UNIVERSITÉ SORBONNE PARIS CITÉ

Thèse préparée

à l'UNIVERSITÉ PARIS DIDEROT

École doctorale STEP'UP - ED N°560

Laboratoire de Physique Nucléaire et de Hautes Energies - UMR 7585

Search for New Physics with Two Photons in the Final State with the ATLAS Detector

par

Yee Chinn Yap

presentée et soutenue publiquement le

23 janvier 2017

Thèse de doctorat de Physique de l'Univers

dirigée par Lydia Roos

devant un jury composé de :

Beate Heinemann Professeure (DESY et University of Freiburg)	Rapporteuse
Paris Sphicas Senior scientist (CERN) et professeur (National and Kapodistrian University of Athens)	Rapporteur
Abdelhak Djouadi Directeur de recherche (LPT, Orsay)	Membre
Sotiris Loucatos Professeur (APC, Paris)	Membre
Lydia Roos Directrice de recherche (LPNHE, Paris)	Directrice de thèse

Abstract

This thesis reports on the search for new physics in the diphoton decay channel with the proton-proton collision data collected by ATLAS at a centre-of-mass energy of $\sqrt{s} = 8$ TeV in 2012 and $\sqrt{s} = 13$ TeV in 2015 and 2016.

A feasibility study of the search for a pseudoscalar A decaying to a Z boson and a 125 GeV Higgs boson in the context of an extended Higgs sector, namedly the two-Higgs-doublet models, is presented. The search is performed with a final state of two jets and two photons using 20.3 fb^{-1} of data at $\sqrt{s} = 8$ TeV. The expected sensitivity is found to be competitive with the analysis with a final state of two electrons or muons and two τ leptons, but less sensitive to the other searches with the Higgs decaying to a pair of b-quarks.

Search for high mass resonances decaying to two photons at $\sqrt{s} = 13$ TeV is also presented. The analysed dataset corresponds to an integrated luminosity of 3.2 fb^{-1} in 2015 and 12.2 fb^{-1} in 2016. Two searches are performed: one optimised for a generic spin-0 particle, and the other for a spin-2 particle using Randall-Sundrum graviton model as benchmark. Particular attention is paid to the understanding and the parameterisation of the line shape of a large-width resonance. In the search using the collision data recorded in 2015, a moderate excess was observed at a diphoton invariant mass around 750 GeV with a best-fit width of around 45 GeV, corresponding to a local significances of 3.8-3.9 standard deviations, depending on the spin of the resonance being searched for. The global significances are estimated to be 2.1 standard deviations. The search is updated with 2016 data and no significant excess is observed. Limits on the production cross-section times branching ratio to two photons for the two resonance types are reported: for a spin-0 resonance, the fiducial cross-section ranges from 12 fb (51 fb) at 200 GeV to 0.2 fb (0.3 fb) at 2400 GeV, with the narrow (10% of the mass) width hypothesis.

A part of this document is also devoted to the study of photons that convert into electron-positron pairs before entering the calorimeter using their electromagnetic shower information. The conversion study has two objectives: improving the identification of photon conversion at high pile-up rate and providing an orthogonal classification method to assess the performance of the standard track-based algorithm.

Résumé

Le sujet de cette thèse est la recherche de signaux de physique au-delà du modèle standard, dans le canal de désintégration en deux photons, dans les collisions proton-proton enregistrées par le détecteur ATLAS à des énergies dans le centre de masse de $\sqrt{s} = 8$ TeV en 2012 et $\sqrt{s} = 13$ TeV en 2015 et 2016.

Une étude de faisabilité de la recherche d'un boson pseudo-scalaire A se désintégrant en un boson Z et un boson de Higgs d'une masse de 125 GeV est présentée. Cette étude est effectuée dans le contexte de modèles prédisant deux doublets de Higgs. L'état final recherché est constitué de deux jets hadroniques et deux photons et les données utilisées sont celles enregistrées en 2012 à $\sqrt{s} = 8$ TeV, correspondant à une luminosité intégrée de 20.3 fb^{-1}. La sensibilité attendue est comparable à celle atteinte avec une autre recherche effectuée dans la collaboration ATLAS, basée sur l'état final constitué d'une paire d'électrons ou muons et d'une paire de leptons τ. En revanche, elle est inférieure à celle des études dans lesquelles le Higgs est recherché dans sa désintégration en une paire de quarks beaux.

La recherche de nouvelles résonances de haute masse se désintégrant en deux photons dans les collisions à 13 TeV est également présentée. Les échantillons analysés correspondent à des luminosités intégrées de 3.2 fb^{-1} en 2015 et 12.2 fb^{-1} en 2016. Deux études sont développées : l'une est optimisée pour la recherche, aussi générale que possible, d'une particule de spin 0; la seconde cible une particule de spin 2 telle que le graviton prédit par le modèle de Randall-Sundrum. Une attention particulière est portée à la compréhension et la paramétrisation de la forme de la résonance dans l'hypothèse où celle-ci aurait une largeur naturelle non négligeable. L'analyse des données enregistrées en 2015 conclut à l'observation d'un léger excès autour d'une masse de 750 GeV, d'une largeur d'environ 45 GeV, et dont la significativité statistique locale est estimée à 3,9 et 3,8 écarts standard selon le spin de la résonance recherchée. La significativité globale est estimée à 2,1 écarts standard. L'analyse des données enregistrées en 2016 ne mène à l'observation d'aucun excès significatif. Des limites sur le produit de la section efficace de production d'une résonance par le taux de désintégration en deux photons sont établies : pour une résonance de spin nul, elles vont de 12 fb (51 fb) à une masse de 200 GeV à 0,2 fb (0,3 fb) à 2400 GeV, dans l'hypothèse d'une largeur naturelle de 4 MeV (10% de la masse).

Une étude de la reconnaissance des photons se convertissant en une paire électron-positon avant d'interagir avec les calorimètres, basée sur la forme de la gerbe électromagnétique détectée est également documentée. Cette étude a deux objectifs : d'une part, améliorer l'identification des photons convertis dans le contexte de taux d'empilement élevés ; d'autre part, fournir une méthode d'identification basée sur des informations indépendantes (calorimétriques) de celles utilisées par la méthode standard (trajectoires des particules chargées) afin d'evaluer les performances de cette dernière.

“Moderation is a fatal thing. Nothing succeeds like excess.”

— Oscar Wilde

Acknowledgements

As cliché as it may sound, this thesis would not have been possible without my supervisor Lydia Roos. These past three years have been a phenomenal journey, and I thank you for showing me so much patience and kindness. I have received so much help on various issues, from battling with the bureaucracy, to defending my contributions and securing opportunities for myself. Thank you for believing in my potential, for pushing me because you believe I am capable of doing better, and for encouraging me when I am in despair. I have learned a great deal from you, and best of all, we have learned so much together. We make an incredible team. Thank you for always being there for me and for being so much more than just a supervisor. You have exceeded every expectation.

This thesis has benefited from the rigorous reviews of Bertrand Laforge, Bogdan Malaescu, Frédéric Derue, José Ocariz, Martin White and Sandrine Laplace, besides the admirable work of Lydia. Thank you all, I know it must not have been easy to read my often sloppy writing. Your careful reading and practical suggestions are very much appreciated.

I also thank all the members of my thesis jury, especially the *rapporteurs*, for accepting the responsibilities and doing a great job.

I have enjoyed my time at LPNHE, mostly thanks to the wondrous support I get from the ATLAS group. In particular, I have interacted more with Bertrand, Bogdan, Dimitris Varouchas, Frédéric, Irena Nikolic-Audit, José, Paolo Francavilla, Renjie Wang and Sandrine, and they have been the most lovely colleagues, always ready to extend me a helping hand. Thank you Bogdan for always being so nice and I really appreciate our discussions about statistics and jets. I am indebted to Bertrand, José and Sandrine for educating me on Higgs and photon physics. Thanks to many of you for the French culture and language immersion during lunchtime.

The LPNHE ATLAS students (Carlo Pandini, Álvaro López, Dilia Portillo, Stefano Manzoni, Audrey Ducourthial, Changqiao Li, Ahmed Tarek, Ilaria Luise, Louis D'eramo, Robert Hankache) are the best bunch with whom I have gone through this journey together. To those who have yet to finish, all the best! It has been marvellous having my BFF Dilia and also Stefano as officemates. Thanks for livening up the office with your colourful personalities. Special thanks also go to many of the students at the lab for the after-work beers and *soirées*.

Working in the diphoton resonance analysis team has been really intense, especially after the unblinding when the initial team of about ten increased manyfold. Nonetheless, I have had a terrific time working in the team. We have been lucky to have excellent conveners over the course of the past year: Marco Delmastro, Liron Barak, Leonardo Carminati, Tancredi Carli, Elisabeth Petit and German Carrillo-Montoya. Within the HFITTER team, Nicolas Berger, Liron and Kirill Grevtsov are wonderful to work with. Nicolas has taught me so much about statistics and fitting, while Liron has been incredibly supportive and efficient in coordinating our efforts. Chris Meyer, Simone Mazza and Hongtao Yang also prove to be valuable collaborators. While working on the scalar singlet signal model, Larry Lee and Martin have been the best teammates one could ask for. I am also thankful for the occasional QFT lessons from Martin.

More broadly at CERN, I have the privilege to work along or discuss with many exceptional physicists. My discussions with Andreas Hoecker and Bruno Mansoulié have always been enlightening. I am very grateful to Bruno Lenzi for guiding my work on photon conversion like a second supervisor and for teaching me so much about photon reconstruction.

I am glad to have a worthy *comité de suive de thèse*. Zhiqing Zhang and Arnaud Robert, thank you for taking an interest and concern in my welfare and thesis progression.

Last but not least, I am grateful to my parents for being so supportive of every decision I made in life. I am sorry my work has kept me from spending much time at home. Thank you for being my biggest supporters.

Contents

Introduction

The Standard Model of particle physics encapsulates our best understanding of how the fundamental particles and three of the four known forces are related. For decades, the Higgs boson remained the last missing piece of the Standard Model. Its discovery was one of the key objectives of the Large Hadron Collider (LHC) at CERN. LHC is the world's largest and most powerful particle collider. It was constructed between 1998 and 2008, with operation commenced in November 2009. The period between 2009 and early 2013 is known as Run 1. LHC delivered proton-proton collision data samples at centre-of mass energies of 7 TeV and 8 TeV in 2011 and 2012 respectively. On the 4th July 2012, the discovery of a new particle in the mass region around 125-126 GeV consistent with the Standard Model Higgs boson was announced by the ATLAS and CMS experiments. The discovery was the most important legacy of the Run 1. The decay channel of the Higgs to two photons played a significant role thanks to the fully reconstructed final state and an excellent mass resolution. Following Run 1, the machine was shut down for a two-year period for planned maintenance, known as Long Shutdown 1 (LS1). The LHC restarted in 2015 with an increase in the collision energy, from 8 TeV to 13 TeV. The period from 2015 is known as Run 2.

Although the Standard Model has been successful in describing the phenomena within its domain, it is still an incomplete theory. It does not explain gravity nor provide a candidate for dark matter. Another example of its inadequacy is the lack of explanation of the matter-antimatter asymmetry in the universe. While the SM allows for charge conjugation and parity violation in the quark sector, it is insufficient by many orders of magnitude to account for the primordial matter-antimatter asymmetry of the universe. One other shortcoming is what is known as the hierarchy problem, where there exists a vast discrepancy between aspects of the weak nuclear force and gravity. Thus, the search for new physics is the central theme of this thesis performed within the ATLAS collaboration. ATLAS is one of the four major experiments at the LHC. It is a general-purpose particle physics experiment and, together with CMS, is designed to exploit the full discovery potential and the huge range of physics opportunities that the LHC provides.

The present thesis work started in December 2013, during the LS1 period. Photon object is the focus of the research activity in this thesis. A study on the conversion of photon into electron-positron pair was carried out in the context of the technical task required to qualify as an ATLAS author. The difference in shape of the electromagnetic shower created by the unconverted photon or the converted photon in the calorimeter is studied to better identify the photons that have converted. The study consisted in investigating the use of electromagnetic shower variables to either improve the present conversion classification algorithm or to evaluate its performance on data. This is in light of the expected degradation of performance of the present algorithm with the number of collisions per bunch crossing, which is expected to increase towards the end of Run 2. In parallel, a feasibility study of an $A \to Zh$ search with the h boson decaying to $\gamma\gamma$ and the Z boson decaying to hadrons was also performed. The pseudoscalar A is predicted by models with extended Higgs sector. In this search, h is assumed to be the 125 GeV boson discovered in 2012. The study aimed at assessing the sensitivity of this search compared to other search channels in ATLAS.

From the summer of 2015, I moved from the $A \to Zh$ study to a search for high mass diphoton resonances. The increase in the LHC collision energy boosts the sensitivity to new physics at high mass scale. The diphoton final state provides a clean experimental signature for a generic search for new resonances. Unlike the previous study, the analysis was performed in a team. I was first responsible to extend the search to large-width resonances by studying the resonance shape under several model assumptions. I was also in charge of producing the statistical results and investigating analysis improvement as well as several checks on the signal and background modelling.

The search using 13 TeV data recorded in 2015 was first presented in December 2015 at a seminar at CERN (also known as the End Of the Year Event), then updated at the *51st Rencontres de Moriond on Electroweak Interactions and Unified Theories* at La Thuile in March 2016 and subsequently published in the Journal of High Energy Physics. The 2015 result triggered a great interest in the scientific community. 2016 was a productive year for the LHC, and at the time of the *38th International Conference on High Energy Physics* (ICHEP) conference at Chicago in August 2016, the 2016 data available for analysis was almost four times that of 2015. The high mass search was updated using 2016 data for the ICHEP conference.

In the second half of 2015, I have also taken shifts in the ATLAS control room manning the calorimeter and forward detectors desk.

This thesis is organised as follows: Chapter 1 gives an overview of the Standard Model and a brief introduction to some beyond Standard Model theories relevant to the new physics searches carried out in the thesis. A general description of the LHC and the ATLAS detector is outlined in Chapter 2. The reconstruction of photon candidates and the techniques to identify photons are discussed in Chapter 3. Chapter 4 chronicles my work on the photon conversion classification and measurement study. Moving on to the search analyses, the statistical method used to interpret the data is presented in Chapter 5. The feasibility study of the $A \to Zh$ search in the $jj\gamma\gamma$ channel is reported in Chapter 6. The search for high mass diphoton resonances is described over the next three chapters. Chapter 7 provides a theory-oriented study of the high-mass, large-width resonance line shape under several model assumptions. The analysis and the results from the search using 2015 data are reported in Chapter 8. Finally, the updated analysis and the results using the first half of 2016 data are summarised in Chapter 9.

Chapter 1

Theoretical Context

The theoretical context of the present thesis work is outlined in this chapter. The chapter starts with a brief description of the Standard Model (SM), which is currently our best available description of the subatomic world, with a focus on the electroweak sector and the Higgs mechanism. Non-resonant SM diphoton production is also discussed as it constitutes the main background of the searches performed in this thesis. Some beyond the SM (BSM) theories are also introduced: two-Higgs-doublet models (2HDMs) and models that include a high mass diphoton resonance.

1.1 The Standard Model

The Standard Model of particle physics was developed in the early 1970s. It provides a theoretical description of the elementary particles that make up matter, and three of the four forces that govern their interaction. In the SM, the fundamental particles are divided into matter particles and force particles. The constituents of matter have spin 1/2, hence they are fermions. The force particles on the other hand are mediator of the fundamental forces and have spin 1. A third category of particle is the Higgs boson which has spin 0. Figure 1.1 shows the particles in the SM and their properties (mass, charge and spin).

The SM includes the electromagnetic, weak and strong interactions, and all their carrier particles, but does not incorporate gravity. Practically, this has a negligible effect on any particle physics experiment since the gravitational force between particles is so much weaker than even the weak force. If one attempts to describe gravity with a quantum field theory, the theory appears to be not renormalisable. Although not yet found, the corresponding force-carrying particle of gravity, called the graviton, would have spin 2.

The SM is the most successful theory of particle physics to date, with almost all of its predictions confirmed experimentally. It predicted the existence of W and Z bosons, gluon, and top and charm quarks before these particles were observed. The last missing piece, the Higgs boson, was observed by the ATLAS and CMS experiments, and announced on 4 July 2012. The SM also provides an excellent description of the quark model. Moreover, the violation of the charge-parity (CP) symmetry in the weak interaction is well explained by the SM.

The SM is a gauge theory, based on the symmetry groups $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$. The strong force is based on the $SU(3)_C$ group, and is described by Quantum Chromodynamics (QCD). The electroweak theory which unifies the electromagnetic and weak interactions relies on the $SU(2)_L$ and $U(1)_Y$ groups. The subscripts C and Y refers to the colour and hypercharge quantum numbers respectively, and L indicates coupling to only left-handed fields.

Local gauge invariance is required for the renormalisability and unitarity of the SM [1]. The SM Lagrangian is invariant under local transformations of the gauge group and can be written as a sum of QCD and EW terms:

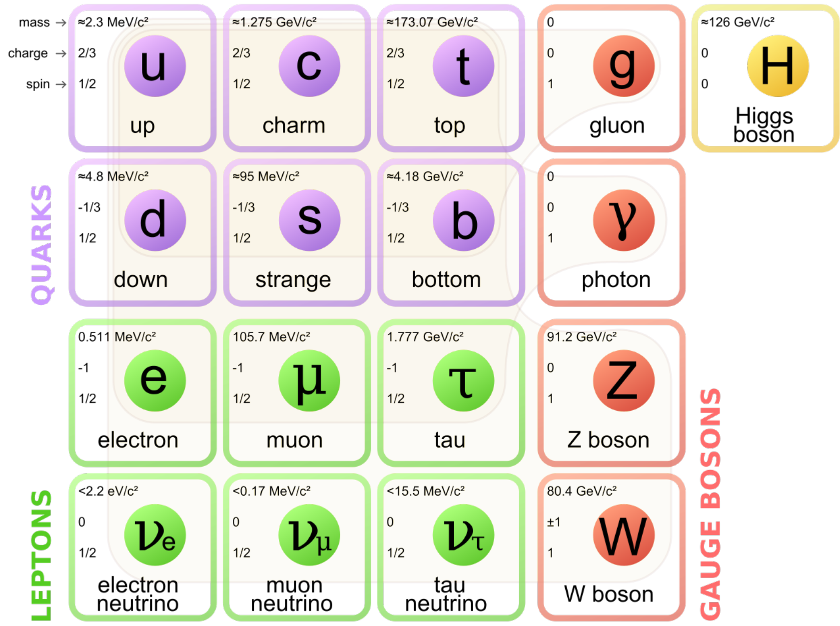


Figure 1.1: Elementary particles in the Standard Model. Mass, charge and spin of the quarks, leptons, gauge bosons and Higgs boson are also shown in the figure.

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{EW}}. \tag{1.1}$$

1.1.1 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is a non-Abelian gauge theory of strong interactions that describes interactions between quarks and gluons. Quarks are electrically charged and can exist in one of three *colour* states (conventionally called red, green and blue). Colour is the charge of the strong interaction, and can be thought of as a QCD analogue of the electric charge. Mesons and baryons are colour-singlet combinations of $q\bar{q}$ and qqq respectively, where q represents quarks and $\bar{q}$ anti-quarks, since only colourless objects can exist in nature as free particles, a phenomenon called *colour confinement*. In the SM, there are six flavours of quarks: up, down, strange, charm, top and bottom.

The Lagrangian of QCD can be written as [2]

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4} G_a^{\mu\nu} G^a_{\mu\nu} + \sum_f \bar{q}_f (i(\gamma^\mu D_\mu) - m_f) q_f, \tag{1.2}$$

where q_f is a quark field of flavour f and γ^μ are the Dirac matrices.

The covariant derivative D_μ is

$$D_\mu = \partial_\mu + i g_s \frac{\lambda_a}{2} G^a_\mu, \tag{1.3}$$

where λ_a are the Gell-Mann matrices and $G^a_{\mu\nu}$ is the field strength tensor of the gluon fields G^a_μ given by

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f^{abc} G^b_\mu G^c_\nu, \tag{1.4}$$

with g_s being the strong coupling constant and $f^{abc} (a, b, c = 1, 2, ..., 8)$ the structure constants of $SU(3)_C$.

1.1.2 Electroweak Theory

The electroweak theory (also known as the Glashow-Weinberg-Salam theory) [3–5] is a unified description of the electromagnetism and weak interactions. The gauge group of the electroweak theory can be written as $SU(2)_L \otimes U(1)_Y$, where the $SU(2)_L$ gauge transformation acts on weak isospin fields and the $U(1)_Y$ acts on weak hypercharge fields.

The Lagrangian for the electroweak interactions can be divided into four parts:

$$\mathcal{L}_{\text{EW}} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_f + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}. \tag{1.5}$$

The first term is the kinetic term for the gauge boson fields [6]:

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4} W^a_{\mu\nu} W^{\mu\nu}_a - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \tag{1.6}$$

where $W^a_{\mu\nu}(a = 1, 2, 3)$ and $B_{\mu\nu}$ are the field strength tensors for the weak isospin and weak hypercharge fields, W^a_μ and B_μ:

$$W^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu - g_W \epsilon^{abc} W^b_\mu W^c_\nu, \tag{1.7}$$

and

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu. \tag{1.8}$$

with g_W being the coupling constant of the weak interaction and ϵ^{abc} is the permutation symbol.

The physical weak boson fields $W^\pm_\mu$ and Z_μ and the photon field A_μ are linear combinations of the W^a_μ and B_μ gauge fields:

$$W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \mp i W^2_\mu), \tag{1.9}$$

$$Z_\mu = W^3_\mu \cos\theta_W - B_\mu \sin\theta_W, \tag{1.10}$$

$$A_\mu = W^3_\mu \sin\theta_W + B_\mu \cos\theta_W, \tag{1.11}$$

where θ_W is the weak mixing (or Weinberg) angle defined as

$$\sin\theta_W = \frac{g_Y}{\sqrt{g^2_W + g^2_Y}}, \tag{1.12}$$

with g_Y being the $U(1)_Y$ coupling constant.

The second term of the electroweak Lagrangian is the kinetic term for the SM fermions:

$$\begin{aligned} \mathcal{L}_f =& (\bar{u}_{iL}, \bar{d}_{iL}) i\gamma^\mu D_\mu \begin{pmatrix} u_{iL} \\ d_{iL} \end{pmatrix} + (\bar{\nu}_{iL}, \bar{l}_{iL}) i\gamma^\mu D_\mu \begin{pmatrix} \nu_{iL} \\ l_{iL} \end{pmatrix} + \\ & \bar{u}_{iR} i\gamma^\mu D_\mu u_{iR} + \bar{d}_{iR} i\gamma^\mu D_\mu d_{iR} + \bar{l}_{iR} i\gamma^\mu D_\mu l_{iR} + \text{h.c.} \end{aligned} \tag{1.13}$$

where the subscript i runs over the generation of fermions, $\begin{pmatrix} u_{iL} \\ d_{iL} \end{pmatrix}$ and $\begin{pmatrix} \nu_{iL} \\ l_{iL} \end{pmatrix}$ are $SU(2)_L$ doublets of quark and lepton fields, and u_{iR}, d_{iR} and l_{iR} are right-handed singlets of up- and down-type quark and charged lepton fields. In the electroweak theory, the left-handed fields are $SU(2)_L$ doublets, while their right-handed partners transform as $SU(2)_L$ singlets. Since the left-handed and the right-handed fermions form different types of multiplets, the weak interaction does not respect the parity or the charge-conjugation symmetries. The left-handed and the right-handed fermions also have different $U(1)_Y$ hypercharges, which is needed to give them similar electric charges $Q = Y/2 + I_3$ where Y is the hypercharge and I_3 is the weak isospin.

For instance, left-handed up-type quarks have isospin of 1/2 and hypercharge 1/3, consequently an electric charge of 2/3, while right-handed up-type quarks have isospin 0 and hypercharge 4/3 hence an electric charge of 2/3.

The gauge covariant derivative D_μ is given by

$$D_\mu = \partial_\mu + ig_W \frac{\sigma_a}{2} W^a_\mu + ig_Y \frac{Y}{2} B_\mu, \tag{1.14}$$

where σ_a are the Pauli matrices.

The Lagrangian terms introduced so far described the free propagation and interaction of the fermion and gauge fields. However, there is no mass term, and therefore the weak gauge boson and fermions are massless according to the above Lagrangian. Indeed, mass terms for the gauge bosons like $\frac{1}{2}m^2 B_\mu B^\mu$ are not gauge invariant. It is also impossible to write mass terms for the fermions since the left-handed and right-handed fermions transform differently under $SU(2)_L$.

The two additional Lagrangian terms $\mathcal{L}_{\text{Higgs}}$ and $\mathcal{L}_{\text{Yukawa}}$ are needed to give a mass to fermions and weak bosons. This is described in the next section.

1.1.3 The Higgs Mechanism

In order to generate masses for the weak bosons while respecting gauge invariance, the $SU(2)_L \otimes U(1)_Y$ gauge symmetry needs to be spontaneously broken. The mechanism of the EW symmetry breaking was developed by three groups in 1964: Robert Brout and François Englert [7], Peter Higgs [8] and Gerald Guralnik, Carl Richard Hagen and Tom Kibble [9]. It is popularly known as the Higgs mechanism.

One can introduce a $SU(2)_L$ doublet of complex scalar fields

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}, \tag{1.15}$$

and add the Lagrangian density terms

$$\mathcal{L} = (D^\mu \Phi)^\dagger (D_\mu \Phi) - V(\Phi), \tag{1.16}$$

where the covariant derivative D^μ is given in Eq. 1.14.

The potential $V(\Phi)$ is defined as

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda(\Phi^\dagger \Phi)^2, \tag{1.17}$$

with $\mu^2 < 0$ and $\lambda > 0$ corresponding to a potential with a 'Mexican hat' shape, as shown in Figure 1.2. The minimum of the potential is reached for complex field values that lie in a circle around the origin, with radius $v = \sqrt{-\mu^2/\lambda}$, which is the vacuum expectation value of the scalar field.

Generally, the minimum of the potential is chosen to be along the real part of the lower component of Φ:

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{1.18}$$

to ensure that the photon remains massless. In choosing this minimum, the underlying symmetry is lost, i.e. the electroweak symmetry is broken.

Writing the scalar field as an expansion around the vacuum expectation value in the unitary gauge:

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h \end{pmatrix}, \tag{1.19}$$

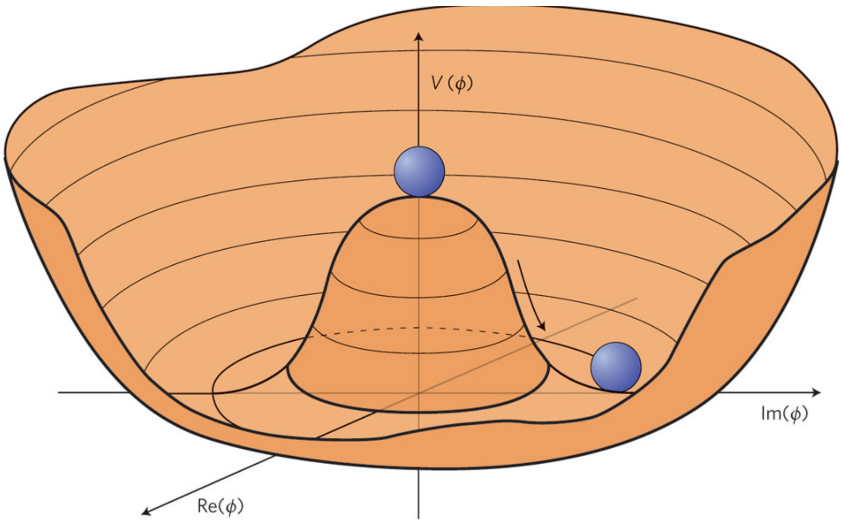


Figure 1.2: Shape of the Higgs potential for $\mu^2 < 0$.

one can expand the Lagrangian in Eq. 1.16 near this minimum. Expanding the first term in the Lagrangian, one obtains [6]

$$\begin{aligned}(D^\mu\Phi)^\dagger(D_\mu\Phi) &= \left|(\partial_\mu + ig_W\frac{\sigma_a}{2}W_\mu^a + ig_Y\frac{Y}{2}B_\mu)\frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v+h\end{pmatrix}\right|^2 \\ &= \frac{1}{2}\partial_\mu h\partial^\mu h + \frac{(v+h)^2}{8}\left(2g_W^2W_\mu^-W^{+\mu} + (g_W^2+g_Y^2)Z_\mu Z^\mu\right).\end{aligned} \tag{1.20}$$

From the terms $\frac{1}{4}v^2g_W^2W_\mu^-W^{+\mu}$ and $\frac{1}{8}v^2(g_W^2+g_Y^2)Z_\mu Z^\mu$, the masses of W and Z bosons can be identified as

$$m_{W^\pm} = \frac{vg_W}{2}, \tag{1.21}$$

and

$$m_Z = \frac{v\sqrt{g_W^2+g_Y^2}}{2} = \frac{vg_W}{2\cos\theta_W}. \tag{1.22}$$

The photon remains massless. The vacuum expectation value v can be obtained from the measurement of the muon lifetime [10]:

$$v = \frac{2m_W}{g_W} = \frac{1}{\sqrt{\sqrt{2}G_F}} = 246 \text{ GeV}, \tag{1.23}$$

where G_F is the Fermi coupling constant.

The expanded Lagrangian from Eq. 1.16 is

$$\mathcal{L}_{\text{Higgs}} = \frac{1}{2}\partial_\mu h\partial^\mu h + \frac{(v+h)^2}{8}\left(2g_W^2W_\mu^-W^{+\mu} + (g_W^2+g_Y^2)Z_\mu Z^\mu\right) + \frac{\mu^4}{4\lambda} + \mu^2h^2 + \lambda vh^3 + \frac{\lambda}{4}h^4. \tag{1.24}$$

Further expanding the terms that describe the coupling of the Higgs boson with the W and Z:

$$\frac{1}{2}hvg_W^2W_\mu^-W^{+\mu} = \frac{2m_W^2}{v}hW_\mu^-W^{+\mu}, \tag{1.25}$$

$$\frac{1}{4}hv(g_W^2+g_Y^2)Z_\mu Z^\mu = \frac{m_Z^2}{v}hZ_\mu Z^\mu, \tag{1.26}$$

one recognises that the coupling of vector boson is proportional to the square of the boson mass.

After the symmetry breaking, three of the four degrees of freedom in the Higgs field are absorbed to give the $W^{\pm}$ and Z bosons their longitudinal components. The remaining degree of freedom becomes the Higgs boson—a new scalar particle. The mass of the Higgs boson appeared in the above Lagrangian and is $\sqrt{2}|\mu^2| = \sqrt{2\lambda}v$. However, since λ is a free parameter of the theory, the Higgs boson mass is not predicted.

One can also introduce fermion masses by coupling the fermion fields to the scalar field. This is achieved with gauge-invariant Yukawa terms [11]:

$$\begin{aligned}\mathcal{L}_{\text{Yukawa}} = -\sum_{\text{families}} \Big(&\Gamma^u_{ij}(\bar{u}_{iL}, \bar{d}_{iL})\tilde{\Phi}u_{jR} + \Gamma^d_{ij}(\bar{u}_{iL}, \bar{d}_{iL})\Phi d_{jR} + \Gamma^l_{ij}(\bar{\nu}_{iL}, \bar{l}_{iL})\Phi l_{jR} \\ &+ \Gamma^\nu_{ij}(\bar{\nu}_{iL}, \bar{l}_{iL})\tilde{\Phi}\nu_{jR} + \text{h.c.}\Big),\end{aligned} \tag{1.27}$$

where the matrices Γ^f_{ij} describe the so-called Yukawa couplings between the Higgs doublet and the fermions. As the neutrino has no right-handed counterpart in the SM, it can not acquire a mass term through Yukawa coupling. Substituting $\Phi = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v+h \end{pmatrix}$ and $\tilde{\Phi} = \frac{1}{\sqrt{2}}\begin{pmatrix} v+h \\ 0 \end{pmatrix}$, one obtains the Lagrangian terms after symmetry breaking:

$$\begin{aligned}\mathcal{L}_{\text{Yukawa}} = -\sum_{\text{families}} \Bigg(&\frac{\Gamma^u_{ij}v}{\sqrt{2}}(\bar{u}_{iL}, \bar{d}_{iL})(1+\frac{h}{v})u_{jR} + \frac{\Gamma^d_{ij}v}{\sqrt{2}}(\bar{u}_{iL}, \bar{d}_{iL})(1+\frac{h}{v})d_{jR} \\ &+ \frac{\Gamma^l_{ij}v}{\sqrt{2}}(\bar{\nu}_{iL}, \bar{l}_{iL})(1+\frac{h}{v})l_{jR} + \text{h.c.}\Bigg).\end{aligned} \tag{1.28}$$

At this point, the Yukawa couplings have not been specified. In order to generate the correct mass for every fermion, one identifies the coupling of each fermion to the Higgs as being proportional to its mass, i.e. $\frac{\Gamma^f_{ij}v}{\sqrt{2}} = m_f$. Figure 1.3 shows a parameterisation of the (square root of the) coupling to fermions (bosons) as a function of the mass, from a combined measurement in the ATLAS and CMS experiments [12]. The parameterisation is chosen so that the behaviour expected in the SM is linear with the same slope for fermions and bosons.

1.1.4 Higgs Production and Decay at the LHC

In the SM, the main production mechanisms for Higgs particles at hadron colliders are: the associated production with W/Z boson, the weak vector boson fusion processes, the gluon-gluon fusion mechanism and the associated Higgs production with top quark. This reflects the fact that the Higgs boson couples preferentially to the heavy particles, that is the massive W and Z vector bosons, the top quark and to a lesser extent, the bottom quark. Of these, the dominant production mode at the LHC is the gluon-gluon fusion, followed by the vector boson fusion. This is due to the much higher gluon-gluon luminosity even though the process is mediated by triangular loops of heavy quarks. The Feynman diagrams of these processes are displayed in Figure 1.4.

The properties of the SM Higgs particle depend on its mass m_H. Since the Higgs couplings to gauge bosons and fermions grow with the masses of the particles, the Higgs boson will tend to decay into the pair of particles with highest mass that are kinematically allowed. The decay branching ratios of the SM Higgs as a function of its mass are shown in Figure 1.5.

In the Handbook of LHC Higgs Cross Sections 4 [16], most up-to-date predictions of Higgs boson production cross-sections and decay branching ratios are presented. The production cross-section of a SM Higgs with mass 125 GeV [17] at centre-of-mass energies of 8 and 13 TeV and its branching ratios are listed in Table 1.1 and 1.2 respectively.[1]

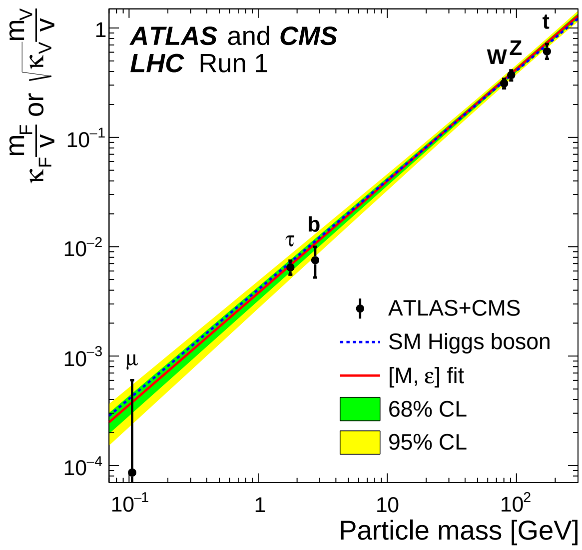


Figure 1.3: Coupling strength parameterised as $\kappa_F m_F/v$ for the fermions and as $\sqrt{\kappa_V} m_V/v$ for the weak vector bosons as a function of particle mass from the combination of ATLAS and CMS Run 1 data. The coupling modifiers κ_F and κ_V are equal to one in the SM case. The dashed (blue) line indicates the predicted dependence on the particle mass in the case of the SM Higgs boson. The solid (red) line indicates the best fit result to the phenomenological model of Ref. [13] with the corresponding 68% and 95% CL bands [12].

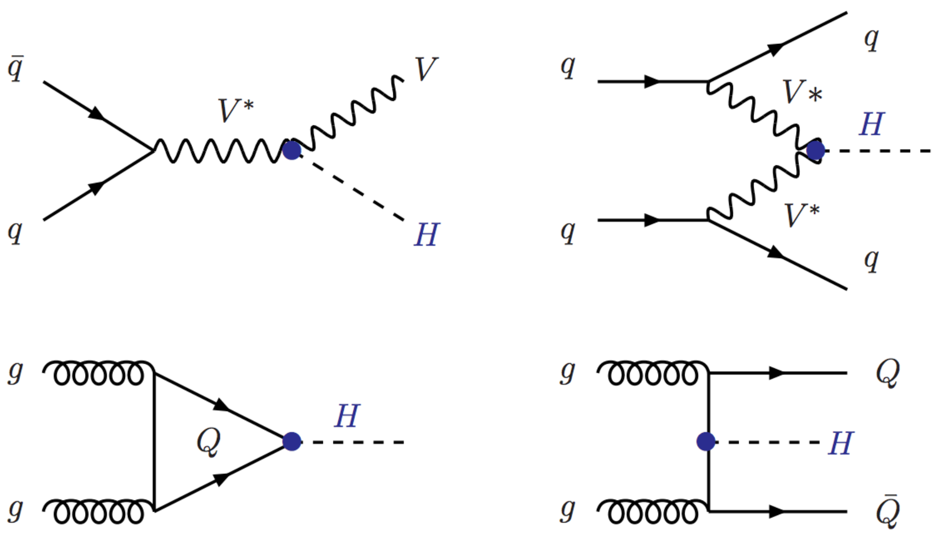


Figure 1.4: Dominant SM Higgs boson production mechanisms in hadronic collisions: associated production with W/Z (top left), vector boson fusion (top right), gluon-gluon fusion (bottom left) and associated production with heavy quarks (bottom right). Q represents heavy quarks [14].

For $m_H = 125$ GeV, the Higgs decays predominantly to two b-quarks, with a branching ratio of 58.24%. On the other hand, the branching ratio of the $H \to \gamma\gamma$ decay, which is of central interest in this thesis, is very low. The decay occurs via the loop-induced processes depicted in Figure 1.6, mediated by W boson and charged fermions. Since the Higgs couplings to fermions is proportional to the fermion mass, the contribution of light fermions is negligible, hence only

[1] The combined mass measurement from ATLAS and CMS gives $125.09 \pm 0.21(stat.) \pm 0.11(syst.)$ GeV [17], but we take it as 125 GeV here.

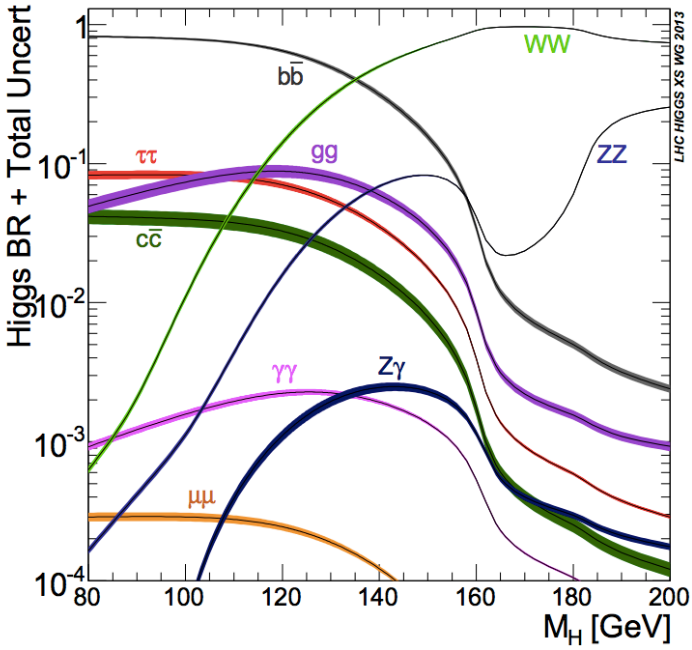


Figure 1.5: Standard Model Higgs boson decay branching ratios [15].

Production mode	$\sigma_{8\ \mathrm{TeV}}$ [pb]	$\sigma_{13\ \mathrm{TeV}}$ [pb]
Gluon-gluon fusion	21.42	48.58
Vector boson fusion	1.601	3.782
WH	0.7026	1.373
ZH	0.4208	0.8839
$t\bar{t}H$	0.1330	0.5071

Table 1.1: SM Higgs production cross-sections at $\sqrt{s} = 8$ TeV and 13 TeV with $m_H = 125$ GeV, as presented in the Handbook of LHC Higgs Cross Sections 4 [16].

Decay mode	Branching ratio [%]
$H \to b\bar{b}$	58.24
$H \to WW$	21.37
$H \to \tau\tau$	6.272
$H \to ZZ$	2.619
$H \to \gamma\gamma$	0.227
$H \to Z\gamma$	0.153
$H \to \mu\mu$	0.022
$H \to$ other	11.1

Table 1.2: SM Higgs branching ratios with $m_H = 125$ GeV, as presented in the Handbook of LHC Higgs Cross Sections 4 [16].

the top quark effectively contributes. Moreover, the W boson loop process and the top quark loop process interfere destructively, leading to a branching ratio of 0.227%. Despite the low branching fraction, the diphoton decay channel constitutes one of the main discovery channel thanks to the excellent diphoton mass resolution of the LHC detectors and clean experimental signature.

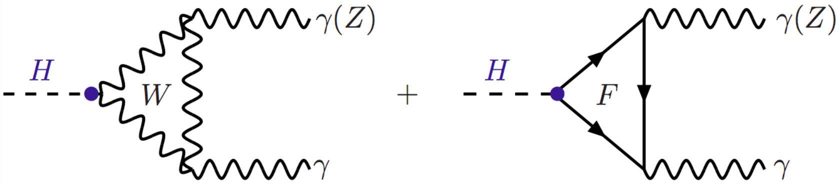


Figure 1.6: Loop-induced Higgs boson decays into two photons (or $Z\gamma$) [14].

1.2 Non-Resonant Diphoton Production

Non-resonant diphoton production is the main background to the high mass diphoton resonances search (Chapters 7 to 9). It also contributes to the background of the $A \to Zh$ search in the $jj\gamma\gamma$ final state reported in Chapter 6 if additionally at least two jets are present.

The SM prompt diphoton production can be split into direct photon production and production of fragmentation photons, where one or both of the two final state photons actually comes from the fragmentation of a quark or a gluon.[2]

There are three main processes contributing to the non-resonant diphoton production:

- the Born process $qq \to \gamma\gamma$, of order $\mathcal{O}(\alpha^2)$,
- the box process $gg \to \gamma\gamma$, of order $\mathcal{O}(\alpha^2\alpha_s^2)$, and
- the bremsstrahlung process $qg \to q\gamma\gamma$, of order $\mathcal{O}(\alpha^2\alpha_s)$.

where α is the fine structure constant defined as $e^2/4\pi$ (e is the electric charge of the electron) and $\alpha_S = g_S^2/4\pi$.

Although the box process is next-to-leading order (NLO), its cross-section is significant due to the enhanced gluon-gluon luminosity at the LHC. Figure 1.7 shows the leading order Feynman diagrams of the above processes including processes with photons produced in parton fragmentation.

In addition, SM photon-jet (γj) and dijet (jj) production also contributes to the total background when the jet(s) is(are) misidentified as a photon. These processes have cross-sections that are many orders of magnitude higher than the diphoton cross-section.

[2] Due to colour confinement, a quark or a gluon cannot be directly observed. It instead fragments into many partons and can radiate a photon in the process.

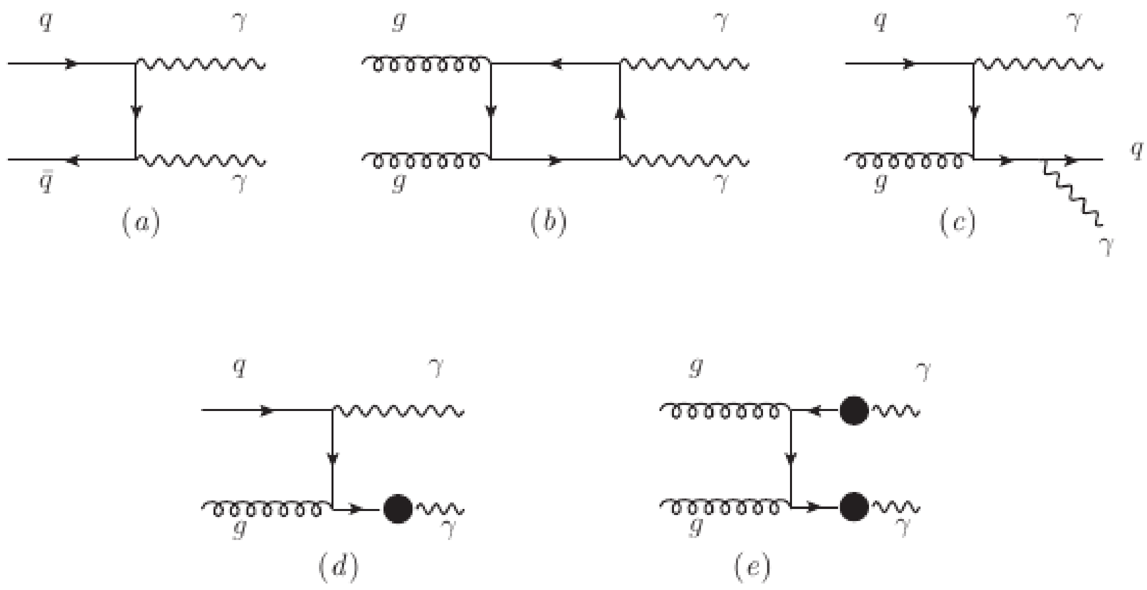
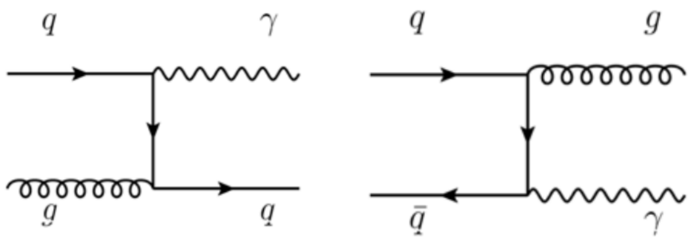
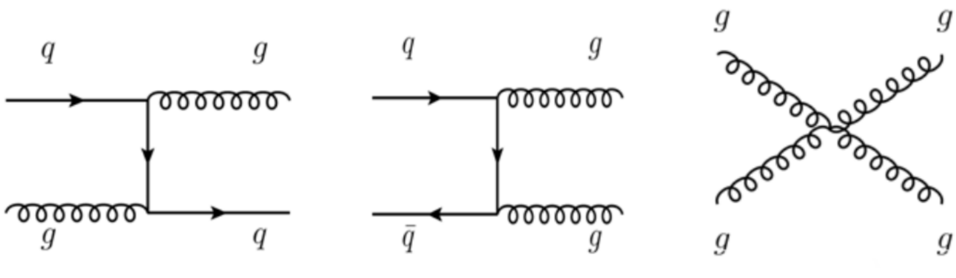


Figure 1.7: Feynman diagrams at the lowest order of (a) the Born process $qq \to \gamma\gamma$, (b) the box process $gg \to \gamma\gamma$, (c) the bremsstrahlung process $qg \to q\gamma\gamma$ and leading order fragmentation processes where (d) one or (e) two partons fragment into high-p_T photons.



(a)



(b)

Figure 1.8: Leading order Feynman diagrams of (a) photon-jet and (b) jet-jet processes.

1.3 Beyond the Standard Model

While the SM is self-consistent, it is an incomplete theory. There are a number of observations that cannot be explained by SM. It does not explain gravity, the matter-antimatter asymmetry and the hierarchy problem, nor provide a dark matter candidate. In addition, according to the Standard Model, neutrinos are massless particles, which is contradicted by the observation of neutrino oscillations [18] in 1998 by the Super-Kamiokande collaboration. This can be accounted by a modification to the Standard Model to include neutrino mass, see Ref. [19] for an example of formulation.

There are many theoretical developments attempting at explaining the deficiencies of the Standard Model. Most consists of various extensions of the Standard Model since the SM already describes the phenomena in its domain very well. The BSM theories related to the searches performed in this thesis are presented in the following, namely the two-Higgs doublet models (2HDMs) and models that include a high mass diphoton resonance.

1.3.1 Two-Higgs-Doublet Models

Two-Higgs doublet models [20] are some of the simplest possible extensions of the SM. They are built by extending the SM scalar sector from a single scalar $SU(2)$ doublet to two scalar doublets.

With two complex scalar $SU(2)$ doublets there are eight fields. Three of the degrees of freedom become longitudinal degrees of freedom of the massive gauge bosons; the remaining five are physical scalar (Higgs) fields. The five physical particles that survive electroweak symmetry breaking are the CP-even neutral Higgs bosons h and H (conventionally, h is lighter than H), the CP-odd pseudoscalar A and two charged Higgs bosons $H^{\pm}$.

A feature of general 2HDMs is the existence of tree-level flavour-changing neutral currents (FCNC). One can avoid these potentially dangerous interactions (since tree-level FCNC is not observed) by imposing discrete symmetries. The Paschos-Glashow-Weinberg theorem [21, 22] states that requiring that each group of right-handed fermions (up-type quarks, down-type quarks and charged leptons) couples exactly to one of the two doublets is necessary and sufficient to avoid FCNC at tree level.

The 2HDMs can be classified into five classes: Type-I, II, III, *lepton-specific* and *flipped* model. Type-III 2HDMs include tree-level FCNCs, while the other types lead to natural flavour conservation. Table 1.3 lists the types of 2HDMs and which of the two doublets (Φ_1 and Φ_2) a fermion type couples to. The Type-I and Type-II 2HDMs are the most studied, thus only these two are considered in the following.

Type	Description	u_{iR}	d_{iR}	e_{iR}	remarks
Type-I	fermiophobic	Φ_2	Φ_2	Φ_2	Φ_2 couples to fermions, Φ_1 to gauge bosons
Type-II	MSSM-like	Φ_2	Φ_1	Φ_1	up- and down-type quarks couple to separate doublets
X	lepton-specific	Φ_2	Φ_2	Φ_1	
Y	flipped	Φ_2	Φ_2	Φ_1	
Type-III		Φ_1, Φ_2	Φ_1, Φ_2	Φ_1, Φ_2	flavor-changing neutral currents at tree level

Table 1.3: Description of two-Higgs-doublet models and the doublet(s) that couple(s) to different fermion types (up-type right-handed quarks u^i_R, down-type right-handed quarks d^i_R and charged right-handed leptons e^i_R). By convention, Φ_2 is the doublet to which up-type quarks couple.

A 2HDM has six free parameters: four Higgs masses (m_h, m_H, m_A, $m_{H^{\pm}}$), the ratio of the two vacuum expectation values ($\tan\beta \equiv v_2/v_1$) and the mixing angle α of the neutral CP-even 2HDM Higgs bosons.

In all models, the coupling of the light Higgs h to either WW or ZZ is the same as the SM coupling multiplied by $\sin(\beta-\alpha)$ and the coupling of the heavier Higgs H is the same as the SM coupling times $\cos(\beta-\alpha)$. To obtain the SM-like limit (also known as decoupling limit) where the couplings of the h boson are SM-like, one sets $\sin(\beta-\alpha) \to 1$.

Coupling of the pseudoscalar A to vector bosons is absent at tree level. The pseudoscalar A decays predominantly to $b\bar{b}$ if the top-quark threshold is not reached. For a mass of A boson above two times the mass of top quark, then the decays to $t\bar{t}$ dominate instead. If the mass of the pseudoscalar is above $m_h + m_Z \approx 220$ GeV, but below 350 GeV, the $A \to Zh$ decay mode can be a good probe for new physics search. Figure 1.9 shows the branching ratio of $A \to Zh$ decay as a function of the pseudoscalar mass for $\sin(\beta - \alpha) = 0.99$ and Type-I 2HDMs.

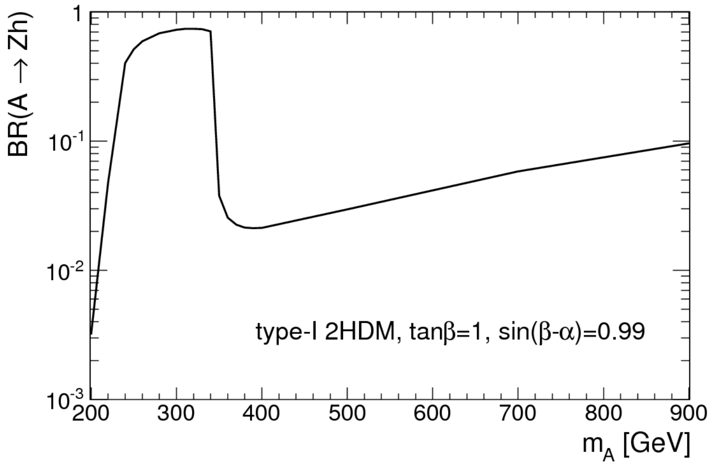


Figure 1.9: Branching ratio of $A \to Zh$ decay as a function of the mass of the pseudoscalar m_A for $\sin(\beta - \alpha) = 0.99$ and Type-I 2HDMs [23].

The A boson is produced mainly through gluon fusion and in association with b-quarks. The production of the A boson in association with b-quarks is relevant for Type-II 2HDMs in the high $\tan\beta$ region.

1.3.2 Models that Include a High Mass Diphoton Resonance

A high mass diphoton resonance can be either spin-0 or spin-2. Spin-1 state decaying to two spin-1 photons is excluded due to the Landau-Yang theorem [24,25] and higher spins are disfavoured theoretically.

Spin-0 resonances are predicted by models with extended Higgs sectors [26–32], including 2HDMs. In these models, the resonances can be either scalar or pseudoscalar, and can have significant branching ratios into two photons in certain scenarios. In the specific case of the 2HDMs [20], the couplings of the heavy scalar and pseudoscalar to the W and Z bosons can be very small (when one assumes SM-like limit) or vanish. Under these assumptions, sizable branching fractions in diphoton final states are possible. In most cases of extended Higgs sectors, the additional Higgs bosons are produced predominantly in the LHC via gluon-gluon fusion due to the much higher gluon-gluon luminosity compared to quark-antiquark.

Another simple extension to the SM is to introduce a scalar that is a singlet under the SM gauge group. In the scalar singlet model introduced in Ref. [33], a discovery in the $\gamma\gamma$ channel is possible while it could be missed in other final states. The Lagrangian for the scalar singlet X interaction with SM fields can be written as

$$\mathcal{L}_X = c_3 \frac{g_X^2}{\Lambda} G^a_{\mu\nu} G^{\mu\nu a} X + c_2 \frac{g^2}{\Lambda} W^i_{\mu\nu} W^{\mu\nu i} X + c_1 \frac{g'^2}{\Lambda} B_{\mu\nu} B^{\mu\nu} X + \sum_f c_f \frac{m_f}{\Lambda} \bar{f} f X, \tag{1.29}$$

where Λ is an energy scale above which BSM physics shows up, the $c_i (i = 1, 2, 3, f)$ are dimensionless coefficients and f is any SM fermion of mass m_f. The partial width of X to two photons is a function of the coefficients of the Lagrangian and is proportional to the mass of the scalar m_X cubed [33]:

$$\Gamma(X \to \gamma\gamma) = \frac{(c_1 + c_2)^2 e^4 m_X^3}{4\pi\Lambda^2}. \tag{1.30}$$

In the spin-2 case, a popular example of theories predicting a new high-mass diphoton resonance is the Randall-Sundrum (RS) model [34]. The RS model attempts to solve the hierarchy problem of the SM, where the electroweak scale ($M_{\mathrm{EW}} \sim 10^3$ GeV) is much lower than the Planck scale ($M_{\mathrm{Pl}} \sim 10^{19}$ GeV).

The model postulates the existence of a fourth spatial dimension that is compactified in a warped geometry space. It involves a finite five-dimensional bulk that contains two (3+1)-dimensional branes: the Planck brane and the TeV brane. SM particles are restricted to the TeV brane, whilst gravity is free to propagate in the bulk.

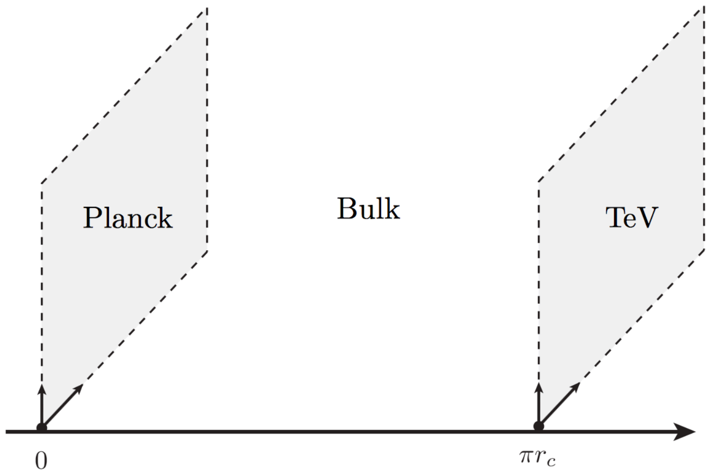


Figure 1.10: Randall-Sundrum scenario with two branes enclosing a five-dimensional bulk.

Given this space-time configuration, TeV scales are naturally generated from the Planck scale thanks to a geometric *warp* factor that relates the fundamental Planck scale on one brane to the apparent scale on the other with a coupling scale defined as

$$\Lambda_\pi = \overline{M}_{\mathrm{Pl}} \exp(-k r_c \pi), \tag{1.31}$$

where $\overline{M}_{\mathrm{Pl}} = M_{\mathrm{Pl}}/\sqrt{8\pi}$ is the reduced Planck scale, and k and r_c are the curvature and compactification radius of the extra dimension, respectively.

The propagation of the graviton in the bulk leads to a series of massive graviton excitations, known as a Kaluza-Klein (KK) tower, that is visible on the TeV brane. These KK gravitons should have spin 2, a mass splitting between successive KK levels on the TeV scale, and a universal dimensionless coupling $k/\overline{M}_{\mathrm{Pl}}$ to the SM fields. The degrees of freedom of the RS model can be expressed in terms of $k/\overline{M}_{\mathrm{Pl}}$ and the mass m_{G^*} of the lightest KK graviton excitation G^*. As opposed to the concept of Large Extra Dimensions (also known as ADD model after Nima Arkani-Hamed, Savas Dimopoulos and Gia Dvali) [35], where a continuous spectrum of Kaluza-Klein states is predicted, the Randall-Sundrum model predicts a series of narrow heavy graviton resonances.

