that only statistical uncertainties are shown.

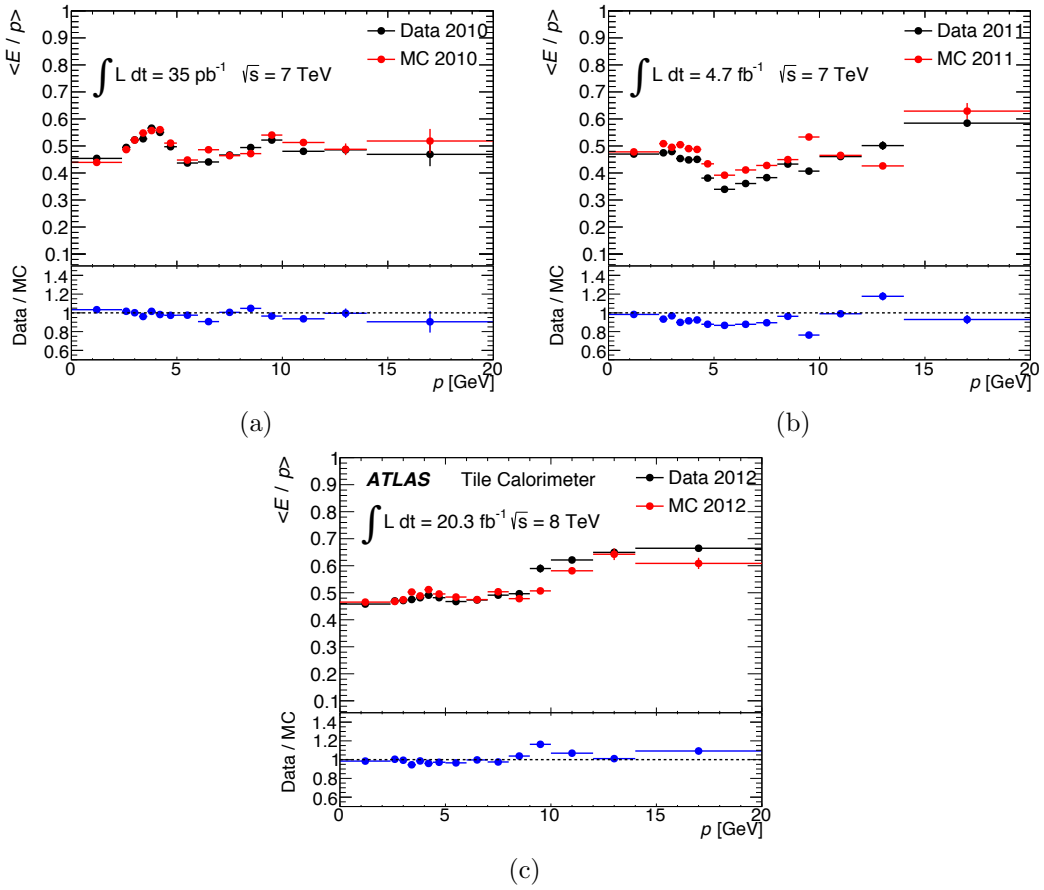
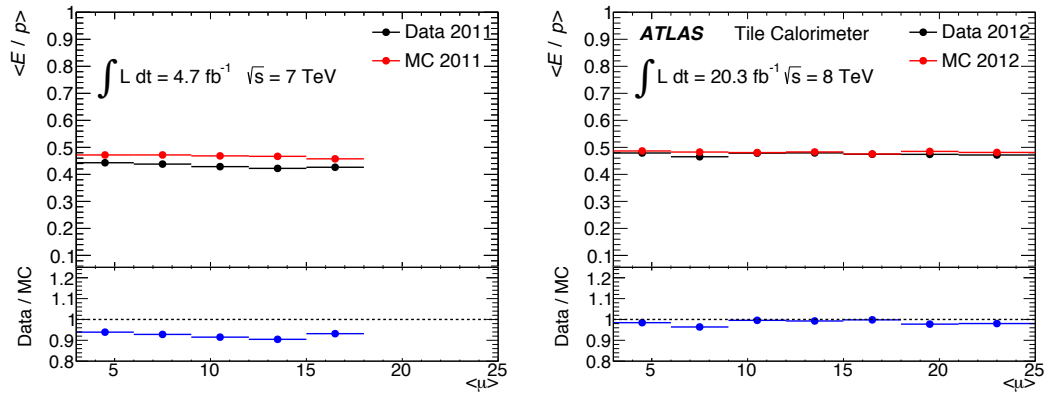


Figure 4.9: Distributions of $\langle E/p \rangle$ for single charged hadrons as a function of the momentum with $|\eta| < 1.7$ and $p_{\mathrm{T}} > 2$ GeV after full event selection (see Section 4.2) for a) 2010, b) 2011 and c) 2012 data sets. Note that only statistical uncertainties are shown.



(a) (b)

Figure 4.10: Distributions of $\langle E/p \rangle$ for single charged hadrons as a function of μ with $|\eta| < 1.7$ and $p_{\mathrm{T}} > 2$ GeV after full event selection (see Section 4.2) for a) 2011 and b) 2012 data sets. Note that only statistical uncertainties are shown.

4.4 Systematic uncertainties

The systematic uncertainties on the measurement of $\langle E/p \rangle$ are related to the uncertainties on track and energy reconstruction. The results shown in this section are calculated using the 2012 data set only.

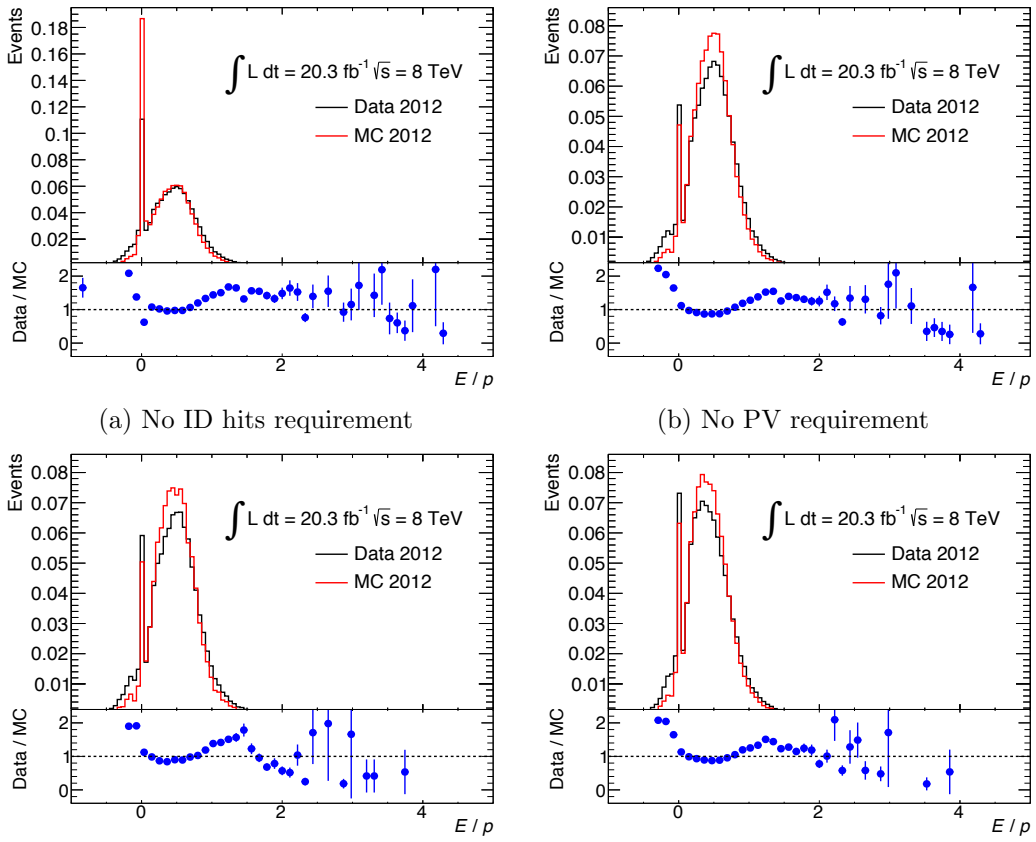
4.4.1 Bin to bin impact

In order to estimate the impact of the different selection cuts over the E/p, δs (see equation 4.1) is defined as the difference between the nominal and the variation of the quantity (s) defined as the weighted sum over all the bins of the normalized double $\langle E/p \rangle$ ratio, computed as the bin value from the E/p distribution in data over the bin value from the E/p distribution in MC, and the weight (w_i) corresponds to the sum of the number of entries in each bin from the E/p distribution from data and the E/p distribution from MC. The further this quantity is from zero, the stronger the impact of the selection cut.

$$\delta s = s^{variation} - s^{nominal} \text{ where } s = \frac{\sum_{i}^{bins} \left(1 - \frac{E/p_i^{\text{data}}}{E/p_i^{\text{MC}}}\right) w_i}{\sum_{i}^{bins} w_i} \tag{4.1}$$

This quantity is very sensitive to shape variations. For example in the extreme cases of the removal of the inner detector hits or primary vertex requirements, it yields a value of 6.4% and 2.0% respectively, as shown in Figures 4.11a and 4.11b. Shifting the energy deposit requirement in LAr down by 1σ of the distribution of the energy deposit (300 MeV) has an impact of 4.5% (see Figure 4.11d). Increasing the number of ID hits from one to three in the Pixel, from six to eight in the SCT and from one to fourteen in the TRT has an equivalent impact of 1.3%. These values are chosen from

the distribution of hits before applying the full event selection. They correspond to the central value of the distributions in the number of Pixel and SCT hits, and to the end of the plateau region in the number of TRT hits. The resulting E/p distribution in this case is shown in Figure 4.11c.



(a) No ID hits requirement

(b) No PV requirement

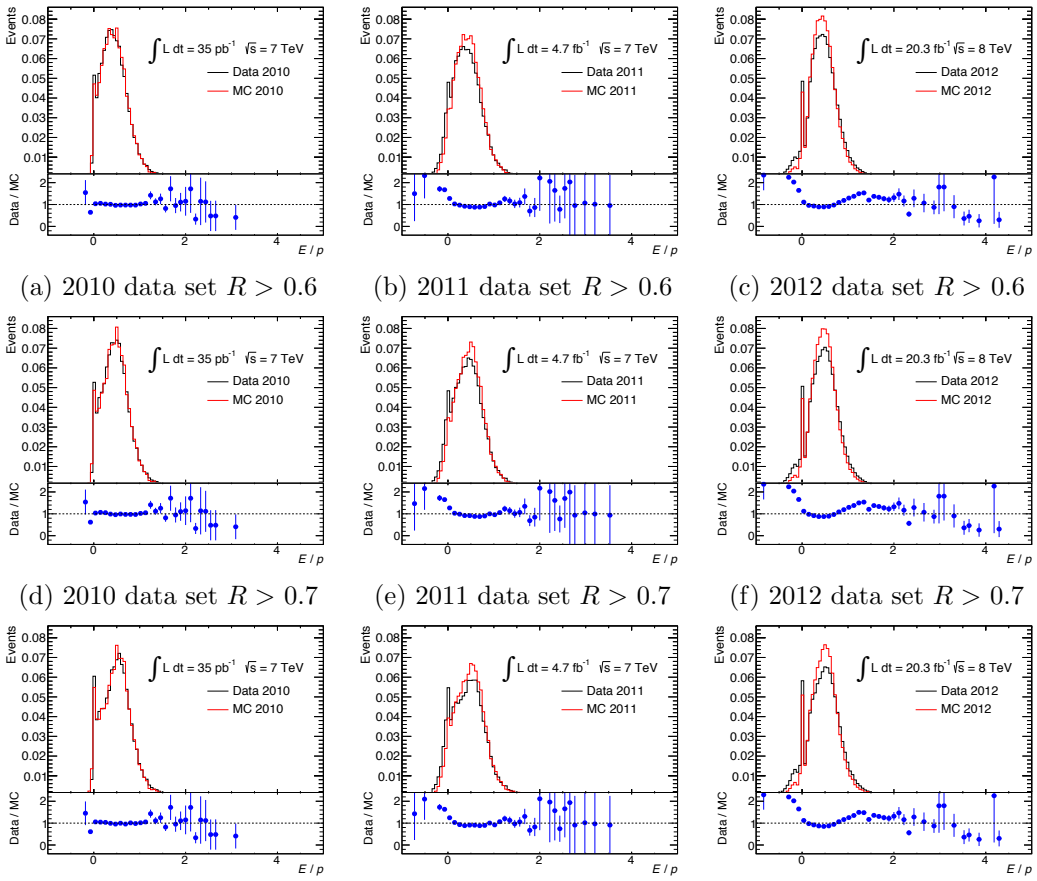
(c) Hits: Pixel >3 & SCT >8 & TRT > 14

(d) $\Delta E(\text{LAr}) < 300$ MeV

Figure 4.11: Distributions of $\langle E/p \rangle$ for single charged hadrons after full event selection (see Section 4.2) for different variations of the cuts using the 2012 data set. Note that only statistical uncertainties are shown.

Similarly δs are used to calculate the impact of the variation of the fraction of energy deposited in the TileCal ($R = \Delta E_{\text{Tile}}/\Delta E_{\text{Tile+LAr}}$), where the largest impact is 1.1% for ratios 5% higher and lower than the nominal value. Figure 4.12 shows the

distribution of the E/p for each value of R for each data set.



(a) 2010 data set $R > 0.6$ (b) 2011 data set $R > 0.6$ (c) 2012 data set $R > 0.6$

(d) 2010 data set $R > 0.7$ (e) 2011 data set $R > 0.7$ (f) 2012 data set $R > 0.7$

(g) 2010 data set $R > 0.8$ (h) 2011 data set $R > 0.8$ (i) 2012 data set $R > 0.8$

Figure 4.12: Distributions of $\langle E/p \rangle$ for single charged hadrons after full event selection (see Section 4.2) for different relative values of energy deposited in TileCal ($R = \Delta E(\text{Tile})/\Delta E(\text{Tile+LAr})$) using the 2010, 2011 and 2012 data sets. Note that only statistical uncertainties are shown.

Table 4.10 shows the value of δs for the variation of each cut. According to δs, the requirements on the energy deposit in LAr and the number of ID hits have a much more important impact than the cut in R.

Variation	δs (%)
Removal of inner detector hits requirement	6.4
Removal of primary vertex requirement	2.0
$\Delta E(\text{LAr}) < 300$ MeV	4.5
Pixel hits >3 & SCT hits >8 & TRT hits > 14	1.3
$\Delta E(\text{Tile})/\Delta E(\text{Tile+LAr}) > 0.6$	0.6
$\Delta E(\text{Tile})/\Delta E(\text{Tile+LAr}) > 0.7$	0.3
$\Delta E(\text{Tile})/\Delta E(\text{Tile+LAr}) > 0.8$	1.1

Table 4.10: Bin to bin impact of the variation of the selection cuts.

4.4.2 Likelihood method

To better quantify the relative contribution of each systematic a binned likelihood function of the expected E/p distribution to the E/p distribution from data is built (see equation 4.2). Different uncertainties enter in the expected E/p distribution as corrections to the nominal E/p distribution extracted from the MC. Each correction term is the product of a parameter α times the difference of the nominal $\langle E/p \rangle$ in MC and the variational $\langle E/p \rangle$ in MC. These variations take into account the following uncertainties: ΔE_{LAr}, R, no primary vertex cut and no inner detector hits cut. The likelihood function is minimized to extract the best value of the α parameters.

$$\begin{aligned}\mathcal{L}\left(E/p^{\text{data}}, E/p^{\text{expected}}\right) &= \prod_{i}^{bins} Poiss\left(E/p_i^{\text{data}}, E/p_i^{\text{expected}}\right) \\ &= \prod_{i}^{bins} Poiss\left(E/p_i^{\text{data}}, E/p_i^{\text{MC,nominal}} + \sum_{j}^{variations} \alpha_j \left(E/p_i^{\text{MC,nominal}} - E/p_i^{\text{MC},j}\right)\right)\end{aligned} \tag{4.2}$$

The fit value for the α parameters along with the expected $\langle E/p \rangle$ and the difference with respect to the nominal $\langle E/p \rangle$ from MC are shown in Table 4.11.

The systematic uncertainty calculated as the quadratic sum of all the differences yields a value of $\varepsilon(\langle E/p \rangle^{\text{MC}}) = -0.038$. Thus the value of the $\langle E/p \rangle$ in MC is $\langle E/p \rangle^{\text{MC}} = 0.505 \pm 0.001(\text{stat.}) \pm 0.038(\text{sys.})$. This is compatible with the value of

Systematic	α	$\langle E/p\rangle^{\mathrm{MC}}_{exp}$	$\langle E/p\rangle^{\mathrm{MC}}_{exp} - \langle E/p\rangle^{\mathrm{MC}}_{nom}$
$\Delta E(\mathrm{LAr}) < 300$ MeV	0.73	0.458	-0.047
$R > 0.6$	0.26	0.491	-0.014
No PV cut	0.05	0.502	-0.003
No Hits cut	0.01	0.454	-0.051

Table 4.11: Fit value for the α parameters, the expected $\langle E/p\rangle$ and the difference with respect to the nominal $\langle E/p\rangle$ from MC for the different parameters considered in the likelihood fit.

the $\langle E/p\rangle$ in data, which corresponds to $\langle E/p\rangle^{\mathrm{data}} = 0.502 \pm 0.001(\mathrm{stat.})$.

4.4.3 Up and down method

The systematic uncertainty on the double $\langle E/p\rangle$ ratio, $\langle E/p\rangle^{\mathrm{data}}/\langle E/p\rangle^{\mathrm{MC}}$ is evaluated by shifting each variational component up and down by one standard deviation. Increased and decreased values of the following event selection cuts are used: energy loss in LAr, relative value of the energy deposited in TileCal, R, and the number of hits in the inner detector.

The criteria used to choose the thresholds for the energy loss in LAr is based on the population of the distribution, the lower limit is set 1σ below the mean value at 300 MeV and the upper limit is set at 1σ above the mean value at 1.5 GeV. Similarly the criteria for the threshold on number of the hits in the ID is to cut on the relative maximum of their distributions.

The upper and lower limit for the systematic uncertainty is obtained from the sum in quadrature of the different up and down contributions. Table 4.12 shows the difference between the variational ratio and the nominal ratio.

Double ratios of the experimental to simulated data of $\langle E/p\rangle$ are used to validate the agreement between the data and the MC. The overall double ratio for the 2012 data set is measured and corresponds to a value of $\langle E/p\rangle^{\mathrm{data}}/\langle E/p\rangle^{\mathrm{MC}} = 0.994 \pm 0.003\ (\mathrm{stat.})\ ^{+0.009}_{-0.014}(\mathrm{sys.})$. This result shows a good calibration of cell energy to the

Systematic	$\langle E/p\rangle_{var}^{\text{data}}$	$\langle E/p\rangle_{var}^{\text{MC}}$	$(\langle E/p\rangle^{\text{data}}/\langle E/p\rangle^{\text{MC}})_{var}$	$\langle E/p\rangle^{\text{data}}/\langle E/p\rangle_{var}^{\text{MC}}$ $-\langle E/p\rangle^{\text{data}}/\langle E/p\rangle_{nom}^{\text{MC}}$
ΔE(LAr)<300 MeV	0.448	0.456	0.982	-0.012
ΔE(LAr)<1.5 GeV	0.501	0.503	0.996	0.002
$R > 0.7$	0.496	0.497	0.998	0.004
$R > 0.8$	0.503	0.509	0.988	-0.006
Less Hits in ID	0.464	0.464	1.000	0.006
More Hits in ID	0.498	0.503	0.990	-0.004

Table 4.12: Difference between the variational double $\langle E/p\rangle$ ratio and the nominal double $\langle E/p\rangle$ ratio.

electromagnetic scale and also a good agreement between experimental data and MC predictions for the single hadrons.

4.4.4 Summary

Table 4.13 shows a summary of the results obtained for the systematic uncertainty on the double ratio of $\langle E/p\rangle$ for the 2012 data set using the three different methods detailed in Section 4.4. Note that the likelihood method only provides a systematic on $\langle E/p\rangle^{\text{MC}}$. The up and down method constitutes the most robust and accurate procedure to estimate this uncertainty and is therefore chosen as the final systematic uncertainty on the double ratio of the experimental to simulated data of $\langle E/p\rangle$.

	Bin to bin impact	Likelihood method	Up and down method
Systematic uncertainty	0.023	0.038	$^{+0.009}_{-0.014}$

Table 4.13: Results of the systematics on the double ratio of $\langle E/p\rangle$ for the 2012 data set, except in the case of the likelihood method, where the calculation refers to $\langle E/p\rangle^{\text{MC}}$ only.

4.5 Conclusions

Overall E/p studies show that ATLAS has a good understanding of the energy scale for isolated hadrons using the three years of Run 1 data-taking. The 2010 data validate

the work done in detector calibration and simulation tuning with the calorimeter test beam [130, 131].

The statistical uncertainty of the double $\langle E/p \rangle$ ratio was extracted and found to be smaller than 1% for all the years, being 0.4% in 2012, 0.7% in 2011 and 0.4% in 2010. The systematic uncertainty of the same observable was extracted and found to be less than 2%.

This result is compatible with the one obtained for central isolated hadrons by the combined calorimeter E/p studies [132,133]. In the Long Barrel region ($|\eta| < 0.7$) the 3% level agreement is maintained despite sizeable changes in beam conditions, showing that in this region pile-up is well simulated and energy reconstruction techniques are robust against pile-up.

The data to MC ratios are compatible with 1 in the 2010 and 2012 data sets and lower than 1 in the 2011 data set with a difference of 5%. The η and p transition regions corresponding to $0.7 < |\eta| < 1.1$ and momentum range between 9 and 10 GeV, as well as the high momentum region ($p > 16$ GeV) show some data / MC discrepancies not covered by the systematic uncertainty which leaves open the possibility of further simulation development.

Chapter 5

Search for di-Higgs production in the $\gamma\gamma b\bar{b}$ final state

This chapter presents the search for Higgs boson pair production in the ATLAS $\gamma\gamma b\bar{b}$ final state using a data set of 36.1 fb^{-1} at $\sqrt{s} = 13$ TeV. It covers the event selection and categorization of the analysis as well as the strategy followed to model the signal and background processes. Final results including expected and observed limits are presented. The chapter concludes with a summary of the main findings.

5.1 Introduction

A search for the production of pairs of Higgs bosons in the $\gamma\gamma b\bar{b}$ final state is presented, and considers both resonant (see Section 1.5.2) and non-resonant (see Section 1.5.1) contributions.

For the non-resonant search, deviations in the Higgs boson self-coupling are characterized using κ_λ ($\kappa_\lambda = \lambda^{HHH}/\lambda^{HHH}_{SM}$) while κ_t characterises deviations in the Yukawa coupling between the top quark and the Higgs boson ($\kappa_t = y_t/y_{t,SM}$). In

both cases the SM subscript refers to the SM value of these parameters. For the resonant search, the narrow-width approximation is used, focusing on a resonance with mass (m_X) in the range 260 GeV $< m_X <$ 1000 GeV. The lower limit (260 GeV) is chosen for being close to the kinematic threshold for HH production, and the upper limit (1000 GeV) corresponds to the boundary where a reliable estimate of the background could be performed with the current data set.

Although this search is for a generic scalar decaying to a pair of Higgs bosons, the simulated samples used to optimize the search were produced for the gluon-fusion production mode. Previous searches have been carried out by the ATLAS [56] and CMS [134, 135] collaborations in the $\gamma\gamma b\bar{b}$ channel at both $\sqrt{s} = 8$ TeV and $\sqrt{s} = 13$ TeV, as well as in other final states [55, 136–138].

5.2 Data samples and MC simulation

The search discussed in this thesis uses the full p-p collisions data set of 36.1 fb^{-1} collected in 2015 and 2016 at $\sqrt{s} = 13$ TeV. Only events that pass the high data-quality standards are analyzed in order to reject events recorded during periods with sub-detector failures or defects.

Data events are selected using a diphoton trigger requiring the presence in the EM calorimeter of two clusters of energy depositions with transverse energy above 35 GeV and 25 GeV for the leading (highest transverse energy) and subleading (second highest transverse energy) cluster. In the high-level trigger the shape of the energy deposition of both clusters is required to be loosely consistent with that expected from an electromagnetic shower initiated by a photon. The diphoton trigger has an efficiency greater than 99% for events that satisfy the final event selection (described in Section 5.3).

A set of MC simulated samples is used to model the signal and background pro-

cesses considered in the search. Non-resonant production of Higgs boson pairs via the gluon-gluon fusion process is simulated at NLO accuracy in QCD using form factors for the top-quark loop from HPAIR [139, 140] to approximate finite top-quark mass effects. The simulated events are reweighted to reproduce the m_{HH} spectrum obtained in [141] and [142], which calculated the process at NLO in QCD while fully accounting for the top-quark mass. The total cross section is normalized to 33.41 fb, in accordance with a calculation at NNLO in QCD [34, 143].

Non-resonant BSM Higgs boson pair production with varied κ_λ is simulated at LO accuracy in QCD [144] for eleven values of κ_λ in the range -10 $< \kappa_\lambda <$ 10. The total cross sections for these samples is computed as a function of κ_λ at LO accuracy in QCD and a constant NNLO/LO k-factor (2.283) computed at $\kappa_\lambda = 1$, is then applied. Theoretical studies have shown that the k-factor as a function of κ_λ is constant to within a few percent [145]. As the amplitude for Higgs boson pair production can be expressed in terms of κ_λ and the top Yukawa coupling, weighted combinations of the simulated samples can produce predictions for other values of κ_λ.

Resonant BSM Higgs boson pair production, using a narrow width approximation, is simulated at NLO accuracy for ten different mass points (260, 275, 300, 325, 350, 400, 450, 500, 750 and 1000 GeV).

This analysis is affected by backgrounds from single Higgs boson production and non-resonant backgrounds with continuum spectra. Background estimation is carried out using data-driven methods whenever possible; in particular, data is used to estimate the continuum background contribution from SM processes with multiple photons and jets, which forms the dominant background for this search. Monte Carlo event generators are used for the simulation of the background from SM single Higgs boson production, as shown in Table 5.1. The major single Higgs bosons production channels contributing to the background are gluon-gluon fusion (ggH),

associated production with a Z-boson (ZH), associated production with heavy flavor ($t\bar{t}H$) and associated production with a single top-quark (tH). Among these, the largest contributions come from $t\bar{t}H$ and ZH.

A list of the signal and background samples used in this search together with the programs used for matrix element generation, for parton showering, hadronization and underlying event and the PDF sets used is shown in Table 5.1.

Table 5.1: Summary of the event generators and PDF sets used to model the signal and the background processes. The SM cross sections σ for the Higgs boson production processes with $m_H = 125.09$ GeV are also given separately for $\sqrt{s} = 13$ TeV, together with the orders of the calculations corresponding to the quoted cross sections, which are used to normalize the samples, after multiplication by the Higgs boson branching ratio to diphotons, 0.227%.

Process	Generator	Showering	PDF set	σ [pb] $\sqrt{s} = 13$ TeV	Order of calculation of σ
Non-resonant SM HH	MG5_aMC@NLO	Herwig++	CT10	33.41	NNLO+NNLL
Non-resonant BSM HH	MG5_aMC@NLO	Pythia8	NNPDF 2.3 LO	-	LO
Resonant BSM HH	MG5_aMC@NLO	Herwig++	CT10	-	NLO
ggH	Powheg NNLOPS	Pythia8	PDF4LHC15	48.52	N^{3}LO(QCD)+NLO(EW)
VBF	Powheg-Box	Pythia8	PDF4LHC15	3.78	NNLO(QCD)+NLO(EW)
WH	Powheg-Box	Pythia8	PDF4LHC15	1.37	NNLO(QCD)+NLO(EW)
$q\bar{q}' \to ZH$	Powheg-Box	Pythia8	PDF4LHC15	0.76	NNLO(QCD)+NLO(EW)
$gg \to ZH$	Powheg-Box	Pythia8	PDF4LHC15	0.12	NLO+NLL(QCD)
$t\bar{t}H$	MG5_aMC@NLO	Pythia8	NNPDF3.0	0.51	NLO(QCD)+NLO(EW)
$b\bar{b}H$	MG5_aMC@NLO	Pythia8	CT10	0.49	5FS(NNLO)+4FS(NLO)
t-channel tH	MG5_aMC@NLO	Pythia8	CT10	0.07	4FS(LO)
W-associated tH	MG5_aMC@NLO	Herwig++	CT10	0.02	5FS(NLO)
$\gamma\gamma$	Sherpa	Sherpa	CT10	-	LO

5.3 Object and event selection

5.3.1 Photons

Events are preselected by requiring at least two photon candidates with $E_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.37$ (excluding the transition region between the barrel and endcap calorimeters of $1.37 < |\eta| < 1.52$) that fulfill the loose photon identification criteria [81]. The two photon candidates with the highest E_{T} are chosen as the diphoton candidate, and used to identify the diphoton primary vertex among all reconstructed vertices, using a neural-network algorithm based on track and primary vertex information, as well as the directions of the two photons measured in the calorimeter and inner detector [146]. The neural-network algorithm selects a diphoton vertex within 0.3 mm of the true $H \to \gamma\gamma$ production vertex in 79% of simulated gluon-gluon fusion events. For the other Higgs boson production modes this fraction ranges from 84% to 97%, increasing with jet activity or the presence of charged leptons. The performance of the diphoton primary vertex neural-network algorithm is validated using $Z \to e^+e^-$ events in data and simulation, by ignoring the tracks associated with the electron candidates and treating them as photon candidates. Sufficient agreement between the data and the simulation is found. The diphoton primary vertex is used to redefine the direction of the photon candidates, resulting in an improved diphoton invariant mass resolution. The invariant mass of the two photons is given by $m_{\gamma\gamma} = \sqrt{2E_1E_2(1-\cos\alpha)}$, where E_1 and E_2 are the energies of the leading and subleading photons and α is the opening angle of the two photons with respect to the selected production vertex. Studies performed using $HH \to \gamma\gamma b\bar{b}$ events show that the selection efficiency difference between using the neural-network diphoton vertex and the highest $\sum_i^{N_{trk}} p_{\mathrm{T},i}^2$ primary vertex is about 1–2%.

Following the identification of the diphoton primary vertex, the leading and subleading photon candidates in the diphoton candidate are respectively required to have

$E_{\mathrm{T}}/m_{\gamma\gamma} > 0.35$ and 0.25, and to both satisfy the tight identification criteria as well as the calorimeter and track isolation requirements.

5.3.2 Jets, *b*-jets and muons

Jets

Jets are required to satisfy $p_{\mathrm{T}} > 25$ GeV for $|\eta| < 2.5$. Jets that do not originate from the diphoton primary vertex are rejected, for $|\eta| < 2.4$, using the JVT algorithm, which combines tracking information into a multivariate likelihood. For jets with $p_{\mathrm{T}} < 60$ GeV and $|\eta| < 2.4$ a medium working point is used, with an efficiency greater than 92% for non-pileup jets with $p_{\mathrm{T}} > 20$ GeV. The efficiency of the JVT algorithm is corrected in the simulation to match that observed in the data. Jets are discarded if they are within $\Delta R = 0.4$ of an isolated photon candidate, or within $\Delta R = 0.2$ of an isolated electron candidate.

b-jets

Jets consistent with the decay of a b-hadron are identified using a multivariate discriminant, having as input information from track impact parameters and secondary vertices. Working points are defined by requiring the discriminant output to exceed a particular value that is chosen to provide a specific b-jet efficiency in an inclusive $t\bar{t}$ sample. Correction factors derived from $t\bar{t}$ events with final states containing two leptons are applied to the simulated event samples to compensate for differences between data and simulation in the b-tagging efficiency. The analysis uses two working points which have a b-tagging efficiency of 70% (60%), a c-jet rejection factor of 12 (35) and a light-jet rejection factor of 380 (1540) respectively.

Muons

Muons within $\Delta R = 0.4$ of a b-tagged jet are used to correct for energy losses from semileptonic b-hadron decays. This correction improves the energy measurement of b-jets and improves the signal acceptance by 5–6%.

5.3.3 Categorization

Events are selected for analysis if there are at least two photons and at least two jets, one or both tagged as b-jets. The diphoton invariant mass is initially required to fall within a broad mass window of 105 GeV $< m_{\gamma\gamma} <$ 160 GeV. In order to remain orthogonal to the ATLAS search for $HH \to b\bar{b}b\bar{b}$ [136], any event with more than two b-jets using the 70% efficient working point is rejected before the remaining events are divided into three categories.

- The 2-tag signal category consists of events with exactly two b-jets satisfying the requirement for the 70% efficient working point.

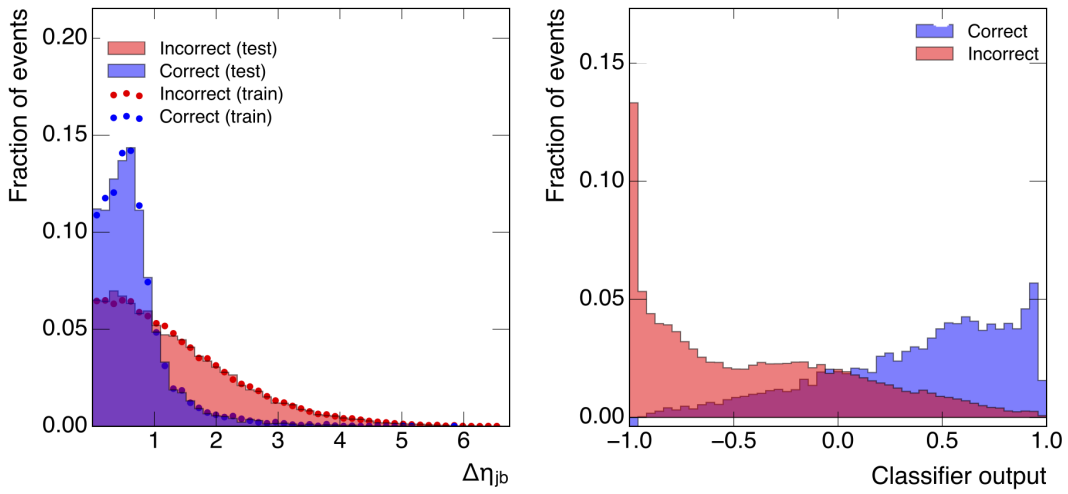
- The 1-tag signal category is defined using events failing the above requirement but nevertheless containing exactly one b-jet identified using a more stringent (60% efficient) working point. Here the second jet, which is in this case not identified as a b-jet, is chosen using a boosted decision tree (BDT). The BDT uses kinematic variables, namely jet p_{T}, dijet p_{T}, dijet mass, jet η, dijet η and the $\Delta\eta$ between the selected jets, as well as information about whether each jet satisfied less stringent b-tagging criteria. The ranking of the jets from best to worst in terms of closest match between the dijet mass and m_H, highest jet p_{T} and highest dijet p_{T} are also used as inputs. The jet with the highest BDT score is selected and the event is included in the 1-tag signal category. The efficiency with which the correct jet is selected by this BDT is 60–80% across the range of

resonant and non-resonant signal hypotheses considered in this search. The $\Delta\eta$ between the jets and the classification output from the BDT for the resonant 260 GeV sample is shown in Figure 5.1.

- The 0-tag control category uses events that contain no b-jet from either working point. This category is not directly used in the analysis, but is instead used to provide data-driven estimates of the background shape in the signal categories.

Loose and Tight selections

As a result of a number of optimization studies two different kinematic selections, called loose and tight, are defined, where the tight selection is a strict subset of the loose one. The searches for low-mass resonances and for non-SM values of the Higgs boson self-coupling both use the loose selection, as the average transverse momentum (p_{T}) of the Higgs bosons is lower in these cases [147] resulting in a higher sensitiv-



(a) (b)

Figure 5.1: a) The jet pair $\Delta\eta$ and b) the classification output from the BDT for the 260 GeV resonant signal sample.

ity to these analyses. Analogously, the tight selection is used for signals where the Higgs bosons typically have larger average p_{T}, namely higher-mass resonances and SM non-resonant HH production. Table 5.2 summarizes the use of the loose and tight selections in the different non-resonant and resonant analyses carried out in the scope of this Higgs boson pair production search.

Table 5.2: Summary of the kinematic selections used for the non-resonant and resonant analyses.

Analysis		Selection
Non-resonant	σ_{HH}^{SM}	Tight
	λ_{HHH}	Loose
Resonant	$m_X < 500$ GeV	Loose
	$m_X > 500$ GeV	Tight

Note that both the loose and tight selections have their own dedicated BDT to choose the second jet in the 1-tag category. It is optimized using simulated continuum background events as well as signal events from lower-mass or higher-mass resonances, respectively.

Further requirements are then made on the p_{T} of the jets and on the mass of the dijet system, which differ for the loose and tight selections. In the loose selection, the highest-p_{T} jet is required to have $p_{\mathrm{T}} > 40$ GeV, and the next-highest-p_{T} jet must satisfy $p_{\mathrm{T}} > 25$ GeV, with the invariant mass of the jet pair (m_{jj}) required to lie between 80 and 140 GeV. For the tight selection, the highest-p_{T} and the next-highest-p_{T} jets are required to have $p_{\mathrm{T}} > 100$ GeV and $p_{\mathrm{T}} > 30$ GeV, respectively, with 90 GeV $< m_{jj} <$ 140 GeV. Finally, in the resonant search, the diphoton invariant mass is required to be within 4.7 (4.3) GeV of the Higgs boson mass for the loose (tight) selection. This additional selection on $m_{\gamma\gamma}$ is optimized to contain at least 95% of the simulated Higgs boson pair events for each mass hypothesis. The various

possible combinations of signal category and selection are summarized in Tables 5.3 and 5.4 for the non-resonant and resonant analyses, respectively.

Table 5.3: Summary of the selection requirements made for each combination of (i) 1-tag or 2-tag signal category, and (ii) loose or tight selection, for the non-resonant analysis. Note that the search for non-SM values of the Higgs boson self-coupling, λ_{HHH}, uses the loose selection, while the tight selection is used for the search for SM non-resonant HH production.

	Non-resonant			
	1-tag		2-tag	
	Loose (λ_{HHH})	Tight (σ_{HH}^{SM})	Loose (λ_{HHH})	Tight (σ_{HH}^{SM})
$m_{\gamma\gamma}$ range [GeV]	105–160	105–160	105–160	105–160
Jet b-tagging WPs used	60% + BDT	60% + BDT	70%	70%
Leading jet p_{T} [GeV]	>40	>100	>40	>100
Subleading jet p_{T} [GeV]	>25	>105	>25	>30
m_{jj} range [GeV]	80–140	90–140	80–140	90–140

Table 5.4: Summary of the selection requirements made for each combination of (i) 1-tag or 2-tag signal category, and (ii) loose or tight selection, for the resonant analysis.

	Resonant			
	1-tag		2-tag	
	Loose	Tight	Loose	Tight
$m_{\gamma\gamma}$ range [GeV]	120.39–129.79	120.79–129.39	120.39–129.79	120.79–129.39
Jet b-tagging WPs used	60% + BDT	60% + BDT	70%	70%
Leading jet p_{T} [GeV]	>40	>100	>40	>100
Subleading jet p_{T} [GeV]	>25	>30	>25	>30
m_{jj} range [GeV]	80–140	90–140	80–140	90–140

5.3.4 Optimization

All optimizations are performed using the resonant Higgs boson pair production LO and non-resonant Higgs boson pair production NLO signal samples. The resonant signal cross section is scaled to the average limit on $HH \to \gamma\gamma b\bar{b}$ set with early Run 2 data [2]: 5 pb $\times$ $\mathcal{B}$ $(H \to \gamma\gamma)$ $\times$ $\mathcal{B}$ $(H \to b\bar{b})$ $\times$ 2 = 12.89 fb. The background is

taken from the 2-tag and 1-tag data with Loose-not-Tight, non-isolated photons. A transfer factor is obtained from the 0-tag data to extrapolate from Loose-not-Tight, non-isolated photons to Tight, isolated photons. The optimizations are performed expecting a total 2015+2016 dataset of 40 fb^{-1}. The figure of merit used for the optimization is defined by Asimov's formula [148] given below:

$$Z = \sqrt{2\Big((s+b)ln(1+\frac{s}{b}) - s\Big)}, \tag{5.1}$$

where s and b are the number of signal and background events that pass the analysis selection.

Photon ID and isolation optimization

Studies are performed to select the photon identification criteria and isolation working points in this search. `Tight` identification and `FixedCutLoose` isolation are the default choices for the $H \to \gamma\gamma$ analyses. Figure 5.2 shows the Asimov significance for the resonant signal sample at $m_X = 275$ GeV for a variety of photon ID and isolation choices, for the 2-tag category (1-tag category), using the chosen 70% (60%) b-tagging working point. The photon ID choices are shown in Table 5.5 and the photon isolation choices are shown in Table 5.6. In this optimization, jets are required to have $p_{\mathrm{T}} >$ 25 GeV and $|\eta| < 2.5$, with a dijet mass between 60 and 180 GeV.

The `FixedCutLoose` photon isolation working point matches the photon isolation definition in Chapter 3.2.3, where the calorimeter (track) isolation in a cone of ΔR = 0.2 is required to be less than 6.5% (5%) of the photon E_{T}. In `FixedCutTight`, the calorimeter isolation in a cone of $\Delta R = 0.4$ is required to be less than 2.45 GeV + 2.2% of the photon E_{T}, while the track isolation requirement remains the same as in `FixedCutLoose`. The analogous `CaloOnly` working points do not include the track isolation requirements, as their names suggest.

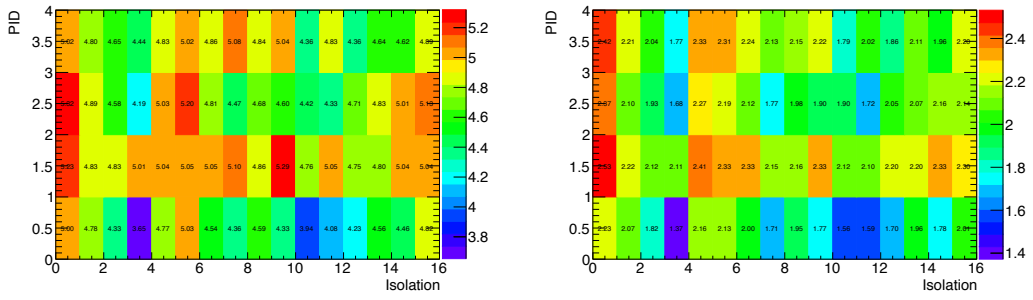
Table 5.5: Photon ID choices for the leading and subleading photons used to define the y-axes in Figure 5.2.

Bin	Leading photon	Subleading photon
0	`Loose`	`Loose`
1	`Tight`	`Tight`
2	`Tight`	`Loose`
3	`Loose`	`Tight`

These studies confirm the choice of `Tight` photon ID and `FixedCutLoose` photon isolation for both leading and subleading photons which corresponds to bin (0,1) in both histograms in Figure 5.2.

b-tagging optimization

The b-tagging working points are optimized simultaneously to the photon identification and isolation. A category is defined first where both jets must pass the same b-tagging working point (2-tag). For the events that fail this criteria another category is defined where only one jet must pass the defined b-tagging working point (1-tag). The Asimov significance is calculated using the working points of the MV2c10 tag-



(a) (b)

Figure 5.2: Asimov significance, for different photon ID and isolation choices, using the 70% and 60% b-tagging working point for the a) 2-tag category and b) 1-tag category, respectively, for resonant production at $m_X = 275$ GeV.

Table 5.6: Photon isolation choices for the leading and subleading photons used to define the x-axes in Figure 5.2.

Bin	Leading photon	Subleading photon
0	FixedCutLoose	FixedCutLoose
1	FixedCutTight	FixedCutTight
2	FixedCutTightCaloOnly	FixedCutTightCaloOnly
3	FixedCutLooseCaloOnly	FixedCutLooseCaloOnly
4	FixedCutLoose	FixedCutTight
5	FixedCutTight	FixedCutLoose
6	FixedCutLoose	FixedCutTightCaloOnly
7	FixedCutTight	FixedCutLooseCaloOnly
8	FixedCutTight	FixedCutTightCaloOnly
9	FixedCutLoose	FixedCutLooseCaloOnly
10	FixedCutLooseCaloOnly	FixedCutTightCaloOnly
11	FixedCutTightCaloOnly	FixedCutLooseCaloOnly
12	FixedCutLooseCaloOnly	FixedCutTight
13	FixedCutTightCaloOnly	FixedCutTight
14	FixedCutLooseCaloOnly	FixedCutLoose
15	FixedCutTightCaloOnly	FixedCutLoose

ger with an efficiency of 60%, 70%, 77% or 85%, separately for the 1-tag and 2-tag categories.

Figure 5.3 shows the optimizations for the 2-tag category with the four different b-tagging working points. It is observed that the highest significance is achieved with the 70% working point.

Figures 5.4, 5.5, 5.6 and 5.7 show the optimizations for the 1-tag category assuming that the 2-tag category uses the 60%, 70%, 77% or 85% working points, respectively. Note that the categories are exclusive (events can only fall into one of the categories) and that by definition only tighter b-tagging working points are allowed for the 1-tag category with respect to the working point being used in the 2-tag category.

The significance in the 1-tag category increases with a tighter b-tagging selection,

e.g. is maximum when using the 60% working point. A combined significance is calculated considering the two signal categories in this search. The result shows that the significance is highest when the 2-tag category uses the 70% working point and the 1-tag category uses the 60% working point.

Jet p_{T} and dijet mass optimization

After choosing the b-tagging working points for the 1- and 2-tag categories, the jet p_{T} and the dijet mass cuts are optimized simultaneously with events that contain two photons and two central jets. The NLO resonant signal samples and the $\gamma\gamma$+jets background MC are used for this study.

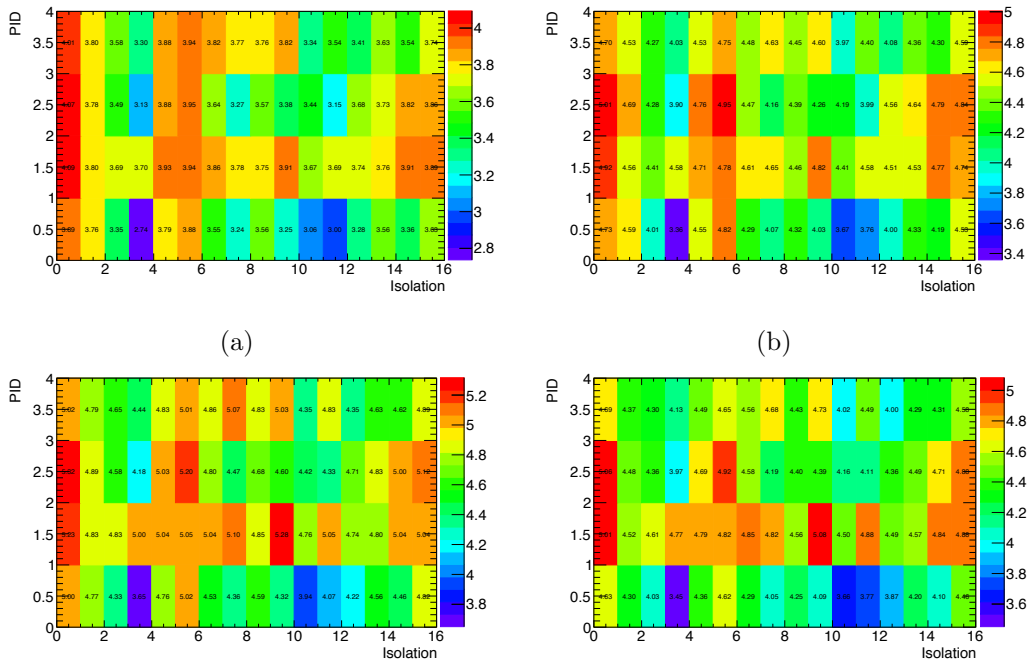


Figure 5.3: Asimov significance, for different photon ID and isolation choices, for the 2-tag category and using the a) 85%, b) 77%, c) 70% and d) 60% b-tagging working points for resonant production at $m_X = 275$ GeV.

Separate optimizations are performed for the 1- and 2-tag categories, and for the loose and tight selections. Figure 5.8 (5.9) shows the optimizations for the 2- and 1-tag categories with the tight (loose) selection.

The results show that the optimal loose selection corresponds to requiring a $p_{\mathrm{T}} >$ 40 (25) GeV for the highest- (next-highest)-p_{T} jet, respectively, as well as a dijet mass between 80 and 140 GeV. For the tight selection, the best significance is achieved by requiring a $p_{\mathrm{T}} >$ 100 (30) GeV for the highest- (next-highest)-p_{T} jet, respectively, with a dijet mass between 90 and 140 GeV.

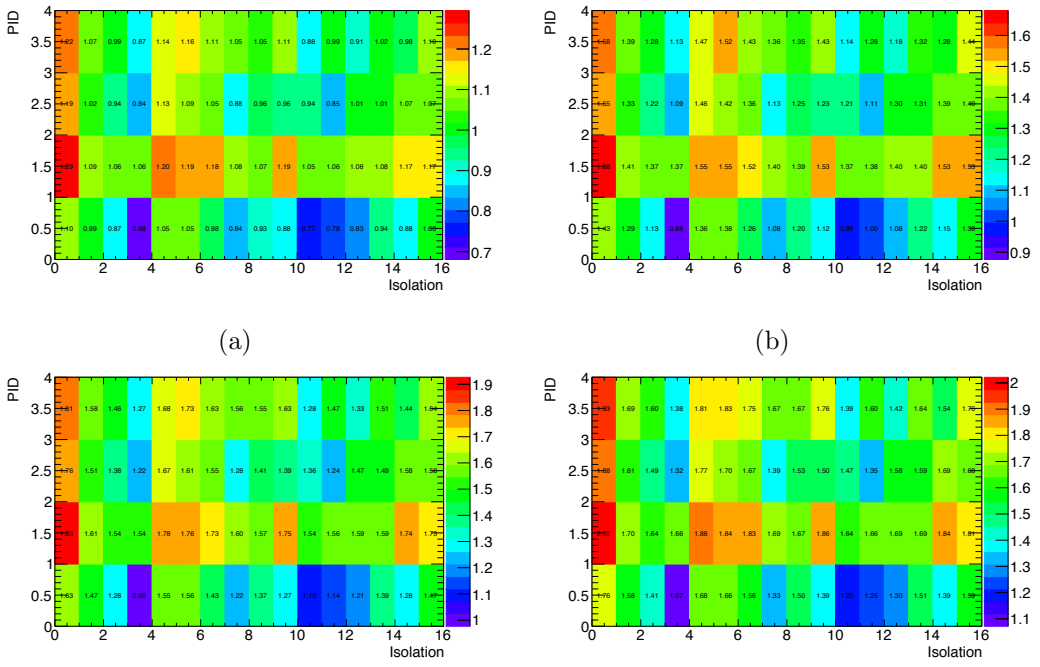


Figure 5.4: Asimov significance, for different photon ID and isolation choices, for the 1-tag category (assuming that the 2-tag category uses the 85% working point) and using the a) 85%, b) 77%, c) 70% and d) 60% b-tagging working points for resonant production at $m_X =$ 275 GeV.

Non-resonant analysis

For non-resonant Higgs boson pair production, the efficiency of events to satisfy the 2-tag selection criteria is 10% and 5.8% in the loose and tight selections, respectively. In the 1-tag category, the corresponding efficiencies are 7.2% and 3.9%, which are slightly lower than for the 2-tag category due to the lower probability of selecting the correct jet pair.

Due to the differing jet kinematics, the signal acceptance is lower in all cases for the generated NLO signal than for a LO signal. The acceptance of the LO prediction is approximately 15% higher when using the tight selection and 10% higher when

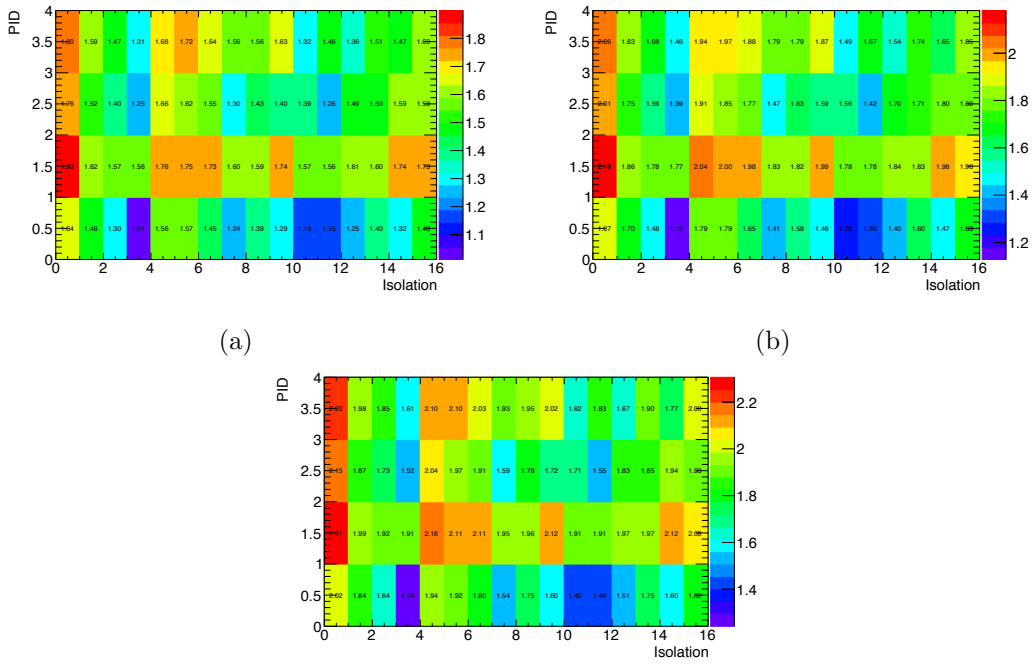
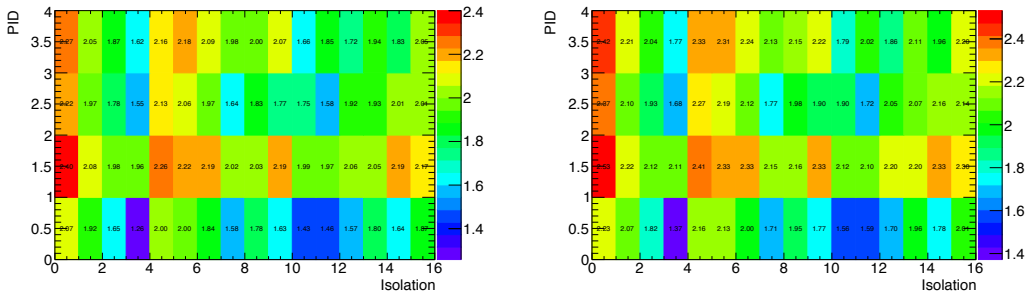


Figure 5.5: Asimov significance, for different photon ID and isolation choices, for the 1-tag category (assuming that the 2-tag category uses the 77% working point) and using the a) 77%, b) 70% and c) 60% b-tagging working points for resonant production at $m_X = 275$ GeV.

using the loose selection.

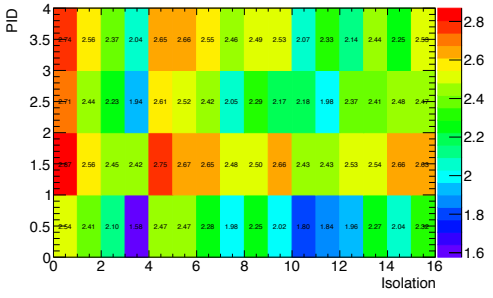
Resonant analysis

In the resonant analysis, before reconstructing the four-object mass, $m_{\gamma\gamma jj}$, the four-momentum of the dijet system is scaled by m_H/m_{jj}. As shown in Figure 5.10, this improves the four-object mass resolution by 60% on average across the resonance mass range of interest. It also modifies the shape of the non-resonant background in the



(a) (b)

Figure 5.6: Asimov significance, for different photon ID and isolation choices, for the 1-tag category (assuming that the 2-tag category uses the 70% working point) and using the a) 70% and b) 60% b-tagging working points for resonant production at $m_X = 275$ GeV.



(a)

Figure 5.7: Asimov significance, for different photon ID and isolation choices, for the 1-tag category (assuming that the 2-tag category uses the 60% working point) and using the 60% b-tagging working points for resonant production at $m_X = 275$ GeV.

region below 270 GeV.

For the resonant analysis, efficiencies range from 6% to 15.4% in the 2-tag category and from 5.1% to 12.3% in the 1-tag category for 260 GeV $< m_X <$ 1000 GeV.

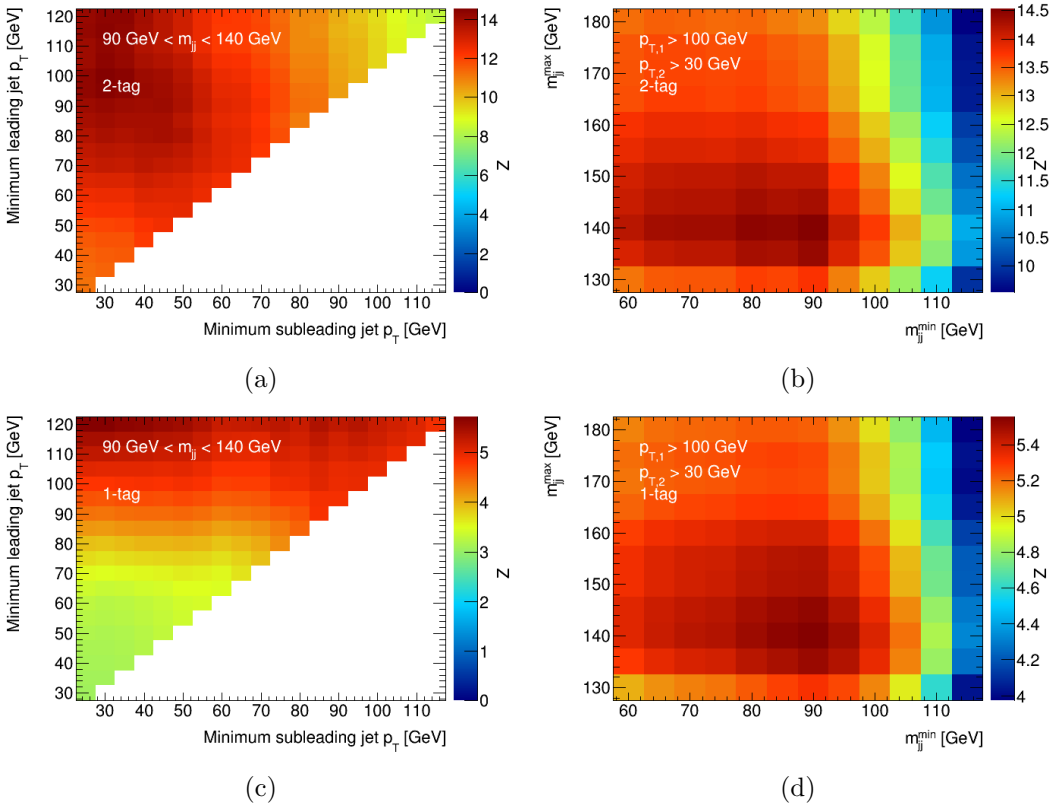


Figure 5.8: Asimov significance is shown for varied jet p_{T} cuts for the a) 2-tag and c) 1-tag categories, and for diverse dijet mass cuts for the b) 2-tag and d) 1-tag categories, using the tight selection.