1.3.3 750 GeV Models

The preliminary result of the high mass diphoton resonances search using 2015 data showing a broad excess at 750 GeV (see Chapter 8) attracted significant theoretical interest. In the few months after the result was presented in a seminar at CERN in December 2015, a huge amount (about five hundred) of papers attempting at explaining the origin of the excess were published. Many models were proposed, and it would be impossible to list them all. Ref. [36] gives a review of the various models that can explain the 750 GeV diphoton excess. In most models, the new resonance is assumed to be spin-0, as spin-2 graviton is disfavoured because it couples universally to SM fields and no excess has been seen in the lepton channels. Gluon-initiated production is also preferred due to the higher increase in the gluon-gluon luminosity at $\sqrt{s} = 13$ TeV compared to 8 TeV, which is more compatible with the absence of excess in Run 1 data.

In order to obtain the observed rate, dimension-5 non-renormalisable effective Lagrangian can be used. This is a good approximation to a generic theory with extra particles in the loop of the $X \to \gamma\gamma$ decay, provided such particles are much heavier than the mass of the resonance. This is similar to the scalar singlet model described previously.

In what is known as the *everybody's model*, a neutral scalar X and extra charged fermions or scalars Q are added to the SM. $X \to \gamma\gamma$ and $gg \to X$ are both mediated by Q (replacing the Feynman diagram in Figures 1.4 and 1.6 by non-SM particles Q in the production and decay loop). Q cannot be a SM particle since X would also decay into a $Q\bar{Q}$ pair at tree level. For instance, if the loop is mediated mainly by the top quark, then X could decay mainly to a pair of top quarks and such a final state would have been observed.

Other possibilities that have been proposed include spin-2 graviton that couples more strongly to photons than leptons, composite resonance, dark matter mediator and many others.

Chapter 2

The Large Hadron Collider and the ATLAS Detector

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [37] is the world's largest particle accelerator, designed to collide protons and heavy ions at a centre-of-mass energy of $\sqrt{s} = 14$ TeV. It lies in a tunnel previously used by the Large Electron-Positron collider (LEP) which is 27 kilometres in circumference and around 50 to 175 metres beneath the France-Switzerland border near Geneva, Switzerland.

The construction of the LHC began in 1998. The re-use of existing LEP tunnel and its injection chain helped to cut the costs and time needed for the construction. LHC consists of two rings with counter-rotating beams, that collide at four points where the two rings of the machine intersect. There are four main LHC experiments: ATLAS [38], CMS [39], LHCb [40] and ALICE [41], one at each collision point.

The first LHC beam was circulated through the collider on 10 September 2008. However, a faulty connection between two magnets caused the machine to be stopped for repairs. On 20 November 2009, the LHC started operations again. The beam energy was 3.5 TeV in 2010 and 2011, then increased to 4 TeV in 2012 which corresponds to $\sqrt{s} = 8$ TeV. Between Run 1 (2009-2013) and Run 2(from 2015), The LHC was shut down in early 2013 for a planned two year upgrade, a period known as Long Shutdown 1 (LS1). In 2015, the machine restarted and the first stable beam collisions at $\sqrt{s} = 13$ TeV happened on 3 June 2015.

2.1.1 The LHC Injection Chain

The LHC relies on a series of accelerators to accelerate protons and heavy ions to high energies. Each machine boosts the energy of the particles before injecting them into the next machine. The accelerator complex at CERN is shown in Figure 2.1.

The proton source is a bottle of hydrogen gas, from which electrons are stripped off by an electric field. The protons are injected into the LINAC2, a 30 metres long linear accelerator, which further accelerates them to the energy of 50 MeV. The beam is then transferred to the Proton Synchrotron Booster (PSB) which brings the beam energy to 1.4 GeV, followed by the Proton Synchrotron (PS), which accelerates the beam to 25 GeV. The protons are then injected to the Super Proton Synchrotron (SPS) where their energy is accelerated to 450 GeV.

Finally, the proton bunches are transferred to the two beam pipes of the LHC. The beam in one beam pipe circulates clockwise while the beam in another circulates anticlockwise. The LHC then accelerates the beams to the desired energy, which is 4 (6.5) TeV in the 2012 (2015-2016) data-taking.

The beams are guided around the accelerator ring by superconducting magnets. Thousands

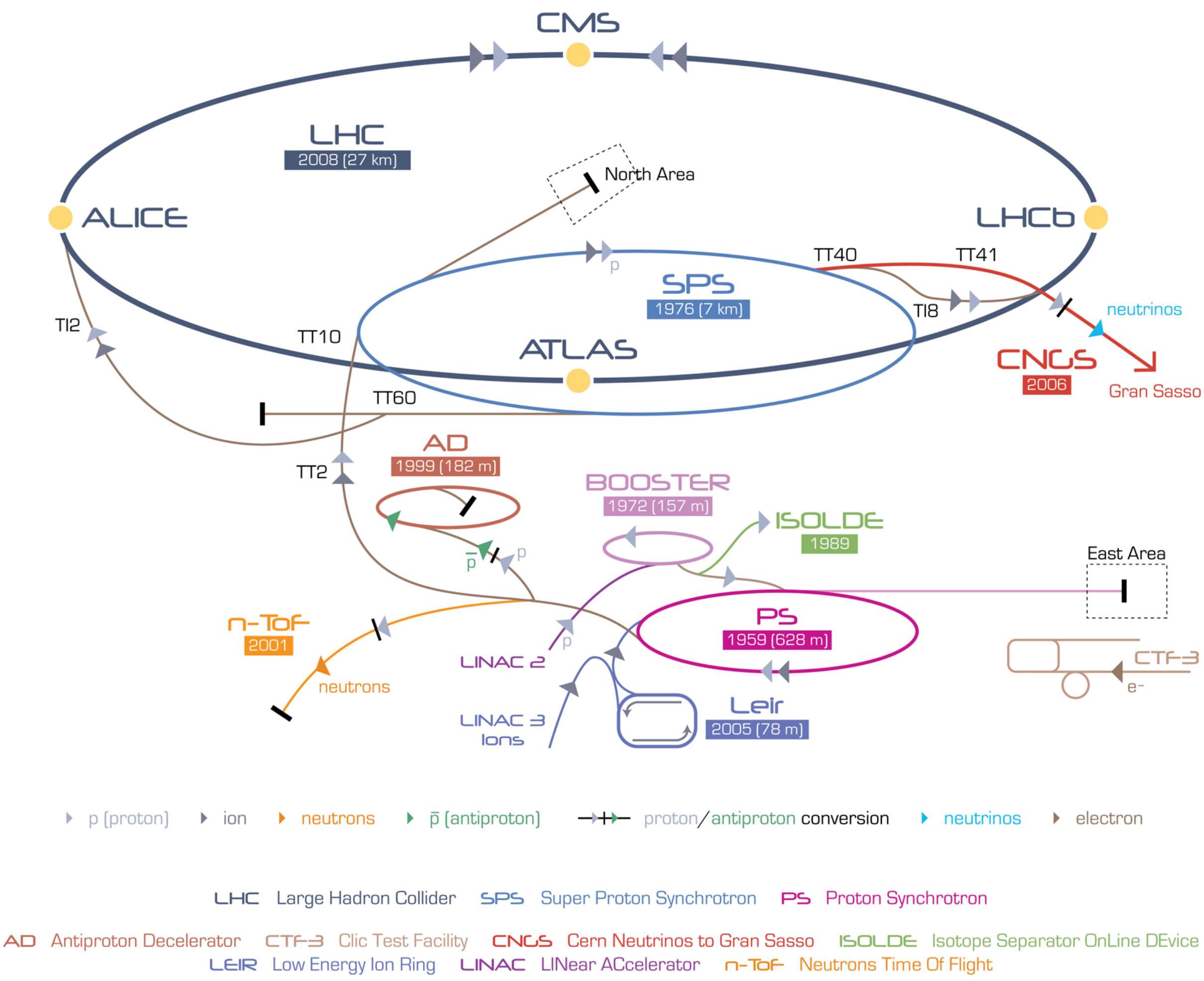


Figure 2.1: CERN accelerator complex.

of magnets of different varieties and sizes are used to direct them around the accelerator. These include 1232 dipole magnets 15 metres in length which bend the beams, and 392 quadrupole magnets, each 57 metres long, which focus the beams. The dipoles of the LHC represent the most important technological challenge for the LHC design. They are able to provide a nominal magnetic field of 8.3 T and require an operating temperature of 1.9 K. Radiofrequency cavities with a maximum oscillation of 400 MHz are used to accelerate the particles from 450 GeV to a maximum of 7 TeV and then keep them at a constant energy by compensating for energy losses.

After injection and acceleration, protons are circulated and collided for a period of time—usually several hours. This is known as a fill. The protons of the LHC circulate around the ring in well-defined bunches. In the LHC, under nominal operating conditions, each proton beam has 2808 bunches, with each bunch containing about 10^{11} protons.

2.1.2 Luminosity

The performance of the LHC can be characterised not only by its beam energy but also by the luminosity. Luminosity is a measure of the number of interactions or events per second. For a specific type of event, its rate is given by $L \cdot \sigma_{\text{event}}$, where L is the instantaneous luminosity and σ_{event} the cross-section of the process.

The luminosity of two Gaussian beams colliding depends on the parameters of the beam:

$$L = \frac{N_b^2 n_b f_{\text{rev}}}{A} = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \epsilon_n \beta^*} F, \tag{2.1}$$

where A is the cross-sectional area, N_b is the number of particles per bunch, n_b the number of bunches per beam, f_{rev} the revolution frequency, γ_r the relativistic γ factor, ϵ_n the normalised transverse beam emittance, β^* the beta function at the collision point and F the geometric luminosity reduction factor due to the crossing angle at the interaction point (IP). The instantaneous luminosity has unit of $\text{cm}^{-2}\text{s}^{-1}$. The beta function relates the beam size σ to the emittance:

$$\beta^* = \frac{\sigma^2 \gamma_r}{\epsilon_n}, \tag{2.2}$$

and the geometric luminosity reduction factor F is given by

$$F = \left(1 + (\frac{\theta_c \sigma_z}{2\sigma^*})^2\right)^{-1/2}, \tag{2.3}$$

with θ_c being the crossing angle at the interaction point, σ_z the RMS bunch length and σ^* the transverse RMS beam size at the interaction point.

The LHC has a design instantaneous luminosity of 10^{34} $\text{cm}^{-2}\text{s}^{-1}$ and bunch spacing of 25 ns. While the instantaneous luminosity is important, integrated luminosity is often used to describe the performance of the collider as it measures the amount of data available to analyses. The aim of the operation of an accelerator is to optimise the integrated luminosity, which is simply the integral of the instantaneous luminosity with respect to time, $\mathcal{L} = \int L dt$. The integrated luminosity is often expressed in inverse femtobarn (fb^{-1}) where a barn is a unit of area equal to 10^{-24}cm^2.

Figure 2.2 shows the expected cross-sections and event rates for several processes of interest at hadron colliders as a function of the centre-of-mass energy.

2.1.3 Performance and Pile-up

Figure 2.3a shows the cumulative integrated luminosity delivered to ATLAS during the period where the beams are stable as a function of month in the year 2011, 2012, 2015 and 2016. The peak instantaneous luminosity as a function of time in 2016 is shown in Figure 2.3b. The design instantaneous luminosity has been exceeded in 2016 where the peak luminosity reaches 1.37×10^{34} $\text{cm}^{-2}\text{s}^{-1}$.

A disadvantage of the high number of closely-spaced bunches in the beam each containing a large number of particles, and therefore high instantaneous luminosity, is the occurrence of additional *pp* collisions in addition to the collision of interest, referred to as *pile-up*. The in-time pile-up refers to the occurrence of multiple collisions in the same bunch-crossing whereas out-of-time pile-up refers to the the additional collisions occurring in bunch-crossings just before and after the collision of interest. Pile-up events affect the measurement of the physics objects used in the analysis from additional energy contributions and track misreconstruction.

In 2012, the mean number of interactions per crossing $\langle\mu\rangle$ was 20.7. In the beginning of Run 2, $\langle\mu\rangle$ was 13.7 in 2015 and 24.2 in 2016. The pile-up is expected to be higher in the following years.

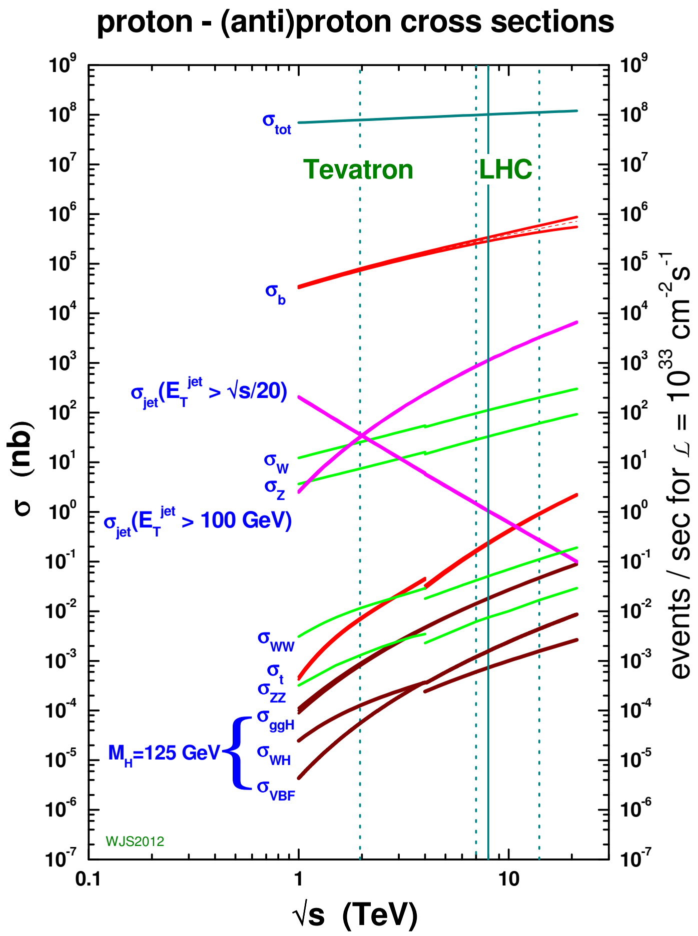
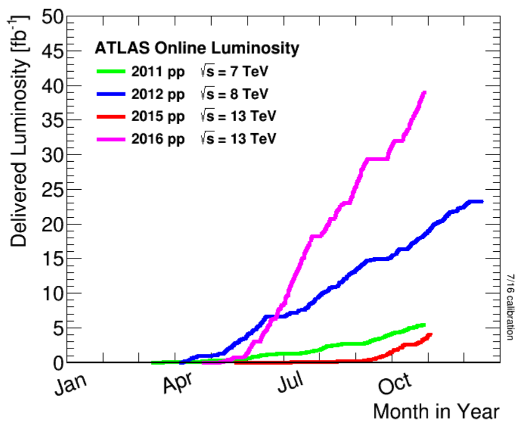
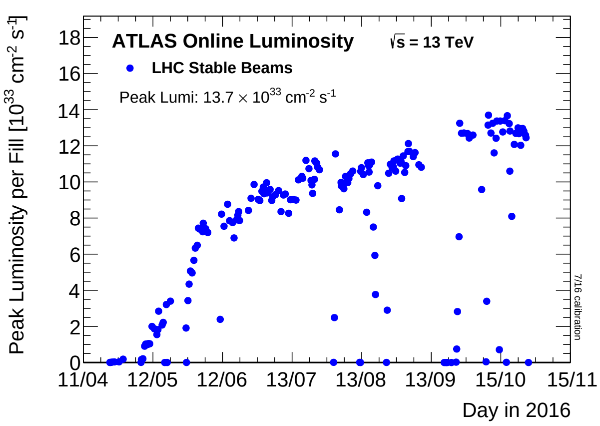


Figure 2.2: Expected cross-sections and event rates for signal and background processes at hadron colliders, as a function of the centre-of-mass energy $\sqrt{s}$. The discontinuity in some of the cross-sections at 4 TeV is due to the switch from proton-antiproton (Tevatron) to proton-proton (LHC) collisions at that energy [42].



(a)



(b)

Figure 2.3: (a) Cumulative luminosity versus month in the year delivered to ATLAS during stable beams and for high energy pp collisions. (b) The peak instantaneous luminosity delivered to ATLAS during stable beams for each LHC fill as a function of time in 2016.

2.2 The ATLAS Detector

The ATLAS (A Toroidal LHC ApparatuS) detector [38] is one of the experiments at the LHC. It is a multi-purpose particle detector with a forward-backward symmetric cylindrical geometry.

2.2.1 Coordinate System

The coordinate system used by ATLAS is a right-handed system with its origin at the nominal interaction point in the centre of the detector. The z-axis is along the beam pipe while the x-y plane is transverse to the beam direction. The x-axis points from the IP to the centre of the LHC ring, and the y-axis is defined as pointing upwards. The side of the detector with positive z is named side-A while the opposite side with negative z is side-C.

Cylindrical coordinates (r, ϕ) are used in the transverse plane, with the azimuthal angle ϕ measured as usual around the z-axis. The polar angle θ is rarely used; instead, pseudorapidity or rapidity is preferred since differences in rapidity are Lorentz invariant under boosts along the longitudinal axis. The rapidity is defined as $y = 1/2 \ln[(E + p_z)/(E - p_z)]$. In the case of massless objects, such as photons, the rapidity is equivalent to the pseudorapidity, which is defined in terms of θ as $\eta = -\ln \tan(\theta/2)$. The angular separation between two objects is typically given by $\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

2.2.2 Detector Overview

The ATLAS detector is composed of three subsystems: the Inner Detector (ID), the calorimeter and the Muon Spectrometer (MS). Each system is typically composed of a cylindrical section centred at the interaction point called barrel, which is sandwiched between two end-cap components, further along the beam line, to provide a large coverage.

The ID is a tracker surrounded by a superconducting solenoid that measures precisely the tracks and vertices of charged particles. The ID consists of the pixel detector, the silicon microstrip Semiconductor Tracker (SCT) detector and the Transition Radiation Tracker (TRT). For the Run 2 of the LHC, a new pixel layer, named Insertable B-Layer (IBL) [43], has been added close to the beam line.

The calorimeter consists of four subdetectors providing electromagnetic and hadronic energy measurement: the electromagnetic (EM) calorimeter, the tile calorimeter, the Hadronic End-Cap Calorimeter (HEC) and the Forward Calorimeter (FCal). Liquid Argon (LAr) technology is used in the EM calorimeter, and also for the hadronic calorimeters in the end-caps and forward regions. The hadronic calorimeter consists of the tile calorimeter in the barrel, which uses iron absorbers and plastic scintillator tiles as active material and the LAr-based HEC and FCal.

The MS comprises of separate trigger and tracking systems, and is immersed in a superconducting air-core toroidal magnet.

The detector measures 25 metres in diameter and 44 metres in length, and weighs approximately 7000 tonnes. The detector layout is shown in Figure 2.4.

2.2.2.1 The Inner Detector and Solenoid

The Inner Detector [44] provides accurate reconstruction of the origins (through the primary or secondary vertices), positions and momenta of charged particles. A sketch of the ID is shown in Figure 2.5.

The ID is surrounded by a thin superconducting solenoid operating at 4.5 K that provides a 2 T axial magnetic field. The field curves the trajectory of charged particles and allows for a precise measurement of their momenta. For optimal calorimeter performance, the solenoid

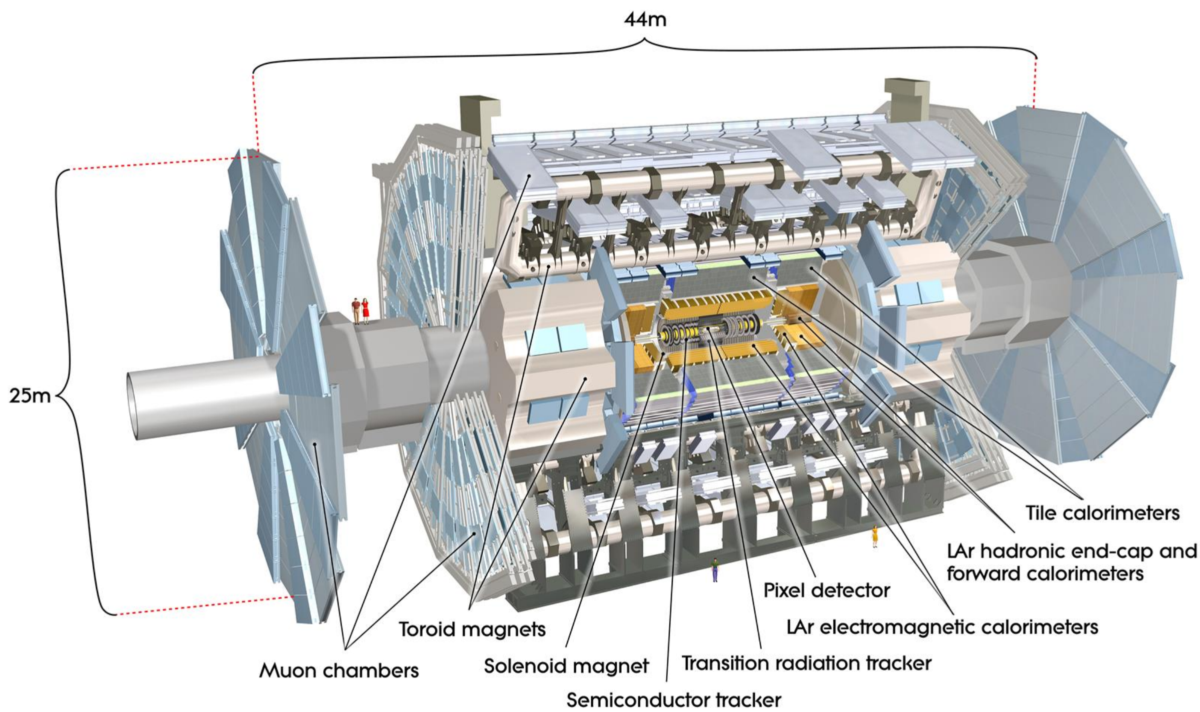


Figure 2.4: Cutaway view of the ATLAS detector. [38]

material was optimised to be as low as possible, amounting to 0.66 radiation lengths (X_0) at normal incidence.[1]

Since the LHC operates at high instantaneous luminosity and there are many collisions within the same bunch crossings, there is a very large track density in the detector. This challenge is met by the fine granularity and high precision measurements of the detector. The pixel detector and the SCT are made of silicon pixel and microstrip sensors that offer precise tracking, while the TRT is a straw tracker with also particle identification capabilities using transition radiation. Anticipating the challenges posed by high luminosities during Run 2 and to mitigate the impact of radiation damage to the innermost layer of the pixel detector, a fourth layer, the IBL has been added to ATLAS during the shutdown period between Run 1 and Run 2 at 33 mm away from the beam line.

In the barrel region, the silicon trackers are arranged on concentric cylinders around the beam axis whereas in the end-caps they are located on disks transverse to the beam axis. The pixel detector has high granularity, the minimum pixel size in $R - \phi \times z$ being 50×400 μm^2. The pixel modules are arranged in three concentric cylinders in the barrel (R = 50.5, 88.5, 122.5 mm) in addition to the IBL and three disks in each end-cap (z = 495, 580, 650 mm). The intrinsic accuracies are 10 μm in R-ϕ and 115 μm in z (R) in the barrel (end-cap). The SCT consists of four double layers of silicon microstrips in the barrel (R = 299, 371, 443, 514 mm), and nine disks of varying sizes in each end-cap (from z = 854 to 2720 mm). Each microstrip is 6.4 cm long with a pitch of 80 μm. The intrinsic resolution is 17×580 μm in $R - \phi \times z$ (R in the end-cap). The pixel detector and the SCT cover the region $|\eta| < 2.5$.

The TRT comprises many layers (73 in the barrel) of gaseous straw tubes interleaved with transition radiation material, providing typically 36 hits per track. The barrel part provides only R-ϕ information. In the barrel, the straws are parallel to the beam axis whereas in the

[1]The radiation length is the mean distance over which a high-energy electron loses all but $1/e$ of its energy by bremsstrahlung, and 7/9 of the mean free path for pair production by a high-energy photon.

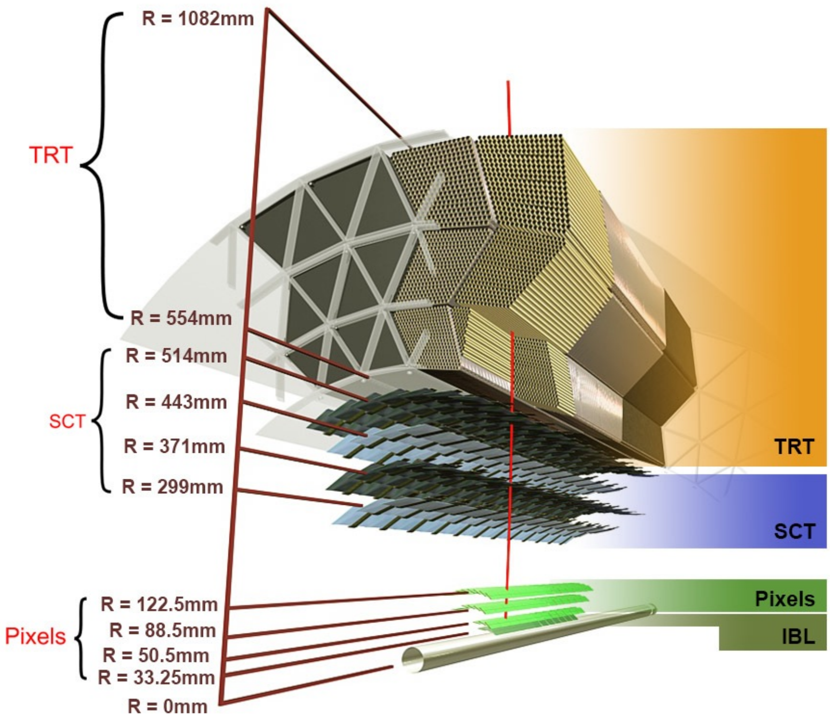


Figure 2.5: Sketch of the ATLAS inner detector showing all its subsystems, including the new IBL. The distances to the IP are also shown. [45]

.

end-cap they are arranged radially in wheels. The straw tubes of the TRT provides tracking up to $|\eta| = 2.0$.

The combination of precise silicon trackers at small radii with the TRT at larger radius gives robust pattern recognition and high precision in both R-ϕ and z coordinates. In addition to tracking, electron identification capability of the TRT is complementary to that of the calorimeter.

2.2.2.2 Calorimetry

The calorimeters are situated outside the solenoidal magnet that surrounds the ID. There are two basic calorimeter systems: an inner electromagnetic calorimeter and an outer hadronic calorimeter. Both are sampling calorimeters, i.e. they consist of alternating layers of an absorber and an active material that provides the detectable signal.

The barrel part of ATLAS calorimetry system is composed of a high-granularity lead/liquid-argon (LAr) EM calorimeter and a steel/scintillating-tile hadronic calorimeter. In the end-caps, the hadronic calorimetry is implemented with a LAr design instead of with scintillating tiles. A schematic of the sampling calorimetry system is presented in Figure 2.6.

The LAr technology requires cryogenic installations, as the operation temperature is typically 87 K above which argon is a gas. The barrel EM calorimeter is contained in a barrel cryostat, which surrounds the tracking detectors. Two end-cap cryostats house the end-cap EM and hadronic calorimeters, as well as the forward calorimeter.

Liquid Argon Electromagnetic Calorimeter

The LAr electromagnetic calorimeter [46] has an accordion geometry. It is divided into a barrel section covering the pseudorapidity region $|\eta| < 1.45$ and two end-caps covering the pseudorapidity region $1.375 < |\eta| < 3.2$. The EM calorimeter has three longitudinal layers.

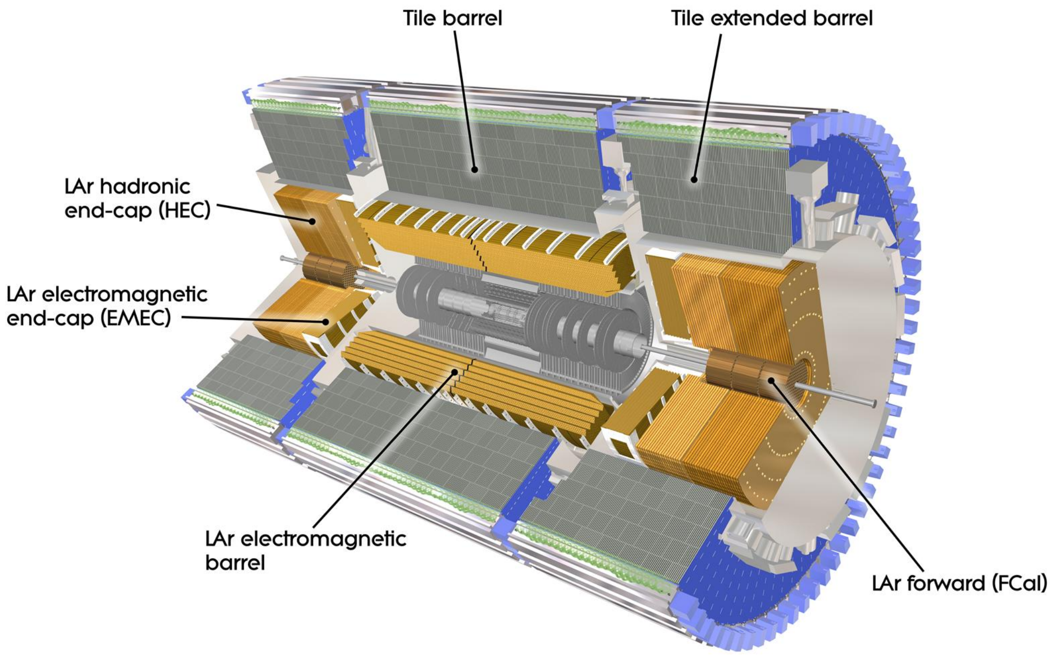


Figure 2.6: Cutaway view of the ATLAS calorimetry system. [38]

The first one, also called the *strip layer*, has a high granularity in η to discriminate between showers initiated by single photons and by photon pairs from the decay of neutral mesons in jets. The second layer collects most of the energy of the electromagnetic showers and has a granularity of 0.025×0.0245 in $\eta \times \phi$ at $|\eta| < 1.4$.[2] A third layer has coarser segmentation in η and is used to correct for leakage of high energy showers. The third layer covers only the pseudorapidity region $|\eta| < 2.5$. In the region $|\eta| < 1.8$, a thin presampler layer is installed in front of the first layer and is used to estimate the energy losses from material upstream of the calorimeter. A sketch of the EM calorimeter in the barrel region is shown in Figure 2.7. A summary of the pseudorapidity coverage, granularity and segmentation of the EM calorimeter is provided in Table 2.1.

The relative energy resolution of the EM calorimeter can be parameterised as

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \tag{2.4}$$

where a, b and c are η-dependent parameters; a is the sampling term, b the noise term, and c the constant term. The design value of the sampling term is around 10% in the barrel and worsens as the amount of material upstream of the calorimeter increases at larger $|\eta|$. The noise term describes the fluctuations from pile-up and electronic noises. The constant term is due to non-uniformities in the response of the calorimeter; it has a design value of 0.7%. The sampling term and the noise term are important for low energy particles. At higher energies, the relative energy resolution tends asymptotically to the constant term.

The thickness of the EM calorimeter is more than 22 X_0 in order to fully contain the showers of photons and electrons. Although the EM calorimeter is designed for electrons and photons, typically a significant fraction of jet energy is deposited there.

[2]The length in ϕ is sometimes quoted as 0.025 and sometimes as 0.0245; in reality, it is defined as $2\pi/256 = 0.02454$.

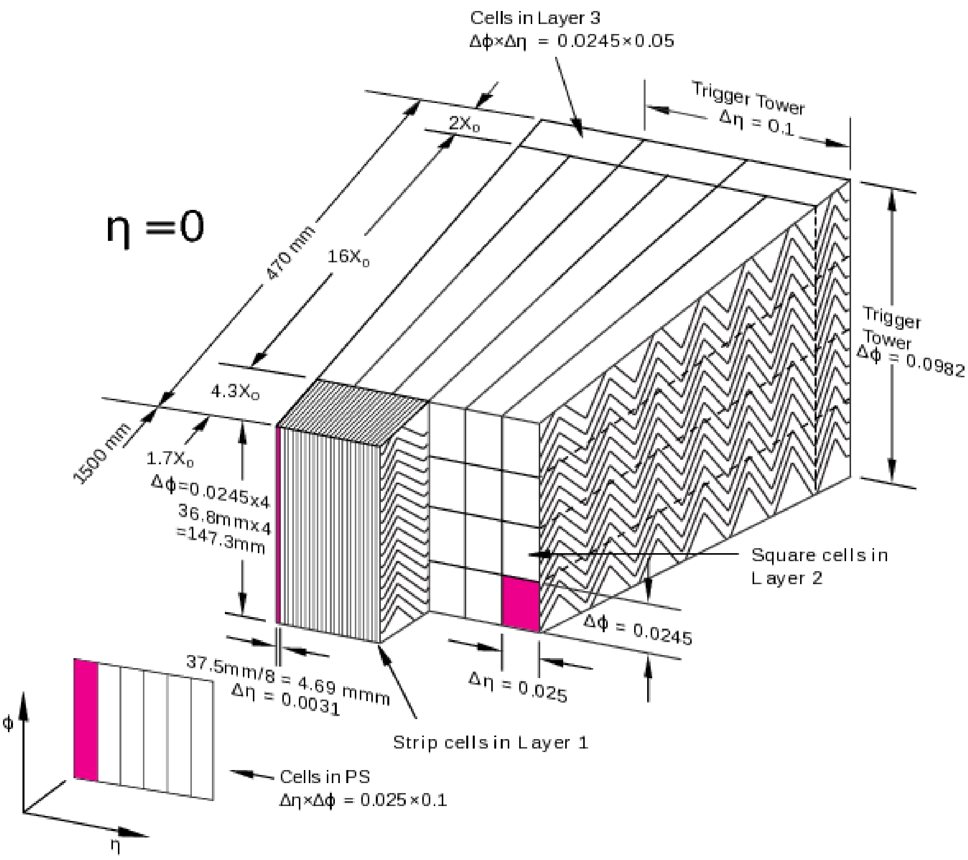


Figure 2.7: Schematic view of a barrel module of the EM calorimeter showing the accordion geometry. The different longitudinal layers (presampler and three layers in the accordion calorimeter) are depicted. The granularity of the cells in each of the longitudinal layers is shown.

	Barrel	End-cap
Presampler	0.025×0.1 ($\|\eta\| < 1.52$)	0.025×0.1 ($1.5 < \|\eta\| < 1.8$)
Layer 1	$0.025/8 \times 0.1$ ($\|\eta\| < 1.4$)	0.050×0.1 ($1.375 < \|\eta\| < 1.425$)
	0.025×0.025 ($1.4 < \|\eta\| < 1.475$)	0.025×0.1 ($1.425 < \|\eta\| < 1.5$)
		$0.025/8 \times 0.1$ ($1.5 < \|\eta\| < 1.8$)
		$0.025/6 \times 0.1$ ($1.8 < \|\eta\| < 2.0$)
		$0.025/4 \times 0.1$ ($2.0 < \|\eta\| < 2.4$)
		0.025×0.1 ($2.4 < \|\eta\| < 2.5$)
		0.1×0.1 ($2.5 < \|\eta\| < 3.2$)
Layer 2	0.025×0.025 ($\|\eta\| < 1.4$)	0.050×0.025 ($1.375 < \|\eta\| < 1.425$)
	0.075×0.025 ($1.4 < \|\eta\| < 1.475$)	0.025×0.025 ($1.425 < \|\eta\| < 2.5$)
		0.1×0.1 ($2.5 < \|\eta\| < 3.2$)
Layer 3	0.05×0.25 ($\|\eta\| < 1.35$)	0.050×0.025 ($1.5 < \|\eta\| < 2.5$)

Table 2.1: Granularity in $\Delta\eta \times \Delta\phi$ of the EM calorimeter [38].

Tile Calorimeter

The tile calorimeter [47] is a hadronic calorimeter in the barrel that uses steel absorbers and plastic scintillator tiles as active medium. The ultraviolet scintillation light that is produced when a charged particle crosses the active medium is collected by wavelength-shifting optical fibres that feed into photomultiplier tubes. The tile long barrel covers the region $|\eta| < 1.0$, and its two extended barrels cover the range $0.8 < |\eta| < 1.7$. The tile calorimeter is divided azimuthally into 64 modules and segmented in depth into three layers. The granularity is 0.1×0.1 in $\eta \times \phi$ in its first two layers and 0.1×0.2 in the third. The hadronic calorimetry is extended to larger pseudorapidities by the HEC and the FCal detectors.

Hadronic End-Cap Calorimeter

The HEC is a liquid argon sampling calorimeter using copper plates as absorbers. It is located in the end-cap region and covers $1.5 < |\eta| < 3.2$. The HEC comprises two wheels in each end-cap cryostat: a front wheel (HEC1) and a rear wheel (HEC2), each wheel containing two longitudinal sections. It has a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ at $1.5 < |\eta| < 2.5$ and 0.2×0.2 for $2.5 < |\eta| < 3.2$.

Forward Calorimeter

The LAr FCal is located in the same cryostat as the end-cap calorimeters and covers the forward region of $3.1 < |\eta| < 4.9$. Each FCal is split into three modules. Copper is used as the absorber for the first layer, and tungsten for the other two.

2.2.2.3 Muon Spectrometer

The MS [48] is a system of tracking detectors surrounding the calorimeters, designed to identify muons and measure their momenta. Muons interact little with the calorimeters, so their tracks are measured precisely by the MS located on the outer part of the ATLAS detector system. The muon tracks are deflected by the large superconducting air-core toroidal magnet [49, 50], that gives ATLAS its name. The ATLAS toroidal system includes a barrel toroid (composed of 8 separate coils) and two end-cap toroids, which together provide a 0.5 T field in the central region and a 1 T field in the end-caps.

The MS consists of two tracking systems and two triggering systems. The Monitored Drift Tubes (MDTs) and the Cathode Strip Chambers (CSCs) are designed for precision tracking. The MDTs cover the pseudorapidity range $|\eta| < 2.7$ and consist of three to eight layers of drift tubes. The MDTs are similar to the TRT's straws, but with larger tube diameters. In the forward region of $2 < |\eta| < 2.7$, the CSCs which are multiwire proportional chambers with cathode planes segmented into strips in orthogonal directions are used in the innermost tracking layer due to their higher rate capability and time resolution.

The trigger system covers the pseudorapidity region of $|\eta| < 2.4$. Resistive Plate Chambers (RPCs) are used in the barrel and Thin Gap Chambers (TGCs) in the end-caps. RPCs are gaseous parallel electrode-plate detectors while TGCs are made of effectively the same technology as the CSCs. Muon triggers are based on coincidences in two or more layers with a pattern fitting the description of a muon track above a minimum p_T threshold.

2.2.2.4 Forward detectors

The ATLAS forward detectors are a set of smaller detectors in the forward region, whose main purpose is to study inelastic *pp* scattering at small angles:

- LUCID (Luminosity measurements Using Cherenkov Integrating Detector), a Cherenkov detector consisting of two symmetric modules installed at ± 17 m from the interaction point. LUCID is dedicated to monitoring the LHC luminosity via the measurement of *pp* inelastic collisions.

- ZDC (Zero Degree Calorimeters) located at ± 140 m from the interaction point is dedicated to the measurement of centrality for heavy-ions collisions by the detection of forward spectator neutrons. It consists of six tungsten/quartz calorimeter modules where the light from the quartz fibers is read out by photomultiplier tubes.

- ALFA (Absolute Luminosity For ATLAS), built with scintillating fiber trackers and located at ± 240 m from the IP. It estimates the absolute luminosity of ATLAS by measuring the elastic-scattering amplitude at small angles and by comparing it to the total cross-section.

2.2.3 Trigger and Data Acquisition System

The LHC is designed with a maximum bunch crossing rate of 40 MHz. A multi-level (three-level for Run 1, two-level for Run 2) trigger system [51, 52] is used to select events to be recorded for offline analyses.

The Level-1 (L1) trigger [53] finds regions of interest (RoIs) in the calorimeters and MS by using fast, dedicated hardware. It reduces the event rate to the 100 kHz level. Events retained by the L1 trigger are passed to the high-level trigger (HLT) [54]. The HLT is further divided into two parts during the Run 1 data taking: the Level-2 (L2) and the Event Filter (EF) triggers. These are combined during Run 2. The HLT is software-based and has access to the full detector information. The full event is read out and the algorithms used to reconstruct and select objects are as close to the offline reconstruction software as possible. The HLT further reduces the event rate to the 1 kHz level.

A trigger selection can be *prescaled* or *un-prescaled*: prescaled triggers only retain a fraction of the events that would pass the trigger selection, to further reduce the data-flow rate and avoid over-burdening the data taking system.

The Data Acquisition (DAQ) system is the framework in which the Trigger System operates. After an event is accepted by the L1 trigger, the data are transferred to the Readout Drivers (ROD). The digitised signals are converted into raw data and transmitted to the HLT. The events selected by the HLT are transferred to mass storage. In addition to moving data down to the trigger selection chain, the DAQ provides an interface for configuration, control and monitoring of the ATLAS detector during data taking. Supervision of the detector hardware is handled by a separate framework called the Detector Control System (DCS).

2.2.4 Computing and Software

The Worldwide LHC Computer Grid (WLCG) is a network connecting many computer centres in many countries. It provides computing resources to store, distribute and analyse the data generated by the LHC. The WLCG is made up of four layers, or *tiers*. The primary event (or raw data) processing occurs at CERN in the Tier-0 facility. The raw data format is transformed into the Event Summary Data (ESD) format, which contain all the reconstruction information (calorimeter cells, tracks, etc) of the event. In addition, the data format can also be stored in terms of the reconstructed physics objects instead of the detector information, named Analysis Object Data (AOD).

Then, the datasets are copied to Tier-1 facilities around the world, which are responsible for the safe-keeping of the raw and reconstructed data. Finally, the data are copied to the Tier-2 facilities, typically located in universities and other scientific institutes. Individual scientists can access these facilities through local (also sometimes referred to as Tier-3) computing resources.

2.2.5 Data Quality

Only data recorded with stable beam and in optimal functional conditions of the ATLAS detector are used in the analysis, ensuring all essential elements of the ATLAS detector were operational with good efficiency. A Data Quality (DQ) assessment system is introduced in order to identify problems that can occur during the data taking. Once the quality of the data for a given period is assessed and the problems identified, lumiblocks[3] or even entire runs that cannot be used for physics due to a severe problem are rejected. A list of runs and lumiblocks that are available for physics analysis is established and called the Good Run List (GRL).

On top of the GRL request, a series of standard cleaning cuts are applied. Data events that are incomplete, corrupted, or have problems in the calorimetric system are removed. Events with a trip in the tile calorimeter are also rejected.

The relative fraction of luminosity associated to data of good quality delivered by the various subdetectors during the 2016 data taking up to 11 July 2016 is shown in Table 2.2. The toroid magnet was off for some runs, but this does not affect the dataset used in the high mass diphoton resonances search in Chapter 9. In this period of data taking, 91 to 98% of the total luminosity delivered by the LHC is good for physics, depending on whether or not the data with toroid magnet off can be used.

Inner Tracker			Calorimeters		Muon Spectrometer				Magnets	
Pixel	SCT	TRT	LAr	Tile	MDT	RPC	CSC	TGC	Solenoid	Toroid
98.9	99.9	100	99.8	100	99.6	99.8	99.8	99.8	99.7	93.5
Good for physics: 91-98% ($10.1-10.7$ fb^{-1})										

Table 2.2: Luminosity weighted relative fraction of good quality data delivery by the various components of the ATLAS detector/trigger subsystems during LHC fills with stable beams in 2016 data taken between 28 April and 10 July 2016, corresponding to a recorded integrated luminosity of 11.0 fb^{-1}. The toroid magnet was off for some runs, leading to a loss of 0.7 fb^{-1}. Analyses that do not require the toroid magnet can use that data.

[3] During the data taking, the events are split into lumiblocks that corresponds to roughly one or two minutes.

Chapter 3

Photon Reconstruction and Performance

Several physics processes in pp collisions produce final states with photons. In particular, photons were instrumental in the discovery of the Higgs boson in the $H \to \gamma\gamma$ channel in 2012. The analyses that I have participated in as well as the study on photon conversion in Chapter 4 also involve photon objects. Thus, it is necessary to discuss how the photon candidates are reconstructed.

A photon candidate can be either converted (produces an electron-positron pair when it interacts with the material upstream of the EM calorimeter) or unconverted. If the conversion occurs early enough in the inner detector, the associated charged tracks can be reconstructed. A photon is classified as converted when one or two tracks match the energy cluster in the electromagnetic calorimeter and the track(s) are consistent with those originating from a conversion vertex. On the other hand, a photon that does not convert in the inner detector does not leave tracks. Converted and unconverted photon is distinguished from electron which has cluster matched to a well-reconstructed inner detector track originating from a vertex in the beam interaction region. Due to certain ambiguities between the converted photon and the electron candidates, the reconstruction of electrons is also briefly discussed in the following.

A large fraction of the reconstructed photon candidates are actually jets from hadronic decays misreconstructed as photons. To reject these background photons, two types of selections are applied on photon candidates: *identification* and *isolation* requirements.

The description on photon reconstruction and selection in this chapter are the foundations of the following chapters on the photon conversion studies and the search analyses. The reconstruction process of the photon candidates is first outlined in Section 3.1. Section 3.2 then describes the energy calibration process. Photon identification and isolation are summarised in Section 3.3 and 3.4 respectively.

3.1 Photon Reconstruction

Photons and electrons are reconstructed in ATLAS from clusters of energy deposit in the EM calorimeter. The EM calorimeter is divided into towers of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.0245$. Inside each tower, the energy of all cells in the four longitudinal layers is summed, giving the energy of the tower. Clusters are seeded by towers with total transverse energy above 2.5 GeV, and searched for by a sliding window algorithm [55], with a window size of $\Delta\eta \times \Delta\phi = 0.075 \times 0.123$ (3×5 cells in $\eta \times \phi$). The sliding window algorithm [55] selects a rectangular window of this size that maximises the contained energy by adjusting the centre of the window in η and ϕ. After an energy comparison, duplicate clusters of lower energy are removed from nearby seed clusters. The initial cluster reconstruction efficiency for photons with transverse energy $E_T > 25$ GeV is estimated from simulation to be greater than 99%.

Once seed clusters are reconstructed, a search is performed for inner detector tracks [56, 57] that are loosely matched to the clusters, in order to identify and reconstruct electrons and photon conversions.

Tracks are considered as loosely matched to a cluster if the distance between the cluster barycentre and the extrapolation of the tracks to the middle layer of the EM calorimeter is within 0.05 along η (for tracks with at least 4 hits in the silicon detectors) and 0.05 or 0.2 in ϕ depending on which side the track is bending towards. To recover low momentum tracks that potentially suffered significant bremsstrahlung loss before reaching the calorimeter, the extrapolated track after rescaling its momentum to the measured cluster energy is required to be within 0.1 in ϕ to the cluster on the side the track is bending towards or within 0.05 on the opposite side. Only tracks loosely matched to a cluster are retained for the reconstruction of electrons and converted photons. Finally, the track parameters, except stand-alone TRT tracks, are precisely re-estimated using an optimised electron track fitter, the Gaussian Sum Filter (GSF) algorithm [58].

Double-track conversion vertex candidates are reconstructed from pairs of oppositely charged tracks with transverse momentum $p_T > 500$ MeV in the inner detector that are likely to be electrons. The pairs are selected such that the tracks are close in space and have a small opening angle. The selected track pairs are fitted to a common vertex under the constraint that they are parallel at the vertex. There are three categories of track pairs forming a double-track conversion: two tracks with at least four hits in the silicon detector (Si-Si), two tracks without silicon hits (TRT-TRT) and pairs with one track with silicon hits and one stand-alone TRT track (Si-TRT).

The efficiency to reconstruct photon conversions as double-track vertex candidates decreases significantly for conversions taking place in the outermost layers of the inner detector. This is due to late conversions where the two tracks are not well separated. Considering single-track conversions allows to recover such late conversions as well as asymmetric conversions where one of the two tracks has low transverse momentum. A candidate is reconstructed as a single track conversion if the track does not have any hits in the b-layer, and if the tracks have hits in the TRT, they are required to be consistent with the pattern expected from an electron. Since a vertex fit cannot be performed, the conversion vertex is defined as the location of the hit associated to the track that is the closest to the beams. The position of the conversion vertex reconstructed in this way can be off by as much as a detector layer, but this is less severe if the conversion occurs inside the TRT thanks to the high straw density. Tracks which pass through a dead region in the b-layer are not considered as single-track conversions unless they are missing a hit in the next layer.

The classification of tracks as single-track or double-track conversion candidates described above is applied to all tracks that are loosely matched to an EM cluster. Once this classification is done, a new matching procedure is performed. It varies according to the type of the conversion vertex candidates:

- Symmetric double-track conversion vertex (the ratio of the p_T of the higher p_T track to that of the lower p_T track is less than a factor of four). Each track is extrapolated individually to the second layer of the calorimeter and both tracks are required to match the same cluster.

- Asymmetric double-track conversion vertex (the p_T difference of the two tracks is more than a factor of four). The converted photon direction is reconstructed from the electron and positron track parameters at the vertex determined by the constrained vertex fit and a straight line extrapolation is performed from the vertex position to the second layer.

- Single-track conversion vertex. The track is extrapolated to the second layer from its last measurement.

If the impact point on the calorimeter is within a certain window in (η, ϕ) from the cluster centre, then the conversion vertex candidate is considered matched to that cluster. The requirement differs for different types of conversion candidates and whether the associated tracks have hits in the silicon (Si) detector.

If multiple conversion vertices are matched to the same cluster, preference is given to double-track candidates over single-track candidates, and to the candidate with more tracks with hits on the silicon detectors. If the remaining vertex candidates are formed from the same number of tracks with Si hits, preference is given to the candidate with the vertex closer to the beam axis.

The converted candidate is built from the best conversion vertex candidate and the cluster, whereas unconverted photon candidate is built from cluster to which neither a conversion vertex candidate nor any track has been matched. At this stage of the reconstruction, most of the unconverted photon candidates have been identified, but the converted ones are mostly still treated as electron candidates.

Converted photon candidates that are also electron candidates are classified as converted photon if the electron track coincides with the track(s) originating from the conversion vertex candidate matched to the same cluster. The only exception is in the case of a double-track conversion vertex candidate where the coinciding track has a b-layer hit, while the other track lacks one (missing hit in a disabled b-layer module is counted as a hit). If the track does not coincide with any of the tracks from the conversion vertex candidate, the converted photon candidate is rejected unless its p_T is larger than the electron track p_T. Single-track converted photon candidates are recovered from electron candidates if the electron candidate has a best matched track that is made of only TRT hits with $p_T > 2$ GeV and $E/p < 10$ (E being the cluster energy and p the track momentum), regardless of whether a conversion vertex candidate has been matched to its cluster or not. Electron candidates having corresponding stand-alone TRT track and the track p_T is less than 2 GeV are recovered as unconverted photons. In addition, electron candidates that are not considered as single-track converted photon candidates and with a matched track that has $p_T < 2$ GeV or $E/p > 10$ are also recovered as unconverted photon candidates. Using this procedure, around 85% of the unconverted photons erroneously categorised as electrons are recovered.

The procedure to recover converted and unconverted photons from the electrons candidates is simplified in Run 2. If no track with at least four hits in the silicon detector is matched to the cluster, it is recovered as a photon. If the vertex candidate is made of two tracks with hits in the silicon detector, one of which being the electron track and has no hit in the Pixel detector, it is also considered as a photon. Some criteria is applied to choose the electron candidate, and the rest are considered as both photons and electrons.

Figure 3.1 shows the conversion reconstruction efficiency as a function of pseudorapidity, as estimated from MC simulation. The conversion reconstruction is inefficient around $|\eta| = 1$ due to the gap at the transition from the barrel to the end-cap TRT and at $|\eta| > 2$ as the limit of the TRT pseudorapidity range is approached. The total efficiency in reconstructing prompt photons with $E_T > 25$ GeV in the pseudorapidity acceptance as photon candidates, either converted or unconverted, is found to be 96% [59]. The remaining are incorrectly reconstructed as electrons. At high E_T, the photon reconstruction efficiency decreases due to a reduction in the efficiency to reconstruct photon conversions when the two conversion tracks are close by and hard to separate.

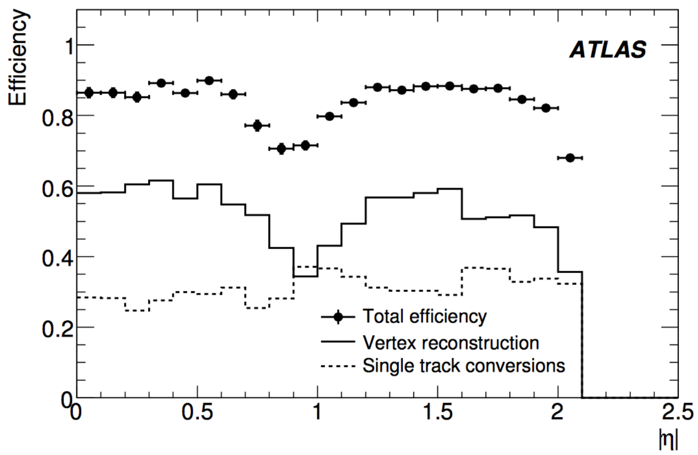


Figure 3.1: Reconstruction efficiencies for conversions from 20 GeV p_T photons as a function of conversion radius (left) and pseudorapidity (right) [60]. The points with error bars show the total reconstruction efficiency, the solid histogram shows the conversion vertex reconstruction efficiency, and the dashed histogram shows the single-track conversion reconstruction efficiency.

3.2 Photon Energy Reconstruction and Calibration

3.2.1 Energy Reconstruction

The photon energy is measured using information from the calorimeter cluster whose size depends on the classification of the photon. In Run 1, clusters of 3×5 cells in $\eta \times \phi$ are used for unconverted photon candidates in the barrel while for converted photon candidates, the size increases to 3×7. Converted photons are allowed a broader window in $\Delta\phi$ due to the opening of the conversion products under the magnetic field. In the end-cap, the same cluster size of 5×5 cells is used for all candidates. In Run 2, the cluster size of photon candidates in the barrel has changed to 3×7 cells for both converted and unconverted photon candidates.

Before the EM cluster energy is calibrated, a correction to the η and ϕ positions of the photon candidates is applied. The position is measured independently for each calorimeter layer, then weighted according to the position of the cells in the cluster and corrected for systematic effects known from MC and test beam studies, and taking into account the finite granularity of each layer.

The photon transverse momentum E_T is computed from the photon cluster's calibrated energy E and the pseudorapidity η_{S2} of the barycentre of the cluster in the second layer of the EM calorimeter as $E_T = E/\cosh(\eta_{S2})$.

The relative energy resolution of the EM calorimeter can be parameterised as

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \tag{3.1}$$

where a, b and c are η-dependent parameters; a is the sampling term, b is the noise term, and c is the constant term. The design value of the sampling term is around 10% in the barrel and worsens as the amount of material upstream of the calorimeter increases at larger $|\eta|$. The noise term describes the fluctuations from pile-up and electronic noises. The constant term is due to non-uniformities in the response of the calorimeter; it has a design value of 0.7%. The sampling term and the noise term are important for low energy particles. At higher energies, the relative energy resolution tends asymptotically to the constant term.

3.2.2 Energy Calibration

The procedure to calibrate the energy response [61] is summarised below and illustrated in Figure 3.2. The energy calibration procedure differs between converted and unconverted photon candidates.

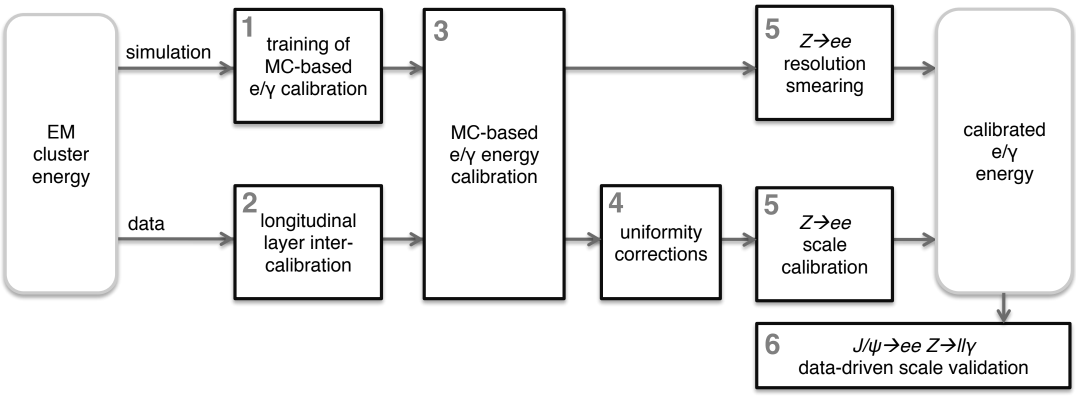


Figure 3.2: Overview of the electron and photon energy calibration procedure. [61]

The EM cluster properties, including its longitudinal development, and additional information from the ATLAS inner tracking system, are calibrated to the original energy in simulated MC samples using multivariate techniques (step 1), which constitutes the core of the MC-based e/γ response calibration (step 3).

The energy scales of the different longitudinal layers in the electromagnetic calorimeter are equalised in data with respect to simulation before the overall energy scale is determined (step 2). The correction from the intercalibration of the first and second calorimeter layers is applied to the energy in layer 2. The presampler energy scale is also determined and applied to data. No intercalibration of the third layer is performed [61].

Multivariate techniques are used to correct for the energy lost in the material upstream of the calorimeter, the energy deposited in the neighbouring cells, the energy lost beyond the LAr calorimeter, etc (step 3). The calibration constants are determined using a multivariate algorithm, with uncalibrated energy, cluster properties and position, shower development and information from inner tracking detector as inputs. A separate optimisation is performed according to bins of $|\eta_{cluster}|$ (the cluster barycentre pseudorapidity in the ATLAS coordinate system), E_T^{calo} (transverse energy measured in the calorimeter) and type (unconverted or converted photon). This MC-based calibration relies on an accurate description of detector geometry and interactions of particles with matter in the MC simulation. The distribution of E_1/E_2, the ratio of energy deposited in the strip layer to the energy in layer 2, in data is used to quantify the amount of detector material upstream of the calorimeter. E_1/E_2 is a measure of the longitudinal shower development. If E_1/E_2 is higher in data compared to simulation, it indicates earlier shower development, which is due to an excess of material compared to the simulation. Modifications in the detector material description in simulation were made based on the amount of material determined from measuring E_1/E_2. The MC-based calibration is applied to the cluster energies in data as well as in MC samples.

Non-uniformities in the calorimeter response are also corrected (step 4). The overall electron response in data is calibrated so that it agrees with the expectation from simulation, using a large sample of $Z \to ee$ events. Energy scale factors are extracted from electrons from $Z \to ee$ decay samples and applied to electron and photon candidates in data (step 5). The energy miscalibration is parameterised as

$$E_i^{data} = E_i^{MC}(1 + \alpha_i) \tag{3.2}$$

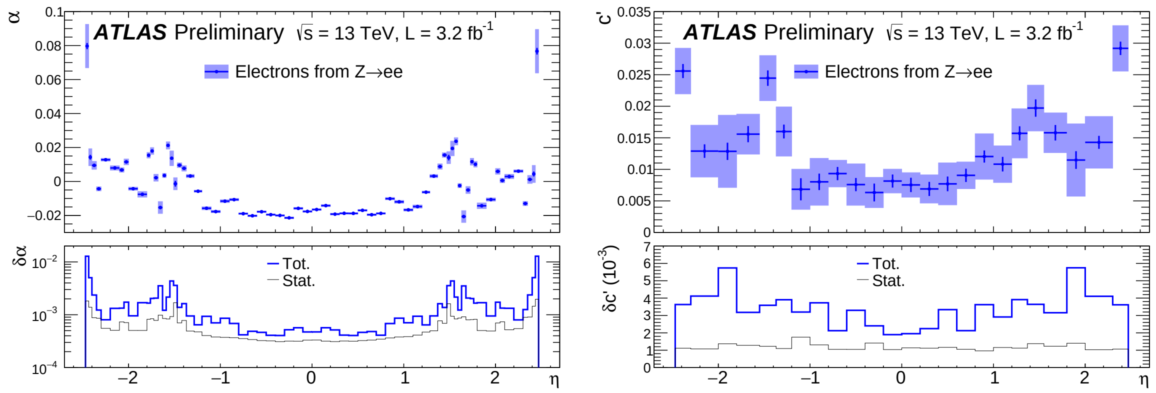
where E_i^{data} and E_i^{MC} are the electron energies in data and simulation, and α_i the departure from optimal calibration, in a given pseudorapidity bin labelled i.

From the $Z \to ee$ events, it is found that the resolution in data is slightly worse than that in simulation. The difference in energy resolution can be modelled by an additional effective constant term (c_i'):

$$\left(\frac{\sigma_E}{E}\right)_i^{data} = \left(\frac{\sigma_E}{E}\right)_i^{MC} \oplus c_i', \tag{3.3}$$

for a given pseudorapidity region labelled i.

The energy scale correction (α_i) and the effective constant term (c_i) are also estimated using $Z \to ee$ events. The results from 2015 data, corresponding to an integrated luminosity of 3.2 fb^{-1}, are shown in Figure 3.3.



(a) (b)

Figure 3.3: (a) Energy scale factor α and (b) additional effective constant term c derived from $Z \to ee$ events using the template method, as a function of η. The error bands on the top plots include both statistical and systematic uncertainties. The statistical and total uncertainties on these quantities are represented by black and blue lines in the bottom plots. [62]

The main sources of uncertainty on the energy scale are from $Z \to ee$ calibration, calibration of the presampler layer, layer intercalibration, layer 2 gain and material. Potential mismodelling of the resolution sampling term, uncertainty in deriving the additional constant term, electronic noise and detector material uncertainty are the main contributors to the resolution uncertainty.

Finally, the calibration is cross-checked using electron candidates from $J/\psi \to ee$ events and photon candidates from $Z \to ll\gamma$ in data (step 6).

In Run 2, improvements have been made to the MC-based calibration with respect to Run 1 using the latest simulation and reconstruction algorithms. In particular, in the region $1.4 < |\eta| < 1.6$, the energy measured by the scintillators in the transition region is used as one of the variable of the multivariate calibration. In this pseudorapidity region, an improvement in energy resolution of about 15% for converted photons and 8% for unconverted photons is achieved [62].

The instantaneous luminosity delivered by the LHC as well as the pile-up has increased in 2016 compared to 2015. The stability of the calibration as a function of μ and of time has been studied. Figure 3.4 shows the ratio of the average invariant mass of the electron pair from $Z \to ee$ event to the average value observed in the whole 2015 data sample as a function of μ and of time. The calorimeter response is stable at the per mille level, lower than the systematic uncertainty of the calibration procedure.

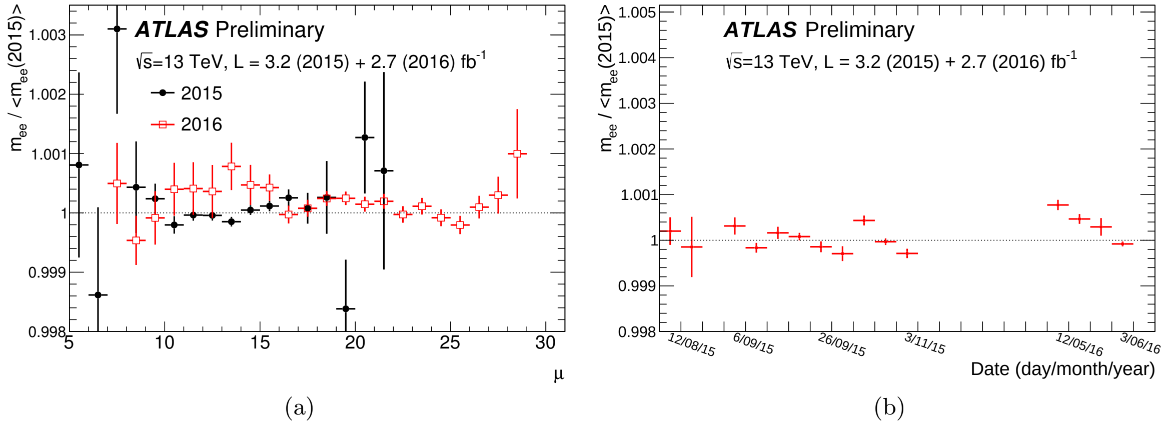


Figure 3.4: Stability of the average invariant mass for electron pair from $Z \to ee$ events (a) as a function of the average number of interactions per bunch crossing, μ, and (b) as a function of the time of data taking using data collected in 2015 and first part of 2016. Each bin shows the average invariant mass divided by the average invariant mass measured in 2015. [62]

3.3 Photon Identification

Photon identification refers to a set of selections based on the EM shower shape that is designed to separate prompt photons from fake candidates, mainly from hadronic decays. Prompt photons refer to photons that are produced directly in the hard scattering process (direct photons) or during final state parton fragmentation (fragmentation photons).

3.3.1 Identification Variables and Selection Definition

The shower deposit in the EM calorimeter by prompt photons is narrower and there is less leakage in the hadronic calorimeter compared to fake photons. The finely segmented strips in the first layer of the EM calorimeter also allow to discern the presence of two peaks in the energy profile of the shower from $\pi^0 \to \gamma\gamma$ decays. A set of variables that characterise the shower development (both lateral and longitudinal) in the EM calorimeter and the shower leakage in the hadronic calorimeter is defined. The variables are listed in Table 3.1. Two cut-based selection working points, *loose* and *tight*, are established.

The *loose* selection is based only on the shower shape in the EM layer 2 and the energy deposited in the hadronic calorimeter. This selection is used by the photon triggers and provides high prompt photon identification efficiency, ranging from 97% for photons with $E_T = 20$ GeV to above 99% for photons with $E_T > 40$ GeV [63]. The *tight* selection uses additionally the information from the strip layer and tighter requirements on the layer 2 and hadronic calorimeter variables to provide maximum background rejection. Different cut values are defined for converted and unconverted photons in the *tight* selection in order to optimise the performance.

The photon shower shape distributions in the ATLAS MC simulation do not perfectly match the ones observed in data. Corrections are thus applied to the simulated values, by applying a shift to each of the variables that need to be corrected [59]. The value of these shifts are optimised separately for unconverted and converted photons, and as a function of the pseudorapidity. Two examples of such a correction on shower shape variables are shown in Figure 3.5 for unconverted photons in $1.52 < |\eta| < 2.37$. The agreement between data and simulation is much improved after the shifts have been applied.

Pile-up broadens the distributions of the discriminating variables and thus reduces the separation between prompt and fake photon candidates. Some variables are more sensitive to pile-up than others. These are usually the variables in which the energy is calculated across many cells.

Type	Name	Description	loose
Hadronic leakage	R_{had_1}	Ratio of the transverse energy in the first sampling layer of the hadronic calorimeter to the transverse energy of the photon EM cluster (used for the range of $\|\eta\| < 0.8$ or $\|\eta\| > 1.37$).	✓
	R_{had}	Ratio of the transverse energy in the hadronic calorimeter to the transverse energy of the photon EM cluster (used in the range $0.8 < \|\eta\| < 1.37$).	✓
EM layer 2	R_η	Ratio of energy deposited in 3x7 (ηxϕ) versus 7x7 cells.	✓
	$w_{\eta 2}$	Lateral width of the shower in η, over a region of 3x5 cells in $\Delta\eta \times \Delta\phi$ around the centre of the photon cluster.	✓
	R_ϕ	Ratio of energy deposited in 3x3 versus 3x7 cells.	
EM layer 1	w_{s3}	Shower width calculated from three strips around the strip with maximum energy deposit.	
	w_{stot}	Total lateral shower width, measured over all strips in a region of $\delta\eta \times \delta\phi = 0.0625 \times 0.2$ (20×2 strips).	
	F_{side}	Energy outside the 3 central strips but within 7 strips divided by energy within the 3 central strips.	
	ΔE	Difference between the energy of the strip that has the second greatest energy and the energy of the strip with the least energy found between the strips with the greatest and the second-greatest energies.	
	E_{ratio}	Ratio of the energy difference between the largest and the second largest energy deposits to the sum of both energies.	

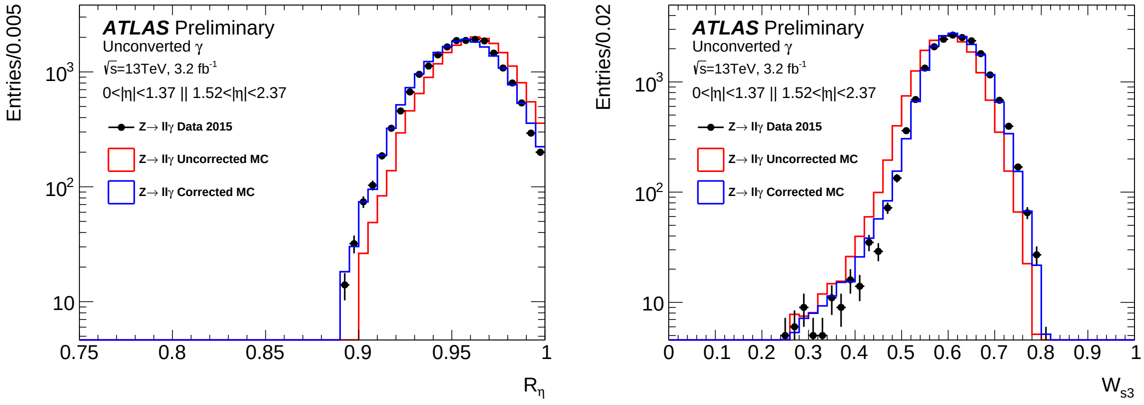
Table 3.1: Discriminating variables used for tight photon identification. The variables used in loose identification selection are also marked.

The cuts on the discriminating variables can be tuned according to pile-up conditions expected in order to maintain the performance of photon identification.

3.3.2 Identification Efficiencies

Photon identification efficiencies are measured using three data-driven techniques: using a high purity sample of radiative Z decays; extrapolating photon shower shape properties from electrons in a $Z \to ee$ sample; and measuring the photon purity in an inclusive photon sample using isolation properties (referred to as matrix method). All three measurements use photon candidates within the pseudorapidity acceptance ($|\eta| < 1.37$ and $1.52 < |\eta| < 2.37$).

For the final 2012 dataset, photon candidates in the identification studies are also required to be isolated with E_T^{iso} in a cone of size $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ lower than 4 GeV (see next Section for definition). The tight photon identification efficiency is measured to be 50-65% (45-55%) for unconverted (converted) photons with $E_T = 10$ GeV and up to 95-100% at photon $E_T > 100$ GeV. For the typical photon energies ($E_T > 40$ GeV) that are relevant for the high



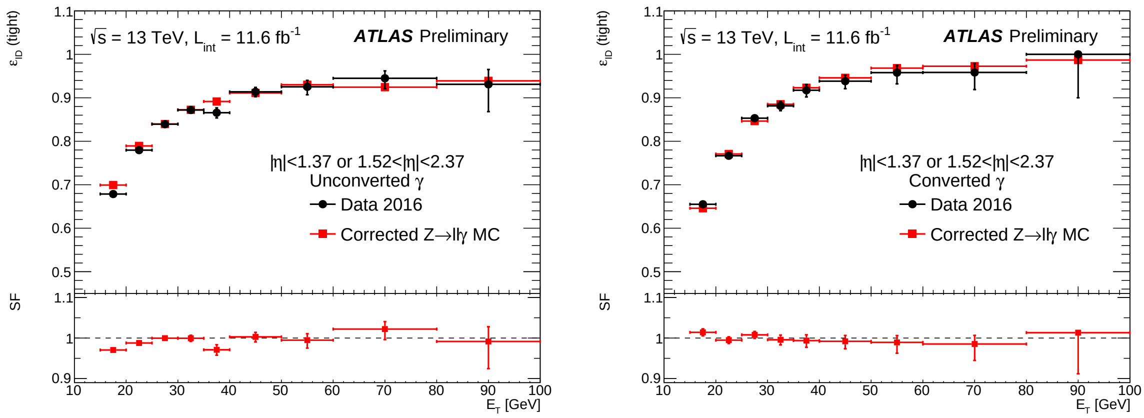
(a) (b)

Figure 3.5: Distributions of two discriminating variables (R_η and w_{s3}) for unconverted photons in $1.52 < |\eta| < 2.37$. The black points are for the photons from radiative Z decay events in 2015 data, the red histogram for the photons from radiative Z events in simulation, and the blue histogram after applying the shift to the simulated values [64].

mass resonance searches reported in this thesis ($A \to Zh$ with $h \to \gamma\gamma$ and $X \to \gamma\gamma$), the tight photon identification efficiency is higher than 90% [59].

Photon identification cuts have been re-optimised in Run 2. Photon identification efficiencies have also been measured in 2015 data. In this study, the photon candidates are required to pass $E_T^{iso} < 0.065 \cdot E_T$ and $p_T^{iso} < 0.05 \cdot E_T$, in a cone with an opening $\Delta R < 0.2$ (see next Section for definition of p_T^{iso}. The tight selection efficiencies range from 53-64% (47-61%) for unconverted (converted) photons with $E_T = 10$ GeV to 88-92% (96-98%) at photon $E_T > 100$ GeV [64].

The tight identification efficiency has also been measured using only the first method of radiative Z boson decays on 11.6 fb^{-1} of data recorded in 2016. The comparison of data-driven tight identification efficiency measurements to simulation as a function of E_T of the photon is shown in Figure 3.6.



(a) (b)

Figure 3.6: Measured tight identification efficiency in 2016 data compared to MC as a function of photon E_T using radiative Z boson decays for $|\eta| < 1.37$ and $1.52 < |\eta| < 2.37$. The bottom panel shows the ratio of the data-driven results to the MC predictions, namely scale factor. Only the statistical uncertainties are shown [65].

In all of the measurements, no significant differences are observed between the data-driven efficiency measurements and the simulations. The difference is taken into account by computing data-to-MC efficiency ratios, also referred to as scale factors (SF). Most SF values are close to unity. In analyses with photons in final state, the combined scale factors are then applied to simulated photons as weights to provide an even better description of the identification efficiency.

3.4 Photon Isolation

Isolation properties of photon candidates can provide further separation of prompt photons from fake photons. Prompt photons are usually isolated, with not much energy activity around them. Isolation variables are defined from the calorimetric deposits or the tracking information provided by the inner detector, around the photon candidate. The track-based isolation variable is computed by summing the transverse momentum of tracks with $p_T > 1$ GeV that originate from the primary vertex within a cone centred around the photon direction. For converted photons, the one or two tracks associated with the conversion are excluded from the track isolation computation.[1]

The calorimeter-based isolation variable is computed as the scalar sum of the transverse energy of all topological clusters [55] within a cone around the photon position, removing the energy from the photon cluster itself. Topological clusters are formed starting from cells with energy greater than 500 MeV and with signal-to-noise ratio (S/N) larger than 4, then adding iteratively surrounding cells with a $S/N > 2$ to the cluster. Figure 3.7 illustrates the computation of the calorimeter isolation variables, where all the topological clusters with barycentre within the isolation cone are included while the photon cluster energy is subtracted. The expected photon energy deposit outside the rectangular cluster and contributions from pile-up and the underlying event are also subtracted in the computation.

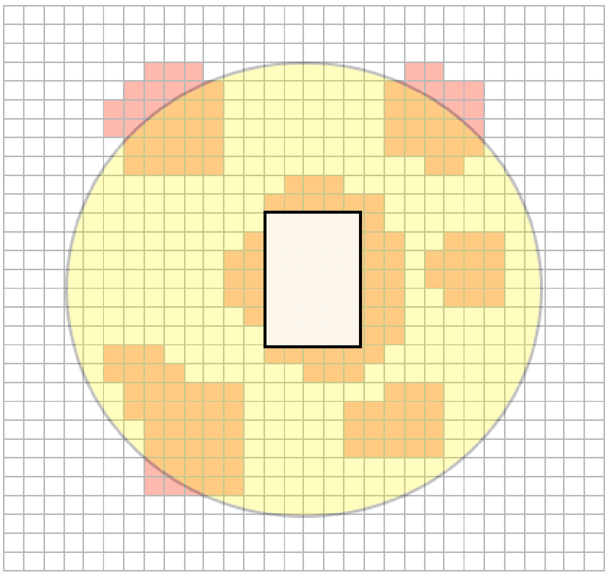


Figure 3.7: Schema of the photon calorimeter isolation. The grid represents the calorimeter cells in the η and ϕ direction. The photon is located at the centre of the yellow cone representing the isolation cone. All topological clusters (represented in red) with barycentre inside the isolation cone are considered in the isolation computation. The rectangle represents the core energy of the photon candidate that is subtracted from the isolation computation [66].

Cut values on track and calorimeter isolation variables are analysis-dependent. For instance, in the high mass diphoton resonance search at $\sqrt{s} = 13$ TeV, the track isolation in a cone of

[1]From 2016, all the tracks associated to reconstructed photon conversion vertices are excluded from the p_T^{iso} computation.

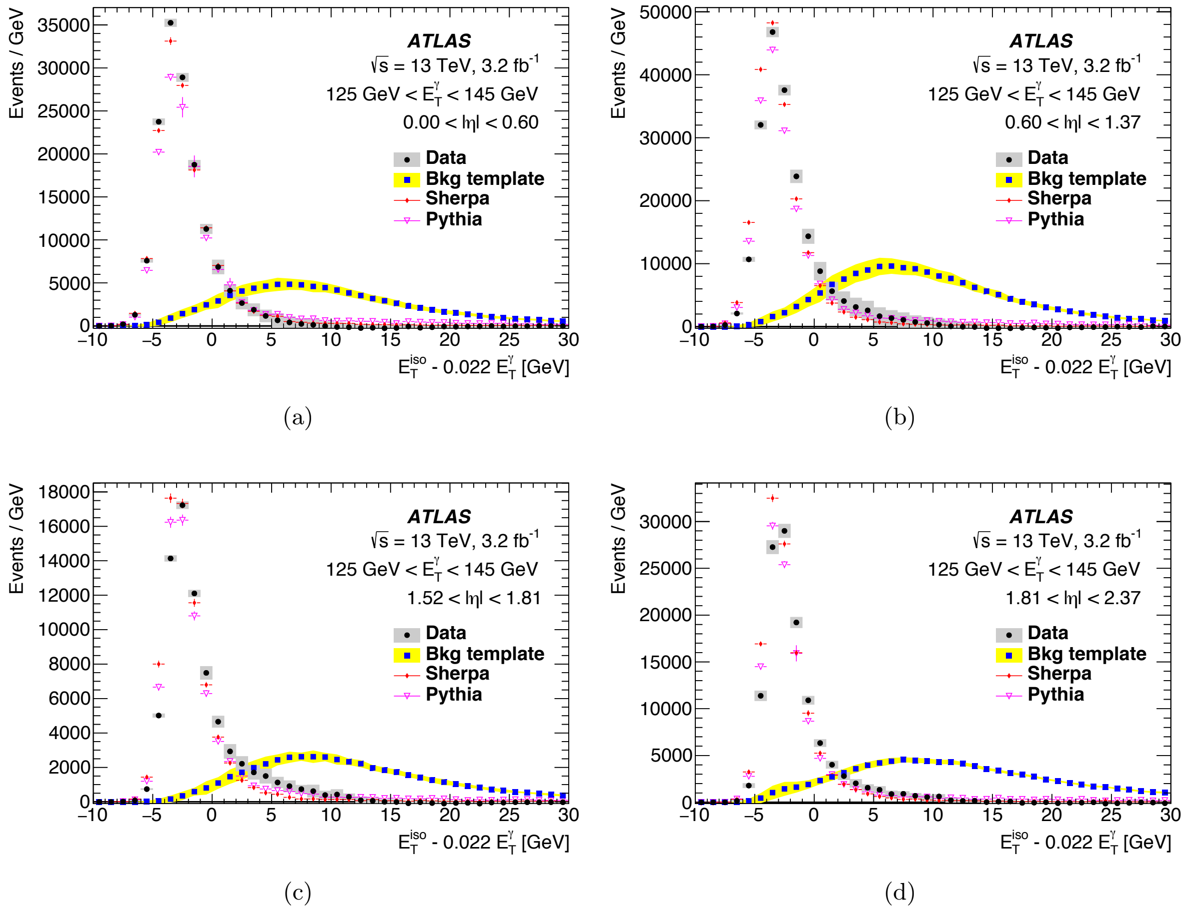
size $\Delta R = 0.2$ is required to be below $0.05 \times E_T$ GeV, and calorimeter isolation energy in a cone of size $\Delta R = 0.4$ is required to be smaller than $0.022 \times E_T + 2.45$ GeV.

Inclusive photon samples generated with Sherpa [67] and Pythia8 [68] are used to check the efficiency of the isolation requirement in a wide E_T range from 50 GeV up to 1000 GeV. The isolation efficiency is measured in data by fitting the isolation distribution in data with signal and background photon isolation template. The distribution of the isolation energy for fake photons is obtained from a data control region by reverting some photon identification criteria while the distribution for signal photons is take from simulation.

For the calorimeter-based isolation, the background contribution is subtracted from data. The resulting distributions are compared with simulation to obtain a shift of the calorimetric isolation. This data-driven shift is obtained from the difference in the peak position of the function parameterising the calorimeter isolation distribution in data and that of the inclusive photon simulated samples. This shift is generally a few hundred MeV, depending on p_T and η of the photon candidates. The uncertainty of the calorimeter-based isolation efficiency is taken from the difference between the nominal value and that from applying a data-driven shift to the isolation energy in simulation. Figure 3.8 shows the distributions of the calorimeter isolation variable $(E_T^{iso} - 0.022E_T(\gamma))$ in 2015 data compared to simulations for photon candidates with 125 GeV $< E_T <$145 GeV and fulfilling the tight identification criteria.

The track isolation efficiency is measured in data by fitting the isolation distribution in data with signal and background photon isolation template. The residual disagreement between the fit model and the data is used to derive the uncertainty in the isolation efficiency.

The calorimeter- and track-based isolation distributions in the data are very similar to the Monte Carlo, resulting in scale factor within 1 or 2% of 1. The efficiency of the isolation requirement used in high mas diphoton resonance search for photons fulfilling tight identification selection in signal MC samples is 90% to 96% in the E_T range 100 GeV to 500 GeV, with an uncertainty between 1% and 2% [69]. The isolation requirement reduces the rate at which jets are misidentified as photons by about one order of magnitude.



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

Figure 3.8: Distributions of the calorimeter isolation variable ($E_T^{iso} - 0.022E_T(\gamma)$) for photon candidates fulfilling the tight identification criteria for 125 GeV $< E_T <$145 GeV and four η regions. The background contribution to the 2015 data, shown as 'Bkg template', has been subtracted. It has been determined using a control sample with a subset of the identification requirements inverted and normalized to the data in the region $E_T^{iso} - 0.022E_T(\gamma) > 12$ GeV. The 2015 data distributions are compared to predictions from simulation using either Sherpa or Pythia8 to generate inclusive photon events. The error bars represent the statistical uncertainties. The bands around the background and the subtracted data distributions represent the estimate of the systematic uncertainties on the background estimate [69].

Chapter 4

Photon Conversion Classification and Measurement Using Shower Information

As described in the previous chapter, the classification of a photon candidate as a converted or an unconverted photon is based on tracking and vertex fitting algorithms. The cluster sizes are different for converted and unconverted photons in Run 1. As energy calibration and photon identification requirements are optimised for converted and unconverted photons separately, misclassification can have a significant impact on physics analysis. The impact is expected to be reduced at Run 2, since cluster sizes are no longer different for converted and unconverted photons. However, the performance of tracking algorithm degrades with the pile-up rate, which is expected to increase at Run2, due to mismatching of pile-up tracks to clusters and the presence of fake tracks.

This chapter documents the task I have completed in order to qualify as an ATLAS author. The studies were performed between February 2014 and July 2015 and therefore are based on Run 1 data. They are documented in an ATLAS Internal Note (Ref. [70]) of which I am the main author. Based on an earlier preliminary study, I have investigated the use of electromagnetic shower variables to either improve the present conversion classification algorithm or to evaluate its performance on data.

Section 4.1 gives a summary of the data and MC samples used in these studies. The choice of the shower shape variables sensitive to the photon conversion status is discussed in Section 4.2. I then compare the performance of the alternative classification to the well established track-based method that has been in use since Run 1 in Section 4.3. Finally, I apply a similar EM shower shape-based method to measure the conversion fractions in 2012 data using a template fit method, as described in Section 4.4.

4.1 Data and MC samples

The study presented in this chapter addresses the issue of a potential deterioration of the track-based conversion reconstruction due to the expected pile-up increase at Run 2. However, since it was performed before the start of Run 2, pp collision data collected in 2012 at $\sqrt{s} = 8$ TeV is used, corresponding to an integrated luminosity of 20.7 fb^{-1}. Sources of relatively pure photons include $Z \to ll\gamma$ and diphoton events.

The $Z \to ll\gamma$ decay refers to a decay of the Z boson into an electron or a muon pair with the radiation of a photon in the final state. It allows the selection of a very pure and unbiased source of photons by requiring cuts on the dilepton and dilepton+photon invariant masses. In the following, only the $Z \to \mu\mu\gamma$ decay is used. The events are selected with the muon triggers and must contain two like-flavor, opposite-charge leptons. The invariant mass of the $\mu\mu\gamma$ system must be consistent with the mass of the Z boson. A cut of $80 < m_{\mu\mu\gamma} < 96$ GeV is applied to select final state radiation (FSR) events. The photons are required to have $p_T > 10$ GeV and pass *loose* identification criteria. Since shower shape variables are used in this study, tight identification requirements are not applied since the requirements differ for converted and unconverted photons. Details on the selection criteria are described in Ref. [59]. In 2012 data, there are about 73 thousand selected $Z \to \mu\mu\gamma$ candidate events. The photons from such decay have typically low transverse momenta, with most of the photons having $10 < p_T < 50$ GeV. The selected $Z \to \mu\mu\gamma$ sample has a high purity of around 96%.

High-p_T photon candidates are also a relatively pure source of photons. In order to increase the purity, one can select diphoton events where one of the photon passes the tight selection and the other photon, passing loose identification, is used to probe the conversion classification. The data sample is recorded with a diphoton trigger, `g35_loose_g25_loose` with 35 GeV and 25 GeV p_T thresholds on the leading-p_T and subleading-p_T photon candidates. In the high-level trigger, clusters of energy in the electromagnetic calorimeter are reconstructed and required to satisfy the loose shower shape criteria. Events are selected if the leading (subleading) photon has a transverse momentum greater than 70 (50) GeV. The purity of the selected sample has not been measured but it is expected to be below 80

A number of different MC samples are used in this study, to study the effects of pile-up or to provide a description of the data. MC event generators are used to produce events at the particle level, which are then passed through a Geant4 [71] simulation of the ATLAS detector and reconstructed using the same analysis chain used for the collision data. They are listed in Table 4.1.

Generator	$\sqrt{s}$	Process	N_{events}	Pile-up	Remarks
ParticleGenerator	14 TeV	Single photon	1 million	0 40 60 80	flat E_T spectrum between 7 and 80 GeV.
ParticleGenerator	8 TeV	Single photon	40 million	0	E_T up to 3 TeV.
Sherpa	8 TeV	$Z \to \mu\mu\gamma$	5 million 1 million 1 million 1 million	2012 condition	nominal geometry geometry config. A geometry config. E'+L' geometry config. F'+M+X
Sherpa	8 TeV	$Z \to \mu\mu$	10 million	2012 condition	

Table 4.1: MC samples used in this study.

Single photon samples simulated at four different pile-up conditions, characterised by μ, are used to study the effects of pile-up. These samples are produced during the Long Shutdown 1 (LS1) period with $\sqrt{s} = 14$ TeV and using expected Run 2 conditions, i.e. with IBL inserted and with a bunch spacing of 25 ns. The events are generated with a flat transverse momentum spectrum between 7 and 80 GeV.

Another single photon sample with a wide range of p_T (up to 3 TeV) and high statistics is used. This sample was simulated for the MVA calibration [61] at $\sqrt{s} = 8$ TeV and zero pile-up. It can be used for validation with 2012 data.

Selection criteria applied on the photon candidates in the single photon samples are:

- loose identification cuts,
- pseudorapidity within $|\eta| < 2.37$, excluding $1.37 < |\eta| < 1.52$, and

- the reconstructed photon matches a true photon.[1]

Simulated $Z \rightarrow \mu\mu\gamma$ samples are used to study the $Z \rightarrow \mu\mu\gamma$ events in 2012 data. Hence, the samples are simulated at $\sqrt{s} = 8$ TeV with 2012 pile-up conditions. All samples mentioned so far are simulated with the final Run 1 geometry, named ATLAS-GEO-21-02-02. Additionally, $Z \rightarrow \mu\mu\gamma$ samples using distorted geometries are also used to study the effect of different amount of material upstream of the electromagnetic calorimeter. Distorted geometry is often similar to the nominal one, which has the best estimation of the material distributions, but with additional material added in certain parts of the detector. The amount of additional material in the distorted geometry samples used in this study and their location is shown in Table 4.2. Furthermore, nominal geometry $Z \rightarrow \mu\mu$ sample is used for studying the background photons in selected $Z \rightarrow \mu\mu\gamma$ events in data. The $Z \rightarrow \mu\mu\gamma$ and $Z \rightarrow \mu\mu$ samples are generated with the Sherpa generator using the CT10 PDF set.

Configuration	ID	SCT/TRT EC	ID end-plate	Cryostat 1	PS/S1	B-EC transition
A	5%					
E'+L'		7.5% X_0		5% X_0		
F'+M+X			7.5% X_0		5% X_0	50%

Table 4.2: Fraction of additional material or radiation lengths (X_0) added in various places of the detectors, for the different MC geometry configurations. The different places are in the whole Inner Detector (ID), at the end of SCT/TRT end-caps (SCT/TRT EC), in the Inner Detector end-plate (ID end-plate), radially in barrel cryostat before the calorimeter (Cryostat 1), radially between the presampler and the strips in barrel (PS/S1) and in the calorimeter barrel/end-cap transition region (B-EC transition).

Shifts, also known as fudge factors, are applied on the shower shape variables in MC samples that are also used in photon identification ($w_{\eta 2}$, R_ϕ, F_{side} and w_{s3}) to improve the agreement with the data shower shape variables, as explained in Section 3.3. For comparison between simulated samples where no real data is involved, such as those in Section 4.3.1, these shifts do not need to be applied.

4.2 Conversion Classification

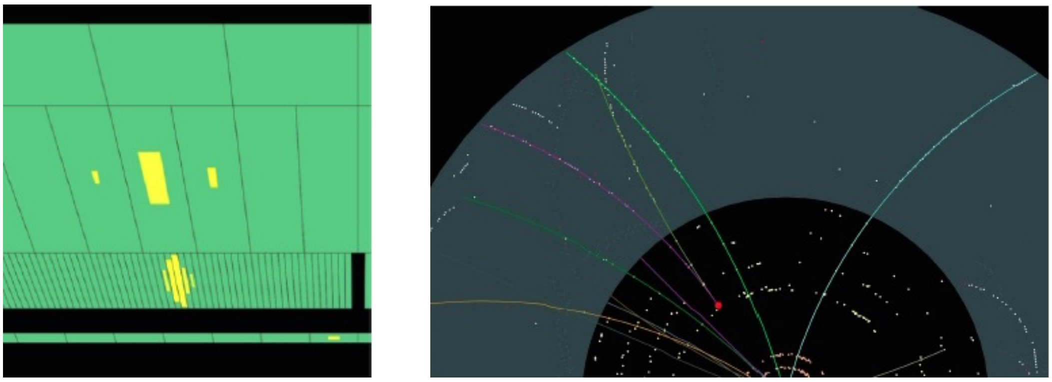
The classification of photon candidates as converted or unconverted using information on the electromagnetic shower shape is described in this section. A boosted decision tree (BDT) is trained from a number of discriminating variables that characterise the shower profile of a photon candidate, known as shower shape variables.

4.2.1 Discriminating Variables

In photon identification, selection criteria are applied on the shower profile of the photon candidate to distinguish true photons from hadrons. Similarly, one could use shower shape variables to separate converted from unconverted photons. The EM shower shapes exhibit some differences between converted and unconverted photons, which are largely due to the earlier shower development for converted photons (being electrons when interacting with the EM calorimeter), and the separation of electron-positron pairs under the influence of the magnetic field. Therefore, shower shape variables that describe the depth or the width in ϕ of the electromagnetic shower are expected to be discriminant. The typical shower shape of an unconverted photon

[1]This is based on ΔR matching between the cluster and the extrapolated point of impact of the true particle to the calorimeter.

and the tracks left in the inner detector by a photon converting into an electron-positron pair are shown in Figure 4.1.



(a) (b)

Figure 4.1: (a) Shower of an unconverted photon in the electromagnetic calorimeter and (b) tracks in the inner detector left by a photon converting into an electron-positron pair.

Various shower shape variables are explored in this study, each with varying degree of discrimination. Eight variables are chosen as further additions do not bring further discrimination in a significant manner. Descriptions of the chosen variables are presented in Table 4.3.

Type	Layer	Name	Description
Lateral	Layer 2	$w_{\eta 2}$	See Table 3.1.
		R_ϕ	See Table 3.1.
		F_{side}	See Table 3.1.
	Layer 1	w_{s3}	See Table 3.1.
		E_{maxs1}	Energy of strip with maximal energy deposit.
Longitudinal	Mixture	f_1	Fraction of energy in the strip layer.
		E_1/E_2	Ratio of energy deposited in the strip layer to the energy in layer 2.
		E_0/E_{acc}	Ratio of energy in the presampler to the energy in the accordion layers (layer 1, 2 and 3.)

Table 4.3: Shower shape variables that offer discrimination between converted and unconverted photons.

The distribution of the shower shape variables being used and their correlations are presented in Figures 4.2 and 4.3 for converted and unconverted photons. To use tracking information as well as shower shape in identifying photon conversion, the numbers of hits in SCT and in TRT from the associated tracks are added to the list of input variables.[2]

4.2.2 Choice of the Multivariate Methods

The eight shower shape variables defined in Table 4.3 are input variables to the multivariate classification. The Toolkit for Multivariate Analysis (TMVA) [72] framework is used to implement the multivariate techniques for conversion classification. A typical classification example is to determine whether an event comes from a signal or background process, and in this study, we

[2]There are a few other variables related to tracking information in conversion reconstruction, but it was found that the number of hits in SCT and in TRT already contain sufficient information.

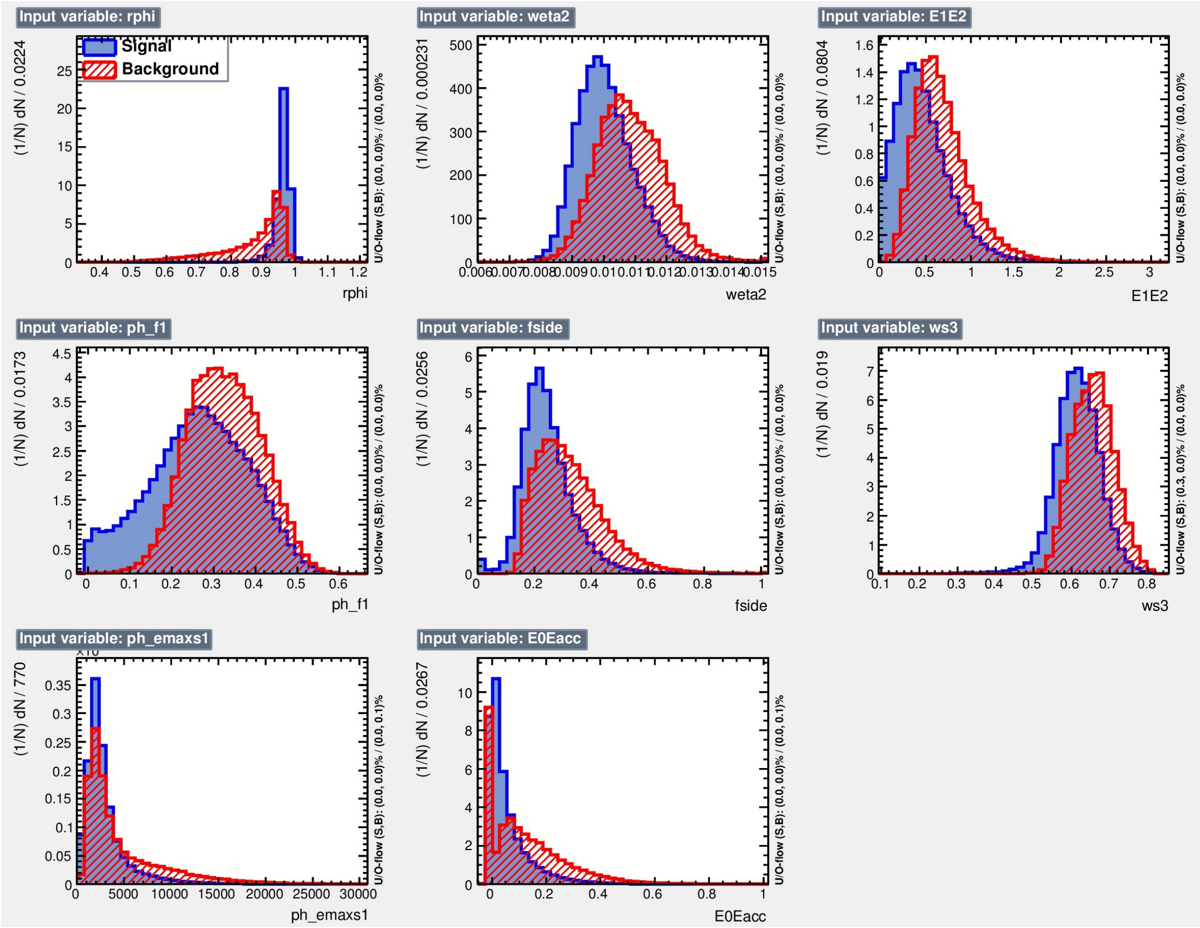
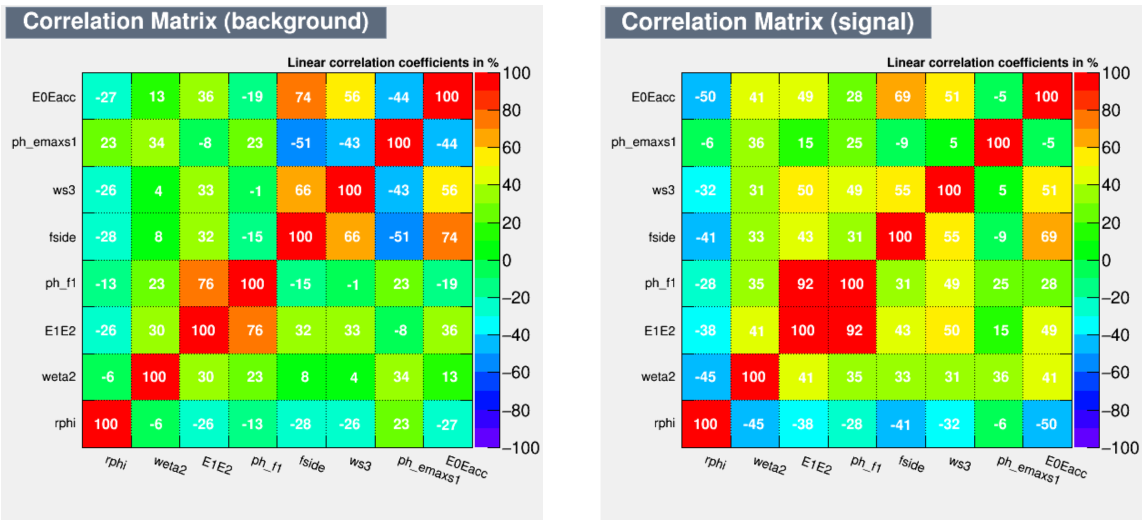


Figure 4.2: Distributions of the shower shape variables for unconverted (signal) and converted (background) photons in the $Z \to \mu\mu\gamma$ MC sample.



(a) (b)

Figure 4.3: Correlation matrices of the shower shape variables for (a) converted photons and (b) unconverted photons.

1119 are interested in classifying a photon as converted or unconverted. A true unconverted photon
1120 is arbitrarily defined as signal and true converted photon as background.

A typical TMVA classification or regression analysis consists of two independent phases: the training phase, where the multivariate methods are trained, tested and evaluated, and an application phase, where the chosen methods are applied to the classification or regression problem they have been trained for. In the training phase, the sample is divided into two subsamples of equal size, one is used for training and the other as a test sample. This is to check for overtraining or other possible biases.

Numerous multivariate methods are available in TMVA, several of them are tested for conversion classification: the computation of a likelihood, a Fisher discriminant and boosted decision trees (BDT). A decision tree is a binary tree-structured classifier similar to the one sketched in Figure 4.4. Repeated left/right (yes/no) decisions are taken on one single variable at a time until a stop criterion is fulfilled. The phase space is split this way into many regions that are eventually classified as signal or background. We use a BDT classifier with two types of boosting: adaptive boost (AdaBoost) and gradient boost. The former is denoted BDT and the latter BDTG. Boosted decision trees require little tuning to obtain reasonably good results and are insensitive to the inclusion of poorly discriminating variables.

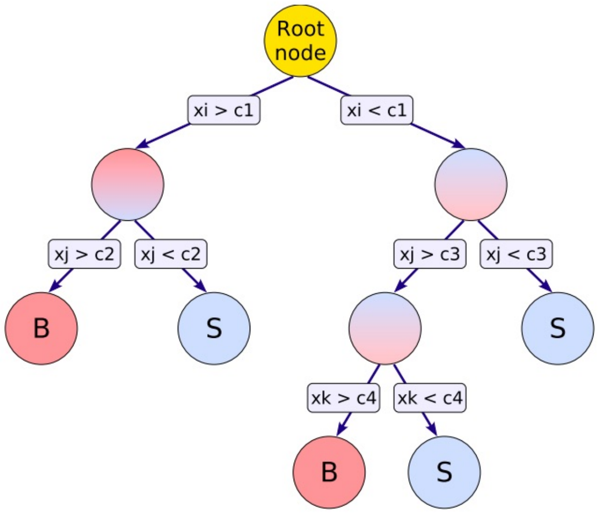


Figure 4.4: Schematic view of a decision tree.

Figure 4.5 shows the ROC curve (background rejection factor versus signal classification efficiency) of the four multivariate methods with the shower shape variables in Table 4.3 as input. The curve is created by varying the cut on the classifier output in order to scan the background rejection and signal classification efficiency. The best performance is achieved with the BDT and BDTG methods. Hence, only these two methods are retained in the following.

4.2.3 Categorisation

The probability of a photon to convert increases with the amount of material it traverses, which varies with η, as shown in Figure 4.6. In order to improve the performance of the multivariate method, one can take into account the different probability of a photon to convert at different η. This is achieved by using categorised classifiers, where the training sample is split into disjoint categories based on η. Several classifiers are trained independently, one in each category.

Seven $|\eta|$ categories are defined according to the variation of the amount of material: $|\eta| < 0.6$, $0.6 < |\eta| < 0.8$, $0.8 < |\eta| < 1.15$, $1.15 < |\eta| < 1.37$, $1.52 < |\eta| < 1.81$, $1.81 < |\eta| < 2.01$ and $2.01 < |\eta| < 2.37$.

The distributions of BDT and BDTG with $|\eta|$ categories from training and test samples superimposed are shown in Figure 4.7. The distributions agree well between the training and the

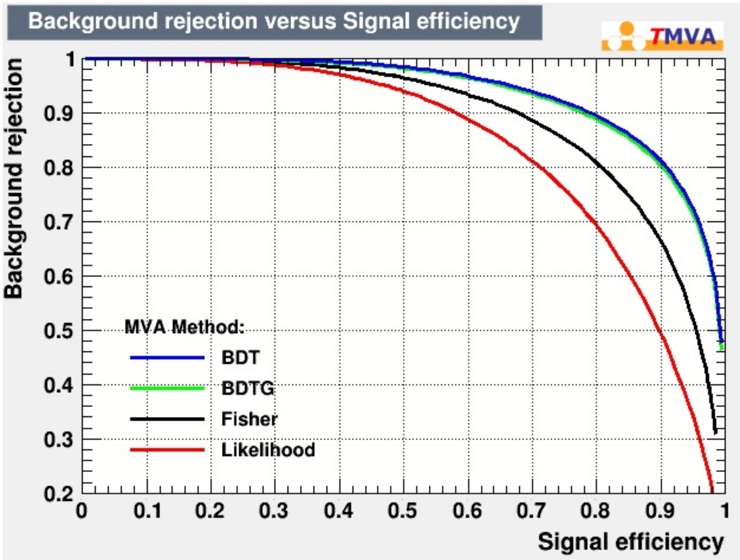


Figure 4.5: Background rejection versus signal efficiency for the four methods being tested.

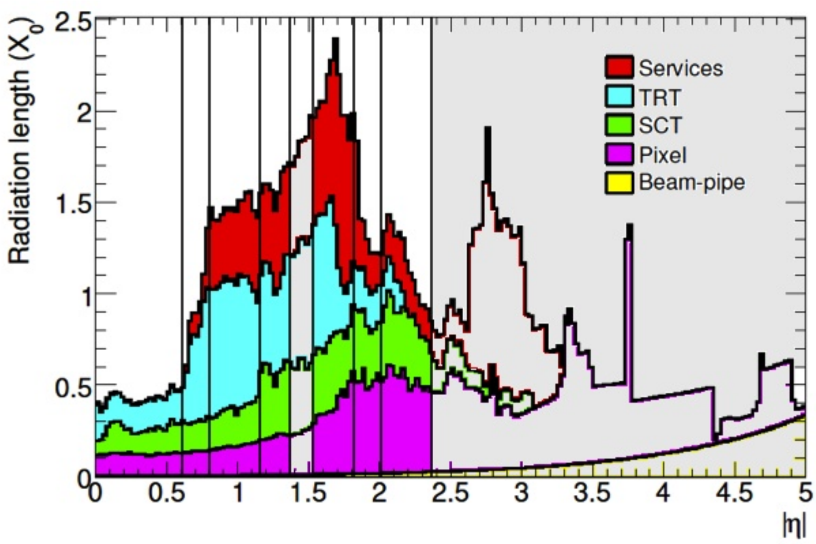
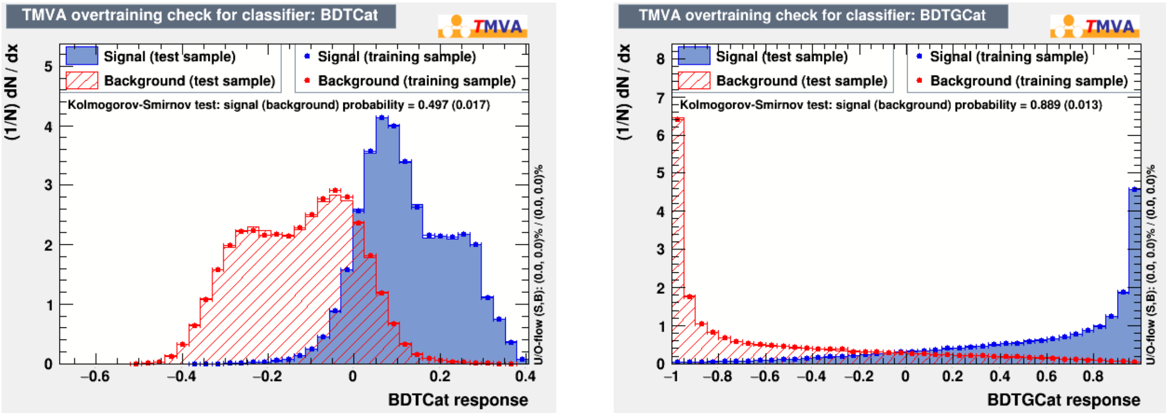


Figure 4.6: Amount of material in terms of radiation length in the inner detector as a function of $|\eta|$. The vertical lines drawn here indicate the bin boundaries chosen and the gray shaded areas correspond to the $|\eta|$ region commonly excluded in analyses involving photons.

test samples. Since the two independent samples have identical performance, we can conclude that there is no overtraining. In the figure, the BDTG distributions are smooth whereas the BDT distributions show double peak structures for both converted and unconverted photons. The double peak seen in the converted photon distribution is due to two populations of converted photons, one with large value of R_ϕ and the other with low value. R_ϕ is one of the most discriminant variable in the BDT. As seen in Figure 4.2, it is ambiguous if a photon is converted or not if the value of R_ϕ is large, between 0.9 and 1. The case is similar for the BDT distribution of unconverted photons, but it is due to a combination of variables, most notably E_1/E_2 and f_1. Further information about the composition of the BDT distributions can be found in the next section.

The improvements in the background rejection and signal efficiency by introducing $|\eta|$ categories are illustrated in Figure 4.8. The BDT with adaptive boost performs slightly better



(a) (b)

Figure 4.7: Distribution of the output variable of the (a) BDTCat and (b) BDTGCat classifiers. The histogram shows the distributions of background/converted (red) and signal/unconverted (blue) from the training sample whereas the points with error bars correspond to the distribution in the test sample.

than BDTG (gradient boost) but its irregular shape may introduce some difficulties in the fitting procedure presented in Section 4.4. Thus, both boostings are retained, even though no significant difference between the two is observed aside from the shape of their distributions.

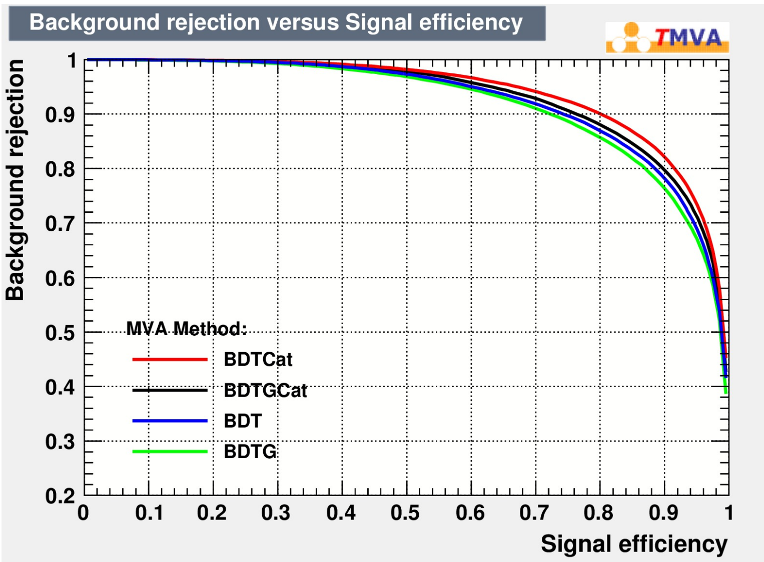


Figure 4.8: Background rejection versus signal efficiency for BDTCat (BDT with $|\eta|$ categories), BDTGCat (BDTG with $|\eta|$ categories), BDT and BDTG.

In the following, the categorised boosted decision tree classifiers are always used. The BDT with adaptive boost and $|\eta|$ categories (BDTCat) is used in Section 4.3 whereas the BDT with gradient boost and $|\eta|$ categories (BDTGCat) is used in Section 4.4 since the former performs slightly better and the latter has a more regular shape. BDTCat is shortened to BDT and BDTGCat to BDTG in the text, but occasionally the variables with suffix 'Cat' are also used, notably in the figures.

Another categorisation is based on the photon transverse momentum p_T as shower shape variables depend strongly on both η and p_T. In this study, the p_T binning is used only in Section 4.3.3.2 where a simulated single photon sample is used in the training and the classification is applied to diphoton events in data, which has different kinematics to the training sample. An advantage of introducing p_T binning along with η binning is that the classification trained can be applied to photons from any type of physics processes in data. The single photon sample with high statistics, zero pile-up and p_T up to 3 TeV is used to train this BDT. Nine categories are defined, covering a whole range of p_T, with the bin edges listed below:

- 0, 10, 20, 40, 60, 80, 120, 500, 1000, 5000 GeV.

In addition, the sample is split in the 7 $|\eta|$ categories as defined in Section 4.2.3, leading to a total of 7×9 categories.

The BDT trained on this single photon Monte Carlo sample with p_T and η categories has been applied to data for validation. This is performed for both radiative $Z \to \mu\mu\gamma$ decays and diphoton events. This will be demonstrated in Section 4.3.3.2 on diphoton events in data.

In the study where the MC training sample has the exact kinematics as the simulation or data to which the classification is applied on, the p_T categorisation is not used in order to maximise the statistics available for training.

4.2.4 Sensitivity of Shower Shape Variables to the True Radius of Conversion

The later a conversion occurs (i.e. the higher the conversion radius), the less able the classification is to separate conversion as it resembles unconverted photon more. As depicted in Figure 4.9, the BDT output value varies as a function of the radius of conversion, R_{conv}, which is defined as the distance of the conversion vertex to the beam axis. The radii chosen for the figure approximate the different parts of the inner detector: the Pixel detector at $R < 200$ mm and the TRT at $550 < R < 800$ mm.

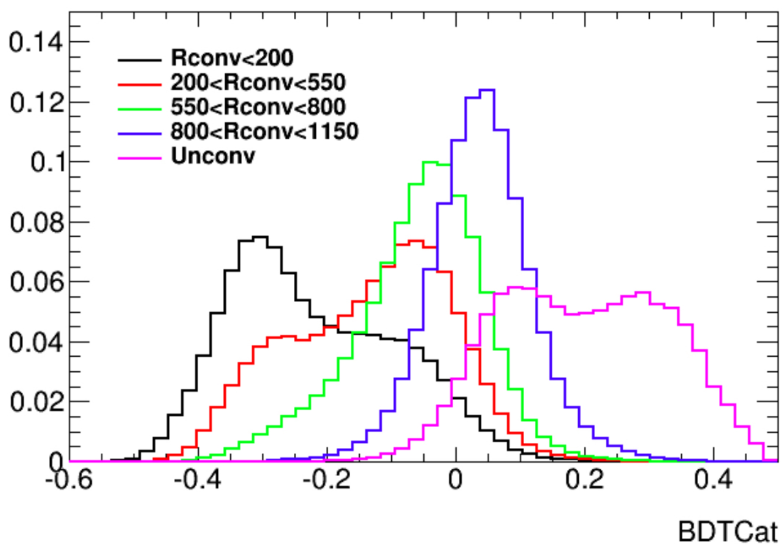


Figure 4.9: BDT output distributions for different true radii of conversion in millimetre. The curves have been normalised to unity. The double peak structure seen in early conversions and for photons that did not convert before the electromagnetic calorimeter are explained in Section 4.2.3.

The top plot in Figure 4.10 shows the BDTG output value at each true conversion vertex, which corresponds to the position of material in the inner detector region (bottom plot). The BDTG is used instead of BDT for visualisation purpose as it has a more regular shape with early conversions that peak at -1 and unconverted photons at +1. From this plot, one can clearly see that the different classifier outputs correspond to different regions of the inner detector.

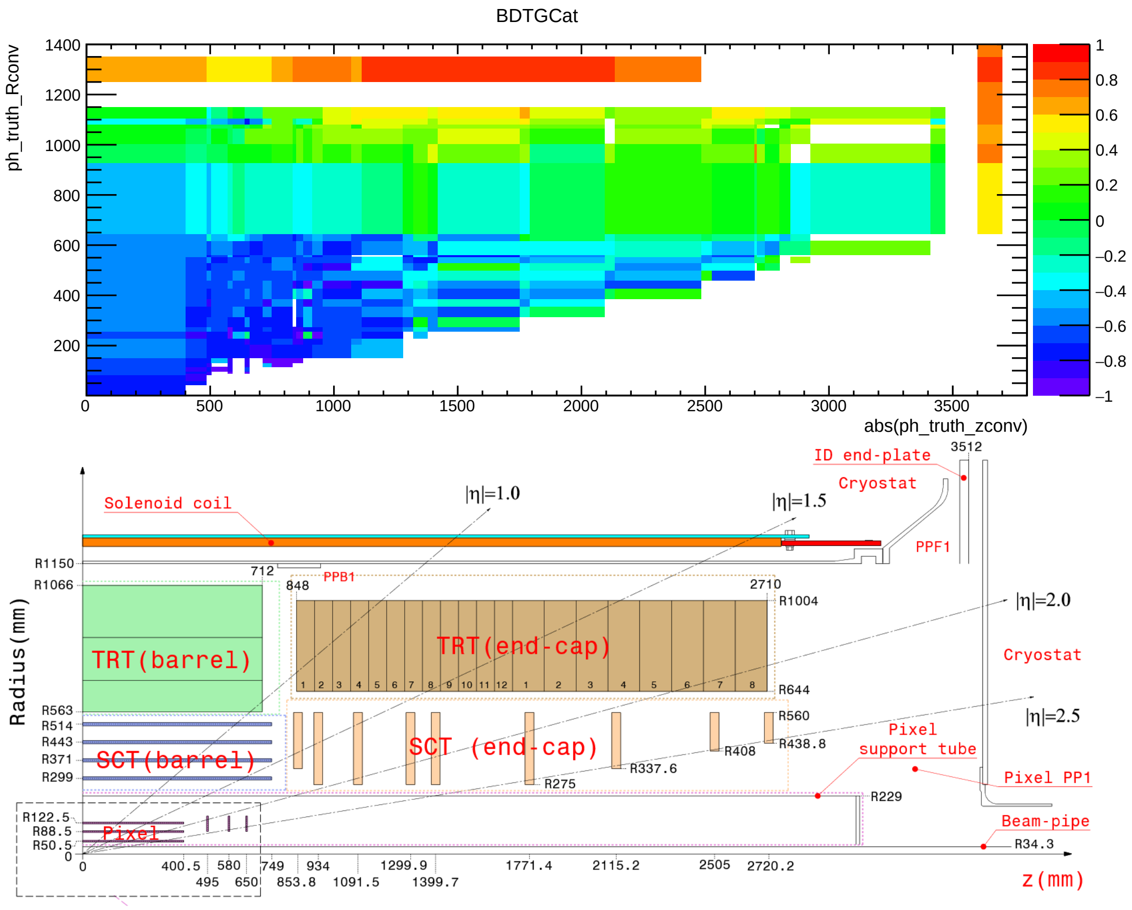


Figure 4.10: (Top) BDTG output in a plot of true radius of conversion against z position. The BDTG output of true unconverted photons are plotted at a radius of 1300 mm for barrel and at a z position of 3650 mm for end-cap. The position of the unconverted photons are determined from interpolating their cluster position to the origin. (Bottom) Diagram of the inner detector.