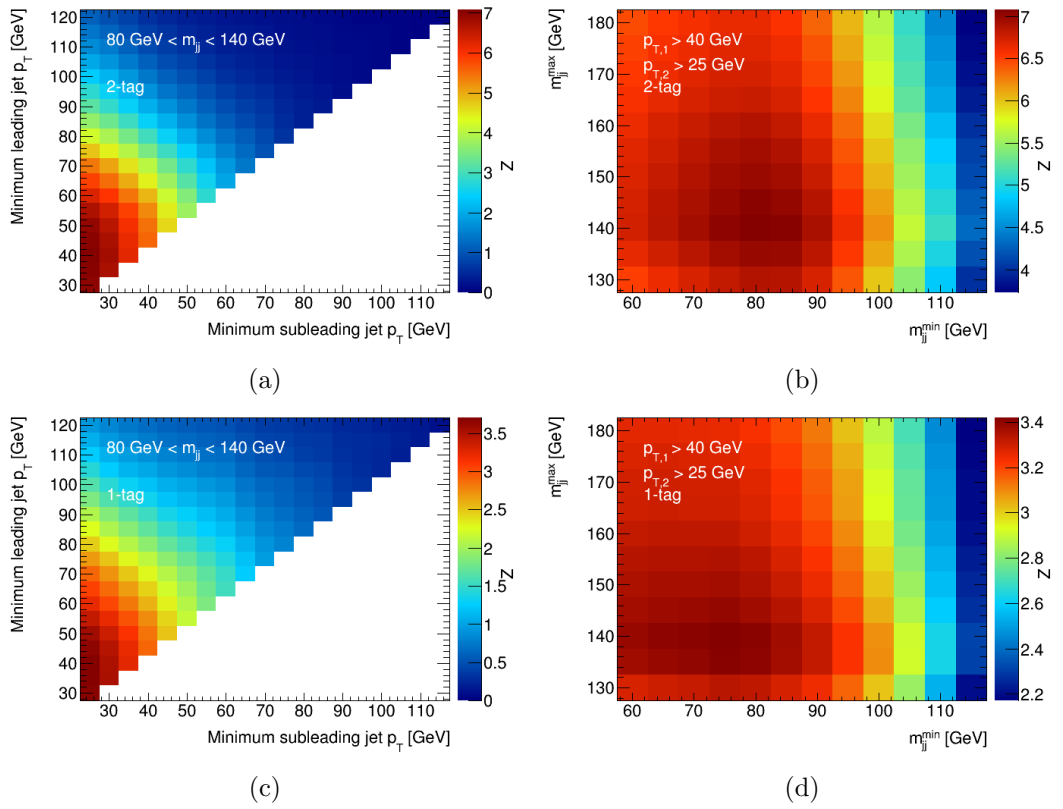
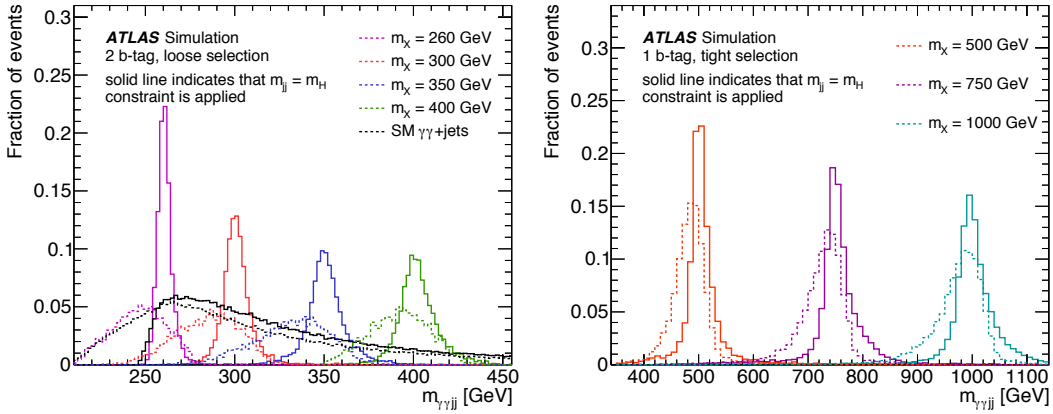


Figure 5.9: Asimov significance is shown for varied jet p_{T} cuts for the a) 2-tag and c) 1-tag categories, and for diverse dijet mass cuts for the b) 2-tag and d) 1-tag categories, using the loose selection.



(a) (b)

Figure 5.10: Reconstructed $m_{\gamma\gamma jj}$ with (solid lines) and without (dashed lines) the dijet mass constraint, for a subset of the mass points used in the resonant analysis. The examples shown here are for a) the 2-tag category with the loose selection and b) the 1-tag category with the tight selection. The effect on the continuum background is also shown in a) [3].

5.4 Signal and background modeling

Both the resonant and non-resonant searches for Higgs boson pairs proceed by performing unbinned maximum-likelihood fits to the data in the 1-tag and 2-tag signal categories simultaneously. The non-resonant search involves a fit to the $m_{\gamma\gamma}$ distribution, while the search for resonant production uses the $m_{\gamma\gamma jj}$ distribution. The signal-plus-background fit to the data uses parametrized forms for both the signal and background probability distributions. These parametrized forms are determined through fits to simulated samples.

As the loose selection is used for resonances with $m_X \leq 500$ GeV and the tight selection for resonances with $m_X \geq 500$ GeV, different ranges of $m_{\gamma\gamma jj}$ are used in each case. For the loose (tight) selection, only events with $m_{\gamma\gamma jj}$ in the range 245 GeV $< m_{\gamma\gamma jj} <$ 610 GeV (335 GeV $< m_{\gamma\gamma jj} <$ 1140 GeV) are considered. These ranges are the smallest ones containing over 95% of all of the simulated signal sample events with m_X below, or above, 500 GeV respectively.

5.4.1 Background composition

Contributions to the continuum diphoton background originate from $\gamma\gamma$, γj, $j\gamma$ and jj sources produced in association with jets, where j denotes jets misidentified as photons and γj and $j\gamma$ differ by the jet faking the subleading or the leading photon candidate respectively. Their relative fractions are determined from data using a double two-dimensional sideband method (2x2D) [149, 150]. In the 2x2D sideband method, a loose' identification criteria is defined, in which some of the cuts required for tight photon identification are loosened. Both photons in the data are required to pass this loose' criteria with no requirement made on the isolation. The total observed yield, $W_{tot}^{L'L'}$ is the sum of the diphoton signal yield $W_{\gamma\gamma}^{L'L'}$ and the background from

misidentified photons ($W_{\gamma j}^{L'L'}$, $W_{j\gamma}^{L'L'}$, $W_{jj}^{L'L'}$):

$$W_{tot}^{L'L'} = W_{\gamma\gamma}^{L'L'} + W_{\gamma j}^{L'L'} + W_{j\gamma}^{L'L'} + W_{jj}^{L'L'} \tag{5.2}$$

Each leading (subleading) photon in the sample can be divided into a region, A(A'), B(B'), C(C') and D(D') depending on whether they satisfy the tight identification criteria and/or the isolation requirement as shown in Figure 5.11. Each leading region can be combined with a subleading region to give 16 combinations that a pair of photons could belong to. The yield for each combination is linked to the terms in equation 5.2 by means of efficiencies and fake rates. In total, 16 simultaneous equations can be constructed which can be solved to give the efficiencies and fake rates and from this the relative fractions of $\gamma\gamma$, γj, $j\gamma$, jj can be predicted in data for each category in the analysis. In the case of the 1- and 2-tag categories the contribution from $\gamma\gamma$ events is in the range 80-90%.

The choice of functional form used to fit the background in the final likelihood models is derived using simulated events. Diphoton continuum events were simulated using the Sherpa event generator as described in Section 5.2. As this prediction from Sherpa does not provide a good description of the $m_{\gamma\gamma}$ spectrum in data, the mismodeling is corrected for using a data-driven reweighting function.

In the 0-tag control category, the number of events in data is high enough that the 2x2D method can be applied in bins of $m_{\gamma\gamma}$. The events generated by Sherpa can also be divided into $\gamma\gamma$, γj, $j\gamma$, and jj sources based on the same photon identification and isolation criteria as used in data. For each of these sources, the $m_{\gamma\gamma}$ distributions for both Sherpa and the data are fit using an exponential function and the ratio of the two fit results is taken as an $m_{\gamma\gamma}$-dependent correction function. The size of the correction is less than 5% for the majority of events.

These reweighting functions are then applied in the 1-tag and 2-tag signal cat-

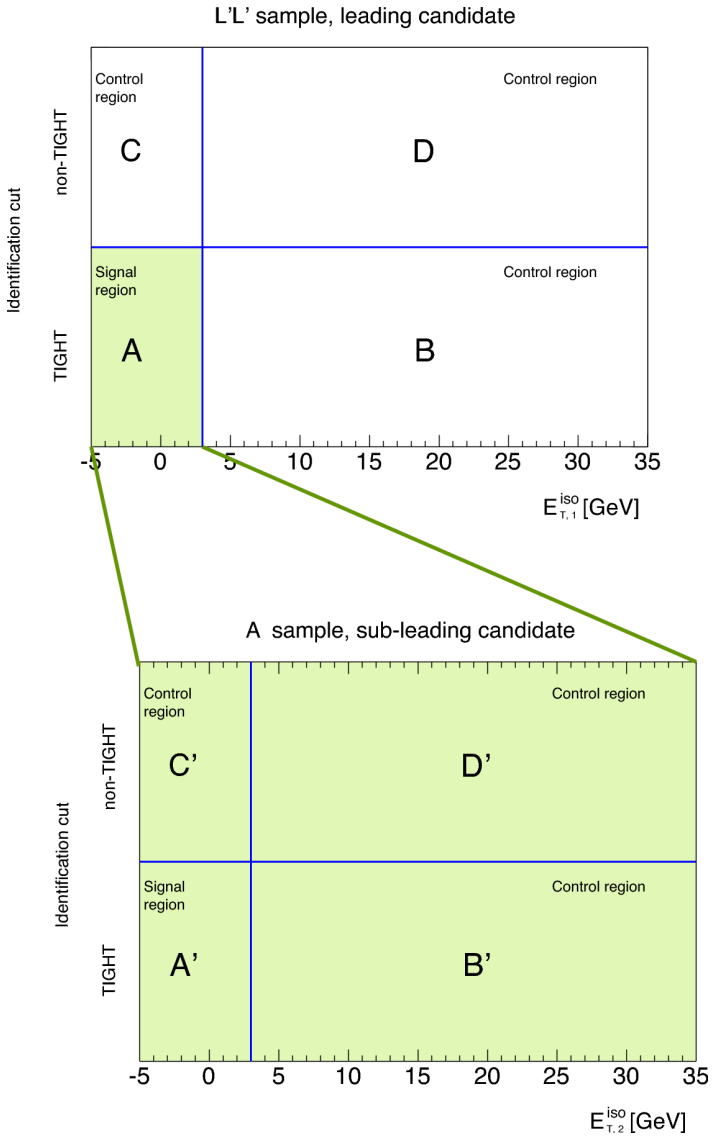


Figure 5.11: Schematic representation of the 2x2D sideband method. The top (bottom) plane displays the isolation on the x-axis and the identification criteria on the y-axis for the leading (subleading) photon.

egories to correct the shape of the Sherpa prediction. The fractional contribution from the different continuum background sources is fixed to the relative proportions derived in data with the 2x2D method. Finally, the overall normalization is chosen such that, in the disjoint sideband region 105 GeV $< m_{\gamma\gamma} <$ 120 GeV and 130 GeV $< m_{\gamma\gamma} <$ 160 GeV, the total contribution from all backgrounds is equal to that from data.

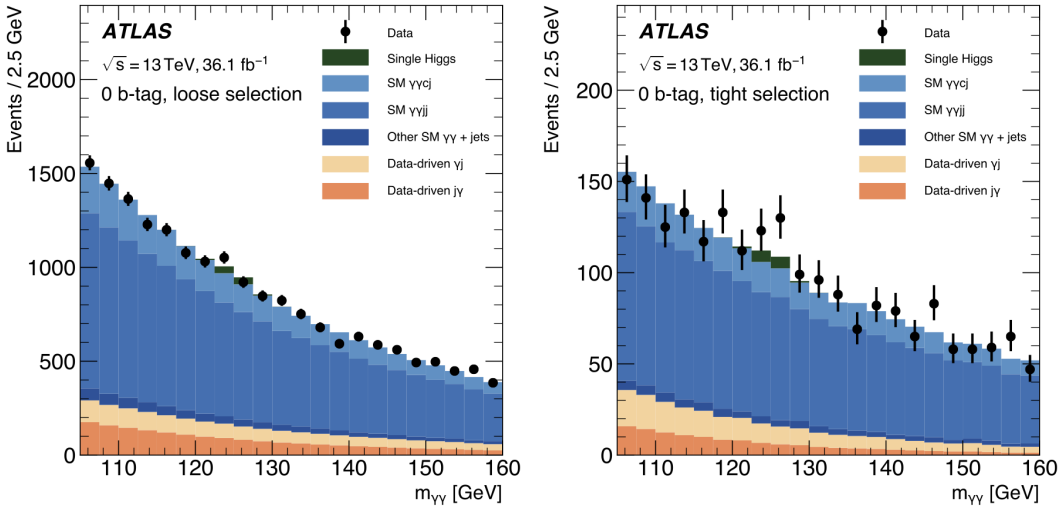
The contribution from $\gamma\gamma$ produced in association with jets is further divided according to the flavors of the two jets (for example bb, bc, c + light jet). This decomposition is taken directly from the proportions predicted by the Sherpa event generator and no attempt is made to classify the data according to jet flavor. The continuum background in the 1-tag category comes primarily from $\gamma\gamma bj$ events (60%) and in the 2-tag category from $\gamma\gamma bb$ events (80%). A comparison between data in the 0-tag control category and this data-driven prediction of the total background can be seen in Figures 5.12a and 5.12b for the $m_{\gamma\gamma}$ distribution with the loose and tight selections, and in Figures 5.12c and 5.12d for the $m_{\gamma\gamma jj}$ distribution with the loose and tight selection, respectively.

5.4.2 Non-resonant analysis

Signal modeling

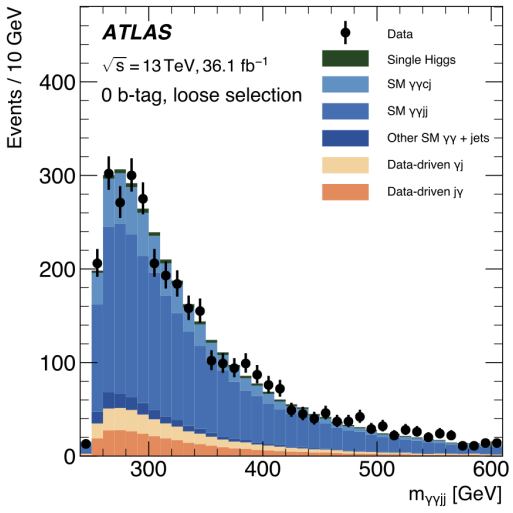
The shape of the diphoton mass distribution in $HH \to \gamma\gamma b\bar{b}$ events is described by the double-sided Crystal Ball function [151], consisting of a Gaussian core with power-law tails on either side. The parameters of this model are determined through fits to the simulated non-resonant SM HH sample described in Section 5.2. The effect of the reweighting to the full top-quark mass prediction on the signal model is negligible, with the resolution of the double-sided Crystal Ball varying by 1%.

The parametric form for a given category i for a Higgs boson mass m_H, can be

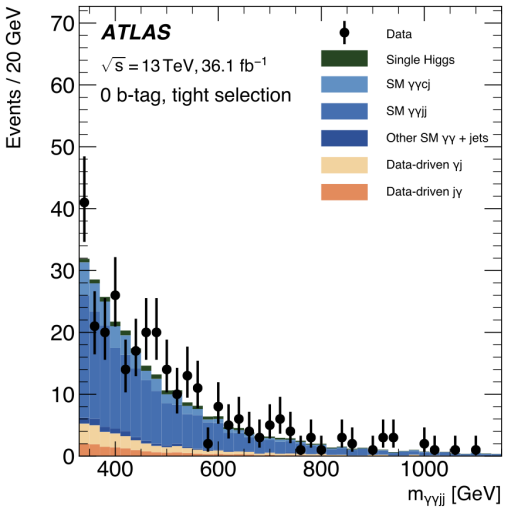


(a)

(b)



(c)



(d)

Figure 5.12: The predicted number of background events from continuum diphoton plus jet production (blue), other continuum photon and jet production (red and orange) and single Higgs boson production (green) is compared with the observed data (black points) in the 0-tag control category for the $m_{\gamma\gamma}$ distribution with the a) loose selection, b) tight selection, and for the $m_{\gamma\gamma jj}$ distribution with the c) loose selection and d) tight selection [3].

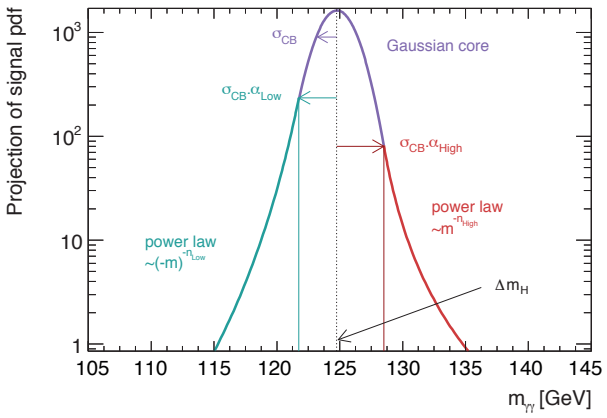


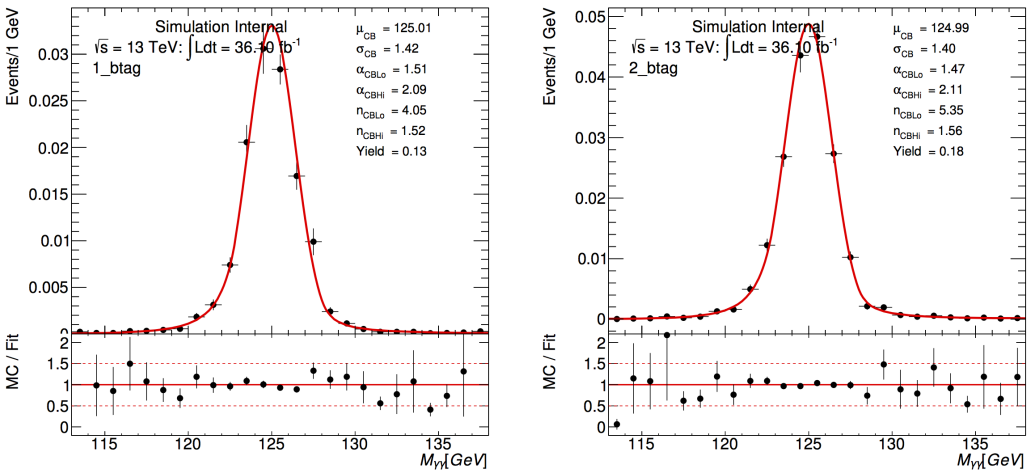
Figure 5.13: A visualization of the double-sided Crystal Ball function [152].

written as:

$$f_i^{\text{sig}}(m_{\gamma\gamma};\Delta\mu_{\text{CB},i},\sigma_{\text{CB},i},\alpha^{\pm}_{\text{CB},i},n^{\pm}_{\text{CB},i}) = \mathcal{N}_c \begin{cases} e^{-t^2/2} & -\alpha^{-}_{\text{CB},i} \leq t \leq \alpha^{+}_{\text{CB},i} \\ \dfrac{\left(\frac{n^{-}_{\text{CB},i}}{|\alpha^{-}_{\text{CB},i}|}\right)^{n^{-}_{\text{CB},i}} e^{-|\alpha^{-}_{\text{CB},i}|^2/2}}{\left(\frac{n^{-}_{\text{CB},i}}{\alpha^{-}_{\text{CB},i}} - \alpha^{-}_{\text{CB},i} - t\right)^{n^{-}_{\text{CB},i}}} & t < -\alpha^{-}_{\text{CB},i} \\ \dfrac{\left(\frac{n^{+}_{\text{CB},i}}{|\alpha^{+}_{\text{CB},i}|}\right)^{n^{+}_{\text{CB},i}} e^{-|\alpha^{+}_{\text{CB},i}|^2/2}}{\left(\frac{n^{+}_{\text{CB},i}}{\alpha^{+}_{\text{CB},i}} - \alpha^{+}_{\text{CB},i} - t\right)^{n^{+}_{\text{CB},i}}} & t > \alpha^{+}_{\text{CB},i} \end{cases}, \tag{5.3}$$

where $t = (m_{\gamma\gamma} - m_H - \Delta\mu_{\text{CB},i})/\sigma_{\text{CB},i}$, and $\mathcal{N}_c$ is a normalization factor. The non-Gaussian parts are parametrized by $\alpha^{\pm}_{\text{CB},i}$ and $n^{\pm}_{\text{CB},i}$ separately for the low- $(-)$ and high-mass $(+)$ tails.

The parameters of this model are determined through fits to the simulated non-resonant SM HH sample described in Section 5.2. Figure 5.14 shows the fits to the $m_{\gamma\gamma}$ distribution for the 1-tag and 2-tag categories with the tight selection.



(a) (b)

Figure 5.14: Parametrization with a double-sided Crystal Ball function of the $m_{\gamma\gamma}$ distribution from the non-resonant signal SM HH sample in the a) 1-tag and b) 2-tag categories. Events are required to pass the tight selection.

Background modeling

For the non-resonant analysis, the continuum $m_{\gamma\gamma}$ background is modeled using a functional form obtained from a fit to the data. The potential bias arising from this procedure, called "spurious signal", is estimated by performing signal-plus-background fits to the combined continuum background from simulation, including the $\gamma\gamma$, γj, $j\gamma$ and jj components. The maximum absolute value of the extracted signal, for a signal in the range 121 GeV $< m_{\gamma\gamma} <$ 129 GeV, is taken as the bias. This method is used to discriminate between different potential fit functions - the function chosen is the one with the smallest spurious signal bias. If multiple functions have the same bias, the one with the smallest number of parameters is chosen.

The functions that have been considered are the following:

- Exponential, including exponential of 2^{nd} and 3^{rd} order polynomial. These

Table 5.7: Number of spurious signal events, N_{spur}, and its ratio to the statistical uncertainty on the fitted number of signal events, Z_{spur}, for the 2-tag category with the loose selection. The χ^2/ndof obtained from performing a background-only fit is also shown. The number of parameters for each function is given by $nPars$.

Model	Z_{spur} [%]	N_{spur}	$nPars$	χ^2/ndof
Exponential	-23.7	-0.894	1	1.65
Exp. of 2$^{\text{nd}}$ order	24.6	1.13	2	1.67
Exp. of 3$^{\text{rd}}$ order	17.4	0.85	3	1.70
Dijet	21.9	0.949	2	1.67
Bernstein of order 3	-18.6	0.833	3	1.69
Bernstein of order 4	-20.1	0.898	4	1.71
Bernstein of order 5	30.2	1.6	5	1.66

functions take the form:
$Exp^N(m_{\gamma\gamma};\theta^{bkg}) = exp\left(\sum_{j=0}^{N}\theta_j^{bkg}\cdot m_{\gamma\gamma}^j\right)$

- Bernstein polynomials of order N. These functions take the form:
  $Bernstein^N(m_{\gamma\gamma};\theta^{bkg}) = \sum_{j=0}^{N}\theta_j^{bkg}b_{j,N}$ with $b_{j,N} = C_n^j x^j(1-x)^{N-j}$, where $x = (m_{\gamma\gamma}\ [\text{GeV}] - 100)/60$.
- Dijet function: $Dijet(m_{\gamma\gamma};\theta^{bkg}) = \theta_1^{bkg}(1-m_{\gamma\gamma})^{\theta_2^{bkg}} m_{\gamma\gamma}^{\theta_3^{bkg}}$

Tables 5.7–5.10 show the number of expected spurious signal events as well as its ratio to the statistical uncertainty on the fitted number of signal events for the 2-tag and 1-tag categories for the loose and tight selections.

Figure 5.15 shows the number of spurious signal events as a function of $m_{\gamma\gamma}$ for the loose and tight selections in the 2-tag and 1-tag categories. The first-order exponential function has the smallest bias among the seven functions considered and is therefore chosen. The background from single Higgs boson production is described using a double-sided Crystal Ball function, with its parameters determined through fits to the appropriate simulated samples.

Table 5.8: Number of spurious signal events, N_{spur}, and its ratio to the statistical uncertainty on the fitted number of signal events, Z_{spur}, for the 2-tag category with the tight selection. The χ^2/ndof obtained from performing a background-only fit is also shown. The number of parameters for each function is given by $nPars$.

Model	Z_{spur} [%]	N_{spur}	$nPars$	χ^2/ndof
Exponential	-34.0	-0.626	1	2.15
Exp. of 2$^{\text{nd}}$ order	-25.1	-0.488	2	1.80
Exp. of 3$^{\text{rd}}$ order	-24.3	-0.507	3	2.00
Dijet	-31.2	-0.574	2	2.18
Bernstein of order 3	-21.7	-0.460	3	1.88
Bernstein of order 4	-27.2	-0.593	4	2.29
Bernstein of order 5	-23.8	-0.534	5	2.63

Table 5.9: Number of spurious signal events, N_{spur}, and its ratio to the statistical uncertainty on the fitted number of signal events, Z_{spur}, for the 1-tag category with the loose selection. The χ^2/ndof obtained from performing a background-only fit is also shown. The number of parameters for each function is given by $nPars$.

Model	Z_{spur} [%]	N_{spur}	$nPars$	χ^2/ndof
Exponential	-29	-2.91	1	1.18
Exp. of 2$^{\text{nd}}$ order	-31.8	-3.36	2	1.21
Exp. of 3$^{\text{rd}}$ order	-32.1	-3.68	3	1.23
Dijet	13.8	1.33	2	1.48
Bernstein of order 3	-28	-3.02	3	1.20
Bernstein of order 4	-33.8	-3.94	4	1.22
Bernstein of order 5	-25.3	-3.05	5	1.10

Table 5.10: Number of spurious signal events, N_{spur}, and its ratio to the statistical uncertainty on the fitted number of signal events, Z_{spur}, for the 1-tag category with the tight selection. The χ^2/ndof obtained from performing a background-only fit is also shown. The number of parameters for each function is given by $nPars$.

Model	Z_{spur} [%]	N_{spur}	$nPars$	χ^2/ndof
Exponential	-6.68	-0.248	1	0.917
Exp. of 2$^{\text{nd}}$ order	-8.86	-0.355	2	0.939
Exp. of 3$^{\text{rd}}$ order	-7.79	-0.327	3	0.958
Dijet	6.61	0.251	2	0.959
Bernstein of order 3	-6.35	-0.269	3	0.956
Bernstein of order 4	6.54	0.272	4	0.951
Bernstein of order 5	11.5	0.525	5	0.965

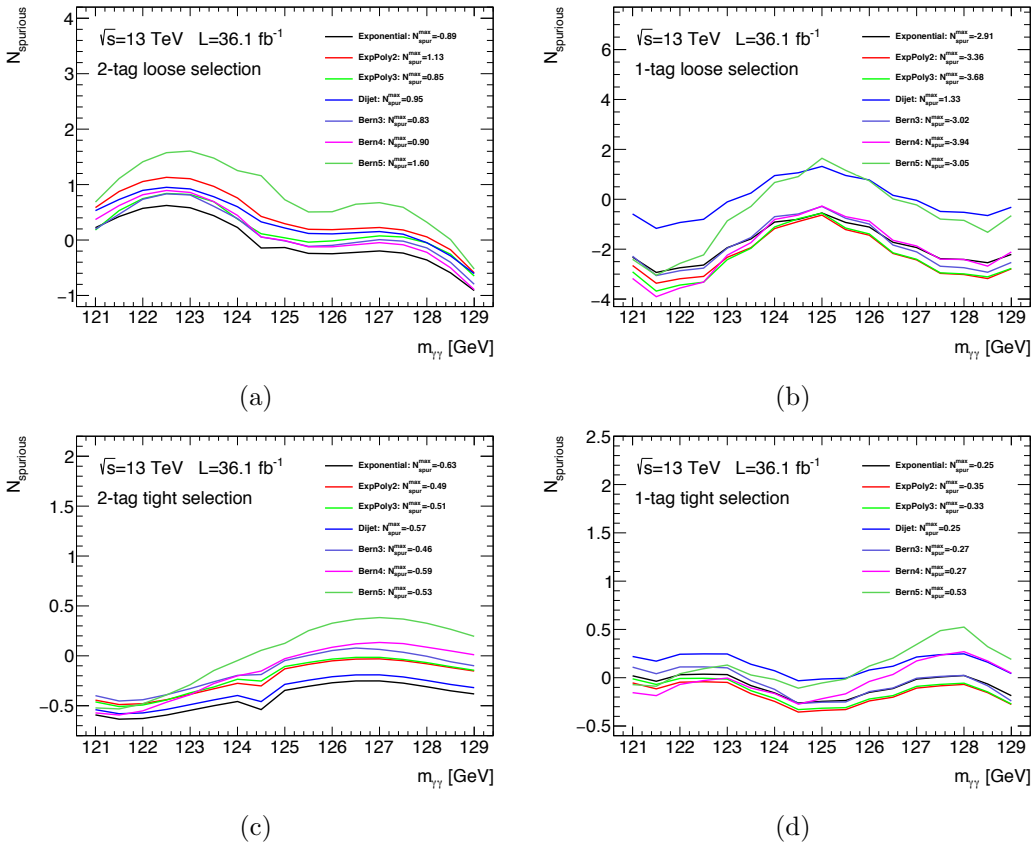


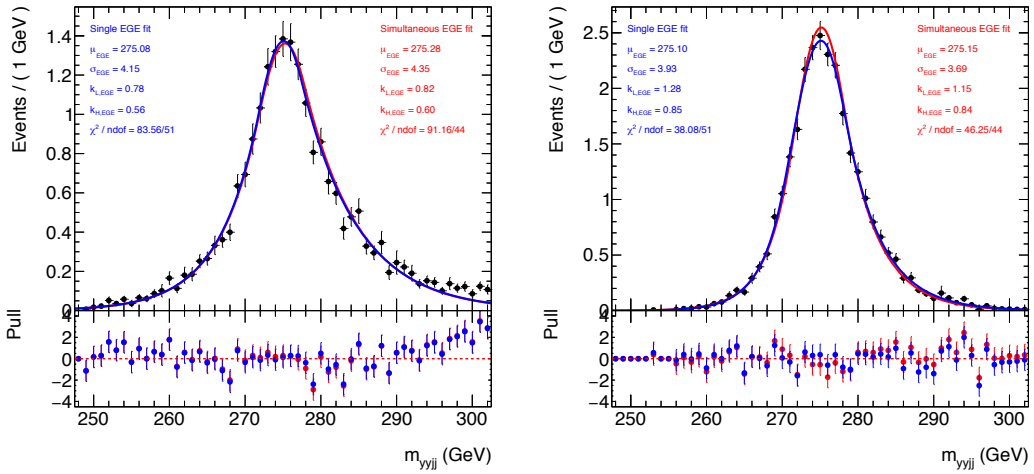
Figure 5.15: Number of spurious signal events for different fit functions for the non-resonant case and for the a) 2-tag and b) 1-tag categories with the loose selection. Similarly, the tight selection is used in c) for the 2-tag and d) 1-tag categories.

5.4.3 Resonant analysis

Signal modeling

For each resonant hypothesis, a fit is performed to the $m_{\gamma\gamma jj}$ distribution of the simulated events in a window around the nominal m_X. The shape of this distribution is described using a function consisting of a Gaussian core with exponential tails on either side, ExpGaussExp (EGE) [153]. A simultaneous fit to all signal samples is carried out in which each of the model parameters is further parametrized in terms

of m_X. This allows the model to provide a prediction for any mass satisfying 260 GeV $< m_X <$ 1000 GeV, where these boundaries reflect the smallest and largest m_X values among the generated samples described in Section 5.2. A single mass point fit, providing concrete values of the model parameters of the EGE, is performed for validation of the signal parametrization obtained from the simultaneous fit. Figures 5.16 and 5.17 show a loose and a tight selection example, respectively, of the simultaneous and single fits to the $m_{\gamma\gamma jj}$ distribution.

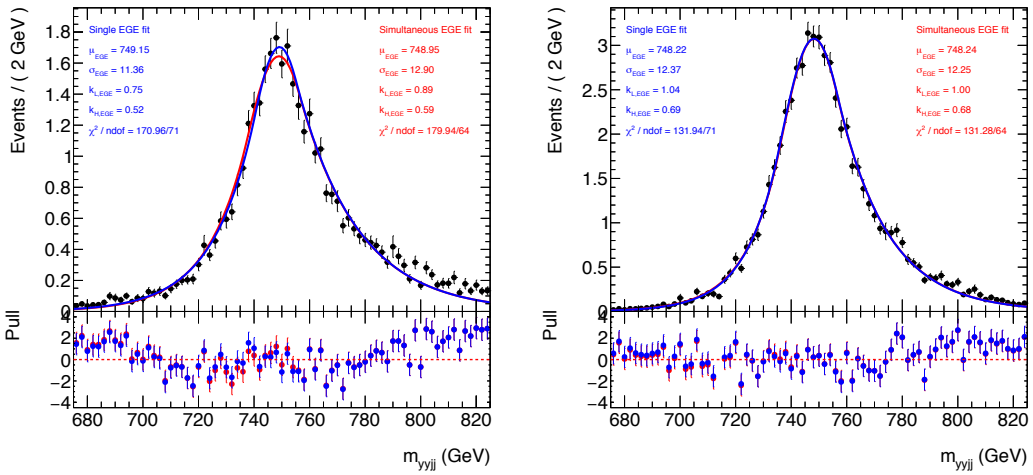


(a) (b)

Figure 5.16: Parametrization with an ExpGaussExp function of the $m_{\gamma\gamma jj}$ distribution from the resonant signal HH sample at $m_X = 275$ GeV in the a) 1-tag and b) 2-tag categories. Events are required to pass the loose selection.

Background modeling

For the resonant analysis, a spurious signal study is also carried out, performing a signal-plus-background fit to the combined continuum background from simulation using the $m_{\gamma\gamma jj}$ distribution for events within the $m_{\gamma\gamma}$ window described in Section 5.3.3. The background from single Higgs boson production is included as part of



(a) (b)

Figure 5.17: Parametrization with an ExpGaussExp function of the $m_{\gamma\gamma jj}$ distribution from the resonant signal HH sample at $m_X = 750$ GeV in the a) 1-tag and b) 2-tag categories. Events are required to pass the tight selection.

this continuum distribution.

Due to the different $m_{\gamma\gamma jj}$ ranges used with the loose and tight selections, the shape of the $m_{\gamma\gamma jj}$ distribution differs between these two cases and hence different background functions are considered.

The three different background functions considered for the loose selection are the following:

- Novosibirsk: $P(x) = e^{-0.5^{(lnq_y)^2/\Lambda^2+\Lambda^2}}$, $q_y = 1 + \Lambda(x - x_0)/\sigma \times \frac{sinh\Lambda\sqrt{ln4}}{\Lambda\sqrt{ln4}}$ [154].
- Modified Gamma: a Gamma distribution where both the shape and scale parameters are replaced by linear functions of the mass (adding two additional degrees of freedom).
- Modified Landau: a Landau distribution where the scale parameter is replaced by linear functions of the mass (adding one additional degree of freedom).

Four different background functions, each with only one free parameter, are con-

Table 5.11: Number of spurious signal events, N_{spur}, and its ratio to the statistical uncertainty on the fitted number of signal events, Z_{spur}, for the 2-tag category with the loose selection. The χ^2/ndof obtained from performing a background-only fit is also shown. The number of parameters for each function is given by $nPars$.

Model	Z_{spur} [%]	N_{spur}	$nPars$	χ^2/ndof
Novosibirsk	56.97	0.58	3	33.9 / 13
Modified Gamma	82.02	0.77	5	36.3 / 11
Modified Landau	65.62	1.09	3	45.8 / 13

Table 5.12: Number of spurious signal events, N_{spur}, and its ratio to the statistical uncertainty on the fitted number of signal events, Z_{spur}, for the 1-tag category with the loose selection. The χ^2/ndof obtained from performing a background-only fit is also shown. The number of parameters for each function is given by $nPars$.

Model	Z_{spur} [%]	N_{spur}	$nPars$	χ^2/ndof
Novosibirsk	33.64	2.06	3	36.7 / 22
Modified Gamma	53.67	3.42	5	45.2 / 20
Modified Landau	96.93	7.54	3	128.3 / 24

sidered for the tight selection:

- Exponential (x): $Exp.[x](m_{\gamma\gamma bb}; p_0) = exp\left(p_0 \cdot m_{\gamma\gamma bb}\right)$
- Exponential (x^2): $Exp.[x^2](m_{\gamma\gamma bb}; p_0) = exp\left(p_0 \cdot m^2_{\gamma\gamma bb}\right)$
- Inverse polynomial (x^{-2}): $Inv.poly[x^{-2}](m_{\gamma\gamma bb}; p_0) = 1 + p_0 / \left(m^2_{\gamma\gamma bb}\right)$
- Inverse polynomial (x^{-3}): $Inv.poly[x^{-3}](m_{\gamma\gamma bb}; p_0) = 1 + p_0 / \left(m^3_{\gamma\gamma bb}\right)$
- Power law: $Powerlaw(m_{\gamma\gamma bb}; p_0) = m^{p_0}_{\gamma\gamma bb}$

Figure 5.18 shows the background-only fits of the $m_{\gamma\gamma jj}$ distribution for the loose and tight selections in the 2-tag and 1-tag categories. Tables 5.11–5.14 show the number of expected spurious signal events as well as its ratio to the statistical uncertainty on the fitted number of signal events for the 2-tag and 1-tag categories for the loose and tight selections.

For the loose (tight) selection, the Novosibirsk function (exponential function) has the smallest bias among the three (four) functions considered and is therefore chosen.

Table 5.13: Number of spurious signal events, N_{spur}, and its ratio to the statistical uncertainty on the fitted number of signal events, Z_{spur}, for the 2-tag category with the tight selection. The χ^2/ndof obtained from performing a background-only fit is also shown. The number of parameters for each function is given by $nPars$.

Model	Z_{spur} [%]	N_{spur}	$nPars$	χ^2/ndof
Exponential	98.94	0.21	1	36.4 / 15
Exp. of 2^{nd} order	123.18	0.27	1	59.0 / 15
Inv. polynomial of 2^{nd} order	88.33	0.25	1	42.9 / 15
Inv. polynomial of 3^{rd} order	116.84	0.13	1	22.2 / 15
Powerlaw	72.46	0.17	1	25.9 / 15

Table 5.14: Number of spurious signal events, N_{spur}, and its ratio to the statistical uncertainty on the fitted number of signal events, Z_{spur}, for the 1-tag category with the tight selection. The χ^2/ndof obtained from performing a background-only fit is also shown. The number of parameters for each function is given by $nPars$.

Model	Z_{spur} [%]	N_{spur}	$nPars$	χ^2/ndof
Exponential	63.66	0.89	1	32.5 / 26
Exp. of 2^{nd} order	73.64	0.97	1	53.9 / 26
Inv. polynomial of 2^{nd} order	355.52	1.51	1	354.7 / 26
Inv. polynomial of 3^{rd} order	146.35	1.03	1	68.8 / 26
Powerlaw	78.00	0.96	1	45.0 / 26

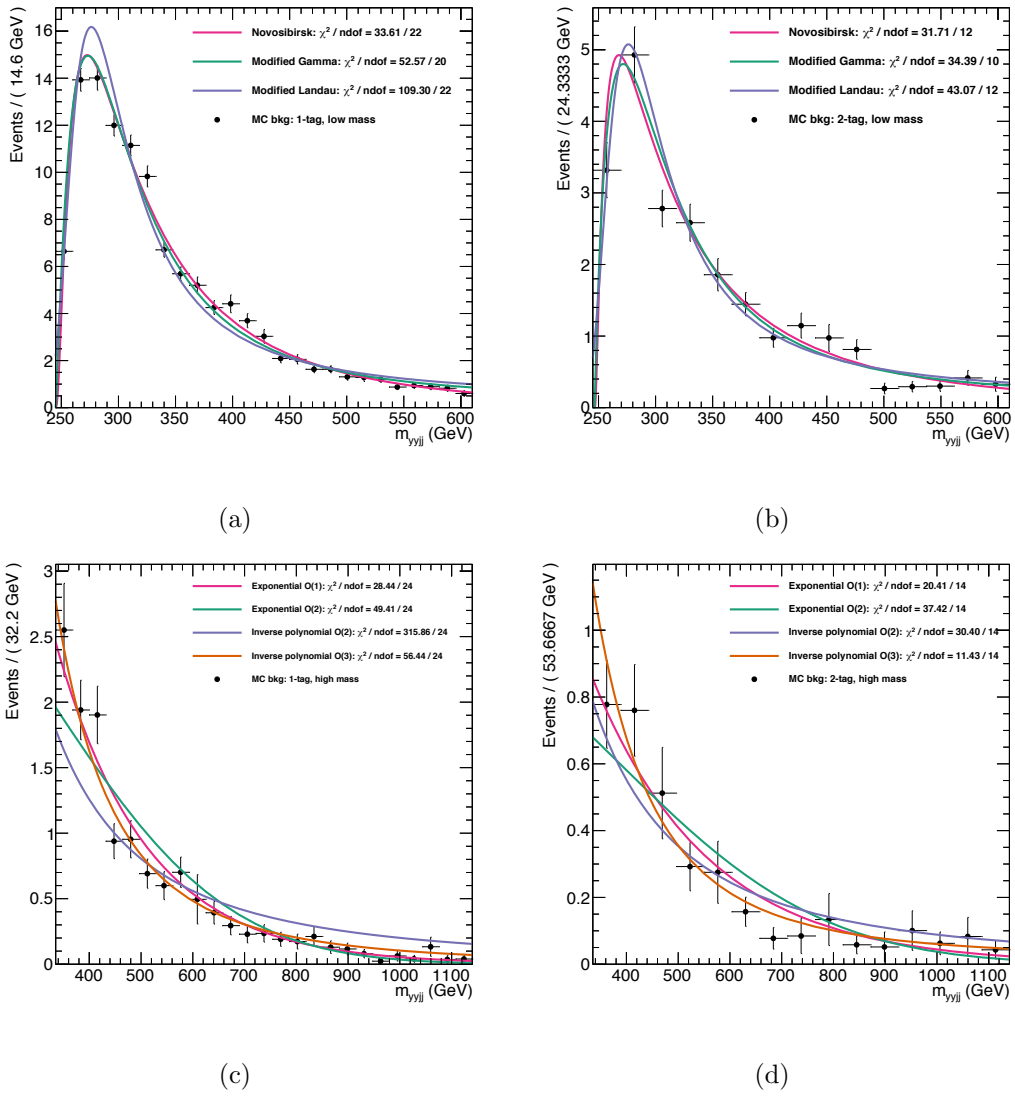


Figure 5.18: Background-only fit to the $m_{\gamma\gamma jj}$ distribution for the a) 1-tag and b) 2-tag categories with the loose selection and the c) 1-tag and d) 2-tag categories with the tight selection.

As a result, for low-mass resonances both the signal and background fit functions have a characteristic peaked shape. This degeneracy could potentially introduce a bias in the extracted signal cross section. In order to stabilize the background fit,

nominal values of the shape parameters are estimated by fitting to the simulated events described in Section 5.4.1. The shape is then allowed to vary in the likelihood to within the statistical covariance of this template fit. Experimental systematics on the background shape have a small effect and are neglected. The normalization of the background is estimated by interpolating the $m_{\gamma\gamma}$ sideband data. Additionally, a simple bias test is performed by drawing pseudo-data sets from the overall probability distribution created by combining the Novosibirsk background function with the signal function. For each mass point and each value of the injected signal cross section, fits are performed on the ensemble of pseudo-data sets and the median extracted signal cross section is recorded. For resonances with masses below 400 GeV, a small correction is applied to remove the observed bias. The correction is less than ± 0.05 pb everywhere and a corresponding uncertainty of ± 0.02 pb in this correction is applied to the extracted signal cross section. The corresponding uncertainty in the number of events in each category is roughly half that of the spurious signal.

Cut-and-count method

An alternative method based on a counting technique is used to estimate the signal and background events in the search for resonant Higgs boson pair production, motivated by the lack of statistics in this final state. The signal region (SR) in the 2-tag and 1-tag categories are defined as events that are within the tight $m_{\gamma\gamma}$ window defined in Section 5.3.3 and fall inside the smallest possible $m_{\gamma\gamma jj}$ window containing 95% of signal events. Figure 5.19 shows the parametrization of the $m_{\gamma\gamma jj}$ windows as a function of m_X for the 2-tag and 1-tag categories with the loose and tight selections. The number of background events is extrapolated using the following formula:

$$N_{SR}^{B} = N_{SB} \times \frac{N_{SR}^{MC}}{N_{SB}^{MC}} \times \epsilon_{m_{\gamma\gamma jj}}^{B}, \tag{5.4}$$

where N_{SB} is the number of observed events in the $m_{\gamma\gamma}$ sidebands, N_{SR}^{MC} (N_{SB}^{MC}) is the number of events predicted by the simulated events described in Section 5.4.1 in the $m_{\gamma\gamma}$ signal region (sidebands), and $\epsilon^{B}_{m_{\gamma\gamma jj}}$ is the efficiency to pass the $m_{\gamma\gamma jj}$ window cuts.

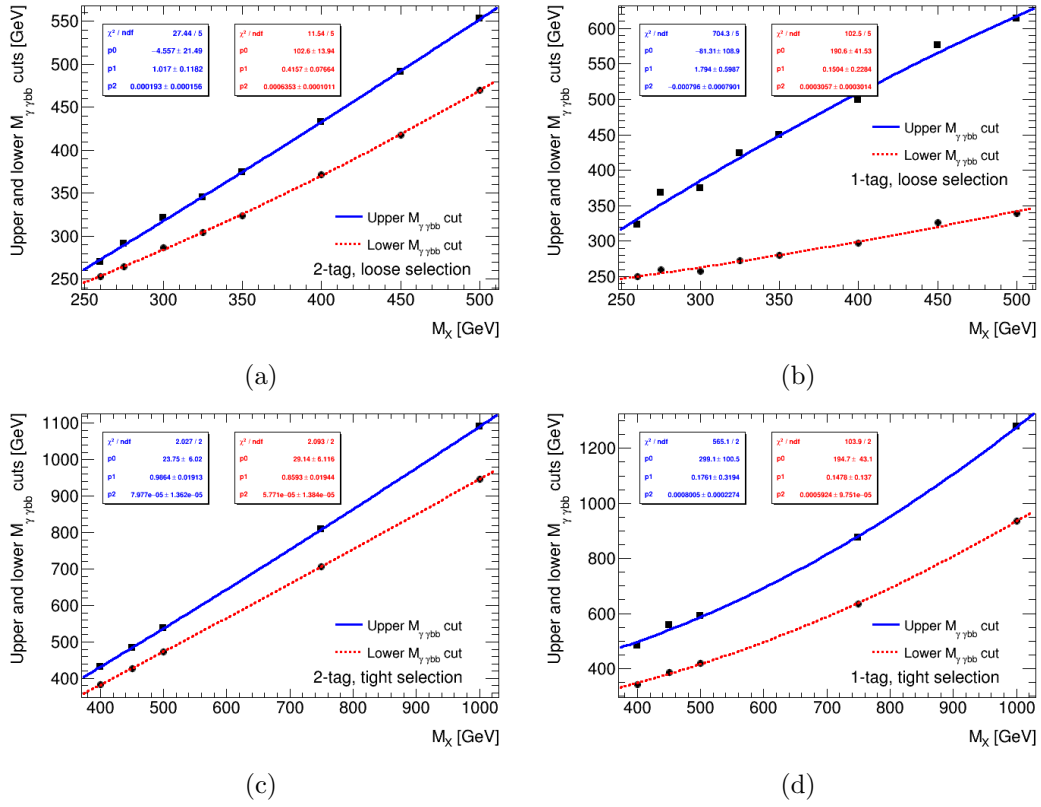


Figure 5.19: Parametrization with a 2$^{\text{nd}}$ order polynomial function of the $m_{\gamma\gamma jj}$ windows as a function of m_X for the a) 2-tag and b) 1-tag categories with the loose selection and the c) 2-tag and d) 1-tag categories with the tight selection.

The $\epsilon^{B}_{m_{\gamma\gamma jj}}$ efficiency is determined from the $m_{\gamma\gamma jj}$ spectrum in the simulated 2-tag and 1-tag categories. This approach assumes the kinematics are similar in the data and the simulated events which introduces an uncertainty. This is discussed in Section 5.5.3. The $m_{\gamma\gamma jj}$ spectrum, after applying the $m_{\gamma\gamma}$ selection, is fitted with a

Novosibirsk function for the 1-tag and 2-tag categories with the loose selection, and with an exponential function for the tight selection categories. The contribution from the single Higgs boson production, approximately 5–7% (12–13%) with loose (tight) selection, is removed from the distribution before fitting. The values of $\epsilon^{B}_{m_{\gamma\gamma jj}}$ are calculated for each resonance mass using the integral of the corresponding function in the full mass range and in the 95% window. The values range from 0.38±0.04 (0.08±0.03) until 0.68±0.03 (0.21±0.04) in the 1-tag (2-tag) category with the loose selection for different resonance mass hypotheses. In the case of the tight selection, where the event statistics are less, the values of $\epsilon^{B}_{m_{\gamma\gamma jj}}$ range from 0.02±0.02 (0.03±0.04) until 0.53±0.05 (0.15±0.05) in the 1-tag (2-tag) category for different resonance mass hypotheses.

The contribution from the SM single Higgs boson production in the signal region are calculated in a similar way by using:

$$N^{SM}_{SR} = N^{SM} \times \epsilon^{SM}_{m_{\gamma\gamma jj}}, \tag{5.5}$$

where N^{SM} denotes the total number of single Higgs boson events in the $m_{\gamma\gamma}$ signal region. The rates are normalized to the SM expectations. The values of $\epsilon^{SM}_{m_{\gamma\gamma jj}}$ are calculated for each resonance mass using the integral of the corresponding function (Novosibirsk (exponential) for loose (tight) selection categories) in the full mass range and in the 95% window. The values range from 0.38±0.06 (0.07±0.02) until 0.67±0.06 (0.22±0.04) in the 1-tag (2-tag) category with the loose selection for different resonance mass hypotheses. In the case of the tight selection the values of $\epsilon^{SM}_{m_{\gamma\gamma jj}}$ range from 0.04±0.03 (0.02±0.03) until 0.50±0.06 (0.19±0.06) in the 1-tag (2-tag) category for different resonance mass hypotheses.

5.5 Systematic uncertainties

Although statistical uncertainties dominate the sensitivity of this analysis given the small number of events, care is taken to make the best possible estimates of all systematic uncertainties, as described in more detail below.

5.5.1 Theoretical uncertainties

Theoretical uncertainties in the production cross section of single Higgs bosons are estimated by varying the renormalization and factorisation scales. In addition, uncertainties due to the PDF and the running of the QCD coupling constant (α_S) are considered. The scale uncertainties reach a maximum of $^{+20\%}_{-24\%}$ and the PDF+α_S uncertainty is not more than $\pm 3.6\%$ [34]. An uncertainty in the rate of Higgs boson production with associated heavy flavor jets is also considered. A 100% uncertainty is assigned to the ggH and WH production modes, motivated by studies of heavy flavor production in association with top-quark pairs [155] and W boson production in association with b-jets [156]. No heavy flavor uncertainty is assigned to the ZH and $t\bar{t}H$ production modes, where the dominant heavy flavor contribution is already accounted for in the LO process. Finally, additional theoretical uncertainties in single Higgs boson production from uncertainties in the $H \to \gamma\gamma$ and $H \to b\bar{b}$ branching fractions are $^{+2.9\%}_{-2.8\%}$ and $\pm 1.7\%$, respectively [34].

The same sources of uncertainty are considered on the SM HH signal samples. The effect of scale and PDF+α_S uncertainties on the NNLO cross section for SM Higgs boson pair production are 4–8% and 2–3% respectively. In addition, an uncertainty of 5% arising from the simplifications used in the EFT approximation is taken into account [141].

In the search for resonant Higgs boson pair production, uncertainties arising from scale and PDF uncertainties, which primarily affect the signal yield, are neglected.

For this search, the SM non-resonant HH production is considered as a background, with an overall uncertainty on the cross section of $^{+7\%}_{-8\%}$. Interference between SM HH and the BSM signal is neglected.

For all samples, systematic differences between alternative models of parton showering and hadronisation were considered and found to have a negligible impact.

5.5.2 Experimental uncertainties

The systematic uncertainty in the integrated luminosity for the data in this analysis is 2.1%. It is derived following a methodology similar to that detailed in [157], using beam-separation scans performed in 2015 and 2016.

The efficiency of the diphoton trigger is measured using bootstrap methods [158], and is found to be 99.4% with a systematic uncertainty of 0.4%. Uncertainties associated with the vertex selection algorithm have a negligible impact on the signal selection efficiency.

Differences between data and simulation give rise to uncertainties in the calibration of the photons and jets used in this analysis. As the continuum backgrounds are estimated from data, these uncertainties are applied only to the signal processes and to the single Higgs boson background process. In order to calculate the impact of the experimental uncertainties, signal and background fits are performed as described in Section 5.4, with the relevant observables varied within their uncertainties. Changes in the peak location (m_{peak}), width (σ_{peak}) and expected yield in $m_{\gamma\gamma}$ ($m_{\gamma\gamma jj}$) for the non-resonant (resonant) model, relative to the nominal fits, are extracted. The tail parameters are kept at their nominal values in these modified fits. For the resonant analysis, systematic uncertainties are evaluated for each m_X and the maximum across the range is taken as a conservative uncertainty.

The dominant yield uncertainties are listed in Table 5.15. Uncertainties in the

photon identification and isolation directly affect the diphoton selection efficiency; jet energy scale and resolution uncertainties affect the m_{bb} window acceptance [159–161], while flavor tagging uncertainties lead to migration of events between categories. Uncertainties in the peak location (width), which are mainly due to uncertainties in the photon energy scale (energy resolution), are about 0.2–0.6% (5–14%) for both the single Higgs boson and Higgs boson pair samples in the resonant and non-resonant analyses.

The spurious signal for the chosen background model, as defined in Sections 5.4.2 and 5.4.3, is assessed as an additional uncertainty in the total number of signal events in each category. In the 2-tag (1-tag) category, the uncertainty corresponds to 0.63 (0.25) events for the non-resonant analysis, 0.58 (2.06) events for the resonant analysis with the loose selection, and 0.21 (0.89) events for the resonant analysis with the tight selection.

Finally, as described in Section 5.4.3, an m_X-dependent correction to the signal cross section, together with its associated uncertainty, is applied in the case of the resonant analysis at low masses to adjust for a small degeneracy bias.

5.5.3 Background modeling uncertainties

The uncertainties from modeling the continuum background shape are only applied in the resonant cut-and-count method. The uncertainty on $\epsilon^B_{m_{\gamma\gamma jj}}$, in the resonant Higgs boson pair production, has two components which are summed in quadrature: the heavy flavor uncertainty and the statistical uncertainty from the fitting procedure. The heavy flavor uncertainty originates from the assumption that the shape of the 2-tag and 1-tag categories in the simulated events can be used to estimate the $m_{\gamma\gamma jj}$ shape in the data. A Novosibirsk (exponential) is fitted to the four object distribution ($m_{\gamma\gamma jj}$) that falls in the $m_{\gamma\gamma}$ signal region for the loose (tight) selection categories. For

Table 5.15: Summary of dominant systematic uncertainties affecting expected yields in the resonant and non-resonant analyses. For the non-resonant analysis, uncertainties in the Higgs boson pair signal and SM single Higgs boson backgrounds are presented. For the resonant analysis, uncertainties on the Higgs boson pair signal for the loose and tight selections are presented. Sources marked "-" and other sources not listed in the table are negligible by comparison. No systematic uncertainties related to the continuum background are considered, since this is derived through a fit to the observed data.

Source of systematic uncertainty		% effect relative to nominal in the 2-tag (1-tag) category							
		Non-resonant analysis				Resonant analysis: BSM HH			
		SM HH signal		Single-H bkg		Loose selection		Tight selection	
Luminosity		±2.1	(±2.1)	±2.1	(± 2.1)	±2.1	(±2.1)	±2.1	(±2.1)
Trigger		±0.4	(±0.4)	±0.4	(± 0.4)	±0.4	(±0.4)	±0.4	(±0.4)
Pile-up modeling		±3.2	(±1.3)	±2.0	(± 0.8)	±4.0	(±4.2)	±4.0	(±3.8)
Photon	identification	±2.5	(±2.4)	±1.7	(± 1.8)	±2.6	(±2.6)	±2.5	(±2.5)
	isolation	±0.8	(±0.8)	±0.8	(± 0.8)	±0.8	(±0.8)	±0.9	(±0.9)
	energy resolution	-		-		±1.0	(±1.3)	±1.8	(±1.2)
	energy scale	-		-		±0.9	(±3.0)	±0.9	(±2.4)
Jet	energy resolution	±1.5	(±2.2)	±2.9	(± 6.4)	±7.5	(±8.5)	±6.4	(±6.4)
	energy scale	±2.9	(±2.7)	±7.8	(± 5.6)	±3.0	(±3.3)	±2.3	(±3.4)
Flavor tagging	b-jets	±2.4	(±2.5)	±2.3	(± 1.4)	±3.4	(±2.6)	±2.5	(±2.6)
	c-jets	±0.1	(±1.0)	±1.8	(±11.6)	-		-	
	light-jets	<0.1	(±5.0)	±1.6	(± 2.2)	-		-	
Theory	PDF+α_S	±2.3	(±2.3)	±3.1	(± 3.3)	n/a		n/a	
	Scale	+4.3	(+4.3)	+4.9	(+ 5.3)	n/a		n/a	
		−6.0	(−6.0)	+7.0	(+ 8.0)	n/a		n/a	
	EFT	±5.0	(±5.0)	n/a		n/a		n/a	

low-mass resonances the background fit function has a characteristic peaked shape. For this reason the peak of the Novosibirsk fit in the simulated events is constrained to the extracted data-driven value within its statistical uncertainty. A simple bias test is performed by drawing pseudo-data sets from the overall probability distribution created by combining the Novosibirsk (exponential) background function with the signal function. For each mass point and each value of the injected signal cross section, fits are performed on the ensemble of pseudo-data sets and the median extracted signal cross section is recorded. This is repeated for each resonance mass resulting in an uncertainty ranging from 4% ($m_X = 300$ GeV) to 13% ($m_X = 400$ GeV). The largest

error is used for all hypotheses.

Modified Landau (Novosibirsk + Modified Landau) fits are used to parametrize the window efficiency in simulated events for the 1-tag (2-tag) categories with the loose selection. The standard deviation gives the statistical uncertainty on the estimated efficiency, which is dependant on the mass hypothesis. The largest error is used for all hypotheses and corresponds to 6% (5%) in the 1-tag (2-tag) category with the loose selection. The same procedure is followed in the tight selection categories but using a 2^{nd} order polynomial fit function. In this case the largest error corresponds to 7% in both the 1-tag and 2-tag categories. The uncertainty on $\epsilon^{SM}_{m_{\gamma\gamma jj}}$ also affects the SM single Higgs boson process which contributes to the background for the resonant analysis. The parametrized signal mass window efficiency for these samples is used to extract its standard deviation. Again, the largest difference within mass hypotheses is chosen which is found to be 12% (32%) for the 1-tag (2-tag) categories with the loose selection and 7% (11%) for the 1-tag (2-tag) categories with the tight selection.

5.6 Statistical analysis

The data is interpreted following the statistical procedure summarized in [148, 162]. An extended likelihood function is built from the number of events observed and the invariant masses observed. For the non-resonant analysis, the diphoton mass $m_{\gamma\gamma}$ is used while the four body mass $m_{\gamma\gamma jj}$ is used in the resonant analysis.

The likelihood for a given category i is given by

$$\mathcal{L}_i = \text{Pois}(n_i|N_i(\theta)) \cdot \prod_{j=1}^{n_i} f_i(m^j_{\gamma\gamma(jj)}, \theta) \cdot G(\theta), \tag{5.6}$$

where n_i (N_i) is the observed (expected) number of selected candidates, $f_i(m^j_{\gamma\gamma(jj)})$ is the value of the probability density function (pdf) of the invariant mass distribution

evaluated for each candidate j, θ are nuisance parameters and $G(\theta)$ is a set of unit Gaussian constraints on a subset of the nuisance parameters. The likelihood is given by the product of the likelihood function for the 1-tag and 2-tag categories.

5.6.1 Non-resonant analysis

In the non-resonant search, the number of events expected N_i is found by summing the contributions from the continuum background $N_{\text{bkg},i}$, the expected yields from single Higgs boson production $N_{\text{Higgs},i}$, non-resonant HH production, $N_{\text{HH},i}$ and the spurious signal yield, $N_i^{\text{spur}} \cdot \theta_i^{\text{spur}}$.

$$N_i = \mu \cdot N_{\text{HH},i}(\boldsymbol{\theta}_i^{\text{yield}}, \boldsymbol{\theta}_i^{\text{migr}}) + N_{\text{Higgs},i}(\boldsymbol{\theta}_i^{\text{yield}}, \boldsymbol{\theta}_i^{\text{migr}}) + N_{\text{bkg},i} + N_{spur,i} \cdot \boldsymbol{\theta}_{\text{spur},i} \quad (5.7)$$

The yield expected from single Higgs boson production is fixed to the prediction from the Monte Carlo samples while the yield from HH production is parametrized by μ, defined as the measured HH yield divided by the prediction. The shape $f_i(m_{\gamma\gamma}^j)$ is given by

$$\begin{aligned} f_i(m_{\gamma\gamma}^j) =& [(\mu \cdot N_i^{\text{HH}} + N_i^{\text{Higgs}} + N_i^{\text{spur}} \cdot \theta_i^{\text{spur}}) \cdot f_i^{\text{sig}}(m_{\gamma\gamma}^j, \theta_i^{\text{sig}}) \\ &+ N_i^{\text{bkg}} \cdot f_i^{\text{bkg}}(m_{\gamma\gamma}^j, \theta_i^{\text{bkg}})]/N_i, \end{aligned} \quad (5.8)$$

where $f_i^{\text{sig}}(m_{\gamma\gamma}^j)$ is the double-sided Crystal Ball function and $f_i^{\text{bkg}}(m_{\gamma\gamma}^j)$ is the background function which has been selected with the spurious signal test.

5.6.2 Resonant analysis

In the resonant analysis the fit is performed to the four body mass, $m_{\gamma\gamma jj}$. The single Higgs boson background does not have a resonant structure in this variable so it is included in the prediction of the continuum background. The number of events

expected N_i is

$$N_i = \mu \cdot N_{\mathrm{HH},i}(\boldsymbol{\theta}_i^{\mathrm{yield}}, \boldsymbol{\theta}_i^{\mathrm{migr}}) + N_{\mathrm{bkg},i} + N_{spur,i} \cdot \boldsymbol{\theta}_{\mathrm{spur},i} \tag{5.9}$$

and the shape $f_i(m^j_{\gamma\gamma jj})$ is given by

$$f_i(m^j_{\gamma\gamma jj}) = [(\mu \cdot N_i^{\mathrm{HH}} + N_i^{\mathrm{spur}} \cdot \theta_i^{\mathrm{spur}}) \cdot f_i^{\mathrm{sig}}(m^j_{\gamma\gamma jj}, \theta_i^{\mathrm{sig}}) + N_i^{\mathrm{bkg}} \cdot f_i^{\mathrm{bkg}}(m^j_{\gamma\gamma jj}, \theta_i^{\mathrm{bkg}})]/N_i, \tag{5.10}$$

where $f_i^{\mathrm{sig}}(m^j_{\gamma\gamma jj})$ is the ExpGaussExp [153] function and $f_i^{\mathrm{bkg}}(m^j_{\gamma\gamma})$ is the Novosibirsk (exponential) for the loose (tight) selection.

The alternative cut-and-count method for the resonant analysis is based on the $m_{\gamma\gamma}$ and $m_{\gamma\gamma jj}$ windows as discussed in Section 5.4.3. In this approach the likelihood is a product of two Poisson bins for the 2-tag and 1-tag categories.

The number of events in the $m_{\gamma\gamma}$ sidebands remains the same for all hypotheses and the number of observed events in the different resonance mass windows changes. The number of events in the sideband and signal regions can be split into contributions from the background continuum, the SM and BSM as:

$$\begin{aligned} N_{Sideband} &= N_{SB}^{Continuum} + N_{SB}^{SM} + N_{SB}^{BSM}, \\ N_{SR} &= N_{SR}^{Continuum} + N_{SR}^{SM} + N_{SR}^{BSM}, \end{aligned} \tag{5.11}$$

where each of the signal region terms are defined as:

$$
\begin{aligned}
N_{SR}^{Continuum} &= N_{SB}^{Continuum} \times \frac{N_{SR}^{MC}}{N_{SB}^{MC}} \times \epsilon_{m_{\gamma\gamma jj}}^{B}, \\
N_{SR}^{SM} &= N^{SM} \times \epsilon_{m_{\gamma\gamma jj}}^{SM}, \\
N_{SR}^{BSM} &= N^{BSM} \times \epsilon_{m_{\gamma\gamma jj}}^{BSM},
\end{aligned} \tag{5.12}
$$

where $\epsilon_{m_{\gamma\gamma jj}}^{BSM}$ is 0.95 by definition for any mass hypothesis. The number of BSM events after the tight $m_{\gamma\gamma}$ cut is given by:

$$
N^{BSM} = \mathcal{B}_{\gamma\gamma} \times \mathcal{B}_{b\bar{b}} \times 2 \times 1000 \times \mathcal{L} \times \text{resonant sel. eff.} \times \text{npbBSM}, \tag{5.13}
$$

where $\mathcal{L}$ is given in fb^{-1}. The fitted parameter, in addition to the systematic uncertainties, is npbBSM from Equation 5.13, and the fit parameters coming from the 2-tag and 1-tag signal categories are correlated.

5.6.3 Limit setting

The measured HH signal yields in both the non-resonant and resonant analysis can be determined with the profile likelihood ratio test statistic:

$$
\Lambda(\nu) = -2 \ln \frac{\mathcal{L}(\nu, \hat{\theta}_\nu)}{\mathcal{L}(\hat{\nu}, \hat{\theta})}, \tag{5.14}
$$

where $\hat{\nu}$ and $\hat{\theta}$ are the values of the parameter of interest and the nuisance parameters that unconditionally maximize the likelihood while $\hat{\theta}_\nu$ are the values of the nuisance parameters that maximize the likelihood on the condition that the parameter of interest is held fixed to a given value ν.

In the absence of any significant excess, upper limits will be established. For

the non-resonant analysis, upper limits can be established on the production cross section and BSM values of the Higgs boson self-coupling can also be excluded. For resonant HH production, upper limits can be placed on the production cross section as a function of m_X.

The 95% confidence level (CL) upper limit on the parameter of interest ν is determined using the $\mathrm{CL_s}$ prescription [163]. For this, the agreement between data and the expected yield for the hypothesized value of the parameter of interest ν is quantified by the test statistic, q_ν, defined as

$$q_\nu = \begin{cases} \Lambda(\nu) & 0 < \hat{\nu} \leq \nu \\ 0 & \nu < \hat{\nu} \end{cases}, \tag{5.15}$$

where $\hat{\nu} \geq 0$ is the fitted parameter of interest. The observed value of the test statistic, q_{obs}, is determined from the ratio of the likelihood obtained by fixing the number of signal events to that predicted for a given value of the parameter of interest, to the likelihood normalized by allowing the number of signal events to float in the fit.

Since both the expected and observed numbers of events are small in the analysis, test-statistic distributions are evaluated by pseudo-experiments generated by profiling the nuisance parameters of the likelihood model on the observed data, as described in [164]. Better limits on κ_λ are expected with the loose selection, whereas for the SM value $\kappa_\lambda = 1$ the strongest limits on the Higgs boson pair production cross section are derived from the tight selection.

5.7 Results

The observed data is in good agreement with the data-driven background expectation, as summarized in Table 5.16. Across all categories, the number of observed events in data is compatible with the number of expected background events within the

calculated uncertainties. Therefore, in the absence of any significant excess, upper limits are set.

Table 5.16: Expected and observed numbers of events in the 1-tag and 2-tag categories for events passing the selection for the resonant analysis, including the $m_{\gamma\gamma}$ requirement. The event numbers quoted for the SM Higgs boson pair signal assume that the total production cross section is 33.41 fb. The uncertainties on the continuum background are those arising from the fitting procedure. The uncertainties on the single Higgs boson and Higgs boson pair production yields are the systematics from experimental and theoretical sources. The loose and tight selections are not orthogonal.

	1-tag		2-tag	
	Loose selection	Tight selection	Loose selection	Tight selection
Continuum background	118 ±5	15.7 ± 1.6	21 ± 2	3.7 ± 0.8
SM single Higgs boson background	5.51 ±0.10	2.20 ± 0.05	1.63 ± 0.04	0.56 ± 0.02
Total background	123 ±5	17.9 ± 1.6	23 ± 2	4.3 ± 0.8
SM Higgs boson pair signal	0.219±0.006	0.120± 0.004	0.305± 0.007	0.175± 0.005
Data	125	19	21	3

Figure 5.20 shows the observed diphoton invariant mass spectra for the non-resonant analysis with the loose (top) and tight (bottom) selections. The best-fit Higgs boson pair cross section is $0.04^{+0.43}_{-0.36}$ ($0.21^{+0.33}_{-0.25}$) pb for the loose (tight) selection. Figure 5.21 shows the observed four-body invariant mass spectra for the resonant analysis in the loose (top) and tight (bottom) selections. Maximum-likelihood background-only fits are also shown. The largest discrepancy between the background-only hypothesis and the data manifests as an excess at 480 GeV with a local significance of 1.2σ. The results are also interpreted as upper limits on the relevant Higgs boson pair production cross sections.

5.7.1 Exclusion limits on non-resonant *HH* production

The 95% CL upper limit for the non-resonant Higgs boson pair production cross section is obtained using the tight selection. Figure 5.22a shows this upper limit,

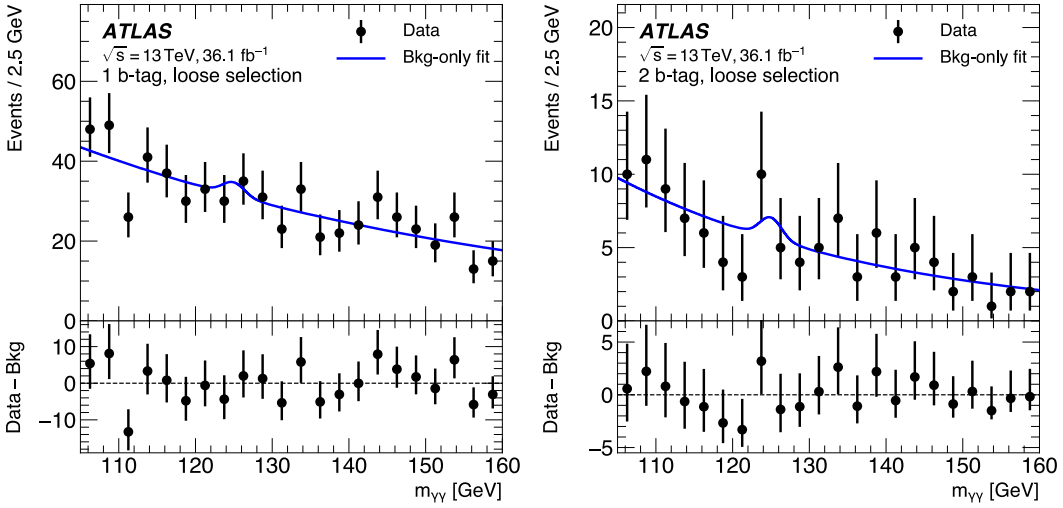


Figure 5.20: For the non-resonant analysis, data (black points) are compared with the background-only fit (blue solid line) for $m_{\gamma\gamma}$ in the 1-tag (left) and 2-tag (right) categories with the loose (top) and tight (bottom) selections. Both the continuum $m_{\gamma\gamma}$ background and the background from single Higgs boson production are considered. The lower panel shows the residuals between the data and the best-fit background [3].

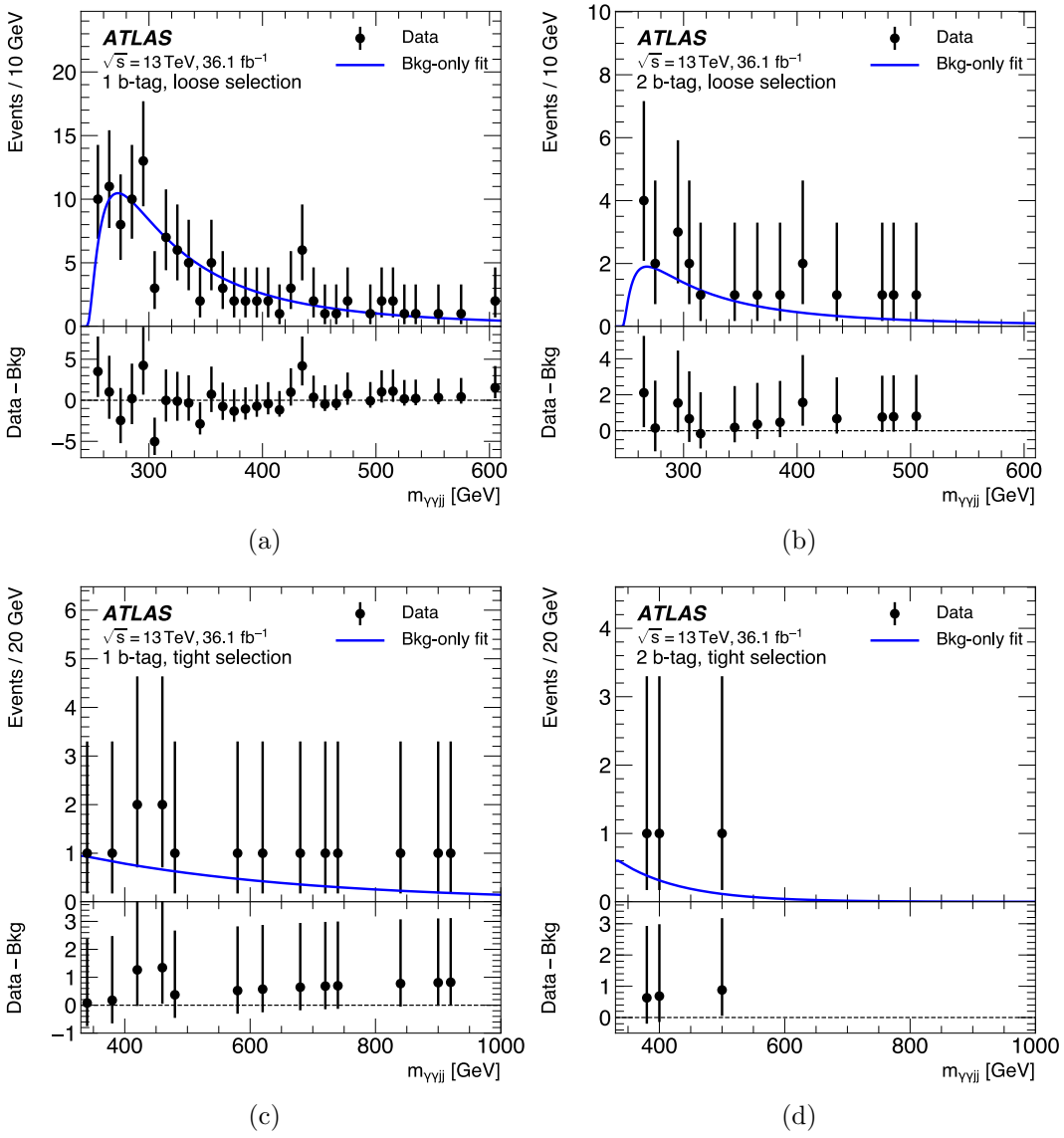


Figure 5.21: For the resonant analysis, data (black points) are compared with the background-only fit (blue solid line) for $m_{\gamma\gamma jj}$ in the 1-tag (left) and 2-tag (right) categories with the loose (top) and tight (bottom) selections. The lower panel shows the residuals between the data and the best-fit background [3].

together with $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands. The observed (expected) value is 0.73 (0.93) pb. As a multiple of the SM production cross section, the observed (expected) limits are 22 (28). The limits and the $\pm 1\sigma$ band around each expected limit are presented in Table 5.17.

Table 5.17: The 95% CL observed and expected limits on the Higgs boson pair cross section in pb and as a multiple of the SM production cross section. The $\pm 1\sigma$ band around each 95% CL limit is also indicated.

	Observed [pb]	Expected [pb]	-1σ [pb]	$+1\sigma$ [pb]
$\sigma_{gg \to HH}$	0.73	0.93	0.66	1.4
As a multiple of σ_{SM}	22	28	20	40

5.7.2 Exclusion limits on the Higgs boson self-coupling

Varying the Higgs boson self-coupling, λ_{HHH}, affects both the total cross section of the non-resonant Higgs boson pair production and the event kinematics, affecting the signal selection efficiency. In the non-resonant analysis, results are interpreted in the context of κ_λ, using the loose selection, which is more sensitive for the range of κ_λ values accessible with this data set. The 95% CL limits on $\sigma_{gg \to HH}$ are shown together with $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands around the expected limit in Figure 5.22b. The limits are calculated using the asymptotic approximation [148] for the profile-likelihood test statistic. Fixing all other SM parameters to their expected values, the Higgs boson self-coupling is constrained at 95% CL to $-8.2 < \kappa_\lambda < 13.2$ whereas the expected limits are $-8.3 < \kappa_\lambda < 13.2$.

5.7.3 Exclusion limits on resonant *HH* production

The 95% CL limits on resonant Higgs boson pair production are shown in Figure 5.23, using the loose selection for $m_X \leq 500$ GeV and the tight selection for $m_X \geq 500$ GeV. The SM HH contribution is considered as part of the background in this search although its inclusion has a negligible impact on the results. For resonance masses in the range 260 GeV $< m_X <$ 1000 GeV, the observed (expected) limits range between 1.14 (0.90) pb and 0.12 (0.15) pb.

Cut-and-count method

The expected and observed number of events in the resonant search with the cut-and-count method are given in Table 5.18 for the 1-tag and 2-tag categories with the loose and tight selections. They are obtained using the parametrization of the $m_{\gamma\gamma jj}$

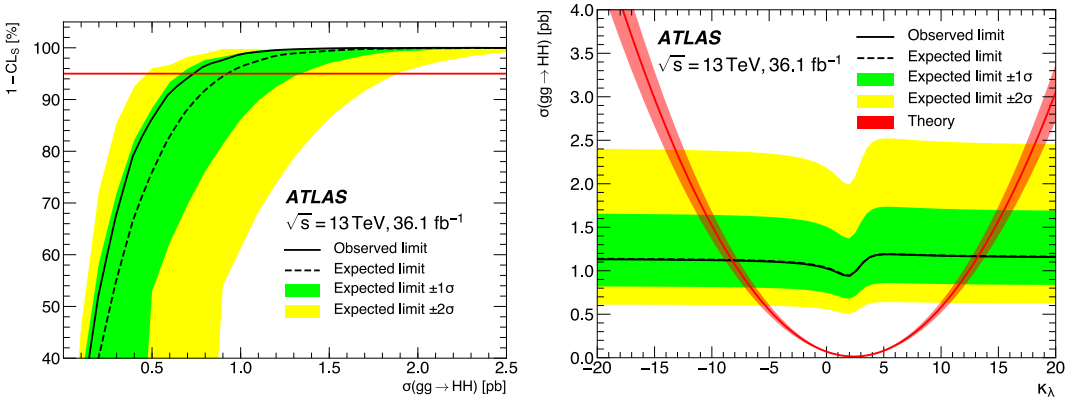
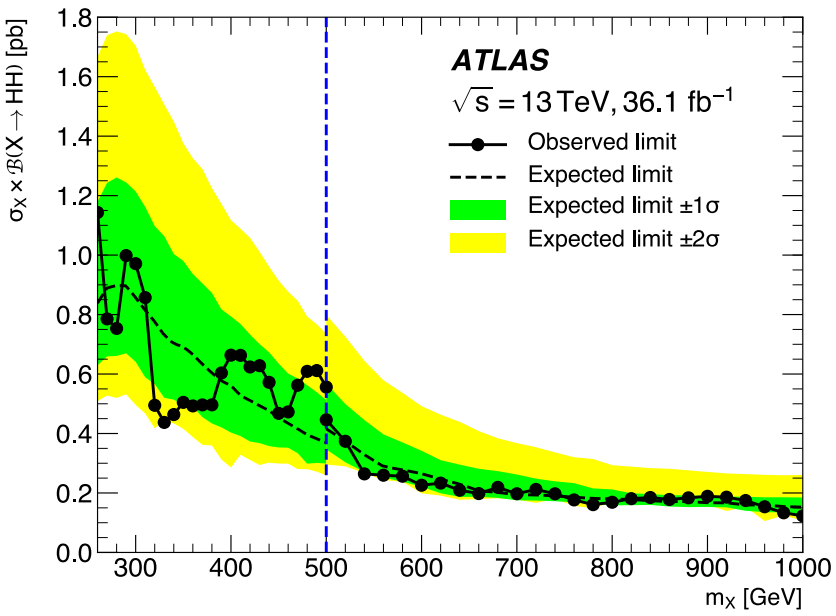


Figure 5.22: The expected and observed 95% CL limits on the non-resonant production cross section $\sigma_{gg\to HH}$ a) for the SM-optimized limit using the tight selection and b) as a function of κ_λ using the loose selection. In a) the red line indicates the 95% confidence level. The intersection of this line with the observed, expected, and $\pm1\sigma$ and $\pm2\sigma$ bands is the location of the limits. In b) the red line indicates the predicted HH cross section if κ_λ is varied but all other couplings remain at their SM values. The red band indicates the theoretical uncertainty of this prediction [3].

windows for each combination of category and selection, as shown in Figure 5.19. Figure 5.24 shows the 95% CL limits on resonant Higgs boson pair production as a function of m_X with the cut-and-count method as well as a comparison with the limits obtained with the fitting procedure.

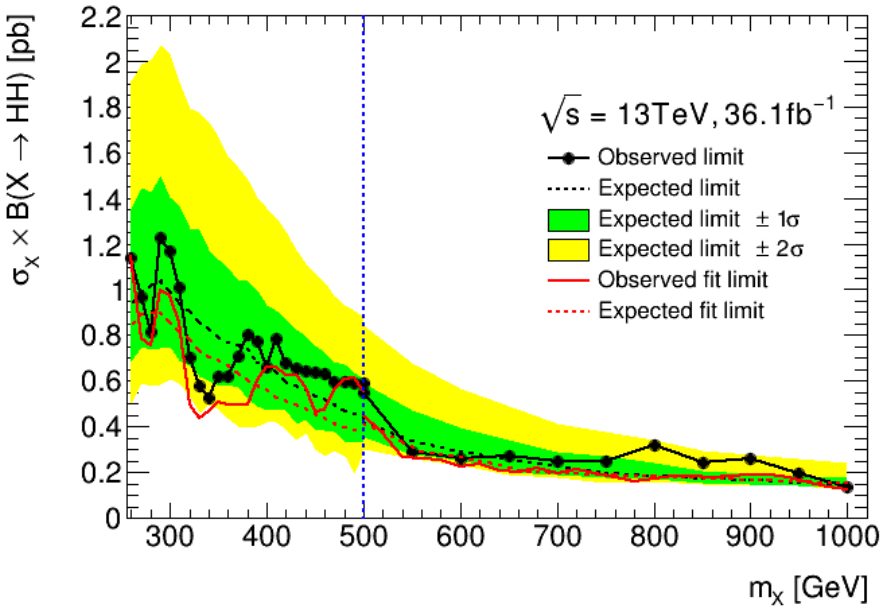
For resonance masses in the range 260 GeV $< m_X <$ 1000 GeV, the observed (expected) limits range between 1.23 (1.04) pb and 0.14 (0.15) pb. The expected limits show approximately a 10% increase in the range 260 GeV $< m_X <$ 500 GeV with respect to the results obtained with the fitting approach. The comparison of the observed limits shows a similar trend for both approaches with statistical fluctuations at some mass points. The expected and observed limits in the range 500 GeV $< m_X <$



(a)

Figure 5.23: The expected and observed 95% CL limits on the resonant production cross section, $\sigma_X \times \mathcal{B}\ (X \to HH)$ as a function of m_X. The loose selection is used for $m_X \leq 500$ GeV, while the tight selection is used for $m_X \geq 500$ GeV. This is delineated with the blue dashed line [3].

1000 GeV show a much higher compatibility between the fitting and the cut-and-count methods. The discrepancy in the observed limit at $m_X = 800$ GeV is a result of the differences between parametrizing the $m_{\gamma\gamma jj}$ windows and the parametrized signal model used in the fitting approach, where the 1-tag fit presents some imperfections at higher resonance masses, which translates into an underestimate of the number of signal events accounted by the model.



(a)

Figure 5.24: The expected and observed 95% CL limits on the resonant production cross section, $\sigma_X \times \mathcal{B}\,(X \to HH)$ as a function of m_X, with the cut-and-count method (black) and with the fitting approach (red). The loose selection is used for $m_X \leq 500$ GeV, while the tight selection is used for $m_X \geq 500$ GeV. This is delineated with the blue dashed line.

	2-tag		1-tag	
Resonance mass (GeV)	Expected	Observed	Expected	Observed
260	3.36	4	63.48	68
270	4.44	4	69.10	70
280	4.66	3	73.13	73
290	4.65	6	75.76	73
300	4.56	6	77.18	71
310	4.43	5	77.64	71
320	4.28	3	77.32	70
330	4.16	2	76.42	67
340	4.03	1	75.08	71
350	3.88	2	73.43	74
360	3.71	2	71.57	73
370	3.54	3	69.58	70
380	3.36	4	67.50	64
390	3.19	4	65.39	58
400	3.02	3	63.30	58
410	2.85	4	61.23	59
420	2.70	3	59.21	60
430	2.55	3	57.25	56
440	2.41	3	55.35	55
450	2.28	3	53.52	52
460	2.15	3	51.78	51
470	2.04	3	50.10	48
480	1.93	3	48.50	47
490	1.83	3	46.98	46
500	1.73	3	45.52	45
500	0.48	1	5.93	6
550	0.44	0	4.92	4
600	0.39	0	4.00	3
650	0.35	0	3.21	5
700	0.30	0	2.50	4
750	0.26	0	1.89	4
800	0.23	0	1.39	5
850	0.19	0	1.00	3
900	0.16	0	0.70	3
950	0.13	0	0.51	2
1000	0.11	0	0.42	0

Table 5.18: Observed and expected number of events in the resonant analysis for the 2-tag and 1-tag categories with the cut-and-count method. The loose selection is used for $m_X \leq$ 500 GeV, while the tight selection is used for $m_X \geq 500$ GeV.

5.8 Comparisons with other results

5.8.1 ATLAS *HH* results

ATLAS has performed searches for resonant and non-resonant Higgs boson pair production in other final states with data collected during Run 2 at $\sqrt{s} = 13$ TeV. The most sensitive analyses ($b\bar{b}b\bar{b}$, $\gamma\gamma b\bar{b}$ and $b\bar{b}\tau\tau$) have also been combined [165]. No evidence of HH production has been found in either the resonant and non-resonant searches.

Figure 5.25 shows the combined observed and expected 95% CL upper limits on the resonant production cross section as a function of the resonance mass, m_X. For resonance masses below 300 GeV, the $\gamma\gamma b\bar{b}$ analysis dominates the sensitivity while in the region 300–500 GeV, all three analyses contribute significantly. For masses greater than 500 GeV, the $b\bar{b}b\bar{b}$ drives the sensitivity of the combination.

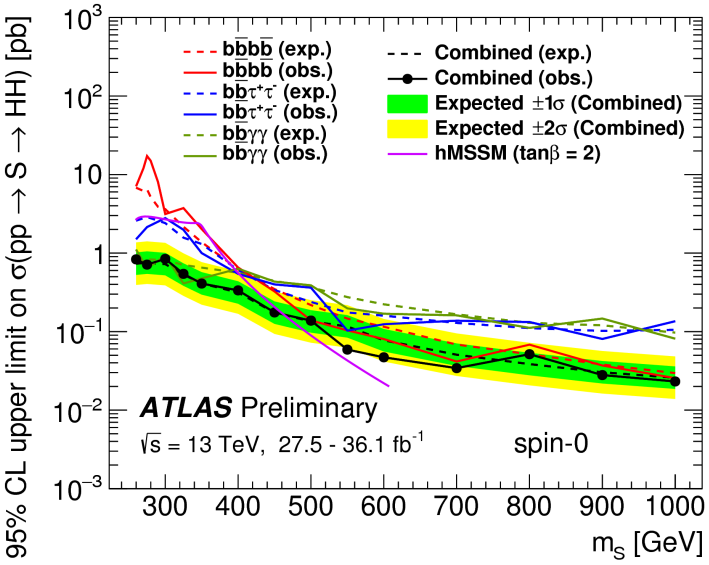


Figure 5.25: The combined observed and expected 95% CL upper limits on spin-0 resonant production cross section, $\sigma\ (p\text{-}p \to X) \times \mathcal{B}\ (X \to HH)$ as a function of m_X. The observed and expected limits from the individual analyses are also shown [165].

No significant excess is observed across the mass range probed. The observed (expected) combined 95% CL upper limits on the cross section for resonant HH production range from 0.83 (0.73) pb at 260 GeV to 0.02 (0.03) pb at 1 TeV.

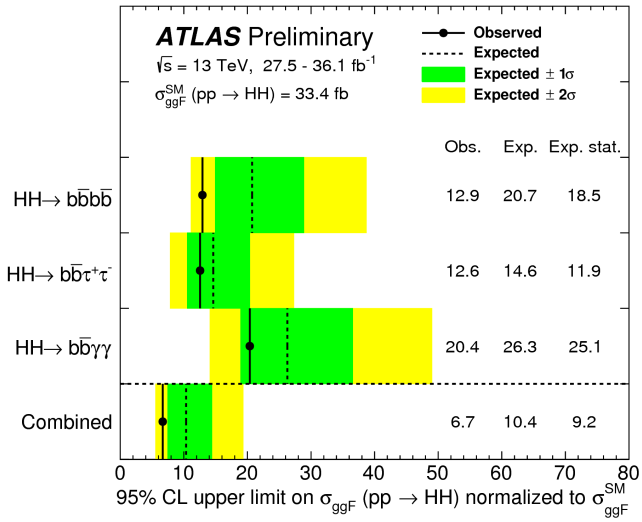
The 95% CL observed and expected upper limits on the signal strength of SM non-resonant Higgs boson pair production are shown for each individual analysis and for the combination in Figure 5.26a. The observed (expected) combined limit is 0.22 pb (0.35 pb), corresponding to 6.7 (10.4) times the SM prediction. The observed upper limit set is stronger than expected due to all three analyses having a deficit of data with respect to the background-only prediction.

The combined results are also interpreted in terms of the Higgs boson self-coupling, λ_{HHH}. Figure 5.26b shows the 95% CL observed and expected upper limits on the cross section as a function of κ_λ. The value of κ_λ is observed (expected) to be constrained at 95% CL to -5.0 $< \kappa_\lambda <$ 12.1 (-5.8 $< \kappa_\lambda <$ 12.0). It can be seen in Figure 5.26b that the $b\bar{b}b\bar{b}$ and $b\bar{b}\tau\tau$ analyses have a worse expected constraint for κ_λ than the $\gamma\gamma b\bar{b}$ analysis despite both having a better exclusion limit for SM HH production. The kinematics change significantly as a function of κ_λ, in particular for higher values of κ_λ the physics objects have lower transverse momentum which results in a significant decrease in efficiency in the $b\bar{b}b\bar{b}$ and $b\bar{b}\tau\tau$ analyses. For this reason, the $\gamma\gamma b\bar{b}$ analysis dominates the sensitivity at higher values of κ_λ.

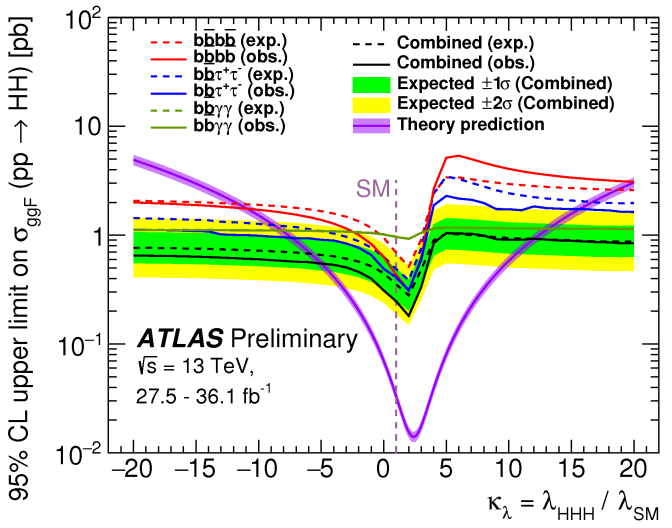
5.8.2 CMS $H \to \gamma\gamma b\bar{b}$ results

The CMS collaboration has also searched for resonant and non-resonant Higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state using a data set of 35.9 fb^{-1} at $\sqrt{s} = 13$ TeV [135]. CMS also observes no evidence of Higgs boson pair production in this final state.

Figure 5.27a shows the observed and expected 95% CL upper limits on the res-



(a)



(b)

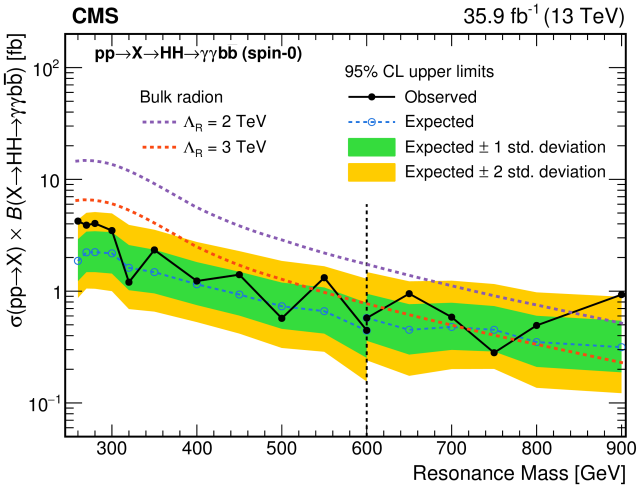
Figure 5.26: a) Observed and expected upper limits at 95% CL on the cross section of non-resonant SM Higgs boson pair production from the individual analyses and the combination. b) Observed and expected upper limits at 95% CL on the cross section of non-resonant Higgs boson pair production as a function of κ_λ [165].

onant production cross section as a function of the resonance mass. For resonance masses below 300 GeV, the expected 95% CL upper limit on the production cross section times branching ratios including that of $H \to \gamma\gamma$ and $H \to b\bar{b}$ is approximately 2.0 fb. It can be seen in Figure 5.24 that for resonance masses below 300 GeV, the expected 95% CL upper limit from ATLAS is approximately 0.9 pb. This result does not include the branching ratios of $H \to \gamma\gamma$ and $H \to b\bar{b}$. By including these, an expected limit of 2.3 fb is obtained, which is comparable to the mentioned CMS result.

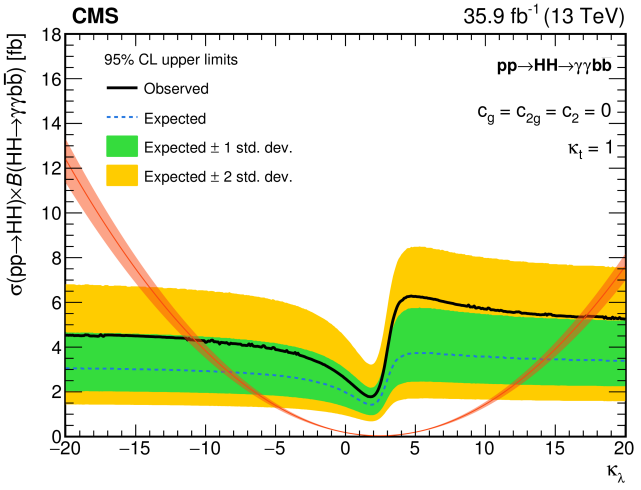
The observed (expected) 95% CL upper limit for SM non-resonant HH production is 0.79 (0.63 pb), which corresponds to 24 (19) times the Standard Model value. This expected limit is better than the expected limit reported by ATLAS (28 times the SM prediction). This is as a result of the CMS collaboration having a dedicated selection for the non-resonant analysis, an aspect ATLAS can improve on in future in order to improve the SM HH sensitivity. Figure 5.27b shows the observed and expected upper limits on the non-resonant cross section as a function of κ_λ. The value of κ_λ is observed (expected) to be constrained at 95% CL to $-11 < \kappa_\lambda < 17$ ($-8 < \kappa_\lambda < 15$). The expected κ_λ constraint reported by ATLAS is $-8.3 < \kappa_\lambda < 13.2$, showing a good agreement between both results.

5.8.3 Prospects for the HL-LHC

In the near future the LHC will undergo a number of upgrades in order to increase the luminosity to approximately 7–10 times the design luminosity. This upgraded LHC is referred to as the High-Luminosity LHC (HL-LHC) which aims to deliver 3000 fb^{-1} of p-p collisions to the ATLAS detector. One of the primary goals of the HL-LHC is to find evidence of HH production. The ATLAS collaboration has studied the prospects for observing HH with the HL-LHC in the $\gamma\gamma b\bar{b}$ final state [166].



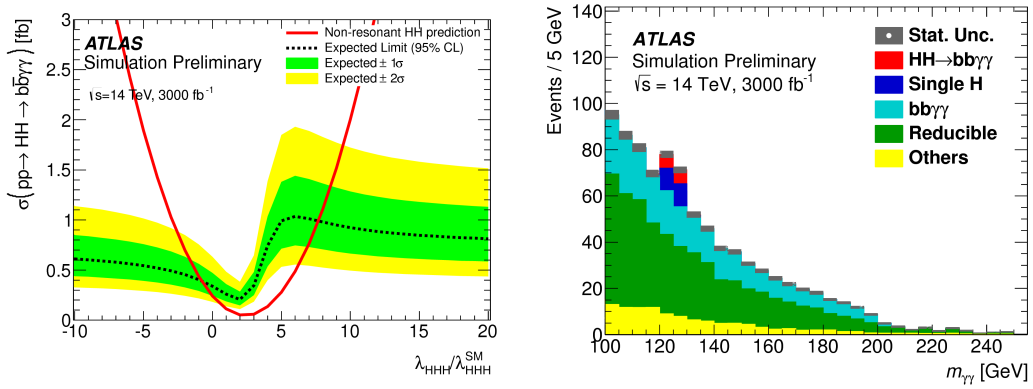
(a)



(b)

Figure 5.27: a) Observed and expected 95% CL upper limits on resonant production cross section as a function of the resonance mass for a spin-0 hypothesis. b) Observed and expected upper limits at 95% CL on the cross section of non-resonant Higgs boson pair production as a function of κ_λ [135].

The $\gamma\gamma b\bar{b}$ HL-LHC study is performed by generating large MC samples and smearing the truth physics objects to emulate the response of an upgraded ATLAS detector in the HL-LHC pile-up conditions. An expected statistical significance of 1.05σ is found for this final state. The Higgs boson self-coupling is expected to be constrained to -0.8 $\leq \kappa_\lambda \leq$ 7.7 at 95% CL, as shown in Figure 5.28a. Figure 5.28b shows the diphoton invariant mass distribution after all selections.



(a) (b)

Figure 5.28: a) Expected 95% CL upper limits on the the cross section $\sigma(HH \to \gamma\gamma b\bar{b})$ as a function of κ_λ. b) The diphoton invariant mass distribution after all the selections. The shaded area corresponds to the MC statistical uncertainty [166].

The event selection used differs considerably from the current Run 2 analysis. The HL-LHC study can use much tighter selections, without the concern of still being able to obtain a data-driven background estimate. There are also differences in photon identification and isolation, where the efficiencies are expected to get worse at the HL-LHC, while the b-tagging performance is expected to improve with the upgraded ATLAS inner tracker.

5.9 Conclusions

Searches for resonant and non-resonant Higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state are performed using 36.1 fb^{-1} of p-p collision data collected at $\sqrt{s} = 13$ TeV with the ATLAS detector at the LHC in 2015 and 2016.

No significant deviations from the Standard Model predictions are observed. A 95% CL upper limit of 0.73 pb is set on the cross section for non-resonant production, while the expected limit is 0.93 pb. This observed (expected) limit is 22 (28) times the predicted SM cross section. The Higgs boson self-coupling is constrained at 95% CL to -8.2 $< \kappa_\lambda <$ 13.2 whereas the expected limits are -8.3 $< \kappa_\lambda <$ 13.2. For resonant production of $X \to HH \to \gamma\gamma b\bar{b}$, a limit is presented for the narrow-width approximation as a function of m_X. The observed (expected) limits range between 1.14 pb (0.90 pb) and 0.12 pb (0.15 pb) in the range 260 GeV $< m_X <$ 1000 GeV with the fitting approach, and between 1.23 (1.04) pb and 0.14 (0.15) pb in the same mass range with the cut-and-count method.

Chapter 6

Measurement of Higgs boson properties in the diphoton decay channel

This chapter describes the measurement of the Higgs boson production-mode cross sections, signal strengths, and simplified template cross sections (STXS), as well as the measurement of the fiducial and differential cross sections in the $H \to \gamma\gamma$ decay channel with the ATLAS detector. Special emphasis is made to cover the event reconstruction and selection of the analysis as well the strategy followed to estimate the uncertainty in the modeling of the parton shower, underlying event and hadronization, and how it affects the final results. The chapter concludes with a brief summary of the main findings.

6.1 Higgs boson production-mode cross sections and signal strengths

Cross sections times branching ratio of the Higgs boson decaying to two photons are measured for inclusive Higgs boson production, as well as for several individual production processes: gluon-gluon fusion (ggH), vector-boson fusion (VBF), Higgs boson production in association with a vector boson (VH), and production of a Higgs boson in association with a top-antitop quark pair ($t\bar{t}H$) or a single top quark (t-channel and W-associated, respectively denoted as tHq and tHW, or in their sum as tH) (see Section 1.4.1). The signal strength μ is defined as:

$$\mu = \frac{\sigma \times \mathcal{B}(H \to \gamma\gamma)}{\sigma_{SM} \times \mathcal{B}(H \to \gamma\gamma)_{SM}}, \tag{6.1}$$

where $\sigma \times \mathcal{B}(H \to \gamma\gamma)$ is the measured cross section times branching ratio of the Higgs boson decaying to two photons and $\sigma_{SM} \times \mathcal{B}(H \to \gamma\gamma)_{SM}$ corresponds to the SM prediction. Measurements of μ are given for different production processes as well as for the inclusive production.

Cross sections in specific regions of phase space using the framework of the "simplified template cross sections" (STXS) [167] are also reported. These regions of phase space are truth bins which are described further in Section 6.1.1.

6.1.1 Simplified template cross sections

The measurements of cross sections separated by the production mode are extended to measurements in specific regions of phase space using the framework of the "simplified template cross sections" [34, 168]. This framework was developed in order to reduce the theoretical uncertainties that are directly folded into measurements and provide more granular measurements but still allow the use of advanced analysis like

multivariable techniques.

The measurements are reported as cross section times $\mathcal{B}(H \to \gamma\gamma)$ for a Higgs boson absolute rapidity $|y_H|$ less than 2.5 and with further particle-level requirements. The different production modes ggH, VBF, VH and top quark associated production serve as starting points for the measurements and are labelled as "stage-0" STXS bins. This corresponds closely to the signal strength measurements which were reported in Run 1. However, the definition of the production modes has been modified, since for example there is little distinction between VH when the vector boson decays hadronically and VBF production, both are added together and the production mode is labelled as VBF + VH. The stage-0 bins are split further using the kinematics of the hadronic jets, Higgs boson and vector bosons into 31 "stage-1" STXS bins.

Since the current data set is not large enough to probe all of the 31 stage-1 cross sections with sufficiently small statistical uncertainties, regions with poor sensitivity or with large anti-correlations are merged together. Table 6.1 summarizes the merged stage-1 cross sections (middle column) as well as the full 31 stage-1 cross sections (right column) and the stage-0 cross sections (left column).

6.2 Fiducial integrated and differential cross sections

Fiducial cross sections are determined in a variety of phase-space regions sensitive to inclusive Higgs boson production and to explicit Higgs boson production mechanisms. The measurement of these cross sections provides an alternative way to study the properties of the Higgs boson and to search for physics beyond the Standard Model. For each fiducial region of an integrated cross section measurement or bin of a differential distribution, the $H \to \gamma\gamma$ signal is extracted using a fit to the corresponding diphoton invariant mass spectrum. The cross sections are determined by correcting these yields for experimental inefficiencies and resolution effects, and by

Table 6.1: The particle-level kinematic regions of the stage-1 simplified template cross sections, along with the intermediate set of regions used for the measurements presented in this paper. The transverse momentum of the Higgs boson and the leading and subleading jet is denoted as p_{T}^{Hjj} with the subleading jet being the second highest momentum jet in a given event. Events are considered "VBF-like" if they contain at least two jets with an invariant mass of $m_{jj} > 400$ GeV, and a rapidity separation between the two jets of $|\Delta y_{jj}| > 2.8$. Events are considered "VH-like" if they contain at least two jets with an invariant mass of 60 GeV $< m_{jj} < 120$ GeV. All $q\bar{q}' \to Hq\bar{q}'$ VBF and VH events (with the vector boson V decaying hadronically) which are neither VBF nor VH-like are part of the "Rest" selection. For the $p_{\mathrm{T}}^{H} > 200$ GeV gluon-gluon fusion and $p_{\mathrm{T}}^{j} > 200$ GeV VBF + VH regions, only the sum of the corresponding cross sections is reported while the difference of the two is profiled in the fit. In total, the cross sections for nine kinematic regions are measured. The process $gg \to ZH$ refers only to box and loop processes dominated by top and bottom quarks.

Process	Measurement region	Particle-level stage-1 region
$ggH + gg \to Z(\to qq)H$	0-jet	0-jet
	1-jet, $p_{\mathrm{T}}^{H} < 60$ GeV	1-jet, $p_{\mathrm{T}}^{H} < 60$ GeV
	1-jet, $60 \le p_{\mathrm{T}}^{H} < 120$ GeV	1-jet, $60 \le p_{\mathrm{T}}^{H} < 120$ GeV
	1-jet, $120 \le p_{\mathrm{T}}^{H} < 200$ GeV	1-jet, $120 \le p_{\mathrm{T}}^{H} < 200$ GeV
	$\ge$ 1-jet, $p_{\mathrm{T}}^{H} > 200$ GeV	1-jet, $p_{\mathrm{T}}^{H} > 200$ GeV $\ge$ 2-jet, $p_{\mathrm{T}}^{H} > 200$ GeV
	$\ge$ 2-jet, $p_{\mathrm{T}}^{H} < 200$ GeV or VBF-like	$\ge$ 2-jet, $p_{\mathrm{T}}^{H} < 60$ GeV $\ge$ 2-jet, $60 \le p_{\mathrm{T}}^{H} < 120$ GeV $\ge$ 2-jet, $120 \le p_{\mathrm{T}}^{H} < 200$ GeV VBF-like, $p_{\mathrm{T}}^{Hjj} < 25$ GeV VBF-like, $p_{\mathrm{T}}^{Hjj} \ge 25$ GeV
$qq' \to Hqq'$ (VBF + VH)	$p_{\mathrm{T}}^{j} < 200$ GeV	$p_{\mathrm{T}}^{j} < 200$ GeV, VBF-like, $p_{\mathrm{T}}^{Hjj} < 25$ GeV $p_{\mathrm{T}}^{j} < 200$ GeV, VBF-like, $p_{\mathrm{T}}^{Hjj} \ge 25$ GeV $p_{\mathrm{T}}^{j} < 200$ GeV, VH-like $p_{\mathrm{T}}^{j} < 200$ GeV, Rest
	$p_{\mathrm{T}}^{j} > 200$ GeV	$p_{\mathrm{T}}^{j} > 200$ GeV
VH (leptonic decays)	VH leptonic	$q\bar{q} \to ZH$, $p_{\mathrm{T}}^{Z} < 150$ GeV $q\bar{q} \to ZH$, $150 < p_{\mathrm{T}}^{Z} < 250$ GeV, 0-jet $q\bar{q} \to ZH$, $150 < p_{\mathrm{T}}^{Z} < 250$ GeV, $\ge$ 1-jet $q\bar{q} \to ZH$, $p_{\mathrm{T}}^{Z} > 250$ GeV $q\bar{q} \to WH$, $p_{\mathrm{T}}^{W} < 150$ GeV $q\bar{q} \to WH$, $150 < p_{\mathrm{T}}^{W} < 250$ GeV, 0-jet $q\bar{q} \to WH$, $150 < p_{\mathrm{T}}^{W} < 250$ GeV, $\ge$ 1-jet $q\bar{q} \to WH$, $p_{\mathrm{T}}^{W} > 250$ GeV $gg \to ZH$, $p_{\mathrm{T}}^{Z} < 150$ GeV $gg \to ZH$, $p_{\mathrm{T}}^{Z} > 150$ GeV, 0-jet $gg \to ZH$, $p_{\mathrm{T}}^{Z} > 150$ GeV, $\ge$ 1-jet
Top-associated production	top	$t\bar{t}H$ W-associated tH (tHW) t-channel tH ($tHqb$)
$b\bar{b}H$	merged w/ ggH	$b\bar{b}H$

taking into account the integrated luminosity of the data. No attempt is made to separate individual production modes in favor of presenting fiducial regions enriched with a given production mode.

Particle-level fiducial definition

The inclusive fiducial region is defined at the particle level by two photons, not originating from the decay of a hadron, that have absolute pseudorapidity $|\eta| < 2.37$, excluding the region $1.37 < |\eta| < 1.52$, with the leading (subleading) photon transverse momentum greater than 35% (25%) of $m_{\gamma\gamma}$. The two photons are required to be isolated from hadronic activity by imposing that the summed transverse momentum of charged stable particles with $p_{\mathrm{T}} > 1$ GeV, within a cone of $\Delta R = 0.2$ centered on the photon direction, be less than 5% of the photon transverse momentum. This selection is applied to all the presented fiducial integrated and differential cross section results and the isolation criterion is tuned to mimic the detector level selection.

Measurements of cross sections of fiducial integrated regions

Cross sections in five fiducial integrated regions are measured that target either specific Higgs boson production mechanisms or are sensitive to the presence of physics beyond the Standard Model. The selection criteria defining these regions are based on the particle-level fiducial definition previously defined. One additional cross section and three cross section limits are reported in smaller fiducial regions sensitive to specific Higgs boson production mechanisms:

- a VBF-enhanced region with two jets with large invariant mass and rapidity separation,
- a region of events containing at least one charged lepton,

- a region of events with large missing transverse momentum,
- and a region of events with a topology matching the presence of a top-antitop quark pair.

Measurements of differential cross sections

Eleven fiducial differential cross sections are reported and compared to theoretical predictions, for events belonging to the inclusive fiducial region as a function of the following observables:

- $p_{\mathrm{T}}^{\gamma\gamma}$ and $|y^{\gamma\gamma}|$, the transverse momentum and rapidity of the diphoton system,
- $p_{\mathrm{T}}^{j_1}$ and $|y^{j_1}|$, the transverse momentum and rapidity of the leading jet,
- $p_{\mathrm{T}}^{j_2}$ and $|y^{j_2}|$, the transverse momentum and rapidity of the subleading jet,
- $|\cos\theta^*|$, the cosine of the angle between the beam axis and the diphoton system in the Collins-Soper frame [169],
- $\Delta\phi_{jj}$ and $|\Delta y_{jj}|$, the difference in azimuthal angle and in rapidity between the leading and subleading jets,
- $|\Delta\phi_{\gamma\gamma,jj}|$, the difference in azimuthal angle between the dijet system formed by the leading and subleading jets and the diphoton system,
- and m_{jj} the invariant mass of the leading and subleading jets.

Inclusive Higgs boson production is dominated by gluon-gluon fusion, for which the transverse momentum of the Higgs boson is largely balanced by the emission of soft gluons and quarks. Measuring $p_{\mathrm{T}}^{\gamma\gamma}$ probes the perturbative QCD modeling of this production mechanism which is mildly sensitive to the bottom and charm quark Yukawa couplings of the Higgs boson [170]. The distribution at high transverse

momentum is sensitive to new heavy particles coupling to the Higgs boson and to the top quark Yukawa coupling.

The rapidity distribution of the Higgs boson is also sensitive to the modeling of the gluon-gluon fusion production mechanism, as well as to the parton distribution functions (PDFs) of the colliding protons.

6.3 Data samples and MC simulation

The results in this chapter use a data set of 36.1 fb^{-1} of p-p collisions collected in 2015 and 2016 as described in Section 5.2.

Signal samples are generated for the main Higgs boson production modes using Monte Carlo event generators as described in Section 1.3.2. The mass and width of the Higgs boson are set in the simulation to $m_H = 125$ GeV and $\Gamma_H = 4.07$ MeV [33], respectively. The samples are normalized with the latest available theoretical calculations of the corresponding SM production cross sections, as summarized in Section 1.4.1. The normalization of all Higgs boson samples also accounts for the $H \to \gamma\gamma$ branching ratio of 0.227%.

Higgs boson production via ggH is simulated at NNLO accuracy in QCD using the Powheg NNLOPS program [171], with the PDF4LHC15 PDF set [172]. The parton-level events produced by the Powheg NNLOPS program are passed to Pythia8 [173] to provide parton showering, hadronization and underlying event, using the AZNLO set of parameters that are tuned to data [174]. The sample is normalized with the N^{3}LO in QCD and NLO EW cross section [175–178].

Higgs boson production via VBF and VH are generated at NLO accuracy in QCD using the Powheg-Box program [179–182] with the PDF4LHC15 PDF set. Higgs boson production in association with top or bottom quark(s) are generated at NLO accuracy in QCD using MG5_aMC@NLO [183] with the NNPDF3.0 PDF set [26] in

the case of $t\bar{t}H$ and with the CT10 PDF set [184] in the rest of the production modes.

Additional proton-proton interactions (pile-up) are included in the simulation for all generated events such that the average number of interactions per bunch crossing reproduces that observed in the data. A summary of the used signal and background samples is shown in Table 6.2.

Table 6.2: Summary of the event generators and PDF sets used to model the signal and the main background processes. The SM cross sections σ for the Higgs production processes with $m_H = 125.09$ GeV are also given separately for $\sqrt{s} = 13$ TeV, together with the orders of the calculations corresponding to the quoted cross sections, which are used to normalize the samples, after multiplication by the Higgs boson branching ratio to diphotons, 0.227%. The following versions were used: Pythia8 version 8.212 (processes) and 8.186 (pile-up overlay); Herwig++ version 2.7.1; Powheg-Box version 2; MG5_aMC@NLO version 2.4.3; Sherpa version 2.2.1

Process	Generator	Showering	PDF set	σ [pb] $\sqrt{s} = 13$ TeV	Order of calculation of σ
ggH	Powheg NNLOPS	Pythia8	PDF4LHC15	48.52	N^{3}LO(QCD)+NLO(EW)
VBF	Powheg-Box	Pythia8	PDF4LHC15	3.78	NNLO(QCD)+NLO(EW)
WH	Powheg-Box	Pythia8	PDF4LHC15	1.37	NNLO(QCD)+NLO(EW)
$q\bar{q}' \to ZH$	Powheg-Box	Pythia8	PDF4LHC15	0.76	NNLO(QCD)+NLO(EW)
$gg \to ZH$	Powheg-Box	Pythia8	PDF4LHC15	0.12	NLO+NLL(QCD)
$t\bar{t}H$	MG5_aMC@NLO	Pythia8	NNPDF3.0	0.51	NLO(QCD)+NLO(EW)
$b\bar{b}H$	MG5_aMC@NLO	Pythia8	CT10	0.49	5FS(NNLO)+4FS(NLO)
t-channel tH	MG5_aMC@NLO	Pythia8	CT10	0.07	4FS(LO)
W-associated tH	MG5_aMC@NLO	Herwig++	CT10	0.02	5FS(NLO)
$\gamma\gamma$	Sherpa	Sherpa	CT10	-	LO
$V\gamma\gamma$	Sherpa	Sherpa	CT10	-	LO

6.3.1 Alternative MC simulation

A number of alternative signal samples are also generated in order to evaluate theory uncertainties. Signal samples for ggH, VBF and VH are generated by switching the parton showering algorithm from Pythia8 to Herwig7 and from Pythia8 to Herwig++ in the $t\bar{t}H$ sample.