Since the combination of shower shape variables is sensitive to the radius of conversion, this study could be extended in the future as a tool to estimate the amount of material in the ATLAS inner detector.

4.3 Classification Performance and Validation

In this section, the performance of the conversion classification developed in the previous section is presented along with its validation using data. The BDT with adaptive boost is used here but using BDTG would lead to very similar results and the same conclusions.

Two figures of merit are used to measure the performance of the various conversion classification methods: the conversion reconstruction efficiency and the fake conversion rate, defined in Eq. 4.1 and 4.2:

$$\text{Efficiency} = \frac{N(true\,conv \cap reco\,conv)}{N(true\,conv)}, \tag{4.1}$$

$$\text{Fake conversion rate} = \frac{N(true\,unconv \cap reco\,conv)}{N(true\,unconv)}. \tag{4.2}$$

4.3.1 Classification Performance at Various Pile-up Conditions

To probe how the various conversion classification methods perform as a function of pile-up, 14 TeV single photon samples simulated at four different pile-up conditions are used. Each Monte Carlo sample is divided in two parts, with the first half used to train the BDT and the other half used as a test sample. The performances are evaluated on the test sample. To ease comparison, the efficiency of the BDT in reconstructing conversion is fixed at 90% for each $|\eta|$ category. Similarly, the fake conversion rate can also be fixed instead of the efficiency.

Since the samples used in this study are from LS1, many changes have been made since then to photon reconstruction, so it should be noted that tracking performance in this section does not correspond to what is achieved in Run 2. The reconstruction in these samples was made for 50 ns bunch spacing and for pile-up of up to 40. It was not tuned for Run 2 conditions. The algorithm for 2015 data taking is much better adapted to the running conditions than what was used in these samples.

4.3.1.1 Comparison of calorimeter-based and track-based conversion classification

Conversion reconstruction efficiencies and fake conversion rates are plotted in Figure 4.11 against μ for the two methods of conversion classification: calorimeter-based and track-based. With the efficiency of the calorimeter-based tagging fixed at 90%, the increase of fake conversion rate with μ is relatively small, going from 18% at $\mu = 0$ to 27% at $\mu = 80$. Obvious degradation with the pile-up rate is seen in the default track-based conversion tagging. The conversion reconstruction efficiency drops from 83% at $\mu = 0$ to 62% at $\mu = 80$ and the fake conversion rate rises to 44% at $\mu = 80$. Due to many pile-up tracks at high μ, it is expected that the fake conversion rate would increase as mismatching of pile-up tracks to electromagnetic cluster becomes more likely. Fake tracks also contribute to this degradation.

One can also fix the fake rate instead of the efficiency. Fake conversion rate of 2% is chosen as it is the typical value that is aimed for in conversion classification. By fixing the fake rate to this value, the efficiency is 62% at $\mu = 0$ and drops to 51% at $\mu = 80$. Although the efficiencies are not high, they are quite stable.

4.3.1.2 Using both calorimeter and tracking information

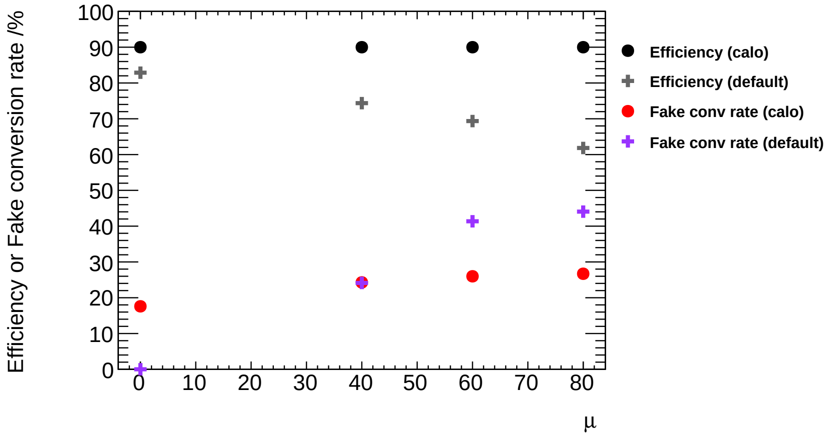
Through the addition of tracking information in the BDT, the performance of the classifier is further improved. The fake conversion rate is lowered when the tracking information is included especially at lower values of μ where the tracking performance is good, as shown in Figure 4.12. The fake conversion rate at $\mu = 0$ is 2.5% with 90% conversion reconstruction efficiency whereas the fake conversion rate in the default track-based tagging is lower at 0.01%, albeit with a lower efficiency. For the same efficiency as the default track-based method, the fake rate obtained from the BDT using both calorimeter and tracking information, is found to be 0.01%, thus recovering the excellent performance of the default method at $\mu = 0$.

Further comparisons of the performance of the three conversion tagging methods (default track-based, BDT of calorimeter variables and BDT of calorimeter and tracking variables) as a function of pseudorapidity, transverse momentum, radius of conversion and number of primary vertices for $\mu = 0$, 40 and 80 samples can be found in Appendix A.1.

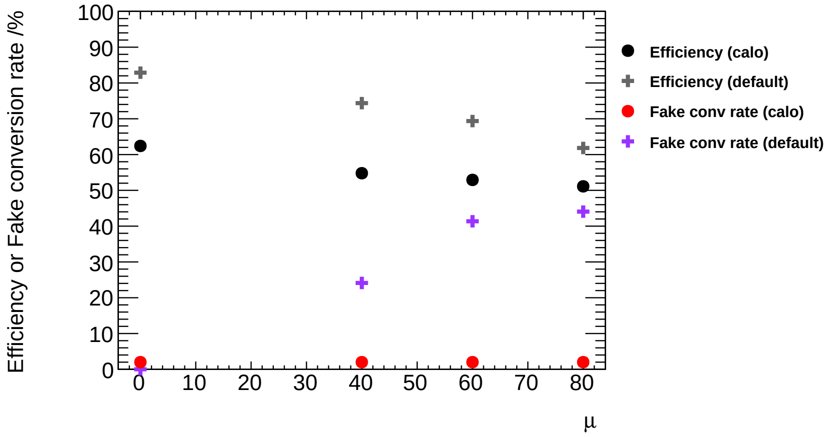
4.3.1.3 Choice of the Training Sample

In Sections 4.3.1.1 and 4.3.1.2, the BDT classifiers are trained and tested on samples with the same distribution of μ. This however does not imply that it is imperative to use a training sample with the same pile-up as the test sample or data.

In the case of using only calorimeter variables, the classifier trained with $\mu = 0$ can safely be applied to samples with higher μ. It has been checked that this yields comparable results, as



(a)



(b)

Figure 4.11: Conversion reconstruction efficiency and fake conversion rate of the calorimeter-based and the track-based tagging as a function of μ. The efficiency of the calorimeter-based tagging is fixed at 90% in (a). The fake conversion rate is fixed at 2% in (b).

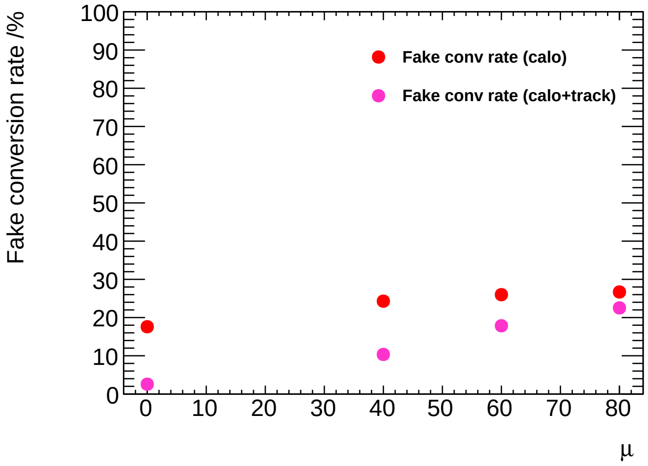


Figure 4.12: Comparison of the fake conversion rate against μ between a BDT using only calorimeter information and a BDT using both calorimeter and tracking information. The conversion reconstruction efficiency is fixed at 90% in both cases.

shown in Figure 4.13. Using the same cuts on the classifier output to obtain a 90% efficiency on the $\mu = 0$ sample, for higher μ samples, the efficiencies are slightly smaller but fake conversion rates are also slightly smaller. To investigate quantitatively the implications of different μ training samples, matching the efficiency to 90% shows that the difference between the classifier trained at different μ amounts to only 2-3% in fake rates.

The same does not apply to the case when tracking variables are included in the BDT since tracking performance in these samples varies significantly with pile-up, unlike the separation of converted and unconverted photons in calorimeter variables.

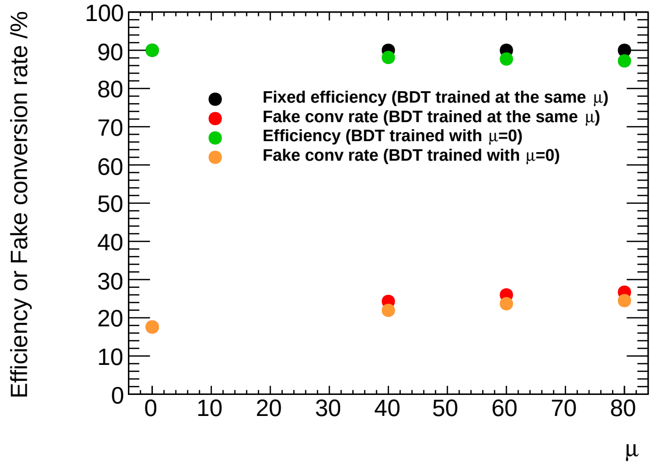


Figure 4.13: Comparison of the conversion reconstruction efficiency and the fake conversion rate between a BDT trained with the same μ as the application sample and a BDT trained with $\mu = 0$ Monte Carlo sample.

4.3.2 Classification Performance with 2012 Run Conditions

The performance of the tracking-based conversion classification shown in Section 4.3.1 is based on conditions for which the tracking-based algorithms were not tuned. Hence, the comparisons were not fair as the reconstruction in those LS1 samples was made for 50 ns bunch spacing and pile-up of up to 40.

For a fair comparison of the track-based and of the calorimeter-based classification, their respective performances with 2012 run conditions are explored in this section. An 8 TeV simulated sample of radiative $Z \to \mu\mu\gamma$ decays is used.

The efficiency and fake rate of the calorimeter-based classification compared to the default method are presented in Table 4.4 where the BDT has been tuned to have the same fake rate as the default method. The classification based on shower shape information is not competitive with the default method with 2012 run conditions. By training the BDT with additional tracking information, the efficiency of reconstructing conversion can be increased while the fake conversion rate remains the same.

4.3.3 Validation on Data

Using simulated samples, we have demonstrated the feasibility of employing a multivariate classifier to combine information about shower shape to separate converted from unconverted photons. Nevertheless, this method needs to be tested on data in order to validate the description of Monte Carlo. In particular, the Monte Carlo does not describe the shower shape in data well. This could be problematic, as the fudge factors applied on the shower shape variables are not perfect and are dependent on the reconstructed conversion status. In collision

Default track-based		Shower information only		Shower + track information	
Efficiency	Fake rate	Efficiency	Fake rate	Efficiency	Fake rate
79%	2.3%	62%	2.3%	86%	2.3%

Table 4.4: The conversion reconstruction efficiency and fake conversion rate of the various classification methods (the default algorithm, the classification using only calorimeter information and the classification using both calorimeter and tracking information) with 2012 LHC run conditions.

data, sources of relatively pure photons include $Z \to ll\gamma$ and diphoton events. To validate the calorimeter-based conversion classification, the distributions of the BDT outputs from Monte Carlo and data are compared for both photons reconstructed as converted and unconverted by the track-based algorithm.

4.3.3.1 Radiative $Z \to \mu\mu\gamma$ decays

In order to check that the distributions of the calorimeter-based classifier output in data are reasonably well described by the simulation, one has to rely on the track-based classification to separate the converted and unconverted photon candidates in the data sample. The data distributions are compared to MC for converted and unconverted photons reconstructed from the track-based algorithms. The distributions for true converted and unconverted photons are also displayed to show the imperfection of the track-based classification.

Figure 4.14 shows the comparison of the classifier output distributions in data and MC. The simulated sample used in this plot and to train the classifier is the $Z \to \mu\mu\gamma$ sample. The plots show good agreement between data and Monte Carlo distributions of the classifier output for reconstructed converted and unconverted photons. Tiny differences between the two are due to the imperfection of Monte Carlo at describing shower shape in data and possible difference in efficiency of the track-based reconstruction in data compared to MC. Comparing the distributions of true converted and unconverted photons to their reconstructed (from tracking) counterparts, one observes that they are not very different in some $|\eta|$ regions where the tracking is efficient. In $0.8 < |\eta| < 1.15$ and $|\eta|$ above 1.8, the inefficiency of tracking (see Section 3.1) is evident as only a fraction of the true converted photons are reconstructed as converted.

4.3.3.2 Diphoton events

The high-p_T loosely identified photon from diphoton events in data (see Section 4.1 for the selection criteria) are also used to validate the classification method. Figure 4.15 shows the agreement of the BDT output on data compared to Monte Carlo. The agreement between data and Monte Carlo is reasonably good, though slightly worse than $Z \to \mu\mu\gamma$ in Figure 4.14. The BDT is trained with single photon Monte Carlo with binning in $|\eta|$ and p_T as described in Section 4.2.3. The training sample has zero pile-up while the data has pile-up, so the track-based separation of converted and unconverted photons has different efficiency for the data and the Monte Carlo. There are also significant jet contamination in the diphoton data.

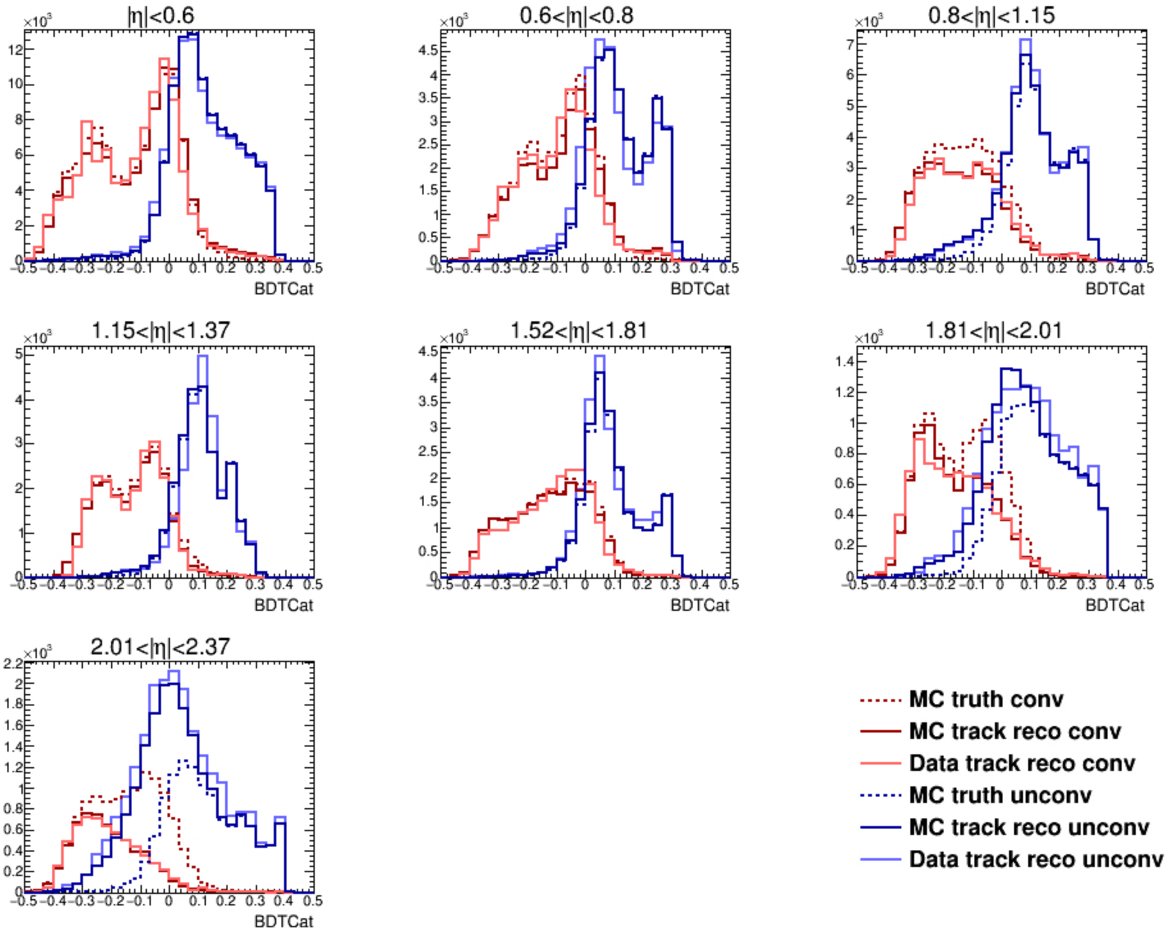


Figure 4.14: Comparison of the distributions of the BDT classifier output in $Z \rightarrow \mu\mu\gamma$ data and MC. The data (MC) distributions are drawn with light (dark) colours. The conversion status is taken from the default track-based classification. The distributions in MC converted and unconverted photons separated based on the true conversion status are plotted with dashed lines.

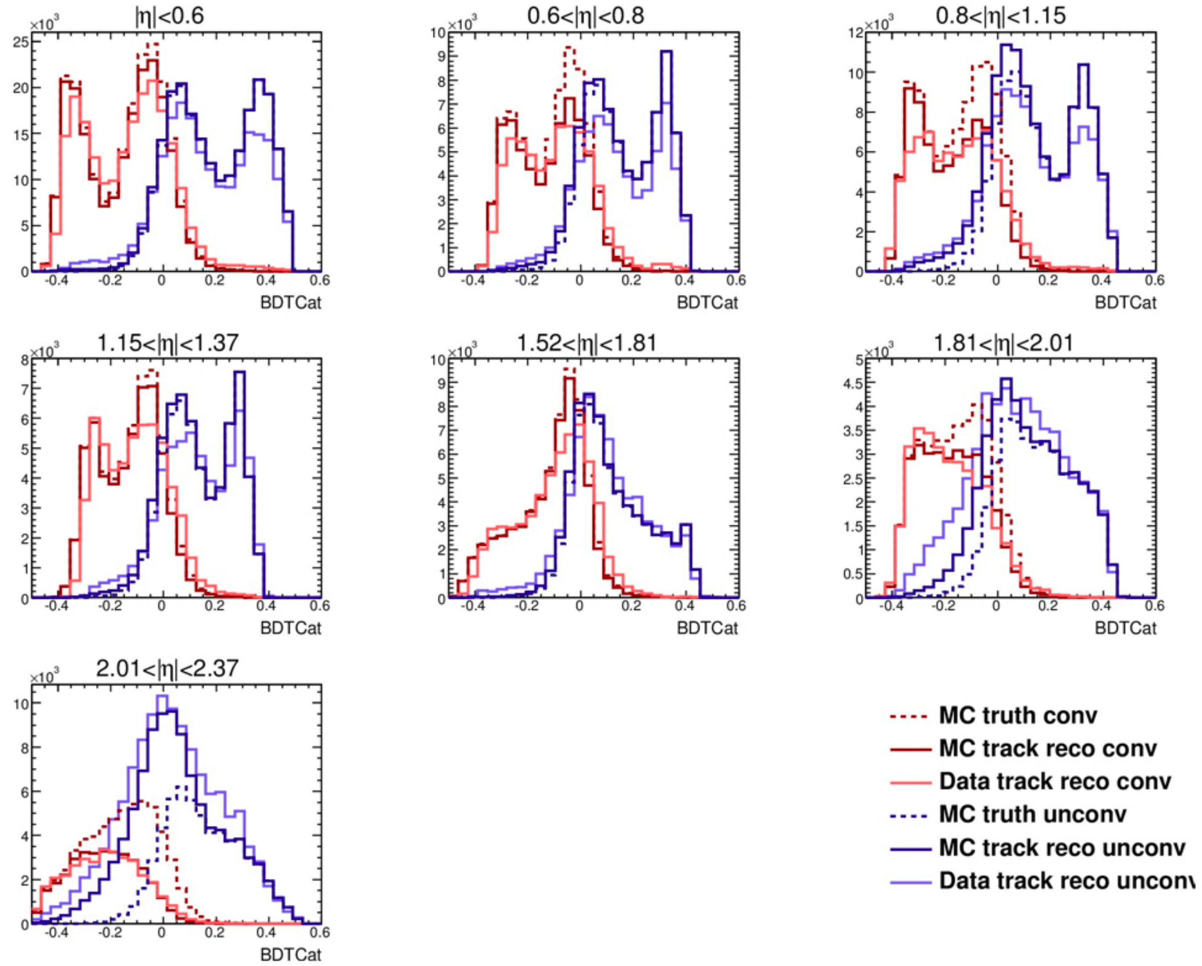


Figure 4.15: Comparison of the distributions of the BDT classifier output in diphoton data and single photon MC sample. The data (MC) distributions are drawn with light (dark) colours. The conversion status is taken from the default track-based classification. The distributions in MC converted and unconverted photons separated based on the true conversion status are plotted with dashed lines.

4.4 Template Fit to Measure Conversion Fractions in $Z \to \mu\mu\gamma$ Data

Thus far, the goal of the study was to improve on the existing method of separating converted and unconverted photons in a high pile-up environment. Another interesting and divergent study is to be able to measure its performance in the real data. The efficiencies and fake rates of the track-based conversion tagging method are known from studies using Monte Carlo, such as those presented in Section 4.3 and Ref. [60]. However, the reconstruction in the data may not be exactly the same, and it is important to measure the efficiencies in data as it affects the photon energy scale, identification efficiency, etc. In the photon calibration [61], the energy of true converted photons reconstructed as unconverted is typically underestimated by about 2%, depending on p_T, $|\eta|$ and the radius of the conversion. On the other hand, fake conversions induce around 2% overestimation of the energy.

In Section 12.1 of Ref. [61], a fit of the E_1/E_2 distributions, using templates of true converted and unconverted photons extracted from Monte Carlo, is performed on data to estimate the conversion reconstruction efficiency and fake conversion rate. Systematic uncertainties of the absolute photon energy scale were estimated by comparing the efficiency and fake rate measured in data to Monte Carlo simulation.

In this section, an extension of the work in Ref. [61] to measure the performance of the track-based conversion tagging in data with higher precision is presented. The E_1/E_2 distribution used in the fit is replaced by the distribution of a BDT classifier output. As in the previous sections, the BDT is based on several shower shape variables and therefore is expected to bring a better separation of converted and unconverted photons compared to the original E_1/E_2 variable alone. This is studied to gauge the improvement that it may bring, taking into account the potential increased systematic uncertainties from the additional variables. In the following, BDT with gradient boost, i.e. BDTG is used instead of BDT with adaptive boost in the fit.

The previous study in Ref. [61] uses leading photons from diphoton events with the same event selection as adopted by the $H \to \gamma\gamma$ analysis. In the selection, both photons in the diphoton event pass tight photon identification criteria. As the shower shape used in tight photon identification is included in the template fit, the tight identification requirements cannot be applied, to avoid bias from the shower shape cut values that differ for converted and unconverted photons. Diphoton events have a sizeable jet background that is difficult to subtract or estimate without resorting to using the tight identification criteria. Hence, in this study, the $Z \to \mu\mu\gamma$ events sample is used as it has a high purity even without applying tight photon identification criteria. To assess the possible improvement, the template fits with E_1/E_2 variable on $Z \to \mu\mu\gamma$ events as well as with BDTG are both performed and then compared.

4.4.1 Method

The measurements use $Z \to \mu\mu\gamma$ events in the 2012 data. Simulated $Z \to \mu\mu\gamma$ samples using nominal and distorted geometries are used for fitting and systematics studies respectively, whereas the $Z \to \mu\mu$ sample is used for background study. Reweightings are applied to account for the differences in p_T, η, z vertex and pile-up distributions in the Monte Carlo samples compared to the data. Fudge factors are also applied to the shower shape variables in Monte Carlo to better match data distributions. The efficiency of selecting candidate events passing all the cuts is around 14% [73], which means the nominal Monte Carlo has roughly 10 times the statistics of data (approximately 73 thousand events).

The conversion reconstruction efficiency and fake conversion rate are extracted by fitting Monte Carlo templates to data distributions for each of the two reconstructed conversion status. This study is performed in four $|\eta|$ bins (0-0.6, 0.6-1.37, 1.52-1.81, 1.81-2.37) as was previously done in Ref. [61]. The BDTG is trained with 7 $|\eta|$ categories as in Section 4.2.3. The 7 categories are a superset of the 4 categories defined in Ref. [61] since the bin boundaries are the same in certain bins and in others, two bins are merged into one. Hence, there is no problem in using 7 $|\eta|$ categories in training and 4 in fitting. The BDTG is trained with the $Z \to \mu\mu\gamma$ MC sample with nominal geometry and applied to data.

The HistFactory [74] tool is used to build the statistical model based on input template histograms. The template histograms are obtained from the nominal Monte Carlo sample for each combination of true and reconstructed conversion status.

In each $|\eta|$ bin, the numbers of reconstructed unconverted photons and converted photons are given by equations 4.3 and 4.4:

$$N^{reco}_{unconv} = N \cdot [f_{\text{conv}} \cdot (1 - f_{\text{reco}}) + (1 - f_{\text{conv}}) \cdot (1 - f_{\text{fake}})] \tag{4.3}$$

$$N^{reco}_{conv} = N \cdot [f_{\text{conv}} \cdot f_{\text{reco}} + (1 - f_{\text{conv}}) \cdot f_{\text{fake}}] \tag{4.4}$$

where N is the total number of candidate photons, f_{conv} the probability of a photon to convert into an electron-positron pair, f_{reco} the conversion reconstruction efficiency (Eq. 4.1), and f_{fake} the fake conversion rate (Eq. 4.2).

The variable to be fitted has to be sensitive to the true conversion status. The templates are fitted simultaneously for reconstructed unconverted and converted photons in each $|\eta|$ bin according to equations 4.3 and 4.4. There are 8 channels, corresponding to 4 $|\eta|$ bins and 2 reconstructed conversion status. Channels refer to the regions of the data with disjoint event selections. In each channel, there are 3 components, true converted, true unconverted and background (if included). In each $|\eta|$ bin, the two channels corresponding to reconstructed converted and unconverted are linked by f_{conv}, f_{fake} and f_{reco}.

As f_{conv} and f_{reco} were found to be highly anti-correlated, f_{conv} is fixed to be the value in the nominal Monte Carlo with some uncertainties given from the relative fraction taken from the MC with geometry A (see Section 4.1) while f_{fake} and f_{reco} are let free in the fit. Relative normalisations of the templates are allowed to vary by a few percent determined by the uncertainties in f_{conv}. Other sources of systematic uncertainties vary the shape of the templates. This will be discussed further in Section 4.4.4.

4.4.2 Choice of Variables

To choose which variable to include in the boosted decision tree, firstly one has to check the data and Monte Carlo distributions to ensure that they agree well (with the fudge factors applied). When a difference is observed, one cannot disentangle whether it arises from the difference in conversion reconstruction performance or from the lack of ability of MC to describe the shower shapes in data due to imperfect calorimeter simulation. Hence, only the variables that are well-modelled by the MC are used in order to ensure that the distributions on those variables in data are reliable to study the performance of the track-based conversion reconstruction.

In order to quantify the agreement, the ratio of the mean in data divided by the mean in MC is plotted in bins of $|\eta|$. This is illustrated in Figures 4.16 and 4.17 for the variables E_1/E_2 and R_ϕ. The E_1/E_2 variable has been corrected with the layer intercalibration procedure described in Section 3.2. The level of deviation observed in the E_1/E_2 MC distribution with respect to data is taken as a reference. This leads to a condition on the difference in mean between data and MC to be no more than 5% in the whole range of η.

Out of the eight variables detailed in Table 4.3, five variables pass this specification: E_1/E_2, R_ϕ, f_1, w_{s3} and $w_{\eta 2}$. E_{maxs1}, which is the energy of the strip with maximal energy deposit, does not show a good agreement between data and Monte Carlo. E_0/E_{acc} is not corrected for data and MC agreement unlike E_1/E_2 hence the data and MC distributions do not agree well. Finally, there are some disagreements in the end-cap region for F_{side}.

Different combinations of these five variables are tested in terms of the error on the fitted values of f_{reco} and f_{fake} on Monte Carlo pseudo-data.[3] It was found that a BDTG with only 2 variables (E_1/E_2 and R_ϕ) as input performs just as well as having more variables, or even all 5. The shape of the BDTG distributions with only 2 variables as input is not very different from the distributions with more variables as they both peak at -1 and 1. Even though the separation of the distributions of true converted and unconverted is better, it is marginal from the point of view of the template fit. Hence it was decided to use the combination of E_1/E_2 and R_ϕ in order to avoid complicating the fit with too many sources of systematic uncertainties which would result in loss of sensitivity.

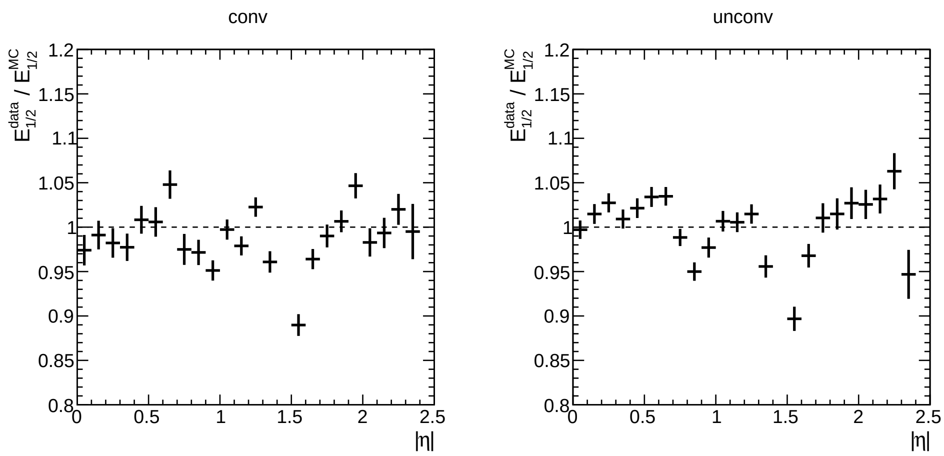


Figure 4.16: Ratios of the mean of E_1/E_2 in data divided by the Monte Carlo distribution as a function of $|\eta|$ for converted (left) and unconverted (right) photons after layer recalibration has been applied on MC [61].

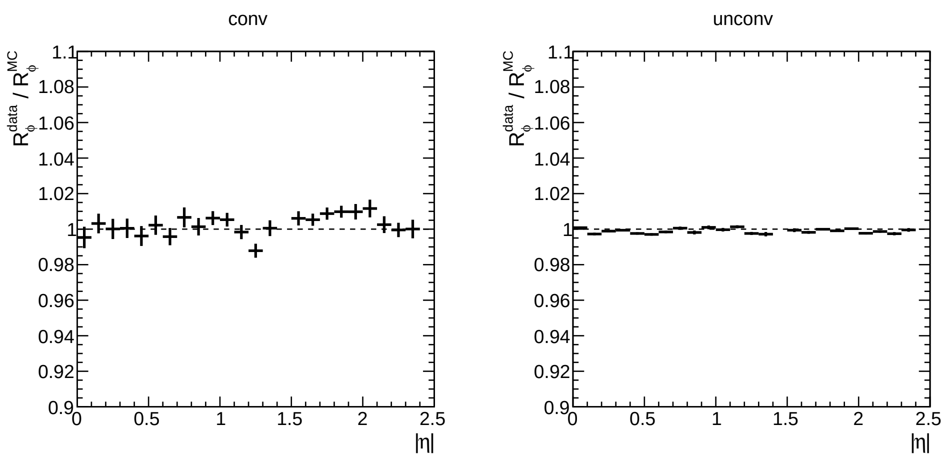


Figure 4.17: Ratios of the mean of R_ϕ in data divided by the Monte Carlo distribution as a function of $|\eta|$ for converted (left) and unconverted (right) photons after applying fudge factor on MC.

[3]In the template fit, the MC distributions are used in place of data distributions. This gives the expected sensitivity.

4.4.3 Background Templates from Purity Estimation

The source of background in the $Z \to \mu\mu\gamma$ sample is $Z \to \mu\mu$ decays with at least an additional jet in the final state which fakes a photon. To extract the background templates, first one needs to know the purity in data, which is expected to be different in each bin of E_1/E_2 or BDTG. The purities for each of the 64 bins (8 bins in the fit variables $\times$ 4 $|\eta|$ bins $\times$ 2 conversion status) are extracted by fitting the distributions of $\mu\mu\gamma$-system invariant mass, $m_{\mu\mu\gamma}$.

The $Z \to \mu\mu\gamma$ signal peaks sharply at the mass of the Z boson whereas the background, $Z \to \mu\mu + jet$ has a flatter distribution. To perform the fit, the $m_{\mu\mu\gamma}$ mass requirement is loosened from $80 < m_{\mu\mu\gamma} < 96$ GeV to $64 < m_{\mu\mu\gamma} < 114$ GeV.

The signal $m_{\mu\mu\gamma}$ templates are obtained from the nominal Sherpa $Z \to \mu\mu\gamma$ Monte Carlo sample with the usual selections except for the $m_{\mu\mu\gamma}$ mass requirement. The background $m_{\mu\mu\gamma}$ templates are obtained from an inclusive Sherpa $Z \to \mu\mu$ sample with the same selection as well. Details on these two samples are given in Section 4.1. However, as the Sherpa $Z \to \mu\mu$ sample contains $Z \to \mu\mu\gamma$ events, the overlap between the two needs to be removed. Selections are applied to choose only true prompt photons from the signal MC sample and fake photons from the inclusive background sample, based on the matching of photon candidates to the true particle.

Ideally, there should be separate $m_{\mu\mu\gamma}$ templates for each bin of E_1/E_2 or BDTG. However, since the fit is performed on a large number of bins, the statistics pose a problem. Out of the ten million events generated in the Sherpa $Z \to \mu\mu$ sample, only 3800 events pass all the required selection cuts and have $m_{\mu\mu\gamma}$ between 64 and 114 GeV. The statistics are also not uniform, with some bins having as few as 5 events. In order to increase the background statistics in each bin, background templates are taken from 4 bins in the fit variable. The fit is still done independently in each bin, but a background template is shared by 4 neighbouring bins.

The sum of signal and background $m_{\mu\mu\gamma}$ templates, with floating normalisations, is fitted to the data distribution in each bin. The purity is calculated from the fitted numbers of background and signal events in each bin, rescaled from the wide $m_{\mu\mu\gamma}$ window to the signal region ($80 < m_{\mu\mu\gamma} < 96$ GeV). The ultimate background templates (for both E_1/E_2 and BDTG) are given by the product of the purity in each bin and the corresponding statistics in data. This yields both shape and normalisation of the background templates.

Figure 4.18 shows the result of the fit to $m_{\mu\mu\gamma}$ distributions in 8 bins of BDTG for the $|\eta| < 0.6$ channel with photons reconstructed as unconverted, as an example. The bottom right plot shows the BDTG template obtained from the eight $m_{\mu\mu\gamma}$ fits, corresponding to the fake unconverted photons with $|\eta| < 0.6$.

These templates are then compared with the ones obtained directly from the background MC sample. The templates from MC are normalised to the number of background events expected in data from MC cross-sections and acceptances. Figure 4.19 shows the two templates superimposed on the same plots. Good agreement is observed. The background templates all peak at -1 regardless of whether these fake photons are reconstructed as converted or not as they have shower shapes that are similar to those of a true converted photon.

4.4.4 Systematic Uncertainties

Sources of systematic uncertainties considered in this study are those from the limited statistics in the Monte Carlo and background templates, the uncertainties in the amount of material upstream of the electromagnetic calorimeter and the uncertainties due to the residual difference between data and MC description of the shower shape after corrections have been applied. The statistical uncertainties on the signal and background templates affect the normalisation in a bin-by-bin basis whereas other uncertainties considered here vary the overall normalisation and the coherent shape of the templates.

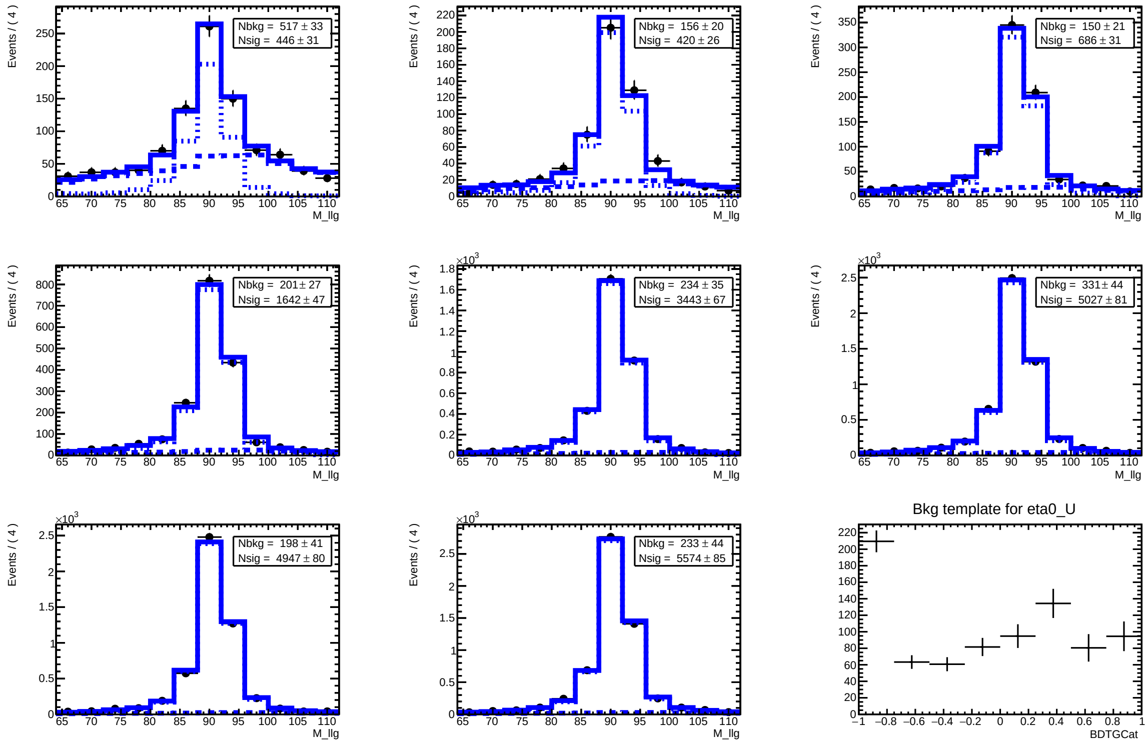


Figure 4.18: From top left to bottom right, the first 8 plots show the $m_{\mu\mu\gamma}$ fit results for photons (true photons and jets faking photons) with $|\eta| < 0.6$ and reconstructed as unconverted by the default algorithm in 8 bins of BDTG. The solid blue line represents the results of the fit to data distribution with the sum of the signal (blue dotted line) and background (blue dashed line). The final plot on bottom right is the resulting BDTG background template.

4.4.4.1 Uncertainties due to limited statistics in the Monte Carlo and background templates

The uncertainties on the templates need to be taken into account due to the limited statistic of the samples used to extract the templates. In particular, large uncertainties are present in the background templates from the fit of low statistics $m_{\mu\mu\gamma}$ templates as the data is divided into 64 bins.

To maximise the number of photon probes in the Monte Carlo, both the training and test samples are used. Since it was checked that there is no overtraining, the two independent samples can be safely combined. In all the fits, statistical uncertainties on the signal templates are taken into account in the form of bin-by-bin statistical variations. One nuisance parameter per bin is associated with the total statistical uncertainty in that bin.

The statistical errors from the background templates are not included in the fits with background components due to a technical problem. In HistFactory, when uncertainties from both signal and background are included, the results look as if the uncertainties are much larger than they really are. Therefore, it was decided to include only the statistical errors in the signal true converted and unconverted templates. It was checked that the statistical errors from the signal components are indeed larger than those from the background components since the background contributes to only about 4% in data. Hence, ignoring the background template uncertainties should not have a significant impact. This is shown in Figure 4.20 where the total signal errors are the sum of errors in true converted and unconverted templates in quadrature and the total error is the sum of the errors in all the components, including background. The errors are calculated from the templates scaled to match the fractions in data from the fit without

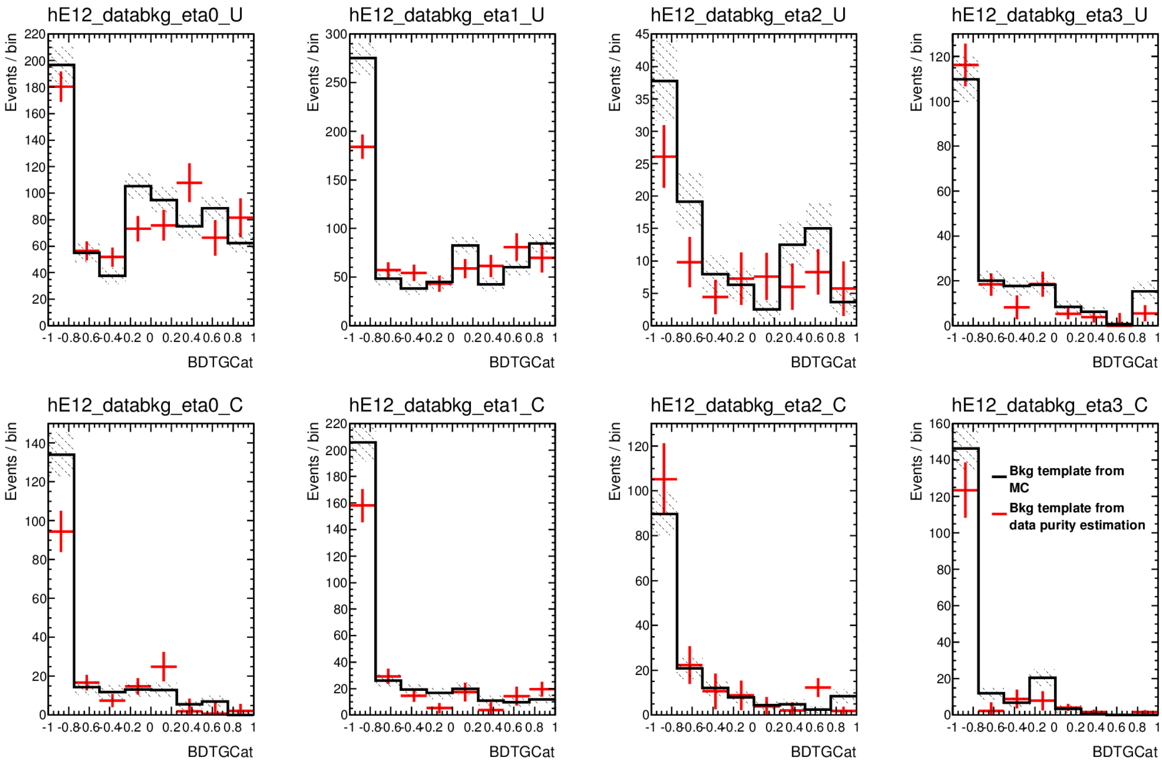


Figure 4.19: The BDTG background templates that are extracted from the Monte Carlo (black) superimposed with the background templates from the data purity estimation (red). The plot on the bottom right of Figure 4.18 is the same as the red plot on the top left. The top four plots correspond to the channels of unconverted photons as reconstructed by the default algorithm at four $|\eta|$ bins, denoted eta0, eta1, eta2 and eta3. The bottom four plots correspond to the channels of reconstructed converted photons.

uncertainties from the statistical errors on the signal and background templates.

4.4.4.2 Uncertainties due to material description

The imperfect description of the material before the electromagnetic calorimeter in the ATLAS simulation results in an uncertainty which can affect both the probability of a photon to convert ($f_{\rm conv}$) and the shape of the shower shape variables and thus of the BDTG output. In order to take into account this uncertainty, $Z \to \mu\mu\gamma$ samples simulated with three different distorted geometries listed in Section 4.1 are used.

In the distorted geometry sample with configuration A, there are 5% extra material in the inner detector. This directly affects both the true conversion fraction and the shape of the templates. The additional materials in configuration E'+L' and F'+M+X are close to the calorimeter, so they do not change the true conversion fraction but should have a sizeable impact on the shape. Other distorted geometries exist but these are thought to be the most relevant ones.

The difference in the shape of the template between the nominal and the distorted geometry samples is taken as a $+1\sigma$ systematic variation, whereas a -1σ variation is derived by inverting this difference. For the normalisation uncertainty, the ratio of $f_{\rm conv}$ in geometry A to that in the nominal geometry is taken as $+1\sigma$ variation, and is also symmetrised.

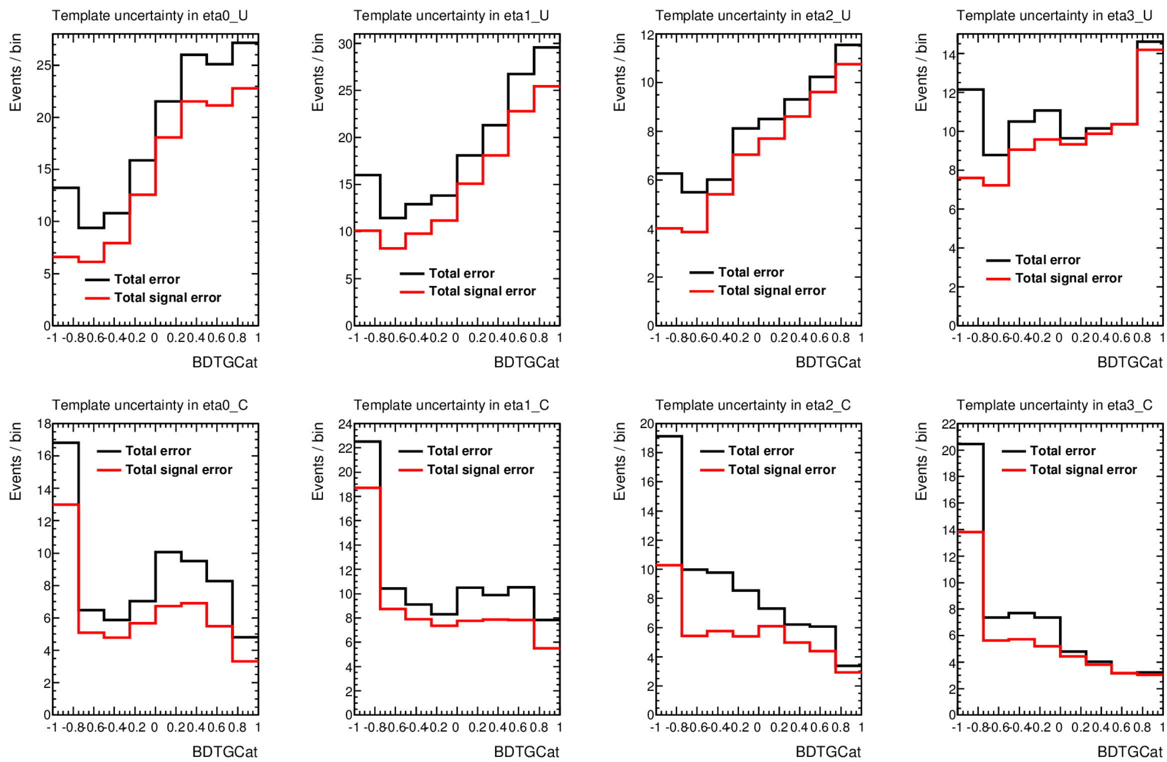


Figure 4.20: The sum in quadrature of the errors from all components (black) and the errors from only signal components (red). The top four plots correspond to the channels of unconverted photons at four $|\eta|$ bins, denoted eta0, eta1, eta2 and eta3. The bottom four plots correspond to the channels of converted photons.

4.4.4.3 Uncertainties due to imperfect description of shower shape in MC

There are other systematic uncertainties that have an effect on the shape of E_1/E_2, such as the cross-talk in the electromagnetic calorimeter. They were included in the previous study in Ref. [61] and were found to be small compared to the systematic uncertainties due to material. For this reason, they are not included here.

For R_ϕ, additional systematic uncertainties on the variable need to be considered. The statistical uncertainty on the fudge factor of a given shower shape variable is taken as 1σ uncertainty. This uncertainty accounts for the residual difference between data and MC distributions after the fudge factor has been applied.

As E_1/E_2 is not used in photon identification, no fudge factor has been derived for the variable. However, this energy ratio has been corrected via a layer recalibration using muons and the associated uncertainties are negligible compared to the other sources [61].

4.4.4.4 Parameterisation of the normalisation and shape variations due to systematic effects

In summary, normalisation systematics are taken from geometry A as the increase in material in geometry A impacts on the true fraction of conversion. In addition, all of the 3 distorted geometries and fudge factor uncertainties contribute to the shape systematics. Figure 4.21 gives an example of the difference in shape between the nominal and the distorted samples.

In order to properly take into account the uncertainties described above, nuisance parameters are introduced for each source of systematic uncertainty, i.e. material distortion in geometry A,

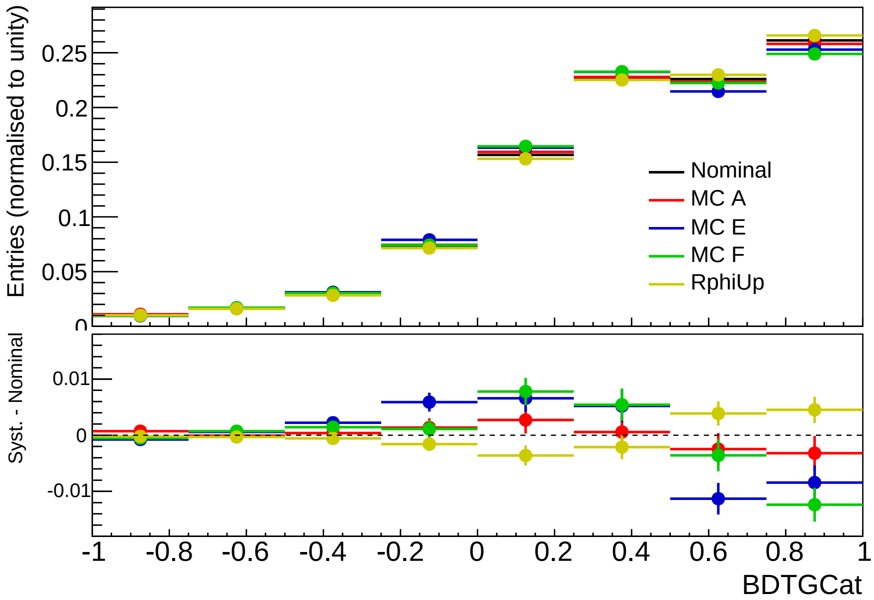


Figure 4.21: An example of shape systematics for true unconverted photons with $|\eta| < 0.6$ and reconstructed as unconverted. RphiUp refers to the BDTG obtained when the 1σ uncertainty is added to the R_ϕ variable.

E'+L' and F'+M+X and the uncertainty from the R_ϕ fudge factors. For the material-related systematic uncertainties, independent nuisance parameters are defined for each $|\eta|$ bin. Whereas for the uncertainties from the fudge factors, a nuisance parameter for each $|\eta|$ bin and conversion status is used since different fudge factors are applied depending on the reconstructed conversion status.

Since the distortion in geometry A affects both f_{conv} and the shape of the templates, one can either use (in each $|\eta|$ category) one nuisance parameter for each of the two effects, or one parameter taking into account both effects in a correlated way. The two options are tested, using the geometry A MC sample as pseudo-data, scaled to match the number of events in the real data sample. The fit is performed using templates extracted from the nominal geometry MC sample, adding nuisance parameters to include all the systematic effects. No background component is taken into account in the test since the pseudo-data contains only signal.

The nuisance parameters corresponding to geometry A are expected to be at 1σ and the others at 0σ. Figure 4.22 shows the fitted value of each nuisance parameter where the uncertainties from geometry A on the shape and on f_{conv} are separated.

We conclude that the pseudo-data which has the same statistics as the data sample, does not have enough statistical power to pull the normalisation nuisance parameters. This is confirmed by artificially scaling the pseudo-data to higher statistics. The test is repeated by correlating the shape and normalisation effects in a single nuisance parameter per $|\eta|$ bin. The pulls of the nuisance parameters are shown in Figure 4.23. The nuisance parameters related to geometry A are all pulled at 1σ and the other nuisance parameters are scattered around 0, as expected. In the following, one single nuisance parameter per $|\eta|$ bin is used for the material distortion in geometry A.

Using one parameter to take into account shape distortion and another for normalisation when the two effects come from the same source is not realistic. This was initially tested in order to separate the two effects and to know which effect dominates the pull on the systematics related to the amount of material in the inner detector (geometry A). However, it became apparent afterwards that using one single nuisance parameter per $|\eta|$ bin to parameterise both effects is

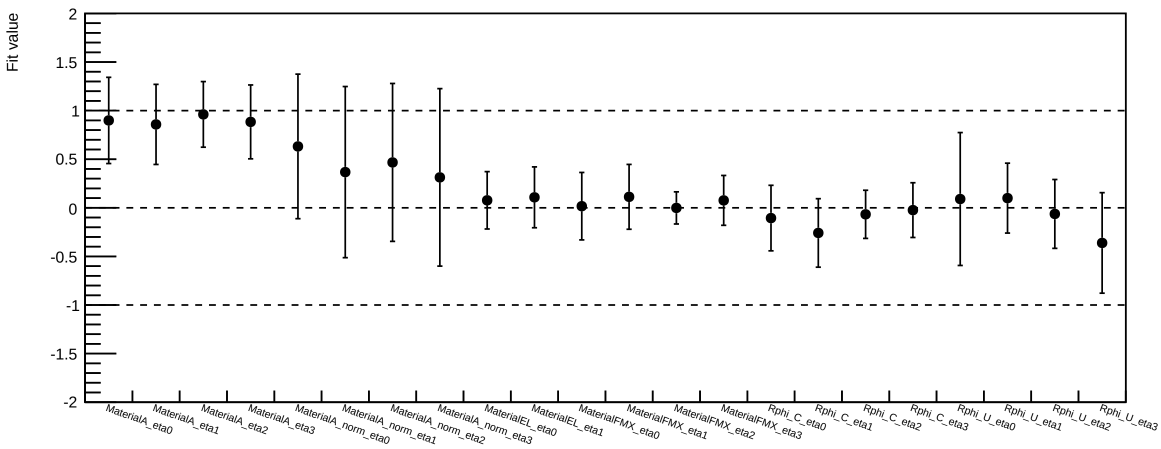


Figure 4.22: Nuisance parameters from the fit using the BDTG output to the distorted geometry A pseudo-data. The nuisance parameters denoted MaterialA correspond to the shape-related systematics and the MaterialA_norm nuisance parameters correspond to the normalisation systematics.

the correct way.

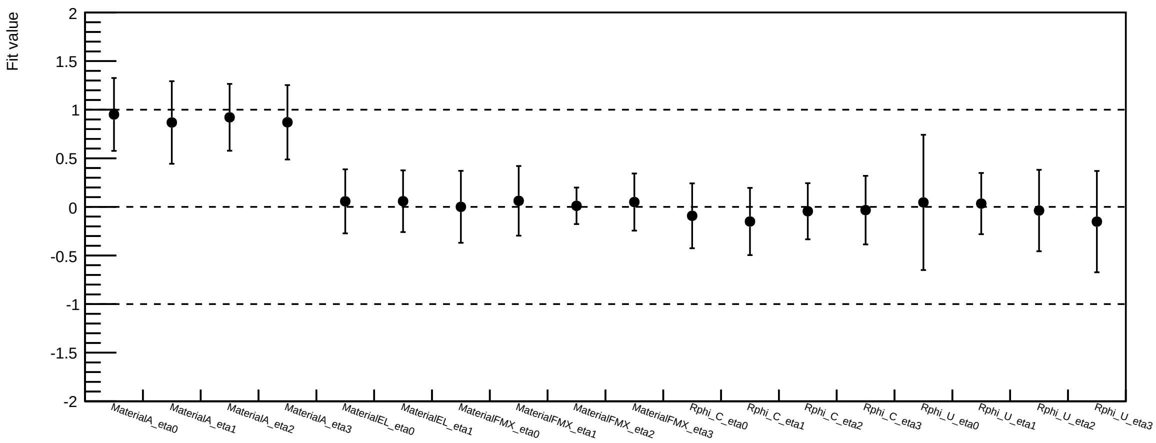


Figure 4.23: Nuisance parameters from the fit using the BDTG output to the distorted geometry A pseudo-data. The MaterialA nuisance parameters parameterise both variations in shape and normalisation.

4.4.5 Closure Test on Monte Carlo Sample

MC templates can be used as pseudo-data by a simple scaling to match the statistics in data. Using MC as pseudo-data provides a check on the methods and nuisance parameters. The checks to be performed are whether the fitted values of f_{reco} and f_{fake} match those obtained directly from the Monte Carlo truth, as presented in Table 4.5, perfect fit of templates and no pull on nuisance parameters. The fit to pseudo-data also shows how much each nuisance parameter is constrained. The errors on the fitted values of f_{reco} and f_{fake} are also a measure of the sensitivity of such a fit. In this closure test, the background component is set to zero. Perfect fit is seen with the Monte Carlo as pseudo-data, as expected. This is illustrated in Figure 4.24 for the fit using E_1/E_2.

Figure 4.25 shows the fitted values of the nuisance parameters from the fit of E_1/E_2 and BDTG on pseudo-data. The values are all at zero with varying degrees of constraint. Systemat-

η	f_{conv}	f_{fake}	f_{reco}
$\lvert\eta\rvert < 0.6$	0.168	0.021	0.892
$0.6 < \lvert\eta\rvert < 1.37$	0.310	0.033	0.804
$1.52 < \lvert\eta\rvert < 1.8$	0.444	0.017	0.939
$1.8 < \lvert\eta\rvert < 2.37$	0.560	0.008	0.613

Table 4.5: The fractions of conversion, conversion reconstruction efficiencies and fake conversion rates obtained from the Monte Carlo truth in the four $|\eta|$ bins.