The measured cross sections are compared to a range of theoretical predictions of

gluon-gluon fusion production:

- The parton-level N^{3}LO QCD and NLO EW prediction of [34, 172, 175–178].
- The parton-level JVE+N^{3}LO prediction of [185], which includes NNLL resummation in QCD of the p_{T} of the leading jet which is matched to the N^{3}LO total cross section.
- The parton-level STWZ-BLPTW predictions of [186, 187], which include NNLL'+NNLO resummation for the p_{T} of the leading jet in QCD, combined with a NLL'+NLO resummation in QCD for the subleading jet.[1] The numerical predictions for $\sqrt{s} = 13$ TeV are taken from [34].
- The parton-level NNLOJET prediction of [188, 189] is a fixed-order NNLO prediction in QCD for inclusive H + one-jet production.
- The parton-level GoSam prediction of [190, 191], which provides the fixed-order loop contributions accurate at NLO in QCD in the inclusive H + zero-jet, H + one-jet, H + two-jet, and H + three-jet regions. The real-emission contributions at fixed order in QCD are provided by Sherpa [192].
- The Powheg NNLOPS prediction [171].
- The Sherpa (Meps@Nlo) prediction of [191–202] is accurate to NLO in QCD in the inclusive H + zero-jet, H + one-jet, H + two-jet, and H + three-jet regions and includes top-quark mass effects. The one-loop corrections are incorporated from GoSam [190, 191] and the different jet multiplicity regions are merged using the Meps@Nlo multijet merging technique.

[1] The prime indicates that the leading contributions from N^{3}LL (resp. NNLL) are included along with the full NNLL (resp. NLL) corrections.

- The MG5_aMC@NLO prediction of [183, 203], which includes up to two jets at NLO accuracy using the FxFx merging scheme [204]. The central merging scale is taken to be 30 GeV. The generated events are passed to Pythia8 [173] to provide parton showering and hadronization to create the full final state (without underlying event).

All predictions but NNLOJET and Sherpa (Meps@Nlo) use the NNLO PDF set following the PDF4LHC15 recommendations. The NNLOJET prediction uses the CT14 NNLO PDF set [205] and Sherpa (Meps@Nlo) uses the NNPDF3.0PDF set [26]. GoSam, Sherpa (Meps@Nlo), and NNLOJET apply the kinematic selection on the final-state photons. For all other predictions, the fiducial acceptance is determined using Powheg NNLOPS.

6.4 Object and event selection

6.4.1 Object selection

The majority of the object selection is identical to the one described in Section 5.3. The main differences are described in the following:

- Jets: jets are required to have $p_{\mathrm{T}} > 25$ GeV for $|\eta| < 2.4$ and $p_{\mathrm{T}} > 30$ GeV for $2.4 < |\eta| < 4.4$. The jet selection is tightened to $p_{\mathrm{T}} > 30$ GeV within $|y| < 4.4$ for most event reconstruction categories and the measurement of fiducial integrated and differential cross sections.

- b-jets: the chosen identification criterion has an efficiency of 70% for identifying jets originating from a b-quark.

- Electrons: electron candidates are required to have $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.47$, excluding the region $1.37 < |\eta| < 1.52$. For the measurements of fiducial cross

sections the electron selection is tightened to $p_{\mathrm{T}} > 15$ GeV. Electrons must satisfy the medium identification criteria as well as loose criteria for the calorimeter and track isolation, aimed at a combined efficiency of 99% independently of the candidate transverse momentum.

- Muons: muon candidates are required to have $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.7$, and satisfy medium identification criteria based on the number of hits in the silicon detectors, in the TRT and in the muon spectrometer. For the measurements of fiducial cross sections the muon selection is tightened to $p_{\mathrm{T}} > 15$ GeV. Muon candidates are required to satisfy loose criteria for the calorimeter and track isolation, in this case depending on the candidate transverse momentum, and aimed at a combined efficiency ranging from 95–97% at $p_{\mathrm{T}} = 10$–60 GeV to 99% for $p_{\mathrm{T}} > 60$ GeV.

- Missing E_{T}: the magnitude of the missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$ is measured from the negative vectorial sum of the transverse momenta of all photon, electron and muon candidates and of all hadronic jets after accounting for overlaps between jets, photons, electrons, and muons, as well as an estimate of soft contributions based on tracks originating from the diphoton vertex which satisfy a set of quality criteria. A full description of this algorithm can be found in Section 3.7.

6.4.2 Total production-mode cross sections, signal strengths and simplified template cross sections

Event categorization

The events satisfying the diphoton selection discussed in Section 5.3 are classified into 31 exclusive categories. A combined fit to the event reconstruction categories

is performed to determine nine simplified template cross sections (with $|y_H| < 2.5$), as well as production-mode cross sections and signal strengths interpretations of the data. A summary of the selection requirements defining each category is provided in Table 6.3. In the sequence of the classification, priority is given to categories aimed at selecting signal events from processes with smaller expected cross sections. The fraction of signal events in each reconstructed category is shown in Figure 6.1. The following quantities are used to describe the event selection of the categories defined in Table 6.3:

- N_{lep}: Number of selected electrons and muons.
- $N_{\text{jets}}^{\text{cen}}$: Number of selected central ($|\eta| < 2.5$) jets.
- N_{jets}: Number of selected jets.
- $N_{\text{jets}}^{\text{fwd}}$: Number of selected forward ($|\eta| > 2.5$) jets.
- $N_{b-\text{tag}}$: Number of selected b-tagged jets.
- $Z_{\ell\ell}$ veto: Events with same flavor dilepton candidates with an invariant mass within 10 GeV of the Z boson mass are vetoed.
- BDT_{ttH}: A boosted decision tree (BDT) trained to separate $t\bar{t}H$ signal. The BDT exploits five kinematic variables: the scalar sum of jet transverse momenta H_{T}, the mass of all jets $m_{\text{all jets}}$, the jet multiplicity, the central jet multiplicity and the number of b-tagged jets.
- $m_{e\gamma}$: The invariant mass of the selected electron and any of the two signal photons.
- $p_{\text{T}}^{\ell+E_{\text{T}}^{\text{miss}}}$: The transverse momentum of the combined lepton and $E_{\text{T}}^{\text{miss}}$ system.

- $E_{\mathrm{T}}^{\mathrm{miss}}$ significance: $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{\sum E_{\mathrm{T}}}$.

- $\mathrm{BDT_{VH}}$: A BDT trained to separate VH signal. The BDT uses the following information: the dijet invariant mass, the component $p_{\mathrm{Tt}}^{\gamma\gamma}$ of the diphoton $\vec{p}_{\mathrm{T}}$ transverse to its thrust axis in the transverse plane, the rapidity difference between the dijet and the diphoton system, and the cosine $\cos\theta^*_{\gamma\gamma,jj}$, where $\theta^*_{\gamma\gamma,jj}$ is the angle between the diphoton system's momentum and the direction of motion of the diphoton-dijet system in the Collins-Soper frame.

- p_{T}^{Hjj}: The vector sum of the momenta of the Higgs candidate and the two leading jets.

- $\Delta\eta_{jj}$: The difference in pseudorapidity between the two leading jets.

- $\mathrm{BDT_{VBF}}$: A BDT trained to separate the VBF signal. The BDT exploits six kinematic variables: m_{jj}, $\Delta\eta_{jj}$, $p_{\mathrm{Tt}}^{\gamma\gamma}$, the absolute azimuthal difference of the diphoton and the dijet system $|\Delta\phi_{\gamma\gamma,jj}|$, the minimum angular separation between either of the two signal photons and either of the two leading jets $\Delta R_{\gamma j}^{\mathrm{min}}$, and $|\eta_{\gamma\gamma} - 0.5(\eta_{j1} + \eta_{j2})|$.

6.4.3 Fiducial integrated and differential cross sections

The fiducial volumes are defined to closely mimic the detector-level photon and object selections described in Section 5.3. This reduces the model-dependence of the quoted cross sections in contrast to the per production mode simplified template cross section measurements. The cross sections are determined by correcting measured signal yields for experimental inefficiencies and resolution effects, and by taking into account the integrated luminosity of the data. Fiducial cross sections are determined in a variety of phase-space regions sensitive to inclusive Higgs boson production and to explicit Higgs

Table 6.3: Shorthand label and event selection defining each of the 31 event reconstruction categories for the measurement of the signal strengths and simplified template cross sections. The labels denote the predominant production process or kinematic properties the category targets. Jets are required to have $p_{\mathrm{T}} > 30$ GeV unless otherwise noted. The categories are mutually exclusive and the criteria are applied in descending order of the shown categories.

Category	Selection
tH lep 0fwd	$N_{\mathrm{lep}} = 1$, $N_{\mathrm{jets}}^{\mathrm{cen}} \leq 3$, $N_{b-\mathrm{tag}} \geq 1$, $N_{\mathrm{jets}}^{\mathrm{fwd}} = 0$ ($p_{\mathrm{T}}^{\mathrm{jet}} > 25$ GeV)
tH lep 1fwd	$N_{\mathrm{lep}} = 1$, $N_{\mathrm{jets}}^{\mathrm{cen}} \leq 4$, $N_{b-\mathrm{tag}} \geq 1$, $N_{\mathrm{jets}}^{\mathrm{fwd}} \geq 1$ ($p_{\mathrm{T}}^{\mathrm{jet}} > 25$ GeV)
ttH lep	$N_{\mathrm{lep}} \geq 1$, $N_{\mathrm{jets}}^{\mathrm{cen}} \geq 2$, $N_{b-\mathrm{tag}} \geq 1$, $Z_{\ell\ell}$ veto ($p_{\mathrm{T}}^{\mathrm{jet}} > 25$ GeV)
ttH had BDT1	$N_{\mathrm{lep}} = 0$, $N_{\mathrm{jets}} \geq 3$, $N_{b-\mathrm{tag}} \geq 1$, $\mathrm{BDT}_{\mathrm{ttH}} > 0.92$
ttH had BDT2	$N_{\mathrm{lep}} = 0$, $N_{\mathrm{jets}} \geq 3$, $N_{b-\mathrm{tag}} \geq 1$, $0.83 < \mathrm{BDT}_{\mathrm{ttH}} < 0.92$
ttH had BDT3	$N_{\mathrm{lep}} = 0$, $N_{\mathrm{jets}} \geq 3$, $N_{b-\mathrm{tag}} \geq 1$, $0.79 < \mathrm{BDT}_{\mathrm{ttH}} < 0.83$
ttH had BDT4	$N_{\mathrm{lep}} = 0$, $N_{\mathrm{jets}} \geq 3$, $N_{b-\mathrm{tag}} \geq 1$, $0.52 < \mathrm{BDT}_{\mathrm{ttH}} < 0.79$
tH had 4j1b	$N_{\mathrm{lep}} = 0$, $N_{\mathrm{jets}}^{\mathrm{cen}} = 4$, $N_{b-\mathrm{tag}} = 1$ ($p_{\mathrm{T}}^{\mathrm{jet}} > 25$ GeV)
tH had 4j2b	$N_{\mathrm{lep}} = 0$, $N_{\mathrm{jets}}^{\mathrm{cen}} = 4$, $N_{b-\mathrm{tag}} \geq 2$ ($p_{\mathrm{T}}^{\mathrm{jet}} > 25$ GeV)
VH dilep	$N_{\mathrm{lep}} \geq 2$, 70 GeV $\leq m_{\ell\ell} \leq$ 110 GeV
VH lep High	$N_{\mathrm{lep}} = 1$, $\lvert m_{e\gamma} - 89\ \mathrm{GeV}\rvert > 5$ GeV, $p_{\mathrm{T}}^{\ell+E_{\mathrm{T}}^{\mathrm{miss}}} > 150$ GeV
VH lep Low	$N_{\mathrm{lep}} = 1$, $\lvert m_{e\gamma} - 89\ \mathrm{GeV}\rvert > 5$ GeV, $p_{\mathrm{T}}^{\ell+E_{\mathrm{T}}^{\mathrm{miss}}} < 150$ GeV, $E_{\mathrm{T}}^{\mathrm{miss}}$ significance > 1
VH MET High	150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV, $E_{\mathrm{T}}^{\mathrm{miss}}$ significance > 9 or $E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV
VH MET Low	80 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 150 GeV, $E_{\mathrm{T}}^{\mathrm{miss}}$ significance > 8
jet BSM	$p_{\mathrm{T,j1}} > 200$ GeV
VH had tight	60 GeV $< m_{\mathrm{jj}} <$ 120 GeV, $\mathrm{BDT}_{\mathrm{VH}} > 0.78$
VH had loose	60 GeV $< m_{\mathrm{jj}} <$ 120 GeV, $0.35 < \mathrm{BDT}_{\mathrm{VH}} < 0.78$
VBF tight, high p_{T}^{Hjj}	$\lvert\Delta\eta_{jj}\rvert > 2$, $\lvert\eta_{\gamma\gamma} - 0.5(\eta_{\mathrm{j1}} + \eta_{\mathrm{j2}})\rvert < 5$, $p_{\mathrm{T}}^{Hjj} > 25$ GeV, $\mathrm{BDT}_{\mathrm{VBF}} > 0.47$
VBF loose, high p_{T}^{Hjj}	$\lvert\Delta\eta_{jj}\rvert > 2$, $\lvert\eta_{\gamma\gamma} - 0.5(\eta_{\mathrm{j1}} + \eta_{\mathrm{j2}})\rvert < 5$, $p_{\mathrm{T}}^{Hjj} > 25$ GeV, $-0.32 < \mathrm{BDT}_{\mathrm{VBF}} < 0.47$
VBF tight, low p_{T}^{Hjj}	$\lvert\Delta\eta_{jj}\rvert > 2$, $\lvert\eta_{\gamma\gamma} - 0.5(\eta_{\mathrm{j1}} + \eta_{\mathrm{j2}})\rvert < 5$, $p_{\mathrm{T}}^{Hjj} < 25$ GeV, $\mathrm{BDT}_{\mathrm{VBF}} > 0.87$
VBF loose, low p_{T}^{Hjj}	$\lvert\Delta\eta_{jj}\rvert > 2$, $\lvert\eta_{\gamma\gamma} - 0.5(\eta_{\mathrm{j1}} + \eta_{\mathrm{j2}})\rvert < 5$, $p_{\mathrm{T}}^{Hjj} < 25$ GeV, $0.26 < \mathrm{BDT}_{\mathrm{VBF}} < 0.87$
ggH 2J BSM	≥ 2 jets, $p_{\mathrm{T}}^{\gamma\gamma} \geq 200$ GeV
ggH 2J High	≥ 2 jets, $p_{\mathrm{T}}^{\gamma\gamma} \in$ [120, 200] GeV
ggH 2J Med	≥ 2 jets, $p_{\mathrm{T}}^{\gamma\gamma} \in$ [60, 120] GeV
ggH 2J Low	≥ 2 jets, $p_{\mathrm{T}}^{\gamma\gamma} \in$ [0, 60] GeV
ggH 1J BSM	$= 1$ jet, $p_{\mathrm{T}}^{\gamma\gamma} \geq 200$ GeV
ggH 1J High	$= 1$ jet, $p_{\mathrm{T}}^{\gamma\gamma} \in$ [120, 200] GeV
ggH 1J Med	$= 1$ jet, $p_{\mathrm{T}}^{\gamma\gamma} \in$ [60, 120] GeV
ggH 1J Low	$= 1$ jet, $p_{\mathrm{T}}^{\gamma\gamma} \in$ [0, 60] GeV
ggH 0J Fwd	$= 0$ jets, one photon with $\lvert\eta\rvert > 0.95$
ggH 0J Cen	$= 0$ jets, two photons with $\lvert\eta\rvert \leq 0.95$

Figure 6.1: The expected composition of the selected Higgs boson events, in terms of the different production modes, for each reconstructed category [4].

boson production mechanisms. The measurement of these cross sections provides an alternative way to study the properties of the Higgs boson and to search for physics beyond the Standard Model. For each fiducial region of an integrated cross section measurement or bin of a differential distribution, the $H \to \gamma\gamma$ signal is extracted using a fit to the corresponding diphoton invariant mass spectrum. The cross sections are

determined by correcting these yields for experimental inefficiencies and resolution effects, and by taking into account the integrated luminosity of the data. No attempt is made to separate individual production modes in favour of presenting fiducial regions enriched with a given production mode.

Particle-level fiducial definition

The fiducial volume at particle level is defined using particles with a mean lifetime $c\tau > 10\,\text{mm}$. The photon definition is described in Section 6.2. The lepton four-momentum is defined as the combination of an electron (or muon) and all nearby photons within $\Delta R < 0.1$ that do not originate from the decay of a hadron. Muons are required to have $p_{\text{T}} > 15$ GeV and $|\eta| < 2.7$. Electrons are required to have $p_{\text{T}} > 15$ GeV and $|\eta| < 2.47$, excluding the region $1.37 < |\eta| < 1.52$, and are rejected if the distance ΔR to a photon with $p_{\text{T}} > 15$ GeV is less than 0.4. Jets are reconstructed from all particles, excluding muons and neutrinos, using the anti-k_{T} algorithm with a radius parameter of 0.4. Jets are required to have $p_{\text{T}} > 30$ GeV, $|y| < 4.4$ and to be well separated from photons with $p_{\text{T}} > 15$ GeV ($\Delta R > 0.4$) and electrons ($\Delta R > 0.2$). Jets are considered to originate from a b-hadron if there is a b-hadron with $p_{\text{T}} > 5$ GeV within a cone of size $\Delta R = 0.4$ around the jet. The missing transverse momentum is defined as the vector sum of neutrino transverse momenta, for neutrinos that do not originate from the decay of a hadron. The particle-level fiducial definition is summarized in Table 6.4.

The cross section (σ_i) in a fiducial integrated region, and the differential cross section ($\mathrm{d}\sigma_i/\mathrm{d}x$) in a bin of variable x, are given by

$$\sigma_i = \frac{N_i^{\text{sig}}}{c_i \int L\,\mathrm{d}t} \qquad \text{and} \qquad \frac{\mathrm{d}\sigma_i}{\mathrm{d}x} = \frac{N_i^{\text{sig}}}{c_i\,\Delta x_i \int L\,\mathrm{d}t}, \tag{6.2}$$

where N_i^{sig} is the number of signal events, $\int L\,\mathrm{d}t$ is the integrated luminosity of the

Table 6.4: Summary of the particle-level definitions of the five fiducial integrated regions described in the text. The photon isolation $p_{\mathrm{T}}^{\mathrm{iso},0.2}$ is defined analogously to the reconstructed-level track isolation as the transverse momentum of the system of charged particles within $\Delta R < 0.2$ of the photon.

Objects	Definition
Photons	$\lvert\eta\rvert < 1.37$ or $1.52 < \lvert\eta\rvert < 2.37$, $p_{\mathrm{T}}^{\mathrm{iso},0.2}/p_{\mathrm{T}}^{\gamma} < 0.05$
Jets	anti-k_{T}, $R = 0.4$, $p_{\mathrm{T}} > 30$ GeV, $\lvert y\rvert < 4.4$
Leptons, ℓ	e or μ, $p_{\mathrm{T}} > 15$ GeV, $\lvert\eta\rvert < 2.47$ for e (excluding $1.37 < \lvert\eta\rvert < 1.52$) and $\lvert\eta\rvert < 2.7$ for μ
Fiducial region	**Definition**
Diphoton fiducial	$N_{\gamma} \geq 2$, $p_{\mathrm{T}}^{\gamma_1} > 0.35\, m_{\gamma\gamma} = 43.8$ GeV, $p_{\mathrm{T}}^{\gamma_2} > 0.25\, m_{\gamma\gamma} = 31.3$ GeV
VBF-enhanced	Diphoton fiducial, $N_j \geq 2$, $m_{jj} > 400$ GeV, $\Delta y_{jj} > 2.8$, $\lvert\Delta\phi_{\gamma\gamma,jj}\rvert > 2.6$
$N_{\mathrm{lepton}} \geq 1$	Diphoton fiducial, $N_{\ell} \geq 1$
High $E_{\mathrm{T}}^{\mathrm{miss}}$	Diphoton fiducial, $E_{\mathrm{T}}^{\mathrm{miss}} > 80$ GeV, $p_{\mathrm{T}}^{\gamma\gamma} > 80$ GeV
$t\bar{t}H$-enhanced	Diphoton fiducial, ($N_j \geq 4$, $N_{b\text{-jets}} \geq 1$) or ($N_j \geq 3$, $N_{b\text{-jets}} \geq 1, N_{\ell} \geq 1$)

data set, c_i is a correction factor that accounts for detector inefficiency and resolution, and Δx_i is the bin width. The correction factors are determined using the simulated samples discussed in Section 6.3. This bin-by-bin unfolding method showed similar performance to that of the non-regularized inversion of the full migration matrix and of regularized methods [206–208] within the current statistical accuracy and systematic uncertainties. A description of unfolding is given in Appendix A.

6.5 Signal and background modeling

The Higgs boson signal yield is measured through an unbinned maximum-likelihood fit to the diphoton invariant mass spectrum in the range 105 GeV $< m_{\gamma\gamma} <$ 160 GeV for each event reconstruction category, fiducial region, or each bin of a fiducial differential cross section. The mass range is chosen to be large enough to allow a reliable determination of the background from the data, and at the same time small enough to avoid large uncertainties from the choice of the background parametrization. The signal and background shapes are modeled as described below, and the background model parameters are freely floated in the fit to the $m_{\gamma\gamma}$ spectra.

6.5.1 Signal model

The Higgs boson signal manifests itself as a narrow peak in the $m_{\gamma\gamma}$ spectrum. The signal distribution is modeled by the double-sided Crystal Ball function as described in Section 5.4.2.

The parameters of the model that define the shape of the signal distribution are determined through fits to the simulated signal samples. The parametrization is derived separately for each reconstructed category or fiducial region of the integrated or differential cross section measurement.

6.5.2 Background composition and model

The diphoton invariant mass model for the background used to fit the data is determined from studies of the bias in the signal yield in signal-plus-background fits to large control samples of data or simulated background events.

Continuum $\gamma\gamma$ production is simulated with the SHERPA event generator as explained in Section 6.3. The relative fractions of $\gamma\gamma$, γj and jj components contributing to the overall background are found by using the 2x2D sideband method (see Section 5.4.1) in 1 GeV bins of $m_{\gamma\gamma}$. Linear functions are then fitted to these fractions as a function of $m_{\gamma\gamma}$. The $\gamma\gamma$ sample is then used to produce the background only template but is reweighted according to the linear functions derived.

The measurement of the background fractions in data is performed for each category or fiducial region. The fractions of these background sources in the inclusive diphoton sample are $(78.7\,^{+1.8}_{-5.2})\%$, $(18.6\,^{+4.2}_{-1.6})\%$ and $(2.6\,^{+0.5}_{-0.4})\%$, respectively. The uncertainties in the measured background fractions are systematically dominated. The main sources originate from the definition of the background control regions.

The functional form used to model the background $m_{\gamma\gamma}$ distribution in the fit to the data is chosen, in each region, to ensure a small bias in the extracted signal yield

relative to its experimental precision, following the procedure described in [37]. The potential bias (spurious signal) is estimated as the maximum of the absolute value of the fitted signal yield, using a signal model with mass between 121 and 129 GeV, in fits to background control regions.

The spurious signal is required, at 95% CL, to be less than 10% of the expected SM signal yield or less than 20% of the expected statistical uncertainty in the SM signal yield. In the case when two or more functions satisfy those requirements, the background model with the least number of parameters is chosen. The background distribution of all regions is found to be well modeled by at least one of the following functions: an exponential of a first- or second-order polynomial, a power law, or a third-order Bernstein polynomial. These functions are detailed in Sections 5.4.2 and 5.4.3.

6.6 Statistical model

The statistical model is similar to that of Section 5.6. It is detailed again here, with a stronger emphasis on systematic uncertainties since they have a bigger impact in this analysis.

An extended likelihood function is built from the number of observed events and invariant diphoton mass values of the observed events using the analytic functions describing the distributions of $m_{\gamma\gamma}$ in the range 105–160 GeV for the signal and the background.

The likelihood for a given reconstructed category, fiducial region, or differential bin i of the integrated or differential cross section measurement is a marked Poisson

probability distribution,

$$\mathcal{L}_i = \text{Pois}(n_i|N_i(\theta)) \cdot \prod_{j=1}^{n_i} f_i(m_{\gamma\gamma}^j, \theta) \cdot G(\theta), \tag{6.3}$$

where n_i (N_i) is the observed (expected) number of selected candidates, $f_i(m_{\gamma\gamma}^j)$ is the value of the probability density function (pdf) of the invariant mass distribution evaluated for each candidate j, θ are nuisance parameters and $G(\theta)$ is a set of unit Gaussian constraints on a subset of the nuisance parameters. The likelihood for the measurements of the total Higgs boson production-mode cross sections and signal strengths is given by the product of the likelihood functions of each event reconstruction category. For the fiducial integrated and differential cross section measurements the likelihood of all bins i in a spectrum is taken.

The number of expected candidates is the sum of the signal and background yields, denoted by N_i^{sig} and N_i^{bkg}, and the fitted spurious signal yield, $N_i^{\text{spur}} \cdot \theta_i^{\text{spur}}$,

$$N_i = N_i^{\text{sig}} + N_i^{\text{bkg}} + N_i^{\text{spur}} \cdot \theta_i^{\text{spur}}. \tag{6.4}$$

In more detail, the invariant mass distribution for each category has signal and background components,

$$f_i(m_{\gamma\gamma}^j) = \left[(N_i^{\text{sig}} + N_i^{\text{spur}} \cdot \theta_i^{\text{spur}}) \cdot f_i^{\text{sig}}(m_{\gamma\gamma}^j, \theta_i^{\text{sig}}) + N_i^{\text{bkg}} \cdot f_i^{\text{bkg}}(m_{\gamma\gamma}^j, \theta_i^{\text{bkg}}) \right] / N_i, \tag{6.5}$$

where θ_i^{sig} and θ_i^{bkg} are nuisance parameters associated with systematic uncertainties affecting the resolutions and positions (see Section 6.7.1) of the invariant mass distributions of the signal f_i^{sig} or the shape of the background f_i^{bkg}.

Systematic uncertainties are incorporated into the likelihood function by multi-

plying the relevant parameter of the statistical model by a factor

$$F_{\mathrm{G}}(\sigma, \theta) = (1 + \sigma \cdot \theta) \tag{6.6}$$

in the case of a Gaussian pdf for the effect of an uncertainty of size σ or, for cases where a negative model parameter does not make physical sense (e.g. the uncertainty in the integrated luminosity), by a factor

$$F_{\mathrm{LN}}(\sigma, \theta) = e^{\sqrt{\ln(1+\sigma^2)}\theta} \tag{6.7}$$

for a log-normal pdf. In both cases the corresponding component of the constraint product $G(\theta)$ is a unit Gaussian centered at zero for the nuisance parameter θ. The systematic uncertainties affecting the yield and mass resolution use the log-normal form while a Gaussian form is used for all others. When two uncertainties are considered fully correlated they share the same nuisance parameter. Systematic uncertainties with partial correlations are decomposed into their uncorrelated and fully correlated components before being assigned to nuisance parameters.

All measured Higgs boson signal yields are determined with the profile likelihood ratio test statistic

$$\Lambda(\nu) = -2 \ln \frac{\mathcal{L}(\nu, \hat{\theta}_\nu)}{\mathcal{L}(\hat{\nu}, \hat{\theta})}, \tag{6.8}$$

where $\hat{\nu}$ and $\hat{\theta}$ are the values of the parameter of interest (e.g. a signal strength or a simplified template cross section) and nuisance parameters that unconditionally maximize the likelihood while $\hat{\theta}_\nu$ are the values of the nuisance parameters that maximize the likelihood on the condition that the parameter of interest is held fixed to a given value ν. In the asymptotic approximation, which is valid for all the results presented in this chapter, $\Lambda(\nu)$ may be interpreted as an increase in χ^2 from its minimum value [148] such that approximate confidence intervals are easily constructed.

The total uncertainty in ν is thus obtained from the ν values such that $\Lambda(\nu) = 1$ with all other parameters "profiled" (*i.e.* set to the values that maximize the likelihood for those values of ν). Theory uncertainties in the parameters of interest are found by fixing the nuisance parameters associated with experimental uncertainties and subtracting in quadrature the statistical uncertainty. The statistical uncertainty is similarly determined, by fixing all nuisance parameters to their best-fit values, except for those describing the background shape and normalization. The experimental uncertainty is found by subtracting in quadrature the theory and the statistical uncertainties from the total uncertainty.

6.7 Systematic uncertainties

Several sources of systematic uncertainty are considered in this measurement, which are extensively covered in [209, 210]. They can be grouped into three categories:

- Uncertainties associated with the parametrization of the signal and background when fitting the $m_{\gamma\gamma}$ spectrum.
- Experimental uncertainties arising either from the extraction of the signal in a given category or from migrations between categories.
- Theoretical and modeling uncertainties in each category, causing migrations between categories, or affect the fiducial acceptance.

The origin of the uncertainties and their treatment are discussed in detail below and summarized in Table 6.5.

The analysis based on event reconstruction categories and those of fiducial cross sections treat yield and migration uncertainties differently: whereas the former incorporate them directly into the likelihood function, the latter incorporate them at a

later stage as part of the correction factor or the luminosity. Details on the theoretical and modeling uncertainties are given in Section 6.7.3.

Table 6.5: Summary of the sources of systematic uncertainties for results based on event reconstruction categories or fiducial integrated and differential cross sections. The columns labels "Category Likelihood" and "Fiducial Likelihood" provide an overview about which terms are part of the Likelihood (✓) or incorporated at a later stage (-). Both sets of results incorporate uncertainties associated with the Higgs boson mass, photon energy scale and resolution, and uncertainties associated with the choice of the background function into the likelihood function, either using log normal ($F_{\mathrm{LN}}(\sigma_i, \theta_i)$) or Gaussian constraints ($F_{\mathrm{G}}(\sigma_i, \theta_i)$) with σ_i denoting the systematic uncertainty (i is the index to each of the unique nuisance parameters θ). When acting on $N_{\mathrm{S}}^{\mathrm{tot}}$ the uncertainty value is the same for all processes, whereas the uncertainty has a different value for each signal process for the case denoted $N_{\mathrm{S}}^{\mathrm{p}}$. The number of nuisance parameters, N_{NP}, for the spurious signal uncertainty varies, e.g. for the category-based results 31 independent error sources are present and for the differential measurements one source per measured bin is included.

		Systematic uncertainty source	N_{NP}	Constraint	Category Likelihood	Fiducial Likelihood
Theory		ggH QCD	9	$N_{\mathrm{S}}^{\mathrm{ggH}}\, F_{\mathrm{LN}}(\sigma_i, \theta_i)$	✓	-
		Missing higher orders (non-ggH)	6	$N_{\mathrm{S}}^{\mathrm{p}}\, F_{\mathrm{LN}}(\sigma_i, \theta_i)$	✓	-
		$\mathcal{B}(H \to \gamma\gamma)$	1	$N_{\mathrm{S}}^{\mathrm{tot}}\, F_{\mathrm{LN}}(\sigma_i, \theta_i)$	✓	-
		PDF	30	$N_{\mathrm{S}}^{\mathrm{p}}\, F_{\mathrm{LN}}(\sigma_i, \theta_i)$	✓	-
		α_{S}	1	$N_{\mathrm{S}}^{\mathrm{p}}\, F_{\mathrm{LN}}(\sigma_i, \theta_i)$	✓	-
		Underlying Event / Parton Shower	5	$N_{\mathrm{S}}^{\mathrm{p}}\, F_{\mathrm{LN}}(\sigma_i, \theta_i)$	✓	-
Experimental	Yield	Heavy flavor content	1	$N_{\mathrm{S}}^{\mathrm{p}}\, F_{\mathrm{LN}}(\sigma_i, \theta_i)$	✓	-
		Luminosity	1	$N_{\mathrm{S}}^{\mathrm{tot}}\, F_{\mathrm{LN}}(\sigma_i, \theta_i)$	✓	-
		Trigger	1	$N_{\mathrm{S}}^{\mathrm{tot}}\, F_{\mathrm{LN}}(\sigma_i, \theta_i)$	✓	-
		Photon identification	1	$N_{\mathrm{S}}^{\mathrm{p}}\, F_{\mathrm{LN}}(\sigma_i, \theta_i)$	✓	-
		Photon isolation	2	$N_{\mathrm{S}}^{\mathrm{p}}\, F_{\mathrm{LN}}(\sigma_i, \theta_i)$	✓	-
	Migration	Flavor tagging	14	$N_{\mathrm{S}}^{\mathrm{p}}\, F_{\mathrm{LN}}(\sigma_i, \theta_i)$	✓	-
		Jet energy scale	20	$N_{\mathrm{S}}^{\mathrm{p}}\, F_{\mathrm{LN}}(\sigma_i, \theta_i)$	✓	-
		Jet flavor composition	7	$N_{\mathrm{S}}^{\mathrm{p}}\, F_{\mathrm{LN}}(\sigma_i, \theta_i)$	✓	-
		Jet flavor response	7	$N_{\mathrm{S}}^{\mathrm{p}}\, F_{\mathrm{LN}}(\sigma_i, \theta_i)$	✓	-
		Electron	3	$N_{\mathrm{S}}^{\mathrm{p}}\, F_{\mathrm{LN}}(\sigma_i, \theta_i)$	✓	-
		Muon	11	$N_{\mathrm{S}}^{\mathrm{p}}\, F_{\mathrm{LN}}(\sigma_i, \theta_i)$	✓	-
		Missing transverse momentum	3	$N_{\mathrm{S}}^{\mathrm{p}}\, F_{\mathrm{LN}}(\sigma_i, \theta_i)$	✓	-
		Pileup	1	$N_{\mathrm{S}}^{\mathrm{p}}\, F_{\mathrm{LN}}(\sigma_i, \theta_i)$	✓	-
		Photon energy scale	40	$N_{\mathrm{S}}^{\mathrm{p}}\, F_{\mathrm{LN}}(\sigma_i, \theta_i)$	✓	-
Mass		ATLAS-CMS m_H	1	$\mu_{\mathrm{CB}}\, F_{\mathrm{G}}(\sigma_i, \theta_i)$	✓	✓
		Photon energy scale	40	$\mu_{\mathrm{CB}}\, F_{\mathrm{G}}(\sigma_i, \theta_i)$	✓	✓
		Photon energy resolution	9	$\sigma_{\mathrm{CB}}\, F_{\mathrm{LN}}(\sigma_i, \theta_i)$	✓	✓
Background		Spurious signal	Varies	$N_{\mathrm{spur},c}\, \theta_{\mathrm{spur},c}$	✓	✓

6.7.1 Systematic uncertainties in the signal and background modeling from fitting the $m_{\gamma\gamma}$ spectrum

Systematic uncertainties associated with the signal and background parametrizations are treated in a similar way for all the measurements. These include systematic uncertainties in the photon energy scale and resolution, and the uncertainties due to the specific choice of background model.

The uncertainties in the photon energy scale and resolution impact the signal model, as the photon energy scale shifts the position of the peak by between 0.21% and 0.36%, and the assumed energy resolution broadens or narrows the signal peak relative to its nominal width by between 6% and 13%. The size of both uncertainties is dependent on the energy, rapidity and jet activity of the selected photon pair.

An additional uncertainty in the signal peak position is added as a nuisance parameter in the fit, reflecting the uncertainty in the measurement of the Higgs boson mass of 0.24 GeV [211]. A variation of the signal mass by this quantity is found to impact the measured global signal strength or the diphoton fiducial cross section by less than 0.1%.

The uncertainty due to the choice of background function is taken to be the spurious signal yield obtained when fitting the $m_{\gamma\gamma}$ spectrum reconstructed from background-only simulated samples (or signal-suppressed control regions in data), as discussed in Section 6.5.2.

6.7.2 Experimental systematic uncertainties affecting the expected event yields

There are two categories of uncertainties: (i) those in the expected overall signal yield and (ii) those that cause migrations of events between categories and bins, as well as into and out of the photon fiducial selection.

The sources of uncertainties in the expected signal yield consist of: the luminosity delivered to the ATLAS experiment [157] in 2015 and 2016 (3.2%), the efficiency of the diphoton trigger [158] (0.4%), the photon identification efficiency (1.6%), the photon track isolation efficiency (0.8%) and the photon calorimeter isolation efficiency (0.1%).

Uncertainties which affect the calibration of photons, jets, and leptons cause migrations between categories and bins, as well as migrations into and out of the fiducial acceptance. Their sources include: the modeling of pileup in the simulation (1.4–5.6%), the photon energy scale and resolution (0.2–1.9%, depending on the diphoton p_{T}), the jet energy calibration and the jet energy resolution (2.8–15%), the efficiency of the jet vertex tagger (0–0.3%), the efficiency of the b-tagging algorithm (3%), the electron and muon reconstruction, identification and isolation efficiencies (0.6% and 0.5%, respectively), the electron and muon energy and momentum scale and resolution (negligible for all measurements) and the energy scales and resolutions of photons, jets and leptons propagated to the missing transverse momentum uncertainty (4.0–4.8%).

6.7.3 Theoretical and modeling uncertainties

The overall theoretical cross section uncertainties affect the signal strength measurements, which are ratios of the observed to predicted event yields, but not the cross section measurements which do not rely on absolute predictions. Modeling uncertainties that alter the kinematic properties of the events, such as the Higgs boson transverse momentum or the jet multiplicity, have an impact on both types of measurements. These uncertainties are estimated with different approaches as discussed further in Sections 6.7.4 and 6.7.5.

6.7.4 Theoretical and modeling uncertainties for results based on event reconstruction categories

PDF acceptance uncertainties

The theoretical modeling uncertainties in the per-category acceptance of each production process affect the measurement of production-mode cross sections and signal strengths. Uncertainties due to the choice of parton distribution functions and the value of α_S are estimated by reweighting the MC samples to the twenty-nine alternative PDF variations and to the two α_S variations from the nominal PDF4LHC_NLO_30_AS PDF set, following the PDF4LHC15 recommendations [172].

Perturbative QCD uncertainties

The perturbative QCD uncertainties are evaluated by varying the factorization and renormalization (QCD) scales. They are calculated for each production mode and provided by the LHC Higgs boson cross section working group [34]. For the gluon-gluon fusion process, the total production-mode cross section has been calculated at N^3LO precision in QCD and has an uncertainty of 3.9%, as determined by QCD-scale variations and including top, bottom, and charm quark mass effect uncertainties. However, the perturbative uncertainty becomes significantly larger in different kinematic regions, e.g. when requiring additional jets or high Higgs boson p_T. To take this effect into account nine uncertainty sources are included:

- Four sources [34] account for uncertainties in the jet multiplicities due to missing higher-order corrections: two accounting for yield uncertainties (with uncertainties up to 8.9% in each STXS region) and two accounting for migrations between jet multiplicity bins (with uncertainties up to 18% in each STXS region), using the STWZ [186] and BLPTW [186, 187, 212] predictions as an input.

- Three uncertainty sources parameterize modeling uncertainties in the Higgs boson p_{T}. The first two encapsulate the migration uncertainty between the intermediate and high p_{T} region with events with at least one jet. The third uncertainty parameterizes top-quark mass effects in the gluon-gluon fusion loop, where the difference between the LO and NLO predictions is taken as an uncertainty due to missing higher-order corrections. This introduces a negligible uncertainty at low Higgs boson p_{T} and a sizable uncertainty of the order of 30% at $p_{\mathrm{T}} > 500$ GeV.

- Two sources account for the uncertainty in the acceptance of gluon-gluon fusion events in the VBF categories, due to missing QCD higher-orders in the calculation. The uncertainty estimation uses an extension of the Stewart-Tackmann method [213, 214] and typically ranges between 20% and 32%.

Underlying event, parton shower and hadronization uncertainties

The uncertainty in the modeling of the parton shower, underlying event and hadronization affects all measurements (referred to as "Underlying Event / Parton Shower" in the following). It is estimated by taking the difference in acceptance after switching the parton showering algorithm from Pythia8 to Herwig7 in the ggH, VBF, and VH samples at truth level, and from Pythia8 to Herwig++ in the $t\bar{t}H$ sample at reconstruction level, respectively. These differences are treated as four independent uncertainty sources. Additionally, for ggH the effect of the eigenvector tunes from the AZNLO set are merged to provide one additional uncertainty component.

In the case of the ggH, VBF, and VH samples, the relative acceptance between the different parton showering algorithms is estimated at truth level by mimicking the reconstructed $H \to \gamma\gamma$ category selections. Identification efficiencies are ignored and BDTs are evaluated using truth level quantities. In all cases, the Underlying Event /

Parton Shower uncertainty for each category is obtained as the absolute value of the difference in acceptance. Events with Dalitz decays are removed in order to avoid a possible bias in the estimate of the uncertainties when comparing the two different parton showering algorithms. The corresponding cross section is applied and results are normalized to 36.1 fb^{-1} of integrated luminosity.

In Table 6.6 the ggH and VBF differences in acceptance between Herwig7 and Pythia8 (AZNLO) are shown. These results give an Underlying Event / Parton Shower uncertainty of $0.66 \pm 0.18\%$ for the inclusive fiducial region when using the NNLOPS ggH sample and of $0.09 \pm 0.18\%$ for the inclusive fiducial region when using the VBF sample.

Analogously, the VH (ZH, $ggZH$, W^-H and W^+H) differences in acceptance between Herwig7 and Pythia8 (AZNLO) are shown in Table 6.7. The results give an Underlying Event / Parton Shower uncertainty of $0.1 \pm 0.5\%$ for the inclusive fiducial region when using the ZH sample, of $0.7 \pm 0.8\%$ when using the $ggZH$ sample, of $0.8 \pm 0.6\%$ when using the W^-H sample and of $0.7 \pm 0.7\%$ when using the W^+H sample.

The Underlying Event / Parton Shower uncertainties in $t\bar{t}H$ are extracted similarly to those above, however showering of $t\bar{t}H$ is done with Herwig++ in place of Herwig7. These uncertainties are estimated at reconstructed level due to the availability of each sample with full simulation. The differences in acceptance between Herwig++ and Herwig7 for the $t\bar{t}H$ sample are reported in Table 6.8.

The eigentune variations from the AZNLO set are studied for ggH only. The differences in acceptance for the ISR variations are shown in Table 6.9. Table 6.10 includes analogue results for the FSR and MPI eigentune variations. It is found that the MPI and ISR variations are significantly smaller than those from FSR. The uncertainties from FSR variations are comparable to those from switching the parton

shower model from PYTHIA8 to HERWIG7.

The biggest challenge in determining accurate estimates of the Underlying Event / Parton Shower uncertainties is obtaining large enough MC samples to have sufficient events in all category selections. For example, in the VBF tight, low p_{T}^{Hjj} category, which is the category most pure in VBF production, the difference in acceptance for ggH production when switching from PYTHIA8 to HERWIG7 is $7.9 \pm 4.8\%$.

Comparison of higher-order FxFx and NNLOPS *ggH*

Acceptance uncertainties are obtained in the VBF and VH categories for the SM ggH process by using the sample corresponding to the MG5_AMC@NLO FxFx process and the nominal POWHEG NNLOPS ggH sample. Because the FxFx merging scheme merges NLO calculations at 0, 1, and ≥ 2 jets, this study evaluates the sensitivity of the dijet BDT classifiers to higher-order effects, as NNLOPS produces dijet predictions at LO. Only shape differences in the BDT response are considered as uncertainty.

In Table 6.11 the BDT acceptance uncertainties and a comparison with the QCD uncertainties for the relevant categories are shown. It is found that the BDT acceptance uncertainties for each generator are below the level of the QCD uncertainties, meaning that an additional shape uncertainty is not necessary to account for this discrepancy.

Heavy flavor content uncertainties in $t\bar{t}H$ categories

In the $t\bar{t}H$ categories, a conservative uncertainty of 100% is assigned to the expected number of events from ggH, VBF and WH production in association with additional b-jets. These uncertainties are included to account for the Higgs boson + heavy flavor content uncertainty and are motivated by the ggH and VBF Underlying Event / Parton Shower studies, as seen in Table 6.6. There are very few ggH events entering

	ggH		**VBF**	
Categories	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
ggH 0J Cen	-1.5	0.4	21.5	1.1
ggH 0J Fwd	-1.4	0.3	20.1	0.8
ggH 1J Low	2.0	0.4	0.3	0.5
ggH 1J Med	-6.4	0.5	7.3	0.5
ggH 1J High	-6.4	1.1	9.9	0.9
ggH 1J BSM	-3	5	21	4
ggH 2J Low	18.0	1.1	-10.3	0.7
ggH 2J Med	4.7	0.9	-6.9	0.7
ggH 2J High	-0.1	1.2	-0.9	1.0
ggH 2J BSM	-3	2	3	2
VBF loose, low p_{T}^{Hjj}	6	3	-10.9	0.5
VBF tight, low p_{T}^{Hjj}	8	5	-10.7	0.5
VBF loose, high p_{T}^{Hjj}	4	2	-1.1	1.1
VBF tight, high p_{T}^{Hjj}	6	2	0.3	0.7
VH had loose	2.8	1.7	7	2
VH had tight	-1	2	7	3
jet BSM	-1.7	1.4	-5.3	0.7
VH MET Low	-32	19	10	17
VH MET High	27	116	7	60
VH lep Low	-	-	-	-
VH lep High	-	-	-	-
VH dilep	-	-	-	-
tH had 4j2b	-16	12	-55	7
tH had 4j1b	-16	8	-54	6
ttH had BDT4	-23	12	-65	7
ttH had BDT3	34	47	-62	21
ttH had BDT2	-39	19	-77	11
ttH had BDT1	-21	35	-89	12
ttH lep	27	116	93	272
tH lep 1fwd	-	-	20	49
tH lep 0fwd	-19	31	6	40

Table 6.6: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the NNLOPS ggH and VBF $H \to \gamma\gamma$ samples for the different event reconstruction categories, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of ggH and VBF events passing each category.

	ZH		$ggZH$		W^-H		W^+H	
Categories	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
ggH 0J Cen	1	2	-6	7	-1	2	-1	3
ggH 0J Fwd	1.0	1.7	-23	5	0.6	1.7	-3.4	1.8
ggH 1J Low	0.7	1.5	-12	5	-0.9	1.6	0.5	1.8
ggH 1J Med	2.4	1.7	-12	3	4.4	1.9	2	2
ggH 1J High	-1	3	-14	4	-2	4	1	4
ggH 1J BSM	30	16	-11	15	7	17	-6	15
ggH 2J Low	-0.1	1.6	14	4	-2	2	-2	2
ggH 2J Med	1.1	1.8	12	3	0	2	0	2
ggH 2J High	4	3	11	3	5	4	1	4
ggH 2J BSM	6	6	4	5	-6	7	0	7
VBF loose, low p_{T}^{Hjj}	-5	13	46	17	-35	12	15	18
VBF tight, low p_{T}^{Hjj}	-66	21	14	24	0	38	77	70
VBF loose, high p_{T}^{Hjj}	2	6	41	7	0	7	3	8
VBF tight, high p_{T}^{Hjj}	12	8	35	6	3	9	4	10
VH had loose	-1.8	1.7	-9	3	-2	2	0	2
VH had tight	-6.2	1.6	-11	2	-9	2	-4	2
jet BSM	0	3	-4	3	-2	3	-4	3
VH MET Low	-4	2	0	3	2	3	3	3
VH MET High	1	3	-2	3	-6	8	-4	8
VH lep Low	-	-	-	-	-	-	-	-
VH lep High	-	-	-	-	-	-	-	-
VH dilep	-	-	-	-	-	-	-	-
tH had 4j2b	5	9	22	10	-65	27	-6	48
tH had 4j1b	2	9	-17	10	-4	20	-33	16
ttH had BDT4	10	9	-9	8	-14	23	-10	24
ttH had BDT3	8	24	-37	16	-60	43	155	171
ttH had BDT2	-2	15	-10	14	-35	35	99	80
ttH had BDT1	7	25	-80	10	-68	34	-54	34
ttH lep	-16	37	-	-	8	56	-20	40
tH lep 1fwd	-	-	-	-	-	-	-	-
tH lep 0fwd	-39	14	-54	19	-22	14	-7	15

Table 6.7: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the VH $H \to \gamma\gamma$ samples for the different event reconstruction categories, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of VH events passing each category.

$t\bar{t}H$			
Categories	**Variation [%]**	**Categories**	**Variation [%]**
ggH 0J Cen	36.4		
ggH 0J Fwd	12.3	VH MET Low	-2.9
ggH 1J Low	23.7	VH MET High	-2.0
ggH 1J Med	-2.9	VH lep Low	3.6
ggH 1J High	30.8	VH lep High	9.6
ggH 1J BSM	-15.3	VH dilep	12.5
ggH 2J Low	11.4		
ggH 2J Med	1.7	tH had 4j2b	0.9
ggH 2J High	-1.2	tH had 4j1b	0.4
ggH 2J BSM	5.9	ttH had BDT4	-5.5
VBF loose, low p_{T}^{Hjj}	27.1	ttH had BDT3	-8.0
VBF tight, low p_{T}^{Hjj}	17.2	ttH had BDT2	-6.7
VBF loose, high p_{T}^{Hjj}	4.3	ttH had BDT1	-8.2
VBF tight, high p_{T}^{Hjj}	4.4	ttH lep	-4.4
VH had loose	8.1	tH lep 1fwd	-6.6
VH had tight	7.8	tH lep 0fwd	2.9
jet BSM	2.9		

Table 6.8: Differences in acceptance between Herwig++ and Pythia8 (AZNLO) obtained with the MG5_aMC@NLO $t\bar{t}H$ $H \to \gamma\gamma$ sample for the different event reconstruction categories, as described in [215]. This difference corresponds to (Herwig++ -Pythia8 (AZNLO)) / Pythia8 (AZNLO).

	ISR_1 up		**ISR_1 down**		**ISR_2 up**		**ISR_2 down**	
Categories	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
ggH 0J Cen	-0.4	0.3	0.6	0.3	-0.5	0.3	-0.5	0.3
ggH 0J Fwd	0.1	0.2	1.1	0.2	0.0	0.2	0.2	0.2
ggH 1J Low	0.2	0.3	1.4	0.3	0.5	0.3	0.4	0.3
ggH 1J Med	-0.1	0.4	1.1	0.4	-0.2	0.4	0.0	0.4
ggH 1J High	0.4	0.8	1.6	0.8	-0.4	0.8	0.2	0.8
ggH 1J BSM	2	3	0	3	-2	3	0	3
ggH 2J Low	0.0	0.7	1.3	0.7	0.1	0.7	0.3	0.7
ggH 2J Med	0.4	0.6	0.8	0.6	0.0	0.6	0.3	0.6
ggH 2J High	1.4	0.9	1.9	0.9	1.4	0.9	1.1	0.9
ggH 2J BSM	1.3	1.8	2.5	1.8	2.0	1.8	1.9	1.8
VBF loose, low p_{T}^{Hjj}	1.2	1.8	2.0	1.8	0.4	1.8	0.5	1.8
VBF tight, low p_{T}^{Hjj}	-1	3	1	3	-1	3	-2	3
VBF loose, high p_{T}^{Hjj}	-0.9	1.5	-0.4	1.6	-1.2	1.5	-1.6	1.5
VBF tight, high p_{T}^{Hjj}	0.1	1.6	2.0	1.6	1.3	1.6	2.3	1.6
VH had loose	0.0	1.2	0.8	1.2	0.1	1.2	0.3	1.2
VH had tight	2.2	1.6	1.2	1.6	0.4	1.5	1.3	1.5
jet BSM	0.2	1.0	0.6	1.0	-0.2	0.9	0.5	1.0
VH MET Low	-27	13	-24	13	-12	15	-28	13
VH MET High	-20	57	21	81	-87	15	-47	41
VH lep Low	-	-	-	-	-	-	-	-
VH lep High	-	-	-	-	-	-	-	-
VH dilep	-	-	-	-	-	-	-	-
tH had 4j2b	-6	9	-7	9	-7	9	-8	9
tH had 4j1b	2	6	5	7	1	6	-3	6
ttH had BDT4	3	10	2	10	8	10	-11	9
ttH had BDT3	-8	25	3	27	16	30	10	29
ttH had BDT2	-2	18	0	18	0	18	10	20
ttH had BDT1	13	32	5	30	11	31	18	33
ttH lep	-47	41	8	73	-73	24	20	80
tH lep 1fwd	-	-	-	-	-	-	-	-
tH lep 0fwd	11	25	-4	23	9	25	3	24

Table 6.9: Differences in acceptance between AZNLO (ISR eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \rightarrow \gamma\gamma$ sample for the different event reconstruction categories, as described in [215]. This difference corresponds to (AZNLO (ISR eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends on the number of ggH events passing each category.

	FSR up		FSR down		MPI up		MPI down	
Categories	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
ggH 0J Cen	-0.7	0.3	1.1	0.3	-0.3	0.3	-0.6	0.3
ggH 0J Fwd	-0.1	0.2	1.5	0.2	0.3	0.2	-0.1	0.2
ggH 1J Low	3.7	0.3	-1.9	0.3	0.0	0.3	0.6	0.3
ggH 1J Med	-0.1	0.4	1.4	0.4	0.0	0.4	0.1	0.4
ggH 1J High	-1.1	0.8	2.8	0.9	0.8	0.8	0.4	0.8
ggH 1J BSM	-5	3	3	3	0	3	-1	3
ggH 2J Low	5.8	0.8	-4.6	0.7	-0.9	0.7	0.3	0.7
ggH 2J Med	3.7	0.6	-3.6	0.6	-0.6	0.6	0.3	0.6
ggH 2J High	3.1	0.9	-0.7	0.9	0.4	0.9	0.5	0.9
ggH 2J BSM	2.2	1.8	1.5	1.8	1.1	1.8	1.5	1.8
VBF loose, low p_{T}^{Hjj}	8.9	1.9	-6.0	1.7	0.7	1.8	1.6	1.8
VBF tight, low p_{T}^{Hjj}	5	3	-6	3	-3	3	-3	3
VBF loose, high p_{T}^{Hjj}	0.7	1.6	-3.1	1.5	-1.8	1.5	-2.0	1.5
VBF tight, high p_{T}^{Hjj}	3.4	1.6	0.2	1.6	1.8	1.6	1.2	1.6
VH had loose	1.7	1.2	-1.3	1.2	0.6	1.2	-0.6	1.2
VH had tight	0.5	1.6	1.7	1.6	0.4	1.5	0.2	1.6
jet BSM	2.4	1.0	-1.2	0.9	0.2	1.0	-0.2	0.9
VH MET Low	-16	14	-19	14	-24	13	-21	14
VH MET High	-46	41	61	104	7	72	47	96
VH lep Low	-	-	-	-	-	-	-	-
VH lep High	-	-	-	-	-	-	-	-
VH dilep	-	-	-	-	-	-	-	-
tH had 4j2b	2	9	-1	9	-14	8	4	10
tH had 4j1b	-4	6	5	6	3	6	3	6
ttH had BDT4	10	10	11	10	0	10	-5	9
ttH had BDT3	17	30	19	31	-2	26	6	28
ttH had BDT2	10	19	14	20	5	19	3	18
ttH had BDT1	-14	25	11	31	4	30	8	31
ttH lep	-87	16	-73	24	-47	41	-73	24
tH lep 1fwd	-	-	-	-	-	-	-	-
tH lep 0fwd	-21	19	-6	22	14	26	-3	23

Table 6.10: Differences in acceptance between AZNLO (FSR / MPI eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for the different event reconstruction categories, as described in [215]. This difference corresponds to (AZNLO (FSR / MPI eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends on the number of ggH events passing each category.

Categories	**BDT FxFx uncertainty (%)**	**QCD uncertainty (%)**
VBF loose, low p_{T}^{Hjj}	5	20
VBF tight, low p_{T}^{Hjj}	3	26
VBF loose, high p_{T}^{Hjj}	1	11
VBF tight, high p_{T}^{Hjj}	11	17
VH had loose	7	19
VH had tight	4	23

Table 6.11: BDT acceptance uncertainties obtained with the POWHEG NNLOPS ggH $H \to \gamma\gamma$ and the MG5_aMC@NLO FxFx ggH $H \to \gamma\gamma$ samples for the VBF and VH event reconstruction categories. Comparison with the QCD uncertainties for the same even reconstruction categories.

these categories, and the change in the predicted yield is consistent with the statistical uncertainty of 24–80%, which is strongly category dependent. In order to stay conservative, the uncertainty of 100% is maintained on the ggH contribution. This is consistent with recent measurements of $t\bar{t}$ and vector-boson production in association with b-jets [216–218]. For VBF, the uncertainty ranges from 65-90%, with statistical uncertainties between 7–21%.

Due to the small contributions of these production modes to the $t\bar{t}H$ event reconstruction categories the large uncertainties do not strongly impact the $t\bar{t}H$ sensitivity.

Normalization uncertainties

The theoretical modeling uncertainties in the measurement of signal strengths include all of the sources that affect the measurement of the production-mode cross sections, plus additional uncertainties in the overall normalization of each production mechanism. Uncertainties in the overall normalization of each production process from missing higher-order QCD effects and the choice of parton distribution function are specified as part of the theoretical calculations used to normalize the simulated samples.

The normalization uncertainty from the $H \to \gamma\gamma$ branching ratio corresponds to a total value of $^{+2.90\%}_{-2.84\%}$.

6.7.5 Theoretical and modeling uncertainties for fiducial integrated and differential results

The theoretical modeling uncertainty in the detector correction factor used to measure the fiducial integrated and differential cross sections is taken to be the envelope of the following three sources:

- The uncertainty in the relative contributions of the different Higgs boson pro-

duction mechanisms. This uncertainty is estimated by varying the fraction of the ggH, VBF, VH and $t\bar{t}H$ processes by an amount commensurate with the 68% confidence levels of the measured production mode cross section ratios [39]. The variations of each production mechanism are carried out simultaneously and include the known correlations between the measured production mode cross section ratios. These uncertainties range from 0.1% to 31%, depending on the fiducial region or differential variable, increasing typically in bins and regions sensitive to $t\bar{t}H$-production.

- The uncertainty in the detector correction factor due to a possible mismodeling of the Higgs boson transverse momentum and rapidity distributions is estimated by reweighting the Higgs boson distributions in simulation to match those observed in the data. The resulting uncertainties range from 0.1% to 4.5%, increasing in fiducial regions and bins with high jet multiplicities.

- The uncertainty in the modeling of the parton shower, underlying event, and hadronization. This uncertainty is derived as described in Section 6.7.4 and the size of this uncertainty ranges from from 0.1% up to 30%, with the highest uncertainties in fiducial regions with large missing transverse energy. The Underlying Event / Parton Shower tune acceptance uncertainties in the five fiducial regions and in each bin of the eleven fiducial differential cross sections reported in this thesis are detailed in Appendix B.

Typically differential measurements involving only the photon kinematics are less affected by these model uncertainties than measurements using selections on jets or missing transverse momentum.

6.8 Results

6.8.1 Observed data

The observed invariant mass distribution of the selected diphoton pairs of all categories as defined in Table 6.3, is shown in Figure 6.2. Figure 6.3 shows the invariant mass distributions for the sums of the categories most sensitive to the different production modes. For illustration purposes, events in each category are weighted according to the expected signal (S_{90}) to background (B_{90}) ratio in a $m_{\gamma\gamma}$ region containing 90% of the expected signal yield, using a weight of the form $\ln(1 + S_{90}/B_{90})$. The results of signal-plus-background fits to these spectra, displaying both the total sum and the background-only components, are shown, as well as the residuals between the data and the background component. Both the signal-plus-background and background-only distributions shown are obtained from the sum of the individual distributions in each category weighted in the same way as the data points. In the fit a single signal strength μ affecting simultaneously all production modes has been assumed. The observed mass peak of the Higgs boson, constrained in the fit as $m_H = 125.09 \pm 0.24$ GeV, is well within 68% CL of the Run 1 ATLAS+CMS combined measurement.

6.8.2 Signal strengths

A global signal strength μ is measured assuming the ratios between different production processes to be as predicted by the SM. The profile of the negative log-likelihood ratio $\lambda(\mu)$ of the global signal strength of all Higgs boson processes μ for $m_H = 125.09$ $\pm$ 0.24 GeV is shown in Figure 6.4.

The measured central value and 68% CL interval for μ is found to be:

$$\mu = 0.99\ ^{+0.15}_{-0.14} = 0.99\ \pm 0.12\,(\text{stat.})\ ^{+0.06}_{-0.05}\,(\text{exp.})\ ^{+0.07}_{-0.05}\,(\text{theo.})\ ,$$

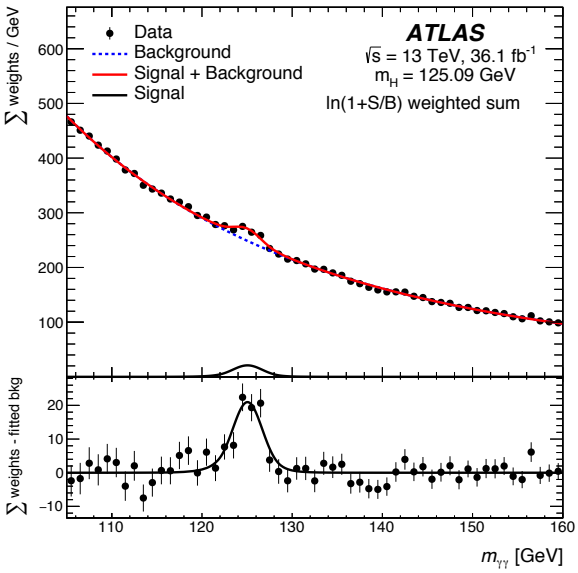


Figure 6.2: Weighted diphoton invariant mass spectrum observed in the 2015 and 2016 data at 13 TeV. The error bars represent 68% confidence intervals of the weighted sums. The solid red curve shows the fitted signal-plus-background model when the Higgs boson mass is constrained to be 125.09 ± 0.24 GeV. The background component of the fit is shown with the dotted blue curve. The signal component of the fit is shown with the solid black curve [4].

well compatible with the SM prediction ($\mu = 1$). This result confirms the ATLAS Run 1 diphoton signal strength measurement of $\mu = 1.17 \pm 0.23\,(\text{stat.})\,^{+0.10}_{-0.08}\,(\text{exp.})\,^{+0.12}_{-0.08}\,(\text{theo.})$ with around a factor of two improvement in each component of the uncertainty. The Run 1 result was obtained using the NNLO SM prediction for ggH production [33, 219], which is about 10% lower than the N^{3}LO calculation used here (see Section 6.3). The impact of the main sources of systematic uncertainty (presented in Table 6.5) in the measured global signal strength is summarized in Table 6.12. The distinction between yield and migration uncertainties adopted in Table 6.5 is used and the uncertainties are grouped into theory uncertainties, experimental uncertainties, mass resolution and scale, background shape, and luminosity.

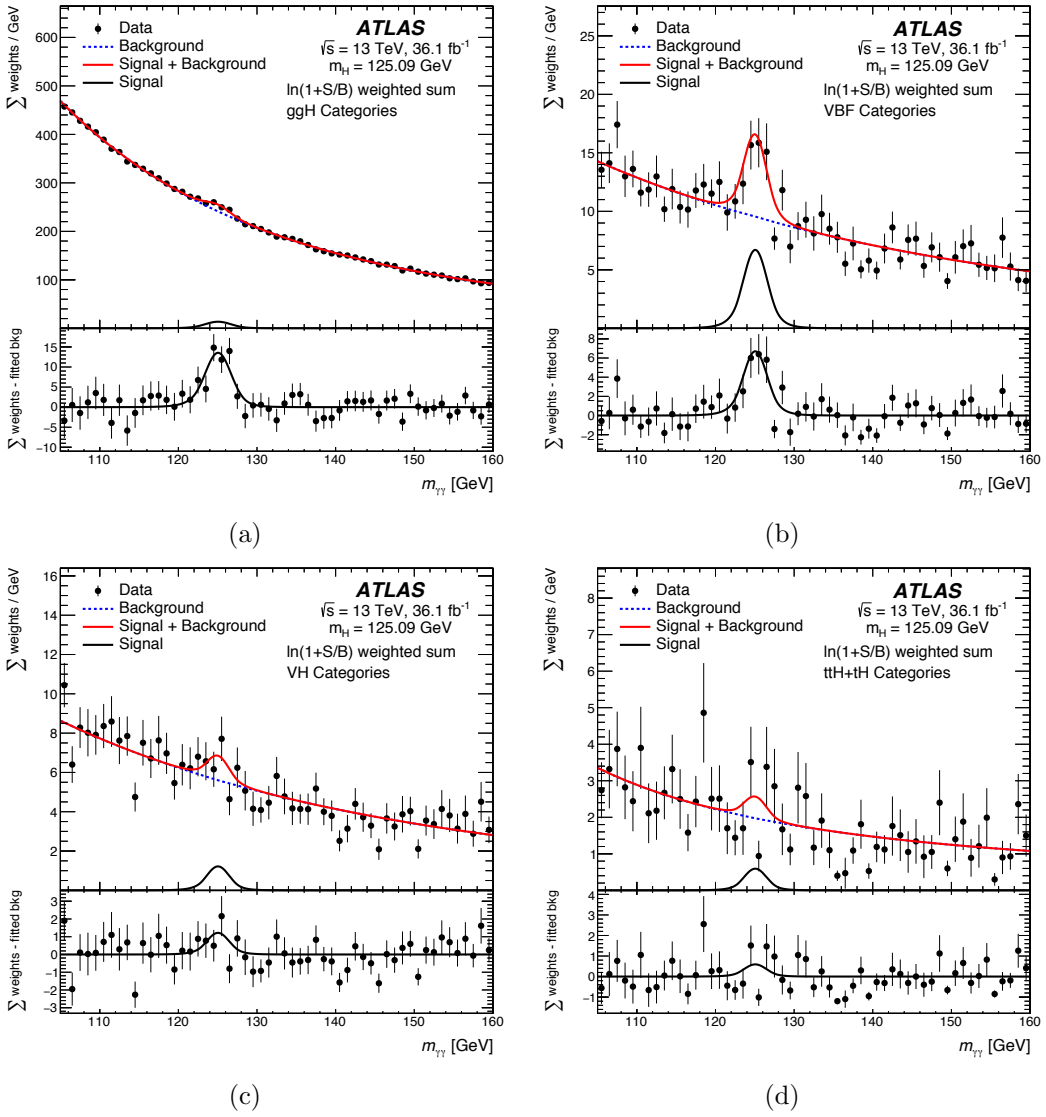


Figure 6.3: Weighted diphoton invariant mass spectra observed in the 13 TeV data for events belonging to: a) "untagged" categories and the "jet BSM" category, in which the expected signal is produced mainly through gluon–gluon fusion, b) VBF categories, c) VH categories and d) $t\bar{t}H$ categories. The error bars represent 68% confidence intervals of the weighted sums. The solid red curve shows the fitted signal-plus-background model when the Higgs boson mass is constrained to be 125.09 ± 0.24 GeV. The background component of the fit is shown with the dotted blue curve. The signal component of the fit is shown with the solid black curve [4]

Table 6.12: Main systematic uncertainties σ_μ^{syst} in terms of absolute value in the combined signal strength parameter μ. The values for each group of uncertainties are determined by subtracting in quadrature from the total uncertainty the change in the 68% CL range of μ when the corresponding nuisance parameters are fixed to their best fit values. The experimental uncertainty in the yield does not include the luminosity contribution, which is accounted for separately. The uncertainties correspond to the sources detailed in Table 6.5.

Uncertainty Group	σ_μ^{syst}
Theory (QCD)	0.041
Theory ($\mathcal{B}(H \to \gamma\gamma)$)	0.028
Theory (PDF+α_S)	0.021
Theory (Underlying Event / Parton Shower)	0.026
Luminosity	0.031
Experimental (yield)	0.017
Experimental (migrations)	0.015
Mass resolution	0.029
Mass scale	0.006
Background shape	0.027

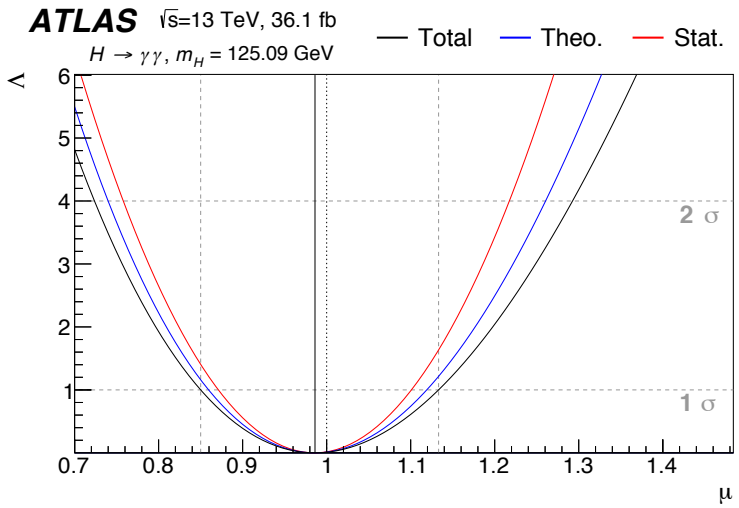


Figure 6.4: Observed negative log-profile likelihood Λ of the global signal strength μ. The three likelihood contours shown correspond to all theory and experimental nuisance parameters fixed (stat.), all experimental nuisance parameters fixed (theo.), and with all nuisance parameters floating (total). The intersections of the solid curves and horizontal lines at $\Lambda = 1$ and $\Lambda = 4$ indicate the 1σ and 2σ confidence intervals of the corresponding result [4].

In addition to the global signal strength, the signal strengths of the primary production processes are evaluated by exploiting the sensitivities of the analysis categories of Table 6.3 to specific production processes. The measured signal strengths are shown together with the global signal strengths discussed in Figure 6.5 and found to be:

$$\mu_{\mathrm{ggH}} = 0.81\ ^{+0.19}_{-0.18} = 0.81\ \pm 0.16\,(\mathrm{stat.})\ ^{+0.07}_{-0.06}\,(\mathrm{exp.})\ ^{+0.07}_{-0.05}\,(\mathrm{theo.})$$

$$\mu_{\mathrm{VBF}} = 2.0\ ^{+0.6}_{-0.5} = 2.0\ \pm 0.5\,(\mathrm{stat.})\ ^{+0.3}_{-0.2}\,(\mathrm{exp.})\ ^{+0.3}_{-0.2}\,(\mathrm{theo.})$$

$$\mu_{\mathrm{VH}} = 0.7\ ^{+0.9}_{-0.8} = 0.7\ \pm 0.8\,(\mathrm{stat.})\ ^{+0.2}_{-0.2}\,(\mathrm{exp.})\ ^{+0.2}_{-0.1}\,(\mathrm{theo.})$$

$$\mu_{\mathrm{top}} = 0.5\ ^{+0.6}_{-0.6} = 0.5\ ^{+0.6}_{-0.5}\,(\mathrm{stat.})\ ^{+0.1}_{-0.1}\,(\mathrm{exp.})\ ^{+0.1}_{-0.0}\,(\mathrm{theo.})$$

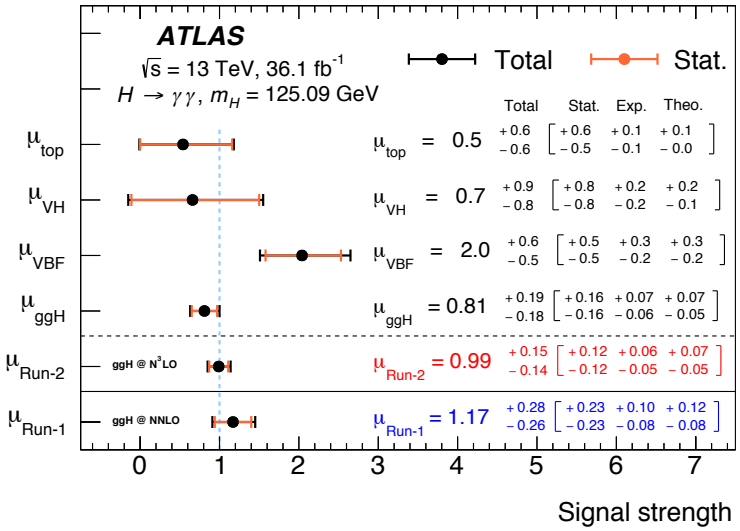


Figure 6.5: Summary of the signal strengths measured for the different production processes (ggH, VBF, VH and top) and globally ($\mu_{\text{Run 2}}$), compared to the global signal strength measured at 7 and 8 TeV ($\mu_{\text{Run 1}}$) [146]. The black and orange error bars show the total and statistical uncertainties. The signal strength $\mu_{\text{Run 1}}$ was derived assuming the Higgs boson production-mode cross section based on [33, 219]. Uncertainties smaller than 0.05 are displayed as 0.0. In the more recent theoretical predictions used in this analysis [34, 177], the gluon-gluon fusion production-mode cross section is larger by approximately 10% [4].

For Higgs boson production via VH the signal strength is assumed to be scaled by a single parameter (*i.e.* $\mu_{\text{VH}} = \mu_{\text{ZH}} = \mu_{\text{WH}}$). The $b\bar{b}H$ contributions are scaled with ggH (*i.e.* $\mu_{\text{b}\bar{\text{b}}\text{H}} = \mu_{\text{ggH}}$), and the tH and $t\bar{t}H$ productions are measured together rather than separately (*i.e.* $\mu_{\text{top}} = \mu_{\text{t}\bar{\text{t}}\text{H+tH}}$).

6.8.3 Production-mode and Simplified template cross sections

The production-mode cross sections for $m_H = 125.09 \pm 0.24$ GeV in a region with Higgs-boson rapidity $|y_H| < 2.5$, multiplied by the branching ratio of the Higgs boson decay to diphotons, are evaluated in the following way. The fitted value of σ_{top} corresponds to the sum of $t\bar{t}H$, tHq, and tHW production-mode cross sections un-

der the assumption that their relative ratios are as predicted by the SM. The VH production-mode cross section value is fitted under the assumption that the ratio of the WH and ZH production mode cross sections is as predicted by the SM and includes both production from quark and gluon initial states. The $b\bar{b}H$ contributions are merged with ggH. Such results are obtained through signal-plus-background fits to the diphoton invariant mass distribution in each category by expressing, in the likelihood, the signal yield $N^i_{\text{sig},m}$ in each category i for a particular production mode m as $N_{\text{sig},m} = \int L\,\mathrm{d}t \times \sigma^{\text{SM}}_m \times \mathcal{B}(H \to \gamma\gamma) \times \epsilon^i_m$ using the same notation as in Section 6.8.2. Results for the production-mode cross sections are shown in Table 6.13. In Figure 6.6, the results are shown divided by their SM expectation but no additional theory uncertainties have been added to the ratio. All observed cross sections are in agreement with the SM values.

Table 6.13: Best-fit values and uncertainties of the production-mode cross sections times branching ratio. The SM predictions [34] with their uncertainties are shown for each production process. Uncertainties smaller than 0.05 are displayed as 0.0.

Process	Result	Uncertainty [fb]				SM prediction
($\lvert y_H\rvert < 2.5$)	[fb]	Total	Stat.	Exp.	Theo.	[fb]
ggH	82	$^{+19}_{-18}$	$\left(\pm 16\right.$	$^{+7}_{-6}$	$\left.^{+5}_{-4}\right)$	102^{+5}_{-7}
VBF	16	$^{+5}_{-4}$	$\left(\pm 4\right.$	± 2	$\left.^{+3}_{-2}\right)$	8.0 ± 0.2
VH	3	± 4	$\left(^{+4}_{-3}\right.$	± 1	$\left.^{+1}_{-0}\right)$	4.5 ± 0.2
Top	0.7	$^{+0.9}_{-0.7}$	$\left(^{+0.8}_{-0.7}\right.$	$^{+0.2}_{-0.1}$	$\left.^{+0.2}_{-0.0}\right)$	1.3 ± 0.1

Since the current data set is not yet sensitive to all of the 31 regions with $|y_H| < 2.5$ (assuming SM acceptance) of stage-1 of the simplified template cross section framework, simplified template cross sections are reported for 10 phase space regions obtained from merging the initial 31 as detailed in Table 6.1. To retain

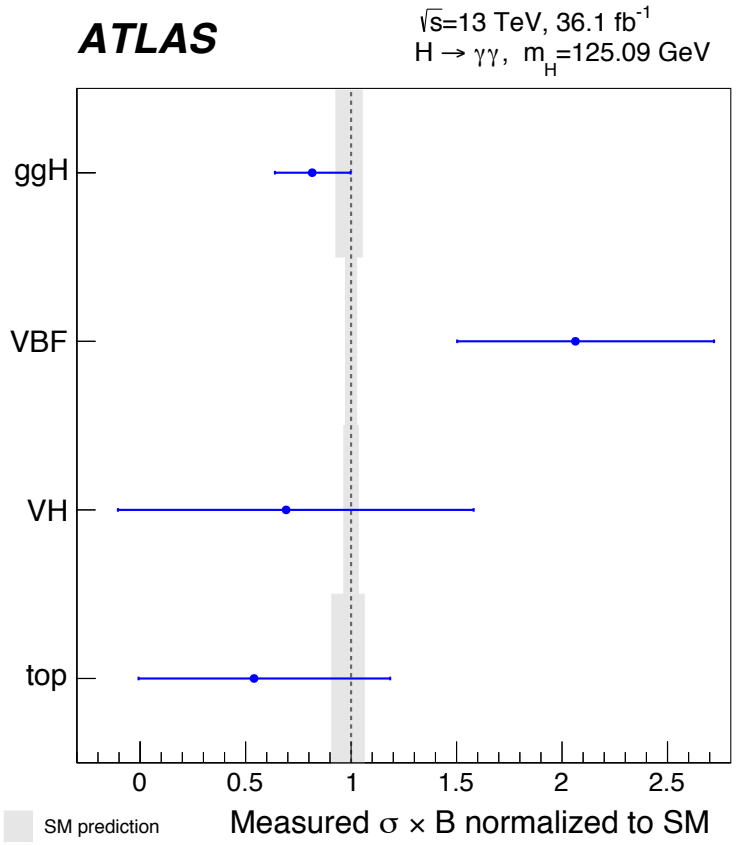


Figure 6.6: Summary plot of the measured production-mode cross sections times the Higgs boson to diphoton branching ratio. For illustration purposes the central values have been divided by their SM expectations but no additional theory uncertainties have been added to the uncertainty of the ratio. The uncertainties in the predicted SM cross sections are shown in gray bands in the plot [4].

sensitivity to BSM Higgs boson production, the $p_{\mathrm{T}}^{H} > 200$ GeV gluon-gluon fusion and $p_{\mathrm{T}}^{j} > 200$ GeV VBF regions are not merged with other regions. In the likelihood, the signal yield N_{sig}^{i} in each reconstructed category i is the sum over the yields $N_{\mathrm{sig},r}^{i}$ expected from each of the 9 regions r of phase space, where $N_{\mathrm{sig},r}^{i} = \int L\,\mathrm{d}t \times \sigma_{r}^{\mathrm{SM}} \times \mathcal{B}(H \to \gamma\gamma) \times \epsilon_{r}^{i}$ where ϵ_{r}^{i} is the probability for an event in a given region r to end up in a given reconstructed category i. The additional region of phase space corresponds to the difference of the cross sections for the $p_{\mathrm{T}}^{H} > 200$ GeV gluon–gluon fusion and $p_{\mathrm{T}}^{j} > 200$ GeV VBF regions. The results are shown in Table 6.14 and in Figure 6.7. The largest deviation (1.7 σ) from the SM prediction is found in the $ggH, 0$ jet bin. The difference of the cross sections for the $p_{\mathrm{T}}^{H} > 200$ GeV ggH and $p_{\mathrm{T}}^{j} > 200$ GeV VBF regions is found to be $4.8^{+2.9}_{-2.7}$ fb. This large difference arises due to a deficit in the $p_{\mathrm{T}}^{j} > 200$ GeV VBF region and an excess in the $p_{\mathrm{T}}^{H} > 200$ GeV ggH region. This result is determined by profiling the difference of the two regions in the fit.

Table 6.14: Best-fit values and uncertainties of the simplified template cross sections times branching ratio. The SM predictions [34] are shown for each region.

Measurement region ($\lvert y_H \rvert < 2.5$)	Result [fb]	Uncertainty [fb] Total	Stat.	Syst.	SM prediction [fb]
ggH, 0 jet	37	$^{+16}_{-15}$	$\left(\pm 14\right.$	$\left.^{+6}_{-5}\right)$	63 ± 5
ggH, 1 jet, $p_{\mathrm{T}}^{H} < 60$ GeV	13	$^{+13}_{-12}$	$\left(\pm 12\right.$	$\left.^{+5}_{-4}\right)$	15 ± 2
ggH, 1 jet, $60 \leq p_{\mathrm{T}}^{H} < 120$ GeV	5	± 6	$\left(\pm 6\right.$	$\left.^{+2}_{-1}\right)$	10 ± 2
ggH, 1 jet, $120 \leq p_{\mathrm{T}}^{H} < 200$ GeV	2.8	$^{+1.7}_{-1.6}$	$\left(^{+1.6}_{-1.5}\right.$	$\left.^{+0.7}_{-0.5}\right)$	1.7 ± 0.3
ggH, ≥ 2 jet	20	$^{+9}_{-8}$	$\left(\pm 8\right.$	$\left.^{+4}_{-3}\right)$	11 ± 2
$qq \to Hqq, p_{\mathrm{T}}^{j} < 200$ GeV	15	$^{+6}_{-5}$	$\left(\pm 5\right.$	$\left.^{+3}_{-2}\right)$	10 ± 0.5
$ggH + qq \to Hqq$, BSM $-$ like	2.0	± 1.4	$\left(\pm 1.3\right.$	$\left.\pm 0.6\right)$	1.8 ± 0.4
VH, leptonic	0.7	$^{+1.4}_{-1.3}$	$\left(^{+1.4}_{-1.2}\right.$	$\left.^{+0.4}_{-0.3}\right)$	1.4 ± 0.1
Top	0.7	$^{+0.8}_{-0.7}$	$\left(^{+0.8}_{-0.7}\right.$	$\left.^{+0.2}_{-0.1}\right)$	1.3 ± 0.1

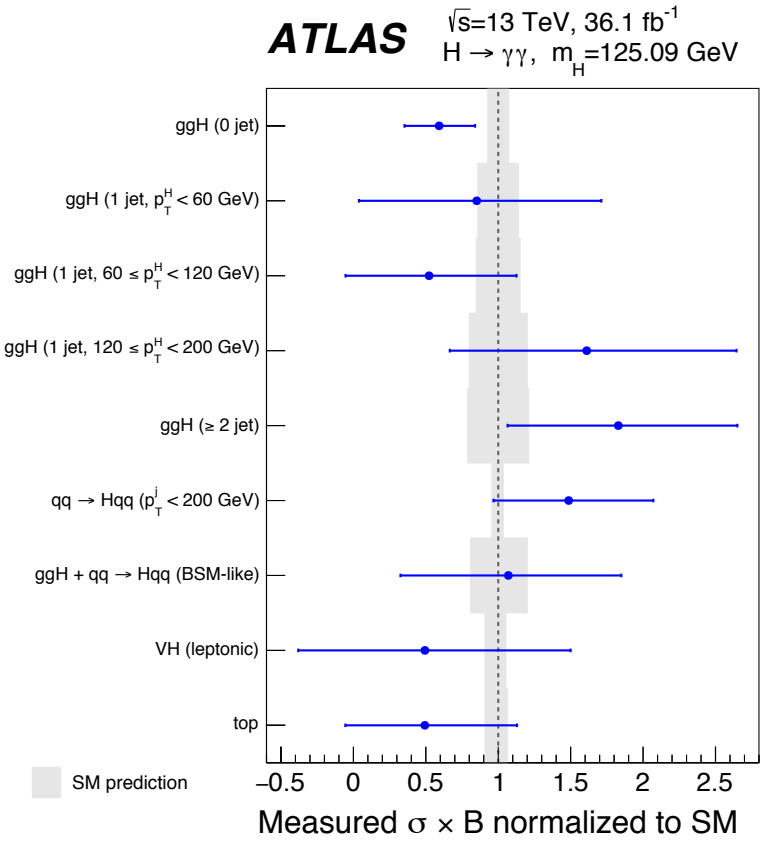


Figure 6.7: Summary plot of the measured simplified template cross sections times the Higgs boson to diphoton branching ratio. For illustration purposes the central values have been divided by their SM expectations but no additional theory uncertainties have been included in the uncertainty of the ratio due to this. The uncertainties in the predicted SM cross sections are shown in gray in the plot [4].

Table 6.15: The measured cross sections in the diphoton, VBF-enhanced, $N_{\text{lepton}} \geq 1$, high $E_{\text{T}}^{\text{miss}}$, and $t\bar{t}H$-enhanced fiducial regions. The SM predictions arising from VBF, VH, $t\bar{t}H$ and $b\bar{b}H$ are collectively labeled as XH.

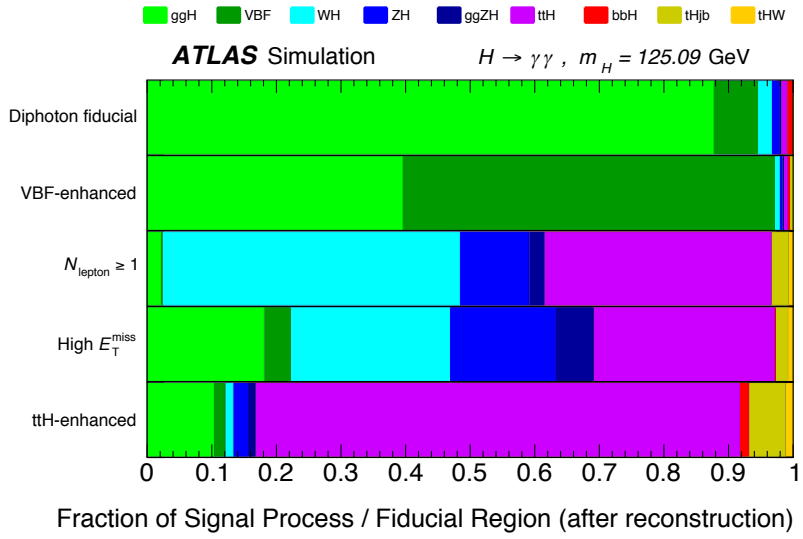
Fiducial region	Measured cross section	SM prediction	
Diphoton fiducial	55 ± 9 (stat.) ± 4 (exp.) ± 0.1 (theo.) fb	64 ± 2 fb	[N^{3}LO + XH]
VBF-enhanced	3.7 ± 0.8 (stat.) ± 0.5 (exp.) ± 0.2 (theo.) fb	2.3 ± 0.1 fb	[default MC + XH]
$N_{\text{lepton}} \geq 1$	≤ 1.39 fb 95% CL	0.57 ± 0.03 fb	[default MC + XH]
High $E_{\text{T}}^{\text{miss}}$	≤ 1.00 fb 95% CL	0.30 ± 0.02 fb	[default MC + XH]
$t\bar{t}H$-enhanced	≤ 1.27 fb 95% CL	0.55 ± 0.06 fb	[default MC + XH]

6.8.4 Measurement of fiducial integrated and differential cross sections

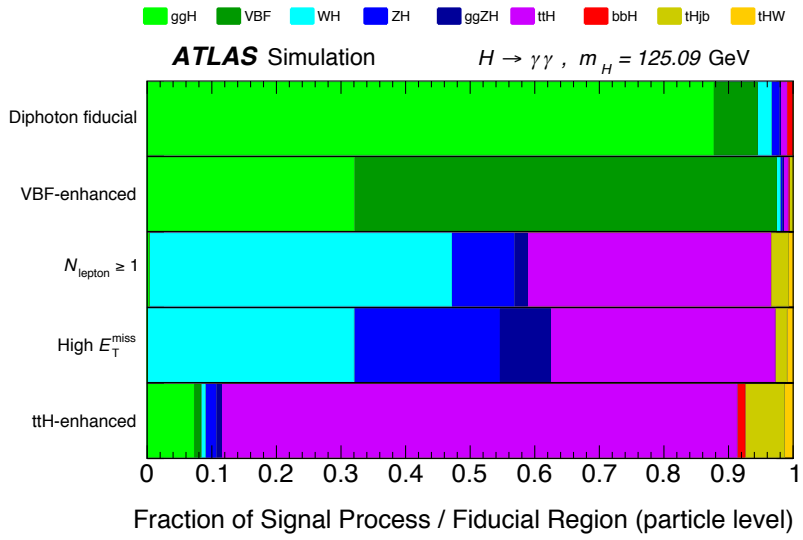
Fiducial integrated cross sections

The cross sections are measured in five fiducial integrated regions that target specific Higgs boson production mechanisms or are sensitive to physics beyond the Standard Model. The regions are summarized in Table 6.4. The expected composition of Higgs boson events in each region after reconstruction and at particle level is summarized in Figure 6.8. At particle level the VBF-enhanced fiducial region contains about 65% VBF and 32% ggH events. The particle-level $N_{\text{lepton}} \geq 1$ region is dominated by WH (47%), ttH (37%) and ZH (13%) production. The particle-level high $E_{\text{T}}^{\text{miss}}$ region is populated by about equal amounts of WH, ZH, and ttH (32%, 30%, and 35%). Finally, the particle-level ttH-enhanced region contains about 80% ttH events.

Figure 6.9 and Table 6.15 summarize measured cross sections of the fiducial regions and limits, and compare both to the Standard Model expectations, constructed as outlined above. The POWHEG NNLOPS prediction, without any additional corrections, is also shown. The uncertainty band is estimated using a set of scale variations and includes PDF uncertainties from eigenvector variations. The Standard Model predictions of all fiducial regions are in agreement with the corresponding measured cross sections.



(a)



(b)

Figure 6.8: The expected composition of Higgs boson events in each fiducial region a) after the reconstruction and b) at particle-level [4].

Differential cross sections

Eleven fiducial differential cross sections are measured that characterize the Higgs boson production kinematics, the kinematics of jets produced in association with the

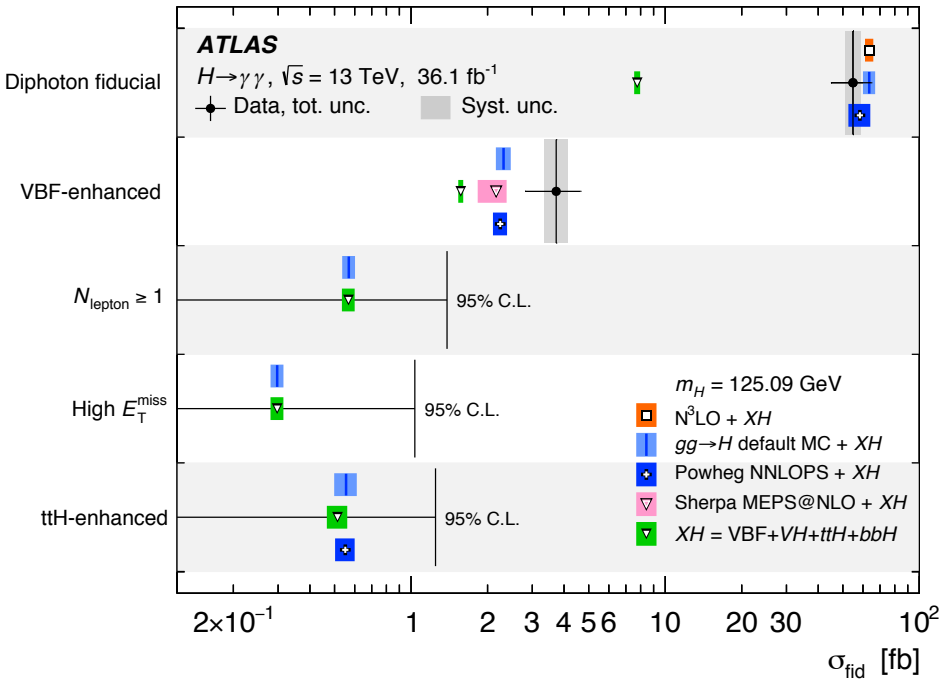


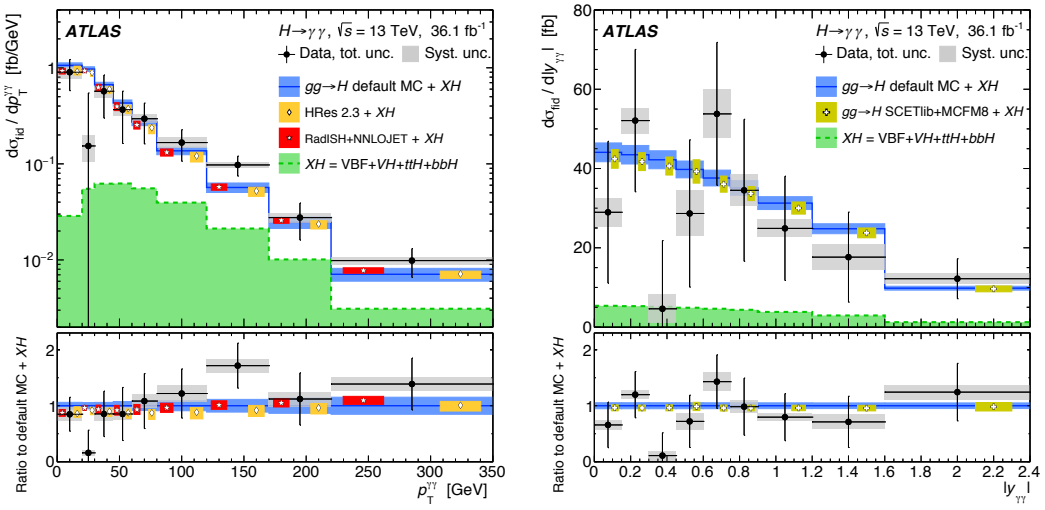
Figure 6.9: The measured cross sections or cross section upper limits of the diphoton, VBF-enhanced, $N_{\text{lepton}} \geq 1$, high $E_{\text{T}}^{\text{miss}}$, and $t\bar{t}H$-enhanced fiducial regions are shown. The intervals on the vertical axis each represent one of these fiducial regions. The data are shown as filled (black) circles. The error bar on each measured cross section represents the total uncertainty in the measurement, with the systematic uncertainty shown as a dark gray rectangle. Each cross section limit is shown at the 95% confidence level. The measured cross sections are compared to a range of predictions and a detailed description of each prediction can be found in Section 6.3. All comparisons include the SM predictions arising from VBF, VH, $t\bar{t}H$, and $b\bar{b}H$, which are collectively labeled as XH [4].

Higgs boson, the spin and CP quantum numbers of the Higgs boson and variables sensitive to the VBF production mechanism. The variables are listed in Section 6.2 and are discussed individually in what follows.

The rapidity distribution of the Higgs boson is also sensitive to the modeling of the gluon-gluon fusion production mechanism, as well as to the parton distribution functions (PDFs) of the colliding protons. The chosen bin widths are a compromise between retaining a sufficiently significant signal and providing spectra with good granularity. Each bin is chosen such that it retains an expected significance of at least two standard deviations, estimated using the Powheg NNLOPS as well as using a fit to $m_{\gamma\gamma}$ sidebands. In Figure 6.10 the SM prediction shows a slight excess at low transverse momentum and low rapidity, and shows a slight deficit at large transverse momentum.

The transverse momentum and absolute rapidity of the leading and subleading jets, shown in Figure 6.11, probe the perturbative QCD modeling and are sensitive to the relative contributions of the different Higgs boson production mechanisms. The transverse momentum distribution of the leading jet probes the emission of energetic quarks and gluons.

The first bin of the leading jet p_{T} spectrum represents zero-jet events that do not contain any jet with $p_{\mathrm{T}} > 30$ GeV. The predicted p_{T} distributions slightly exceed the measured distribution at low transverse momentum and all show a slight deficit at large transverse momentum. They are compatible with the observed slightly harder Higgs boson transverse momentum distribution. The first bin of the subleading jet p_{T} represents one-jet events that do not contain two or more jets with $p_{\mathrm{T}} > 30$ GeV. In events with two jets, the contributions of VBF and VH productions become more important. In the case of the $|y_{j_1}|$ distribution, the predictions show a slight excess at low rapidity.



(a) (b)

Figure 6.10: The differential cross sections for $pp \to H \to \gamma\gamma$ as a function of a) $p_{\mathrm{T}}^{\gamma\gamma}$ and b) $|y^{\gamma\gamma}|$ are shown and compared to the SM expectations [4].

The angular variables $|\cos\theta^*|$ and $\Delta\phi_{jj}$, shown in Figure 6.12, are sensitive to the spin and CP quantum numbers of the Higgs boson. The $|\cos\theta^*|$ distribution shows a strong drop around 0.6 due to the fiducial requirement on the photon system, characteristic in a scalar particle, whereas for a spin-2 particle, an enhancement would be present in precisely this region. On the other hand, the charge conjugation and parity properties of the Higgs boson are encoded in the azimuthal angle between the jets, whose shape is sensitive to the relative contribution of gluon-gluon fusion and vector-boson fusion, as well as to the tensor structure of the interactions between the Higgs boson and gluons or weak bosons [220–223]. The measured $|\cos\theta^*|$ and $\Delta\phi_{jj}$ distributions are consistent with SM predictions for a CP-even scalar particle.

The dijet rapidity separation $|\Delta y_{jj}|$, the dijet mass m_{jj} and the azimuthal difference between the dijet and diphoton system $|\Delta\phi_{\gamma\gamma,jj}|$ are sensitive to the VBF production mechanism. In vector-boson fusion, the t-channel exchange of a W/Z

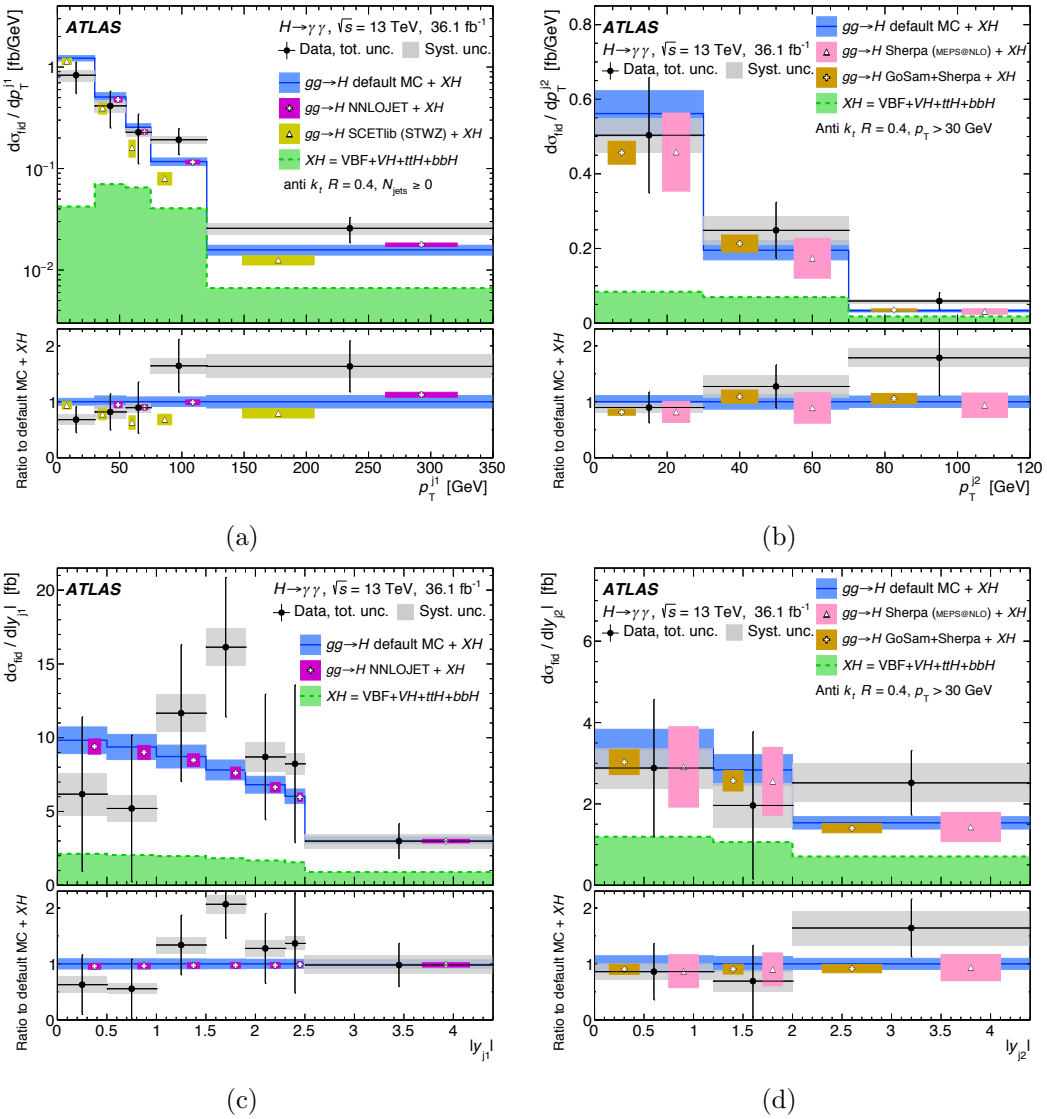
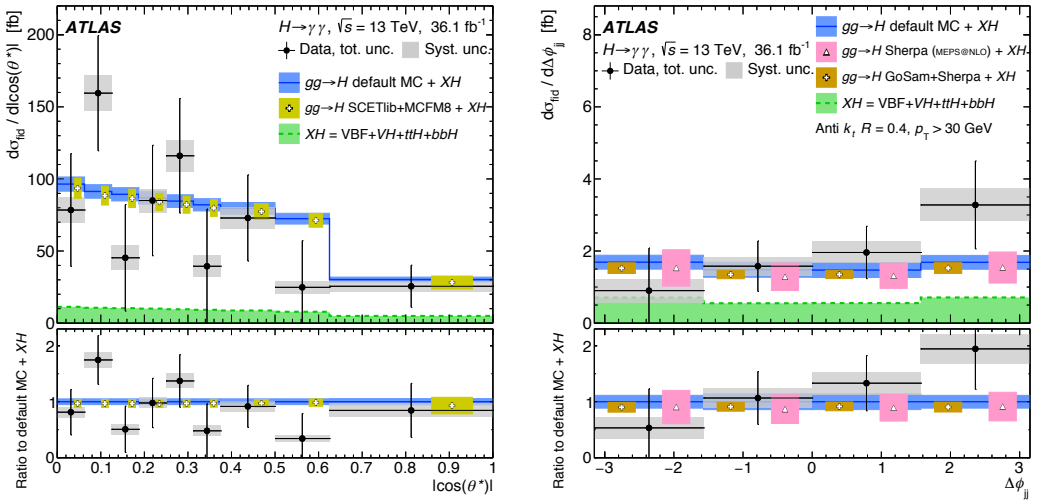


Figure 6.11: The differential cross sections for $pp \to H \to \gamma\gamma$ as a function of a) $p_{\mathrm{T}}^{j_1}$, b) $p_{\mathrm{T}}^{j_2}$, c) $|y^{j_1}|$ and d) $|y^{j_2}|$ are shown and compared to the SM expectations [4].

boson typically results in two moderate-p_{T} jets that are well separated in rapidity. Therefore, the $|\Delta\phi_{\gamma\gamma,jj}|$ distribution for VBF production is expected to be steeper and more peaked towards $|\Delta\phi_{\gamma\gamma,jj}| = \pi$ than for gluon-gluon fusion since quark/gluon



(a) (b)

Figure 6.12: The differential cross sections for $pp \to H \to \gamma\gamma$ as a function of a) $|\cos\theta^*|$ and b) $\Delta\phi_{jj}$ are shown and compared to the SM expectations [4].

radiation in the rapidity interval between the two jets is suppressed in the VBF process because there is no color flow between the two jets. The variables shown in Figure 6.13 are used to discriminate between the ggH and VBF production modes and enter the multivariate classifier that defines the categories used for the simplified template cross section and coupling measurements.

Impact of systematic uncertainties on results

A summary of the uncertainties in the measured cross sections of the fiducial regions are shown in Table 6.16. As an example concerning the differential measurements, a breakdown of the systematic uncertainties in the differential cross sections as a function of $p_{\mathrm{T}}^{\gamma\gamma}$ and N_{jets} is shown in Figure 6.14. The measurements are dominated by the statistical uncertainties. For the systematic uncertainties, the uncertainty in the fitted signal yield, due to the background modeling and the photon energy

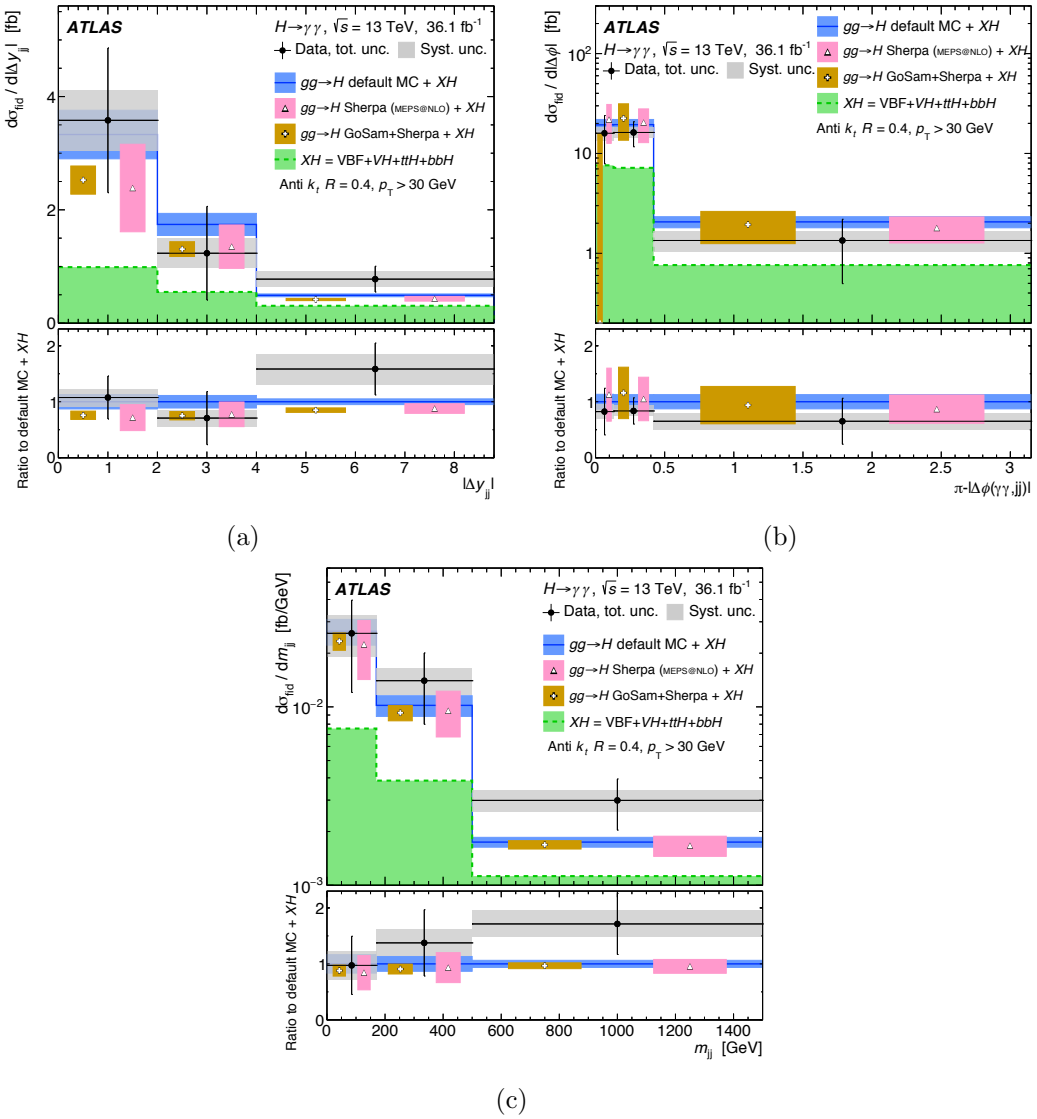


Figure 6.13: The differential cross sections for $pp \to H \to \gamma\gamma$ as a function of a) $|\Delta y_{jj}|$, b) $|\Delta\phi_{\gamma\gamma,jj}|$ and c) m_{jj} are shown and compared to the SM expectations [4].

resolution, is typically more important than the uncertainty in the correction factor due to the theoretical modeling. The jet energy scale and resolution uncertainties become increasingly important for high-jet multiplicities and in the $t\bar{t}H$- and VBF-

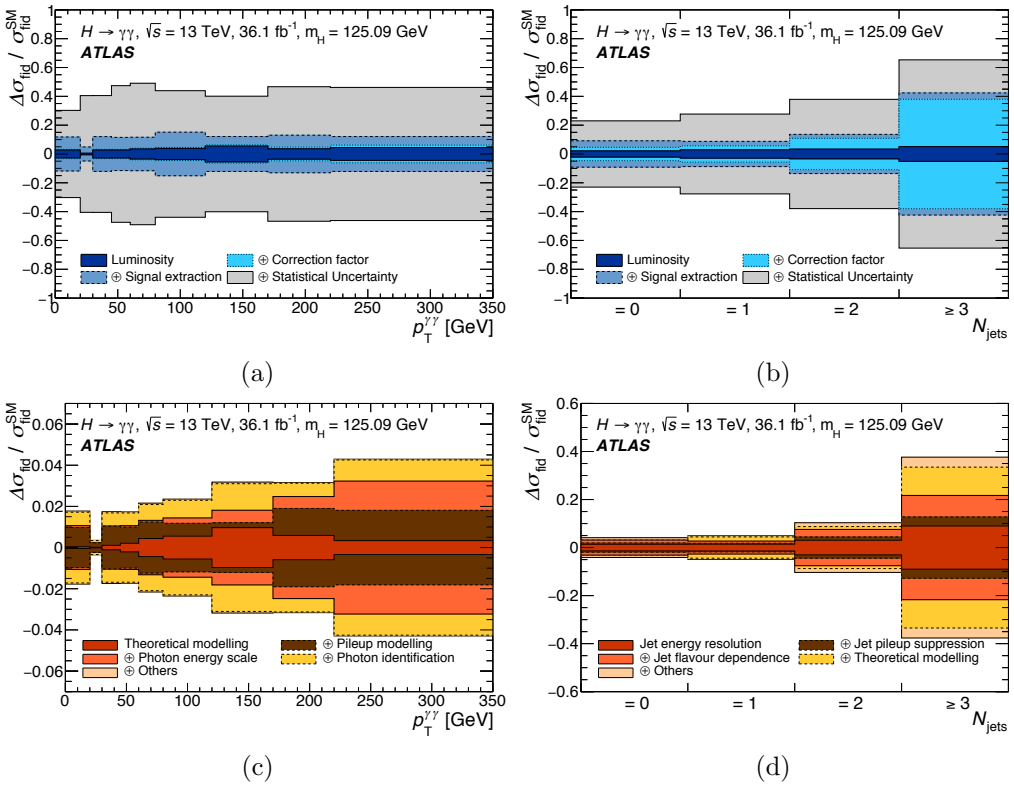


Figure 6.14: The relative size of systematic uncertainties associated with the signal extraction, the correction factors (experimental and theoretical modeling) and the luminosity on the differential cross sections are shown as a function of a) $p_{\mathrm{T}}^{\gamma\gamma}$ and b) N_{jets}. The statistical uncertainty associated with the signal extraction is also shown as a gray band. The relevant components of the uncertainties in the correction factors are shown as a function of c) $p_{\mathrm{T}}^{\gamma\gamma}$ and d) N_{jets} [4].

enhanced phase space. The uncertainty in the modeling of the parton shower and the underlying event dominates in fiducial regions with large missing transverse energy and is very important in measurements using selections on a high number of jets.

Table 6.16: The expected uncertainties, expressed in percent, in the cross sections measured in the diphoton fiducial, VBF-enhanced, $N_{\text{lepton}} \geq 1$, $t\bar{t}H$-enhanced, and high $E_{\text{T}}^{\text{miss}}$ regions. The fit systematic uncertainty includes the effect of the photon energy scale and resolution, and the impact of the background modeling on the signal yield. The theoretical modeling uncertainty is defined as the envelope of the signal composition, the modeling of Higgs boson transverse momentum and rapidity distribution, and the uncertainty of parton shower and the underlying event as described in Section 6.7.4.

Source	Uncertainty in fiducial cross section				
	Diphoton	VBF-enhanced	$N_{\text{lepton}} \geq 1$	$t\bar{t}H$-enhanced	High $E_{\text{T}}^{\text{miss}}$
Fit (stat.)	17%	22%	72%	176%	53%
Fit (syst.)	6%	9%	27%	138%	13%
Photon energy scale & resolution	4.3%	3.5%	3.1%	10%	4.1%
Background modeling	4.2%	7.8%	26.7%	138%	12.2%
Photon efficiency	1.8%	1.8%	1.8%	1.8%	1.9%
Jet energy scale/resolution	-	8.9%	-	4.5%	6.9%
b-jet flavor tagging	-	-	-	3%	-
Lepton selection	-	-	0.7%	0.2%	-
Pileup	1.1%	2.9%	1.3%	2.5%	2.5%
Theoretical modeling	0.1%	4.5%	4.0%	8.1%	31%
Signal composition	0.1%	4.5%	3.1%	8.1%	25%
Higgs boson p_{T}^{H} & $\lvert y_H \rvert$	0.1%	0.9%	0.2%	0.7%	0.1%
Underlying Event / Parton Shower	-	0.3%	0.7%	1.1%	31%
Luminosity	3.2%	3.2%	3.2%	3.2%	3.2%
Total	18%	26%	77%	224%	63%

6.9 Comparisons with other results

6.9.1 ATLAS $H \to ZZ^* \to 4\ell$ results

ATLAS has also reported measurements of Higgs boson properties in the four-lepton decay channel, $H \to ZZ^* \to 4\ell$, where ℓ is an electron or muon, using a data set of 36.1 fb^{-1} [224]. The ratio of the stage-0 cross sections, $(\sigma \times \mathcal{B})/(\sigma \times \mathcal{B})_{SM}$, for $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ are shown in Table 6.17 with the uncertainties given as (stat.)+(exp.)+(theo.). The smaller statistical uncertainty from the $\gamma\gamma$ decay channel results in more precise measurements for ggH and VBF as the systematic uncertainties from both experiment and theory are similar. For the VH and top measurements, the $H \to ZZ^* \to 4\ell$ analysis only sets upper limits at the 95% CL due to poor sensitivity while the $\gamma\gamma$ analysis measures the cross sections with large uncertainties. All results are in agreement with their Standard Model expectations, with the exception of the VBF measurements where the results in both analyses disagree with the SM expectations by 2σ or more.

Table 6.17: Best-fit values and uncertainties for the ratio $(\sigma \times \mathcal{B})/(\sigma \times \mathcal{B})_{SM}$. The upper limits correspond to the 95% CL obtained using the CL$_s$ method. The uncertainties are given as (stat.)+(exp.)+(theo.).

Process	$(\sigma \times \mathcal{B})/(\sigma \times \mathcal{B})_{SM}$ (Observed)	
($\lvert y_H \rvert < 2.5$)	$H \to \gamma\gamma$	$H \to ZZ^* \to 4\ell$
ggH	$0.80^{+0.16+0.07+0.05}_{-0.16-0.06-0.04}$	$1.11^{+0.22+0.07+0.04}_{-0.24-0.06-0.04}$
VBF	$2.0^{+0.5+0.3+0.4}_{-0.5-0.3-0.3}$	$4.0^{+1.7+0.3+0.3}_{-1.4-0.3-0.3}$
VH	$0.7^{+0.9+0.2+0.2}_{-0.7-0.2-0.0}$	< 3.7
Top	$0.6^{+0.6+0.2+0.2}_{-0.5-0.2-0.0}$	< 7.5

The fiducial cross sections for the $H \to ZZ^* \to 4\ell$ analysis are reported depending on lepton flavor. These results are unique to the 4ℓ final state and cannot be compared with the $H \to \gamma\gamma$ results. The total cross section obtained in the $H \to ZZ^* \to 4\ell$

analysis is:

$$\sigma_{4\ell,\text{total}} = 3.62 \ \pm 0.50\,(\text{stat.})^{+0.25}_{-0.20}\,(\text{syst.})\,\text{fb}, \tag{6.9}$$

which can be compared to the theory prediction of 2.91 ± 0.13 fb. The analogous $H \to \gamma\gamma$ result is:

$$\sigma_{\gamma\gamma,\text{total}} = 55 \pm 9\,(\text{stat.}) \pm 4\,(\text{syst.})\,\text{fb}, \tag{6.10}$$

which can be compared to 64 ± 2 fb. Both results are in agreement with the SM expectations and have similar precision.

6.9.2 CMS $H \to \gamma\gamma$ results

The CMS collaboration has also reported measurements of Higgs boson properties in the $\gamma\gamma$ decay channel using a data set of 35.9 fb^{-1} [225]. All results have similar precision to those reported by the ATLAS collaboration. The signal strength measurements reported by CMS are shown in Figure 6.15. The ggH and VBF measurements are in agreement with the SM predictions while the VH and $t\bar{t}H$ measurements disagree by less than 2σ.