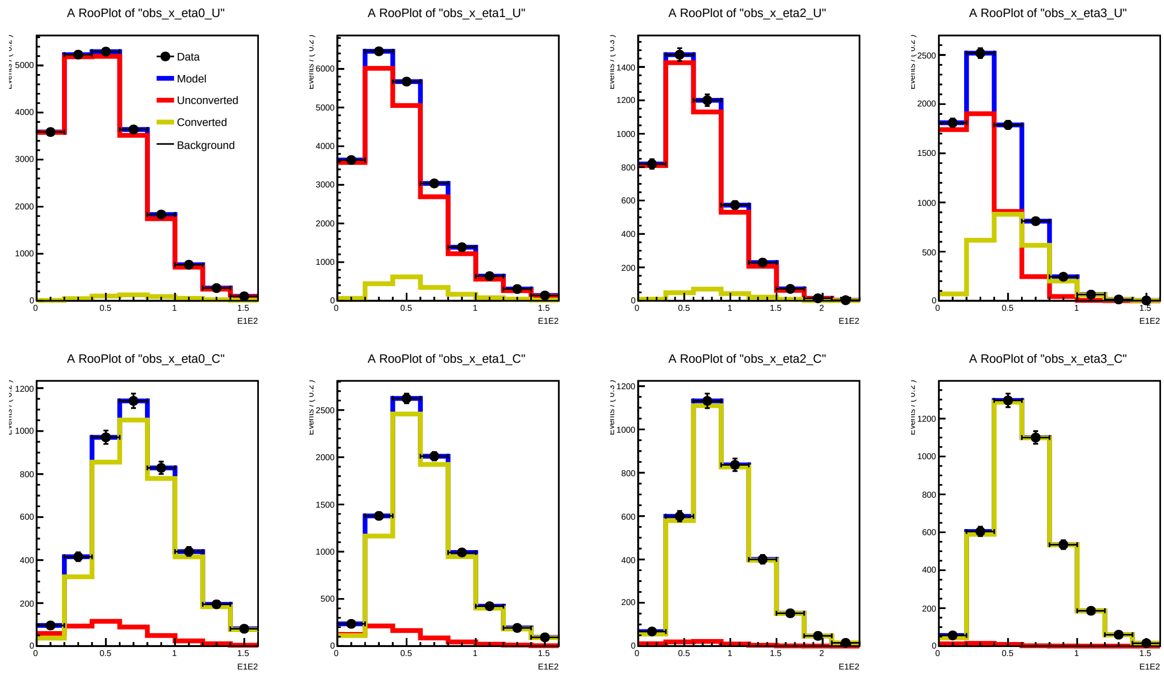


Figure 4.24: Fit to nominal Monte Carlo pseudo-data using E_1/E_2. The blue histograms correspond to the model, which is the sum of true unconverted (red) and converted (yellow). The top four plots correspond to the channels of unconverted photons at four $|\eta|$ bins, denoted eta0, eta1, eta2 and eta3. The bottom four plots correspond to the channels of converted photons.

ics from material difference are constrained, down to 20% of the amount added in the distorted geometry. Similar constraints are also seen in the previous study [75] since E_1/E_2 is sensitive to the amount of material and has been used to estimate it [61]. The fudge factor systematics are also constrained. The uncertainties on the fudge factors are computed from a data sample corresponding to an integrated luminosity of 13 fb^{-1} using both $Z \to ee\gamma$ and $Z \to \mu\mu\gamma$ channels and from Monte Carlo samples simulated with the old geometry where some material excesses were not corrected. Since the number of $ll\gamma$ events in 13 fb^{-1} is not very different from the number of $\mu\mu\gamma$ events in 20 fb^{-1}, the constraints are not expected to be strong.

Table 4.6 shows the results from the fit to pseudo-data. The values of f_{reco} and f_{fake} obtained from the fits on E_1/E_2 and BDTG agree well with each other and with the values expected from Monte Carlo. Errors on these two quantities are also shown.

The uncertainties on f_{reco} from the fit using BDTG are lower by up to a factor of 3 than those from the fit using E_1/E_2 except in the last bin where they are similar. Uncertainties on f_{fake} are of similar sizes for the two fits. This demonstrates that one can extract the conversion reconstruction efficiency from data using the BDTG fit with better precision than using E_1/E_2.

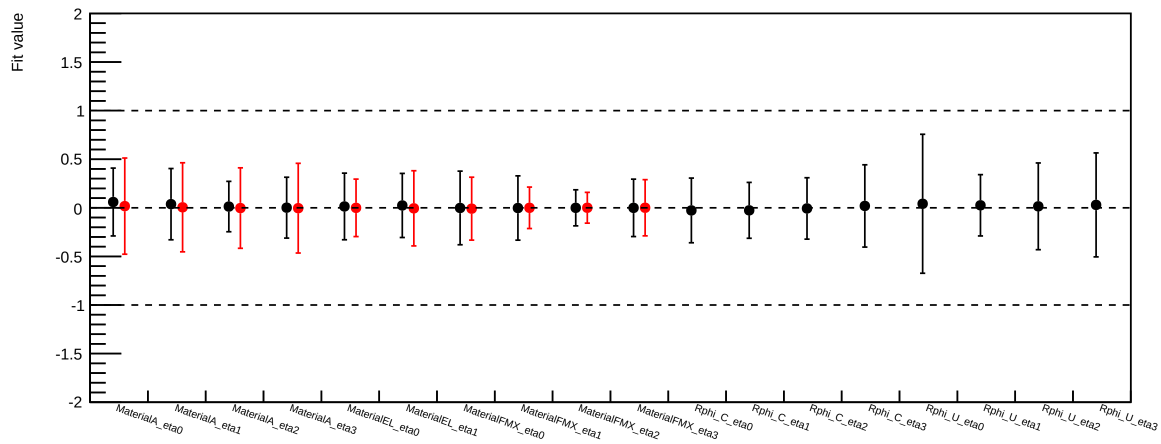


Figure 4.25: Fitted values of the nuisance parameters from the fit using E_1/E_2 (in red) and BDTG (in black) to the nominal MC pseudo-data.

	E_1/E_2		BDTG	
η	f_{fake}	f_{reco}	f_{fake}	f_{reco}
$\lvert\eta\rvert < 0.6$	0.021 ± 0.004	0.891 ± 0.029	0.022 ± 0.003	0.892 ± 0.010
$0.6 < \lvert\eta\rvert < 1.37$	0.033 ± 0.005	0.804 ± 0.015	0.034 ± 0.006	0.804 ± 0.009
$1.52 < \lvert\eta\rvert < 1.8$	0.016 ± 0.015	0.940 ± 0.023	0.017 ± 0.017	0.939 ± 0.016
$1.8 < \lvert\eta\rvert < 2.37$	0.008 ± 0.006	0.613 ± 0.012	0.008 ± 0.008	0.613 ± 0.012

Table 4.6: The results of the fit to pseudo-data using E_1/E_2 (left) and BDTG (right).

4.4.6 Fit to Data Ignoring Background

Since the $Z \to \mu\mu\gamma$ data with photon $p_T > 10$ GeV has a high purity of about 96%, it was initially thought that the background contribution could be neglected. In this section, the fits are performed by fitting the data with the signal templates only. Since the impact of background in data is found to be non-negligible, an attempt to increase the purity by raising the cut on photon p_T is also performed.

The fits are performed on the data sample described in Section 4.1, using E_1/E_2 and BDTG. The quality of the fit is found to be good in both cases. The fitted values of f_{reco} and f_{fake} are presented in Table 4.7. Both fits are performed on the same data with BDTG containing more information on the shower shape than E_1/E_2 alone. However, the value of f_{reco} obtained from the fit on BDTG tends to be lower and the value of f_{fake} tends to be higher, compared to the results from the fit on E_1/E_2.

	E_1/E_2		BDTG	
η	f_{fake}	f_{reco}	f_{fake}	f_{reco}
$\lvert\eta\rvert < 0.6$	0.011 ± 0.006	0.941 ± 0.052	0.023 ± 0.003	0.821 ± 0.013
$0.6 < \lvert\eta\rvert < 1.37$	0.032 ± 0.004	0.795 ± 0.015	0.050 ± 0.006	0.758 ± 0.010
$1.52 < \lvert\eta\rvert < 1.8$	0.016 ± 0.015	0.922 ± 0.024	0.024 ± 0.013	0.899 ± 0.015
$1.8 < \lvert\eta\rvert < 2.37$	0.014 ± 0.007	0.567 ± 0.014	0.020 ± 0.009	0.560 ± 0.012

Table 4.7: Fitted values of the fake conversion rates and the conversion reconstruction efficiencies. The fit is performed using E_1/E_2 (left) and the BDTG (right) variable, neglecting background contribution.

In the nuisance parameter plots in Figure 4.26, discrepancies are seen between the values of the nuisance parameters from the two fits. Although none of them are pulled beyond 1σ, the two fits exhibit different trends in the nuisance parameter pulls. This is most notable in the MaterialA_eta0 nuisance parameter, which is pulled by -0.8σ in the E_1/E_2 fit and by $+0.8\sigma$ in the fit using BDTG variable. This nuisance parameter correlates both the normalisation and shape variations due to material distortion in geometry A.

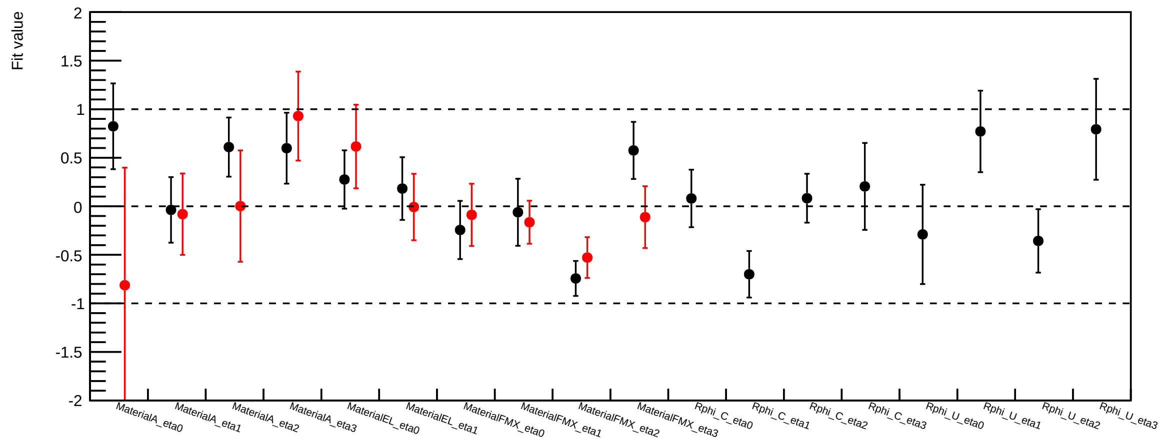


Figure 4.26: Fitted values of the nuisance parameters from the fit using E_1/E_2 (in red) and BDTG (in black) to data, neglecting background contribution.

Since the fit values are expected to be similar, one deduction is that the discrepancies observed could be due to background. From the study in Ref. [59], the photon purity for $p_T > 15$ GeV is found to be about 98%, higher than that for $p_T > 10$ GeV. Therefore, the photon transverse momentum cut is raised to 15 GeV. While some discrepancies remain, f_{reco} from the two fits agree much better in $|\eta| < 0.6$ and $1.52 < |\eta| < 1.8$. This suggests again that the background is responsible for the discrepancies observed between the two fits and by removing background entirely, one may obtain similar fit values for the two fits. Details of this study with higher p_T requirement can be found in Appendix A.3.

4.4.7 Fit to Data with Background Subtraction

To estimate the background templates, the bin-by-bin purity estimation method from Section 4.4.3 is utilised. There are now three components in each channel, true converted, true unconverted and background. A cut on photon $p_T > 10$ GeV is used to maximise statistics since background subtraction is done here.

The signal templates are fitted to data with background subtracted in each channel. The result of the template fit of BDTG is shown in Figure 4.27. Although the background contributions are small, they peak sharply at the value of BDTG=-1, which results in smaller f_{reco} when they are ignored.

Table 4.8 compares the results from the fits to data using E_1/E_2 and BDTG. Good agreement between the two fits is seen, indicating that the discrepancies observed earlier are simply due to background. The error sizes on f_{reco} are also smaller by up to a factor of 3 in the central η region for the fit using BDTG, indicating a more precise measurement, as confirmed by the pseudo-data fit in Section 4.4.5.

In addition, the results obtained using E_1/E_2 with background subtracted are similar to those shown earlier in Table 4.7 where background was ignored. The fit using E_1/E_2 is insensitive to the presence of background but the fit using BDTG is. A test using MC is presented in Appendix A.4 in order to understand the impact of background on these two fits.

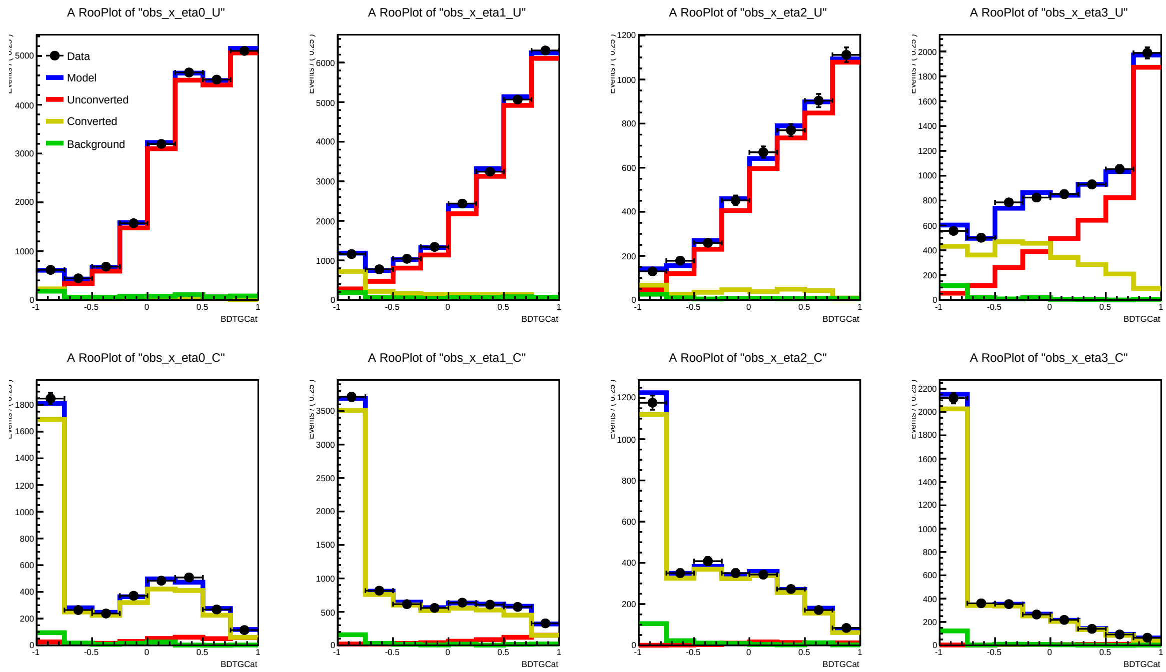


Figure 4.27: Fit to data using BDTG with the background (green) subtracted. The blue histograms correspond to the model, which is the sum of true unconverted (red) and converted (yellow). The top four plots correspond to the channels of unconverted photons at four $|\eta|$ bins, denoted eta0, eta1, eta2 and eta3. The bottom four plots correspond to the channels of converted photons.

The fitted values of the nuisance parameters are shown in Figure 4.28. Comparing only the nuisance parameters related to the material since the R_ϕ-related systematics are not relevant for the fit using E_1/E_2, good agreement is observed. For both fits, the MaterialFMX_eta2 nuisance parameter is pulled in the negative direction, indicating an excess of material in the simulation. This is in accordance with the results of material determination in Ref. [61].

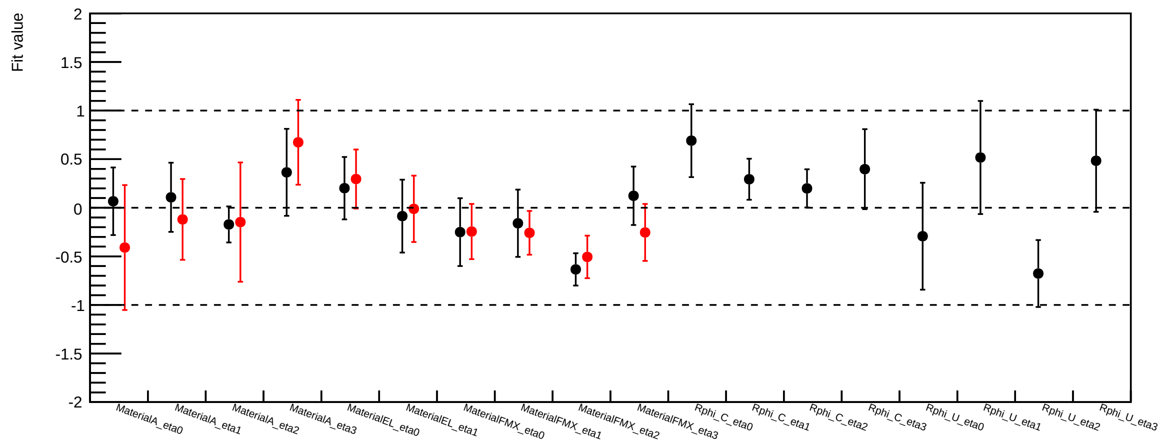


Figure 4.28: Fitted values of the nuisance parameters from the fit using E_1/E_2 (in red) and BDTG (in black) to the data, with the background contribution subtracted.

	E_1/E_2		BDTG	
η	f_{fake}	f_{reco}	f_{fake}	f_{reco}
$\lvert\eta\rvert < 0.6$	0.013 ± 0.005	0.921 ± 0.039	0.015 ± 0.003	0.890 ± 0.012
$0.6 < \lvert\eta\rvert < 1.37$	0.029 ± 0.004	0.801 ± 0.015	0.027 ± 0.005	0.804 ± 0.009
$1.52 < \lvert\eta\rvert < 1.8$	0.016 ± 0.016	0.906 ± 0.026	0.017 ± 0.015	0.904 ± 0.018
$1.8 < \lvert\eta\rvert < 2.37$	0.007 ± 0.007	0.566 ± 0.013	0.014 ± 0.008	0.564 ± 0.012

Table 4.8: The results of the fit to data using E_1/E_2 (left) and BDTG (right) variable, with the background subtracted.

4.5 Conclusion

An alternative method to classify photon as converted or unconverted using information about the electromagnetic shower shape has been studied. It is pile-up robust and performs better than the default track-based conversion reconstruction at high pile-up environment, based on the version of the reconstruction available during the Long Shutdown 1. The conversion reconstruction efficiency is stable by construction and smaller increase is seen in the fake conversion rate compared to the track-based reconstruction. It is also possible to add tracking information to the classification method. The method has been validated on 2012 data.

In the second part of the study, the conversion reconstruction efficiency and fake conversion rate in data are measured, using a template fit method. Compared to using only the distribution of E_1/E_2 in the fit, it is found that using more variables (here E_1/E_2 and R_ϕ combined through a boosted decision tree) yields a more precise measurement on the efficiency of the track-based conversion reconstruction due to the increased separation of true converted and unconverted photons. The same values of the fractions are found using both methods with the combination of variables giving smaller error sizes on the conversion reconstruction efficiency. The increased separation also comes with a caveat. The contribution of jets faking photons has to be subtracted if more variables are used as their shower looks similar to those of true converted photons. A small amount, about 4% of background in $Z \to \mu\mu\gamma$ data, results in a bias in the measured values when it is not taken into account. In the case where the purity cannot be determined precisely or where background subtraction without using shower shape-based identification criteria is difficult or impossible, extraction of conversions fractions by fitting using E_1/E_2 would be preferred.

Sensitivity to the amount of material upstream of the calorimeter is also seen in the shower shape. One can estimate the radius of conversion from the multivariate output. The template fit using a classifier with two variables as input is also able to constrain the amount of material. A known but uncorrected deficit of material in data compared to simulation in $1.5 < |\eta| < 1.6$ is observed in the template fit results.

Chapter 5

Statistics in High Energy Physics

In new physics searches, a critical question to answer is whether the observed anomaly (if present) with respect to the known background prediction is significant or not. This is quantified with a p-value, which is the probability of obtaining a result at least as extreme as observed given the null (background-only) hypothesis is true. When a search is performed in a large range, for example scanning over a large mass range, the so-called *look-elsewhere effect*, taking into account that it is more likely to observe a fluctuation, must be estimated.

Often, due to the large search range, the *look-elsewhere effect* means that the p-value computed is not as significant since fluctuations could occur anywhere within the search range. Global p-value, taking into account this effect is discussed. In the case of no discovery, upper limits are set on the process which can be used to constrain or exclude certain BSM models. Statistical techniques in many ATLAS analyses are based on frequentist approach. A profile likelihood ratio test statistic is used to quantify the agreement with the background-only hypothesis and for determining the exclusion limits.

Section 5.1 presents the profile likelihood method. The computation of the discovery p-values is described in Section 5.2. The look-elsewhere effect and the global p-values are discussed in Section 5.3. Section 5.4 outlines the methods to compute limits. Finally, the computation of p-values and upper limits using asymptotic approximation is described in Section 5.5 and the Asimov dataset in Section 5.6.

The statistical methods reported here are used in the $A \to Zh$ and high mass diphoton resonances search analyses described in Chapters 6, 8 and 9.

5.1 Profile Likelihood Method

Suppose we have a set of n measured quantities $\boldsymbol{x} = (x_1, ..., x_n)$ described by the signal and background probability density functions (pdfs), $f_s(x_i; \boldsymbol{\theta})$ and $f_b(x_i; \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is the set of nuisance parameters, the likelihood of the dataset can be expressed as

$$L(\mu, \boldsymbol{\theta}|\boldsymbol{x}) = e^n \prod_{i=1}^{n} [s(\mu) \cdot f_s(x_i; \boldsymbol{\theta}) + b \cdot f_b(x_i; \boldsymbol{\theta})]. \tag{5.1}$$

In the equation, $s(\mu)$ is the number of signal events and b is the number of background events. The number of signal events s is related to the parameter of interest, μ, which is typically signal strength or cross-section.

The maximum likelihood estimators, $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ are the values of the parameters that maximise the likelihood function $L(\mu, \boldsymbol{\theta})$, or equivalently, minimise $-\ln L(\mu, \boldsymbol{\theta})$. The conditional maximum likelihood estimator $\hat{\hat{\boldsymbol{\theta}}}(\mu)$ is the value of $\boldsymbol{\theta}$ that maximises the likelihood function with μ fixed.

To test a hypothesized value of μ, we consider the profile likelihood ratio

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})}. \tag{5.2}$$

Usually, only $\mu \geq 0$ is considered as a valid signature for signal. However, it is often convenient to allow $\mu < 0$ as long as the pdf is positive everywhere. Obtaining $\hat{\mu} < 0$ indicates a deficit of events in data with respect to the background only hypothesis. The likelihood ratio $\lambda(\mu)$ lies between 0 and 1, with 1 meaning a good agreement between the data and the hypothesised value of μ.

Equivalently, it is convenient to use the test statistic

$$t_\mu = -2 \ln \lambda(\mu). \tag{5.3}$$

5.2 Discovery p-value

When searching for a new phenomenon, one tries to exclude the background-only hypothesis H_0 that the data are consistent with known processes. The p-value of the null hypothesis ($\mu = 0$) quantifies the level of disagreement between the data and the null hypothesis. The probability of obtaining a signal at least as large as the observed one, under the assumption that no signal is present, is denoted p_0. A small value of p_0 suggests the presence of a significant signal. A p_0 of 2% does not mean that there is a 2% probability that the background-only hypothesis is true, but rather that there is a 2% probability to obtain data as observed under the background-only assumption.

To test for discovery, the one-sided test statistic is used:

$$t_0^{\text{uncap}} = \begin{cases} -2\ln \frac{L(0,\hat{\hat{\boldsymbol{\theta}}}(0))}{L(\hat{\mu},\hat{\boldsymbol{\theta}})} & \hat{\mu} \geq 0, \\ +2\ln \frac{L(0,\hat{\hat{\boldsymbol{\theta}}}(0))}{L(\hat{\mu},\hat{\boldsymbol{\theta}})} & \hat{\mu} < 0. \end{cases} \tag{5.4}$$

The p-value is calculated as

$$p_0 = \int_{t_{0,\text{obs}}}^{\infty} f(t_0|0) dt_0, \tag{5.5}$$

where $t_{0,\text{obs}}$ is the test statistic observed in data and $f(t_0|0)$ denotes the pdf of the test statistic.

The discovery p-value is usually expressed as a *significance* Z defined so that a Z standard deviation upward fluctuation of a Gaussian random variable would have an upper tail integral equal to the p-value. That is,

$$Z = \Phi^{-1}\left(1 - p_0(t_0^{\text{uncap}})\right), \tag{5.6}$$

where Φ^{-1} is the inverse of the cumulative distribution of the standard Gaussian. In high energy physics, the standard convention of the level of significance where a discovery can be claimed is Z=5, i.e. a 5σ effect, corresponding to a p-value of 2.87×10^{-7}.

5.3 Look-elsewhere Effect

In the searches presented in this thesis, a wide range of mass or width of the hypothesized resonance is scanned. The p-value obtained at a given mass and/or width is denoted as a local p-value. Fluctuations of data leading to a large Z_0^{local} are more common than the corresponding p_0 probability would suggest since they could occur at any mass or width. This is referred to as *look-elsewhere effect* (LEE).

There are several ways to take into account the look-elsewhere effect. The most straightforward way is to simply run many background-only pseudo-experiments and finding the largest significance for each experiment. This procedure is however unrealistic, as it is extremely CPU consuming.

A simplified method using thousands of pseudo-experiments is employed in the high mass diphoton analysis. In this method, a number of background-only pseudo-experiments are generated from background-only maximum-likelihood fit on the observed data. In the generation of each pseudo-experiment, the global observables (e.g. background parameters and normalisation) are randomised according to a Gaussian constraint pdf with a mean value equal to the profiled value of the corresponding nuisance parameter from the fit to the observed data and a width value defined by the covariance matrix of the fit. In each pseudo-dataset, the maximum Z_0^{local} in the signal search volume is determined. To determine the maximum Z_0^{local}, one usually scans the entire mass and width grid to make sure the fit does not converge on a local minimum. As this is computationally intensive, a number of maximum-likelihood fits are performed on each pseudo-dataset instead. The initial values of the parameters of interest (m_X, signal width and signal cross-section) are randomised within the signal search range prior to each fit. The fit with the minimum negative-log-likelihood value is assumed to correspond to the true maximum Z_0^{local}. From the distribution of Z_0^{local} in the fits to pseudo-experiments, the Z_0^{global} can be inferred by integrating the observed Z_0^{local} to infinity:

$$Z_0^{global} = \int_{Z_0^{local}}^{\infty} f(Z_0^{local})\, dZ_0^{local}. \tag{5.7}$$

Another method which requires even less number of pseudo-experiments (around twenty) is based on random fields. This was studied in Ref. [76] and then generalised to multi-dimensional searches in Ref. [77].

5.4 Upper Limits

Upper limits on μ are computed using the one-sided test statistic defined in Ref. [78]:

$$\tilde{q}_\mu = \begin{cases} -2\ln \frac{L(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(0,\hat{\hat{\boldsymbol{\theta}}}(0))} & \hat{\mu} < 0, \\ -2\ln \frac{L(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(\hat{\mu},\hat{\boldsymbol{\theta}})} & 0 \leq \hat{\mu} \leq \mu, \\ 0 & \hat{\mu} > \mu, \end{cases} \tag{5.8}$$

For $\hat{\mu} > \mu$, the test statistics $\tilde{q}_\mu$ is set to zero because one would not exclude hypotheses below the maximum likelihood estimator of μ (that would be relevant for lower limits). The region $\hat{\mu} < 0$ is also treated in a special way by replacing the likelihood value at $\hat{\mu} < 0$ by the value at $\mu = 0$.

While 5σ is generally required to claim a discovery, upper limit setting on the signal rate is less stringent. Usually, a 95% confidence level (C.L.) is required. A 95% C.L., one-sided upper limit on μ, denoted μ_{up} is obtained by solving for $p_{\mu_{up}} = 5\%$. This is referred to as the CL_{s+b} method as the p-value of the $s+b$ hypothesis is used.

By definition, a 95% C.L. upper limit using the CL_{s+b} procedure accidentally excludes the true parameter value 5% of the time. This results in the possibility that the CL_{s+b} procedure excludes the hypotheses to which one has little or no sensitivity. If the expected number of signal events s is much less than that of background b, and the observed number of events has a downward fluctuation, then this value of s will be excluded with a probability of 5%. This is perceived as undesirable.

The CL_s method mitigates the problem of unrealistic limits that are better than the experimental sensitivity by effectively penalising the p-value of a tested parameter by an amount that increases with decreasing sensitivity [79]. To calculate the CL_s upper limit, the p-value calculation is modified to use[1]

$$\begin{aligned} p_{\mu_{up}}^{CL_s} &= \frac{p_{\mu_{up}}}{p_b}, \\ &= \frac{\int_{\tilde{q}_{\mu_{up}}}^{\infty} f(\tilde{q}_\mu|s+b)d\tilde{q}_\mu}{\int_{\tilde{q}_{\mu_{up}}}^{\infty} f(\tilde{q}_\mu|b)d\tilde{q}_\mu}, \end{aligned} \tag{5.9}$$

where p_b is the p-value for the background-only hypothesis. The CL_s upper limit is therefore more conservative than that from the CL_{s+b} method. The coverage probability of the CL_s procedure can, depending on the true value of μ, be greater than 95%.

5.5 Asymptotic Approximation

The sampling distribution for the test statistic is required in order to compute the p-value using Eq. 5.5. For a discovery, one needs to construct the pdf $f(t_0|0)$, where t_0 is the discovery test statistic. For setting upper limits, the pdf $f(\tilde{q}_\mu|s+b)$ and $f(\tilde{q}_\mu|b)$ for different μ are required. These involve the generation of many pseudo-experiments that are computationally expensive.

In the asymptotic limit of a large data sample, the maximum likelihood estimator of the parameter of interest $\hat{\mu}$ follows a Gaussian distribution with standard deviation σ and the profile log-likelihood ratio is parabolic [80]:

$$-2\ln\lambda(\mu) = \frac{(\mu-\hat{\mu})^2}{\sigma^2} + \mathcal{O}(1/\sqrt{N}). \tag{5.10}$$

The $\mathcal{O}(1/\sqrt{N})$ term can be neglected in the large sample limit.

According to the Wilks' theorem [81], if $\hat{\mu}$ is Gaussian distributed, then in the asymptotic limit, the distribution of the test statistic follows a χ^2 distribution for one degree of freedom. If the assumed strength μ' is different from the tested μ, then the test statistic follows a *noncentral* χ^2 distribution with noncentrality parameter Λ

$$\Lambda = \frac{(\mu-\mu')^2}{\sigma^2}. \tag{5.11}$$

In the asymptotic approximation [78], the probability and cumulative distribution of t_0^{uncap} are:

$$f(t_0^{\text{uncap}}) = \begin{cases} \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{t_0^{\text{uncap}}}}e^{-t_0^{\text{uncap}}/2} & \hat{\mu}\geq 0, \\ \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{-t_0^{\text{uncap}}}}e^{t_0^{\text{uncap}}/2} & \hat{\mu}< 0, \end{cases} \tag{5.12}$$

$$F(t_0^{\text{uncap}}) = \begin{cases} \Phi(\sqrt{t_0^{\text{uncap}}}) & \hat{\mu}\geq 0, \\ 1-\Phi(\sqrt{-t_0^{\text{uncap}}}) & \hat{\mu}< 0. \end{cases} \tag{5.13}$$

Therefore, the p-value of the $\mu=0$ hypothesis is

$$p_0 = \begin{cases} 1-\Phi(\sqrt{t_0^{\text{uncap}}}) & \hat{\mu}\geq 0, \\ \Phi(\sqrt{-t_0^{\text{uncap}}}) & \hat{\mu}< 0, \end{cases} \tag{5.14}$$

and the local discovery significance is simply

$$Z = \Phi^{-1}\left(1-p_0(t_0^{\text{uncap}})\right) = \begin{cases} +\sqrt{t_0^{\text{uncap}}} & \hat{\mu}\geq 0, \\ -\sqrt{-t_0^{\text{uncap}}} & \hat{\mu}< 0. \end{cases} \tag{5.15}$$

[1] In some literature including Ref. [79], a different convention of p_b is used.

Assuming the validity of the Wald approximation, the test statistic used for upper limits can be written as [78]

$$\tilde{q}_\mu = \begin{cases} \frac{\mu^2}{\sigma^2} - \frac{2\mu\hat{\mu}}{\sigma^2} & \hat{\mu} < 0, \\ \frac{(\mu-\hat{\mu})^2}{\sigma^2} & 0 \leq \hat{\mu} \leq \mu, \\ 0 & \hat{\mu} > \mu. \end{cases} \tag{5.16}$$

Using this test statistic, the pdf $f(\tilde{q}_\mu|\mu)$ is found to be [78]

$$f(\tilde{q}_\mu|\mu) = \frac{1}{2}\delta(\tilde{q}_\mu) + \begin{cases} \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\tilde{q}_\mu}}e^{-\tilde{q}_\mu/2} & 0 < \tilde{q}_\mu \leq \mu^2/\sigma^2, \\ \frac{1}{\sqrt{2\pi}(2\mu/\sigma)} \exp\left[-\frac{1}{2}\frac{(\tilde{q}_\mu+\mu^2/\sigma^2)^2}{(2\mu/\sigma)^2}\right] & \tilde{q}_\mu > \mu^2/\sigma^2, \end{cases} \tag{5.17}$$

and the corresponding cumulative distribution is [78]

$$F(\tilde{q}_\mu|\mu) = \begin{cases} \Phi(\sqrt{\tilde{q}_\mu}) & 0 < \tilde{q}_\mu \leq \mu^2/\sigma^2, \\ \Phi(\frac{\tilde{q}_\mu+\mu^2/\sigma^2}{2\mu/\sigma} & \tilde{q}_\mu > \mu^2/\sigma^2. \end{cases} \tag{5.18}$$

Solving for $p^{CL_s}_{\mu_{up}} = \frac{1-F(\tilde{q}_\mu|\mu)}{F(\tilde{q}_\mu|0)} = \alpha$, the median expected CL_s limit on μ is

$$\mu^{\text{med}}_{up} = \sigma\Phi^{-1}(1 - 0.5\alpha), \tag{5.19}$$

and the $\pm N\sigma$ error band is given by

$$\mu_{up\pm N} = \sigma(\Phi^{-1}(1 - \alpha\Phi(\pm N)) \pm N), \tag{5.20}$$

where $\alpha = 0.05$ for a 95% confidence level.

5.6 Asimov Dataset

The median experimental significance and its expected statistical variation can be provided by a single representative dataset [78], referred to as the *Asimov* data set, instead of an ensemble of pseudo-experiments. The Asimov dataset is defined so that when it is used to evaluate the maximum likelihood estimator of a parameter, one obtains the input values of the parameter itself. It is effectively a *perfect* dataset, without any statistical fluctuations.

In the asymptotic regime, the result obtained from the Asimov dataset matches the median of the results obtained with pseudo-experiments. By using the Asimov data set, one can easily obtain the median values of the test statistics.

Chapter 6

Search for a CP-odd Higgs Boson A Decaying to Zh with a $jj\gamma\gamma$ Final State at $\sqrt{s} = 8$ TeV

In the two-Higgs-doublets models (2HDM) [20], the neutral CP-odd pseudoscalar A boson could decay to a Z boson and a h boson. h is assumed to correspond to the recently discovered 125 GeV boson. The decay of $A \to Zh$ is the primary decay mode of the A boson when its mass m_A lies between $m_h + m_Z \approx 220$ GeV and two times the mass of the top quark. Depending on the model parameters, this decay mode could be advantageous even for higher values of m_A. The A boson is mainly produced through gluon-fusion and in association with b-quarks [20]. In this study, only the gluon-fusion production mechanism is considered as it is the dominant mode in most of the 2HDM parameter space, the b-quark associated mode being relevant for type-II 2HDM in the high $\tan\beta$ region.

In ATLAS, searches for an A boson decaying to Zh have been performed with 8 TeV data in the $\ell\ell\tau\tau$, $\ell\ell bb$, $\nu\nu bb$ final states [82], where h is identified by its decay modes, $h \to bb$ and $h \to \tau\tau$. No evidence for the A boson was found. Among the searches, the bb channels have the best sensitivity.

In this chapter, a feasibility study of the search for the A boson decaying to Zh in the 220-1000 GeV mass range is presented, using 20.3 fb^{-1} of $\sqrt{s} = 8$ TeV pp collisions collected by ATLAS in 2012, with jets and photons in the final states. The Z boson decays predominantly to hadrons (jets) whereas the Higgs decay to diphoton offers a clean signature with excellent resolution despite its rarity. If the mass of A m_A is large, one can search for an excess of events in the four-body (two jets and two photons) invariant mass spectrum, over a smoothly falling background. However, since the background spectrum peaks at around 260 GeV [1], a low-m_A signal would appear as a small excess on top of this large peak. Therefore, the search is divided in two ranges of the resonance mass. In the low-mass region ($220 \leq m_A < 400$ GeV), the diphoton invariant mass $m_{\gamma\gamma}$ is used to search for an excess, whereas at high mass ($400 \leq m_A \leq 1000$ GeV), we exploit the four-body invariant mass instead.

The study presented in this chapter was carried out by myself, hence I was involved in all stages of the study from the beginning till the end, starting from the request and validation of the MC samples to the event selection optimisation and the final analysis strategy.

Section 6.1 outlines the data and MC samples used in this study. Event selection and optimisations are described in Section 6.2. The statistical analysis is presented in Section 6.3. Finally, the result is shown in Section 6.4.

[1] This is due to the kinematical requirement on the transverse momentum of the photon and jet candidates as well as requirement on the diphoton and dijet invariant mass.

6.1 Data and Monte Carlo Samples

All Monte Carlo samples are generated with a centre-of-mass energy of 8 TeV.Non-resonant diphoton background is obtained via data-driven methods. The data used in this analysis were recorded by the ATLAS detector during the 2012 run and have an integrated luminosity of 20.3 fb^{-1}. The SM Higgs background is acquired from simulated samples.

6.1.1 Signal Samples

Signal samples for an A boson decaying to Zh are produced for m_A between 220 and 1000 GeV with the final states of $h \to \gamma\gamma$ and $Z \to jj$. The A boson production mode considered here is gluon-gluon fusion. These samples are generated with `Madgraph 5` [83]. The parton showering and hadronisation is performed with `Pythia8` [68]. In the generation, the Higgs mass m_h is fixed at 125 GeV and $m_A = m_H = m_{H^\pm}$, where H and $H^\pm$ are the other heavy Higgs bosons predicted by the model. A narrow width of 1 GeV is assumed for the A, H and $H^\pm$ bosons.

Ten samples are generated, each with a different value of the mass of the A boson: 220, 240, 260, 300, 340, 350, 400, 500, 800 and 1000 GeV. For each mass point, 40 thousand events are simulated.

The expected cross-sections of a gluon-fusion produced A boson, its branching ratios to Zh and the branching ratios of h to $\gamma\gamma$ for some 2HDM scenarios calculated with `SusHi` [84] and `2HDMC` [85] are shown in Table 6.1. The branching ratios of h to $\gamma\gamma$ differ only slightly from the SM value for $m_h = 125$ GeV, which is 0.00228 [15].

In the analysis, the cross-sections predicted by Type-I 2HDM with $\sin(\beta - \alpha) = 0.9998$ and $\tan\beta = 1$ are used to optimise the event selection as $\sin(\beta - \alpha) = 0.9998$ corresponds to the SM-like limit.

6.1.2 Background Samples

There are two sources of background in this search. The first is non-resonant QCD background which is expected to be smoothly falling. The non-resonant background is data-driven. The second source of background is from the SM Higgs, which is resonant in $m_{\gamma\gamma}$. This is relevant for the low-mass analysis.

6.1.2.1 Non-resonant Background

The background consists primarily of non-resonant QCD production of two photons with at least two jets. Additional contributions to the background come from events with one photon and multiple jets or multijet events when one or two jets are misidentified as photons, as well as events with a true Z boson produced together with photons or jets misidentified as photons.

The kinematic properties of these background contributions are studied directly on data by using the sidebands of the $m_{\gamma\gamma}$ spectrum. The signal region is defined as $2\sigma_{\gamma\gamma}$ around m_h, where $\sigma_{\gamma\gamma} = 1.6$ GeV is the diphoton invariant mass resolution and $m_h = 126$ GeV is the value of the mass of the Higgs boson measured in the diphoton decay channel by ATLAS [86]. The sidebands are defined as: $[110 \text{ GeV}, m_h - 2\sigma_{\gamma\gamma}]$ and $[m_h + 2\sigma_{\gamma\gamma}, 160 \text{ GeV}]$.

The amount and the shape of the non-resonant background in the diphoton and four-body invariant mass spectra are extracted from data for the low- and high-mass analyses respectively. The procedure is described in Section 6.3.

6.1.2.2 SM Higgs Background

The 125 GeV $h \to \gamma\gamma$ resonance is a background to the search analysis when there are two reconstructed jets with invariant mass close to the Z boson mass present in the event. While its contribution is quite small relative to the non-resonant background, its resonant nature mimics

m_A [GeV]	$\sigma(gg \to A)$ [pb]	$\mathcal{B}(A \to Zh)$	$\mathcal{B}(h \to \gamma\gamma)$
	Type-I, $\sin(\beta-\alpha) = 0.9998$, $\tan\beta = 1$		
220	16.0784	0.00555	0.00193
240	13.8724	0.07070	0.00194
260	12.3491	0.14160	0.00194
300	11.0507	0.21421	0.00195
340	15.7024	0.05330	0.00195
350	20.0308	0.00227	0.00195
400	8.58149	0.00205	0.00196
500	2.32646	0.00325	0.00196
800	0.12097	0.00941	0.00196
1000	0.02564	0.01539	0.00196
	Type-I, $\sin(\beta-\alpha) = 0.9998$, $\tan\beta = 10$		
220	0.16078	0.35832	0.00208
240	0.12349	0.94284	0.00209
260	0.12349	0.94284	0.00209
300	0.11050	0.96461	0.00209
340	0.15702	0.84919	0.00209
350	0.20030	0.18541	0.00209
400	0.08581	0.17068	0.00210
500	0.02326	0.24593	0.00210
800	0.00120	0.48722	0.00210
1000	0.00025	0.60989	0.00210
	Type-II, $\sin(\beta-\alpha) = 0.9998$, $\tan\beta = 10$		
220	0.02009	0.00091	0.00216
240	0.01362	0.01233	0.00217
260	0.00946	0.02635	0.00218
300	0.00485	0.04261	0.00218
340	0.00265	0.00896	0.00219
350	0.00230	0.00036	0.00219
400	0.00118	0.00032	0.00219
500	0.00036	0.00052	0.00220
800	2.37e-05	0.00151	0.00220
1000	5.53e-06	0.00249	0.00220

Table 6.1: Cross-sections of gluon-fusion produced A boson, its branching ratios to Zh and the branching ratios of h to $\gamma\gamma$ for some 2HDM scenarios. They are calculated with `SusHi` [84] and `2HDMC` [85].

the A boson decay in the low-mass search since the fit is performed on the diphoton invariant mass.

The $m_h = 125$ GeV samples for the five Standard Model Higgs production modes are used: gluon-gluon fusion (ggF), vector boson fusion (VBF), associated production (Wh and Zh), and $t\bar{t}h$ production. The rates for these production modes as well as the branching ratio to two photons are taken from the Handbook of LHC Higgs Cross Sections 3 [15] hence slightly different from the numbers presented in Tables 1.1 and 1.2.

6.2 Event Selection and Optimisations

The object definition and event pre-selection are presented. The kinematical selection is then optimised for sensitivity to the A boson signal by taking advantage of the different kinematics of the A boson decay and the SM background.

6.2.1 Object Definition

6.2.1.1 Photons

Only photons reconstructed within the calorimeter acceptance ($|\eta| < 2.37$ excluding $1.37 < |\eta| < 1.56$) are considered as candidates. The same acceptance is used also in the SM $h \to \gamma\gamma$ analysis in Run 1 and is more stringent than the usual crack-region definition of $1.37 < |\eta| < 1.52$.

They are required to pass the tight identification criteria (see Section 3.3) as well as the default $h \to \gamma\gamma$ track and calorimeter isolation requirements. The track isolation in a cone of $\Delta R = 0.2$ is required to be less than 2.6 GeV, and the calorimeter isolation in a cone of size $\Delta R = 0.4$ around the photon has to be lower than 6 GeV.

6.2.1.2 Jets

Jets are reconstructed from three-dimensional topological clusters of energy in the electromagnetic and hadronic calorimeters using the anti-k_t algorithm [87] with a distance parameter of $R = 0.4$. Jet candidates are required to have a transverse energy greater than 20 GeV and be within $|\eta| < 4.5$. In order to suppress jets from pile-up, the jet vertex fraction (JVF) of the jets within the tracking acceptance ($|\eta| < 2.4$) is required to be less than 0.25. JVF is defined as the ratio of the sum of transverse momenta of tracks associated with the jet that are produced at the diphoton primary vertex and the sum of transverse momenta of tracks associated with the jet from all primary vertices [88].

6.2.2 Event Pre-selection

The candidate events are pre-selected with the following criteria:

- The data are required to satisfy a number of conditions to ensure that the ATLAS detector was operational with good efficiency during data taking.

- Data events that are incomplete, corrupted, or have problems in the calorimetric system are removed. Events with a trip in the tile calorimeter are also rejected.

- Events are selected using a diphoton trigger, `EF_2g20vh_medium`, that requires at least 20 GeV transverse momentum for the two leading photons. The trigger level photon p_T cut is lower than the `g35_loose_g25_loose` trigger used in the standard $h \to \gamma\gamma$ analysis. On the other hand, the requirement on photon identification based on shower shapes is tighter than the loose trigger.[2]

- The events are required to have at least one reconstructed primary vertex with at least three associated tracks.

- At least two reconstructed photons passing tight identification criteria with the leading-p_T photon having transverse momentum greater than 25 GeV and the subleading 22 GeV.

[2]In the *medium* photon identification selection, the loose photon identification selection is applied with additional requirements on the shower shape variables w_{stot} and E_{ratio}.

The two highest p_T photons are chosen. To precisely measure the diphoton invariant mass, an accurate estimate of the location of the diphoton production vertex is necessary. The *photon pointing method* provides the direction of flight of the photons by combining their trajectories, measured using the longitudinal segmentation of the calorimeter, with constraints from the average collision point of the proton beams. A neural-network algorithm selects the diphoton production vertex from the reconstructed collision vertices using a number of information. It combines photon pointing information with the sum of the squared momenta Σp_T^2 and the scalar sum of the momenta Σp_T of the tracks associated with each reconstructed vertex, and the difference in azimuthal angle $\Delta\phi$ between the direction defined by the vector sum of the track momenta and that of the diphoton system.

The photon four-momenta are then recalculated according to the selected diphoton vertex and the diphoton invariant mass $m_{\gamma\gamma}$ is required to be within $110 < m_{\gamma\gamma} < 160$ GeV.

The jet pair is formed by choosing two jets with the highest p_T and a loose invariant mass requirement of $50 < m_{jj} < 130$ GeV is imposed. Further optimisation of the dijet invariant mass window will be performed in the following.

6.2.3 Optimised Kinematic Selection

In order to optimise the kinematic selection, we use significance as a figure of merit. The signal significance is defined as [78]

$$\text{Significance} = \sqrt{2 \cdot \left((S+B) \cdot \ln(1+\frac{S}{B}) - S \right)}. \tag{6.1}$$

Optimisation is performed by varying the relevant cuts and finding the cuts with the largest significance. This is done for each mass hypothesis but in order to keep the selection as simple as possible, a common loose cut is preferred. The value of the cut that gives the maximum expected significance for a relatively low A mass (about $m_A = 240 - 260$ GeV) is chosen since the amount of background at higher A mass is small. Instead of a common absolute cut, if necessary, the cut can be parameterised as a function of the A mass.

In the optimisation procedure, the signal is scaled to the expected number in 2012 collision data assuming Type-I 2HDM with $\sin(\beta - \alpha) = 0.9998$ and $\tan\beta = 1$. The number of background events is taken from both data and MC. The expected non-resonant background contribution is derived from data events in the sidebands outside the signal region in $m_{\gamma\gamma}$. An exponential function is used to characterise the shape of the background in $m_{\gamma\gamma}$ and to estimate the number of background events in the signal region from the number of events in the sidebands. The resonant background which is the SM Higgs background is obtained from MC. The SM Higgs events from the five processes in Section 6.1.2.2 are scaled to their respective expectations. The signal is optimised against the sum of resonant and non-resonant background.

6.2.3.1 Photon p_T

The standard $h \to \gamma\gamma$ analysis uses a relative transverse momentum cut of $p_T/m_{\gamma\gamma} > 0.35$ for the photon with the highest p_T and $p_T/m_{\gamma\gamma} > 0.25$ for the photon with the second-highest p_T. Nevertheless, the Higgs boson from the decay of the A boson has a transverse momentum that increases as a function of the mass m_A, which results in the photons having higher p_T. The p_T distributions for the leading and subleading photon candidates are shown in Figure 6.1 for background events in data (see Section 6.1.2.1) and for signal events from the MC samples with several A mass hypotheses. A tighter p_T requirement than the one used in the standard analysis could be applied to reduce the amount of background.

For different cut values on the photon p_T (either relative or absolute cut), the gain in significance with respect to pre-selection requirement ($p_T > 25$ GeV for the leading photon and

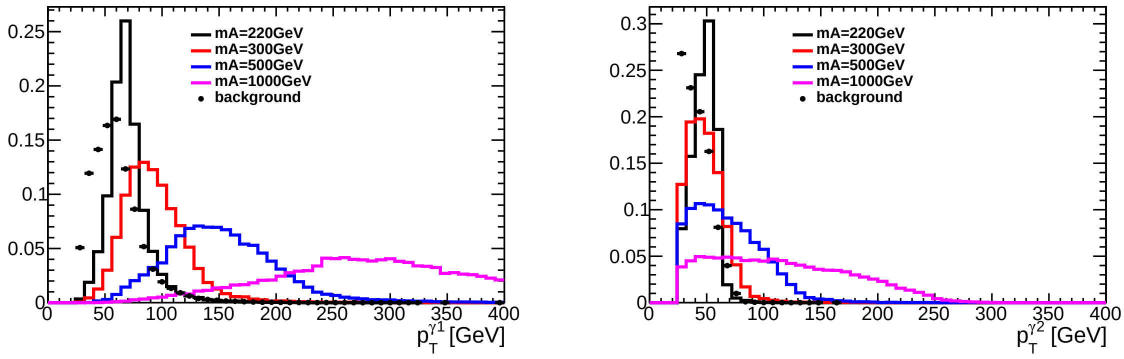


Figure 6.1: Distribution of the transverse momentum of the candidate photons for background events (dots) compared to signal events generated with $m_A = 220$ (black), 300 (red), 500 (green) and 1000 (blue) GeV. Left: leading photon candidate; right: subleading photon candidate.

$p_T > 22$ GeV for the subleading photon) is computed. Varying only the cut on the leading photon p_T and keeping the subleading photon p_T cut of $p_T/m_{\gamma\gamma} > 0.25$, the gain in significance for various low mass hypotheses as a function of the cut is shown in Figure 6.2a. A leading photon p_T cut of $p_T/m_{\gamma\gamma} > 0.5$ is found to be optimal for $m_A = 260$ GeV. The signal and background efficiencies for this p_T requirement are shown in Figure 6.2b where more than half of the background is rejected while the signal efficiency remains relatively high, from 63% at $m_A = 220$ GeV to 96% at higher masses.

In Figure 6.2c, the gain in significance as a function of m_A is shown for two tested cuts on the leading photon p_T, the one used in the standard analysis ($p_T^{\gamma 1}/m_{\gamma\gamma} > 0.35$) and the optimal cut for $m_A = 260$ GeV ($p_T^{\gamma 1}/m_{\gamma\gamma} > 0.5$). The significance with the tighter cut is higher everywhere except at $m_A = 220$ GeV. $p_T/m_{\gamma\gamma} > 0.5$ is optimal for $m_A = 260$ GeV and only results in a tiny loss in significance for $m_A = 220$ GeV. With the cut on the leading photon p_T chosen, a tighter requirement on the subleading photon p_T does not bring significant improvement and thus $p_T^{\gamma 2}/m_{\gamma\gamma} > 0.25$ is kept.

Absolute p_T cuts were also considered, but they do not outperform relative p_T cuts. A tight cut on the absolute photon p_T also distorts the exponential shape of the $m_{\gamma\gamma}$ distributions. In order to keep the background $m_{\gamma\gamma}$ shape exponential, the p_T cuts are restricted to relative cuts. The effect of an absolute p_T cut on the $m_{\gamma\gamma}$ distribution can be seen in Figure 6.2d.

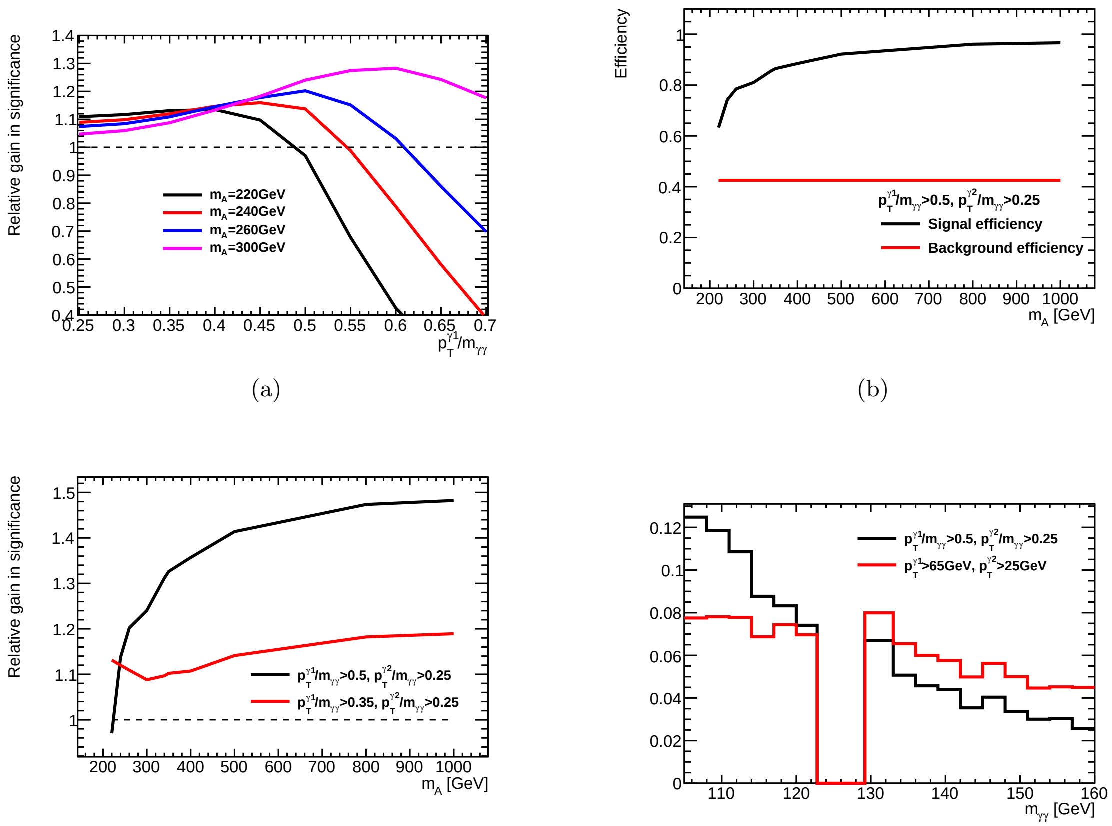
6.2.3.2 Jet p_T

The transverse momentum requirement on the jets can also be tightened to reduce the amount of background, as demonstrated in Figure 6.3. Varying only the cut on the leading jet p_T and keeping subleading jet p_T requirement at $p_T > 25$ GeV, the gain in significance compared to using a pre-selection cut of 22 GeV on both jets is shown in Figure 6.4a for various low mass hypotheses. The optimal p_T cut values for $m_A = 240$ and 260 GeV are 35 GeV for the leading-p_T jet and 25 GeV for the subleading-p_T jet. With these cuts, 42% of the background is eliminated while the relative signal efficiency remains high, from 77% at $m_A = 220$ GeV to 99% at higher masses, as shown in Figure 6.4b.

6.2.3.3 Improving the Resolution of the Reconstructed Mass of the A boson

In order to improve the resolution of the reconstructed mass of the A boson m_A^{rec}, the four-body invariant mass is scaled by the following formula:

$$m_A^{\text{rec}} = m_{jj\gamma\gamma} - m_{jj} + m_Z, \tag{6.2}$$



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

Figure 6.2: (a) Gain in significance with respect to the pre-selection cuts (25 GeV and 22 GeV for the leading and the subleading photon candidate respectively) as a function of $p_T^{\gamma 1}/m_{\gamma\gamma}$ cut for several hypotheses on the A boson mass: $m_A = 220$ (black), 240 (red), 260 (blue) and 300 (violet) GeV. The subleading photon is required to satisfy $p_T/m_{\gamma\gamma} > 0.25$. (b) Signal and background efficiencies across m_A of the optimal cut found for $m_A = 260$ GeV, that is $p_T^{\gamma 1}/m_{\gamma\gamma} > 0.5$ and $p_T^{\gamma 2}/m_{\gamma\gamma} > 0.25$. The efficiency is with respect to the pre-selection cuts. (c) Gain in significance with respect to the pre-selection cuts as a function of m_A comparing two tested cuts: $p_T^{\gamma 1}/m_{\gamma\gamma} > 0.5$ and $p_T^{\gamma 1}/m_{\gamma\gamma} > 0.35$. (d) $m_{\gamma\gamma}$ distribution of the diphoton background for the chosen relative cuts versus an alternative absolute cut of $p_T^{\gamma 1} > 65$ GeV and $p_T^{\gamma 2} > 25$ GeV.

where m_Z is the nominal mass of the Z boson, 91.1 GeV. The mass is scaled by only m_{jj} since the resolution of $\gamma\gamma$ is much better and adding the scaling of $m_{\gamma\gamma}$ by m_h has no noticeable effect on the resolution of the reconstructed A boson mass.

Figure 6.5 shows the effect of this mass scaling on the resolution. The improvement in mass resolution is especially pronounced at lower masses since the resolution is already quite good for high-p_T jets. With the mass scaling, the background shape becomes slightly narrower, which has a small impact on the low A mass search.