6.10 Conclusions

Measurements of Higgs boson cross sections in the Higgs boson to diphoton decay channel are performed using 36.1 fb^{-1} of p-p collision data collected at $\sqrt{s} = 13$ TeV with the ATLAS detector at the LHC in 2015 and 2016. The global signal strength measurement improves on the precision of the previous ATLAS measurement in the diphoton channel by a factor of two [146]. The ggH (VBF) signal strength is measured to be 1σ below (2σ above) the Standard Model expectation. The precision of the

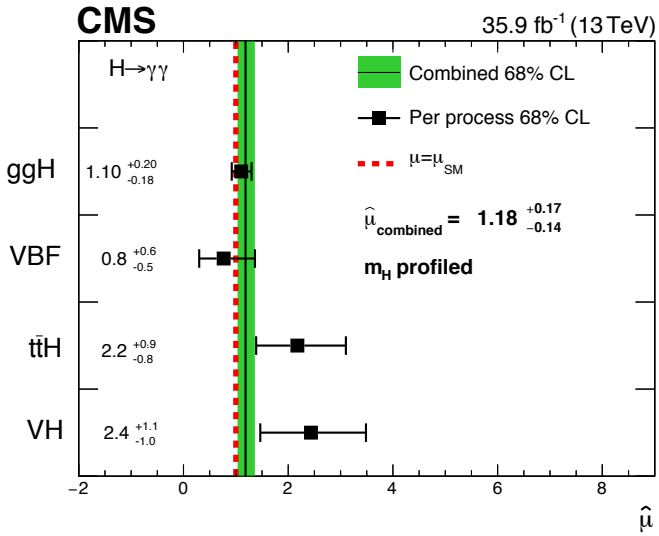


Figure 6.15: Signal strength measurements reported for each process (black points), compared to the global signal strength measurement (green band) and to the SM expectation (dashed red line) [225].

coupling-strength measurements involving top quarks improves by about a factor of three compared to the previous ATLAS measurement in the diphoton channel. These improvements result from a combination of the larger Higgs boson sample collected at $\sqrt{s} = 13$ TeV, from the use of multivariate techniques to target the VBF, VH, and top-quark associated production modes more efficiently, from the improved precision of the ggH Standard Model theory predictions and from a significant reduction of some of the experimental uncertainties such as the photon energy resolution.

Production mode cross section measurements and nine measurements of the so-called simplified template cross sections are also reported. All results show a similar level of agreement as that obtained with the coupling-strength measurements and with the Standard Model expectation. The main theoretical uncertainties in the simplified template cross sections results are the perturbative QCD uncertainties and the uncertainty in the modeling of the parton shower, underlying event and hadroniza-

tion, which cause events to migrate between different bins. The largest experimental uncertainties arise from sources such as the jet energy calibration and the jet energy resolution.

The fiducial cross section measurements reported include the total fiducial cross section, which is in agreement with the Standard Model expectation, as well as cross sections in fiducial regions probing Higgs boson production from vector-boson fusion or associated with large missing transverse momentum, leptons or top quarks. In these measurements, the uncertainty in the fitted signal yield, due to the background modeling and the photon energy resolution, is typically more important than the uncertainty in the correction factor due to the theoretical modeling. The jet energy scale and resolution uncertainties, as well as the uncertainty in the modeling of the parton shower and the underlying event become increasingly important in high-jet multiplicities and in the $t\bar{t}H$- and VBF-enhanced phase space.

Eleven differential cross sections are reported for events belonging to the inclusive diphoton fiducial region, as a function of kinematic variables of the diphoton system or of jets produced in association with the Higgs boson. The reported cross sections are sensitive to the Higgs boson production kinematics, the jet kinematics, the spin and CP quantum numbers of the Higgs boson, and the VBF production mechanism. All measured differential cross sections are compared to predictions and no significant deviation from the Standard Model expectation is observed.

The estimate of the uncertainty in the modeling of the parton shower, underlying event and hadronization presented in this chapter sets a baseline to calculate the impact of this uncertainty in future cross section measurements. It is particularly challenging in categories or regions with large missing transverse energy and in measurements using selections on jets. This effect is observed in the reported VBF signal strength measurement, where the relative theoretical uncertainty is a factor of two

larger than the one observed in the ggH signal strength measurement.

All reported results are statistically limited and their precision will further improve with the full data set to be recorded during Run 2 of the LHC.

Chapter 7

Conclusions

The hunt for the Higgs boson was the centerpiece of the physics programs for the experiments at the Large Hadron Collider during Run 1 of data-taking. The discovery of this particle, announced on July 4^{th} 2012 by the ATLAS and CMS collaborations, represented a milestone in clarifying the mechanism of electroweak symmetry breaking, by which fundamental particles acquire mass. It is now essential that the Higgs boson is extensively studied. Precise measurements of its properties will confirm its nature, and any deviations from the Standard Model prediction will represent a clear sign of new physics.

These goals can not be achieved without a good understanding of the detector. The performance studies reported in this thesis focus on validating the reconstruction and calibration methods of the ATLAS Tile Calorimeter by using the calorimeter response to single hadrons with Run 1 data. Results presented in this thesis show that the double ratio of the mean value ($\langle E/p \rangle$) between data and MC simulation is approximately one, with deviations from unity of less than 5% possibly due to poor electromagnetic scale calibration in the data or differences in the MC description due to a relatively complex hadron shower development. In the Long Barrel region

($|\eta| < 0.7$) the 3% level agreement is maintained despite sizeable changes in beam conditions.

Higgs boson pair production is the simplest production process that is sensitive to the self-coupling λ, and it also provides one with a wealth of possibilities for probing higher-dimensional interactions as well as the existence of heavier states coupled to the Higgs boson. This thesis presented a search for resonant and non-resonant Higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state with data collected at $\sqrt{s} = 13$ TeV with the ATLAS detector. No significant deviations from the Standard Model predictions are observed. The observed (expected) 95% CL upper limit on the cross section for non-resonant production is 0.73 pb (0.93 pb) and corresponds to 22 (28) times the predicted SM cross section, which improves the previous ATLAS Run 2 non-resonant result in a factor of five. The Higgs boson self-coupling is constrained at 95% CL to -8.2 $< \kappa_\lambda <$ 13.2 whereas the expected limits are -8.3 $< \kappa_\lambda <$ 13.2. For resonant production of $X \to HH \to \gamma\gamma b\bar{b}$, a limit is presented for the narrow-width approximation as a function of m_X. The observed (expected) limits range between 1.14 pb (0.90 pb) and 0.12 pb (0.15 pb) in the range 260 GeV $< m_X <$ 1000 GeV.

This thesis also presented the measurement of the total Higgs boson production-mode cross sections, signal strengths, and simplified template cross sections (STXS), as well as the measurement of the fiducial and differential cross sections in the diphoton decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector. All measurements assume a Higgs boson mass of 125.09 ± 0.24 GeV. The measured signal strength relative to the Standard Model expectation is found to be $\mu = 0.99\ ^{+0.15}_{-0.14} = 0.99\ \pm 0.12\,(\text{stat.})\ ^{+0.06}_{-0.05}\,(\text{exp.})\ ^{+0.07}_{-0.05}\,(\text{theo.})$. This result confirms the ATLAS Run 1 diphoton signal strength measurement of $\mu = 1.17 \pm 0.23\,(\text{stat.})\ ^{+0.10}_{-0.08}\,(\text{exp.})\ ^{+0.12}_{-0.08}\,(\text{theo.})$ with around a factor of two improvement in each component of the uncertainty. The total fiducial cross section is measured to be $\sigma_{\text{fid}} = 55 \pm 9\,(\text{stat.})\ \pm 4\,(\text{exp.})\ \pm 0.1\,(\text{theo.})$ fb,

which is in agreement with the Standard Model expectation of 64 ± 2 fb. The precision of these measurements is currently dominated by their statistical uncertainties, but it is expected to improve in the next years of the LHC, as more data is collected the statistical uncertainty will decrease. Accurate estimates of the systematic uncertainties will therefore become essential, in particular of the uncertainty in the modeling of the parton shower, underlying event and hadronization, which is especially challenging due to the difficulty of generating sufficient events for each event reconstruction category, fiducial region, or each bin of a fiducial differential cross section. The work I carried out in the scope of this thesis laid the foundation of the estimate of this uncertainty which will be improved in the upcoming years.

Chapter 8

Resum en valencià

8.1 Fonaments teòrics

8.1.1 El Model Estàndard

El Model Estàndard de física de partícules és una teoria quàntica de camps que descriu les partícules elementals i les seues interaccions fonamentals, i unifica tres de les quatre forces de la natura, electromagnètica, nuclear feble i nuclear forta, en una de sola.

La matèria fermiònica està formada per leptons i quarks d'espín $\frac{1}{2}$, que s'agrupen en tres generacions de massa creixent, cadascun dels quals està compost per dos elements (un carregat electrònicament i l'altre, neutre) que formen un doblet. Aquests doblets de leptons son: l'electró (e) i el neutrí electrònic (ν_e), el muó (μ) i el neutrí muònic (ν_μ), i el tauó (τ) i el neutrí tauònic (ν_τ). Els doblets de quarks reben el nom de *up* i *down* (u,d), *charm* i *strange* (c,s), i *top* i *bottom* (t,b). Els quarks tenen càrrega elèctrica fraccionaria respecte a la de l'electró, $+\frac{2}{3}$ per als quarks de tipus *up* (u,c,t) i $-\frac{1}{3}$ per als quarks de tipus *down* (d,s,b). Els quarks tenen un nombre quàntic addicional, el color, que té tres possibles valors: roig, verd i blau. El confinament de

color és la propietat física per la qual els quarks estan confinats en estats neutres de color, anomenats hadrons.

Les interaccions actuen per mitjà de partícules transmissores, que s'anomenen genèricament bosons de gauge i tenen espín igual a 1. Aquests bosons sorgeixen com una manifestació del grup de simetria de la teoria, sent aquest $SU(3) \otimes SU(2) \otimes U(1)$ per al Model Estàndard, després d'aplicar una invariància gauge local als camps fermiònics. La interacció electromagnètica és transmesa pel fotó (γ). La interacció feble la transmeten els tres bosons febles ($W^{\pm}$ i Z^0). La interacció forta és mitjançada pels vuit gluons (g).

8.1.2 El bosó de Higgs

El 4 de juliol de 2012, els experiments ATLAS i CMS anunciaren el descobriment d'una nova partícula amb una massa al voltant de 125 GeV [37, 38]. La confirmació de l'existència del bosó de Higgs fou fonamental a l'hora d'entendre el mecanisme de trencament espontani de simetria del Model Estàndard, introduint un nou camp de forces, anomenat camp de Higgs, que en interaccionar amb els bosons $W^{\pm}$ i el Z^0 genera la massa d'aquestes dues partícules. El camp de Higgs té els mateixos nombres quàntics que el buit físic, no té càrrega elèctrica ni càrrega forta (color), de manera que no interacciona amb els fotons i gluons. És un camp que impregna tot l'espai-temps i frena el moviment de les partícules que hi interactuen, de manera que aquestes adquireixen massa. L'autointeracció del camp de Higgs amb si mateix genera també una massa per al bosó de Higgs.

Les masses dels leptons carregats i dels quarks es generen també mitjançant la seua interacció amb el camp de Higgs. El variat espectre de masses que mostren els constituents de la matèria reflecteix les molt diferents intensitats que tenen les seues corresponents interaccions amb el camp de Higgs: a més massa més interacció.

Els principals modes de producció del bosó de Higgs en un col·lisionador d'hadrons es mostren en la Figura 8.1 i es detallen a continuació:

- La fusió gluó-gluó (ggH) és el principal mode de producció del bosó de Higgs a l'LHC i representa el 87% de la seua producció total a $\sqrt{s} = 13$ TeV. Aquest mecanisme es troba dominat per bucles de quarks *top* degut a l'alta massa d'aquest tipus de quark.

- La fusió d'un bosó vectorial (VBF) correspon a la segona major contribució a la producció del bosó de Higgs (6.8%) i es caracteritza per tindre dos dolls de partícules (jets) amb alt moment transversal en la direcció pròxima a l'eix del feix, originats pels dos quarks que interaccionen. El bosó de Higgs apareix entre els jets, en la regió central del detector.

- La producció associada amb un bosó vectorial (WH, ZH) és el tercer mecanisme de producció més comú i correspon al 4% de la secció eficaç de producció del bosó de Higgs. Aquest procés també s'anomena Higgs-Strahlung degut a l'analogia existent entre la radiació d'un bosó de Higgs i la de Bremsstrahlung.

- La producció de bosons de Higgs en associació amb quarks *top* ($t\bar{t}H$ i tH) té una de les taxes de producció del bosó de Higgs més baixes a l'LHC. Aquest mode de producció és especialment difícil de mesurar degut a l'existència de leptons i quarks *bottom* en els estats finals.

El bosó de Higgs és inestable; es desintegra molt ràpidament en una partícula i la seua corresponent antipartícula. Els canals de bona resolució de massa, $\gamma\gamma$ i ZZ, són particularment importants doncs permeten mesurar la massa de la nova partícula amb precisió. Al canal $\gamma\gamma$ la massa es determina a partir de les energies i direccions de dos fotons d'alta energia. Al canal ZZ, la massa es determina a partir de la desintegració

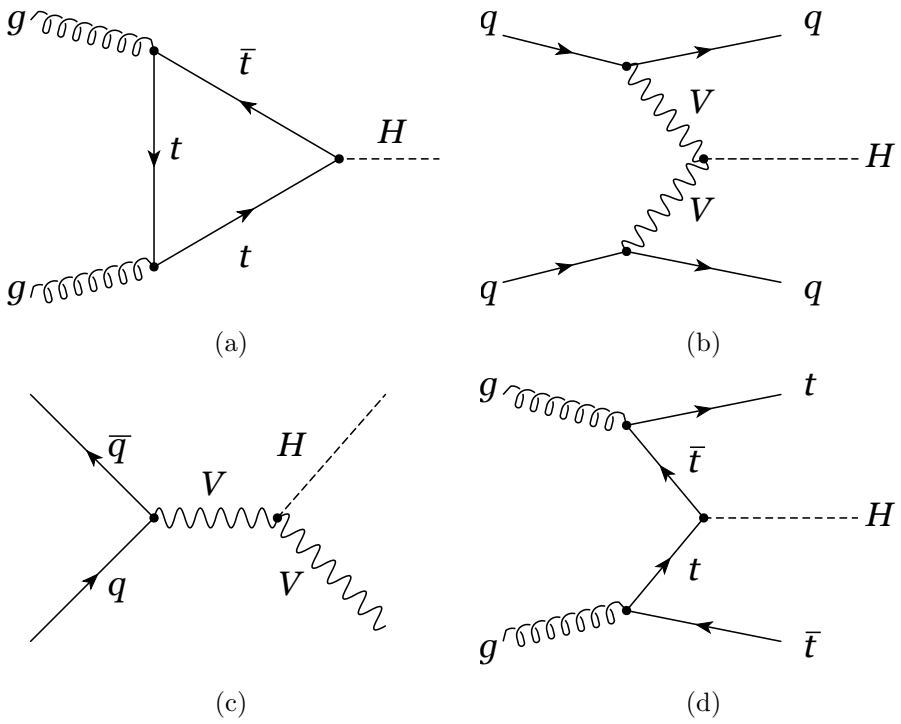


Figure 8.1: Diagrames de Feynman dels modes de producció del bosó de Higgs, mitjançant a) la fusió gluó-gluó a través de bucles de quarks pesats, b) la fusió d'un bosó vectorial (W o Z) amb dos jets, c) la radiació d'un bosó de Higgs (Higgs-Strahlung) a partir d'un bosó V (W o Z) i d) la producció de bosons de Higgs en associació amb quarks *top*.

dels dos bosons Z en dos parells d'electrons o de muons, o en un parell d'electrons i un parell de muons. La mesura del canal WW és més complexa. Cada partícula W s'identifica a partir de la seua desintegració en un electró i un neutrí o en un muó i un neutrí. Els neutrins atravessen els detectors sense ser detectats, de manera que el Higgs es manifesta al canal WW com un excés ampli en la distribució de massa en lloc d'un pic estret. El canal $b\bar{b}$ és el canal de desintegració més probable però té sorolls de fons grans que són deguts a processos del Model Estàndard, de manera que l'anàlisi cerca esdeveniments en els quals el bosó de Higgs es produeix en associació amb un bosó W o Z, que a continuació es desintegra en electrons i/o muons. El canal

$\tau\tau$ es mesura mitjançant l'observació del decaïment en electrons, muons i hadrons.

8.1.3 Producció de parells de bosons de Higgs

La producció de parells de bosons de Higgs (HH) [41] juga un paper clau en la comprensió de l'estructura del potencial subjacent que resulta en el mecanisme de trencament espontani de simetria. No solament és el procés de producció més senzill que és sensible a l'autoacoblament λ, sinó que també proporciona una gran quantitat de possibilitats per investigar interaccions multidimensionals, així com l'existència d'estats més pesats acoblats al Higgs.

Desafortunadament, en el Model Estàndard les taxes de producció de parells de bosons de Higgs en l'LHC es veuen notablement reduïdes, per tant, una mesura de les seccions eficaces de producció d'HH necessita una considerable lluminositat integrada, fins i tot a 13–14 TeV d'energia del centre de massa. Tanmateix, nova física pot produir millores importants en diversos escenaris. Aquestes millores es classifiquen en dues categories bàsiques:

- No ressonant: quan es modifica qualsevol dels acoblaments del Model Estàndard es varia la taxa de producció esperada de parells de bosons de Higgs. D'aquesta manera, desactivar l'autoacoblament λ_{HHH} aconseguiria doblar la taxa de producció de parells de bosons de Higgs en el procés ggH, i canviar el seu signe quadruplicaria aquesta taxa [42]. L'habilitació directa del vertex $tthh$ també conduiria a una millora significativa de la taxa p-$p \to HH$ [43].

- Ressonant: moltes teories més enllà del Model Estàndard (BSM) preveuen l'existència de partícules pesades que poden decaure en un parell de bosons de Higgs. Aquestes podrien identificar-se com una ressonància en l'espectre de la massa invariant del parell de bosons de Higgs. Es podrien produir, per exemple, a través del procés ggH amb l'aparició d'una nova partícula intermèdia,

"X". Els models amb dos doblets de Higgs [45], com l'extensió supersimètrica mínima del Model Estàndard (MSSM) [46], els models de Higgs bessons [47] i els models de Higgs compostos [48, 49] impliquen l'addició d'un segon doblet escalar complex. Això requereix l'existència d'un bosó de Higgs pesat que es desintegraria en dos dels seus socis més lleugers, del tipus del Model Estàndard.

8.2 L'LHC i l'Experiment ATLAS

8.2.1 L'LHC

L'LHC [57], acrònim de *Large Hadron Collider*, és l'accelerador de partícules més gran del món, on dos feixos d'hadrons col·lideixen en el centre d'uns experiments. L'LHC és un túnel de 27 km de circumferència, situat entre les fronteres de França i Suïssa, al nord-oest de la ciutat de Ginebra.

Ha sigut dissenyat per assolir energies del centre de massa de fins a 14 TeV i una lluminositat de 10^{34} cm^{-2} s^{-1} amb una col·lisió cada 25 ns. Els quatre experiments principals, construïts al voltant de l'anell, són: ATLAS [62], CMS [64], LHCb [65] i ALICE [63].

8.2.2 L'Experiment ATLAS

L'experiment ATLAS (A Toroidal LHC Apparatus) és un detector amb objectius generals en operació a l'LHC. Mesura 45 m de llargària i 25 m d'alçària, i pesa més de 7000 tones. Està dividit en sub-detectors i construït amb tecnologies sofisticades i materials especialitzats, tal com es mostra en la Figura 8.2.

ATLAS utilitza un sistema de coordenades a dretes, amb l'origen ubicat al punt d'interacció. La direcció del feix defineix l'eix z. El costat A (costat C) del detector es defineix com aquell que té z positiva (negativa). El sentit positiu de l'eix x es defineix

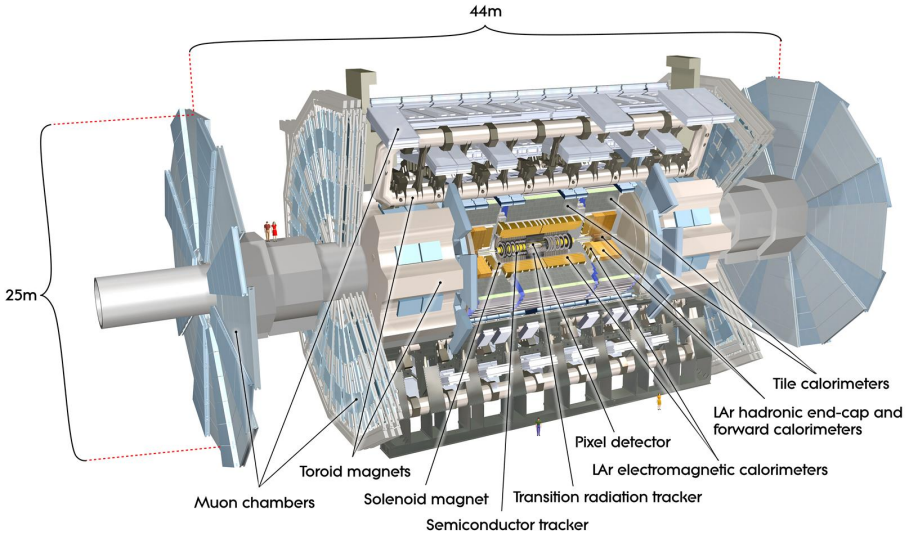


Figure 8.2: Diagrama del detector ATLAS i els seus components principals [69].

com aquell que apunta al centre de l'anell de l'LHC, i el sentit positiu de l'eix y com aquell que apunta cap amunt. El pla $x - y$ és ortogonal a la direcció del feix i es coneix com el pla transversal. Quantitats com el moment transversal, p_{T}, o l'energia transversal, E_{T}, es defineixen en aquest pla transversal. L'angle azimutal ϕ es mesura al voltant de l'eix del feix i l'angle polar θ és l'angle mesurat des de l'eix del feix. En el límit on la partícula viatja quasi a la velocitat de la llum, o equivalentment en l'aproximació que la massa de la partícula és insignificant, la rapidesa, y, i la pseudo-rapidesa, η, es defineixen com:

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \text{ i } \eta = -\ln \tan \left(\frac{\theta}{2} \right), \tag{8.1}$$

on E denota l'energia i p_z el component del moment al llarg de la direcció del feix. El límit de partícules sense massa correspon a $y = \eta$, on ambdues són invariants sota transformacions de Lorentz al llarg de l'eix del feix. La distància ΔR en l'espai de

coordenades $\eta - \phi$ s'utilitza comunament i es defineix com:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \tag{8.2}$$

Des del 2010 fins al 2012 ATLAS recollí un conjunt de dades de 5.4 fb^{-1} i 20.3 fb^{-1} a partir de les col·lisions de protons que tingueren lloc a l'LHC amb energies del centre de massa de 7 i 8 TeV, respectivament. Aquest correspon al primer període de presa de dades de l'experiment ATLAS i rep el nom de Run 1. El segon període de presa de dades, l'anomenat Run 2, començà el 2015 amb col·lisions de protons i energia del centre de massa de 13 TeV. Des del 2015 fins al 2017 ATLAS recollí un conjunt de dades de 86.3 fb^{-1}. El Run 2 està encara en procés i acabarà a final del 2018.

Els sub-detectors d'ATLAS

El detector ATLAS està constituït pels següents sistemes:

- El sistema d'imants està optimitzat per a desviar les trajectòries de les partícules carregades minimitzant els efectes de dispersió múltiple. El sistema magnètic està compost d'un solenoide central per al detector intern, amb un camp axial de 2 T, un toroide de barril i dos toroides externs que generen un camp magnètic tangencial de 0.5 T i 1 T, respectivament.

- El sistema de muons cobreix el rang $|\eta| < 3$, i fa un ús eficient del poder de separació dels imants. Consta de quatre sub-detectors en funció de les necessitats de resolució espacial i temporal: MDT, CSC, RPC i TGC.

- El detector intern (ID) està dissenyat per reconstruir traces i vèrtexs primaris de desintegració amb alta eficiència. Consta de tres sub-detectors: Píxel, SCT i TRT.

- El sistema de calorímetres mesura la posició i l'energia dipositada pels jets, en el rang $|\eta| < 4.9$, utilitzant una àmplia varietat de tècniques. La part més interna correspon al calorímetre d'argó líquid (LAr), el qual es compon d'un cilindre al voltant de l'eix del feix, que actua com a calorímetre electromagnètic (EM), i uns taps als costats (HEC), que actuen com a calorímetres hadrònics. Al voltant del cilindre se situa el TileCal, o calorímetre de teules, que està format per cel·les de material centellejador i acer, i actua com a calorímetre hadrònic.

El sistema de filtrat i d'adquisició de dades

El sistema de filtrat (*trigger*) d'ATLAS té com a finalitat seleccionar els esdeveniments més interessants des del punt de vista físic i reduir d'aquesta manera la quantitat de dades per a emmagatzemar. Així es passa de tindre 40 milions d'esdeveniments per segon a tan sols uns 1000. Hi ha dos nivells en la selecció: el primer es basa en la informació sobre l'electrònica del detector (hardware), en particular dels calorímetres i de les càmeres de muons, mentre que el segon està basat en software i s'executa en un gran clúster situat a prop del detector.

Pel que fa a l'operació del detector ATLAS, existeixen dos sistemes integradors principals complementaris: el sistema DAQ (*Data AdQuisition*) i el sistema DCS (*Detector Control System*). Mentre el primer realitza l'adquisició de les dades que seran emprades en els consegüents estudis de física, el segon gestiona tota la infrastructura relacionada amb l'estat operacional del detector assegurant així la correcta extracció d'informació.

8.3 Reconstrucció d'esdeveniments

ATLAS identifica electrons, fotons, muons, jets i tauons mitjançant un conjunt d'algorismes especialitzats que converteixen les dades provinents del detector en

objectes de física fonamentals. Aquests algorismes reben el nom de reconstrucció d'esdeveniments.

8.3.1 Fotons i electrons

La interacció d'un fotó o un electró amb el calorímetre electromagnètic es troba a l'origen del procés de pluja electromagnètica, en el qual una quantitat significativa d'energia es diposita en una petita quantitat de cel·les properes en el calorímetre. La reconstrucció de fotons i electrons es desenvolupa de la següent manera:

- Construcció de clústers: la reconstrucció de fotons i electrons comença amb la construcció de clústers a partir dels dipòsits d'energia en el calorímetre electromagnètic. Les cel·les del calorímetre s'agrupen en torres de dimensions $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ mitjançant un algorisme de finestra corredissa (*sliding-window*) [82] amb finestres de mida 3×5 en el pla $\eta - \phi$ i moment transversal $p_{\mathrm{T}} > 2.5$ GeV.

- Reconstrucció de traces: una vegada que es reconstrueixen els clústers, es realitza una cerca de les traces en el detector intern que coincideixen amb els clústers, amb la finalitat d'identificar i reconstruir els fotons convertits i els electrons.

- Conversió de traces: aquelles traces compatibles amb les provinents d'un procés de conversió de fotons s'utilitzen per a crear candidats a vèrtexs convertits.

- Concordança entre traces i clústers: es realitza una concordança entre els candidats a vèrtexs convertits i els clústers.

- Creació del clúster final: un algorisme final decideix si un clúster correspon a un fotó no convertit, a un fotó convertit o a un únic electró basant-se en la

concordança amb les traces i vèrtexs convertits, i en el quadrimoment del clúster i les traces.

8.3.2 Muons

Els muons es poden identificar i reconstruir utilitzant informació provinent del sistema de muons, del detector intern i del calorímetre electromagnètic. A més, els muons han de satisfer una sèrie de requisits de qualitat, com el requeriment en el nombre mínim d'impactes (*hits*) presents en cadascun dels sub-detectors. D'aquesta manera es proporcionen quatre seleccions d'identificació de muons (*Loose*, *Medium*, *Tight* i *High-*p_{T}) per fer front a les necessitats específiques de les diferents anàlisis de física. *Loose*, *Medium* i *Tight* són categories inclusives en les quals els muons identificats amb requisits més restrictius també estan inclosos en les categories amb requeriments menys estrictes.

8.3.3 Jets

Els jets, dolls d'hadrons col·limats, són els objectes de física dominants que es produeixen en col·lisions de protons a l'LHC. Els jets es reconstrueixen a partir de les cel·les en els calorímetres utilitzant l'algorisme anti-$\mathrm{k_T}$ [98] amb un paràmetre de distància $R = 0.4$. Posteriorment se'ls aplica totes les correccions d'energia necessàries amb la finalitat de reduir els efectes del pile-up, del soroll electrònic i de les variacions relatives del moment segons la regió del detector, entre d'altres.

8.3.4 *b*-jets

La taxa d'identificació d'un b-jet defineix l'eficiència d'etiquetatge dels jets (b-tagging) de tipus *bottom* per a un determinat algorisme. El sabor d'un jet es defineix a partir de dades simulades utilitzant un con espacial ΔR amb l'objectiu de trobar coincidències

entre hadrons estables i jets reconstruïts. Es realitza una concordança jeràrquica. En primer lloc, es comprova si coincideix amb un hadró de tipus *bottom*, després amb un hadró de tipus *charm* i, per finalitzar, amb un tauó. Aquest procediment de concordança fa que un jet siga classificat, respectivament, com a b-jet, c-jet, τ-jet o *light*-jet (jet lleuger) si no es troba cap coincidència.

L'experiment ATLAS utilitza diversos algorismes [107] basats en informació sobre traces reconstruïdes i vèrtexs secundaris desplaçats en el detector intern per a identificar b-jets gràcies a la vida mitjana relativament llarga dels hadrons de tipus *bottom*.[1] L'algorisme utilitzat en aquest treball rep el nom de MV2c10 i està optimitzat per a discriminar b-jets de *light*-jets i c-jets. D'aquesta manera es defineixen diferents punts de treball depenent de l'eficiència de reconstrucció dels b-jets i la capacitat de rebutjar qualsevol altre tipus de jets.

8.3.5 $E_{\mathbf{T}}^{\mathbf{miss}}$

Els neutrins produïts en col·lisions de protons no interactuaran amb el detector ATLAS i, per tant, no es detectaran. Es pot inferir informació indirecta sobre la presència d'aquestes partícules imposant la conservació del moment en el pla transversal a la direcció del feix. Com que l'estat inicial té component de moment nul en el pla transversal, un desequilibri en el moment transversal total mesurat en l'estat final indicaria la presència d'una partícula invisible. Així, l'energia transversal perduda es defineix com la suma vectorial negativa del moment transversal de tots els objectes seleccionats, més qualsevol altra activitat en el calorímetre al voltant d'aquests.

[1] La vida mitjana dels hadrons de tipus B és de l'ordre de 1.6 ps. Tenint en compte que la massa d'un hadró de tipus B és de 5 GeV i la seua energia és de 30 GeV, la típica trajectòria de vol és d'aproximadament 3 mm.

8.4 Resposta del TileCal a partícules carregades i aïllades

El rendiment del TileCal es prova *in situ* utilitzant la resposta del calorímetre als hadrons aïllats. Aquests objectes, juntament amb els jets, dipositen més energia en el calorímetre hadrònic que els muons i, per tant, es pot aconseguir provar la resposta a energies superiors. Se seleccionen de les dades recollides a partir de les col·lisions de protons que tingueren lloc a l'LHC des del 2010 fins al 2012, amb energies del centre de massa de 7 i 8 TeV.

La resposta es caracteritza pel quocient entre l'energia d'un hadró carregat (E), mesurat pel TileCal, i el moment de l'hadró (p), mesurat pel detector intern. L'energia d'un hadró tant en dades reals com en simulació de Monte Carlo (MC) es calibra a l'escala de l'energia electromagnètica (EM). El quocient doble del valor mitjà ($\langle E/p \rangle$) entre les dades i la simulació de MC hauria de ser proper a la unitat, amb desviacions d'aquesta possiblement a causa d'una mala calibració de l'escala electromagnètica en les dades o de diferències en la descripció del MC per culpa d'un desenvolupament de pluja hadrònica relativament complex.

8.4.1 Selecció d'esdeveniments

El candidat a hadró ha d'associar-se amb un dipòsit d'energia en el calorímetre. L'energia de l'hadró es defineix com la suma de l'energia de les cel·les calibrades a l'escala EM que pertanyen a tots els grups de cel·les (*topo-clústers*) que tenen el seu centre dins d'un con de $\Delta R = 0.2$ al voltant de la traça extrapolada al calorímetre. La selecció d'esdeveniments consta dels següents requeriments:

- Pre-selecció: es requereix que hi haja almenys una traça amb $p_{\mathrm{T}} > 2$ GeV que no s'extrapoli a un mòdul mort, per a la qual cosa s'utilitza l'eina TileTripReader

(TTR) [126].

- Impactes (*hits*) al detector intern: almenys un al Píxel, un al TRT i sis al SCT.
- Les traces han d'estar dins dels límits del TileCal ($|\eta| < 1.7$) quan s'extrapolen a totes les capes del calorímetre.
- Els paràmetres d'impacte (d_0 i z_0) han de complir els requeriments $|d_0| < 1.5$ mm i $|z_0 \sin(\theta)| < 1.5$ mm, respectivament.
- Cal que el nombre mitjà d'interaccions per encreuament de paquets, $\langle\mu\rangle$, estiga entre 3 i 25 per al conjunt de dades del 2012 per tal que es puga realitzar una comparació equitativa entre esdeveniments de dades i de MC per a qualsevol valor de μ. Aquest requeriment no és necessari per als conjunts de dades del 2010 i 2011.
- $\Delta E(\text{LAr}) < 1$ GeV: l'energia dipositada per l'hadró en LAr és compatible amb una partícula de mínima ionització, caracteritzada per dipositar menys d'1 GeV al llarg del seu recorregut sense un tall d'energia límit que permet que els grups d'energia negativa contribuïsquen a l'anàlisi.
- $R = \Delta E(\text{Tile})/\Delta E(\text{LAr+Tile}) > 0.75$: es requereix que la quantitat d'energia dipositada per l'hadró en el TileCal siga almenys el 75% de la seua energia total per a augmentar la puresa d'hadrons en la mostra.

8.4.2 Resultats

La Taula 8.1 mostra el quocient doble de $\langle E/p\rangle$ entre les dades i la simulació de MC, $(\langle E/p\rangle^{dades}/\langle E/p\rangle^{MC})$, utilitzant les dades del 2010, 2011 i 2012, en funció de la pseudo-rapidesa $|\eta|$, l'angle azimutal ϕ, el moment p i el nombre mitjà d'interaccions per encreuament de paquets, $\langle\mu\rangle$.

Table 8.1: Quocients dobles de $\langle E/p \rangle$ en funció de la pseudo-rapidesa $|\eta|$, l'angle azimutal ϕ, el moment p i el nombre mitjà d'interaccions per encreuament de paquets μ per a dades del 2010, 2011 i 2012. Només es mostren les incerteses estadístiques.

	2010	2011	2012
Quocient doble de $\langle E/p \rangle$ vs η	1.003 ± 0.004	0.936 ± 0.006	0.994 ± 0.004
Quocient doble de $\langle E/p \rangle$ vs ϕ	1.003 ± 0.004	0.936 ± 0.007	0.994 ± 0.004
Quocient doble de $\langle E/p \rangle$ vs p	1.003 ± 0.004	0.936 ± 0.006	0.993 ± 0.005
Quocient doble de $\langle E/p \rangle$ vs $\langle \mu \rangle$	-	0.936 ± 0.005	0.994 ± 0.005

La Figura 8.3 mostra les distribucions de $\langle E/p \rangle$ en funció de η, ϕ, el moment i el nombre mitjà d'interaccions, per a dades del 2012.

8.4.3 Incerteses sistemàtiques

La incertesa sistemàtica en el quocient doble de $\langle E/p \rangle$, $\langle E/p \rangle^{\text{dades}} / \langle E/p \rangle^{\text{MC}}$ és avaluada amb les dades del 2012 variant cada component cap amunt i cap avall en una desviació estàndard. S'utilitzen valors augmentats i disminuïts de les següents variables de la selecció d'esdeveniments: pèrdua d'energia en LAr, energia relativa dipositada en TileCal, R i nombre d'impactes al detector intern.

Els criteris utilitzats per triar els llindars de la pèrdua d'energia en LAr es basen en la població de la corresponent distribució. D'aquesta manera, el límit inferior s'estableix a 1σ per sota del valor mitjà (300 MeV) i el límit superior s'estableix en 1σ per sobre del valor mitjà (1500 MeV). De manera semblant, els criteris per a escollir els llindars en el nombre d'impactes al detector intern es basen en utilitzar el màxim (o mínim) relatiu de les seues distribucions.

El límit superior i inferior de la incertesa sistemàtica s'obté de la suma en quadratura de les diferents contribucions superiors i inferiors. El quocient doble de $\langle E/p \rangle$ corresponent a les dades del 2012 és $\langle E/p \rangle^{\text{dades}} / \langle E/p \rangle^{\text{MC}} = 0.994 \pm 0.003$ (estadístic) $^{+0.009}_{-0.014}$ (sistemàtic).

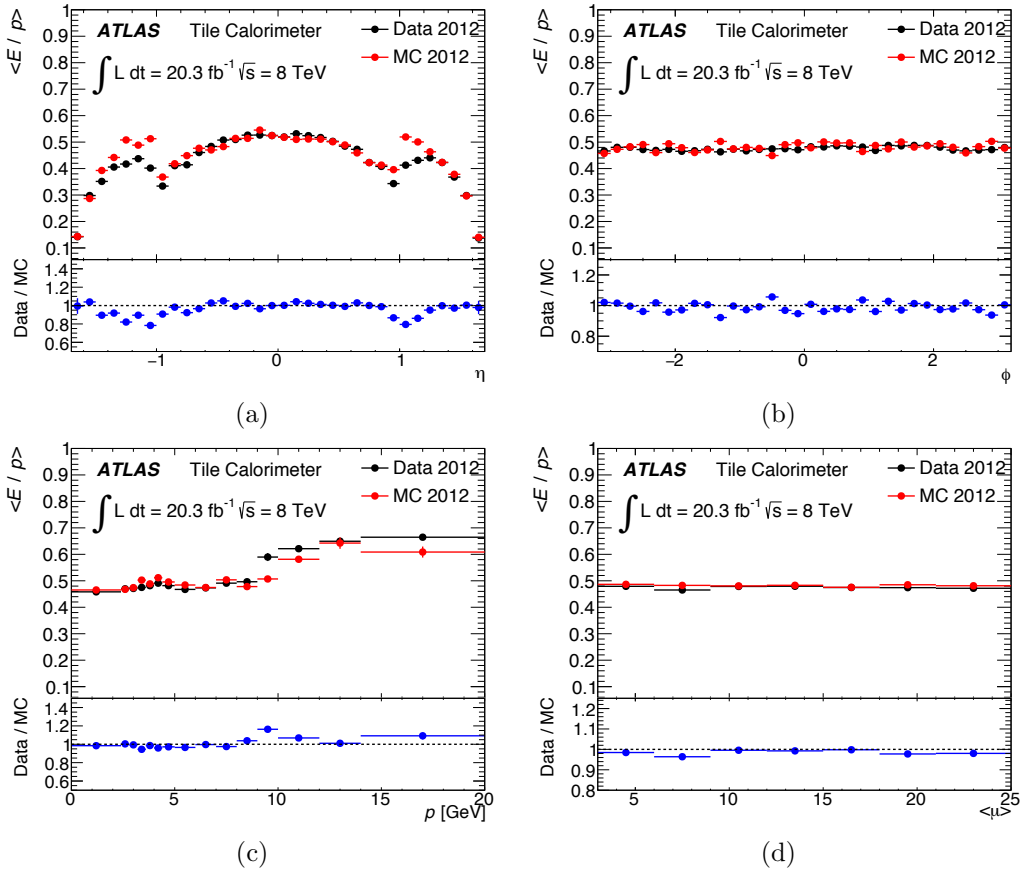


Figure 8.3: Distribucions de $\langle E/p \rangle$ per a hadrons carregats i aïllats en funció de a) la pseudo-rapidesa $|\eta|$, b) l'angle azimutal ϕ, c) el moment p i d) el nombre mitjà d'interaccions per encreuament de paquets $\langle \mu \rangle$, per a dades del 2012. Només es mostren les incerteses estadístiques.

8.5 Recerca de producció de parells de bosons de Higgs en l'estat final $\gamma\gamma b\bar{b}$

En aquesta secció es presenta una recerca de producció ressonant i no ressonant (veure Secció 8.1.3) de parells de bosons de Higgs en l'estat final $\gamma\gamma b\bar{b}$ utilitzant un conjunt de dades de 36.1 fb^{-1} a partir de col·lisions de protons amb una energia del centre de massa de 13 TeV.

En el cas de la cerca no ressonant, qualsevol desviació en l'autoacoblament del bosó de Higgs es caracteritza mitjançant κ_λ ($\kappa_\lambda = \lambda^{HHH}/\lambda^{HHH}_{SM}$) mentre que κ_t caracteritza les desviacions de l'acoblament de Yukawa entre el quark *top* i el bosó de Higgs ($\kappa_t = y_t/y_{t,SM}$). En ambdós casos, el subíndex SM fa referència al valor d'aquests paràmetres en el Model Estàndard. En el cas de la cerca ressonant s'utilitza l'aproximació d'amplària estreta (*narrow-width*), centrant-se en una ressonància de massa (m_X) en el rang 260 GeV $< m_X <$ 1000 GeV.

Encara que aquesta cerca està centrada en la desintegració d'un escalar genèric en un parell de bosons de Higgs, la simulació de MC utilitzada per a optimitzar l'anàlisi correspon al procés ggH.

8.5.1 Selecció i categorització d'esdeveniments

Els esdeveniments se seleccionen per a l'anàlisi si hi ha almenys dos fotons i almenys dos jets, on un o dos dels quals estan etiquetats com a b-jets. Es requereix que el (segon) fotó més energètic del parell complisca la condició $E_{\mathrm{T}}/m_{\gamma\gamma} > 0.35$ (0.25) i que la massa invariant del sistema format pels dos fotons es trobe inicialment dins d'una finestra de massa de 105 GeV $< m_{\gamma\gamma} <$ 160 GeV. Per tal de romandre ortogonal a la cerca de producció de parelles de bosons de Higgs en l'estat final $b\bar{b}b\bar{b}$ amb el detector ATLAS [136], qualsevol esdeveniment amb més de dos b-jets amb una eficiència del 70% és rebutjat. Els esdeveniments restants es divideixen en tres categories:

- La categoria de senyal 2-tag consisteix en aquells esdeveniments amb exactament dos b-jets que satisfan el requisit del 70% d'eficiència.

- La categoria de senyal 1-tag es defineix utilitzant esdeveniments que no compleixen el requisit anterior, però que contenen exactament un b-jet amb una eficiència del 60%. El segon jet de la parella, que en aquest cas no s'identifica com a b-jet, se selecciona amb un mètode anomenat Boosted Decision Tree

(BDT). La BDT utilitza les següents variables cinemàtiques: el moment transversal i la pseudo-rapidesa dels jets, la massa invariant, el moment transversal i la pseudo-rapidesa del sistema format per la parella de jets, i la distància $\Delta\eta$ entre els jets seleccionats. La informació sobre si cada jet satisfà criteris menys estrictes de b-tagging i el rànquing dels jets del millor al pitjor en termes de la coincidència més propera entre la massa invariant del sistema format pels dos jets i m_H també s'utilitzen en aquest model. El jet amb la puntuació de la BDT més alta és seleccionat i l'esdeveniment s'inclou en la categoria de senyal 1-tag. L'eficiència amb la qual la BDT selecciona el segon jet correctament és del 60–80% en tota la gamma d'hipòtesis de senyals ressonants i no ressonants que es consideren en aquesta cerca.

- La categoria de control 0-tag utilitza esdeveniments que no contenen cap b-jet. Aquesta categoria no s'utilitza directament en l'anàlisi, sinó que serveix per a obtindre estimacions basades en dades reals sobre la forma del fons en les categories de senyal.

Seleccions *Loose* i *Tight*

Es defineixen dues seleccions cinemàtiques, *Loose* i $Tight$, on la selecció $Tight$ és un subconjunt més restrictiu de la selecció *Loose*. Les recerques de ressonàncies de massa inferior i de valors més enllà del Model Estàndard de l'autoacoblament del bosó de Higgs utilitzen la selecció Loose, ja que el valor mitjà del moment transversal (p_{T}) dels bosons de Higgs és menor en aquests casos [147]. La selecció $Tight$ s'utilitza per a senyals on els bosons de Higgs solen tenir un valor mitjà de p_{T} més alt, és a dir, per a ressonàncies de massa superior i producció no ressonant de parells de bosons de Higgs.

D'aquesta manera, la selecció *Loose* requereix que el (segon) jet amb moment

transversal més alt complisca la condició $p_{\mathrm{T}} > 40$ (25) GeV, i que la massa invariant del sistema format pel parell de jets (m_{jj}) es trobe entre 80 i 140 GeV. Per a la selecció *Tight*, es requereix que el (segon) jet amb moment transversal més alt complisca la condició $p_{\mathrm{T}} > 100$ (30) GeV, amb 90 GeV $< m_{jj} <$ 140 GeV. Finalment, en la cerca ressonant, la massa invariant del sistema format pels dos fotons ha d'estar dins d'una finestra de massa de 4.7 (4.3) GeV al voltant de la massa del bosó de Higgs per a la selecció *Loose* (*Tight*). Aquesta selecció addicional en $m_{\gamma\gamma}$ està optimitzada per a contenir almenys el 95% dels esdeveniments de parells de bosons de Higgs en la simulació de MC per a cada hipòtesi de massa.

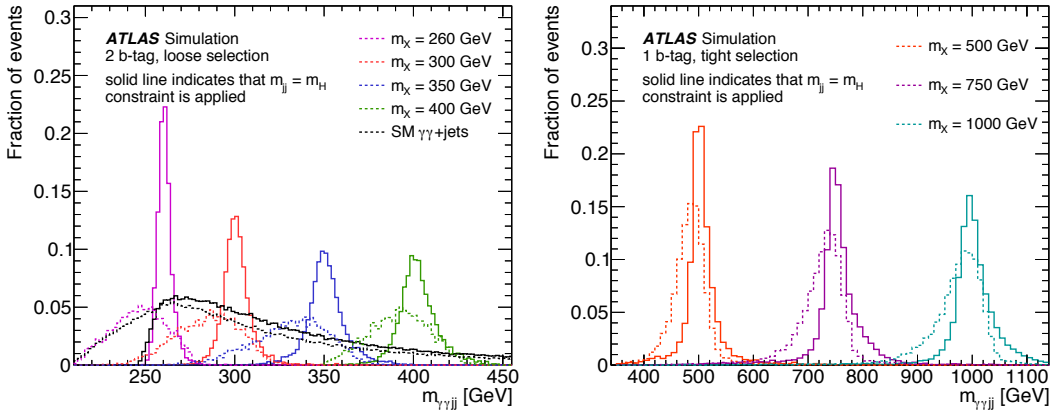
Anàlisi no ressonant

En el cas de la producció no ressonant de parells de bosons de Higgs, l'eficiència de la selecció d'esdeveniments de la categoria 2-tag és del 10% (5.8%) amb la selecció *Loose* (*Tight*). Les eficiències corresponents a la categoria 1-tag tenen un valor del 7.2% i del 3.9% amb les seleccions *Loose* i *Tight*, respectivament. Aquestes són lleugerament inferiors a les de la categoria 2-tag a causa d'una menor probabilitat de seleccionar el parell de jets correctament.

Anàlisi ressonant

En el cas de l'anàlisi ressonant, abans de reconstruir la massa dels quatre objectes, $m_{\gamma\gamma jj}$, el quadrimoment del sistema format pels dos jets s'escala mitjançant un factor m_H/m_{jj}. Tal com es mostra en la Figura 8.4, aquesta transformació millora la resolució de massa dels quatre objectes en un 60% en tot l'interval de ressonàncies de massa d'interès. També modifica la forma del fons continu de tipus no ressonant en la regió $m_X < 270$ GeV.

Per a ressonàncies de massa en el rang 260 GeV $< m_X <$ 1000 GeV, les eficiències



(a) (b)

Figure 8.4: Reconstrucció de la massa dels quatre objectes, $m_{\gamma\gamma jj}$, amb (línies contínues) i sense (línies discontínues) el factor d'escala en m_{jj}, per a un subconjunt dels punts de massa utilitzats en l'anàlisi ressonant. Els exemples que es mostren aquí corresponen a a) la categoria 2-tag amb la selecció *Loose* i b) la categoria 1-tag amb la selecció *Tight*. L'efecte causat en el fons continu també es mostra en a) [3].

de la selecció d'esdeveniments oscil·len entre el 6% i el 15.4% en la categoria 2-tag i entre el 5.1% i el 12.3% en la categoria 1-tag.

8.5.2 Modelització del senyal i del fons

La recerca de parells de bosons de Higgs mitjançant la producció ressonant i la no ressonant es porta a terme utilitzant el mètode estadístic de màxima versemblança a través d'un ajust de les dades per a les categories de senyal 1-tag i 2-tag simultàniament. La recerca de la producció no ressonant implica un ajust de la distribució $m_{\gamma\gamma}$, mentre que la recerca de la producció ressonant utilitza la distribució $m_{\gamma\gamma jj}$. L'ajust del tipus senyal-més-fons de les dades empra funcions parametritzades per a les distribucions de probabilitat de senyal i fons. Aquestes funcions parametritzades es determinen a través d'ajustos de la simulació de MC.

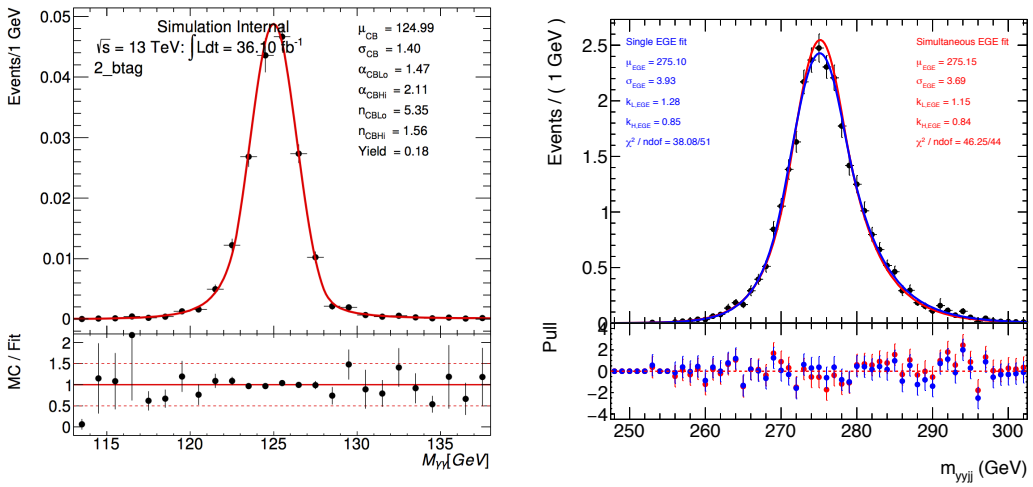
Senyal

La forma de la distribució de la massa invariant del sistema format pels dos fotons, $m_{\gamma\gamma}$, en els esdeveniments de tipus $HH \to \gamma\gamma b\bar{b}$ es descriu amb una funció Crystal Ball de doble cara [151], que consisteix en un nucli gaussià i dues cues definides per lleis potencials en els extrems. Els paràmetres d'aquest model es determinen a través d'ajustos de la simulació de MC corresponent a la producció no ressonant de parells de bosons de Higgs en el Model Estàndard. La Figura 8.5a mostra l'ajust de la distribució $m_{\gamma\gamma}$ per a la categoria 2-tag i la selecció $Tight$.

En el cas de la producció ressonant de parells de bosons de Higgs, per a cada hipòtesi ressonant es realitza un ajust de la distribució $m_{\gamma\gamma jj}$ utilitzant esdeveniments simulats en una finestra de massa al voltant del valor nominal, m_X. La forma d'aquesta distribució es descriu mitjançant una funció anomenada ExpGaussExp [153], que consisteix en un nucli gaussià i dues cues definides per funcions exponencials en els extrems. Es realitza un ajust simultani de totes les hipòtesis de senyal disponibles en la simulació de MC on cadascun dels paràmetres del model es parametritza en termes de m_X. Això permet que el model creat proporcioni una predicció per a qualsevol massa que satisfà el requeriment 260 GeV $< m_X <$ 1000 GeV, on aquests límits reflecteixen els valors mínim i màxim de m_X en la simulació de MC generada. La Figura 8.5b mostra un example de l'ajust simultani de la distribució $m_{\gamma\gamma jj}$ per a la categoria 2-tag i la selecció *Loose*.

Fons

El fons predominant en aquesta cerca rep el nom de fons continu i està format per diferents contribucions provinents de fonts del tipus $\gamma\gamma$, γj, $j\gamma$ i jj produïdes en associació amb jets, on j denota els jets mal identificats com a fotons i γj i $j\gamma$ difereixen en el jet mal identificat com a segon fotó més energètic o com a fotó més



(a) (b)

Figure 8.5: Parametrització a) amb una funció Crystal Ball de doble cara de la distribució $m_{\gamma\gamma}$ utilitzant la simulació de MC corresponent a la producció no ressonant de parells de bosons de Higgs del Model Estàndard, i b) amb una funció ExpGaussExp de la distribució $m_{\gamma\gamma jj}$ utilitzant esdeveniments simulats en una finestra de massa al voltant de $m_X = 275$ GeV per a la producció ressonant. Els esdeveniments corresponen a la categoria de senyal 2-tag amb la selecció a) *Tight* i b) *Loose*.

energètic del parell, respectivament. Aquestes contribucions es determinen a partir de les dades utilitzant un mètode de banda lateral bidimensional doble (2x2D) basat en la variació dels criteris d'identificació i d'aïllament dels fotons [149, 150].

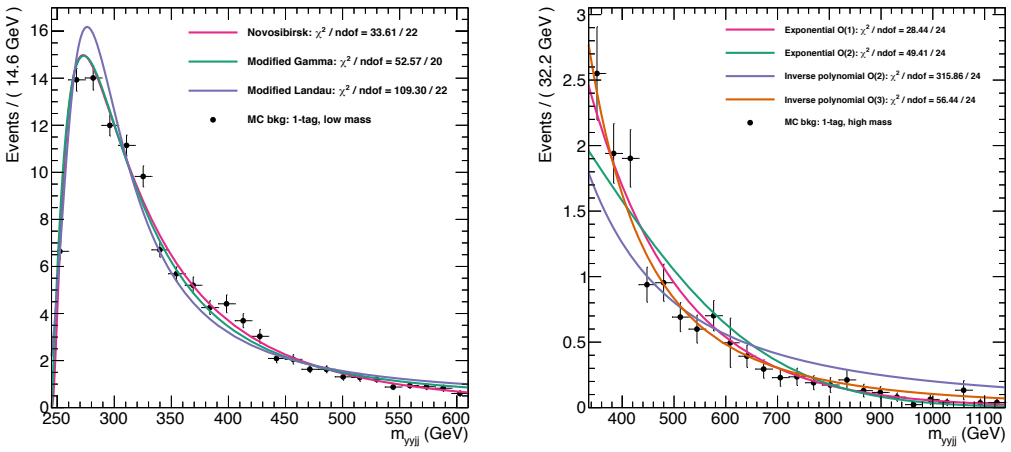
Per a l'anàlisi no ressonant, el fons continu es modelitza utilitzant una funció obtinguda a partir d'un ajust de les dades. El biaix derivat d'aquest procediment, anomenat senyal espuri (*spurious signal*), es calcula mitjançant la realització d'un ajust del tipus senyal-més-fons al fons continu. El valor absolut màxim del senyal, per a un senyal en el rang 121 GeV $< m_{\gamma\gamma} <$ 129 GeV, es pren com a biaix. Aquest mètode s'utilitza per a discriminar entre diferents funcions d'ajust on la funció triada és aquella que té el menor biaix de senyal espuri. Si diverses funcions tenen el mateix biaix, se selecciona aquella que té el menor nombre de paràmetres. La funció expo-

nencial de primer ordre té el menor biaix entre les diferents funcions considerades i, per tant, es tria. El fons de producció de bosons de Higgs individuals es descriu utilitzant una funció de Crystal Ball de doble cara, amb els seus paràmetres determinats mitjançant ajustos de les mostres simulades corresponents.

En el cas de la producció ressonant de parells de bosons de Higgs també s'utilitza el mètode de senyal espuri, realitzant un ajust del tipus senyal-més-fons al fons continu corresponent a la distribució de $m_{\gamma\gamma jj}$ per als esdeveniments que es troben en la finestra estreta de massa $m_{\gamma\gamma}$ descrita en la Secció 8.5.1. El fons de producció de bosons de Higgs individuals s'inclou en el fons continu. A causa dels diferents rangs de $m_{\gamma\gamma jj}$ emprats en les seleccions *Loose* i $Tight$, la forma de la distribució $m_{\gamma\gamma jj}$ difereix entre aquests dos casos i, per tant, es consideren diferents funcions de fons. Per a la selecció *Loose* ($Tight$) es tria la funció Novosibirsk [154] (exponencial de primer ordre) ja que mostra el menor biaix entre les diferents funcions considerades. La Figura 8.6 mostra els ajustos al fons continu de la distribució de $m_{\gamma\gamma jj}$ per als esdeveniments de la categoria de senyal 1-tag amb les seleccions *Loose* i $Tight$.

8.5.3 Incerteses sistemàtiques

Encara que la sensibilitat d'aquesta cerca està dominada per les incerteses estadístiques a causa de l'escàs nombre d'events, les incerteses sistemàtiques es calculen de manera acurada en les dues categories de senyal per a la producció ressonant i no ressonant de parelles de bosons de Higgs tenint en compte diferents contribucions d'origen teòric i experimental. A més, per a la cerca no ressonant també es consideren les incerteses teòriques i experimentals per al fons de bosons de Higgs individuals. Algunes de les més importants són les incerteses en l'escala i la resolució dels jets, en l'etiquetatge dels jets en els diferents sabors i en la modelització del procés ggH.



(a) (b)

Figure 8.6: Ajust al fons continu de la distribució de $m_{\gamma\gamma jj}$ per als esdeveniments de la categoria de senyal 1-tag amb la selecció a) *Loose* i b) *Tight*.