6.2.3.4 Dijet Invariant Mass

The position of the peak of the signal dijet invariant mass distribution varies as a function of m_A, as shown in Figure 6.6a. It is true even for the reconstructed jets that are matched to the true quarks from Z decays (see Figure 6.6b). The matching of jet to the true quark is

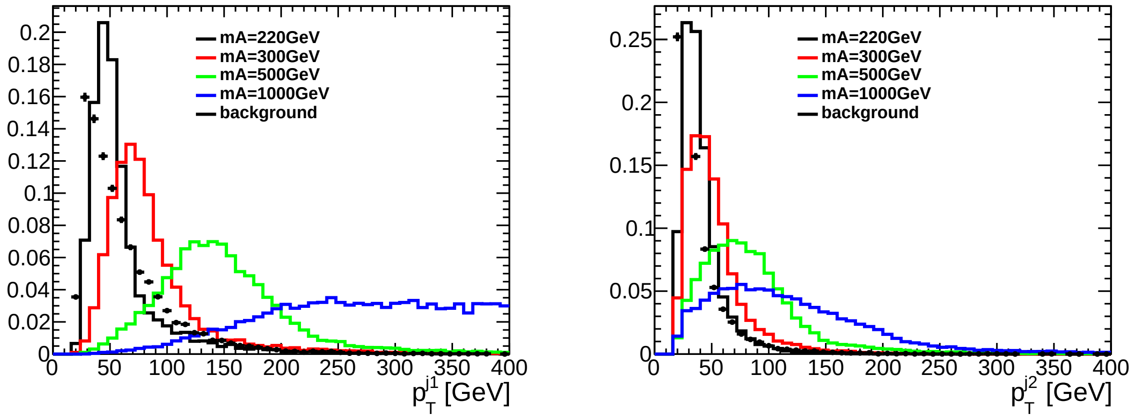
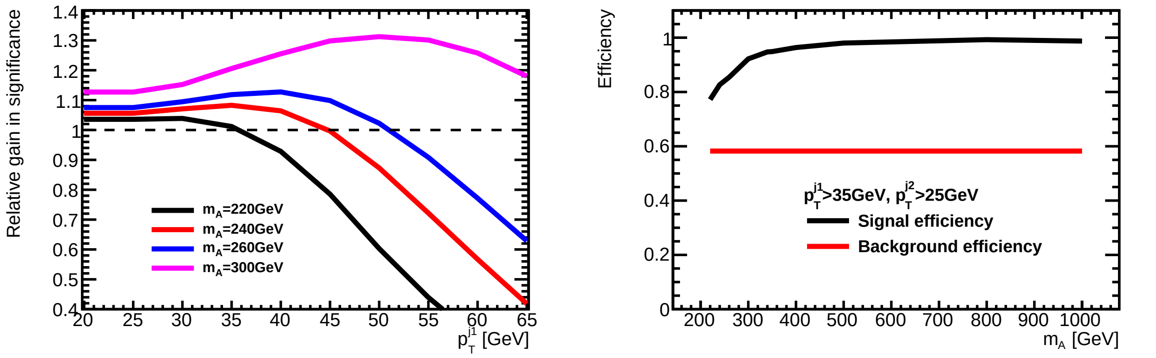


Figure 6.3: Distribution of the transverse momentum of the jet candidates for background events (dots) compared to signal events generated with $m_A = 220$ (black), 300 (red), 500 (green) and 1000 (blue) GeV. Left: leading jet candidate; right: subleading jet candidate.



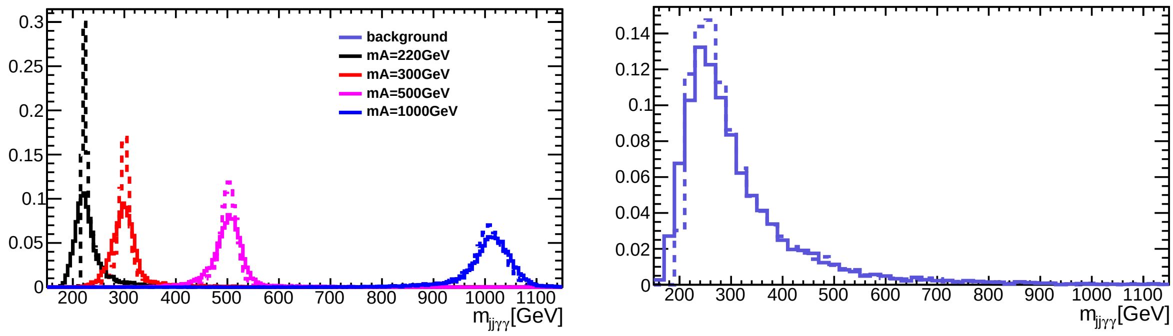
(a) (b)

Figure 6.4: (a) Gain in significance with respect to the pre-selection cuts (22 GeV on both jet candidates) as a function of p_T^{j1} cut for several hypotheses on the A boson mass: $m_A = 220$ (black), 240 (red), 260 (blue) and 300 (violet) GeV. The subleading jet is required to satisfy $p_T > 25$ GeV. (b) Signal and background efficiencies across m_A of the optimal cut found for $m_A = 260$ GeV, that is $p_T^{j1} > 35$ GeV and $p_T^{j2} > 25$ GeV. The efficiency is with respect to the pre-selection cuts as well as the relative photon p_T cut determined in previous section.

accomplished by requiring the selected jet to be within $\Delta R < 0.3$ of a true quark from Z decay. The invariant mass of dijet for low m_A are more spread out since the jets have lower p_T, and hence they have poorer resolution.

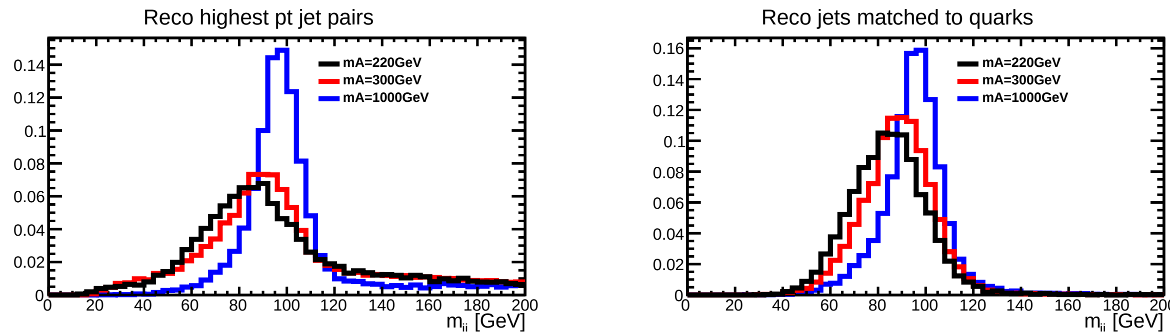
To define an optimum cut on the invariant mass of the dijet system, the peak position of the m_{jj} distribution is parameterised as a function of m_A. It is obtained from the Gaussian mean by fitting the core of the distribution with a Gaussian function. Figure 6.7 shows the dijet mass peak position as a function of m_A. A linear fit of the peak position as a function of m_A gives $m_{jj}^{peak} = 0.0097 \times m_A + 87.4$ GeV.

The window around this peak is then optimised for highest significance for each value of m_A. Figure 6.8a shows the width of m_{jj} window that gives the best significance. The black markers show the width value that gives the highest significance while the vertical error bands show the range of the m_{jj} width cut values that yield a significance within $\pm 1\sigma$ of the statistical uncertainty of the highest significance. A cut value on the width of m_{jj} of 46 GeV is chosen as



(a) (b)

Figure 6.5: Distribution of the reconstructed 4-body mass before scaling (in solid lines) and after scaling (in dashed lines) for (a) signal events generated with $m_A = 220$ (black), 300 (red), 500 (green) and 1000 (blue) GeV, and (b) background events.



(a) (b)

Figure 6.6: Reconstructed dijet mass distributions for $m_A = 220$ (black), 300 (red) and 1000 (blue) GeV choosing (a) the two highest-p_T jets, (b) the jets matching the quarks from the Z decay.

it is optimal for masses of the A boson between 220 and 350 GeV. With higher A masses, the dijet mass distributions are narrower but a common cut is chosen as there is little background in the high mass region. Figure 6.8b shows the gain in significance by optimising the dijet mass window cut compared to the 50 - 130 GeV pre-selection cut. The signal and background efficiencies with a 46 GeV window cut are plotted in Figure 6.8c.

6.2.3.5 Angle between the Jets

The angle in η between the jets is typically small for signal events, especially when the Z boson is boosted (which is true for high A mass). This is demonstrated in Figure 6.9a. $\Delta\eta_{jj} < 1.6$ is found to be the optimal cut for m_A between 220 and 350 GeV. This cut is more than 90% efficient on the signal while removing a substantial amount (about 20%) of background, as seen in Figure 6.9b.

Cuts on the angle in ϕ between the two jets are also tested, but they are difficult to parameterise and do not bring significant improvement. The distributions are shown in Figure 6.10.

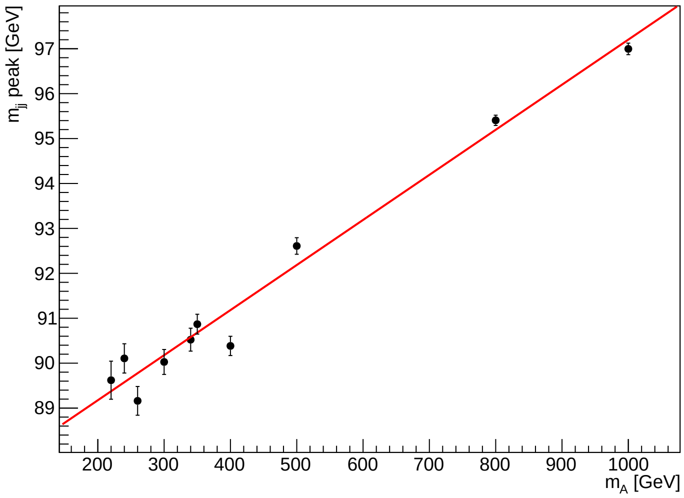
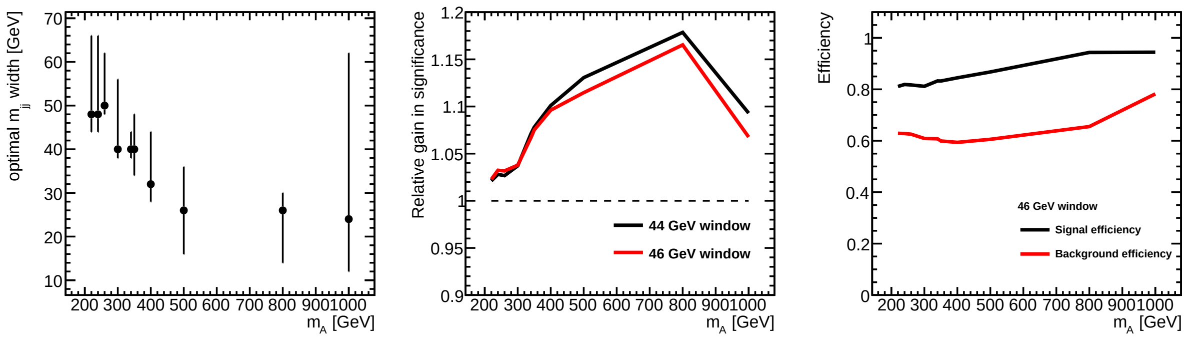


Figure 6.7: Peak position of the m_{jj} distributions as a function of m_A.



(a) (b) (c)

Figure 6.8: (a) Optimal width of dijet invariant mass window as a function of m_A. The points are the optimal cut values and the error bars show the range in which the significance is similar to the optimal cut significance within statistical uncertainties. (b) Gain in significance compared to the $50 < m_{jj} < 130$ GeV pre-selection cut as a function of m_A. (c) Signal and background efficiencies across m_A using 46 GeV dijet mass window cut.

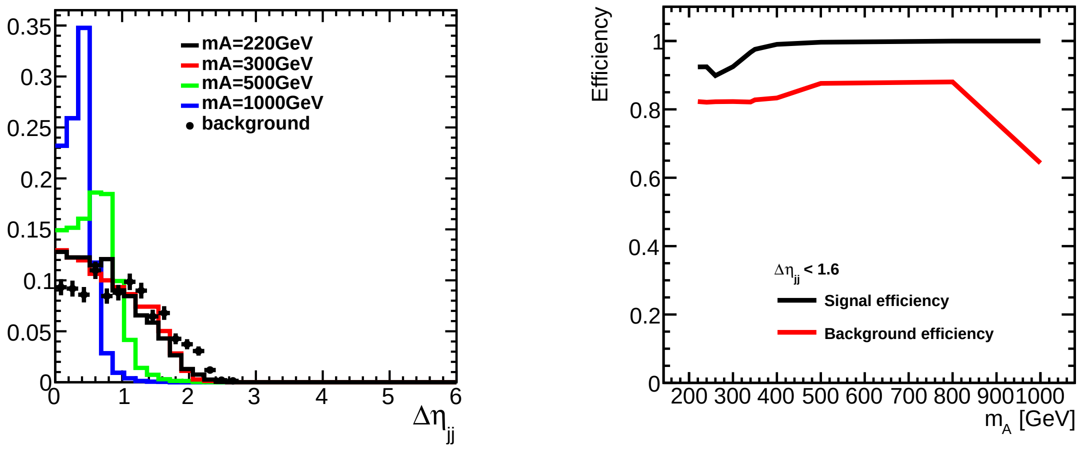
6.2.3.6 Dijet p_T

The transverse momentum distribution of the dijet system varies as a function of m_A. The mass distributions in background events resemble that of a low-mass A boson. By applying a cut on the p_T of the dijet system as a function of m_A^{rec}, the significance can be improved for $m_A > 260$ GeV.

Figure 6.11a shows the distributions of the transverse momentum of the dijet system that forms the Z boson for the non-resonant background and a few signal mass hypotheses. The cut value is chosen to be

$$p_T^{jj} < \text{Min}[163 \text{ GeV}, -2.18 \times 10^{-3} \cdot m_A^2 + 2.23 \cdot m_A - 407 \text{ GeV}] \tag{6.3}$$

from the results of optimisation, as shown in Figure 6.11b. For reconstructed A masses above 500 GeV, a cut of 163 GeV is chosen instead due to the very low background.



(a) (b)

Figure 6.9: (a) Distribution of the difference in η between the two jets ($\Delta\eta_{jj}$) for background events (dots) compared to signal events generated with $m_A = 220$ (black), 300 (red), 500 (green) and 1000 (blue) GeV. (b) Signal and background efficiencies across m_A with a cut on $\Delta\eta_{jj} < 1.6$.

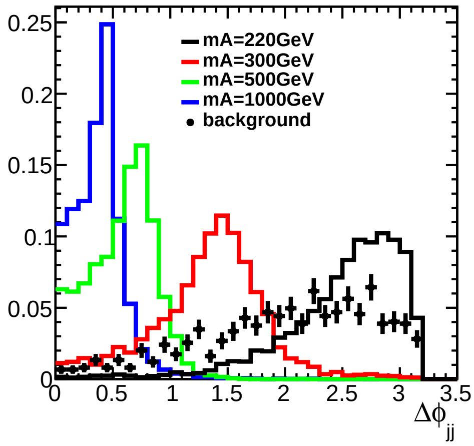
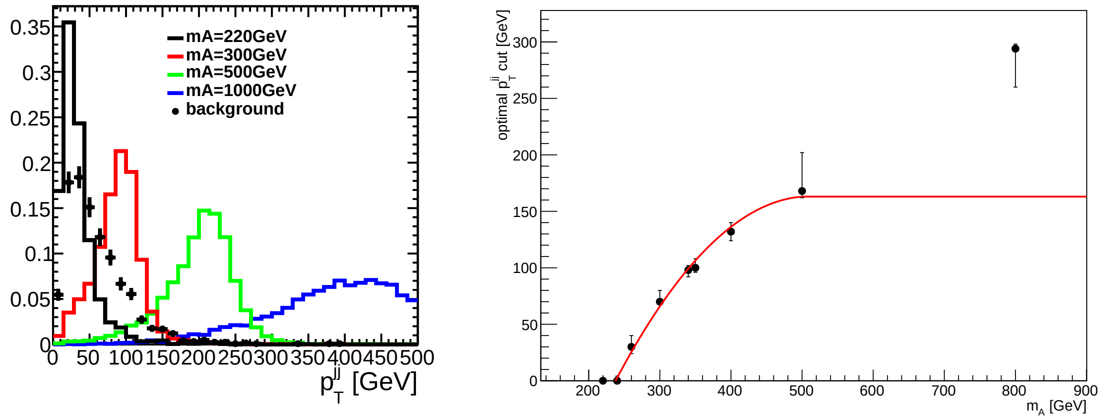


Figure 6.10: Distribution of the difference in ϕ between the two jets ($\Delta\phi_{jj}$) for background events (dots) compared to signal events generated with $m_A = 220$ (black), 300 (red), 500 (green) and 1000 (blue) GeV.

6.2.4 Summary

The event selection has been optimised making use of the different kinematics between the background and the signal events. The final event selection acceptance in the simulated signal samples ranges from 8 to 26% and is shown in Figure 6.12. The selection cuts are summarised below:

- $p_T^{\gamma 1}/m_{\gamma\gamma} > 0.5$ and $p_T^{\gamma 2}/m_{\gamma\gamma} > 0.25$.
- $p_T^{j1} > 35$ GeV and $p_T^{j2} > 25$ GeV.
- m_{jj} within a 46 GeV window centred around $m_{jj}^{peak} = 0.0097 \times m_A + 87.4$ GeV.
- $\Delta\eta_{jj} < 1.6$.
- $p_T^{jj} < \mathrm{Min}[163 \text{ GeV}, -2.18 \times 10^{-3} \cdot m_A^2 + 2.23 \cdot m_A - 407 \text{ GeV}]$.



(a) (b)

Figure 6.11: (a) Distribution of the transverse momentum of the dijet system (p_T^{jj}) for background events (dots) compared to signal events generated with $m_A = 220$ (black), 300 (red), 500 (green) and 1000 (blue) GeV. (b) Optimal p_T^{jj} cut as a function of m_A. The points are the optimal cut values and the error bars show the range in which the significance is similar to the optimal cut significance within statistical uncertainties. The p_T cuts are fitted with a quadratic function for masses up to 500 GeV.

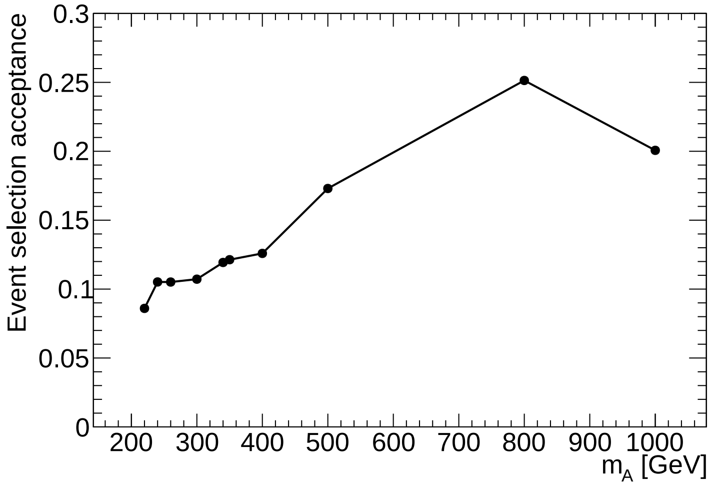


Figure 6.12: Final acceptance (ratio of signal events passing the full selection described previously to the total number of generated events) as a function of generated m_A.

6.3 Statistical Analysis

6.3.1 Fitting Strategy

Figure 6.13 shows the scaled four-body invariant mass m_A^{rec} as defined in Eq. 6.2, for events in data passing the selection described in the previous section, but the cuts on the $m_{\gamma\gamma}$ and m_A^{rec} invariant masses. However, in order to keep the analysis blinded, only the events in the $m_{\gamma\gamma}$ sidebands (see Section 6.1.2) are shown. The peak at around 260 GeV makes it difficult to search for an excess in the $m_A^{\text{rec}} < 400$ GeV range. Therefore, the search is split into two regions: a low-mass region ($220 \leq m_A < 400$ GeV) and a high-mass region ($400 \leq m_A \leq 1000$ GeV).

At low mass, $m_{\gamma\gamma}$ is used to discriminate between signal and background. The search is

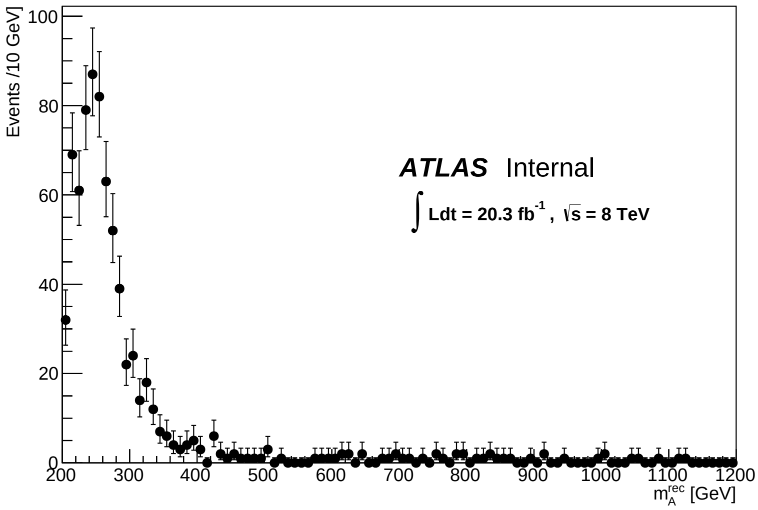


Figure 6.13: Rescaled four-body invariant mass m_A^{rec} for events passing the selection described in Section 6.2, except the cuts on $m_{\gamma\gamma}$ and on m_A^{rec}. The cut on $m_{\gamma\gamma}$ is reversed to select only events in the sidebands: $110 < m_{\gamma\gamma} < m_h - 2\sigma_{\gamma\gamma}$ or $m_h + 2\sigma_{\gamma\gamma} < m_{\gamma\gamma} < 160$ GeV (control region).

performed for six mass points: $m_A = 220, 240, 260, 300, 340$ and 350 GeV, corresponding to the signal Monte Carlo samples (see Section 6.1.1). For each mass point the $m_{\gamma\gamma}$ spectrum is built for events passing the selection, including the cut on m_A^{rec} around m_A as described in Section 6.2.3.3.

For the high mass analysis, the discriminant variable is m_A^{rec}. The signal and control region samples are fitted simultaneously. The signal region sample is composed of events passing the selection, including the $|m_{\gamma\gamma} - m_h| < 2\sigma_{\gamma\gamma}$ cut (Section 6.1.2), whereas the control region sample is built by reversing the cut on $m_{\gamma\gamma}$ to select events in the sidebands ($110 < m_{\gamma\gamma} < m_h - 2\sigma_{\gamma\gamma}$ or $m_h + 2\sigma_{\gamma\gamma} < m_{\gamma\gamma} < 160$ GeV). The control region sample is used to constrain better the background shape in m_A^{rec} by fitting the spectrum in the signal and control regions simultaneously.

6.3.2 Signal and Background Modelling in the Low-mass Analysis

The shape of the $m_{\gamma\gamma}$ spectrum for signal events is parameterised by the sum of a Crystal Ball function [89–91] and a Gaussian as in the SM $h \to \gamma\gamma$ analysis. The Crystal Ball function is defined as

$$N \cdot \begin{cases} \exp\left(-t^1/2\right) , & \text{for } t > -\alpha_{\mathrm{CB}} \\ (n_{\mathrm{CB}}/|\alpha_{\mathrm{CB}}|)^{n_{\mathrm{CB}}} \cdot \exp\left(-|\alpha_{\mathrm{CB}}|^2/2\right) \cdot (n_{\mathrm{CB}}/|\alpha_{\mathrm{CB}}| - |\alpha_{\mathrm{CB}}| - t)^{-n_{\mathrm{CB}}} , & \text{otherwise} \end{cases} \tag{6.4}$$

where N is a normalisation parameter and $t = (m_{\gamma\gamma} - \mu_{\mathrm{CB}})/\sigma_{\mathrm{CB}}$. μ_{CB} and σ_{CB} are the mean and the width of the Gaussian core of the Crystal Ball function, and n_{CB} and α_{CB} parameterise the non-Gaussian tail. The mean and the width of the additional Gaussian function are denoted μ_{G} and σ_{G} in the following. The parameter values are estimated from the fit to the Monte Carlo signal samples (Section 6.1.1). As observed in the SM $h \to \gamma\gamma$ analysis, the correlations between the parameters listed above are large. Therefore, n_{CB} is fixed to 10 and μ_{G} to μ_{CB}. Table 6.2 shows the parameter values fitted on the $m_A = 300$ GeV signal sample. It was checked that the $m_{\gamma\gamma}$ spectrum shape does not vary significantly with m_A. Therefore, this set of parameters is used for any mass point. The result of the fit is shown in Figure 6.14.

parameter	value
$\mu_{\mathrm{CB}} = \mu_{\mathrm{G}}$	124.89 ± 0.03
σ_{CB}	1.48 ± 0.03
α_{CB}	1.7 ± 0.1
n_{CB}	10 (fixed)
σ_{G}	3.9 ± 0.6
f_{G}	0.97 ± 0.02

Table 6.2: Values of the parameters of the Crystal Ball and Gaussian functions extracted from the diphoton spectrum in the Monte Carlo signal sample generated at $m_A = 300$ GeV and used in the low-mass analysis. f_{G} is the fraction of events described by the Gaussian. The meaning of other parameters is given in the text.

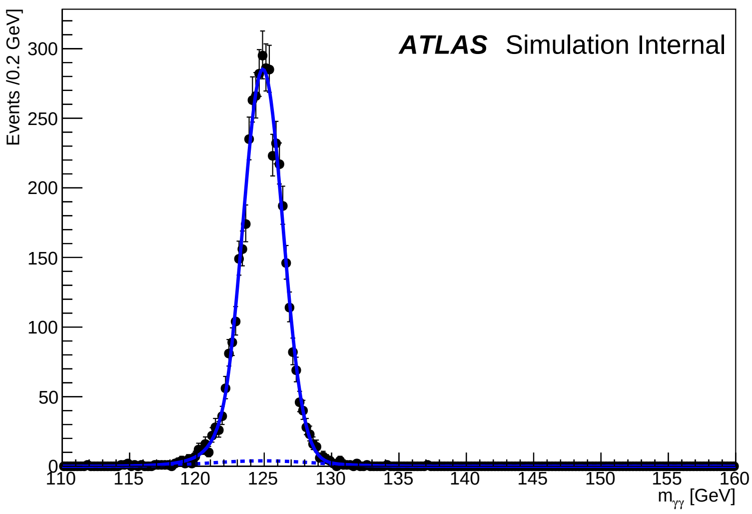


Figure 6.14: Diphoton invariant mass spectrum from signal Monte Carlo events generated for $m_A = 300$ GeV. The spectrum is fitted with a Crystal Ball plus a Gaussian (dotted line).

The continuum background is modelled by an exponential. It is known from the SM $h \to \gamma\gamma$ analysis that an imperfect description of the background by the chosen function can lead to a spurious estimation of signal. This effect has not been studied in the present analysis given that the goal is to roughly estimate the sensitivity reach of this channel. A detailed measurement of the exclusion limit would certainly require a careful study of spurious signal effects and might lead to choosing a better suited function. Figure 6.15 shows the fitted spectrum for each mass point m_A. The signal region is blinded.

Direct production of the SM Higgs h is a background in the low-mass analysis. The expected number of events is estimated from the production cross-sections and the $h \to \gamma\gamma$ decay branching fraction from Ref. [15].The selection efficiencies are computed from the Monte Carlo samples listed in Section 6.1.2.2. The shape is assumed to be the same as for the signal. This assumption has been checked on Monte Carlo. Table 6.3 gives the selection efficiencies and expected number of events for each production channel and each mass point m_A as the selection criteria depend on m_A.

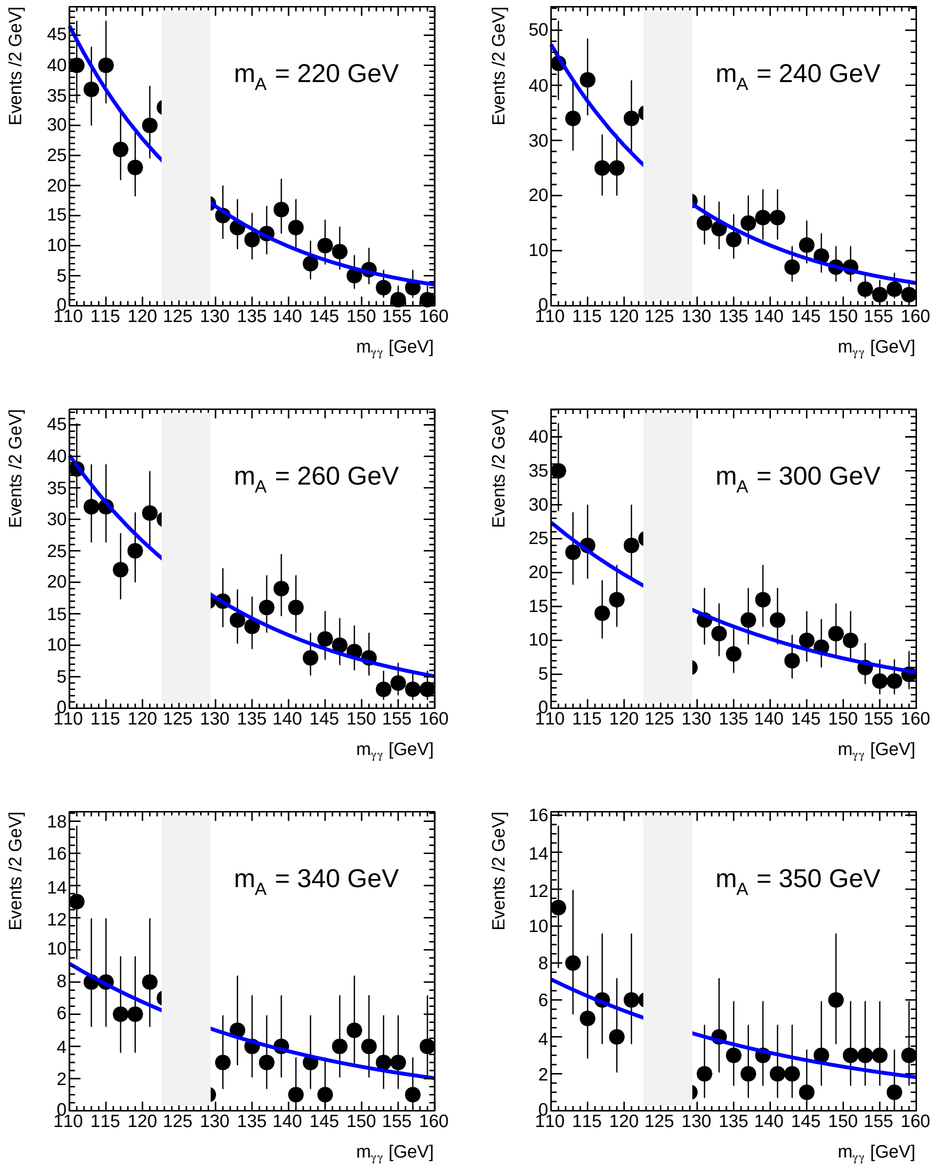


Figure 6.15: Diphoton invariant mass spectra from data events selected for each mass point being tested in the low-mass analysis. The spectra are fitted with an exponential. The signal region is blinded.

6.3.3 Signal and Background Modelling in the High-mass Analysis

The high-mass search is performed for four mass points: m_A = 400, 500, 800 and 1000 GeV. The shape of the signal m_A^{rec} spectrum is parameterised by the sum of a Crystal Ball and a Gaussian, as for the $m_{\gamma\gamma}$ spectrum in the low-mass analysis. The values of the parameters are extracted from the Monte Carlo samples, corresponding to the four mass points mentioned above. The values are given in Table 6.4 and the fitted distributions are shown in Figure 6.16. As in the low-mass search, the number of free parameters is reduced: n_{CB} is fixed to 5, α_{CB} to 1.5 and μ_{G} to μ_{CB}.

The background is modelled by an exponential. In order to increase the statistical accuracy of the parameter estimation, m_A^{rec} sidebands in the signal region are fitted simultaneously with the m_A^{rec} spectrum from events in the control region (Figure 6.13). Because of the steep decrease of the spectrum below 400 GeV changing to an almost flat distribution at higher masses, it is

prod.	220 GeV		240 GeV		250 GeV		260 GeV		300 GeV		350 GeV	
mode	ϵ [%]	evts	ϵ [%]	evts	ϵ [%]	evts	ϵ [%]	evts	ϵ [%]	evts	ϵ [%]	evts
ggF	0.17	1.56	0.19	1.71	0.21	1.90	0.20	1.78	0.11	0.96	0.08	0.78
VBF	0.11	0.08	0.12	0.11	0.13	0.10	0.13	0.10	0.09	0.06	0.08	0.06
WH	2.3	0.74	2.5	0.81	2.7	0.89	2.0	0.90	1.7	0.65	1.0	0.56
ZH	2.8	0.53	3.0	0.58	3.2	0.61	3.1	0.59	2.1	0.41	1.9	0.36
$t\bar{t}H$	0.82	0.05	0.93	0.06	1.2	0.07	1.5	0.09	1.3	0.08	1.2	0.07
total		2.96		3.24		3.57		3.45		2.17		1.83

Table 6.3: Selection efficiencies and expected numbers of events from direct h production, for each SM production mode. Since the selection criteria depends on m_A, the numbers are given for each mass point. The expected numbers of events correspond to an integrated luminosity of 20.3 fb^{-1}.

parameter	$m_A = 400$ GeV	$m_A = 500$ GeV	$m_A = 800$ GeV	$m_A = 1000$ GeV
$\mu_{\mathrm{CB}} = \mu_{\mathrm{G}}$	402.0 ± 0.1	504.4 ± 0.1	807.3 ± 0.2	1008.6 ± 0.3
σ_{CB}	10.8 ± 0.1	13.4 ± 0.1	21.7 ± 0.2	28.9 ± 0.2
α_{CB}	1.5 (fixed)			
n_{CB}	5. (fixed)			
σ_{G}	$76. \pm 8.$	$76. \pm 6.$	$105. \pm 7.$	$230. \pm 50.$
f_{G}	0.918 ± 0.006	0.938 ± 0.004	0.953 ± 0.004	0.961 ± 0.003

Table 6.4: Parameter values of the Crystal Ball and Gaussian functions extracted from the m_A^{rec} spectrum in the Monte Carlo signal samples for the high-mass analysis. f_{G} is the fraction of events described by the Gaussian. The meaning of other parameters is given in the text.

difficult to find one function to describe the full range. Instead, the fit is performed on a window around each mass point, defined as $[m_A^{\mathrm{rec}} - 5 \cdot 10^{-4} \times (m_A^{\mathrm{rec}})^2, m_A^{\mathrm{rec}} + 5 \cdot 10^{-4} \times (m_A^{\mathrm{rec}})^2]$. This choice maximises the number of events entering in the fit for each value of m_A, while restricting the window to a range where the spectrum can reasonably be described by an exponential. Figure 6.17 shows the result of the fit in four windows of the control region m_A^{rec} spectrum.

6.3.4 Limit Extraction

The number of signal and background events are extracted using an extended unbinned likelihood fit. For the low-mass analysis, the fit is performed on the $m_{\gamma\gamma}$ spectra built for the six low-mass m_A points ($m_A = 220, 240, 260, 300, 340$ and 350 GeV). In the high-mass region, a simultaneous fit on the signal and the control region m_A^{rec} spectra is performed in four windows around $m_A = 400, 500, 800$ and 1000 GeV.

The only systematic uncertainty included in this study is the uncertainty on the integrated luminosity of the dataset. The absolute luminosity scale is derived from beam-separation scans performed in November 2012. The uncertainty on the integrated luminosity is ±1.9% [92].

The expected limit is extracted based on the test statistics q_μ defined as[3]

$$q_\mu = \begin{cases} -2\ln\frac{L\left(\mu,\hat{\hat{\theta}}\right)}{L(\hat{\mu},\hat{\theta})}\,, & \text{if } \hat{\mu} \leq \mu, \\ 0\,, & \text{otherwise} \end{cases} \tag{6.5}$$

[3]The test statistics used here is different to the one defined in Section 5.4. In the asymptotic approximation, the two test statistics are equivalent [78].

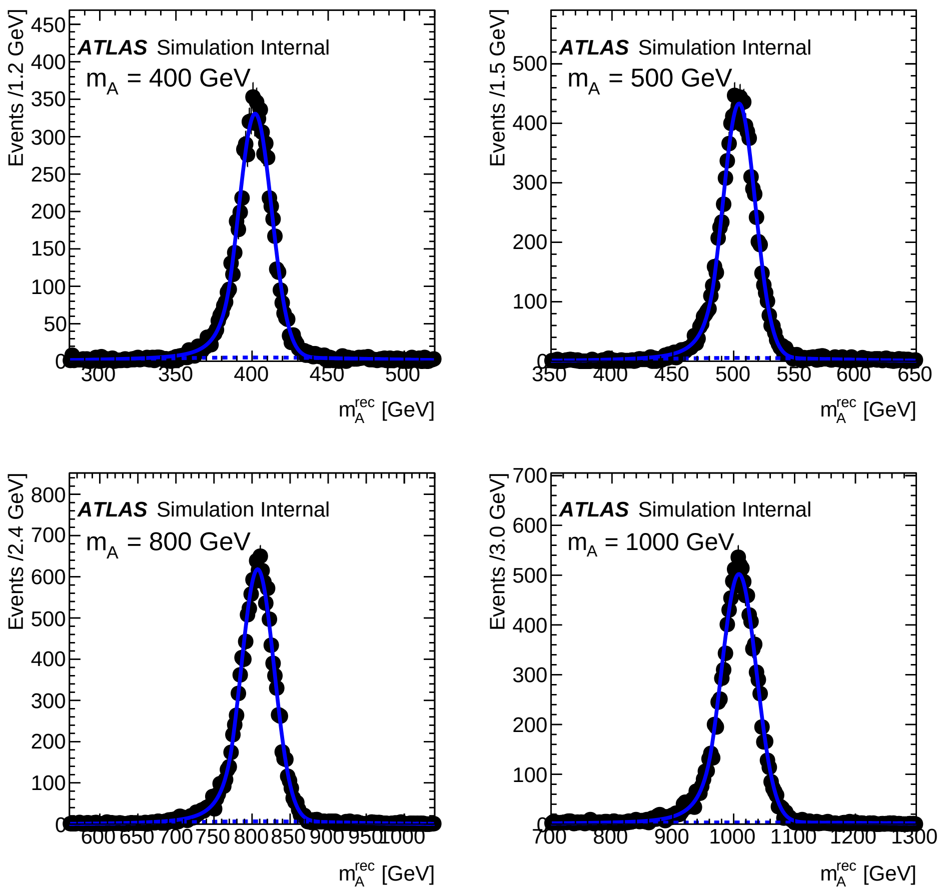


Figure 6.16: Four-body scaled invariant mass m_A^{rec} from signal Monte Carlo events generated for $m_A = 400$, 500, 800 and 1000 GeV. The spectrum is fitted with a Crystal Ball plus a Gaussian (dotted line).

where the signal strength μ is defined as the ratio of a given value of the product of the production cross-section times the $A \to Zh$ decay branching ratio $\sigma(gg \to A) \times \mathrm{BR}(A \to Zh)$ over the value predicted in the 2HDM model (assuming Type-I 2HDM with $\sin(\beta - \alpha) = 0.9998$ and $\tan\beta = 1$). θ includes the slope of the exponential background as well as systematic effects.

Upper limits at 95% confidence level are estimated using the CL_s technique [79]. Only expected limits are derived in this note. They are computed using Asimov datasets [78].

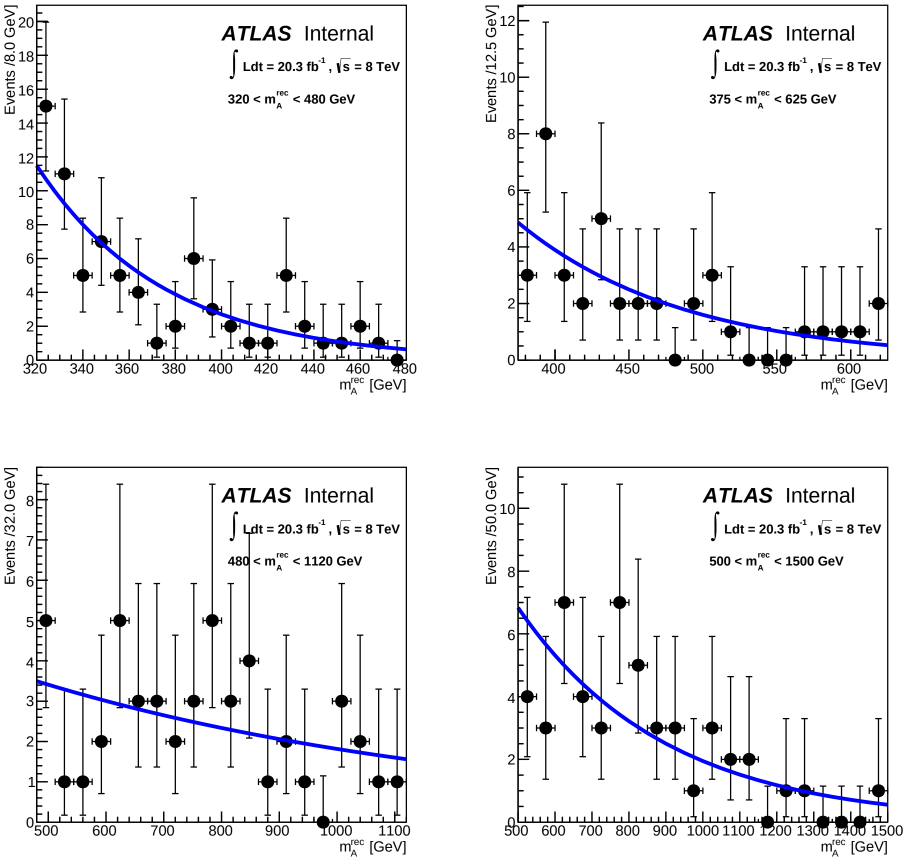


Figure 6.17: Four-body scaled invariant mass m_A^{rec} in the control region within a window around the tested mass hypotheses of $m_A = 400$, 500, 800 and 1000 GeV. The spectrum is fitted with an exponential.

6.4 Results

Figure 6.18 shows the 95% CL expected limits on $\sigma(gg \to A) \times \text{BR}(A \to Zh)$ for the ten mass points already listed above: m_A =220, 240, 260, 300, 340, 350, 400, 500, 800 and 1000 GeV. The values are reported in Table 6.5. We assume here that $\text{BR}(h \to \gamma\gamma) = 2.28 \times 10^{-3}$. For the $Z \to jj$ decay, the experimental value of the branching fraction is used: $\text{BR}(Z \to jj) = (69.911 \pm 0.056\%)$ [93]. In Figure 6.19, the expected limit obtained in other search channels in ATLAS using 2012 data is also shown overlaid. The expected limit in $jj\gamma\gamma$ channel is competitive with $ll\tau\tau$ especially at higher mass. However, the sensitivity of $llbb$ analysis is much better.

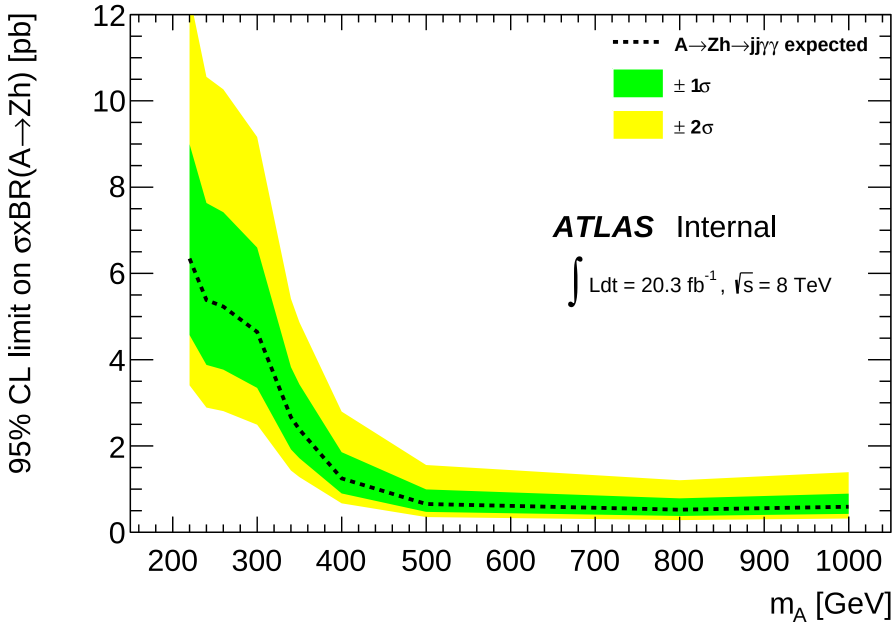


Figure 6.18: Expected limits on $\sigma(gg \to A) \times \text{BR}(A \to Zh)$. The limits are extracted for ten mass points: $m_A = 220\,, 240\,, 260\,, 300\,, 340\,, 350\,, 400\,, 500\,, 800$ and 1000 GeV. The black dashed line represents a simple linear interpolation between two neighbouring points. The green and yellow bands indicate the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties on the limits.

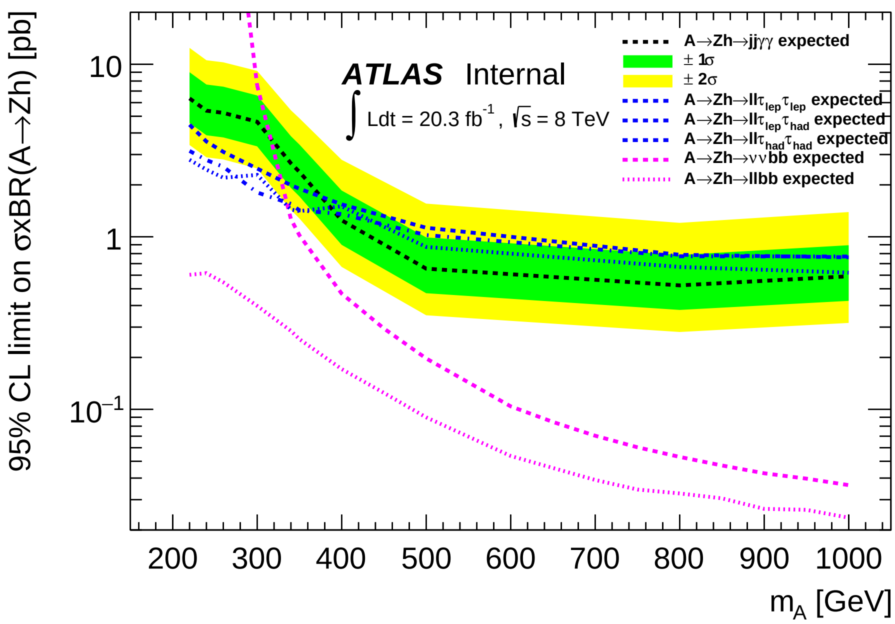


Figure 6.19: Expected limits on $\sigma(gg \to A) \times \text{BR}(A \to Zh)$. The limits are extracted for ten mass points: $m_A = 220\,, 240\,, 260\,, 300\,, 340\,, 350\,, 400\,, 500\,, 800$ and 1000 GeV. The black dashed line represents a simple linear interpolation between two neighbouring points. The green and yellow bands indicate the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties on the limits. Expected limit obtained in other search channels (blue for $ll\tau\tau$ and purple for $ll/\nu\nu bb$) are also shown for comparison.

m_A [GeV]	limit [pb]	$\pm 1\sigma$ [pb]	$\pm 2\sigma$ [pb]
220	6.3	4.6-9.0	3.4-12.4
240	5.4	3.9-7.6	2.9-10.6
260	5.2	3.8-7.4	2.8-10.3
300	4.6	3.3-6.6	2.5-9.2
340	2.7	1.9-3.8	1.4-5.4
350	2.4	1.7-3.4	1.3-4.9
400	1.2	0.9-1.8	0.7-2.8
500	0.6	0.5-1.0	0.3-1.6
800	0.5	0.4-0.8	0.3-1.2
1000	0.6	0.4-0.9	0.3-1.4

Table 6.5: Expected limits on $\sigma(gg \to A) \times \mathrm{BR}(A \to Zh)$ estimated for ten values of m_A. The last two columns give the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty ranges.

6.5 Conclusion

The sensitivity study of a search for a pseudoscalar A in the decay to Zh, with a $jj\gamma\gamma$ final state is presented. The study is performed on 20.3 fb^{-1} of pp collision data at $\sqrt{s} = 8$ TeV as background estimation. The analysis is blinded and the expected sensitivity is given by the expected 95% CL limits on the production cross-section times branching ratio of $A \to Zh$. Comparing the expected limits to those published by ATLAS in other final states, the analysis is found to be competitive with the $ll\tau\tau$ final state but less sensitive to searches with bb final states.

Chapter 7

Large-Width Resonance Shape Studies for High Mass Diphoton Resonances Search

In this chapter, the choice of the signal model to be used in the high mass diphoton analysis as well as the parameterisation of the large-width theoretical line shape is presented. In Run 1, the search was limited to a narrow-width signal where the natural width is negligible compared to the detector resolution. Since certain models predict resonances with widths larger than the mass resolution, an improvement made in the analysis for Run 2 is to extend the search to include large widths. In the case of a narrow-width resonance, the signal shape is dominated by detector resolution. On the contrary, the line shape of a large-width signal is dependent on the model assumption being made, therefore requires more care.

I have studied the line shape of a high mass resonance in the Monte Carlo simulation under several model assumptions. I then developed the parameterisation of the theoretical line shape based on the cross-section computation in the Monte Carlo generators. This gives an analytical line shape for any reasonable resonance mass and width that is then used in the analysis. The parameterisation of the theoretical line shape is checked and validated with simulated samples. New particles decaying to two photons can have spin 0 or spin 2. The resonance shape is studied in the Higgs-like model (spin-0), a scalar singlet model (spin-0) and a Randall-Sundrum graviton model (spin-2). Comparisons of the different approaches to large-width signal generation in various Monte Carlo generators are also made. In these studies, interference with the continuum background is neglected.

Section 7.1 outlines the determination of the theoretical line shape of a resonance. The generation of large-width heavy Higgs-like resonances using the PowHeg generator and its drawbacks are explained in Section 7.2. Section 7.3 then describes the alternative scalar singlet model. The impact of the interference between the spin-0 resonance and the continuum background on the resonance shape is next discussed in Section 7.4. Finally, the parameterisation of the line shape of a spin-2 resonance using the Randall-Sundrum graviton model is presented in Section 7.5.

7.1 Theoretical Line Shape of a Resonance

The line shape of a resonance X is given by the invariant mass distribution of its decay products, i.e. the diphoton invariant mass in this case. This in turn is generated by the differential cross-section $d\sigma/dm_{\gamma\gamma}$. In the analysis, only the $gg \to X \to \gamma\gamma$ process is considered for the spin-0 resonance search, while both $gg \to X \to \gamma\gamma$ and $q\bar{q} \to X \to \gamma\gamma$ processes are taken into account for the spin-2 search. Both processes are collectively denoted $ii \to X \to \gamma\gamma$ where ii refers to

gg and $q\bar{q}$.

7.1.1 Differential Cross-section

The differential cross-section for a two-to-two process in the centre-of-mass frame at a particular energy $\sqrt{\hat{s}}$ is

$$\frac{d\sigma}{d\Omega_{\text{CM}}} = \frac{1}{64\pi^2 \hat{s}} \frac{|\boldsymbol{k}_1'|}{|\boldsymbol{k}_1|} |\mathcal{A}|^2, \tag{7.1}$$

$$|\boldsymbol{k}_1'| = \frac{1}{2\sqrt{\hat{s}}} \sqrt{\hat{s}^2 - 2(m_{1'}^2 + m_{2'})\hat{s} + (m_{1'}^2 - m_{2'})^2}, \tag{7.2}$$

$$|\boldsymbol{k}_1| = \frac{1}{2\sqrt{\hat{s}}} \sqrt{\hat{s}^2 - 2(m_1^2 + m_2)\hat{s} + (m_1^2 - m_2)^2}, \tag{7.3}$$

where $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ are the four-momenta of the two incoming particles, $\boldsymbol{k}_1'$ and $\boldsymbol{k}_2'$ are the four-momenta of the two outgoing particles, $d\Omega_{\text{CM}}$ is the differential solid angle in the centre-of-mass frame and $\mathcal{A}$ is the matrix element.

For $gg \to X \to \gamma\gamma$, both incoming and outgoing particles are massless, so $\frac{|\boldsymbol{k}_1'|}{|\boldsymbol{k}_1|} = 1$. In the $q\bar{q} \to X \to \gamma\gamma$ process, although the quarks are not massless, in the high energy limit $\hat{s}$ is much greater than the mass of a quark and $\frac{|\boldsymbol{k}_1'|}{|\boldsymbol{k}_1|} \approx 1$.

The differential cross-section can then be simplified as

$$\frac{d\sigma}{d\Omega_{\text{CM}}} = \frac{1}{64\pi^2 \hat{s}} |\mathcal{A}|^2. \tag{7.4}$$

The above differential cross-section is for a single event at a particular energy $\sqrt{\hat{s}}$. In the context of proton-proton collisions at the LHC, this is equal to the squared centre-of-mass energy in a given collision. In each collision, each value of $\sqrt{\hat{s}}$ is weighted by the parton luminosity $\mathcal{L}$.

The parton luminosity is defined in terms of the parton distribution functions $f(x, \hat{s})$, given below for gg and $q\bar{q}$:

$$\mathcal{L}_{gg}(\hat{s}) = \frac{1}{s} \int_\tau^1 \frac{dx}{x} f_g(x, \hat{s}) f_g(\tau/x, \hat{s}), \tag{7.5}$$

$$\mathcal{L}_{q\bar{q}}(\hat{s}) = \frac{1}{s} \int_\tau^1 \frac{dx}{x} \sum_{q=d,u,s,c,b} \left(f_q(x, \hat{s}) f_{\bar{q}}(\tau/x, \hat{s}) + f_{\bar{q}}(x, \hat{s}) f_q(\tau/x, \hat{s}) \right), \tag{7.6}$$

where x is the momentum fraction carried by the parton, $\sqrt{s}$ is the centre-of-mass energy of the hadronic collision and $\tau = \hat{s}/s$.

The differential cross-section thus becomes

$$\frac{d\sigma}{d\Omega_{\text{CM}}} = \int d\hat{s} \mathcal{L}(\hat{s}) \frac{1}{64\pi^2 \hat{s}} |\mathcal{A}|^2, \tag{7.7}$$

$$d\sigma = \int d\hat{s} \mathcal{L}(\hat{s}) \frac{1}{64\pi^2 \hat{s}} |\mathcal{A}|^2 d\Omega_{\text{CM}}. \tag{7.8}$$

From this point, the terms that do not depend on $\hat{s}$ shall be dropped in order to simplify the differential cross-section expression. The signal amplitude $\mathcal{A}$ of the $ii \to X \to \gamma\gamma$ resonant process can be expressed as

$$\mathcal{A} = a_{ii\to X}(\hat{s}) \Delta(\hat{s}, m_X) a_{X\to\gamma\gamma}(\hat{s}), \tag{7.9}$$

where m_X is the mass of the resonance and $\Delta(\hat{s}, m_X)$ is the propagator.

Finally, the differential cross-section is written as

$$d\sigma \propto \int d\hat{s} \mathcal{L}(\hat{s}) \frac{1}{\hat{s}} |a_{ii\to X}(\hat{s}) \Delta(\hat{s}, m_X) a_{X\to\gamma\gamma}(\hat{s})|^2. \tag{7.10}$$

7.1.2 Resummed Propagator

The leading order propagator

$$\Delta(\hat{s}, m_X) = 1/(\hat{s} - m_X^2) \tag{7.11}$$

gives divergent results for the cross-section at $\hat{s} = m_X$. This is clearly unphysical so higher order effects in perturbation theory must be included.

The correction to the propagator from the sum of all diagrams that are one-particle irreducible (1PI), i.e. diagrams which are still connected after an internal line is cut, as shown in Figure 7.1, is computed.

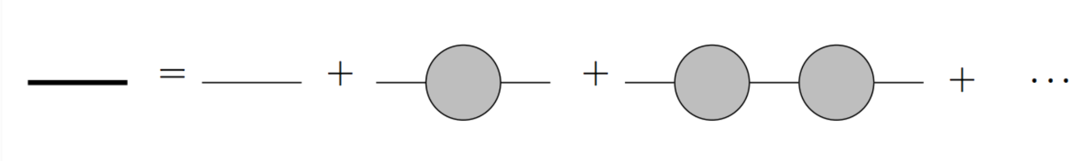


Figure 7.1: Sum of one-particle irreducible (1PI) diagrams with two external lines.

This corresponds to the following geometric series:

$$\begin{aligned}
&\frac{1}{\hat{s} - m^2} + \frac{1}{\hat{s} - m^2}(-\Sigma^2(\hat{s}))\frac{1}{\hat{s} - m^2} + \frac{1}{\hat{s} - m^2}(-\Sigma^2(\hat{s}))\frac{1}{\hat{s} - m^2}(-\Sigma^2(\hat{s}))\frac{1}{\hat{s} - m^2} + ... \\
&= \frac{1}{\hat{s} - m^2 + \Sigma^2(\hat{s})} \\
&= \frac{1}{\hat{s} - m^2 + Re(\Sigma^2(\hat{s})) + iIm(\Sigma^2(\hat{s}))}
\end{aligned} \tag{7.12}$$

where $\Sigma(\hat{s})$ is the self-energy.

The resummation of self energies includes only a subset of higher order diagrams, but neglects other contributions of the same order. This leads to inconsistencies such as violations of gauge invariance, which holds order by order in perturbation theory.

In the resummed propagator above, the pole mass (the real part of the pole of the propagator) is shifted to $m^2 - Re(\Sigma^2)$. In the *on-shell scheme*, the width is defined as

$$\Gamma_X = \frac{1}{m_X} Im(\Sigma^2). \tag{7.13}$$

The optical theorem [94] relates the imaginary part of the forward scattering amplitude to the amplitudes for all possible final states f:

$$Im\mathcal{A}(X \to X) = \frac{1}{2}\sum_f \int d\Pi_f |\mathcal{A}(X \to f)|^2, \tag{7.14}$$

where the integration is performed over Π_f, the phase space of f.

The left-hand side of the equation can be replaced by the imaginary part of the self-energy, therefore

$$Im(\Sigma^2) = \frac{1}{2}\sum_f \int d\Pi_f |\mathcal{A}(X \to f)|^2, \tag{7.15}$$

and the decay width is

$$\Gamma_X = \frac{1}{2m_X}\sum_f \int d\Pi_f |\mathcal{A}(X \to f)|^2. \tag{7.16}$$

Finally, the resummed propagator in the on-shell scheme reads

$$\Delta(\hat{s}, m_X) = \frac{i}{\hat{s} - m_X^2 + i m_X \Gamma_X}. \tag{7.17}$$

The familiar relativistic Breit-Wigner (BW) function can be obtained from the propagator squared:

$$|\Delta(\hat{s}, m_X)|^2 = \frac{1}{(\hat{s} - m_X^2)^2 + m_X^2 \Gamma_X^2}. \tag{7.18}$$

7.1.3 On-shell Calculations

In the approach used by many generators, the computation of cross-sections is done on-shell ($\sqrt{\hat{s}} = m_X$) and in the narrow-width approximation ($\Gamma_X = 0$). The constraint of $\hat{s}$ equal to m_X by is imposed by introducing a term $\delta(\hat{s} - m_X^2)$:

$$d\sigma(\hat{s} = m_X) = \int d\hat{s}\, d\sigma(\hat{s})\, \delta(\hat{s} - m_X^2). \tag{7.19}$$

Replacing $d\sigma(\hat{s})$ from Eq. 7.10, the differential cross-section becomes

$$d\sigma(\hat{s} = m_X) \propto \int d\hat{s} \mathcal{L}(\hat{s}) \frac{1}{\hat{s}} |a_{ii\to X}(\hat{s}) a_{X\to\gamma\gamma}(\hat{s})|^2 \delta(\hat{s} - m_X^2). \tag{7.20}$$

In this approximation, the propagator is replaced by a δ-function and the integration over the virtuality $\hat{s}$ vanishes.

The effect of a non-zero width on the line shape is implemented *a posteriori* by replacing the δ function in the above equation by a Breit-Wigner function. The on-shell approximation effectively factorises the propagator from the equation, enabling the signal cross-section to be described as a separate product of production, BW and decay terms. A BW function is a good approximation of the resonance shape only if the particle's decay width is roughly constant over the width of the resonance. This applies well to narrow resonances but in general, the decay width varies as a function of $\sqrt{\hat{s}}$. Even if the resonance has roughly constant decay branching ratios, the width can vary according to the evolution of the phase space (see Eq. 7.16). In addition, the opening of new decay channels also introduce an additional dependence of the width on $\sqrt{\hat{s}}$.

Only full off-shell calculations of $ii \to X \to \gamma\gamma$ can provide a correct description of a wide resonance. However, since such calculations are not always available, modified forms of propagators have been proposed to be used in on-shell calculation to mimic the full calculations.

The first is the relativistic BW:

$$\delta(\hat{s} - m_X^2) \to \frac{m_X \Gamma_X}{\pi} \frac{1}{(\hat{s} - m_X^2)^2 + m_X^2 \Gamma_X^2}. \tag{7.21}$$

Other forms of propagator aim to improve on the simple relativistic BW by including effects of the non-constant width. They are given below:

$$\delta(\hat{s} - m_X^2) \to \frac{1}{\pi} \frac{\hat{s}\Gamma_X/m_X}{(\hat{s} - m_X^2)^2 + (\hat{s}\Gamma_X/m_X)^2}, \tag{7.22}$$

$$\delta(\hat{s} - m_X^2) \to \frac{1}{\pi} \frac{\sqrt{\hat{s}}\Gamma_{X\to\gamma\gamma}(\hat{s})}{(\hat{s} - m_X^2)^2 + m_X^2 \Gamma_X^2}, \tag{7.23}$$

where $\Gamma_{X\to\gamma\gamma}$ is the partial decay width of the resonance X to the diphoton final state.

Eq. 7.22 is a refinement of Eq. 7.21 as it introduces an $\hat{s}$-dependent width through the substitution of $\Gamma_X \to \Gamma_X \hat{s}/m_X^2$. This is the running-width scheme used by the PowHeg generator.[1] Eq. 7.23 contains a virtuality-dependent partial width to the decay channel being probed. This is the prescription used by the MC@NLO generator [95].

7.2 Large-Width Higgs-like Resonance Models

For a scalar resonance, the most straightforward model is to assume a SM-like Higgs with high mass since the Higgs is the only known scalar particle. In the SM, the width of the Higgs boson is determined by its mass. To keep the search generic, the large-width model being studied here is not restricted to the value of a heavy Higgs width in the SM. To simulate the heavy Higgs boson process, the PowHeg generator is used as it provides NLO accuracy in QCD and has been used for the SM Higgs processes. The MC@NLO implementation of a heavy Higgs is also discussed here.

The line shape of a large-width resonance can be described by substituting the general form of Eq. 7.20 with the relevant propagator. In the PowHeg and MC@NLO prescriptions, only the squared matrix element of the production is needed, while the decay part is absorbed into the propagator.

A review of the line shape of a heavy Higgs boson is given in Section 7.2.1. Section 7.2.2 describes the details in generating the line shape for a Higgs-like resonance as implemented by the PowHeg generator. An alternative prescription using a different propagator from MC@NLO prescription is described next in Section 7.2.3.

7.2.1 Review of Heavy Higgs Line Shape

Prior to the discovery of a Higgs boson at 125 GeV, some studies had been made on the line shape of a heavy Higgs boson, which would have a large width. This is discussed in Ref. [96]. In the following, some remarks and observations on the line shape based on these studies are given.

The running-width Breit-Wigner (Eq. 7.22) was found to describe well the Z boson line shape, for which the $\hat{s}$-dependent width term was suggested by the phase-space and the structure of the electroweak radiative corrections within the SM [97]. Nevertheless, the running-width BW cannot be derived from quantum field theory and hence is not justifiable [98]. In addition, the decay width depends linearly on $\hat{s}$ for the Z boson but not for the Higgs [99]. Therefore, Eq. 7.22 is ill-suited to describe a heavy Higgs.

A SM Higgs with a mass above 135 GeV has a higher decay width to four fermions than to two b-quarks. This implies that in terms of the imaginary part of the Higgs self energy, three-loop diagrams are as important as one-loop diagrams [96]. In the on-shell renormalisation scheme, the perturbation theory is shown to diverge at about 1 TeV [100]. Hence, the on-shell treatment of a heavy Higgs is inadequate.

A proper treatment of the Higgs complex pole is introduced, called the complex-pole scheme (CPS) [96,101]. In the CPS, the denominator of the propagator is $\hat{s} - s_H$ where s_H is the Higgs complex pole, parameterised as

$$s_H = \mu_H^2 - i\mu_H\gamma_H. \tag{7.24}$$

Performing the following transformation,

$$\overline{M}_H^2 = \mu_H^2 + \gamma_H^2,\ \mu_H\overline{\Gamma}_H = \overline{M}_H\gamma_H, \tag{7.25}$$

the Higgs propagator becomes

$$\frac{1}{\hat{s} - s_H} = \left(1 + i\frac{\overline{\Gamma}_H^2}{\overline{M}_H^2}\right)\frac{1}{\hat{s} - \overline{M}_H^2 + i\frac{\overline{\Gamma}_H}{\overline{M}_H}\hat{s}}. \tag{7.26}$$

[1] by setting `bwshape=1`.

Hence, the complex-pole scheme is equivalent to introducing a running width in the propagator with parameters that are not the on-shell ones. The complex pole is gauge-independent to all orders of perturbation theory. In the CPS, the production, propagator and decay terms should be calculated using the complex pole but such calculations are difficult. Instead, the more pragmatic approach of using only the Higgs propagator with its complex pole but not the production and decay terms is usually implemented.[2] This was developed for the search for a high mass Higgs and has also been used to study the off-shell Higgs event yield in the $h \to ZZ^*$ and $h \to WW^*$ channels in ATLAS [102] and CMS [103].

The running-width BW in Eq. 7.22 is able to describe the line shape of a heavy Higgs boson up to $m_H \sim 300 - 400$ GeV [104]. For $m_H \geq 400$ GeV, the CPS provides a more accurate description. However, the drawback of CPS is that it can only be used for a heavy Higgs with SM width, i.e. the width is determined by the value of the mass (see Figure 7.2). To remain agnostic about new physics, no assumption on a value of the resonance width is made in the analysis but rather it is a free parameter.

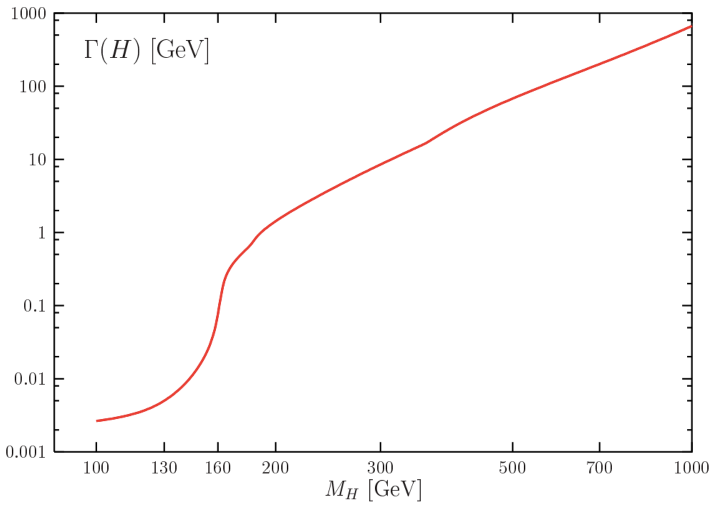


Figure 7.2: The SM Higgs boson total decay width as a function of the Higgs mass m_H [14]. The width of SM Higgs increases rapidly with its mass. For a 400 GeV Higgs boson, its width is 29 GeV, and for a 1 TeV Higgs, its width is as high as 647 GeV.

7.2.2 Heavy Higgs Line Shape in PowHeg

Higgs boson production via gluon fusion (ggH) is simulated at next-to-leading-order (NLO) accuracy in QCD using the PowHeg Box [105] generator, with the CT10 parton distribution function (PDF) for a narrow-width signal and for large widths up to 10% of the resonance mass. The choice of generator was motivated by the SM Higgs analysis [106].

The total cross-section of the production of a Higgs-like resonance through gluon-gluon fusion as implemented in PowHeg [105] is

$$\sigma_{\text{tot}} = \sigma_{\bar{B}} + \sum_q \sigma_{R_{q\bar{q}}}, \tag{7.27}$$

where $\sigma_{\bar{B}}$ includes the contributions from the leading order $gg \to H$ production and the NLO subprocesses $gg \to Hg$, $qg \to Hq$, $gq \to Hq$, while $\sigma_{R_{q\bar{q}}}$ corresponds to the NLO corrections from the subprocess $q\bar{q} \to Hg$.

Showing only the leading contribution from $gg \to H$ process, $\sigma_{\bar{B}}$ is given by

[2] This is known as the OFFP-scheme [96].

$$\sigma_{\bar{B}} = \int [\mathcal{B}_{gg}\mathcal{L}_{gg}] \frac{2\pi}{s}\delta(\hat{s} - m_X^2) d\hat{s} dY, \tag{7.28}$$

where s is the squared center-of-mass energy of the hadronic collider; Y is the rapidity of the resonance X; $\mathcal{B}_{gg}$ is the squared matrix element of the Born contribution to the $gg \to H$ process while $\mathcal{L}_{gg}$ is the gluon-gluon luminosity. In the equation, the on-shell approximation is used and the effect of large width is implemented by replacing the δ-function with the BW in the PowHeg implementation, i.e. Eq. 7.22 that has an $\hat{s}$-dependent width.

The LO differential cross-section for a non-zero width Higgs-like resonance is implemented in PowHeg as[3]

$$\frac{d\sigma_{\text{PowHeg}}}{d\hat{s}} \propto \mathcal{B}_{gg} \cdot \mathcal{L}_{gg} \cdot f_{BW}, \tag{7.29}$$

where f_{BW} is a Breit-Wigner function (in this case, Eq. 7.22).

The line shape generated is the same across all final states since $\mathcal{B}_{gg}$ is the squared matrix element of only the production process and the decay terms are described by the BW function regardless of the final decay products.

Since $\hat{s}$ is measured by the squared of the invariant mass of the final states of the resonance, $\hat{s} = m_{\gamma\gamma}^2$ and $d\hat{s} = 2m_{\gamma\gamma}dm_{\gamma\gamma}$. The differential cross-section can thus be written as

$$\frac{d\sigma_{\text{PowHeg}}}{dm_{\gamma\gamma}} \propto \mathcal{B}_{gg} \cdot \mathcal{L}_{gg} \cdot f_{BW} \cdot m_{\gamma\gamma}. \tag{7.30}$$

The squared matrix element of the production process $\mathcal{B}_{gg}$ is given by [105]

$$\mathcal{B}_{gg} = \frac{\alpha_s^2}{\pi^2}\frac{G_F\hat{s}}{576\sqrt{2}} \left| \frac{3}{2}\sum_Q \tau_Q \left[1 + (1-\tau_Q)f(\tau_Q)\right] \right|^2, \tag{7.31}$$

where $\tau_Q = 4m_Q^2/\hat{s}$ and the sum runs over the heavy quarks with mass m_Q circulating in the gluon-gluon fusion production loop. The function f is

$$f(\tau_Q) = \begin{cases} \arcsin^2 \frac{1}{\sqrt{\tau_Q}} & \tau_Q \geq 1, \\ -\frac{1}{4}\left[\log\left(\frac{1+\sqrt{(1-\tau_Q)}}{1-\sqrt{(1-\tau_Q)}}\right) - i\pi\right]^2 & \tau_Q < 1. \end{cases} \tag{7.32}$$

In the implementation, only the contribution from the top quark is retained.

Figure 7.3 shows the parton luminosities obtained with the CT10 PDF set [107], which is used to produce the PowHeg MC samples. APFEL Web [108, 109] is used to produce the distribution of parton luminosities, given a PDF set. The gluon-gluon luminosity, like the other parton luminosities, falls as a function of mass, resulting in higher cross-sections at lower $\sqrt{\hat{s}}$ than at higher $\sqrt{\hat{s}}$. This leads to a deformation of the invariant mass distributions. The gluon-gluon luminosity distribution in Figure 7.3 is parameterised with an analytical function.

The true invariant mass distributions from the PowHeg MC samples are compared to the line shapes given by Eq. 7.30 for all the mass and width values where MC samples are available. The comparison for $m_X = 400, 1500$ GeV with $\Gamma_X/m_X = 5, 10\%$ are shown as examples in Figures 7.4 and 7.5. In these and following figures showing the true $m_{\gamma\gamma}$ distributions in MC, the diphoton events at the truth level are required to originate from the resonance. The samples used to study the signal shape have a pre-selection applied on the reconstructed photon candidates, where they are required to have $p_T > 25$ GeV, be within the calorimeter pseudorapidity acceptance and pass loose photon identification requirements.

[3] In the PowHeg implementation, the terms $\frac{1}{\hat{s}}|a_{ii\to X}(\hat{s})a_{X\to\gamma\gamma}(\hat{s})|^2\delta(\hat{s} - m_X^2)$ in Eq. 7.20 become $\mathcal{B}_{gg}f_{BW}$.

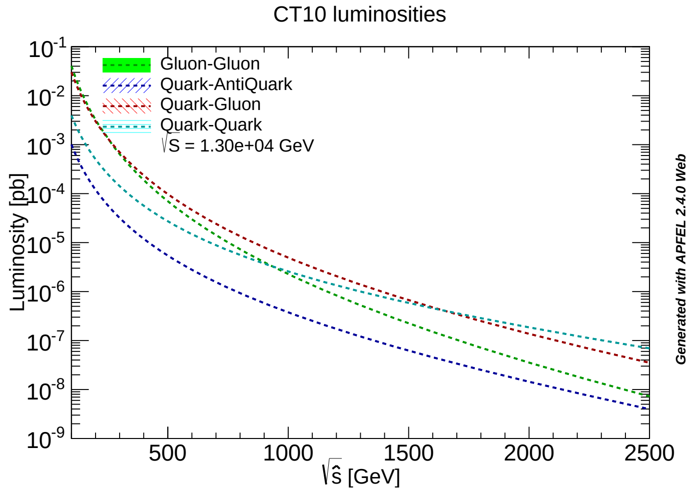
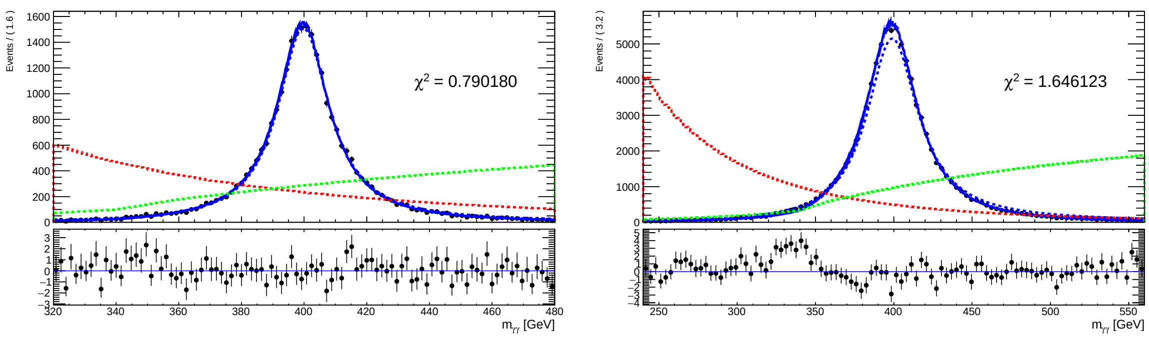


Figure 7.3: The parton luminosities as a function of $\sqrt{\hat{s}}$ obtained with the CT10 PDF.

Very good agreement is found between the line shapes given by Eq. 7.30 and the $m_{\gamma\gamma}$ distributions in MC even though only the LO gluon-gluon initiated process is taken into account in modelling the line shape. Other corrections ($gg \to Hg$, $qg \to Hq$, $gq \to Hq$, etc.) present in the Monte Carlo are found to be insubstantial in changing the line shape.



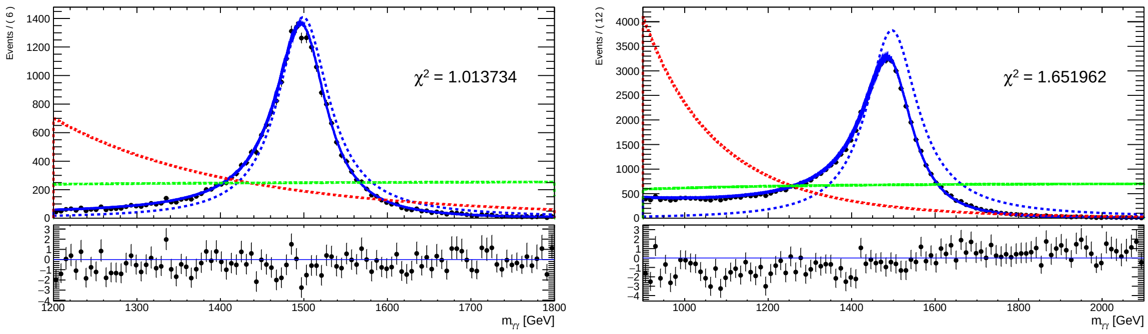
(a) (b)

Figure 7.4: True $m_{\gamma\gamma}$ distributions of the resonance with $m_X = 400$ GeV and width of (a) 5% and (b) 10% of m_X. The dashed red line represents the parameterisation of the gluon-gluon luminosity $\mathcal{L}_{gg}$; the dashed green line represents $\mathcal{B}_{gg} m_{\gamma\gamma}$; the dashed blue line is the running-width BW distribution with the same mass and width as generated in the sample. The product of these three is represented by the solid blue line and agrees well with the true invariant mass distribution generated with PowHeg, shown with black markers.

7.2.3 Problems and Alternatives

Figure 7.5 shows only the signal mass distribution in a restricted mass range around the mass peak. Checking the mass distributions over an extended mass range, one finds that the signal yield is enhanced around 400 GeV, creating a secondary bump in the signal mass distribution away from the resonance mass.[4] This is illustrated for $m_X = 1500$ GeV and width of 150 GeV (10% m_X) in Figure 7.6. This bump is very prominent for large resonance mass and increases in size with the resonance width. The origin of the secondary bump is tracked to the squared

[4] At the analysis-level, the reconstructed $m_{\gamma\gamma}$ distribution also shows a similar bump after the kinematic selection cuts have been applied.



(a) (b)

Figure 7.5: True $m_{\gamma\gamma}$ distributions of the resonance with $m_X = 1500$ GeV and width of (a) 5% and (b) 10% of m_X. The dashed red line represents the parameterisation of the gluon-gluon luminosity $\mathcal{L}_{gg}$; the dashed green line represents $\mathcal{B}_{gg} m_{\gamma\gamma}$; the dashed blue line is the running-width BW distribution with the same mass and width as generated in the sample. The product of these three is represented by the solid blue line and agrees well with the true invariant mass distribution generated with PowHeg, shown with black markers.

matrix element of the production process, as shown in Figure 7.7 where only the top quark contribution is taken into account.

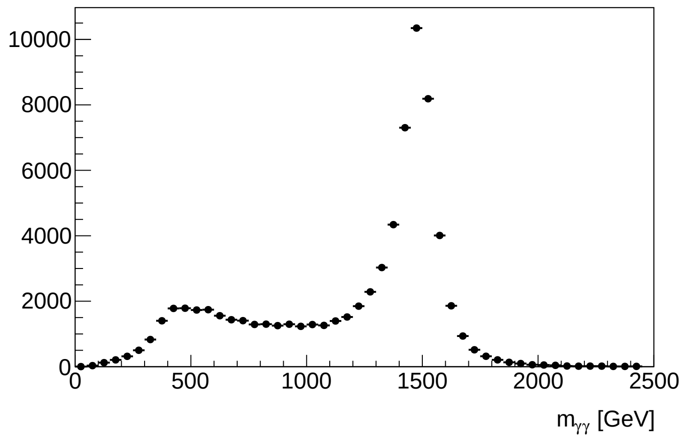


Figure 7.6: The true $m_{\gamma\gamma}$ distribution of the resonance with $m_X = 1500$ GeV and width of 150 GeV (10% m_X) showing also the off-shell region. A prominent bump is seen at around $m_{\gamma\gamma} \approx 400$ GeV.

This problem was also noticed in another analysis in ATLAS searching for high mass resonance decaying to a pair of W bosons. A full (off-shell) calculation of the cross-section was performed, i.e. without assuming narrow width in the first step. Therefore, the full computation does not require the choice, *a posteriori*, of an ad-hoc propagator. This study was done assuming the resonance has properties similar to a SM Higgs, albeit with a higher mass, such that $\Gamma_{X\to WW}$ has the same $\hat{s}$-dependence as the SM Higgs. The resulting line shape was compared to the on-shell computation using the PowHeg (with propagator in Eq. 7.22) and the MC@NLO (with propagator in Eq. 7.23) prescriptions. The MC@NLO prescription is found to reproduce well the line shape from the full calculation. They both exhibit a secondary mass bump around ~ 400 GeV, but much less pronounced than the results obtained with the running-width propagator as it includes the evolution of the partial width $\Gamma(X \to WW)$ with

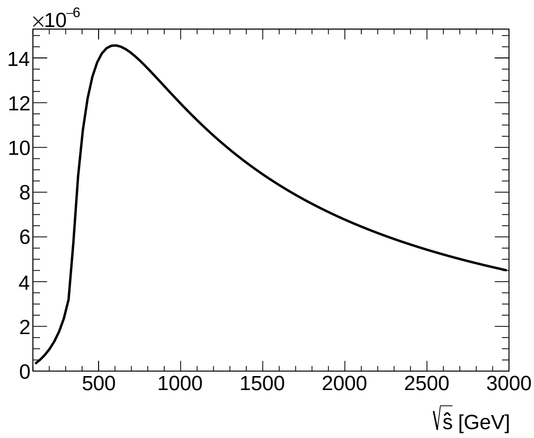


Figure 7.7: The squared matrix element of the gluon-gluon production process $\mathcal{B}_{gg}$ as a function of $\sqrt{\hat{s}}$, assuming only top quarks in the loop. The mass of the top is set to 173 GeV.

the virtuality.

The full calculation for this process as well as for the ZZ final state is available in MG5_aMC@NLO. However, performing full calculation for the $\gamma\gamma$ final state proves to be more difficult due to the loop-induced decay of the Higgs to two photons (see Chapter 1). In fact, it is this loop process that poses yet another problem. Assuming the resonance has couplings like a SM Higgs to other particles, its decay to two photons proceeds via a loop process including W and top contributions. The contribution of the W boson loop interferes destructively with the quark loop and the two contributions nearly cancel each other for a resonance mass of about 640 GeV. This is demonstrated in the partial width for the decay $X \to \gamma\gamma$ as a function of m_X in Figure 7.8.

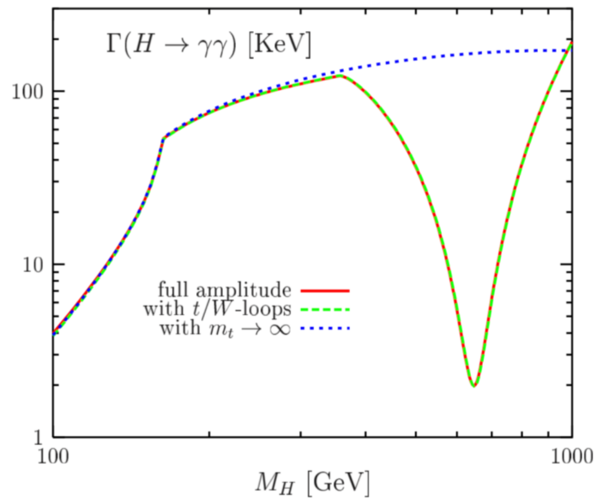


Figure 7.8: The partial width for the decay $X \to \gamma\gamma$ as a function of m_X, assuming X has SM Higgs couplings to the W boson and the top quark. In this figure, the resonance X is actually the SM boson Higgs denoted H. The solid line shows the partial width with the W and all third generation fermion contributions, the dashed green line shows the partial width with W and only the top quark contribution, and the dotted blue line with the W and top quark contributions for $m_t \to \infty$ [14].

Figure 7.9 shows the shape of the $m_{\gamma\gamma}$ invariant mass distribution as described by Eq. 7.30, where the term f_{BW} is taken from Eq. 7.21, 7.22 or 7.23. The mass and width are set to $m_X = 1.5$ TeV and $\Gamma_X/m_X = 10\%$. While the three curves agree reasonably well around the peak, significant differences in the low mass part of the spectrum are observed for all three assumptions. Using f_{BW} from Eq. 7.21 assumes the width stays constant with $\hat{s}$ which is unrealistic, especially for a wide resonance. For Eq. 7.23, the dependence of $\Gamma_X \to \gamma\gamma$ on $\hat{s}$ is assumed to be identical as the one of the SM Higgs as shown in Figure 7.8, which includes a dip around 640 GeV from the destructive interference between the W and the top loop-induced contributions to the decay. This reduces the size of the 400 GeV bump in the diphoton invariant mass distribution, but creating a dip near 640 GeV in the process.

It is clear from Eq. 7.23 and the expression in Eq. 7.10 that the X line shape *does* depend on the final state. Unlike the case of the WW final state where the assumption of tree-level SM Higgs-like coupling leads to a mild distortion of the line shape, one gets a large distortion in the decay to $\gamma\gamma$. It is interesting to note that in this assumption, for $m_X \approx 640$ GeV, the resonance peak is shifted and distorted by a huge amount, as shown in Fig. 7.10.

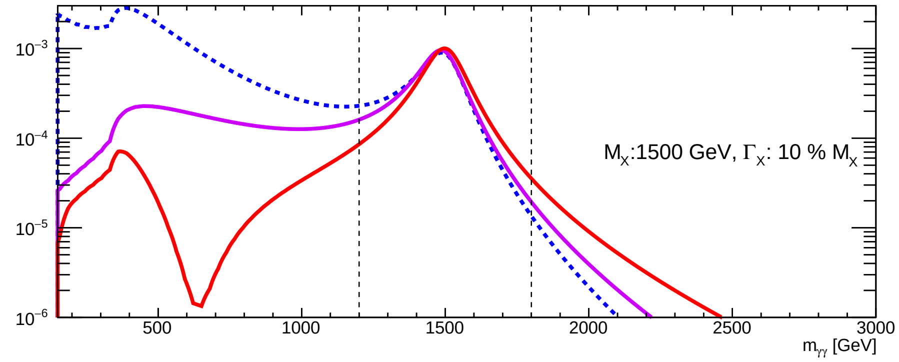


Figure 7.9: The line shape of the true $m_{\gamma\gamma}$ distribution for a resonance with mass $m_X = 1.5$ TeV and $\Gamma_X = 10\%\, m_X$ with different f_{BW} terms. The line shape with BW form taken from Eq. 7.21, 7.22 and 7.23 are shown in dashed blue line, solid violet line and solid red line respectively. Vertical dashed lines are drawn at $m_X \pm 2\Gamma_X$.

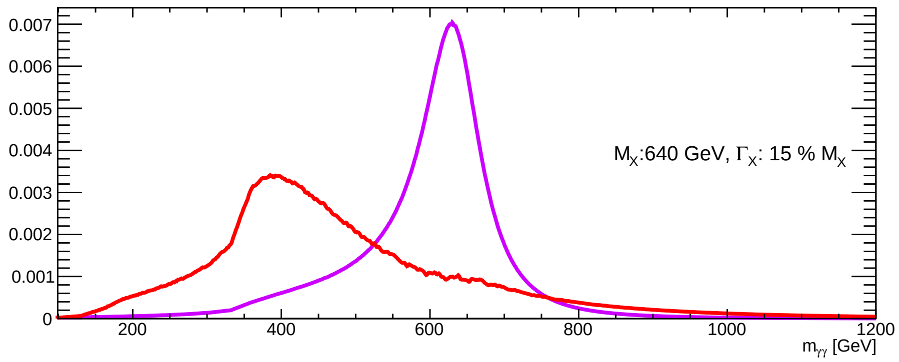


Figure 7.10: The line shape of the true $m_{\gamma\gamma}$ distribution for a resonance with mass $m_X = 640$ GeV and $\Gamma_X = 15\%\, m_X$ with f_{BW} taken from Eq. 7.23 and assuming SM Higgs-like coupling is shown in red. In this assumption, the expected peak at $m_X = 640$ GeV completely disappears and shifts to a much smaller value. The line shape obtained using POWHEG prescription (Eq. 7.22) is shown in violet for comparison.

Figure 7.11 shows, in the dotted green line, how the theoretical line shape of a heavy Higgs decaying in the WW final state would look by substituting the $H \to \gamma\gamma$ partial width for that of $H \to WW$ [14] using the MC@NLO prescription (Eq. 7.23). The theoretical line shape with the $H \to \gamma\gamma$ partial width (Eq. 7.23) as well as that from POWHEG model (Eq. 7.22) are shown in red and violet line respectively for comparison.

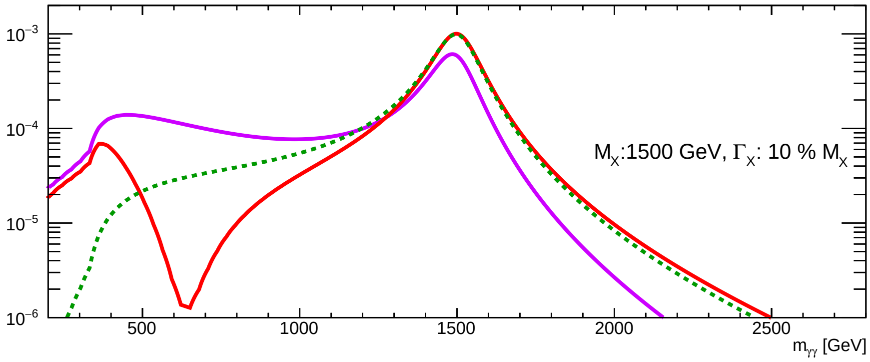


Figure 7.11: Using the BW form in Eq. 7.23, the theoretical line shape of a 1.5 TeV resonance with 150 GeV width decaying to WW final state (plotted in dashed green line) is compared to that of $\gamma\gamma$ final state (shown in solid red line), assuming the resonance has couplings similar to the SM Higgs. The theoretical line shape of a resonance with the same mass and width using PowHeg prescription which is independent of the final state is plotted with the violet line.

The off-shell calculation for the $X \to \gamma\gamma$ decay was not available in MG5_aMC@NLO at the time of this study. However, it would likely lead to the same behaviour of the line shape since the dip is not an artefact but reflects a specific assumption on the decay model. One can slightly modify the assumption by turning off the destructive interference in the decay loop with the top mass set to an arbitrarily high value. By doing so, one restores a smooth behaviour. However, this corresponds to a very specific and unmotivated physics model where the gluon-gluon fusion production proceeds via 173 GeV top quark loop but the decay to diphoton is mediated by W boson and much higher mass fermion loops.[5]

For the search of a high mass resonance decaying to $\gamma\gamma$ in the 2015 data, the PowHeg model is used as the baseline. However, in order to avoid potential effects from the off-shell region, the analysis was restricted to the resonance peak region. A mass window around the peak was therefore included in the definition of the fiducial volume (see Section 8.5.2). To determine the mass window, the line shape with the propagator given by Eq. 7.23, with the top mass set to 1 TeV, was used as an alternative description. A window of $\pm 2\Gamma_X$ was ultimately chosen as it contains a large part of the resonance and is insensitive to the off-shell region particularly where the secondary bump becomes prominent.

7.3 Large-Width Scalar Singlet Model

For the 2016 ICHEP analysis and beyond, the PowHeg Higgs-like signal model has been replaced by an effective field theory (EFT) approach in the Higgs Characterisation model [110], at next-to-leading order in QCD using MG5_aMC@NLO generator. This approach is at least partly generic, since the effective operators at low energies can resolve to a variety of BSM physics scenarios once the cut off scale for the theory is reached (including loops with extra fermions or scalars). At low energies, this new physics is integrated out, and thus one does not need to make explicit assumptions about what the high scale physics is. Furthermore, by operating at NLO in QCD, the benefit of an excellent jet radiation model is retained as is the case with the PowHeg generator.

[5]One could also assume very high mass fermion loops in both production and decay processes. This leads to a line shape that is similar to the EFT approach explained in next Section (see Appendix B for derivation).

7.3.1 The Higgs Characterisation Model

The Higgs Characterisation model was developed to study the properties of the newly discovered 125 GeV boson. To study the spin and CP properties of this boson, the model uses a FEYNRULES [111] implementation of an effective Lagrangian that consist of SM terms (except for the Higgs particle) and a new bosonic state that can couple to SM particles. In the model, the bosonic state could correspond to the SM Higgs with spin-0 and CP-even, or a BSM boson.

Thus, the same theory can be repurposed to apply to a generic new scalar resonance with new physics appearing in loop diagrams at energies higher than the cutoff scale. This corresponds to introducing a scalar that is a singlet under the SM gauge group, which is one of the simplest extensions to the SM. Effective operators are used to model the scalar singlet interactions with SM particles.

The effective Lagrangian for scalar interactions with SM gauge fields is written as [110]:

$$\begin{aligned}
\mathcal{L}_0^V = \Big\{ & c_\alpha \kappa_{\rm SM} \big[\frac{1}{2} g_{HZZ}\, Z_\mu Z^\mu + g_{HWW}\, W_\mu^+ W^{-\mu} \big] \\
& - \frac{1}{4} \big[c_\alpha \kappa_{H\gamma\gamma} g_{H\gamma\gamma}\, A_{\mu\nu} A^{\mu\nu} + s_\alpha \kappa_{A\gamma\gamma} g_{A\gamma\gamma}\, A_{\mu\nu} \widetilde{A}^{\mu\nu} \big] \\
& - \frac{1}{2} \big[c_\alpha \kappa_{HZ\gamma} g_{HZ\gamma}\, Z_{\mu\nu} A^{\mu\nu} + s_\alpha \kappa_{AZ\gamma} g_{AZ\gamma}\, Z_{\mu\nu} \widetilde{A}^{\mu\nu} \big] \\
& - \frac{1}{4} \big[c_\alpha \kappa_{Hgg} g_{Hgg}\, G^a_{\mu\nu} G^{a,\mu\nu} + s_\alpha \kappa_{Agg} g_{Agg}\, G^a_{\mu\nu} \widetilde{G}^{a,\mu\nu} \big] \\
& - \frac{1}{4} \frac{1}{\Lambda} \big[c_\alpha \kappa_{HZZ}\, Z_{\mu\nu} Z^{\mu\nu} + s_\alpha \kappa_{AZZ}\, Z_{\mu\nu} \widetilde{Z}^{\mu\nu} \big] \\
& - \frac{1}{2} \frac{1}{\Lambda} \big[c_\alpha \kappa_{HWW}\, W^+_{\mu\nu} W^{-\mu\nu} + s_\alpha \kappa_{AWW}\, W^+_{\mu\nu} \widetilde{W}^{-\mu\nu} \big] \\
& - \frac{1}{\Lambda} c_\alpha \big[\kappa_{H\partial\gamma}\, Z_\nu \partial_\mu A^{\mu\nu} + \kappa_{H\partial Z}\, Z_\nu \partial_\mu Z^{\mu\nu} + \big(\kappa_{H\partial W}\, W_\nu^+ \partial_\mu W^{-\mu\nu} + h.c. \big) \big] \Big\} X_0 \,,
\end{aligned} \tag{7.33}$$

where $g_{Xyy'}$ is the strength of the coupling of a scalar ($X = H$) or pseudoscalar ($X = A$) boson to the SM bosons ($y, y' = Z, W, \gamma, g$) in the SM and $\kappa_{Xyy'}$ is the dimensionless coupling parameter.

The (reduced) field strength tensors are defined as

$$V_{\mu\nu} = \partial_\mu V_\nu - \partial_\nu V_\mu (V = A, Z, W^\pm), \tag{7.34}$$

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f^{abc} G^b_\mu G^c_\nu, \tag{7.35}$$

and the dual tensor is

$$\widetilde{V}_{\mu\nu} = \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} V^{\rho\sigma}. \tag{7.36}$$

The Lagrangian introduces scalar-SM interactions via dimension-5 operators, with both CP-even and CP-odd interactions included. The parameters c_α and s_α are the cosine and sine of a mixing angle that parameterises the ratio of the CP-even and CP-odd contributions. Given the lack of sensitivity to the CP properties of a new resonance, the model is simplified by turning off CP-odd terms through setting c_α=1.

The process of a gluon-gluon fusion produced scalar decaying to two photons is described by the effective dimension-5 operators[6] in the terms

[6]Operators like $G^a_{\mu\nu} G^{a,\mu\nu} X_0$ are gauge invariant (under the SM gauge group) since $G^a_{\mu\nu} G^{a,\mu\nu}$ is already a gauge invariant operator in the SM and the X_0 does not transform under these gauge transformations.

$$-\frac{1}{4}\left[c_\alpha \kappa_{H\gamma\gamma} g_{H\gamma\gamma}\, A_{\mu\nu}A^{\mu\nu}X_0\right] \tag{7.37}$$

and

$$-\frac{1}{4}\big[c_\alpha \kappa_{Hgg} g_{Hgg}\, G^a_{\mu\nu}G^{a,\mu\nu}X_0. \tag{7.38}$$

The Feynman diagram for this process is shown in Figure 7.12.

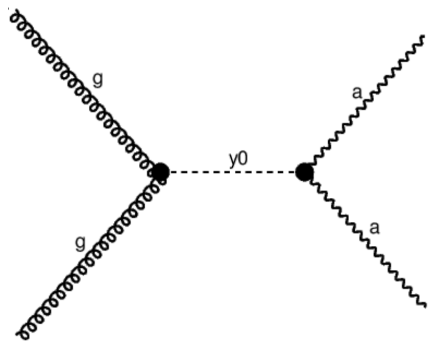


Figure 7.12: Feynman diagram of the process of gluon-gluon fusion produced scalar decaying to two photons with effective ggX and $X\gamma\gamma$ vertices.

The total width of the resonance is computed from all its possible decay processes. However, in the analysis, no assumption on the width is made in order to be model-independent. Instead, a range of width values are generated by setting the width to a given value. Contributions from the decay of the scalar to other particles are hence not relevant. To simplify the model, only the terms that lead to a scalar decaying to photons produced via gluon fusion (7.37 and 7.38) are turned on. Other terms as well as scalar interactions with fermions are set to zero. Since the width is fixed in the generation, the absolute value of the cutoff scale and the coefficients do not affect the line shape, but only the overall normalisation, which is not required in the analysis.

Samples are generated with the NNPDF 3.0 NLO [112] PDF, using the A14 tune of Pythia8 for the parton shower. Range of signal mass between 200 GeV and 2400 GeV with narrow (4.2 MeV) to large ($\leq 15\%\, m_X$) widths are simulated. Ultimately, widths up to only 10% of the resonance mass are considered.

7.3.2 Parameterisation of the Line Shape

Although the on-shell computation such as the one implemented in PowHeg is only an approximation, it allows the line shape for any reasonable resonance mass and width values to be described by simple functions. This requires fewer MC samples to be generated for signal modelling purposes and the line shape can be implemented directly in the statistical model. In this spirit, an attempt to parameterise the line shape from MG5_aMC@NLO in a similar fashion is carried out.

In the MG5_aMC@NLO generator, the relativistic Breit-Wigner distribution as shown in Eq. 7.21 is used. Substituting the propagator in Eq. 7.10, the differential cross-section (in $\hat{s}$) for the $gg \to X \to \gamma\gamma$ process can be written as

$$d\sigma_{EFT} \propto \int d\hat{s}\mathcal{L}_{gg}(\hat{s})\frac{1}{\hat{s}}\frac{|a_{gg\to X}(\hat{s})|^2|a_{X\to\gamma\gamma}(\hat{s})|^2}{(\hat{s}-m_X^2)^2+(m_X\Gamma_X)^2}. \tag{7.39}$$

Using CalcHEP [113] and Mathematica, one can compute the squared of the amplitude $\mathcal{A}$ defined from the effective Lagrangian:

$$|\mathcal{A}|^2 \propto \frac{\hat{s}^4}{(m_X^2 - \hat{s})^2}. \tag{7.40}$$

Replacing the leading-order propagator in CalcHEP by the relativistic Breit-Wigner distribution and relating the squared matrix element to the terms $|a_{gg\to X}(\hat{s})|^2 |a_{X\to\gamma\gamma}(\hat{s})|^2$ through Eq. 7.9, we find that $|a_{gg\to X}(\hat{s})|^2 |a_{X\to\gamma\gamma}(\hat{s})|^2 = \hat{s}^4$. Substituting this in Eq. 7.39, we obtain

$$\frac{d\sigma_{EFT}}{d\hat{s}} \propto \mathcal{L}_{gg} \cdot \frac{\hat{s}^4}{\hat{s}} \cdot f_{BW} = \mathcal{L}_{gg} \cdot \hat{s}^3 \cdot f_{BW}. \tag{7.41}$$

Since $m_{\gamma\gamma}^2 = \hat{s}$ (hence $d\hat{s} = 2m_{\gamma\gamma}dm_{\gamma\gamma}$), the line shape can be parameterised as

$$\frac{d\sigma_{EFT}}{dm_{\gamma\gamma}} \propto \mathcal{L}_{gg} \cdot m_{\gamma\gamma}(m_{\gamma\gamma})^6 \cdot f_{BW} = \mathcal{L}_{gg} \cdot m_{\gamma\gamma}^7 \cdot f_{BW}. \tag{7.42}$$

Figure 7.13 shows the parton luminosities obtained with the NNPDF 3.0 NLO PDF set, which is used to simulate the MC samples. In Figures 7.14, 7.15 and 7.16, the LO analytic description of the line shape is compared to the truth distribution taken from the MC simulation.[7] This confirms that the effect of NLO corrections does not distort the conclusion reached with the LO derivation, as observed in the PowHeg Higgs-like model.

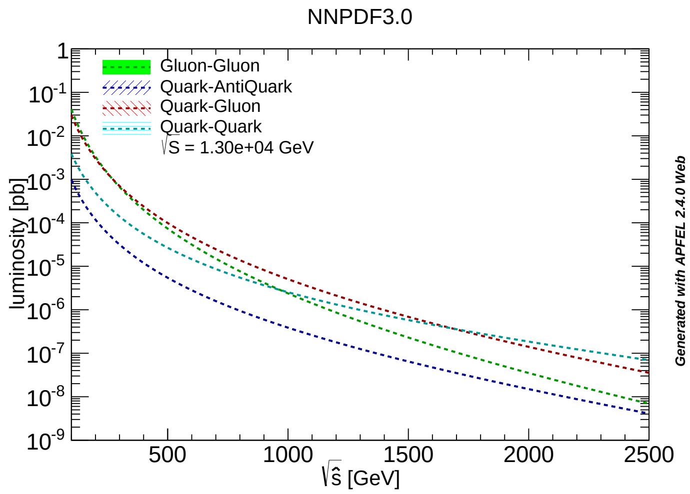
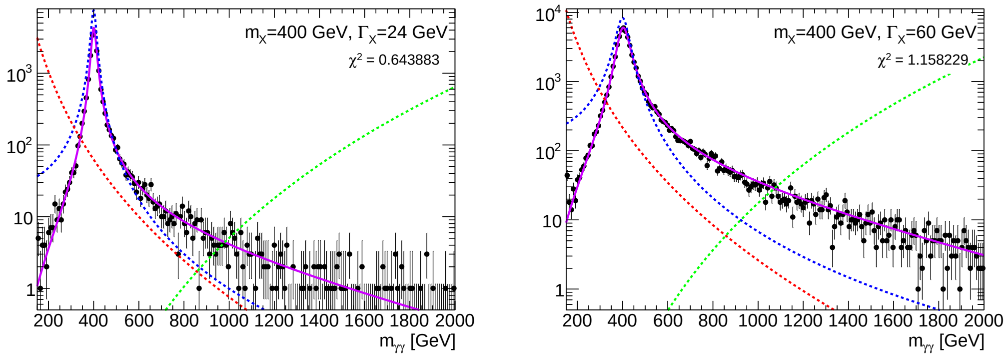


Figure 7.13: The parton luminosities as a function of $\sqrt{\hat{s}}$ obtained with the NNPDF 3.0 NLO PDF. The APFEL Web is used to draw the luminosities [108, 109].

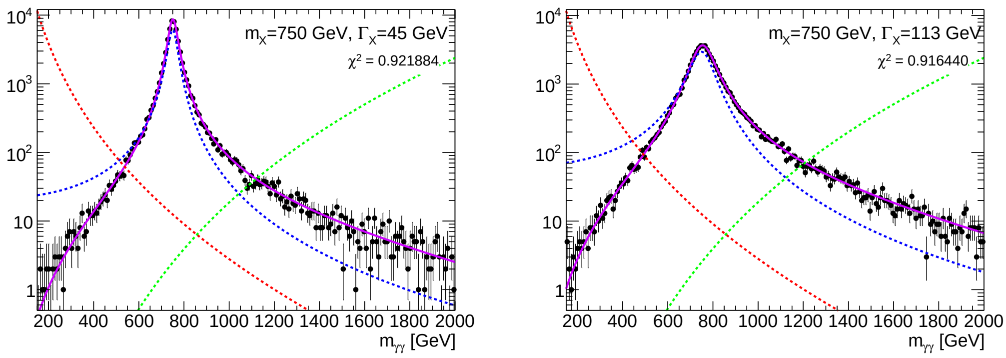
The comparison of the theoretical line shape of the scalar singlet EFT to that of the previously used the PowHeg model is shown in Figure 7.17 under three mass hypotheses, all with $\Gamma_X = 10\%\, m_X$. The secondary bump does not exist in this EFT implementation. The asymmetry of the tail distribution is also changed. Previously in Figure 7.11, the line shape using the PowHeg prescription, which does not take into account the decay width, shows an enhanced tail on the left of the mass distribution and the mass peak is also lower with respect to the generated mass. In the scalar singlet model, the production and decay squared matrix elements, which increase rapidly with mass, compete with the falling gluon-gluon luminosity. This results in a mass shape that is fairly symmetric for large resonance masses but with an enhanced upper tail for low m_X.

[7] The mass and width values being generated in the MG5_aMC@NLO samples are different to those of the PowHeg samples. Hence the examples being presented here do not have the same mass nor width as those in Section 7.2.2.



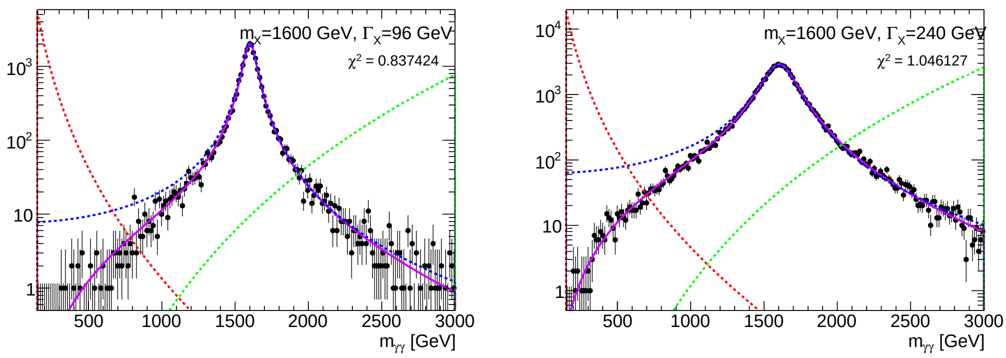
(a) (b)

Figure 7.14: True $m_{\gamma\gamma}$ distributions for a resonance with m_X=400 GeV and width of 6% (left) and 15% (right) of the m_X value. In both plots, the dashed red line represents the parameterisation of the gluon-gluon luminosity, the dashed green line is the functional form $m_{\gamma\gamma}^7$, and the dashed blue line is the Breit-Wigner distribution with the same mass and width as generated in the sample. The product of these three is represented by the solid purple line and agrees well with the true invariant mass distribution generated with MG5_aMC@NLO, shown with black dots.



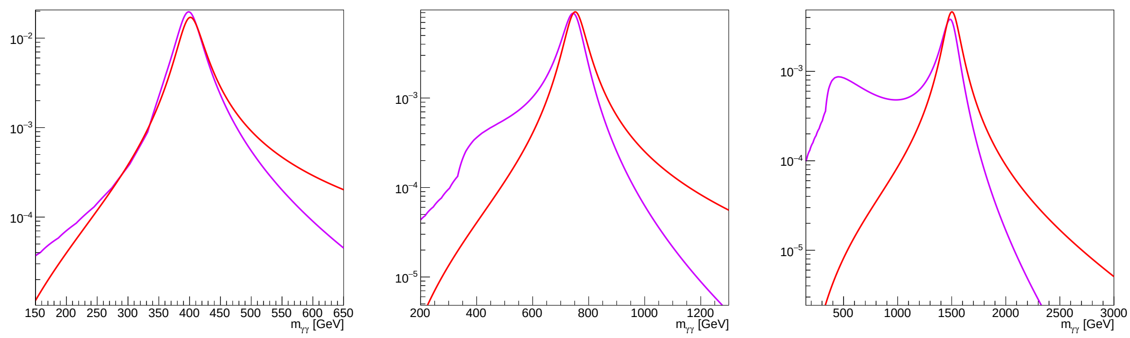
(a) (b)

Figure 7.15: True $m_{\gamma\gamma}$ distributions for a resonance with m_X=750 GeV and width of 6% (left) and 15% (right) of the m_X value. In both plots, the dashed red line represents the parameterisation of the gluon-gluon luminosity, the dashed green line is the functional form $m_{\gamma\gamma}^7$, and the dashed blue line is the Breit-Wigner distribution with the same mass and width as generated in the sample. The product of these three is represented by the solid purple line and agrees well with the true invariant mass distribution generated with MG5_aMC@NLO, shown with black dots.



(a) (b)

Figure 7.16: True $m_{\gamma\gamma}$ distributions for a resonance with m_X=1600 GeV and width of 6% (left) and 15% (right) of the m_X value. In both plots, the dashed red line represents the parameterisation of the gluon-gluon luminosity, the dashed green line is the functional form $m_{\gamma\gamma}^7$, and the dashed blue line is the Breit-Wigner distribution with the same mass and width as generated in the sample. The product of these three is represented by the solid purple line and agrees well with the true invariant mass distribution generated with MG5_aMC@NLO, shown with black dots.



(a) (b) (c)

Figure 7.17: Comparison of the theoretical line shape of a Higgs-like model using PowHeg prescription (violet line) and that of a scalar singlet using MG5_aMC@NLO (red line) under three mass hypotheses: (a) 400 GeV, (b) 750 GeV and (c) 1500 GeV. The resonance widths in these plots are 10% of the m_X value.

7.4 Impact of Interference between the Spin-0 Resonance and Continuum on the Line Shape

Continuum SM background and resonant processes with the same initial and final states can interfere with each other. In fact, in the presence of interference, signal and background do not have a physical meaning. Figure 7.18 shows the Feynman diagrams of the interfering continuum background and resonance signal produced via gluon-gluon fusion.

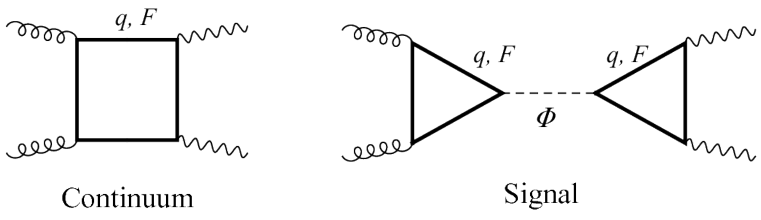


Figure 7.18: Representative Feynman diagrams of the interfering continuum background (left) and a scalar resonance signal (right) in the $gg \to \gamma\gamma$ process. In the EFT as presented in Section 7.3, the triangle loops in the signal process are replaced by effective vertices.

Here, we address the interference between the resonant process $gg \to X \to \gamma\gamma$ and the continuum $gg \to \gamma\gamma$ scattering process. The full $gg \to \gamma\gamma$ amplitude is a sum of resonance and continuum terms:

$$\mathcal{A}_{gg\to\gamma\gamma} = \mathcal{A}_{\text{res}} + \mathcal{A}_{\text{cont}}, \tag{7.43}$$

Writing the propagator explicitly, this becomes

$$\mathcal{A}_{gg\to\gamma\gamma} = \frac{a_{\text{res}}}{\hat{s} - m_X^2 + im_X\Gamma_X} + \mathcal{A}_{\text{cont}}, \tag{7.44}$$

and the squared amplitude is

$$|\mathcal{A}_{gg\to\gamma\gamma}|^2 = \left(\frac{a_{\text{res}}}{\hat{s} - m_X^2 + im_X\Gamma_X} + \mathcal{A}_{\text{cont}}\right)\left(\frac{a^*_{\text{res}}}{\hat{s} - m_X^2 - im_X\Gamma_X} + \mathcal{A}^*_{\text{cont}}\right), \tag{7.45}$$

$$|\mathcal{A}_{gg\to\gamma\gamma}|^2 = \frac{|a_{\text{res}}|^2 + a_{\text{res}}\mathcal{A}^*_{\text{cont}}(\hat{s} - m_X^2 - im_X\Gamma_X) + a^*_{\text{res}}\mathcal{A}_{\text{cont}}(\hat{s} - m_X^2 + im_X\Gamma_X)}{(\hat{s} - m_X^2)^2 + m_X^2\Gamma_X^2} + |\mathcal{A}_{\text{cont}}|^2, \tag{7.46}$$

$$|\mathcal{A}_{gg\to\gamma\gamma}|^2 = \frac{|a_{\text{res}}|^2 + (\hat{s} - m_X^2)(a_{\text{res}}\mathcal{A}^*_{\text{cont}} + a^*_{\text{res}}\mathcal{A}_{\text{cont}}) + im_X\Gamma_X(-a_{\text{res}}\mathcal{A}^*_{\text{cont}} + a^*_{\text{res}}\mathcal{A}_{\text{cont}})}{(\hat{s} - m_X^2)^2 + m_X^2\Gamma_X^2} + |\mathcal{A}_{\text{cont}}|^2. \tag{7.47}$$

Finally, the differential cross-section of $gg \to \gamma\gamma$ can be written as

$$d\sigma(gg \to \gamma\gamma) \propto \int d\hat{s}\mathcal{L}_{gg}\frac{1}{\hat{s}}\left(|\mathcal{A}_{\text{res}}|^2 + N^{\text{Re}}_{\text{int}} + N^{\text{Im}}_{\text{int}} + |\mathcal{A}_{\text{cont}}|^2\right), \tag{7.48}$$

where the real and imaginary part of the interference terms are

$$N^{\text{Re}}_{\text{int}} = \frac{2\mathbb{R}[a_{\text{res}}\mathcal{A}^*_{\text{cont}}](\hat{s} - m_X^2)}{(\hat{s} - m_X^2)^2 + m_X^2\Gamma_X^2}, \tag{7.49}$$

$$N_{\text{int}}^{\text{Im}} = \frac{2\mathbb{I}[a_{\text{res}}\mathcal{A}_{\text{cont}}^{*}]m_X\Gamma_X}{(\hat{s}-m_X^2)^2+m_X^2\Gamma_X^2}. \tag{7.50}$$

Since $N_{\text{int}}^{\text{Re}}$ is proportional to $(\hat{s}-m_X^2)$, it does not contribute to the total cross-section when one integrates over $\hat{s}$. However, it induces a peak-dip or dip-peak structure, distorting the resonance shape, while the imaginary part interference either enhances or reduces the signal rate [114].

In the analysis, the total cross-section is arbitrary since it is left floated in the fit. The distortion of the resonance shape is of interest: depending on the model assumptions being made on the resonance, the interference effect could be either small or could lead to significant distortion of the resonance shape. The size of the interference is expected to increase with the resonance width; however, it may not be negligible even in the narrow-width case.

Potential effects have been examined using a two-Higgs-doublet model (2HDM) in Ref. [115] and [116][8]. The interference effects of a purported 750 GeV signal under two benchmark models (2HDM and singlet scalar) are also studied in Ref. [114] and [117].

The interference effect under a well motivated model may be included in the analysis in the future. For the moment, interference is neglected and the width of the resonance being considered is restricted to less than 10% of the resonance mass.

7.5 Randall-Sundrum Graviton Line Shape

In the search for spin-2 high mass resonance, the Randall-Sundrum (RS) graviton model [34] is used as a benchmark. This entails the lightest Kaluza-Klein (KK) spin-2 graviton excitation (G^*) with a coupling $k/\overline{M}_{\text{Pl}}$, where $\overline{M}_{\text{Pl}}$ is the reduced Planck scale ($\overline{M}_{\text{Pl}} = M_{\text{Pl}}/\sqrt{8\pi}$) and k the curvature scale of the extra dimension. The RS graviton couples to the energy momentum tensor of the SM through a dimension-five operator.

PYTHIA8 is used to generate the Randall-Sundrum (RS) graviton samples at different value of the coupling $k/\overline{M}_{\text{Pl}}$, which changes the resonance width. The generation is performed at LO using the NNPDF23LO PDF set. Only the effect of the first KK excitation is simulated.