8.5.4 Resultats

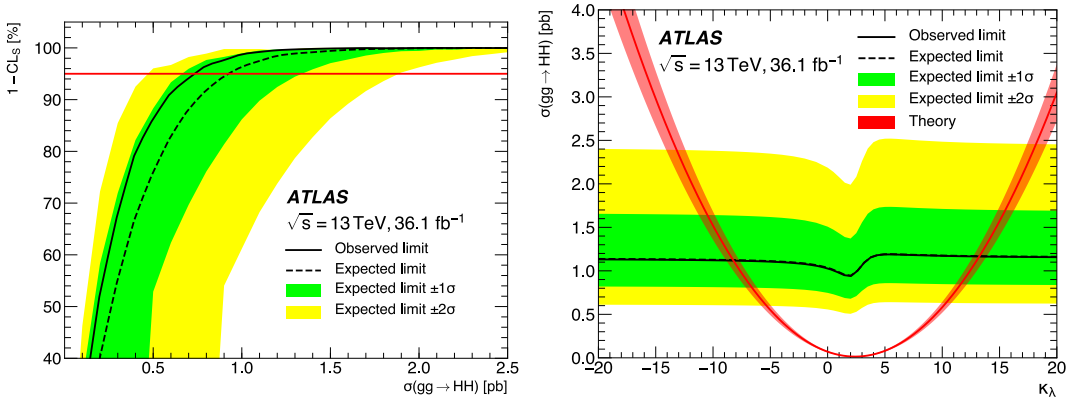
En totes les categories i seleccions el nombre d'esdeveniments observats en les dades és compatible amb el nombre d'esdeveniments de fons predits, dins de les incerteses calculades, tal com es resumeix en la Taula 8.2.

8.5.5 Producció no ressonant de parells de bosons de Higgs

El límit superior en la secció eficaç de producció no ressonant de parells de bosons de Higgs s'obté amb la selecció *Tight*. La figura 8.7a mostra aquest límit superior amb un nivell de confiança (CL) [163] del 95%, juntament amb les bandes d'incertesa a $\pm 1\sigma$ i $\pm 2\sigma$. El valor observat (esperat) és de 0.73 (0.93) pb. Com a múltiple de la secció eficaç de producció en el Model Estàndard, el límit observat (esperat) correspon a un valor de 22 (28).

Table 8.2: Nombre d'esdeveniments predits i observats en les categories 1-tag i 2-tag després d'aplicar els requeriments de la selecció ressonant, incloent el requisit en $m_{\gamma\gamma}$. El nombre d'esdeveniments de senyal de parells de bosons de Higgs en el Model Estàndard es calcula amb l'assumpció que la secció eficaç de producció correspon a un valor de 33.41 fb. Les incerteses en el fons continu provenen del procediment d'ajust utilitzat. Les incerteses en el fons de bosons de Higgs individuals i en el senyal de parells de bosons de Higgs són els errors sistemàtics d'origen teòric i experimental. Les seleccions *Loose* i *Tight* no són ortogonals.

	1-tag		2-tag	
	Selecció *Loose*	Selecció *Tight*	Selecció *Loose*	Selecció *Tight*
Fons continu	117.5 ±4.7	15.7 ± 1.6	21.0 ± 2.0	3.74 ± 0.78
Fons de bosons de Higgs individuals	5.51 ±0.10	2.20 ± 0.05	1.63 ± 0.04	0.56 ± 0.02
Fons total	123.0 ±4.7	17.9 ± 1.6	22.6 ± 2.0	4.30 ± 0.79
Senyal de parells de bosons de Higgs	0.219±0.006	0.120± 0.004	0.305± 0.007	0.175± 0.005
Dades	125	19	21	3



(a) (b)

Figure 8.7: Els límits esperats i observats amb un nivell de confiança del 95% en la secció eficaç de producció no ressonant $\sigma_{gg\to HH}$ a) per al límit en el Model Estàndard utilitzant la selecció *Tight* i b) en funció de κ_λ utilitzant la selecció *Loose*. En a) la línia roja indica el nivell de confiança del 95%. La intersecció d'aquesta línia amb el límit observat, l'esperat i les bandes a $\pm1\sigma$ i $\pm2\sigma$ correspon a la ubicació dels límits. En b) la línia roja indica la predicció de la secció eficaç de producció d'HH si es modifica κ_λ però tots els altres acoblaments romanen als seus valors en el Model Estàndard. La banda roja indica la incertesa teòrica d'aquesta predicció [3].

8.5.6 L'autoacoblament del bosó de Higgs

Variar l'autoacoblament del bosó de Higgs, λ_{HHH}, afecta tant la secció eficaç de producció no ressonant de parells de bosons de Higgs com la cinemàtica de l'esdeveniment, que altera l'eficiència de selecció del senyal. En l'anàlisi no ressonant, els resultats s'interpreten en el context de κ_λ, utilitzant la selecció *Loose*, que és més sensible al rang de valors disponibles de κ_λ amb aquest conjunt de dades. La Figura 8.7b mostra els límits en la secció eficaç $\sigma_{gg\to HH}$ juntament amb les bandes d'incertesa a $\pm1\sigma$ i $\pm2\sigma$ al voltant del límit esperat. Els límits es calculen utilitzant l'aproximació asimptòtica [148] per al test estadístic de versemblança. Si es corregeixen tots els altres paràmetres del Model Estàndard als seus valors esperats, l'autoacoblament del bosó de Higgs es limita amb un nivell de confiança del 95% a -8.2 $< \kappa_\lambda <$ 13.2 mentre que els límits esperats són -8.3 $< \kappa_\lambda <$ 13.2.

8.5.7 Producció ressonant de parells de bosons de Higgs

Els límits en la secció eficaç de producció ressonant de parells de bosons de Higgs es mostren en la Figura 8.8, utilitzant tant la selecció *Loose* com la $Tight$. La contribució provinent de la producció d'HH en el Model Estàndard es considera part del fons d'aquesta cerca encara que la seua inclusió té un impacte insignificant en els resultats. Per a ressonàncies de massa en el rang 260 GeV $< m_X <$ 1000 GeV, els límits observats (esperats) oscil·len entre 1.14 (0.90) pb i 0.12 (0.15) pb.

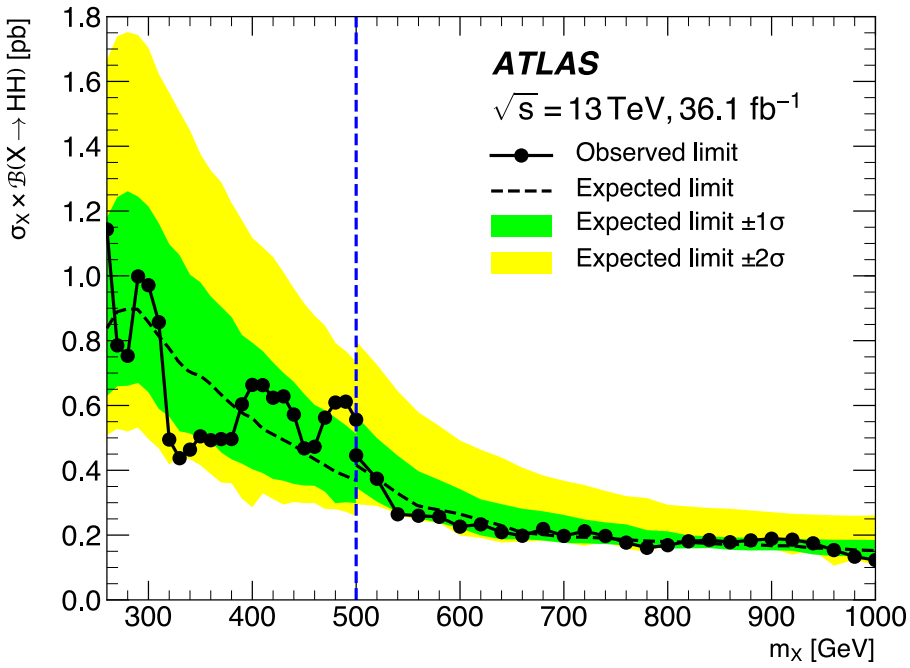
8.6 Mesura de propietats del bosó de Higgs en el canal de desintegració a dos fotons

Aquesta secció descriu la mesura de les seccions eficaces de producció del bosó de Higgs i la força dels senyals, així com la mesura de les seccions eficaces fiducials i diferencials

en el canal de desintegració a dos fotons. Es fa especial èmfasi en l'estratègia seguida per estimar la incertesa en la modelització de la pluja de partons, l'esdeveniment subjacent i l'hadronització.

8.6.1 Selecció i categorització d'esdeveniments

Els esdeveniments seleccionats contenen un parell de fotons que satisfan els criteris d'identificació de fotons de tipus *Tight* així com els requisits d'aïllament de fotons utilitzant informació sobre les traces i el calorímetre. A més, el (segon) fotó més



(a)

Figure 8.8: Els límits esperats i observats amb un nivell de confiança del 95% en la secció eficaç de producció ressonant, $\sigma_X \times \mathcal{B}\,(X \to HH)$ en funció de m_X. La selecció *Loose* s'utilitza per a ressonàncies de massa en el rang $m_X \leq 500$ GeV, mentre que la selecció *Tight* s'utilitza per a ressonàncies de massa en el rang $m_X \geq 500$ GeV. La línia discontínua blava assenyala aquest punt de transició [3].

energètic del parell ha de complir que $E_{\mathrm{T}}/m_{\gamma\gamma} > 0.35$ (0.25).

Els esdeveniments es classifiquen en 31 categories de reconstrucció excloents. Es realitza un ajust simultani d'aquestes categories per a determinar les seccions eficaces per mode de producció i interpretar la força dels senyals de les dades. En la Taula 8.3 es mostra un resum dels requisits de selecció que defineixen cada categoria.

8.6.2 Incerteses sistemàtiques

En la mesura de propietats del bosó de Higgs es consideren diverses fonts d'incertesa sistemàtica. Es poden agrupar en tres categories:

- Les incerteses associades amb la parametrització del senyal i del fons quan s'ajusta l'espectre de la massa invariant dels dos fotons, $m_{\gamma\gamma}$.

- Les incerteses experimentals derivades de l'extracció del senyal en una categoria determinada o de les migracions entre categories.

- Les incerteses teòriques i de modelització en cada categoria, que provoquen migracions entre categories o afecten l'acceptació fiducial.

En particular, el cas de la incertesa en la modelització de la pluja de partons, l'esdeveniment subjacent i l'hadronització afecta totes les mesures. Es calcula com la diferència d'acceptació d'esdeveniments després de modificar l'algoritme de pluja de partons de Pythia8 a Herwig7 en la simulació de MC dels processos ggH, VBF i VH, i de Pythia8 a Herwig++ en la simulació de MC del procés $t\bar{t}H$, respectivament. Aquestes diferències es tracten com quatre fonts independents d'incertesa. A més, en el cas del procés ggH, la diferència d'acceptació entre les diferents modificacions (*tunes*) de l'autovector del conjunt AZNLO proporcionen un component addicional d'incertesa. Els resultats s'obtenen per a cada categoria de reconstrucció

Table 8.3: Selecció d'esdeveniments que defineixen cadascuna de les 31 categories de reconstrucció d'esdeveniments per a la mesura de la força dels senyals i de les seccions eficaces de producció. Les etiquetes denoten el procés de producció predominant o les propietats cinemàtiques de la categoria. Els jets han de complir que $p_{\mathrm{T}} > 30$ GeV tret que s'indiqui el contrari. Les categories són mútuament excloents i els criteris s'apliquen en ordre descendent de les categories mostrades.

Categoria	Selecció
tH lep 0fwd	$N_{\mathrm{lep}} = 1$, $N_{\mathrm{jets}}^{\mathrm{cen}} \leq 3$, $N_{b-\mathrm{tag}} \geq 1$, $N_{\mathrm{jets}}^{\mathrm{fwd}} = 0$ ($p_{\mathrm{T}}^{\mathrm{jet}} > 25$ GeV)
tH lep 1fwd	$N_{\mathrm{lep}} = 1$, $N_{\mathrm{jets}}^{\mathrm{cen}} \leq 4$, $N_{b-\mathrm{tag}} \geq 1$, $N_{\mathrm{jets}}^{\mathrm{fwd}} \geq 1$ ($p_{\mathrm{T}}^{\mathrm{jet}} > 25$ GeV)
ttH lep	$N_{\mathrm{lep}} \geq 1$, $N_{\mathrm{jets}}^{\mathrm{cen}} \geq 2$, $N_{b-\mathrm{tag}} \geq 1$, $Z_{\ell\ell}$ veto ($p_{\mathrm{T}}^{\mathrm{jet}} > 25$ GeV)
ttH had BDT1	$N_{\mathrm{lep}} = 0$, $N_{\mathrm{jets}} \geq 3$, $N_{b-\mathrm{tag}} \geq 1$, $\mathrm{BDT}_{\mathrm{ttH}} > 0.92$
ttH had BDT2	$N_{\mathrm{lep}} = 0$, $N_{\mathrm{jets}} \geq 3$, $N_{b-\mathrm{tag}} \geq 1$, $0.83 < \mathrm{BDT}_{\mathrm{ttH}} < 0.92$
ttH had BDT3	$N_{\mathrm{lep}} = 0$, $N_{\mathrm{jets}} \geq 3$, $N_{b-\mathrm{tag}} \geq 1$, $0.79 < \mathrm{BDT}_{\mathrm{ttH}} < 0.83$
ttH had BDT4	$N_{\mathrm{lep}} = 0$, $N_{\mathrm{jets}} \geq 3$, $N_{b-\mathrm{tag}} \geq 1$, $0.52 < \mathrm{BDT}_{\mathrm{ttH}} < 0.79$
tH had 4j1b	$N_{\mathrm{lep}} = 0$, $N_{\mathrm{jets}}^{\mathrm{cen}} = 4$, $N_{b-\mathrm{tag}} = 1$ ($p_{\mathrm{T}}^{\mathrm{jet}} > 25$ GeV)
tH had 4j2b	$N_{\mathrm{lep}} = 0$, $N_{\mathrm{jets}}^{\mathrm{cen}} = 4$, $N_{b-\mathrm{tag}} \geq 2$ ($p_{\mathrm{T}}^{\mathrm{jet}} > 25$ GeV)
VH dilep	$N_{\mathrm{lep}} \geq 2$, 70 GeV $\leq m_{\ell\ell} \leq 110$ GeV
VH lep High	$N_{\mathrm{lep}} = 1$, $\lvert m_{e\gamma} - 89\ \mathrm{GeV}\rvert > 5$ GeV, $p_{\mathrm{T}}^{\ell+E_{\mathrm{T}}^{\mathrm{miss}}} > 150$ GeV
VH lep Low	$N_{\mathrm{lep}} = 1$, $\lvert m_{e\gamma} - 89\ \mathrm{GeV}\rvert > 5$ GeV, $p_{\mathrm{T}}^{\ell+E_{\mathrm{T}}^{\mathrm{miss}}} < 150$ GeV, $E_{\mathrm{T}}^{\mathrm{miss}}$ significació > 1
VH MET High	150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} < 250$ GeV, $E_{\mathrm{T}}^{\mathrm{miss}}$ significació > 9 o $E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV
VH MET Low	80 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} < 150$ GeV, $E_{\mathrm{T}}^{\mathrm{miss}}$ significació > 8
jet BSM	$p_{\mathrm{T,j1}} > 200$ GeV
VH had tight	$60\,\mathrm{GeV} < m_{\mathrm{jj}} < 120\,\mathrm{GeV}$, $\mathrm{BDT}_{\mathrm{VH}} > 0.78$
VH had loose	$60 GeV < m_{\mathrm{jj}} < 120\,\mathrm{GeV}$, $0.35 < \mathrm{BDT}_{\mathrm{VH}} < 0.78$
VBF tight, high p_{T}^{Hjj}	$\lvert\Delta\eta_{jj}\rvert > 2$, $\lvert\eta_{\gamma\gamma} - 0.5(\eta_{\mathrm{j1}} + \eta_{\mathrm{j2}})\rvert < 5$, $p_{\mathrm{T}}^{Hjj} > 25$ GeV, $\mathrm{BDT}_{\mathrm{VBF}} > 0.47$
VBF loose, high p_{T}^{Hjj}	$\lvert\Delta\eta_{jj}\rvert > 2$, $\lvert\eta_{\gamma\gamma} - 0.5(\eta_{\mathrm{j1}} + \eta_{\mathrm{j2}})\rvert < 5$, $p_{\mathrm{T}}^{Hjj} > 25$ GeV, $-0.32 < \mathrm{BDT}_{\mathrm{VBF}} < 0.47$
VBF tight, low p_{T}^{Hjj}	$\lvert\Delta\eta_{jj}\rvert > 2$, $\lvert\eta_{\gamma\gamma} - 0.5(\eta_{\mathrm{j1}} + \eta_{\mathrm{j2}})\rvert < 5$, $p_{\mathrm{T}}^{Hjj} < 25$ GeV, $\mathrm{BDT}_{\mathrm{VBF}} > 0.87$
VBF loose, low p_{T}^{Hjj}	$\lvert\Delta\eta_{jj}\rvert > 2$, $\lvert\eta_{\gamma\gamma} - 0.5(\eta_{\mathrm{j1}} + \eta_{\mathrm{j2}})\rvert < 5$, $p_{\mathrm{T}}^{Hjj} < 25$ GeV, $0.26 < \mathrm{BDT}_{\mathrm{VBF}} < 0.87$
ggH 2J BSM	≥ 2 jets, $p_{\mathrm{T}}^{\gamma\gamma} \geq 200$ GeV
ggH 2J High	≥ 2 jets, $p_{\mathrm{T}}^{\gamma\gamma} \in [120, 200]$ GeV
ggH 2J Med	≥ 2 jets, $p_{\mathrm{T}}^{\gamma\gamma} \in [60, 120]$ GeV
ggH 2J Low	≥ 2 jets, $p_{\mathrm{T}}^{\gamma\gamma} \in [0, 60]$ GeV
ggH 1J BSM	$= 1$ jet, $p_{\mathrm{T}}^{\gamma\gamma} \geq 200$ GeV
ggH 1J High	$= 1$ jet, $p_{\mathrm{T}}^{\gamma\gamma} \in [120, 200]$ GeV
ggH 1J Med	$= 1$ jet, $p_{\mathrm{T}}^{\gamma\gamma} \in [60, 120]$ GeV
ggH 1J Low	$= 1$ jet, $p_{\mathrm{T}}^{\gamma\gamma} \in [0, 60]$ GeV
ggH 0J Fwd	$= 0$ jets, un fotó amb $\lvert\eta\rvert > 0.95$
ggH 0J Cen	$= 0$ jets, dos fotons amb $\lvert\eta\rvert \leq 0.95$

d'esdeveniments, cada regió fiducial i cada interval de les seccions eficaces diferencials. La incertesa en la regió fiducial inclusiva correspon a un valor de $0.66 \pm 0.18\%$ ($0.09 \pm 0.18\%$) quan s'utilitza la simulació de MC del procés ggH (VBF).

8.6.3 La força dels senyals

La força del senyal global μ és mesurada amb l'assumpció que les ràtios entre els diferents processos de producció són les predites pel Model Estàndard, i correspon a un valor central de $\mu = 0.99\ ^{+0.15}_{-0.14} = 0.99 \pm 0.12\,(\text{estad.})\ ^{+0.06}_{-0.05}\,(\text{exp.})\ ^{+0.07}_{-0.05}\,(\text{teò.})$, compatible amb les prediccions del Model Estàndard ($\mu = 1$). Aquest resultat confirma la mesura de la força del senyal global utilitzant dades recollides per ATLAS durant el Run 1, i correspon a un valor central de $\mu = 1.17 \pm 0.23\,(\text{estad.})\ ^{+0.10}_{-0.08}\,(\text{exp.})\ ^{+0.12}_{-0.08}\,(\text{teò.})$, on s'observa una millora en cada component de la incertesa en un factor dos. La Figura 8.9 mostra un resum de la força del senyal global i la mesurada per als diferents modes de producció.

8.6.4 Les seccions eficaces de producció

Les seccions eficaces per al bosó de Higgs de massa $m_H = 125.09 \pm 0.24$ GeV en una regió definida per $|y_H| < 2.5$, multiplicades per la taxa de desintegració del bosó de Higgs a dos fotons, s'avaluen de la següent manera. El valor ajustat σ_{top} correspon a la suma de les seccions eficaces dels modes de producció $t\bar{t}H$, tHq, i tHW, amb l'assumpció que les seues ràtios relatives són les predites pel Model Estàndard. El valor de la secció eficaç del mode de producció VH està condicionat sota la hipòtesi que la proporció de les seccions eficaces dels modes de producció WH i ZH és la predita pel Model Estàndard i inclou tant la producció d'estats inicials amb quarks com amb gluons. Les contribucions del procés $b\bar{b}H$ es combinen amb ggH. Les seccions eficaces dels diferents modes de producció es mostren en la Taula 8.4.

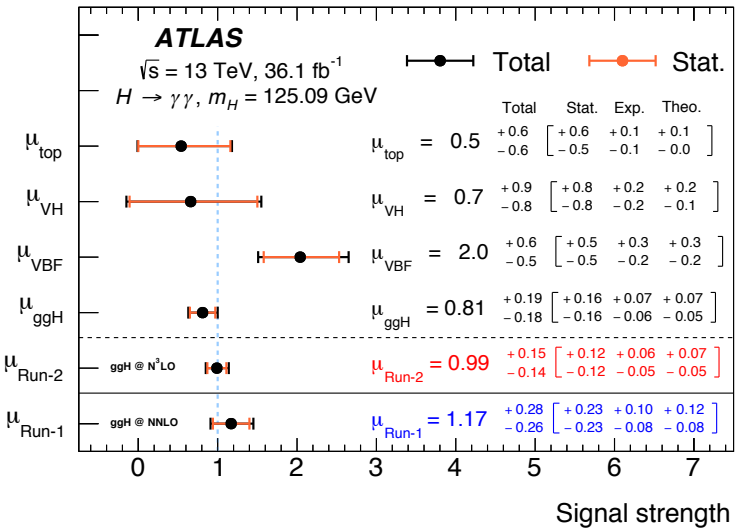


Figure 8.9: Resum de la força dels senyals mesurada per als diferents modes de producció (ggH, VBF, VH i *top*) i de manera global ($\mu_{\text{Run}\,2}$), en comparació amb la força del senyal global mesurada amb energies del centre de massa de 7 i 8 TeV ($\mu_{\text{Run}\,1}$) [146]. Les barres d'error negres i taronges mostren les incerteses totals i estadístiques. La força del senyal $\mu_{\text{Run}\,1}$ fou derivada amb l'assumpció que la secció eficaç de producció del bosó de Higgs es basa en [33, 219]. Les incerteses menors que 0.05 es mostren com a 0.0. En les prediccions teòriques més recents utilitzades en aquesta anàlisi [34, 177], la secció eficaç del mode de producció ggH és major en aproximadament un 10% [4].

Table 8.4: Valors i incerteses del millor ajust de les seccions eficaces de producció multiplicades per la taxa de desintegració del bosó de Higgs a dos fotons. També es mostren les prediccions del Model Estàndard [34] amb les seues incerteses per a cada procés de producció. Les incerteses menors que 0.05 es mostren com a 0.0.

Procés	Resultat	Incertesa [fb]				Predicció del Model Estàndard
($\lvert y_H\rvert < 2.5$)	[fb]	Total	Estad.	Exp.	Teò.	[fb]
ggH	82	$^{+19}_{-18}$	$\left(\pm 16\right.$	$^{+7}_{-6}$	$\left.^{+5}_{-4}\right)$	102^{+5}_{-7}
VBF	16	$^{+5}_{-4}$	$\left(\pm 4\right.$	± 2	$\left.^{+3}_{-2}\right)$	8.0 ± 0.2
VH	3	± 4	$\left(^{+4}_{-3}\right.$	± 1	$\left.^{+1}_{-0}\right)$	4.5 ± 0.2
Top	0.7	$^{+0.9}_{-0.7}$	$\left(^{+0.8}_{-0.7}\right.$	$^{+0.2}_{-0.1}$	$\left.^{+0.2}_{-0.0}\right)$	1.3 ± 0.1

8.6.5 Les seccions eficaces fiducials i diferencials

Els volums fiducials es defineixen per tal d'imitar la selecció a nivell de detector dels fotons i altres objectes de l'anàlisi. D'aquesta manera s'aconsegueix reduir la dependència del model en contrast amb les mesures de la secció eficaç dels diferents modes de producció, descrites en la Secció 8.6.4. La secció eficaç (σ_i) en una regió fiducial, i la secció eficaç diferencial ($\mathrm{d}\sigma_i/\mathrm{d}x$) en un interval d'una certa variable x es determinen corregint el valor del senyal mesurat d'ineficiències experimentals i d'efectes en la resolució, i tenint en compte la lluminositat integrada de les dades, i venen donades per:

$$\sigma_i = \frac{N_i^{\mathrm{sig}}}{c_i \int L\,\mathrm{d}t} \qquad \text{i} \qquad \frac{\mathrm{d}\sigma_i}{\mathrm{d}x} = \frac{N_i^{\mathrm{sig}}}{c_i\,\Delta x_i \int L\,\mathrm{d}t}, \tag{8.3}$$

on N_i^{sig} és el nombre d'esdeveniments de senyal, $\int L\,\mathrm{d}t$ és la lluminositat integrada de les dades, c_i és un factor de correcció que representa la ineficiència i la resolució del detector, i Δx_i és el valor de l'interval o amplitud. Els factors de correcció es determinen utilitzant la simulació de MC.

En aquesta secció les seccions eficaces fiducials es determinen en cinc regions diferents de l'espai de fases sensibles a la producció inclusiva del bosó de Higgs i als mecanismes explícits de producció del bosó de Higgs. La mesura d'aquestes cinc seccions eficaces proporciona una forma alternativa d'estudiar les propietats del bosó de Higgs i de cercar nova física més enllà del Model Estàndard. Per a cada mesura de la secció eficaç en una regió fiducial o interval d'una distribució diferencial, s'extreu el senyal de $H \to \gamma\gamma$ utilitzant un ajust del corresponent espectre de la massa invariant dels dos fotons. La definició a nivell de partícula de les cinc regions fiducials es detalla en la Taula 8.5 mentre que la Taula 8.6 mostra un resum de les seccions eficaces fiducials mesurades i una comparació amb les prediccions del Model Estàndard.

Table 8.5: Resum de les definicions a nivell de partícula de les cinc regions fiducials descrites en el text. L'aïllament dels fotons $p_{\mathrm{T}}^{\text{aïll},0.2}$ es defineix com el moment transversal del sistema de partícules carregades dins d'un con de $\Delta R < 0.2$ al voltant del fotó.

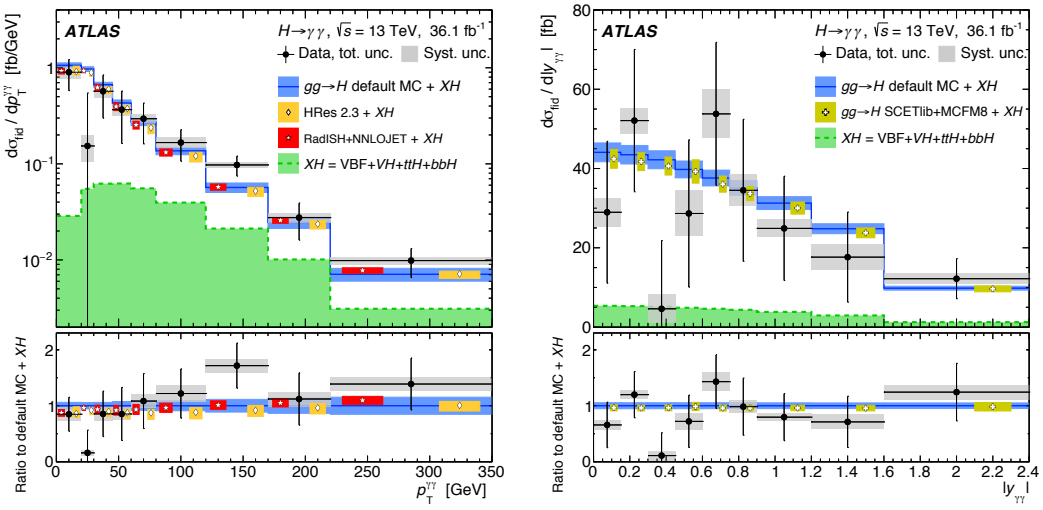
Objecte	Definició
Fotons	$\lvert\eta\rvert < 1.37$ or $1.52 < \lvert\eta\rvert < 2.37$, $p_{\mathrm{T}}^{\text{aïll},0.2}/p_{\mathrm{T}}^{\gamma} < 0.05$
jets	anti-k_{T}, $R = 0.4$, $p_{\mathrm{T}} > 30$ GeV, $\lvert y\rvert < 4.4$
Leptons, ℓ	e o μ, $p_{\mathrm{T}} > 15$ GeV, $\lvert\eta\rvert < 2.47$ per a e (s'exclou $1.37 < \lvert\eta\rvert < 1.52$) i $\lvert\eta\rvert < 2.7$ per a μ
Regió fiducial	Definició
Inclusiva (dos fotons)	$N_\gamma \geq 2$, $p_{\mathrm{T}}^{\gamma_1} > 0.35\, m_{\gamma\gamma} = 43.8$ GeV, $p_{\mathrm{T}}^{\gamma_2} > 0.25\, m_{\gamma\gamma} = 31.3$ GeV
VBF-intensificada	Inclusiva, $N_j \geq 2$, $m_{jj} > 400$ GeV, $\Delta y_{jj} > 2.8$, $\lvert\Delta\phi_{\gamma\gamma,jj}\rvert > 2.6$
$N_{\text{leptons}} \geq 1$	Inclusiva, $N_\ell \geq 1$
Alta $E_{\mathrm{T}}^{\text{miss}}$	Inclusiva, $E_{\mathrm{T}}^{\text{miss}} > 80$ GeV, $p_{\mathrm{T}}^{\gamma\gamma} > 80$ GeV
$t\bar{t}H$-intensificada	Inclusiva, ($N_j \geq 4$, $N_{b\text{-jets}} \geq 1$) o ($N_j \geq 3$, $N_{b\text{-jets}} \geq 1, N_\ell \geq 1$)

Table 8.6: Les seccions eficaces mesurades en les regions fiducials inclusiva (dos fotons), VBF-intensificada, $N_{\text{leptons}} \geq 1$, alta $E_{\mathrm{T}}^{\text{miss}}$, i $t\bar{t}H$-intensificada. Les prediccions del Model Estàndard derivades dels modes de producció VBF, VH, $t\bar{t}H$ i $b\bar{b}H$ estan etiquetades col·lectivament com a XH.

Regió fiducial	Secció eficaç mesurada	Predicció del Model Estàndard	
Inclusiva	55 ± 9 (estad.) $\pm$ 4 (exp.) $\pm$ 0.1 (teò.) fb	64 ± 2 fb	[N^{3}LO + XH]
VBF-intensificada	3.7 ± 0.8 (estad.) $\pm$ 0.5 (exp.) $\pm$ 0.2 (teò.) fb	2.3 ± 0.1 fb	[per defecte MC + XH]
$N_{\text{leptons}} \geq 1$	≤ 1.39 fb 95% CL	0.57 ± 0.03 fb	[per defecte MC + XH]
Alta $E_{\mathrm{T}}^{\text{miss}}$	≤ 1.00 fb 95% CL	0.30 ± 0.02 fb	[per defecte MC + XH]
$t\bar{t}H$-intensificada	≤ 1.27 fb 95% CL	0.55 ± 0.06 fb	[per defecte MC + XH]

Diferents seccions eficaces diferencials es mesuren i es comparen amb diverses prediccions en funció d'onze observables escollits acuradament utilitzant esdeveniments pertanyents a la regió fiducial inclusiva. La producció de bosons de Higgs està dominada pel procés ggH, per la qual cosa el moment transversal del bosó de Higgs està en gran mesura correlacionat amb l'emissió de gluons de baix moment transversal i quarks. La mesura de $p_{\mathrm{T}}^{\gamma\gamma}$ prova la modelització de la Cromodinàmica Quàntica (QCD, de l'anglès *Quantum Chromodynamics*) pertorbativa d'aquest mecanisme de producció que és lleugerament sensible als acoblaments de Yukawa entre quarks *bottom* i *charm*, i el bosó de Higgs. La distribució d'alt moment transversal és sensible a noves partícules pesades que s'acoblen al bosó de Higgs i a l'acoblament de Yukawa del quark *top*. La distribució de la rapidesa del bosó de Higgs també és sensible a la modelització del mecanisme de producció ggH, així com a les funcions de distribució de partons (PDF) dels protons de la col·lisió. En la Figura 8.10 la predicció del Model Estàndard mostra un petit excés per a baix moment transversal i baixa rapidesa, i un petit dèficit per a alt moment transversal.

El moment transversal i la rapidesa absoluta dels dos jets de més alt moment transversal proven la modelització de QCD pertorbativa i són sensibles a les contribucions relatives dels diferents mecanismes de producció del bosó de Higgs. Les variables angulars $|\cos\theta^*|$ i $\Delta\phi_{jj}$ són sensibles als nombres quàntics d'espín i CP (combinació de les transformacions C-conjugació de càrrega- i P-paritat-) del bosó de Higgs. La separació de la rapidesa del sistema dels dos jets, $|\Delta y_{jj}|$, la massa del sistema dels dos jets, m_{jj}, i la diferència azimutal entre el sistema dels dos jets i el sistema dels dos fotons, $|\Delta\phi_{\gamma\gamma,jj}|$, són sensibles al mecanisme de producció VBF. Totes les seccions eficaces diferencials mesurades en funció d'aquests observables es comparen amb les prediccions i no s'observa cap desviació significativa del Model Estàndard.



(a) (b)

Figure 8.10: Les seccions eficaces diferencials per al procés p-$p \to H \to \gamma\gamma$ en funció de a) $p_{\mathrm{T}}^{\gamma\gamma}$ i b) $|y^{\gamma\gamma}|$, i comparacions amb diferents prediccions del Model Estàndard. Les prediccions del Model Estàndard derivades dels modes de producció VBF, VH, $t\bar{t}H$ i $b\bar{b}H$ estan etiquetades col·lectivament com a XH [4].

8.7 Conclusions

La recerca del bosó de Higgs fou l'objectiu principal dels programes de física dels experiments de l'LHC durant el Run 1. El descobriment d'aquesta partícula, anunciat l'any 2012 per les col·laboracions ATLAS i CMS, constituí una fita molt important per a la física de partícules a l'hora d'entendre el mecanisme de trencament espontani de simetria del Model Estàndard pel qual les partícules fonamentals adquireixen massa. Ara és essencial que el bosó de Higgs siga estudiat extensivament. Mesures precises de les seues propietats confirmaran la seua naturalesa, i qualsevol desviació de la predicció del Model Estàndard representarà un signe inequívoc de nova física.

Aquest fi no es pot aconseguir sense un bon enteniment de l'aparell experimental. Els estudis de rendiment descrits en aquesta tesi se centren en validar els mètodes de reconstrucció i calibratge del calorímetre TileCal del detector ATLAS mitjançant l'ús

de la resposta del calorímetre als hadrons aïllats amb dades recollides des del 2010 fins al 2012. Els resultats mostren que el quocient doble del valor mitjà ($\langle E/p \rangle$) entre les dades i la simulació de MC és compatible amb la unitat. Concretament en la regió de barril del TileCal ($|\eta| < 0.7$) s'observa un 3% de discrepància màxima a pesar de canvis importants en les condicions del feix al llarg dels tres anys, la qual cosa mostra que l'efecte del pile-up està ben simulat i que les tècniques emprades per a reconstruir l'energia són robustes.

La producció de parells de bosons de Higgs és el procés de producció més senzill que és sensible a l'autoacoblament λ i proporciona una gran quantitat de possibilitats per investigar interaccions multidimensionals, així com l'existència d'estats més pesats acoblats al Higgs. Aquesta tesi presenta una recerca de la producció de parells de bosons de Higgs en l'estat final $\gamma\gamma b\bar{b}$ amb dades recollides a una energia del centre de massa de 13 TeV amb el detector ATLAS. No s'observen desviacions significatives de les prediccions del Model Estàndard. El límit superior observat (esperat) amb un nivell de confiança del 95% en la secció eficaç de producció no ressonant és 0.73 pb (0.93 pb) i correspon a 22 (28) vegades la predicció del Model Estàndard, el qual millora el resultat precedent publicat per l'experiment ATLAS [2] en un factor de cinc. En el cas de la producció ressonant de $X \to HH \to \gamma\gamma b\bar{b}$, els límits observats (esperats) en la secció eficaç de producció ressonant oscil·len entre 1.14 (0.90) pb i 0.12 (0.15) pb per a ressonàncies de massa en el rang 260 GeV $< m_X <$ 1000 GeV.

Aquesta tesi també presenta la mesura de les seccions eficaces dels modes de producció del bosó de Higgs, la força dels senyals i les seccions eficaces fiducials i diferencials en el canal de desintegració a dos fotons amb dades recollides a $\sqrt{s} = 13$ TeV amb el detector ATLAS. Totes les mesures assumeixen que la massa del bosó de Higgs té un valor de 125.09 ± 0.24 GeV. La força del senyal mesurada respecte al valor predit pel Model Estàndard és $\mu = 0.99\ ^{+0.15}_{-0.14} = 0.99 \pm 0.12\,(\text{estad.})\ ^{+0.06}_{-0.05}\,(\text{exp.})\ ^{+0.07}_{-0.05}\,(\text{teò.})$.

Aquest resultat confirma la mesura realitzada per l'experiment ATLAS amb dades recollides durant el Run 1 [146], $\mu = 1.17 \pm 0.23\,(\text{estad.})\ ^{+0.10}_{-0.08}\,(\text{exp.})\ ^{+0.12}_{-0.08}\,(\text{teò.})$, i la millora en un factor de dos en cada component de la incertesa. La secció eficaç fiducial té un valor de $\sigma_{\text{fid}} = 55 \pm 9\,(\text{estad.}) \pm 4\,(\text{exp.}) \pm 0.1\,(\text{teò.})$ fb, el qual és compatible amb la predicció del Model Estàndard (64 ± 2 fb). Actualment la precisió d'aquestes mesures està dominada per les seues incerteses estadístiques, però s'espera que millori en els propers anys de l'LHC, ja que a mesura que es recopilin més dades la incertesa estadística disminuirà. Aleshores, ser capaços de realitzar estimacions precises de les incerteses sistemàtiques serà essencial, en particular de la incertesa en la modelització de la pluja de partons, l'esdeveniment subjacent i l'hadronització, que suposa un repte a causa de la dificultat de generar esdeveniments suficients per a cada categoria de reconstrucció, regió fiducial o interval d'una secció eficaç diferencial. Aquesta tesi estableix els fonaments per a l'estimació d'aquesta incertesa que s'espera millorar en un futur pròxim.

Appendix A

Unfolding

In the measurement of differential cross sections, there can be non-negligible migrations due to the finite resolution of the ATLAS detector. The reversion of such resolution migrations is called unfolding. The unfolding procedure corrects the reconstructed data to particle level data. This enables the easy comparison of many theory predictions with the measurements.

The problem of unfolding can be formulated by the equation:

$$x_i = \sum_{j=1}^{N} R_{ij} y_j, \tag{A.1}$$

where x_i is the measured reconstructed observable, y_j is the truth observable and the response matrix R_{ij} can be interpreted as the conditional probability:

$$R_{ij} = P(\text{reconstructed in bin i}|\text{true value in bin j}), \tag{A.2}$$

with the sum being equal to:

$$\sum_{i=1}^{N} R_{ij} = P(\text{observed anywhere}|\text{true value in bin j}) = \epsilon_j, \tag{A.3}$$

resulting in the efficiency, ϵ_j. The task of unfolding is to invert equation A.1 to convert the measured values into the true values.

Simple matrix inversion and bin-by-bin unfolding are discussed below. The description of regularized matrix methods can be found in [206–208].

In the $H \to \gamma\gamma$ differential cross section analysis, all of the above methods were studied. It was found that the matrix based methods resulted in a smaller bias than the bin-by-bin unfolding method, however the observed biases of all methods were much smaller than the statistical uncertainties. Since there was no preference for a particular method, the simple bin-by-bin unfolding method was used for this analysis.

A.1 Matrix inversion

One approach to invert equation A.1 would be to determine the inverse of the response matrix, R_{ij}^{-1} in order to recover the particle level distribution. The main disadvantage of this method is that if the off-diagonal elements of the response matrix are too large (e.g. if the chosen bin size is too small compared to the measurement resolution) the resulting expression for the true value

$$y = R_{ij}^{-1} x \tag{A.4}$$

can have extremely large variances. However, the advantage of this inversion approach is that the resulting values for y which are affected by large variances are in fact completely unbiased. Any other unfolding method which aims to reduce the large variances will introduce a bias. The strategy in other unfolding methods is therefore

to accept a small bias (systematic uncertainty) in exchange for a reduction in the variance (statistical uncertainty).

A.2 Bin-by-Bin Unfolding

This method is based on multiplicative correction factors derived from Monte Carlo simulations. The value for an observable y in a given bin i is constructed as:

$$y_i = C_i x_i, \tag{A.5}$$

where the correction factor C_i is given by:

$$C_i = \frac{y_i^{MC}}{x_i^{MC}}, \tag{A.6}$$

where y_i^{MC} and x_i^{MC} are the expected true and reconstructed yields from the simulation. This inversion has a much smaller variance than the inversion of the response matrix, but yields model dependent results, i.e. biased by the choice of Monte Carlo simulation used as the input. The size of this bias is estimated with respect to the Standard Model and the maximal deviation one could possibly expect. This method is satisfactory if the bias is small with respect to the variance obtained. The size of the bias is added to the systematic uncertainty of the measurement.

Appendix B

Underlying event, parton shower and hadronization uncertainties for fiducial integrated and differential results

The uncertainty in the modeling of the parton shower, underlying event and hadronization affects all measurements (referred to as "Underlying Event / Parton Shower" in the following). It is estimated by taking the difference in acceptance at truth level after switching the parton showering algorithm from Pythia8 to Herwig7 in the ggH, VBF, and VH samples. These differences are treated as three independent uncertainty sources. Additionally, for ggH the effect of the eigenvector tunes from the AZNLO set are merged to provide one additional uncertainty component.

The Underlying Event / Parton Shower uncertainty is obtained as the absolute value of the difference in acceptance between the different parton showering algo-

rithms. This difference is estimated in five fiducial regions and in each bin of eleven fiducial differential cross sections, as shown in Tables B.1 to B.4, and in Tables B.5 to B.48, respectively.

	ggH		**VBF**	
Regions	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
Diphoton fiducial	-0.66	0.18	-0.09	0.18
VBF-enhanced	7.2	1.4	-7.3	0.3
$N_{\text{lepton}} \geq 1$	26	34	-52	10
High $E_{\text{T}}^{\text{miss}}$	92	190	170	160
$t\bar{t}H$-enhanced	-14	5	-54	3

Table B.1: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the NNLOPS ggH and VBF $H \to \gamma\gamma$ samples for the five fiducial regions, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of ggH and VBF events passing each region.

	ZH		***ggZH***		***W⁻H***		***W⁺H***	
Regions	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
Diphoton fiducial	-0.1	0.5	-0.7	0.8	-0.8	0.6	-0.7	0.7
VBF-enhanced	-0.2	5	39	5	0	6	2	6
$N_{\text{lepton}} \geq 1$	-3.1	1.9	-2	3	-2.2	1.4	-1.5	1.4
High $E_{\text{T}}^{\text{miss}}$	-2.4	1.7	-2	2	0	2	0	2
$t\bar{t}H$-enhanced	4	5	0	4	-19	10	1	12

Table B.2: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the VH $H \to \gamma\gamma$ samples for the five fiducial regions, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of VH events passing each region.

	ISR_1 up		ISR_1 down		ISR_2 up		ISR_2 down	
Regions	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
Diphoton fiducial	0.04	0.12	0.65	0.12	-0.04	0.12	0.11	0.12
VBF-enhanced	0.7	0.9	2.1	1.0	1.1	1.0	0.5	0.9
$N_{\text{lepton}} \geq 1$	4	20	10	20	-15	17	2	20
High $E_{\text{T}}^{\text{miss}}$	-40	55	-19	70	0	84	-80	25
$t\bar{t}H$-enhanced	-1	4	1	4	-1	4	-4	4

Table B.3: Differences in acceptance between AZNLO (ISR eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for the five fiducial regions, as described in [215]. This difference corresponds to (AZNLO (ISR eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each region.

	FSR up		FSR down		MPI up		MPI down	
Regions	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
Diphoton fiducial	0.41	0.12	0.37	0.12	0.06	0.12	-0.03	0.12
VBF-enhanced	5.5	1.0	-3.2	0.9	0.6	0.9	1.0	1.0
$N_{\text{lepton}} \geq 1$	-10	18	-5	19	-3	20	-2	20
High $E_{\text{T}}^{\text{miss}}$	1	80	-60	40	-20	70	0	80
$t\bar{t}H$-enhanced	-2	4	2	4	-2	4	-1	4

Table B.4: Differences in acceptance between AZNLO (FSR / MPI eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for the five fiducial regions, as described in [215]. This difference corresponds to (AZNLO (FSR / MPI eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each region.

	ggH		**VBF**	
Bins of $p_{\text{T}}^{\gamma\gamma}$ [GeV]	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0–20	1.7	0.3	-0.1	0.8
20–30	-1.0	0.4	1.1	0.7
30–45	-1.2	0.4	-0.2	0.5
45–60	-1.8	0.5	0.9	0.5
60–80	-2.3	0.6	0.2	0.4
80–120	-2.8	0.6	-0.4	0.4
120–170	-2.1	0.8	-0.6	0.5
170–220	-3.0	1.3	-1.4	0.7
220–350	-4.9	1.5	-0.4	0.8
≥ 350	-10	3	-0.2	1.8

Table B.5: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the NNLOPS ggH and VBF $H \to \gamma\gamma$ samples for each bin of the $p_{\text{T}}^{\gamma\gamma}$ observable, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of ggH and VBF events in each bin.

Appendix B. Underlying event, parton shower and hadronization uncertainties for fiducial integrated and differential results

	ZH		$ggZH$		W^-H		W^+H	
Bins of $p_{\mathrm{T}}^{\gamma\gamma}$ [GeV]	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0–20	2	2	-5	7	5	3	-6	3
20–30	5	2	4	7	0	2	1	3
30–45	-0.3	1.5	3	4	-4.3	1.6	-1.5	1.9
45–60	-0.8	1.5	-1	4	-1.4	1.7	-1.4	1.8
60–80	0.4	1.3	-1	3	2.8	1.6	-0.3	1.7
80–120	-1.1	1.2	0	1.7	-0.5	1.4	1.8	1.5
120–170	-0.5	1.5	0	1.5	-3.2	1.7	-1.7	1.8
170–220	-3	2	-1	2	0	3	-1	3
220–350	0	2	-2	2	-1	3	-4	3
≥ 350	0	4	-16	6	-2	5	-2	5

Table B.6: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the VH $H \to \gamma\gamma$ samples for each bin of the $p_{\mathrm{T}}^{\gamma\gamma}$ observable, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of VH events passing each bin.

	ISR_1 up		**ISR_1 down**		**ISR_2 up**		**ISR_2 down**	
Bins of $p_{\mathrm{T}}^{\gamma\gamma}$ [GeV]	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0–20	-0.2	0.2	0.8	0.2	-0.3	0.2	-0.2	0.2
20–30	0.0	0.3	1.1	0.3	0.0	0.3	0.1	0.3
30–45	0.2	0.3	1.4	0.3	0.3	0.3	0.5	0.3
45–60	0.2	0.4	1.3	0.4	0.3	0.4	0.1	0.4
60–80	0.0	0.4	0.9	0.4	-0.4	0.4	-0.1	0.4
80–120	0.1	0.4	1.0	0.4	0.0	0.4	0.3	0.4
120–170	0.9	0.6	1.7	0.6	0.6	0.6	1.1	0.6
170–220	0.9	0.9	1.4	0.9	-0.1	0.9	0.3	0.9
220–350	0.0	1.0	1.3	1.1	0.8	1.1	0.3	1.0
≥ 350	-2	2	-1	2	-4	2	-1	2

Table B.7: Differences in acceptance between AZNLO (ISR eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for each bin of the $p_{\mathrm{T}}^{\gamma\gamma}$ observable, as described in [215]. This difference corresponds to (AZNLO (ISR eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each bin.

	FSR up		**FSR down**		**MPI up**		**MPI down**	
Bins of $p_{\mathrm{T}}^{\gamma\gamma}$ [GeV]	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0–20	0.5	0.2	0.1	0.2	-0.2	0.2	-0.3	0.2
20–30	1.1	0.3	0.4	0.3	0.0	0.3	-0.2	0.3
30–45	1.1	0.3	0.7	0.3	0.3	0.3	0.5	0.3
45–60	1.0	0.4	0.7	0.4	0.4	0.4	0.1	0.4
60–80	0.4	0.4	0.0	0.4	-0.3	0.4	-0.1	0.4
80–120	1.0	0.4	0.8	0.4	0.2	0.4	0.2	0.4
120–170	1.3	0.6	0.8	0.6	0.5	0.6	0.3	0.6
170–220	1.0	0.9	1.1	0.9	0.8	0.9	0.6	0.9
220–350	1.2	1.1	0.4	1.0	0.0	1.0	-0.4	1.0
≥ 350	-2	2	-2	2	-3	2	-3	2

Table B.8: Differences in acceptance between AZNLO (FSR / MPI eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for each bin of the $p_{\mathrm{T}}^{\gamma\gamma}$ observable, as described in [215]. This difference corresponds to (AZNLO (FSR / MPI eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each bin.

	ggH		VBF	
Bins of $\lvert y^{\gamma\gamma}\rvert$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0–0.15	-0.3	0.5	0.1	0.5
0.15–0.30	-0.8	0.5	-0.3	0.6
0.30–0.45	-1.4	0.6	-0.3	0.6
0.45–0.60	-0.5	0.6	0.3	0.6
0.60–0.75	-0.5	0.6	-0.2	0.6
0.75–0.90	-0.6	0.6	-0.1	0.6
0.90–1.20	-0.8	0.5	-0.5	0.5
1.20–1.60	-0.5	0.4	-0.1	0.5
1.60–2.40	-0.3	0.5	0.4	0.5

Table B.9: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the NNLOPS ggH and VBF $H \to \gamma\gamma$ samples for each bin of the $|y^{\gamma\gamma}|$ observable, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of ggH and VBF events in each bin.

	ZH		$ggZH$		W^-H		W^+H	
Bins of $\lvert y^{\gamma\gamma}\rvert$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0–0.15	-1.7	1.7	3	2	-0.4	1.9	-4	2
0.15–0.30	-2.6	1.7	-4	2	-3.9	1.9	2	2
0.30–0.45	0.4	1.7	-3	2	0	2	0	2
0.45–0.60	0.1	1.7	4	3	-3	2	-2	2
0.60–0.75	0.9	1.8	-2	3	2	2	-2	2
0.75–0.90	0.0	1.8	-4	3	2	2	-4	2
0.90–1.20	0.1	1.4	1	2	-0.5	1.6	1.1	1.7
1.20–1.60	1.3	1.3	0	2	0.9	1.5	-0.9	1.6
1.60–2.40	-0.6	1.3	-3	3	-2.2	1.6	-0.5	1.6

Table B.10: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the VH $H \to \gamma\gamma$ samples for each bin of the $|y^{\gamma\gamma}|$ observable, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of VH events passing each bin.

	$\mathbf{ISR_1}$ up		$\mathbf{ISR_1}$ down		$\mathbf{ISR_2}$ up		$\mathbf{ISR_2}$ down	
Bins of $\lvert y^{\gamma\gamma}\rvert$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0–0.15	0.3	0.4	1.4	0.4	0.3	0.4	0.3	0.4
0.15–0.30	-0.1	0.3	1.0	0.4	-0.1	0.4	-0.1	0.4
0.30–0.45	-0.7	0.4	0.2	0.4	-0.7	0.4	-0.8	0.4
0.45–0.60	0.1	0.4	1.2	0.4	0.0	0.4	0.3	0.4
0.60–0.75	0.0	0.4	0.9	0.4	-0.1	0.4	-0.1	0.4
0.75–0.90	0.0	0.4	1.3	0.4	0.1	0.4	0.2	0.4
0.90–1.20	0.4	0.3	1.4	0.3	0.3	0.3	0.3	0.3
1.20–1.60	0.0	0.3	1.1	0.3	0.0	0.3	0.2	0.3
1.60–2.40	0.1	0.3	1.1	0.4	-0.1	0.3	0.4	0.4

Table B.11: Differences in acceptance between AZNLO (ISR eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for each bin of the $|y^{\gamma\gamma}|$ observable, as described in [215]. This difference corresponds to (AZNLO (ISR eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each bin.

	FSR up		FSR down		MPI up		MPI down	
Bins of $\lvert y^{\gamma\gamma}\rvert$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0–0.15	1.2	0.4	0.8	0.4	0.3	0.4	0.2	0.4
0.15–0.30	0.5	0.4	0.3	0.4	-0.2	0.4	-0.3	0.4
0.30–0.45	0.3	0.4	-0.5	0.4	-0.8	0.4	-0.8	0.4
0.45–0.60	0.9	0.4	0.5	0.4	0.4	0.4	0.0	0.4
0.60–0.75	0.7	0.4	0.4	0.4	-0.3	0.4	-0.5	0.4
0.75–0.90	0.9	0.4	0.5	0.4	0.1	0.4	0.2	0.4
0.90–1.20	1.1	0.3	0.7	0.3	0.4	0.3	0.4	0.3
1.20–1.60	0.9	0.3	0.3	0.3	0.1	0.3	-0.1	0.3
1.60–2.40	0.9	0.4	0.5	0.4	0.2	0.3	0.2	0.3

Table B.12: Differences in acceptance between AZNLO (FSR / MPI eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for each bin of the $|y^{\gamma\gamma}|$ observable, as described in [215]. This difference corresponds to (AZNLO (FSR / MPI eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each bin.

	ggH		VBF	
Bins of $p_{\mathrm{T}}^{j_1}$ [GeV]	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0–30	-1.5	0.2	20.5	0.6
30–55	1.2	0.4	4.4	0.4
55–75	0.0	0.6	-2.4	0.4
75–120	-0.3	0.6	-5.1	0.3
120–350	0.1	0.8	-5.7	0.4
≥ 350	-2	4	-4	2

Table B.13: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the NNLOPS ggH and VBF $H \to \gamma\gamma$ samples for each bin of the $p_{\mathrm{T}}^{j_1}$ observable, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of ggH and VBF events in each bin.

	ZH		$ggZH$		W^-H		W^+H	
Bins of $p_{\mathrm{T}}^{j_1}$ [GeV]	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0–30	-0.5	1.1	-17	2	0.6	1.3	-1.4	1.3
30–55	2.1	1.1	-1	2	1.7	1.2	2.4	1.3
55–75	-0.7	1.2	1	2	-0.9	1.5	-1.1	1.6
75–120	-1.1	1.2	2.8	1.5	-3.4	1.5	-1.4	1.5
120–350	-1.8	1.4	1.7	1.5	-3.2	1.7	-4.9	1.7
≥ 350	3	5	-3	9	-5	6	-1	6

Table B.14: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the VH $H \to \gamma\gamma$ samples for each bin of the $p_{\mathrm{T}}^{j_1}$ observable, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of VH events passing each bin.

	ISR_1 up		ISR_1 down		ISR_2 up		ISR_2 down	
Bins of $p_{\mathrm{T}}^{j_1}$ [GeV]	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0–30	-0.09	0.15	0.92	0.16	-0.19	0.15	-0.05	0.15
30–55	0.1	0.3	1.0	0.3	0.2	0.3	0.2	0.3
55–75	-0.1	0.4	0.9	0.4	0.2	0.4	0.4	0.4
75–120	0.4	0.4	1.6	0.4	0.0	0.4	0.4	0.4
120–350	0.8	0.5	1.9	0.5	0.6	0.5	0.6	0.5
≥ 350	1	3	3	3	0	3	1	3

Table B.15: Differences in acceptance between AZNLO (ISR eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for each bin of the $p_{\mathrm{T}}^{j_1}$ observable, as described in [215]. This difference corresponds to (AZNLO (ISR eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each bin.

Appendix B. Underlying event, parton shower and hadronization uncertainties for fiducial integrated and differential results

	FSR up		FSR down		MPI up		MPI down	
Bins of $p_{\mathrm{T}}^{j_1}$ [GeV]	Variation [%]	Error [%]	Variation [%]	Error [%]	Variation [%]	Error [%]	Variation [%]	Error [%]
0–30	-0.31	0.15	1.33	0.16	0.12	0.16	-0.28	0.16
30–55	2.4	0.3	-1.1	0.3	-0.3	0.3	0.3	0.3
55–75	2.2	0.4	-1.1	0.4	-0.1	0.4	0.2	0.4
75–120	2.5	0.4	-0.8	0.4	0.4	0.4	0.4	0.4
120–350	2.8	0.5	-0.7	0.5	0.4	0.5	0.4	0.5
≥ 350	1	3	0	3	1	3	0	3

Table B.16: Differences in acceptance between AZNLO (FSR / MPI eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for each bin of the $p_{\mathrm{T}}^{j_1}$ observable, as described in [215]. This difference corresponds to (AZNLO (FSR / MPI eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each bin.

	ggH		VBF	
Bins of $p_{\mathrm{T}}^{j_2}$ [GeV]	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0–30	-2.1	0.3	5.0	0.3
30–70	5.8	0.6	-5.9	0.3
70–120	6.7	1.3	-12.4	0.5
≥ 120	11	3	-11.9	1.2

Table B.17: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the NNLOPS ggH and VBF $H \to \gamma\gamma$ samples for each bin of the $p_{\mathrm{T}}^{j_2}$ observable, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of ggH and VBF events in each bin.

	ZH		$ggZH$		W^-H		W^+H	
Bins of $p_{\mathrm{T}}^{j_2}$ [GeV]	Variation [%]	Error [%]	Variation [%]	Error [%]	Variation [%]	Error [%]	Variation [%]	Error [%]
0–30	0.9	0.9	-7.5	1.7	0.5	1.1	0.7	1.1
30–70	-0.7	0.9	4.6	1.3	-1.6	1.1	-1.1	1.2
70–120	-1.7	1.8	7.7	1.9	-4	2	-3	2
≥ 120	2	3	-7	3	-4	4	-5	4

Table B.18: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the VH $H \to \gamma\gamma$ samples for each bin of the $p_{\mathrm{T}}^{j_2}$ observable, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of VH events passing each bin.

	$\mathbf{ISR_1}$ up		$\mathbf{ISR_1}$ down		$\mathbf{ISR_2}$ up		$\mathbf{ISR_2}$ down	
Bins of $p_{\mathrm{T}}^{j_2}$ [GeV]	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0–30	0.1	0.2	1.3	0.2	0.2	0.2	0.3	0.2
30–70	0.4	0.3	1.0	0.4	0.3	0.4	0.3	0.4
70–120	-0.2	0.9	1.6	0.9	0.0	0.9	0.9	0.9
≥ 120	1.8	1.7	1.8	1.7	-0.8	1.7	0.3	1.7

Table B.19: Differences in acceptance between AZNLO (ISR eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for each bin of the $p_{\mathrm{T}}^{j_2}$ observable, as described in [215]. This difference corresponds to (AZNLO (ISR eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each bin.

	FSR up		FSR down		MPI up		MPI down	
Bins of $p_{\mathrm{T}}^{j_2}$ [GeV]	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0–30	1.8	0.2	-0.1	0.2	0.1	0.2	0.4	0.2
30–70	3.7	0.4	-2.8	0.4	-0.4	0.4	0.2	0.4
70–120	3.7	0.9	-2.7	0.9	0.2	0.9	0.0	0.9
≥ 120	4.4	1.7	-2.0	1.6	0.7	1.7	0.7	1.7

Table B.20: Differences in acceptance between AZNLO (FSR / MPI eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for each bin of the $p_{\mathrm{T}}^{j_2}$ observable, as described in [215]. This difference corresponds to (AZNLO (FSR / MPI eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each bin.

	ggH		**VBF**	
Bins of $\lvert y^{j_1} \rvert$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0.0–0.5	-2.1	0.6	-4.2	0.6
0.5–1.0	0.3	0.6	-3.4	0.6
1.0–1.5	0.3	0.7	-2.8	0.5
1.5–1.9	0.3	0.8	-3.2	0.5
1.9–2.3	1.3	0.9	-2.8	0.5
2.3–2.5	1.1	1.3	-2.2	0.7
2.5–4.4	3.3	0.6	-0.7	0.3

Table B.21: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the NNLOPS ggH and VBF $H \to \gamma\gamma$ samples for each bin of the $|y^{j_1}|$ observable, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of ggH and VBF events in each bin.

Appendix B. Underlying event, parton shower and hadronization uncertainties for fiducial integrated and differential results

	ZH		$ggZH$		W^-H		W^+H	
Bins of $\lvert y^{j_1}\rvert$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0.0–0.5	-3.2	1.1	-7.5	1.6	-2.4	1.4	-0.8	1.4
0.5–1.0	1.0	1.2	-8.0	1.7	-1.3	1.5	-1.5	1.6
1.0–1.5	1.6	1.3	-1	2	-0.6	1.6	-1.1	1.7
1.5–1.9	0.9	1.8	-2	3	1	2	1	2
1.9–2.3	-2	2	16	3	4	3	-3	3
2.3–2.5	3	4	37	6	-7	4	-2	4
2.5–4.4	3	2	40	4	-2	3	4	3

Table B.22: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the VH $H \to \gamma\gamma$ samples for each bin of the $|y^{j_1}|$ observable, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of VH events passing each bin.

	ISR_1 up		**ISR_1 down**		**ISR_2 up**		**ISR_2 down**	
Bins of $\lvert y^{j_1}\rvert$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0.0–0.5	-0.2	0.4	0.6	0.4	-0.2	0.4	0.1	0.4
0.5–1.0	0.6	0.4	1.9	0.4	0.5	0.4	0.3	0.4
1.0–1.5	0.6	0.5	1.0	0.5	0.3	0.5	0.8	0.5
1.5–1.9	0.3	0.6	1.3	0.6	0.4	0.6	0.3	0.6
1.9–2.3	0.2	0.6	1.5	0.6	0.4	0.6	0.3	0.6
2.3–2.5	-0.1	0.9	0.6	0.9	0.1	0.9	0.8	0.9
2.5–4.4	0.0	0.4	1.4	0.4	0.0	0.4	0.1	0.4

Table B.23: Differences in acceptance between AZNLO (ISR eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for each bin of the $|y^{j_1}|$ observable, as described in [215]. This difference corresponds to (AZNLO (ISR eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each bin.

	FSR up		**FSR down**		**MPI up**		**MPI down**	
Bins of $\lvert y^{j_1}\rvert$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0.0–0.5	1.8	0.4	-1.4	0.4	-0.4	0.4	-0.2	0.4
0.5–1.0	2.9	0.5	-0.1	0.4	0.2	0.4	0.7	0.4
1.0–1.5	2.7	0.5	-0.5	0.5	0.3	0.5	0.6	0.5
1.5–1.9	2.6	0.6	-0.9	0.5	-0.2	0.5	0.0	0.5
1.9–2.3	1.8	0.6	-0.9	0.6	0.3	0.6	0.8	0.6
2.3–2.5	2.2	0.9	-1.6	0.9	-0.4	0.9	0.2	0.9
2.5–4.4	2.6	0.4	-1.7	0.4	-0.2	0.4	0.2	0.4

Table B.24: Differences in acceptance between AZNLO (FSR / MPI eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for each bin of the $|y^{j_1}|$ observable, as described in [215]. This difference corresponds to (AZNLO (FSR / MPI eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each bin.

	ggH		VBF	
Bins of $\lvert y^{j_2}\rvert$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0.0–1.2	5.9	0.8	-9.2	0.5
1.2–2.0	7.9	1.0	-7.9	0.5
2.0–4.4	5.2	0.9	-6.6	0.3

Table B.25: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the NNLOPS ggH and VBF $H \to \gamma\gamma$ samples for each bin of the $|y^{j_2}|$ observable, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of ggH and VBF events in each bin.

	ZH		$ggZH$		W^-H		W^+H	
Bins of $\lvert y^{j_2}\rvert$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0.0–1.2	-1.5	1.0	-0.7	1.3	-2.1	1.3	-1.0	1.4
1.2–2.0	-1.0	1.5	2	2	-2.3	1.9	-3.7	1.9
2.0–4.4	2.5	1.9	21	3	-3	2	-1	2

Table B.26: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the VH $H \to \gamma\gamma$ samples for each bin of the $|y^{j_2}|$ observable, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of VH events passing each bin.

	ISR_1 up		ISR_1 down		ISR_2 up		ISR_2 down	
Bins of $\lvert y^{j_2}\rvert$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0.0–1.2	0.3	0.5	1.1	0.5	0.1	0.5	0.4	0.5
1.2–2.0	1.8	0.7	2.8	0.7	1.4	0.7	1.4	0.7
2.0–4.4	-0.6	0.6	-0.1	0.6	-0.5	0.6	-0.2	0.6

Table B.27: Differences in acceptance between AZNLO (ISR eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for each bin of the $|y^{j_2}|$ observable, as described in [215]. This difference corresponds to (AZNLO (ISR eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each bin.

	FSR up		FSR down		MPI up		MPI down	
Bins of $\lvert y^{j_2}\rvert$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0.0–1.2	3.5	0.5	-2.5	0.5	-0.5	0.5	0.3	0.5
1.2–2.0	4.4	0.7	-1.8	0.7	0.8	0.7	0.9	0.7
2.0–4.4	3.5	0.6	-3.8	0.6	-0.8	0.6	-0.5	0.6

Table B.28: Differences in acceptance between AZNLO (FSR / MPI eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for each bin of the $|y^{j_2}|$ observable, as described in [215]. This difference corresponds to (AZNLO (FSR / MPI eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each bin.

	ggH		**VBF**	
Bins of $\lvert\cos\theta^*\rvert$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0.0–0.0625	0.2	0.6	0.4	0.6
0.0625–0.125	-0.9	0.6	0.6	0.6
0.125–0.1875	-1.7	0.6	0.2	0.6
0.1875–0.25	-0.4	0.6	-0.8	0.6
0.25–0.3125	-0.3	0.6	1.0	0.6
0.3125–0.3750	-1.3	0.6	-0.2	0.6
0.3750–0.5	-0.6	0.4	-0.1	0.5
0.5–0.625	0.3	0.5	1.2	0.5
0.625–1.0	-1.2	0.4	-1.3	0.4

Table B.29: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the NNLOPS ggH and VBF $H \to \gamma\gamma$ samples for each bin of the $\lvert\cos\theta^*\rvert$ observable, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of ggH and VBF events in each bin.

	ZH		***ggZH***		***W*$^-$*H***		***W*$^+$*H***	
Bins of $\lvert\cos\theta^*\rvert$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0.0–0.0625	0.8	1.7	-2	3	-1	2	-5	2
0.0625–0.125	-3.3	1.7	2	3	-3	2	0	2
0.125–0.1875	-0.2	1.8	2	3	4	2	4	2
0.1875–0.25	1.3	1.8	2	3	-4	2	2	2
0.25–0.3125	0.4	1.9	-8	3	1	2	0	2
0.3125–0.3750	0.4	1.9	0	3	-3	2	0	2
0.3750–0.5	1.7	1.4	0	2	-1.2	1.7	-0.8	1.7
0.5–0.625	-1.1	1.5	-5	2	-1.0	1.7	-1.8	1.8
0.625–1.0	-0.8	1.1	0.4	1.4	0.9	1.3	-2.7	1.3

Table B.30: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the VH $H \to \gamma\gamma$ samples for each bin of the $\lvert\cos\theta^*\rvert$ observable, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of VH events passing each bin.

	ISR_1 up		ISR_1 down		ISR_2 up		ISR_2 down	
Bins of $\lvert\cos\theta^*\rvert$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0.0–0.0625	-0.1	0.3	0.7	0.4	-0.2	0.4	-0.2	0.4
0.0625–0.125	-0.3	0.4	0.9	0.4	-0.8	0.4	-0.4	0.4
0.125–0.1875	-0.3	0.4	0.5	0.4	-0.5	0.4	-0.4	0.4
0.1875–0.25	0.0	0.4	1.1	0.4	0.1	0.4	-0.1	0.4
0.25–0.3125	0.3	0.4	1.4	0.4	-0.3	0.4	-0.1	0.4
0.3125–0.3750	-0.1	0.4	0.7	0.4	0.0	0.4	0.1	0.4
0.3750–0.5	-0.1	0.3	0.9	0.3	-0.2	0.3	0.1	0.3
0.5–0.625	0.3	0.3	1.4	0.3	0.3	0.3	0.5	0.3
0.625–1.0	0.3	0.3	1.5	0.3	0.5	0.3	0.7	0.3

Table B.31: Differences in acceptance between AZNLO (ISR eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for each bin of the $\lvert\cos\theta^*\rvert$ observable, as described in [215]. This difference corresponds to (AZNLO (ISR eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each bin.

	FSR up		FSR down		MPI up		MPI down	
Bins of $\lvert\cos\theta^*\rvert$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0.0–0.0625	0.3	0.4	0.1	0.4	-0.3	0.3	-0.1	0.4
0.0625–0.125	0.4	0.4	-0.1	0.4	-0.7	0.4	-0.8	0.4
0.125–0.1875	0.2	0.4	0.0	0.4	-0.2	0.4	-0.7	0.4
0.1875–0.25	0.4	0.4	0.1	0.4	0.0	0.4	0.1	0.4
0.25–0.3125	0.9	0.4	0.3	0.4	0.1	0.4	-0.1	0.4
0.3125–0.3750	1.0	0.4	0.2	0.4	0.1	0.4	0.5	0.4
0.3750–0.5	1.0	0.3	0.5	0.3	0.0	0.3	-0.1	0.3
0.5–0.625	1.3	0.3	1.0	0.3	0.4	0.3	0.3	0.3
0.625–1.0	1.2	0.3	0.6	0.3	0.4	0.3	0.3	0.3

Table B.32: Differences in acceptance between AZNLO (FSR / MPI eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for each bin of the $\lvert\cos\theta^*\rvert$ observable, as described in [215]. This difference corresponds to (AZNLO (FSR / MPI eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each bin.

	ggH		**VBF**	
Bins of $\Delta\phi_{jj}$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
-3.14–(-1.57)	7.6	1.0	-8.7	0.4
-1.57–0.00	3.9	1.0	-5.5	0.5
0.00–1.57	5.1	1.0	-5.1	0.5
1.57–3.14	7.7	1.0	-9.0	0.4

Table B.33: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the NNLOPS ggH and VBF $H \to \gamma\gamma$ samples for each bin of the $\Delta\phi_{jj}$ observable, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of ggH and VBF events in each bin.

	ZH		$ggZH$		W^-H		W^+H	
Bins of $\Delta\phi_{jj}$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
-3.14–(-1.57)	-0.8	1.6	12	2	0	2	-2	2
-1.57–0.00	0.3	1.6	-1.7	1.8	-4.1	1.8	1.2	1.9
0.00–1.57	-1.9	1.5	-1.6	1.8	-2.5	1.8	-2.5	1.9
1.57–3.14	-0.3	1.6	15	2	-2	2	-4	2

Table B.34: Differences in acceptance between HERWIG7 and PYTHIA8 (AZNLO) obtained with the VH $H \to \gamma\gamma$ samples for each bin of the $\Delta\phi_{jj}$ observable, as described in [215]. This difference corresponds to (HERWIG7-PYTHIA8 (AZNLO)) / PYTHIA8 (AZNLO), and the error is statistical only, which depends on the number of VH events passing each bin.

	ISR_1 up		**ISR_1 down**		**ISR_2 up**		**ISR_2 down**	
Bins of $\Delta\phi_{jj}$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
-3.14–(-1.57)	0.8	0.7	0.7	0.7	0.2	0.7	0.5	0.7
-1.57–0.00	-0.5	0.7	-0.3	0.7	-0.7	0.7	-0.7	0.7
0.00–1.57	0.7	0.7	1.2	0.7	0.4	0.7	1.1	0.7
1.57–3.14	0.5	0.7	2.8	0.7	1.0	0.7	0.8	0.7

Table B.35: Differences in acceptance between AZNLO (ISR eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for each bin of the $\Delta\phi_{jj}$ observable, as described in [215]. This difference corresponds to (AZNLO (ISR eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each bin.

	FSR up		**FSR down**		**MPI up**		**MPI down**	
Bins of $\Delta\phi_{jj}$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
-3.14–(-1.57)	4.4	0.7	-3.6	0.6	0.4	0.7	-0.1	0.7
-1.57–0.00	2.2	0.7	-3.1	0.7	-1.7	0.7	-0.9	0.7
0.00–1.57	3.1	0.7	-2.2	0.7	-0.6	0.7	0.4	0.7
1.57–3.14	5.1	0.7	-2.2	0.6	0.7	0.7	1.2	0.7

Table B.36: Differences in acceptance between AZNLO (FSR / MPI eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for each bin of the $\Delta\phi_{jj}$ observable, as described in [215]. This difference corresponds to (AZNLO (FSR / MPI eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each bin.

	ggH		VBF	
Bins of $\|\Delta y_{jj}\|$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0.0–2.0	5.4	0.6	-2.4	0.6
2.0–4.0	6.0	0.9	-9.9	0.4
4.0–8.8	11.1	1.6	-7.2	0.3

Table B.37: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the NNLOPS ggH and VBF $H \to \gamma\gamma$ samples for each bin of the $|\Delta y_{jj}|$ observable, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of ggH and VBF events in each bin.

	ZH		*ggZH*		W^-H		W^+H	
Bins of $\|\Delta y_{jj}\|$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0.0–2.0	-1.2	0.8	-4.5	1.1	-2.4	1.0	-2.2	1.1
2.0–4.0	1	2	32	3	-1	3	1	3
4.0–8.8	9	6	57	7	-3	7	-3	7

Table B.38: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the VH $H \to \gamma\gamma$ samples for each bin of the $|\Delta y_{jj}|$ observable, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of VH events passing each bin.

	ISR_1 up		ISR_1 down		ISR_2 up		ISR_2 down	
Bins of $\|\Delta y_{jj}\|$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0.0–2.0	0.7	0.4	1.4	0.4	0.4	0.4	0.9	0.4
2.0–4.0	-0.3	0.6	0.6	0.6	-0.2	0.6	-0.5	0.6
4.0–8.8	0.5	1.1	0.9	1.1	0.1	1.1	0.5	1.1

Table B.39: Differences in acceptance between AZNLO (ISR eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for each bin of the $|\Delta y_{jj}|$ observable, as described in [215]. This difference corresponds to (AZNLO (ISR eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each bin.

	FSR up		FSR down		MPI up		MPI down	
Bins of $\|\Delta y_{jj}\|$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0.0–2.0	3.6	0.4	-2.0	0.4	-0.1	0.4	0.4	0.4
2.0–4.0	3.5	0.6	-3.7	0.6	-0.8	0.6	-0.4	0.6
4.0–8.8	5.1	1.1	-4.1	1.0	0.1	1.1	0.4	1.1

Table B.40: Differences in acceptance between AZNLO (FSR / MPI eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for each bin of the $|\Delta y_{jj}|$ observable, as described in [215]. This difference corresponds to (AZNLO (FSR / MPI eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each bin.

	ggH		**VBF**	
Bins of $\lvert\Delta\phi_{\gamma\gamma,jj}\rvert$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0.00–0.13	-0.1	0.7	-8.1	0.3
0.13–0.42	7.5	0.8	-3.1	0.5
0.42–3.14	15.3	1.1	-12.9	0.6

Table B.41: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the NNLOPS ggH and VBF $H \to \gamma\gamma$ samples for each bin of the $|\Delta\phi_{\gamma\gamma,jj}|$ observable, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of ggH and VBF events in each bin.

	ZH		***ggZH***		***W⁻H***		***W⁺H***	
Bins of $\lvert\Delta\phi_{\gamma\gamma,jj}\rvert$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0.00–0.13	-1.0	1.3	-5.8	1.6	-5.3	1.5	-3.1	1.6
0.13–0.42	-2.5	1.3	6.2	1.7	0.3	1.7	-1.9	1.8
0.42–3.14	1.8	1.5	14	2	-0.8	1.8	0.5	1.9

Table B.42: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the VH $H \to \gamma\gamma$ samples for each bin of the $|\Delta\phi_{\gamma\gamma,jj}|$ observable, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of VH events passing each bin.

	ISR_1 up		**ISR_1 down**		**ISR_2 up**		**ISR_2 down**	
Bins of $\lvert\Delta\phi_{\gamma\gamma,jj}\rvert$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0.00–0.13	0.0	0.5	0.7	0.5	-0.1	0.5	0.0	0.5
0.13–0.42	0.1	0.6	0.8	0.6	-0.1	0.6	0.5	0.6
0.42–3.14	1.4	0.7	2.3	0.7	1.2	0.7	1.0	0.7

Table B.43: Differences in acceptance between AZNLO (ISR eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for each bin of the $|\Delta\phi_{\gamma\gamma,jj}|$ observable, as described in [215]. This difference corresponds to (AZNLO (ISR eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each bin.

	FSR up		**FSR down**		**MPI up**		**MPI down**	
Bins of $\lvert\Delta\phi_{\gamma\gamma,jj}\rvert$	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0.00–0.13	3.6	0.5	-3.5	0.5	-0.7	0.5	-0.1	0.5
0.13–0.42	3.7	0.6	-2.8	0.5	0.0	0.6	-0.1	0.6
0.42–3.14	4.0	0.7	-1.3	0.7	0.1	0.7	0.9	0.7

Table B.44: Differences in acceptance between AZNLO (FSR / MPI eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for each bin of the $|\Delta\phi_{\gamma\gamma,jj}|$ observable, as described in [215]. This difference corresponds to (AZNLO (FSR / MPI eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each bin.

	ggH		VBF	
Bins of m_{jj} [GeV]	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0–170	6.0	0.6	1.7	0.7
170–500	5.3	0.8	-9.2	0.4
500–1500	8.4	1.5	-8.4	0.3
$\geq$ 1500	20	6	-7.6	0.6

Table B.45: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the NNLOPS ggH and VBF $H \to \gamma\gamma$ samples for each bin of the m_{jj} observable, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of ggH and VBF events in each bin.

	ZH		$ggZH$		W^-H		W^+H	
Bins of m_{jj} [GeV]	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0–170	-0.9	0.9	-3.7	1.3	-2.4	1.1	-2.5	1.2
170–500	-1.6	1.6	8.6	1.9	-2	2	-0.7	2.1
500–1500	7	4	36	4	1	5	4	5
$\geq$ 1500	18	22	129	22	-3	26	25	28

Table B.46: Differences in acceptance between Herwig7 and Pythia8 (AZNLO) obtained with the VH $H \to \gamma\gamma$ samples for each bin of the m_{jj} observable, as described in [215]. This difference corresponds to (Herwig7-Pythia8 (AZNLO)) / Pythia8 (AZNLO), and the error is statistical only, which depends on the number of VH events passing each bin.

	ISR_1 up		ISR_1 down		ISR_2 up		ISR_2 down	
Bins of m_{jj} [GeV]	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0–170	0.6	0.5	1.2	0.5	0.4	0.5	0.8	0.5
170–500	0.3	0.6	1.3	0.6	0.1	0.6	0.1	0.6
500–1500	-0.4	1.0	0.3	1.0	-0.2	1.0	0.2	1.0
$\geq$ 1500	-3	3	0	3	-4	3	-3	3

Table B.47: Differences in acceptance between AZNLO (ISR eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for each bin of the m_{jj} observable, as described in [215]. This difference corresponds to (AZNLO (ISR eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each bin.

	FSR up		FSR down		MPI up		MPI down	
Bins of m_{jj} [GeV]	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**	**Variation [%]**	**Error [%]**
0–170	3.7	0.5	-2.3	0.5	-0.4	0.5	0.4	0.5
170–500	3.6	0.6	-2.9	0.5	0.0	0.6	-0.1	0.6
500–1500	4.1	1.1	-3.7	1.0	-0.5	1.0	0.0	1.0
$\geq$ 1500	1	4	-5	3	-2	3	-1	3

Table B.48: Differences in acceptance between AZNLO (FSR / MPI eigentune variations) and AZNLO obtained with the NNLOPS ggH $H \to \gamma\gamma$ sample for each bin of the m_{jj} observable, as described in [215]. This difference corresponds to (AZNLO (FSR / MPI eigentune variations)-AZNLO) / AZNLO, and the error is statistical only, that depends largely on the number of ggH events passing each bin.

Bibliography

[1] ATLAS Collaboration. Operation and performance of the ATLAS Tile Calorimeter in Run 1. *Eur. Phys. J. C*, 2018. arXiv:1806.02129 [hep-ex].

[2] ATLAS Collaboration. Search for Higgs boson pair production in the $b\bar{b}\gamma\gamma$ final state using pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector. Technical Report ATLAS-CONF-2016-004, CERN, March 2016.

[3] ATLAS Collaboration. Search for Higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state with 13 TeV *pp* collision data collected by the ATLAS experiment. *JHEP*, 2018. arXiv:1807.04873 [hep-ex].

[4] ATLAS Collaboration. Measurements of Higgs boson properties in the diphoton decay channel with 36 fb^{-1} of *pp* collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. D*, 2018. arXiv:1802.04146 [hep-ex].

[5] D. Griffiths. *Introduction to Elementary Particles.* John Wiley & Sons, 1987.

[6] J. Ellis. Higgs Physics, 2013. arXiv:1312.5672 [hep-ph].

[7] A. Pich. The Standard Model of Electroweak Interactions, 2012. arXiv:1201.0537 [hep-ph].

[8] A. Djouadi. The anatomy of electro-weak symmetry breaking, 2005. arXiv:0503172 [hep-ph].

[9] C.N. Yang and R.L. Mills. Conservation of Isotopic Spin and Isotopic Gauge Invariance. *Phys. Rev.*, 96(1):191–195, October 1954.

[10] Particle Data Group. The Review of Particle Physics. *Chin. Phys. C*, 40(100001), 2017.

[11] S.L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22(4):579–588, 1961.

[12] A. Salam. Weak and Electromagnetic Interactions. *Conf. Proc.*, C680519:367–377, 1968.

[13] S. Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19(21):1264–1266, November 1967.

[14] G. t Hooft. Renormalizable Lagrangians for massive Yang-Mills fields. *Nuclear Physics B*, 35(1):167–188, 1971.

[15] P.W. Higgs, Phys. Rev. Lett. 12, 132 (1964),
P.W. Higgs, Phys. Rev. 145, 1156 (1966),
F. Englert and R. Brout, Phys. Rev. Lett. 13, 321 (1964),
G.S. Guralnik, C.R. Hagen, and T.W. Kibble, Phys. Rev. Lett. 13, 585 (1964).

[16] M.E. Peskin and D. V. Schroeder. *An introduction to quantum field theory.* Addison-Wesley Advanced Book Program, USA, 1995.

[17] UA1 Collaboration. Experimental observation of lepton pairs of invariant mass around 95 GeV/c^2 at the CERN SPS collider. *Physics Letters B*, 126:398–410, 1983.

[18] UA2 Collaboration. Evidence for $Z^0 \to e^+e$ at the CERN pp collider. *Physics Letters B*, 129:130–140, 1983.

[19] G.P. Salam. Elements of QCD for hadron colliders, 2010. arXiv:1011.5131 [hep-ph].

[20] J.M. Campbell et al. Hard interactions of quarks and gluons: a primer for LHC physics. *Reports on Progress in Physics*, 70(1):89–193, 2007.

[21] Proton-proton collision image, https://imperialhep.blogspot.com/2011/08/strangeness-at-lhcb.html, 2011.

[22] G. Altarelli and G. Parisi. Asymptotic freedom in parton language. *Nuclear Physics B*, 126(2):298–318, 1977.

[23] S. Alekhin et al. Fixed target Drell-Yan data and NNLO QCD fits of parton distribution functions. *Phys. Rev. D*, 74(054033), 2006.

[24] A.D. Martin et al. Parton distributions for the LHC. *Eur. Phys. J.*, C63:189–285, 2009.

[25] H1 and F.D. Aaron et al. ZEUS collaborations. Combined Measurement and QCD Analysis of the Inclusive $e^{\pm}$ p Scattering Cross Sections at HERA. *JHEP*, 01(109), 2010.

[26] R.D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04(40), 2015.

[27] P.M. Nadolsky et al. Implications of CTEQ global analysis for collider observables. *Phys. Rev. D*, D78(013004), 2008.

[28] T. Sjostrand. Monte Carlo Generators, 2006. arXiv:0611247 [hep-ph].

[29] B. Andersson et al. Parton Fragmentation and String Dynamics. *Phys.Rept.*, 97:31–145, 1983.

[30] G. Marchesini and B. R. Webber. Simulation of QCD jets including soft gluon interference. *Nucl. Phys. B*, 238(1), 1984.

[31] The LHC Higgs Cross Section Working Group. *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*. CERN, Geneva, 2011. arXiv:1101.0593 [hep-ph].

[32] The LHC Higgs Cross Section Working Group. *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*. CERN, Geneva, 2012. arXiv:1201.3084 [hep-ph].

[33] The LHC Higgs Cross Section Working Group. *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*. CERN, Geneva, 2013. arXiv:1307.1347 [hep-ph].

[34] The LHC Higgs Cross Section Working Group. *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*. CERN, Geneva, 2016. arXiv:1610.07922 [hep-ph].

[35] L3 ALEPH, DELPHI and The LEP Working Group for Higgs Boson Searches OPAL Collaborations. Search for neutral MSSM Higgs bosons at LEP. *European Physical Journal C - Particles and Fields*, 47(3):547–587, September 2006.

[36] The TEVNPH Working Group of the CDF and D0 Collaboration. Combined cdf and d0 upper limits on standard model higgs-boson production with up to 6.7 fb^{-1} of data. *FERMILAB-CONF-10-257-E*, 2010.

[37] 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. B*, 716(1):1–29, September 2012.

[38] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30–61, September 2012.

[39] CMS Collaborations ATLAS. Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC *pp* col lision data at $\sqrt{s} = 7$ and 8 TeV. *JHEP*, 08(045), 2016. arXiv:1606.02266 [hep-ex].

[40] Planck Collaboration. Planck 2015 results. XIII. Cosmological parameters. *Astron. Astrophys.*, 594(A13), 2016.

[41] R. Frederix et al. Higgs pair production at the LHC with NLO and parton-shower effects, 2014. arXiv:1401.7340 [hep-ph].

[42] J. Baglio et al. The measurement of the Higgs self-coupling at the LHC: theoretical status. *Journal of High Energy Physics*, 1304(151), 2013.

[43] R. Contino et al. Anomalous Couplings in Double Higgs Production. *Journal of High Energy Physics*, 1208(154), 2012.

[44] G.D. Kribs and A. Martin. Enhanced di-Higgs Production through Light Colored Scalars. *Phys. Rev. D*, D86(95023), 2012.

[45] T. D. Lee. A Theory of Spontaneous T Violation. *Physical Review D*, 8(1226), August 1973.

[46] S. Dimopoulos and H. Georgi. Softly broken supersymmetry and SU(5). *Nuclear Physics B*, 193(1):150–162, December 1981.

[47] M. Papucci Z. Chacko, Y. Nomura and G. Perez. Natural little hierarchy from a partially goldstone twin Higgs. *Journal of High Energy Physics*, 2006(01), January 2006. arXiv:0510273 [hep-ph].

[48] R. Grober and M. Muhlleitner. Composite Higgs boson pair production at the LHC. *Journal of High Energy Physics*, 2011(20), June 2011. arXiv:1012.1562 [hep-ph].

[49] J. Mrazek et al. The other natural two Higgs doublet model. *Nuclear Physics B*, 853(1):1–48, December 2011.

[50] G. Branco et al. Theory and phenomenology of two-Higgs-doublet models, 2011. arXiv:1106.0034 [hep-ph].

[51] J. Baglio et al. Benchmarks for Higgs Pair Production and Heavy Higgs Searches in the Two-Higgs-Doublet Model of Type II, 2014. arXiv:1403.1264 [hep-ph].

[52] B. Bhattacherjee et al. The role of MSSM heavy Higgs production in the self coupling measurement of the 125 GeV Higgs boson at the LHC, 2015. arXiv:1407.6866 [hep-ph].

[53] Chien-Yi Chen et al. Exploring Resonant di-Higgs production in the Higgs Singlet Model, 2015. arXiv:1410.5488 [hep-ph].

[54] L. Randall and R. Sundrum. Large Mass Hierarchy from a Small Extra Dimension. *Phys. Rev. Lett.*, 83(3370), May 1999.

[55] ATLAS Collaboration. Searches for Higgs boson pair production in the $hh \to bb\tau\tau, \gamma\gamma WW^*, \gamma\gamma b\bar{b}, bbbb$ channels with the ATLAS detector. *Phys. Rev. D*, 92, 2015. arXiv:1509.04670 [hep-ph].

[56] ATLAS Collaboration. Search for Higgs Boson Pair Production in the $\gamma\gamma b\bar{b}$ Final State Using pp Collision Data at $\sqrt{s} = 8$ TeV from the ATLAS Detector. *Phys. Rev. Lett.*, 114, 2015. arXiv:1406.5053 [hep-ex].

[57] L. Evans and P. Bryant. LHC Machine. *Journal of Instrumentation*, 3(08):S08003, 2008.

[58] O. S.Brüning et al. LHC Design Report vol.I: The LHC Main Ring. Technical Report CERN-2004-003-V-1, CERN, Geneva, 2004.

[59] O. S.Brüning et al. LHC Design Report vol.II: the LHC Infrastructure and General Services. Technical Report CERN-2004-003-V-2, CERN, Geneva, 2004.

[60] O. S.Brüning et al. LHC Design Report vol.III: the LHC Injector Chain. Technical Report CERN-2004-003-V-3, CERN, Geneva, 2004.

[61] F. Fayette. Strategies for precision measurements of the charge asymmetry of the W boson mass at the LHC within the ATLAS experiment, 2009. arXiv:0906.4260 [hep-ex].

[62] ATLAS Collaboration. *ATLAS Technical Proposal for a General-Purpose pp experiment at the Large Hadron Collider at CERN.* CERN, Geneva, 1994. CERN-LHC-94-43.

[63] ALICE Collaboration. *ALICE : Technical proposal for a Large Ion collider Experiment at the CERN LHC.* CERN, Geneva, 1995. CERN-LHCC-95-71; LHCC-P-3.

[64] CMS Collaboration. *Technical Proposal.* CERN, Geneva, 1994. CERN-LHCC-94-38 ; LHCC-P-1.

[65] LHCb Collaboration. *LHCb: Technical Proposal*. CERN, Geneva, 1998. CERN-LHCC-98-004; LHCC-P-4.

[66] L. Evans. The Large Hadron Collider. *Philosophical Transactions of the Royal Society A*, 2012.

[67] ATLAS Collaboration, "Run 2 Luminosity Public Plots.". `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2/`.

[68] ATLAS Collaboration, "Run 1 Luminosity Public Plots.". `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults/`.

[69] ATLAS collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3(08):S08003, 2008.

[70] J. Goodson. Public images. `http://www.jetgoodson.com/images/thesisImages/magnetSystems.png`.

[71] ATLAS Collaboration. Impact Parameter Resolution, 2016.

[72] ATLAS Collaboration. Public Liquid-Argon Calorimeter Plots on Detector Status.

[73] ATLAS Collaboration. ATLAS detector and physics performance: Technical Design Report, 1. Technical Report ATLAS-TDR-14; CERN-LHCC-99-014 (1999), CERN, Geneva, 1999. https://cds.cern.ch/record/391176.

[74] O. Rifki and B. Abbott. Search for supersymmetry in final states with two same sign leptons or three leptons and jets with the ATLAS detector at the LHC.

[75] T. Berners-Lee. Information management: a proposal. Technical Report CERN-DD-89-001-OC, CERN, 1989.

[76] The Worldwide LHC Computing Grid. Wlcg-tiersjun14, 2016.

[77] T. Cornelissen et al. Concepts, Design and Implementation of the ATLAS New Tracking (NEWT). Technical Report ATL-SOFT-PUB-2007-007, CERN, March 2007.

[78] T. Cornelissen et al. Single Track Performance of the Inner Detector New Track Reconstruction (NEWT). Technical Report ATL-INDET-PUB-2008-002, CERN, March 2008.

[79] R.Fruhwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth.*, A262:444 – 450, 1987.

[80] P. Vanlaer R. Fruhwirth, W. Waltenberger. Adaptive vertex fitting. *J.Phys.*, G34, 2007.

[81] ATLAS Collaboration. Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-1 data. *Eur. Phys. J. C*, 76, 2016.

[82] W. Lampl et al. Calorimeter Clustering Algorithms: Description and Performance. Technical Report ATL-LARG-PUB-2008-002, CERN, 2008.

[83] et al. T. G. Cornelissen. The global χ^2 track fitter in ATLAS. *J. Phys. Conf. Ser.*, 119, 2008.

[84] ATLAS Collaboration. Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung. Technical Report ATLAS-CONF-2012-047, CERN, May 2012.

[85] L. Cerda Alberich. Photon and electron identification with the ATLAS detector. Technical Report ATL-PHYS-PROC-2016-230, CERN, Nov 2016.

[86] ATLAS Collaboration. Photon identification in 2015 ATLAS data, 2016. ATL-PHYS-PUB-2016-014.

[87] ATLAS Collaboration. Photon Identification Efficiencies using 2016 Data with radiative Z boson decays, 2016. EGAM-2016-003.

[88] ATLAS Collaboration. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. *Eur. Phys. J. C*, 77, 2017. arXiv:1603.02934 [hep-ex].

[89] G. P. Salam M. Cacciari and S. Sapeta. On the characterisation of the underlying event. *JHEP*, 04, 2010. arXiv:0912.4926 [hep-ph].

[90] ATLAS Collaboration. Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data. *Eur. Phys. J. C*, 74, 2014.

[91] ATLAS Collaboration. Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton–proton collision data. Technical Report ATLAS-CONF-2014-032, CERN, 2014.

[92] ATLAS Collaboration. Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton–proton collision data. Technical Report ATLAS-CONF-2016-024, CERN, 2016.

[93] ATLAS Collaboration. Electron identification efficiency measured with $Z \to ee$ events using 2016 data. Technical Report EGAM-2016-002, CERN, 2016.

[94] ATLAS Collaboration. Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data. *Eur. Phys. J. C*, 74, 2014. arXiv:1407.3935 [hep-ex].

[95] ATLAS Collaboration. Muon reconstruction performance of the ATLAS detector in protonproton collision data at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 76, 2016. arXiv:1603.05598 [hep-ex].

[96] J. Illingworth and J. Kittler. A survey of the Hough transform. *Computer Vision, Graphics, and Image Processing*, 44:87 – 116, 1988.

[97] ATLAS Collaboration. ATLAS Muon Combined Performance with the full 2016 dataset. Technical Report MUON-2017-001, CERN, 2017.

[98] G. P. Salam M. Cacciari and G. Soyez. The anti-kt jet clustering algorithm. *JHEP*, 04(63), 2008. arXiv:0802.1189 [hep-ph].

[99] ATLAS Collaboration. Jet energy measurement and its systematic uncertainty in proton–proton collisions at $\sqrt{s} = 7$ tev. *Eur. Phys. J. C*, 75(17), 2015. arXiv:1406.0076 [hep-ex].

[100] ATLAS Collaboration. Jet reconstruction and performance using particle flow with the atlas detector. *Eur. Phys. J. C*, 77(466), 2017. arXiv:1703.10485 [hep-ex].

[101] ATLAS Collaboration. Local Hadronic Calibration. Technical Report ATL-LARG-PUB-2009-001, CERN, 2009.

[102] ATLAS Collaboration. Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at $\sqrt{s} = 13$ tev with the atlas detector. *Phys. Rev. D*, 96(072002), 2017. arXiv:1703.09665 [hep-ex].

[103] ATLAS Collaboration. Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV. Technical Report ATLAS-PHYS-PUB-2015-015, CERN, 2015.

[104] ATLAS Collaboration. Jet global sequential corrections with the ATLAS detector in proton–proton collisions at $\sqrt{s} = 8$ TeV. Technical Report ATLAS-CONF-2015-002, CERN, 2015.

[105] ATLAS Collaboration. Relative jet response obtained with eta-intercalibration for 2017 data. Technical Report JETM-2017-008, CERN, 2017.

[106] ATLAS Collaboration. JES Public Plots for Moriond 2017. Technical Report JETM-2017-003, CERN, 2017.

[107] ATLAS Collaboration. Calibration of light-flavour b-jet mistagging rates using ATLAS proton–proton collision data at $\sqrt{s} = 13$ TeV. Technical Report ATLAS-CONF-2018-006, CERN, 2018.

[108] ATLAS Collaboration. Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV. *JHEP*, 08, 2018.

[109] ATLAS Collaboration. Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run. Technical Report ATL-PHYS-PUB-2016-012, CERN, 2016.

[110] ATLAS Collaboration. Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment. Technical Report ATLAS-CONF-2014-004, CERN, 2014.

[111] ATLAS Collaboration. Performance of algorithms that reconstruct missing transverse momentum in $\sqrt{s} = 8$ tev proton–proton collisions in the atlas detector. *Eur. Phys. J. C*, 77(241), 2017. arXiv:1609.09324 [hep-ex].

[112] ATLAS Collaboration. Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ tev using the atlas detector. *Eur. Phys. J. C*, 76(581), 2016. arXiv:1510.03823 [hep-ex].

[113] ATLAS Collaboration. Forward Jet Vertex Tagging: A new technique for the identification and rejection of forward pileup jets. Technical Report ATL-PHYS-PUB-2015-034, CERN, 2015.

[114] ATLAS Collaboration. 2015–2016 TST Systematic and Forward Pileup Suppression in MET. Technical Report JETM-2017-001, CERN, 2017.

[115] ATLAS Collaboration. Missing Transverse Momentum Distribution and Performance in 2016 data. Technical Report JETM-2016-008, CERN, 2016.

[116] Fixed rate random trigger. `https://atlas-tagservices.cern.ch/tagservices/RunBrowser/runBrowserReport/runBrowserReport.php?cn=EF_rd0_filled_NoAlg/`.

[117] ATLAS collaboration. Charged-particle multiplicities in pp interactions at $\sqrt{s} = 900$ GeV measured with the ATLAS detector at the LHC. *Phys. Lett.*, B688:21–42, 2010.

[118] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 0605:026, 2006.

[119] Geant4 collaboration. GEANT4: A Simulation toolkit. *Nucl.Instrum.Meth.*, A506:250–303, 2003.

[120] G. Folger and J.P. Wellisch. String parton models in GEANT4. In *Computing in High Energy and Nuclear Physics, CHEP, La Jolla, 2003*. eConf C0303241 MOMT007 (2003), 2003.

[121] H.W. Bertini. Intranuclear-cascade calculation of the secondary nucleon spectra from nucleon-nucleus interactions in the energy range 340 to 2900 MeV and comparisons with experiment. *Phys. Rev.*, 188:1711–1730, 1969.

[122] M. Blann, B.L. Berman, and T.T. Komoto. Precompound Model Analysis of Photonuclear Reactions. *Phys. Rev.*, C28:2286–2298, 1983.

[123] TileD3PDMaker package. `https://twiki.cern.ch/twiki/bin/viewauth/Atlas/TileD3PD/`.

[124] The ATLAS Track Extrapolation Package. ATL-SOFT-PUB-2007-005.

[125] GoodRunsLists tutorial. `https://twiki.cern.ch/twiki/bin/viewauth/Atlas/GoodRunsListsTutorial/`.

[126] TileTripReader. `https://twiki.cern.ch/twiki/bin/view/Atlas/TileTripReader/`.

[127] ATLAS collaboration. A measurement of the calorimeter response to single hadrons and determination of the jet energy scale uncertainty using LHC Run-1 *pp*-collision data with the ATLAS detector. *Eur. Phys. J. C*, 26(77), 2017.

[128] ATLAS collaboration. Response of the ATLAS calorimeters to single isolated hadrons produced in proton proton collisions at a center of mass energy of $\sqrt{s} = 900$ GeV. Technical report, CERN, 2010. ATLAS-CONF-2010-017.

[129] Pile-up reweighting. `https://svnweb.cern.ch/trac/atlasoff/browser/PhysicsAnalysis/AnalysisCommon/PileupReweighting/`.

[130] ATLAS collaboration. Study of energy response and resolution of the ATLAS barrel calorimeter to hadrons of energies from 20 GeV to 350 GeV. *Nucl. Instrum. Meth.*, A621:134–150, 2010.

[131] ATLAS Liquid Argon HEC collaboration. Performance of the ATLAS hadronic end-cap calorimeter in beam tests. *Nucl. Instrum. Meth.*, A482:94–124, 2002.

[132] ATLAS collaboration. Single hadron response measurement and calorimeter jet energy scale uncertainty with the ATLAS detector at the LHC. *Eur. Phys. J. C*, 73(arXiv:1203.1302. CERN-PH-EP-2012-005):2305, 2012.

[133] Tile Calorimeter Public Results. `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/TileCaloPublicResults/`.

[134] CMS Collaboration. Search for two Higgs bosons in final states containing two photons and two bottom quarks in protonproton collisions at 8 TeV. *Phys. Rev. D*, 94, 2016. arXiv:1603.06896 [hep-ex].

[135] CMS Collaboration. Search for Higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state in pp collisions at $\sqrt{s} = 13$ TeV. *Phys. Lett. B*, 778:101–127, 2018. arXiv:1806.00408 [hep-ex].

[136] ATLAS Collaboration. Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. D*, 94, 2016. arXiv:1804.06174 [hep-ex].

[137] CMS Collaboration. Search for Higgs boson pair production in events with two bottom quarks and two tau leptons in protonproton collisions at $\sqrt{s} = 13$ TeV. *Phys. Lett. B*, 778, 2018. arXiv:1707.02909 [hep-ex].

[138] CMS Collaboration. Search for resonant and nonresonant Higgs boson pair production in the $b\bar{b}\ell\nu\ell\nu$ final state in protonproton collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 01, 2018. arXiv:1708.04188 [hep-ex].

[139] S. Dittmaier S. Dawson and M. Spira. Neutral Higgs-boson pair production at hadron colliders: QCD corrections. *Phys. Rev. D*, 58, 1998. arXiv:9805244 [hep-ph].

[140] M. Spira T. Plehn and P. M. Zerwas. Pair production of neutral Higgs particles in gluongluon collisions. *Nucl. Phys. B*, 479, 1996. arXiv:9603205 [hep-ph].

[141] S.Borowka et al. Higgs Boson Pair Production in Gluon Fusion at Next-to-Leading Order with Full Top-Quark Mass Dependence. *Phys. Rev. Lett.*, 117, 2016. arXiv:1604.06447 [hep-ph].

[142] S.Borowka et al. Full top quark mass dependence in Higgs boson pair production at NLO. *JHEP*, 10, 2016. arXiv:1608.04798 [hep-ph].

[143] M. Grazzini et al. Higgs boson pair production at nnlo with top quark mass effects, 2018. arXiv:1803.02463 [hep-ph].

[144] D. Lopez-Val B. Hespel and E. Vryonidou. Higgs pair production via gluon fusion in the Two-Higgs-Doublet Model. *JHEP*, 09, 2014. arXiv:1407.0281 [hep-ph].

[145] M. Slawinska et al. Phenomenology of the trilinear higgs coupling at proton-proton colliders, 2014. arXiv:1408.5010 [hep-ph].

[146] ATLAS Collaboration. Measurement of Higgs boson production in the diphoton decay channel in *pp* collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector. *Phys. Rev. D*, 90, 2014. arXiv:1408.7084 [hep-ex].

[147] T. Plehn F. Kling and P. Schichtel. Maximizing the significance in Higgs boson pair analyses. *Phys. Rev. D*, 95, 2017. arXiv:1607.07441 [hep-ph].

[148] E. Gross G. Cowan, K. Cranmer and O. Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C*, 71, 2011. arXiv:1007.1727 [physics.data-an].

[149] ATLAS Collaboration. Measurement of the isolated diphoton cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Phys. Rev. D*, 85, 2012. arXiv:1107.0581 [hep-ex].

[150] ATLAS Collaboration. Measurement of the inclusive isolated prompt photon cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Phys. Rev. D*, 83, 2011. arXiv:1012.4389 [hep-ex].

[151] J. E. Gaiser. Charmonium Spectroscopy From Radiative Decays of the J/ψ and ψ', 1982. SLAC-0255.

[152] J. Adelman et al. Supporting note: Selection and performance for the $h \to \gamma\gamma$ and $h \to z\gamma$ analyses, spring 2017, 2017. ATL-COM-PHYS-2017-357.

[153] S. Das. A simple alternative to the Crystal Ball function, 2016. arXiv:1603.08591 [hep-ex].

[154] Belle Collaboration. A detailed test of the CsI(Tℓ) calorimeter for BELLE with photon beams of energy between 20 MeV and 5.4 GeV. *Nucl. Instrum. Meth.*, A441, 2000.

[155] ATLAS Collaboration. Study of heavy flavor quarks produced in association with top quark pairs at $\sqrt{s} = 7$ TeV using the ATLAS detector. *Phys. Rev. D*, 89, 2014. arXiv:1304.6386 [hep-ex].

[156] ATLAS Collaboration. Measurement of the cross-section for W boson production in association with b-jets in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *JHEP*, 06, 2013. arXiv:1302.2929 [hep-ex].

[157] ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC. *Eur. Phys. J. C*, 76, 2016. arXiv:1608.03953 [hep-ex].

[158] ATLAS Collaboration. Performance of the ATLAS Trigger System in 2015. *Eur. Phys. J. C*, 77, 2017. arXiv:1611.09661 [hep-ex].

[159] ATLAS Collaboration. Jet energy scale measurements and their systematic uncertainties in protonproton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. D*, 96, 2017. arXiv:1703.09665 [hep-ex].

[160] ATLAS Collaboration. Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV. Technical Report ATL-PHYS-PUB-2015-015, CERN, 2015.

[161] ATLAS Collaboration. Jet energy measurement and its systematic uncertainty in protonproton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Eur. Phys. J. C*, 75, 2015. arXiv:1406.0076 [hep-ex].

[162] ATLAS Collaboration. Combined search for the Standard Model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Phys. Rev. D*, 86, 2012. arXiv:1207.0319 [hep-ex].

[163] A. L. Read. Presentation of search results: the CL_S technique. *J. Phys. G*, 28, 2002.

[164] ATLAS Collaboration. Procedure for the LHC Higgs boson search combination in summer 2011. Technical Report ATL-PHYS-PUB-2011-011, CERN, 2011.

[165] ATLAS Collaboration. Combination of searches for Higgs boson pairs in pp collisions at 13 TeV with the ATLAS experiment, 2018. ATLAS-CONF-2018-043.

[166] ATLAS Collaboration. Study of the double higgs production channel h($\rightarrow b\bar{b}$) h($\rightarrow \gamma\gamma$) with the atlas experiment at the hl-lhc, 2017. ATL-PHYS-PUB-2017-001.

[167] M. Duehrssen-Debling F. Tackmann, K. Tackmann and P. Francavilla. Simplified template cross sections, Mar 2016. LHCHXSWG-DRAFT-INT-2016-006.

[168] J. R. Andersen et al. Les Houches 2015: Physics at TeV Colliders Standard Model Working Group Report, 2016. arXiv:1605.04692 [hep-ph].

[169] D. E. Soper J. C. Collins and G. F. Sterman. Transverse Momentum Distribution in Drell-Yan Pair and W and Z Boson Production. *Nucl. Phys. B*, 250, 1985.

[170] P. F. Monni F. Bishara, U. Haisch and E. Re. Constraining Light-Quark Yukawa Couplings from Higgs Distributions. *Phys. Rev. Lett.*, 118, 2017. arXiv:1606.09253 [hep-ph].

[171] E. Re K. Hamilton, P. Nason and G. Zanderighi. NNLOPS simulation of Higgs boson production. *JHEP*, 10, 2013. arXiv:1309.0017 [hep-ph].

[172] J. Butterworth et al. PDF4LHC recommendations for LHC Run II. *J. Phys. G*, 43, 2016. arXiv:1510.03865 [hep-ph].

[173] S. Mrenna T. Sjstrand and P. Z. Skands. A brief introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178, 2008. arXiv:0710.3820 [hep-ph].

[174] 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. arXiv:1406.3660 [hep-ph].

[175] C. Anastasiou et al. Higgs Boson Gluon-Fusion Production in QCD at Three Loops. *Phys. Rev. Lett.*, 114, 2015. arXiv:1503.06056 [hep-ph].

[176] C. Anastasiou et al. Mixed QCD-electroweak corrections to Higgs boson production in gluon fusion. *JHEP*, 04, 2009. arXiv:0811.3458 [hep-ph].

[177] C. Anastasiou et al. High precision determination of the gluon fusion Higgs boson cross-section at the LHC. *JHEP*, 05, 2016. arXiv:1602.00695 [hep-ph].

[178] S. Actis et al. NLO electroweak corrections to Higgs boson production at hadron colliders. *Phys. Lett. B*, 670, 2008. arXiv:0809.1301 [hep-ph].

[179] P. Nason. A New method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11, 2004. arXiv:0409146 [hep-ph].

[180] P. Nason S. Frixione and C. Oleari. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *JHEP*, 11, 2007. arXiv:0709.2092 [hep-ph].

[181] C. Oleari S. Alioli, P. Nason and E. Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06, 2010. arXiv:1002.2581 [hep-ph].

[182] P. Nason and C. Oleari. NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG. *JHEP*, 02, 2010. arXiv:0911.5299 [hep-ph].

[183] 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. arXiv:1405.0301 [hep-ph].

[184] H. L. Lai et al. New parton distributions for collider physics. *Phys. Rev. D*, 82, 2010. arXiv:1007.2241 [hep-ph].

[185] A. Banfi et al. Jet-vetoed Higgs cross section in gluon fusion at N^3LO+NNLL with small-R resummation. *JHEP*, 04, 2016. arXiv:1511.02886 [hep-ph].

[186] I. W. Stewart et al. Jet p_T Resummation in Higgs Production at NNLL' + NNLO. *Phys. Rev. D*, 89, 2014. arXiv:1307.1808 [hep-ph].

[187] R. Boughezal et al. Combining resummed Higgs predictions across jet bins. *Phys. Rev. D*, 89, 2014. arXiv:1312.4535 [hep-ph].

[188] X. Chen et al. Precise QCD predictions for the production of Higgs + jet final states. *Phys. Lett. B*, 740, 2015. arXiv:1408.5325 [hep-ph].

[189] X. Chen et al. NNLO QCD corrections to Higgs boson production at large transverse momentum. *JHEP*, 10, 2016. arXiv:1607.08817 [hep-ph].

[190] G. Cullen et al. Automated one-loop calculations with GoSam. *Eur. Phys. J. C*, 72, 2012. arXiv:1111.2034 [hep-ph].

[191] G. Cullen et al. GOSAM-2.0: a tool for automated one-loop calculations within the Standard Model and beyond. *Eur. Phys. J. C*, 74, 2014. arXiv:1404.7096 [hep-ph].

[192] T. Gleisberg et al. Event generation with SHERPA 1.1. *JHEP*, 02, 2009. arXiv:0811.4622 [hep-ph].

[193] S. Schumann and F. Krauss. A Parton shower algorithm based on Catani-Seymour dipole factorisation. *JHEP*, 03, 2008. arXiv:0709.1027 [hep-ph].

[194] N. Greiner et al. Phenomenological analysis of Higgs boson production through gluon fusion in association with jets. *JHEP*, 01, 2016. arXiv:1506.01016 [hep-ph].

[195] M. Schonherr S. Höche, F. Krauss and F. Siegert. A critical appraisal of NLO+PS matching methods. *JHEP*, 09, 2012. arXiv:1111.1220 [hep-ph].

[196] M. Schonherr S. Höche, F. Krauss and F. Siegert. QCD matrix elements + parton showers: The NLO case. *JHEP*, 04, 2013. arXiv:1207.5030 [hep-ph].

[197] M. Buschmann et al. Mass effects in the Higgs-gluon coupling: boosted vs off-shell production. *JHEP*, 02, 2015. arXiv:1410.5806 [hep-ph].

[198] M. Schonherr E. Bothmann and S. Schumann. Reweighting QCD matrix-element and parton-shower calculations. *Eur. Phys. J. C*, 76, 2016. arXiv:1606.08753 [hep-ph].

[199] T. Gleisberg and S. Höche. Comix, a new matrix element generator. *JHEP*, 12, 2008. arXiv:0808.3674 [hep-ph].

[200] R. Kuhn F. Krauss and G. Soff. AMEGIC++ 1.0: A matrix element generator in C++. *JHEP*, 02, 2002. arXiv:0109036 [hep-ph].

[201] S. Höche and M. Schonherr. Uncertainties in next-to-leading order plus parton shower matched simulations of inclusive jet and dijet production. *Phys. Rev. D*, 86, 2012. arXiv:1208.2815 [hep-ph].

[202] F. Krauss S. Höche and M. Schonherr. Uncertainties in MEPS@NLO calculations of h+jets. *Phys. Rev. D*, 90, 2014. arXiv:1401.7971 [hep-ph].

[203] R. Frederix et al. Heavy-quark mass effects in Higgs plus jets production. *JHEP*, 08, 2016. arXiv:1604.03017 [hep-ph].

[204] R. Frederix and S. Frixione. Merging meets matching in MC@NLO. *JHEP*, 12, 2012. arXiv:1209.6215 [hep-ph].

[205] S. Dulat et al. New parton distribution functions from a global analysis of quantum chromodynamics. *Phys. Rev. D*, 93, 2016. arXiv:1506.07443 [hep-ph].

[206] G. D'Agostini. A Multidimensional unfolding method based on Bayes' theorem. *Nucl. Instrum. Meth. A*, 362:487–498, 1995.

[207] A. Hoecker and V. Kartvelishvili. SVD approach to data unfolding. *Nucl. Instrum. Meth. A*, 372:469–481, 1996.

[208] B. Malaescu. An Iterative, dynamically stabilized method of data unfolding, 2009.

[209] A. Hasib et al. Supporting document : Probing higgs production modes and couplings with the $h \to \gamma\gamma$ channel with the run 2 of lhc in the atlas experiment, 2016. ATL-COM-PHYS-2016-1784.

[210] F. Bernlochner et al. Measurement of fiducial and differential cross sections in the $h \to \gamma\gamma$ decay channel with 36.1 fb^{-1} of 13 tev proton-proton collision data with the atlas detector, 2016. ATL-COM-PHYS-2017-145.

[211] ATLAS and CMS Collaborations. Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments. *Phys. Rev. Lett.*, 114, 2015. arXiv:1503.07589 [hep-ex].

[212] X. Liu and F. Petriello. Reducing theoretical uncertainties for exclusive Higgs-boson plus one-jet production at the LHC. *Phys. Rev. D*, 87, 2013. arXiv:1303.4405 [hep-ph].

[213] I. W. Stewart and F. J. Tackmann. Theory uncertainties for Higgs and other searches using jet bins. *Phys. Rev. D*, 85, 2012. arXiv:1107.2117 [hep-ph].

[214] S. Gangal and F. J. Tackmann. Next-to-leading-order uncertainties in Higgs + 2 jets from gluon fusion. *Phys. Rev. D*, 87, 2013. arXiv:1302.5437 [hep-ph].

[215] Theoretical Predictions and systematics. `https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/HiggsPropertiesTheorySystematics`.

[216] ATLAS Collaboration. Measurements of fiducial cross-sections for $t\bar{t}$ production with one or two additional b-jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector. *Eur. Phys. J. C*, 76, 2016. arXiv:1508.06868 [hep-ex].

[217] ATLAS Collaboration. Measurement of differential production cross-sections for a Z boson in association with b-jets in 7 TeV proton-proton collisions with the ATLAS detector. *JHEP*, 10, 2014. arXiv:1407.3643 [hep-ex].

[218] ATLAS Collaboration. Measurement of the cross-section for W boson production in association with b-jets in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *JHEP*, 06, 2013. arXiv:1302.2929 [hep-ex].

[219] D. de Florian and M. Grazzini. Higgs production at the LHC: updated cross sections at $\sqrt{s} = 8$ TeV. *Phys. Lett. B*, 718, 2012. arXiv:1206.4133 [hep-ph].

[220] G. Klamke and D. Zeppenfeld. Higgs plus two jet production via gluon fusion as a signal at the CERN LHC. *JHEP*, 04, 2007. arXiv:hep-ph/0703202.

[221] K. Arnold J. R. Andersen and D. Zeppenfeld. Azimuthal angle correlations for Higgs boson plus multi-jet events. *JHEP*, 06, 2010. arXiv:1001.3822 [hep-ph].

[222] M. Jankowiak M. J. Dolan, P. Harris and M. Spannowsky. Constraining CP-violating Higgs Sectors at the LHC using gluon fusion. *Phys. Rev. D*, 90, 2014. arXiv:1406.3322 [hep-ph].

[223] D. L. Rainwater T. Plehn and D. Zeppenfeld. Determining the structure of Higgs couplings at the LHC. *Phys. Rev. Lett.*, 88, 2002. arXiv:hep-ph/0105325.

[224] ATLAS Collaboration. Measurement of the Higgs boson coupling properties in the $H \to ZZ^* \to 4\ell$ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector. *JHEP*, 03, 2018. arXiv:1712.02304 [hep-ex].

[225] CMS Collaboration. Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 2018. arXiv:1804.02716 [hep-ex].