In the Randall-Sundrum model, the partial decay widths of a graviton G^* with mass m_{G^*} to fermions, gluons or photons are [118]

$$\Gamma(G^*\to f\bar{f}) = N_c\frac{\kappa^2\hat{s}^{3/2}}{320\pi}(1-4r_f)^{3/2}\left(1+\frac{8}{3}r_f\right), \tag{7.51}$$

$$\Gamma(G^*\to gg) = \frac{\kappa^2\hat{s}^{3/2}}{20\pi} \tag{7.52}$$

and

$$\Gamma(G^*\to\gamma\gamma) = \frac{\kappa^2\hat{s}^{3/2}}{160\pi}, \tag{7.53}$$

where f is a fermion with mass m_f and N_c the number of colors, $r_f = m_f^2/m_{G^*}^2$ and $\kappa \propto k/\overline{M}_{\text{Pl}}m_{G^*}$.

At leading-order, a graviton can be produced by gg or $q\bar{q}$ initial state. From Eq. 7.20, one can write the differential cross-section as

$$d\sigma \propto \int d\hat{s}\frac{1}{\hat{s}}\left[\mathcal{L}_{gg}(\hat{s})|a_{gg\to G^*}(\hat{s})a_{G^*\to\gamma\gamma}(\hat{s})|^2 + \mathcal{L}_{q\bar{q}}(\hat{s})|a_{q\bar{q}\to G^*}(\hat{s})a_{G^*\to\gamma\gamma}(\hat{s})|^2\right]\cdot f_{BW}, \tag{7.54}$$

where f_{BW} is the form used in PYTHIA:

[8] Both Type-I and Type-II 2HDM scenarios are explored in these studies.

$$f_{BW} = \frac{m_{G^*}\Gamma_{G^*}}{\pi} \frac{1}{(\hat{s} - m_{G^*}^2)^2 + (\hat{s}\Gamma_{G^*}/m_{G^*})^2}, \tag{7.55}$$

and the Breit-Wigner width Γ_{G^*} can be related to the coupling $k/\overline{M}_{\mathrm{Pl}}$ by

$$\Gamma_{G^*} \approx 1.44(k/\overline{M}_{\mathrm{Pl}})^2 m_{G^*}. \tag{7.56}$$

In Pythia, the cross-section of $ii \to G^* \to \gamma\gamma$ is computed as

$$\Gamma(G^* \to ii) \cdot f_{BW} \cdot \Gamma(G^* \to \gamma\gamma). \tag{7.57}$$

Associating the terms to Eq. 7.54, we obtain

$$|a_{ii\to G^*}(\hat{s})a_{G^*\to\gamma\gamma}(\hat{s})|^2 \propto \hat{s}^4, \tag{7.58}$$

and

$$|a_{q\bar{q}\to G^*}(\hat{s})a_{G^*\to\gamma\gamma}(\hat{s})|^2 \propto \hat{s}^4. \tag{7.59}$$

This has similar form to the squared matrix element of the scalar singlet model. Using the same arguments, the RS graviton line shape can finally be written as

$$\frac{d\sigma_{\mathrm{grav}}}{dm_{\gamma\gamma}} \propto (\mathcal{L}_{gg} + \alpha \cdot \mathcal{L}_{q\bar{q}}) \cdot f_{BW} \cdot m_{\gamma\gamma}^7, \tag{7.60}$$

where α is a constant that determines the contributions from the $q\bar{q}$-induced production relative to the gg-induced production. From the relative LO cross-section of $gg \to G^*$ and $q\bar{q} \to G^*$ processes in literature such as Ref. [118], the value of α is 2/3. However, by checking the cross-section of the two processes from the output of Pythia, a value close to 3/2 is obtained instead. Nevertheless, the line shape is barely affected by the choice of α, and the value of $\alpha = 3/2$ is chosen in the parameterisation of the line shape.

The simulated samples are generated with Pythia version 8.186. It was later found that this version contains a bug in the $q\bar{q} \to G^*$ cross-section that is subsequently fixed in version 8.212. In the $q\bar{q} \to G^*$ cross-section, an extra factor of $\hat{s}/m_{G^*}^2$ was included which resulted in a resonance shape with a much enhanced upper tail than it should.

7.6 Conclusion

The line shape of a large-width high mass resonance under several model assumptions as well as using different MC generator approaches has been presented in this chapter. For spin-0 resonances, the Higgs-like model turns out to be unsatisfactory due to the specific behaviour of the production as well as the decay processes when SM Higgs coupling of the resonance is assumed. The nature of new physics is unknown, hence in such a search for new high mass resonance, a *model-independent* approach is preferred. The scalar singlet EFT model is deemed to be more generic. In addition, when 2015 data showed an excess around 750 GeV, many theory papers attempting at explaining the origin of the excess were published and the scalar singlet model with an EFT approach is used in many of them. The line shape of a RS graviton is also presented, which is similar to the scalar singlet line shape except for the inclusion of $q\bar{q}$-initiated production mode.

In all cases, an analytical parameterisation was derived for the resonance line shape that will be used to model the signal $m_{\gamma\gamma}$ shape in the analysis presented in the next chapters. Excellent agreement is observed between the parameterisation and the true diphoton invariant mass distributions in the MC. This validates the use of the analytic expression to derive distributions for parameter points that lie between the simulated mass and width values.

Chapter 8

Search for Diphoton Resonances at $\sqrt{s} = 13$ TeV with $3.2\ \mathrm{fb}^{-1}$ Recorded in 2015

This chapter describes the analysis that searches for high mass diphoton resonances using $\sqrt{s} =$ 13 TeV data recorded in 2015. The search was first presented in December 2015 at the End Of the Year Event (EOYE) [119] at CERN, then updated at the *51st Rencontres de Moriond on Electroweak Interactions and Unified Theories* 2016 [120] and subsequently published in the Journal of High Energy Physics (JHEP) [69].

Experimentally, new high mass states decaying into two photons provide a clean signature in the diphoton invariant mass $m_{\gamma\gamma}$ thanks to the excellent invariant mass resolution and the smoothly falling background. Such a high mass state can be either spin-0 or spin-2 if it decays to two photons.[1] The search is motivated by several models of new physics. Spin-0 resonances are predicted by models with extended Higgs sectors [26–32]. Diphoton signals could also occur in the decay of heavy spin-2 graviton resonances in models with extra dimensions such as Randall-Sundrum (RS) model. In the analysis, two searches are performed, one targetting a generic spin-0 particle and the other for a spin-2 particle using the RS model as benchmark.

In the spin-0 search, five Higgs-like production modes are considered and the search is performed in such a way that it is applicable to either of these production processes. The search has a tighter kinematical requirement for the photon selection, taking advantage of the isotropic distribution of the decay products in the centre-of-mass frame of a heavy scalar particle. The spin-2 resonance search has a looser kinematic selection requirements to maintain the acceptance, since the decay products of a spin-2 particle tend to be in the more forward region, resulting in them having a lower transverse energy.

Details of the analysed dataset as well as the MC samples that are used to study the resonant signal and non-resonant background are presented in Section 8.1. The event selection is outlined in Section 8.2. The determination of the background $m_{\gamma\gamma}$ shape is described in Section 8.3. The modelling of the signal shape and the signal efficiencies are summarised in Section 8.4 and 8.5. The systematic uncertainties that enter into the analysis and the statistical model used to present the results are described in Section 8.6 and 8.7 respectively. Section 8.8 discusses the results of the search analysis. The re-analysis of 8 TeV data and its compatibility with the observation in 2015 data are reported in Section 8.9. Finally, Section 8.10 presents additional studies on categorisation.

As an active member of the analysis team, I have performed various studies and contributed to many areas of the analysis. My main contribution consisted in providing the signal modelling

[1] Spin-1 state decaying to two spin-1 photons is excluded due to Landau-Yang theorem [24, 25] and higher spins are disfavoured theoretically.

of large-width scalar resonance through the study of the signal shape explained in the previous chapter (Section 8.4.2 and 8.4.3). I also contributed to the computation of large-width fiducial correction factors for the spin-0 analysis (Section 8.5.2). I was also responsible for the statistical analysis on the spin-0 resonance search described in Section 8.8.1 and Section 8.8.3. In addition, I performed some checks on the validity of the narrow-width limit for larger widths and the validity of the asymptotic approximation in the p-value computation, as presented in Section 8.8.4 and Section 8.8.5. Section 8.9.1 was my personal contribution to the 8 TeV re-analysis, where I defined the large-width scalar resonance shape and computed the large-width fiducial correction factor. Finally, I also studied potential analysis improvements using categorisation, as summarised in Section 8.10.

This chapter describes both the spin-0 and the spin-2 analyses. Nonetheless, as I was involved mostly in the spin-0 search, more details are given to the spin-0 analysis in several parts of the chapter.

8.1 Data and Monte Carlo Samples

The analysis uses 3.2 fb^{-1} of pp collision data at $\sqrt{s}$ = 13 TeV collected by ATLAS in 2015. The data sample was recorded using the `HLT_g35_loose_g25_loose` trigger, which is a diphoton trigger with transverse energy E_T thresholds of 35 GeV and 25 GeV for the E_T-leading and subleading photon candidates respectively. This trigger was run unprescaled, and has a signal efficiency close to 99% for events fulfilling the final event selection. Only events taken during stable beam conditions and in which the detector was fully operational are considered. The standard data quality requirements are applied.

Simulated samples are used in the analysis for optimising the event selection, computing signal efficiencies, determining the diphoton spectra for signal and background processes, etc.

8.1.1 Scalar Signal Samples

The spin-0 analysis searches for a generic scalar resonance. In order to study the signal properties, a heavy Higgs (with SM-like couplings) is simulated. Five Higgs-like production modes: gluon-gluon fusion, vector boson fusion (VBF), associated production with a W boson (WH), associated production with a Z boson (ZH) and associated production with a top quark pair ($t\bar{t}H$) are considered.

Gluon-gluon fusion is assumed as baseline since it is the dominant production mode of the SM Higgs due to the much higher gluon-gluon luminosity compared to quark-antiquark. The gluon-gluon fusion heavy Higgs production is simulated with the PowHeg generator [105] interfaced to Pythia8 for showering and hadronisation. The Higgs-like signal samples produced via VBF is also simulated with PowHeg [121] interfaced to Pythia8. The WH, ZH and $t\bar{t}H$ samples are simulated with Pythia8 only. In all samples, only the signal is simulated and the interference with the continuum background is neglected.

The search covers signals with widths ranging from the 4 MeV SM Higgs width (so-called narrow-width approximation or NWA) up to $\Gamma_X/m_X = 10\%$. Signal samples with VBF, WH, ZH and $t\bar{t}H$ production modes are only generated with narrow width. Since the large-width signal shape is determined with the procedure described in Chapter 7, the large-width gluon-gluon fusion samples are mostly used for validation of the signal modelling techniques and in computing the efficiencies.

8.1.2 Graviton Signal Samples

The graviton signal model is simulated using Randall-Sundrum (RS) graviton production processes in pp collision at $\sqrt{s}$ = 13 TeV followed by a decay to two photons, implemented in

Pythia8. Only the effect of the first Kaluza-Klein excitation is simulated. Multiple samples were produced for different values of the couplings with the SM particles $k/\overline{M}_{\rm Pl}$ and mass of the graviton m_{G*}, ranging $0.01 < k/\overline{M}_{\rm Pl} < 0.25$ and $500 < m_{G*} < 5500$ GeV.

In practice, a high-statistics flat mass sample with 1.5 million events simulated is used since it can be reweighted to any mass and coupling values. Starting with the standard Pythia production code for a RS graviton sample of 5 TeV mass and $k/\overline{M}_{\rm Pl} = 0.1$, the resonant effect is undone by dividing the differential cross-section by the Breit-Wigner term used in Pythia (see Eq. 7.55). Then, the differential cross-section is divided by a mass-dependent factor comprising of the parton luminosity and mass-dependent squared matrix element effects, resulting in a flat mass spectrum.

Once the flat graviton mass spectrum has been generated, reweighting to any mass and coupling can be achieved by reversing the procedure on an event-by-event basis using the generator-level information. The reweighting uses the above Breit-Wigner term and the mass-dependent factor fitted on the residual mass dependence in simulation. This is quite similar to the line shape parameterisation derived in Section 7.5, which was only used in the subsequent analysis using 2016 data.

Reweighting the flat sample to a narrow width is ineffective due to the lack of statistics. Hence, for $k/\overline{M}_{\rm Pl} = 0.01$, single mass samples are used while reweighted flat samples are used instead for $k/\overline{M}_{\rm Pl} > 0.05$. The single mass samples at larger values of $k/\overline{M}_{\rm Pl}$ are mostly used for cross-checks.

8.1.3 Background Samples

Events containing two photons with associated jets are simulated using the Sherpa 2.1.1 [67] generator, requiring a photon transverse momentum above 20 GeV. Matrix elements are calculated with up to two partons at LO and merged with the Sherpa parton showering [122] using the ME+PSatLO prescription [123]. The CT10 PDF set is used in conjunction with dedicated parton shower tuning developed by the Sherpa authors. In order to maximise the available statistics over the entire mass range of interest, the simulation is separately performed in slices of diphoton invariant mass.

8.2 Event Selection

The photon candidates are required to be within the calorimeter acceptance of $|\eta| < 2.37$, excluding the transition region between the barrel and the end-cap calorimeters of $1.37 < |\eta| < 1.52$. In the excluded transition region, the photon identification capabilities and energy resolution are degraded.

In order to reject background from jets misidentified as photons, the selected photon candidates have to pass the tight photon identification criteria based on shower shapes in the calorimeter. Furthermore, the photon candidates are required to be isolated. The selection requirement on the calorimeter isolation variable is defined by $E_T^{\rm iso} < 0.022E_T + 2.45$ GeV and on the track isolation variable by $p_T^{\rm iso} < 0.05E_T$. The primary vertex corresponding to the pp collision that produced the diphoton candidate is identified using the algorithm described in Section 6.2.2.

In the selection used to search for a spin-0 resonance, a tight kinematical selection optimised on simulated signal and background samples is applied. The transverse energies of the photon with the highest E_T and second-highest E_T are required to be $E_T > 0.4m_{\gamma\gamma}$ and $E_T > 0.3m_{\gamma\gamma}$ respectively. This selection improves the expected sensitivity by over 20% for masses above 600 GeV. The cut on $E_T/m_{\gamma\gamma}$ selects events in which the photons are preferentially emitted in the central region of the detector. In the spin-2 selection, the transverse energy of each photon is required to be greater than 55 GeV.

Finally, the diphoton invariant mass is required to be larger than 150 GeV for the spin-0 selection and $m_{\gamma\gamma} > 200$ GeV for the spin-2 selection.

8.3 Background Modelling

The diphoton background can be divided into two categories: the irreducible background which is due to SM production of real photon pairs, and the reducible background which includes event in which one or both of the reconstructed photon candidates are from jets misreconstructed as photons. The reducible background can be separated into three components: γj, $j\gamma$ and jj, referring to whether the leading or subleading photon candidates is a fake. Background sources from Drell-Yan, $W\gamma$ or $Z\gamma$ production, with either one or two isolated electrons misidentified as photons, are negligible.

The shape of the mass distribution for the irreducible diphoton background is obtained from the simulation. The shape of the mass distribution for the reducible photon+jet and dijet backgrounds is estimated from data control samples selected with one or two of the photons failing a subset of the tight identification criteria. The final background prediction is obtained by summing the irreducible component and the smoothed estimate of the photon+jet and dijet backgrounds, normalised to their yields estimated in the data. This prediction of the total background shape can be used directly in the form of a background template, or to validate the background modelling.

8.3.1 Irreducible Background

The shape of the irreducible background is obtained from the LO Sherpa samples by applying the analysis selection. NLO prediction of the diphoton invariant mass spectrum can be estimated with the parton-level generator Diphox [124]. Diphox version 1.3.2 is used with the CTEQ66M PDF set. The fragmentation, factorisation and renormalisation scales are set to the mass of the diphoton system. The generator-level photon isolation cut is required in the calculation: it is adjusted to mimic the effect of the reconstruction-level photon isolation cut. The true diphoton invariant mass generated with Sherpa is then reweighted to match that in Diphox. The NLO prediction of reconstructed diphoton invariant mass of the irreducible background can then be obtained from the reweighted sample.

8.3.2 Reducible Background

The shape of the mass distribution for the reducible backgrounds is estimated from background-enriched data control samples. In all control samples, the two photon candidates are required to pass the same isolation cut as for the signal selection, as well as the same kinematical requirements. The control samples are selected with one or two of the photons failing the tight identification criteria and fulfilling a looser set of requirements, called the LoosePrime selection. The LoosePrime cuts are defined by inverting different combinations of the photon shower shape requirements as shown in Table 8.1.

The first set of control samples, named Tight-AntiTight, contain the events where the leading photon passes the tight photon identification while the subleading photon is required to fail the tight photon identification but to pass the LoosePrime requirement. The Tight-AntiTight samples are enriched in $\gamma +$jet events. Similarly, the AntiTight-Tight and AntiTight-AntiTight data control samples are enriched in jet $+\ \gamma$ and jet $+$ jet respectively.

The limited number of data events does not directly allow a precise estimate of the mass distribution for masses above 500 GeV. Thus, the background shapes are smoothened by fitting a function to the mass spectra in the control regions.

Several sets of control samples can be defined based on the chosen LoosePrime selection. One could also use the Loose selection instead of LoosePrime. The control region defined

Definition	Relaxed cuts
LoosePrime2	w_{s3}, F_{side}
LoosePrime3	w_{s3}, F_{side}, ΔE
LoosePrime4	w_{s3}, F_{side}, ΔE, E_{ratio}
LoosePrime5	w_{s3}, F_{side}, ΔE, E_{ratio}, $w_{s,\text{tot}}$

Table 8.1: The definitions of the LoosePrime control regions. Here, LoosePrime$n$ means that the tight identification criteria is applied with the exception of the n shower shape variables. The variables are defined in Section 3.3.

by the Loose selection has the highest statistics and high fraction of fake photons, but may not be representative of the fake photons in the signal region. The control region refined by LoosePrime2 selection is the closest to the signal region but contains a large fraction of true photons. The nominal control region is chosen by finding the right balance between the signal leakage in the control samples, the need for reasonable statistics in the control samples as well as possible correlations between the shower shape variables and the mass distribution. The LoosePrime4 selection is chosen for the nominal region of the background shape estimation of the $\gamma+\text{jet}$ and $\text{jet}+\gamma$ components, while the Loose selection is used for the nominal region of the background shape estimation of the $\text{jet}+\text{jet}$ component. All other LoosePrime variations are used to evaluate the systematic uncertainty on the nominal background shape for the $\gamma+\text{jet}$ and $\text{jet}+\gamma$ components, and only LoosePrime4 and LoosePrime5 variations are used for the $\text{jet}+\text{jet}$ component systematic uncertainty in order to have reasonable statistics in the control samples.

8.3.3 Data Sample Composition

The relative contribution of the true $\gamma\gamma$, γj, $j\gamma$ and jj components in the selected events is measured in data. The diphoton purity (i.e. $\gamma\gamma$ fraction) is expected to be high with the chosen analysis selections. The relative fractions of these background components are used to combine the irreducible and reducible background as defined in the previous sections to study the overall background shape.

Three techniques are used: *two-times-two-dimensional sideband* (2×2DSB) method, *matrix method* and a fit to the two-dimensional distribution of the calorimeter isolation of the two photon candidates (2D isolation fit). Each of these methods makes use of the discriminating power between prompt and fake photons provided by the photon isolation energy and/or shower shape variables in the electromagnetic calorimeter.

The methods are based on control regions, where one or two reconstructed photons fail the *tight* identification and/or the isolation requirements, and the signal region, where the two photon candidates pass both the tight identification and the isolation requirements. In the non-tight control region, the photon that fails the tight identification is required to satisfy the LoosePrime selection defined in Section 8.3.2. Figure 8.1 shows a sketch of the signal and control region, for one candidate photon in terms of its identification and isolation requirements.

The common strategy of the methods to estimate the hadronic background is to extract the amounts of $\gamma\gamma$, γj, $j\gamma$ and jj events from the observed yields (matrix method and 2×2DSB methods) or isolation distributions (2D isolation fit) of the events that populate the signal region and the control regions. The 2×2DSB and matrix methods are based on event counting while 2D isolation fit makes use of the shape of the photon and jet isolation distributions.

The differential purities as a function of the diphoton invariant mass are only measured with the 2×2DSB and matrix methods, while all methods are used to measure the inclusive purities, integrated over the diphoton invariant mass distribution. A good agreement between the results

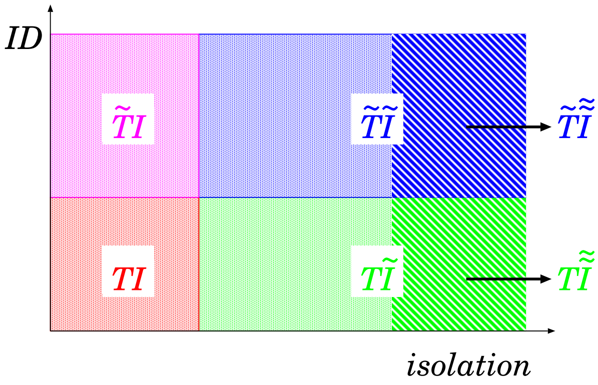
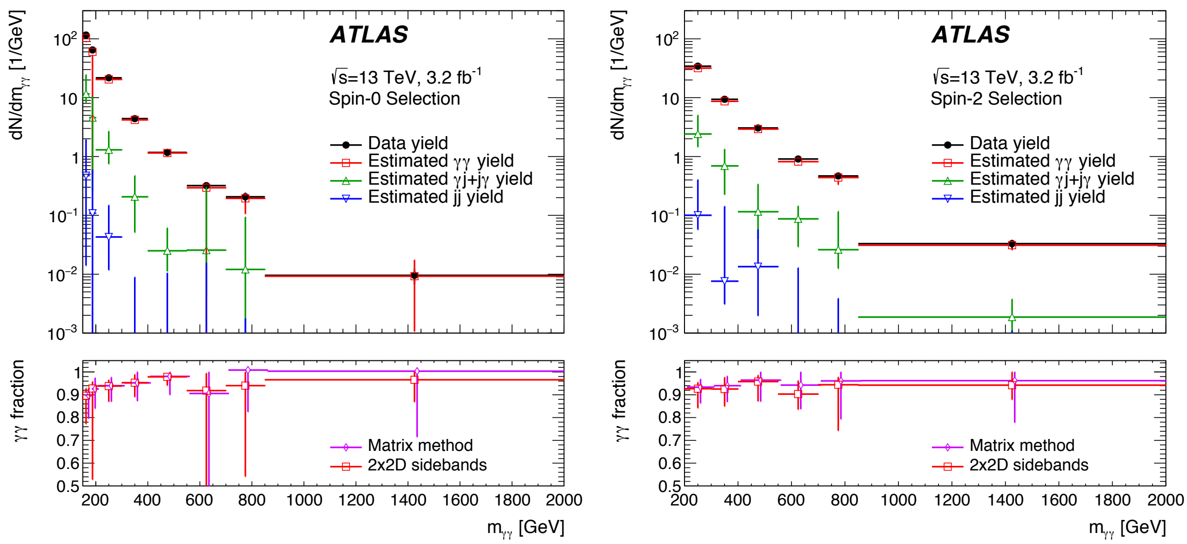


Figure 8.1: Sketch of the signal region (TI) and all the control regions.

of the different methods is observed.

The inclusive purity is $(93^{+3}_{-8})\%$ for the spin-0 selection and $(94^{+3}_{-7})\%$ for the spin-2 selection. The inclusive purity is slightly higher for the spin-2 selection (but within uncertainties) since it is evaluated over $m_{\gamma\gamma} > 200$ GeV for the spin-2 selection and over $m_{\gamma\gamma} > 150$ GeV for the spin-0 selection, and the purity in $150 < m_{\gamma\gamma} < 200$ GeV is smaller than for $m_{\gamma\gamma}$ above 200 GeV.

Figure 8.2 shows the data decomposition and diphoton purity as a function of the diphoton invariant mass for the spin-0 and spin-2 selections.



(a) (b)

Figure 8.2: Diphoton invariant mass distributions of the data for the spin-0 (left) and spin-2 (right) selections and their decomposition into contributions from true diphoton, photon+jet, jet+photon and dijet events. The bottom panels show the purity of the diphoton events as determined from 2×2DSB and matrix method. The total uncertainties are shown, including statistical and systematic components.

8.3.4 Background Modelling Approaches

Two different approaches are used to estimate the background contribution to the $m_{\gamma\gamma}$ distribution in the spin-0 and spin-2 analyses. In both cases, a background template is built by summing the irreducible component and the smoothed estimate of the reducible backgrounds, normalised to their yields estimated in the data.

In the search for scalar resonances, a data-driven estimate is used to model the background. The total background contribution is modelled by a smooth functional form fit to the $m_{\gamma\gamma}$ distributions in data. The functional form must be capable to describe the background over the largest possible mass range and is validated on the background template up to 3 TeV.

The *spurious signal* method is used to both validate the chosen function and also estimate the uncertainties in the background modelling due to the choice of a particular function. Spurious signal refers to the fake fitted signal induced by the systematic differences between the background parameterisation and the background template. To evaluate the amount of spurious signals, a candidate function is used to perform a signal+background fit for a given mass hypothesis m_X on the background template.[2] The number of fitted signal yield for a given mass, $N_{\text{spurious}}(m_X)$, is studied as a function of m_X, varying the signal mass hypothesis in the range of interest in steps of 1 GeV.

In order for a function to be valid, it needs to pass the spurious signal test which requires that the function does not generate $|N_{\text{spurious}}(m_X)|$ larger than 30% of the uncertainty of the fitted signal yield over the full mass range the function is validated for. The number of fitted spurious signal events generally increases with the signal width and the function needs to be valid up to widths of 10% of m_X as defined in the scalar search range. Among the various validated functions, the one with the smallest number of free parameters is chosen as the baseline.

Functions of the form

$$f_{k;d}(x; b, \{a_k\}) = (1 - x^d)^b x^{\sum_{j=0}^{k} a_j \log(x)^j} \tag{8.1}$$

where $x = \frac{m_{\gamma\gamma}}{\sqrt{s}}$ are tested. Finally, a two-parameter function of this form with $d = 1/3$ and $k = 0$ is chosen for the spin-0 search, as given below:

$$f(x; b, a) = (1 - x^{1/3})^b x^a, \tag{8.2}$$

where a and b are the two parameters of the function.

Figure 8.3 shows the number of fitted spurious signal yield N_{spurious} as a function of the tested m_X for several width hypotheses using the above function. The spurious signals used as the uncertainty of the background modelling is obtained from parameterising the distribution of $|N_{\text{spurious}}(m_X)|$ with a conservative envelop for each width hypothesis.

The spin-2 search aims to reach very high masses, where the small number of data events does not constrain effectively the shape of the background invariant mass distribution. The shape of the invariant mass distribution of the total background is thus predicted from a template built from irreducible and reducible background components as described previously. The fractions from the 2D isolation fit method are used to combine the irreducible and reducible background templates. The NLO correction to the irreducible background shape obtained from DIPHOX is applied. The total background prediction is normalised to match the number of data events in a low-mass control region ($200 < m_{\gamma\gamma} < 500$ GeV) and the search range is defined as $m_{\gamma\gamma} > 500$ GeV.[3]

In the background template for the spin-2 search, the uncertainties in the $m_{\gamma\gamma}$ distribution result from the uncertainties in both the shape and the relative normalisation of each component. Sources of systematic uncertainties include the shape of the reducible background, the relative normalisation of the reducible and irreducible backgrounds, the impact of the parton-level isolation requirement in DIPHOX and the effect of the uncertainties in the scales and PDF in the DIPHOX computation. This is discussed in Section 8.6.3.

[2] Details on the signal modelling are given in Section 8.4.

[3] The control region is chosen in the mass region where previous experiments have excluded the presence of a RS graviton signal.

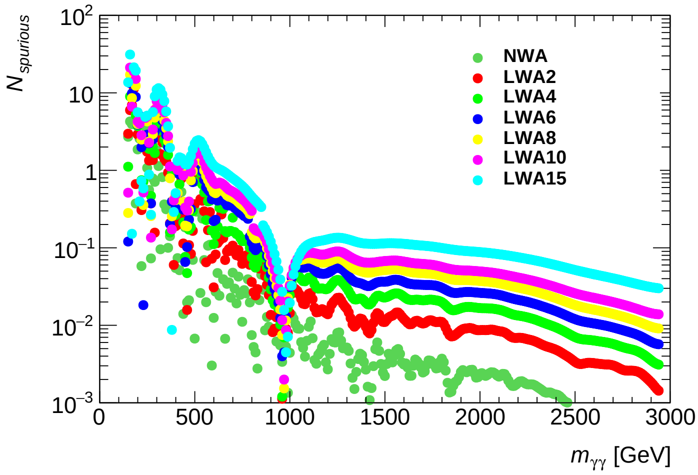


Figure 8.3: Number of fitted spurious signal yield N_{spurious} as a function of tested m_X, shown for NWA (dark green) and $\Gamma_X/m_X = 2$ (red), 4 (green), 6 (blue), 8 (yellow), 10 (violet) and 15% (cyan). The two-parameter function $f(x;b,a) = (1-x^{1/3})^b x^a$ is used and $N_{\text{spurious}}(m_X)$ is required to be lower than 30% of the uncertainty of the fitted signal yield over the full mass and width range.

8.4 Signal Modelling

Besides the modelling of the background $m_{\gamma\gamma}$ shape, the modelling of the signal shape is also of importance. This section describes the modelling of the narrow-width and large-width signal shape.

8.4.1 Narrow-Width Signal Modelling

A double-sided Crystal Ball (DSCB) function is used to parameterise the reconstructed NWA signal shape of a spin-0 or a spin-2 resonance after their respective event selection has been applied. The function is found to be able to describe well the NWA signal shape including the non-Gaussian and asymmetric low and high mass tails. The DSCB function is defined as

$$N \cdot \begin{cases} e^{-t^2/2} & \text{if } -\alpha_{\text{low}} \leq t \leq \alpha_{\text{high}} \\ \dfrac{e^{-0.5\alpha_{\text{low}}^2}}{\left[\frac{\alpha_{\text{low}}}{n_{\text{low}}}\left(\frac{n_{\text{low}}}{\alpha_{\text{low}}} - \alpha_{\text{low}} - t\right)\right]^{n_{\text{low}}}} & \text{if } t < -\alpha_{\text{low}} \\ \dfrac{e^{-0.5\alpha_{\text{high}}^2}}{\left[\frac{\alpha_{\text{high}}}{n_{\text{high}}}\left(\frac{n_{\text{high}}}{\alpha_{\text{high}}} - \alpha_{\text{high}} + t\right)\right]^{n_{\text{high}}}} & \text{if } t > \alpha_{\text{high}}, \end{cases} \tag{8.3}$$

where $t = (m_{\gamma\gamma} - \mu_{CB})/\sigma_{CB}$ with μ_{CB} being the mean of the Gaussian distribution and σ_{CB} represents the width of the Gaussian part of the function; N is a normalisation parameter; α_{low} (α_{high}) parameterises the mass value where the invariant mass resolution distribution becomes a power-law function on the low (high) mass side in units of t; and n_{low} (n_{high}) is the exponent of this power law. The parameter $\Delta m_X = \mu_{CB} - m_X$ is defined as the difference between the mean of the Gaussian and the signal mass value being generated. An illustrative drawing of the double-sided Crystal Ball function is provided in Figure 8.4.

The DSCB parameters are expressed as analytical functions of m_X. The parameterisation is determined in two steps. First, the DSCB shape is fitted to the signal mass distributions for each mass point in the available MC samples (*single mass fits*) yielding a set of DSCB parameters at each mass point. The evolution of these parameters with m_X is then parameterised with either linear or quadratic functions of m_{nX} ($m_{nX} = \frac{m_X - 100\text{ GeV}}{100\text{ GeV}}$).

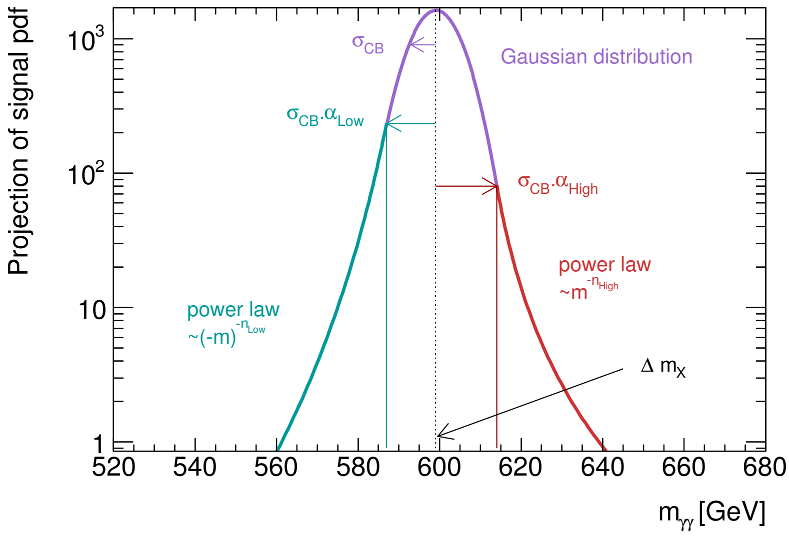


Figure 8.4: Example of DSCB function and illustrative description of its parameters for a signal mass of $m_X = 600$ GeV.

In the second step, a simultaneous fit to all the mass points is performed and the coefficients of the functions describing the evolution of DSCB parameters are obtained. This procedure (*multiple mass fit*) yields a more reliable determination of the coefficients. The *single mass fits* method provides a check on the fitted coefficients as the two methods should give similar results.

Figure 8.5 shows the resulting parameterisation of the DSCB parameters from the multiple mass fit (red line) for the spin-0 signal model. For comparison, the parameters from the fit to the individual mass points are also shown in blue points and the fit of those points is shown with the blue dashed line. Since the parameters are highly correlated, the values of n_{low} and n_{high} values are fixed as constant, using the values first obtained in the *single mass fits*. Excellent agreement between the parameterisation from *multiple mass fit* and *single mass fits* is observed.

8.4.2 Convolution with the Detector Resolution

To obtain the reconstructed mass shape of a large-width signal, the theoretical line shape described in Chapter 7 is convolved with the detector resolution function. The detector resolution is described by the NWA mass shape presented above. This approach allows the determination of the shape of any large-width signal, not restricted to the mass and width values where MC simulations are available.

Fully simulated MC samples are used to test if the convolution approach is able to describe the reconstructed mass distribution of a large-width signal. The DSCB function representing a narrow-width signal of a certain mass is convolved with the true mass distribution of a large-width signal of the same mass, and then compared to the reconstructed mass distribution of the large-width signal. For all the mass and width values for which MC samples are available, the convolution yields a function that describes well the reconstructed mass distribution. Figures 8.6 and 8.7 show the comparison of the reconstructed mass distributions with the outcome of the convolution for a resonance with mass of 400 and 1500 GeV and width of 5 and 10% of m_X.

The detector resolution varies as a function of the photon energy and consequently, the diphoton mass. Nevertheless, taking the detector resolution at the value of the resonance mass in the convolution is sufficient to provide a good description of the reconstructed mass distribution.

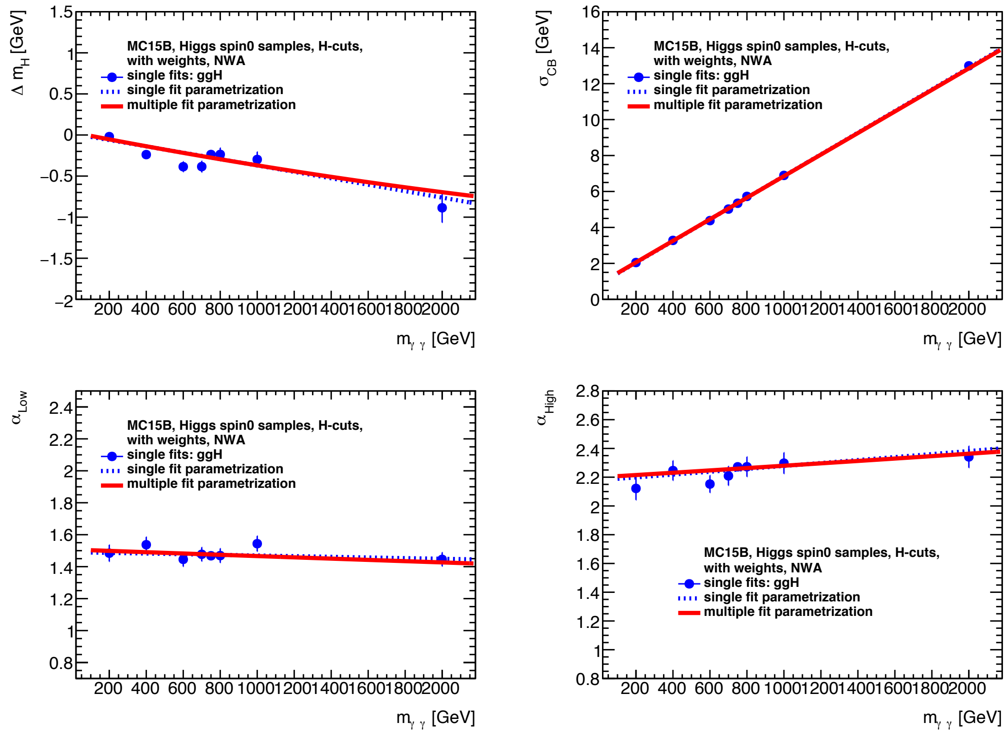
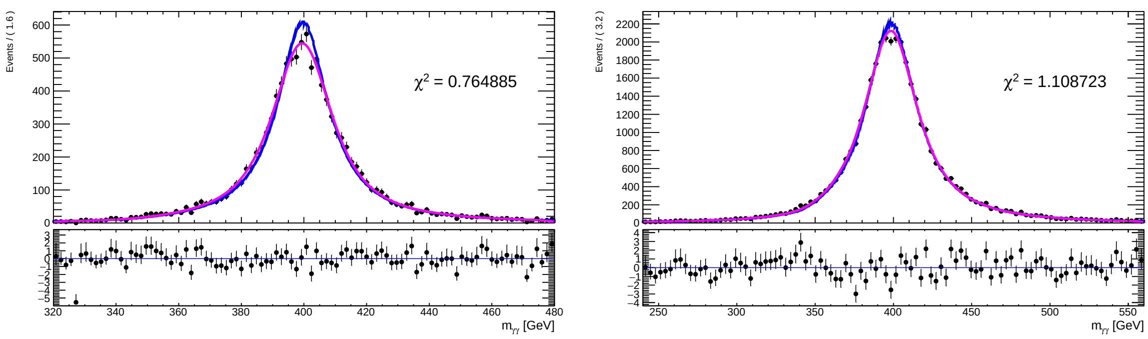


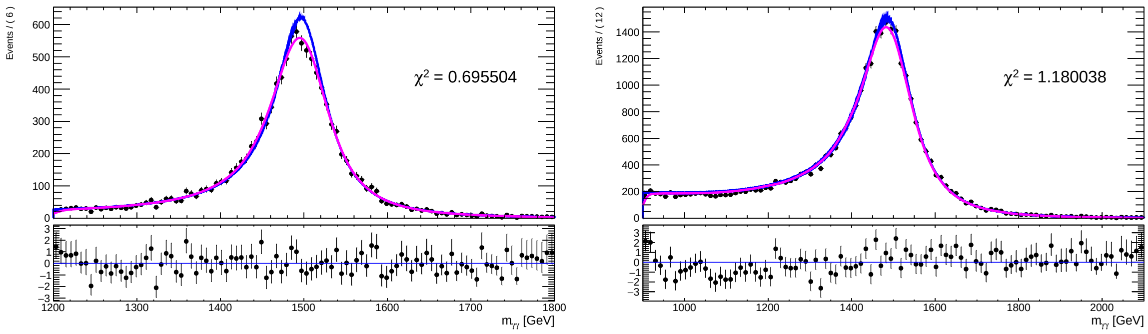
Figure 8.5: Comparison of the multiple mass fit parameterisation (red line) to the parameters from the single mass point fits (blue points) and to the fit of the single mass point parameters (blue dashed line) using the POWHEG spin-0 signal model. Four DSCB parameters are shown as a function of m_X: $\Delta m_X = \mu_{CB} - m_X$ (top left), σ_{CB} (top right) , α_{low} (bottom left) and α_{high} (bottom right).



(a) (b)

Figure 8.6: Reconstructed $m_{\gamma\gamma}$ distributions after selection of the resonance with $m_X = 400$ GeV and width of (a) 5% and (b) 10% of m_X are shown. In each plot, the blue line represents the theoretical line shape given by $\mathcal{B}_{gg} \cdot \mathcal{L}_{gg} \cdot BW \cdot m_{\gamma\gamma}$ while the magenta line represents the outcome of the convolution and the black markers show the reconstructed invariant mass distribution.

The convolution approach provides an elegant and simple signal modelling which is used to fit the data and derive the statistical results. It enables the modelling of the resonance shape of all resonance mass and width within the search range without the use of fully simulated samples.



(a) (b)

Figure 8.7: Reconstructed $m_{\gamma\gamma}$ distributions after selection of the resonance with $m_X = 1500$ GeV and width of (a) 5% and (b) 10% of m_X are shown. In each plot, the blue line represents the theoretical line shape given by $\mathcal{B}_{gg} \cdot \mathcal{L}_{gg} \cdot BW \cdot m_{\gamma\gamma}$ while the magenta line represents the outcome of the convolution and the black markers show the reconstructed invariant mass distribution.

8.4.3 Large-Width Signal Modelling

The convolution of the detector resolution with the theoretical line shape can be used directly in the statistical framework. This is done for the spin-2 signal modelling.

For the spin-0 case however, as discussed in Section 7.2.3, the resonance line shape implemented in PowHeg exhibits a secondary bump at around 400 GeV in the invariant mass distribution. Such features as well as the exact line shape in the off-shell region are highly model-dependent. Since most models predict very similar shapes close to the resonance peak (see Figure 7.9), the analysis is restricted to the resonance peak region. Thus, a DSCB function is used to model the large-width PowHeg signal shape in a similar way as done in the NWA case.[4]

Similarly to the NWA signal modelling where the DSCB function is fitted to the reconstructed mass distributions of the MC samples, the reconstructed mass distributions of the large-width signals are needed. The limited number of large-width MC samples does not allow a reliable signal modelling across mass and width ranges. Instead, *pseudo-samples* of 50,000 events are generated from the convolution of the theoretical line shape in Eq. 7.30 with the DSCB function describing the detector resolution for various values of mass and width.

The DSCB function is fitted to the mass distributions in the pseudo-samples. The fit is restricted to the peak region so that the secondary bump is not taken into account and the signal shape does not extend far to the off-shell region. m_X-dependent parameterisations of the DSCB parameters are first obtained at fixed Γ_X hypothesis following the method outlined in Section 8.4.1. An example of the paramaterisation is shown for $\Gamma_X/m_X = 6\%$ in Figure 8.8. Next, the value of the DSCB coefficients is interpolated between neighbouring values of Γ_X between 0 and 10% to obtain a continuous parameterisation in Γ_X. Figure 8.9 shows the evolution of the DSCB parameter as a function of the resonance width for a fixed resonance mass $m_X = 750$ GeV. The DSCB parameterisation obtained this way is checked on the available MC samples and a good description of the reconstructed mass is found.

[4] One could also cut the mass shape outside the resonant region with a top-hat function, but this is more difficult to implement. Using a DSCB function is easier as it is already used for NWA signal modelling and the function falls steeply away from the peak, effectively removing the off-shell events.

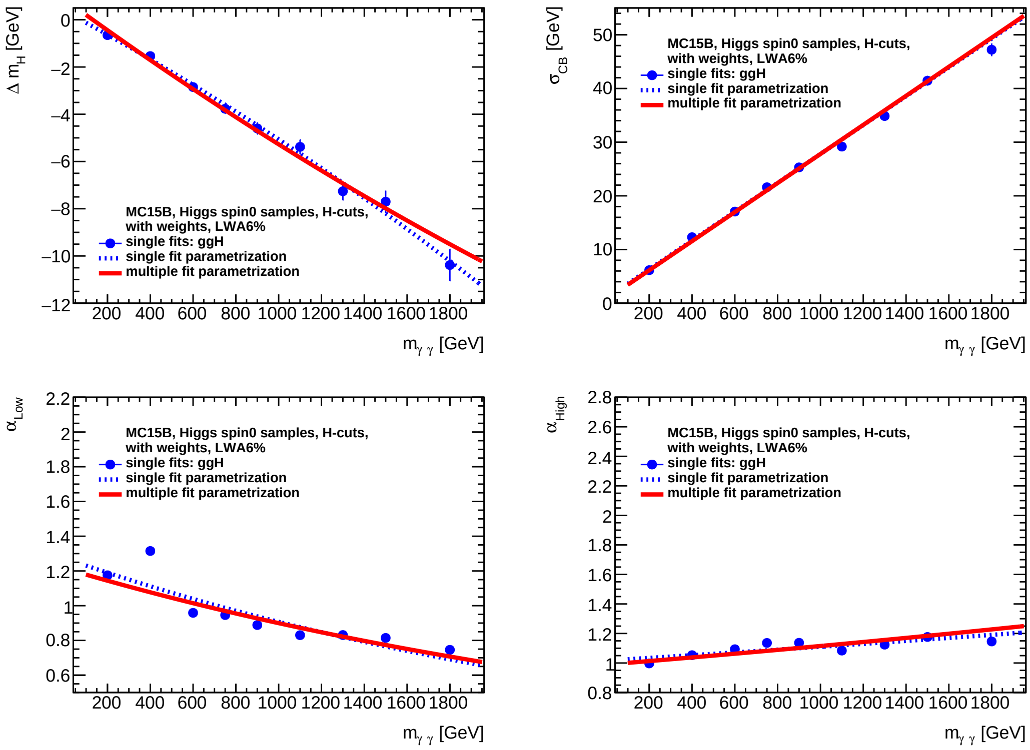
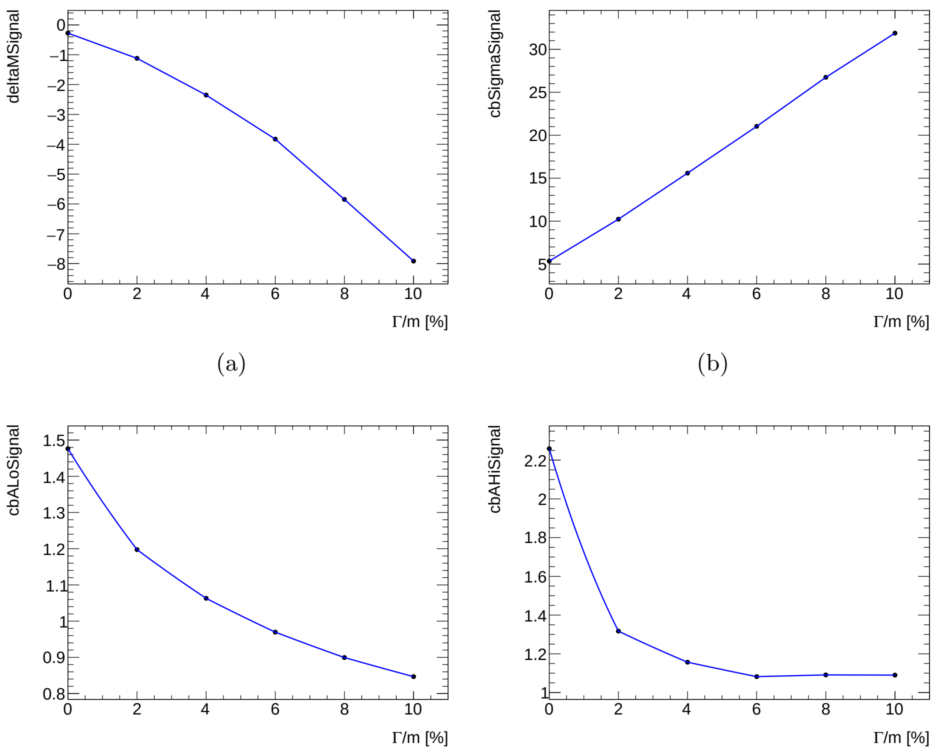


Figure 8.8: Comparison of the multiple mass fit parameterisation (red line) to the parameters from the single mass point fits (blue points) and to the fit of the single mass point parameters (blue dashed line) using the pseudo-samples generated with the parameterisation of PowHeg line shape (Eq. 7.30) with width $\Gamma_X/m_X = 6\%$. Four DSCB parameters are shown as a function of m_X: $\Delta m_X = \mu_{CB} - m_X$ (top left), σ_{CB} (top right) , $\alpha_{\rm low}$ (bottom left) and $\alpha_{\rm high}$ (bottom right).



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

Figure 8.9: Values of the DSCB parameters describing the large-width scalar signal shape as a function of Γ_X/m_X (black points) for the $m_X = 750$ GeV hypothesis. The blue line shows the interpolation between width values. Four DSCB parameters are shown: $\Delta m_X = \mu_{CB} - m_X$ (top left), σ_{CB} (top right) , α_{low} (bottom left) and α_{high} (bottom right).

8.5 Signal Efficiency

In the spin-0 analysis, the search is performed in a model-independent way. In particular, the measured signal cross-section must be reported in a way that is applicable to a range of production processes. For this purpose, the result is provided as a fiducial cross-section, measured within a fiducial volume chosen to minimise the model-dependence of the result. The spin-2 analysis however uses the Randall-Sundrum model as benchmark, and hence the measured number of signal events is corrected to the total acceptance using a simple efficiency correction predicted by the Monte Carlo samples.

Section 8.5.1 outlines the fiducial volume and the correction factor used to extract the fiducial cross-section for the spin-0 analysis. Special treatment of a large-width resonance while using the PowHeg signal model is presented in Section 8.5.2. The signal efficiency of graviton search is presented in Section 8.5.3.

8.5.1 Fiducial Correction Factor for Spin-0 Analysis

To extract the fiducial cross-section, a correction factor is used and is defined as

$$C_X = \frac{N_{\text{selection}}}{N_{\text{fiducial}}}, \tag{8.4}$$

where $N_{\text{selection}}$ is the number of reconstructed signal events passing all the analysis cuts, and N_{fiducial} is the number of signal events at the particle-level generated within the fiducial volume.

The fiducial acceptance defined at truth level is chosen to closely match the experimental acceptance of the measurement in order to minimise the model-dependence of the result. The fiducial volume therefore includes kinematic cuts applied on the truth photon variables so as to mimic the ones used at the reconstruction level. Both photons should be within $|\eta| < 2.37$, although the transition region of $1.37 < |\eta| < 1.52$ is not removed. The transverse energies for the leading and subleading photons are required to satisfy the same mass-dependent cuts as the one used for the reconstructed photons: $\frac{E_T^{\gamma 1}}{m_{\gamma\gamma}} > 0.4$ and $\frac{E_T^{\gamma 2}}{m_{\gamma\gamma}} > 0.3$. In addition, a particle isolation cut is also applied on the photons to reduce the dependence on the production process. The cut is defined by comparing the C_X factors extracted using simulations of other production processes. The particle isolation requirement is $E_T^{\text{iso}}(\Delta R = 0.4) < 0.05 E_T^{\gamma} + 6$ GeV, where the particle isolation, E_T^{iso} is computed using all particles with lifetime greater than 10 ps at generator level in a cone of $\Delta R = 0.4$ around the photon direction.

The C_X factor is calculated from Eq. 8.4 on the available MC samples. The pile-up distributions in the simulated samples are reweighted to match the pile-up condition in data. Figure 8.10 shows the C_X factor as a function of m_X for narrow-width gluon-gluon fusion signal samples. The mass-dependence of C_X factors is fitted using a functional form to ensure smooth interpolation between mass points.

8.5.2 PowHeg Large-Width C_X factor

The resonance line shape implemented in PowHeg shows a secondary bump at around mass of 400 GeV. Although the bump is not included in the statistical modelling due to the use of DSCB to model signal shape, the MC samples we use to extract C_X factors from are not exempted from this effect.

$N_{\text{selection}}$ is given by the fitted number of signal events from a signal+background fit of a sample combining an Asimov background component (see Section 5.6) with the signal events from MC sample at given mass and width. The signal events are weighted so that the total number of signal events is twice the expected upper limit on the event yield. A signal+background

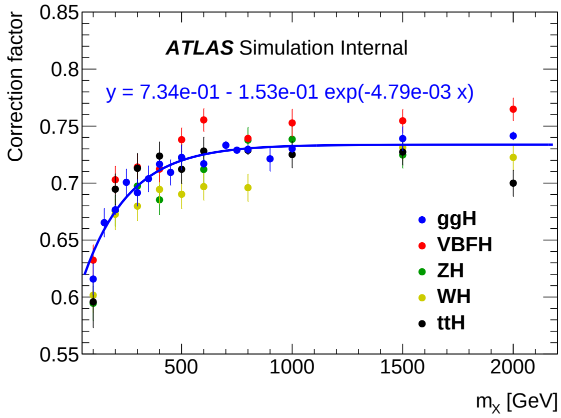


Figure 8.10: C_X factors computed using the PowHeg gluon-gluon fusion samples (blue) as a function of m_X for the 2015 dataset. The C_X factors from other production modes are also shown: VBF (red), ZH (green), WH (yellow) and ttH (black). The function shown in blue line is fitted on C_X factors computed from gluon-gluon fusion samples.

fit is performed since fitting the signal MC distribution with only the signal pdf will yield the same normalisation as the sample.

Figure 8.11 shows the mass distributions from the PowHeg sample for a Higgs-like resonance with $m_X = 1500$ GeV and $\Gamma_X/m_X = 10\%$, where a bump at low mass is visible, compared to the signal parameterisation which is shown in red in the same figure. Only the number of events fitted using this parameterisation should be considered in the $N_{\text{selection}}$ expression in the numerator of the C_X.

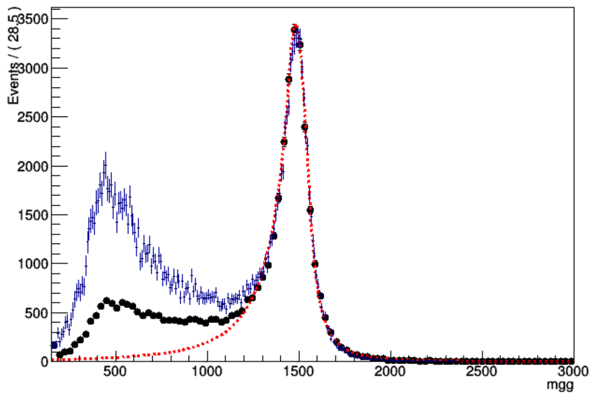


Figure 8.11: The mass distribution of the PowHeg MC sample simulated with $m_X = 1500$ GeV and $\Gamma_X/m_X = 10\%$ is shown in black markers. The signal model at the same mass and width is shown in red line. The mass distribution obtained from re-weighting the MC to an alternative propagator in shown in blue.

Since the fiducial acceptance should match the reconstruction level cuts, the large-width C_X factor using PowHeg model is computed with an additional fiducial cut on top of those defined previously in Section 8.5.1. The fiducial mass volume is chosen by comparing the C_X factor obtained in the standard PowHeg case and with an alternative line shape from the propagator given by Eq. 7.23 with the top quark mass set to 1 TeV in the partial width of $X \to \gamma\gamma$ (shown in blue in the same figure). The $\pm 2\Gamma_X$ window is found to lead to C_X factors that are similar in these two cases implying that a fiducial volume of $m_X \pm 2\Gamma_X$ is rather model-independent. As such, the definition of the C_X factor for the large-width scalar model goes one step further than for the narrow-width model, by defining both the numerator and denominator in terms of events in the main resonance peak rather than the total number of events.

The value of the large-width C_X factor at a given mass is larger than the narrow-width one since the fiducial volume is now smaller due to the line-shape truncation cut. At the reconstruction level, $N_{\text{selection}}$ is computed from the signal parameterisation which has non-zero tails whereas N_{fiducial} includes only events at the truth level inside a truncated mass range. In other words, the fiducial mass acceptance is tighter than the reconstruction level selection. The efficiency of the truncation cut applied to NWA MC samples (i.e. with $\Gamma_X = 4$ MeV) is found to be approximately 85% at $m_X = 800$ GeV and 84% at $m_X = 1500$ GeV, which is consistent with the ratio between the NWA and the large-width C_X factors at these masses.

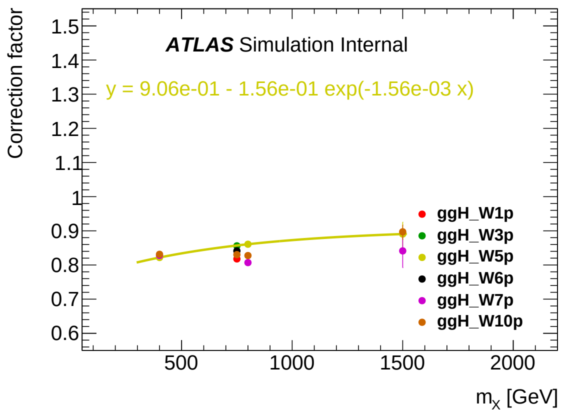


Figure 8.12: C_X factors computed using the PowHeg gluon-gluon fusion large-width samples as a function of m_X for the 2015 dataset for $\Gamma_X/m_X = 1$ (red), 3 (green), 5 (yellow), 6 (black), 7 (violet) and 10% (brown). The fiducial volume includes a mass window of $m_X \pm 2\Gamma_X$. The function shown in yellow line is fitted on C_X factors computed from $\Gamma_X/m_X = 5\%$ samples.

For the large-width case, only gluon-gluon fusion C_X factors are computed. The variation of the C_X factors for different production modes is assumed to be the same as that of the NWA case.

8.5.3 Signal Efficiency for Spin-2 Analysis

The graviton search is performed in the context of the Randall-Sundrum benchmark model. To express the results in terms of the total cross-section, the fitted number of signal events in data is corrected by the total acceptance efficiency:

$$\epsilon_{G*} = \frac{N_{\text{passing}}}{N_{\text{total}}} \tag{8.5}$$

where N_{total} is the total number of events in simulation and N_{passing} counts the number of reconstructed MC diphoton events passing all selections. The efficiency is calculated for different $k/\overline{M}_{\text{Pl}}$ hypothesis, as a function of m_{G*}. The efficiency is computed using single mass point MC samples for the case of narrow width ($k/\overline{M}_{\text{Pl}} = 0.01$) and reweighted flat samples for larger coupling values. The simulated samples have pile-up distributions reweighted to the pile-up condition in the data.

8.6 Systematic Uncertainties

In the following, the relevant systematic uncertainties on signal modelling, signal yield and background modelling are described.

8.6.1 Signal Modelling Uncertainties

In the spin-0 analysis, the signal parameterisation is extracted from gluon-gluon fusion narrow-width MC samples. However, the dependence of the signal shape on the production process is checked with the narrow-width MC samples of the five Higgs-like production processes. While some variations are observed in the tails of the signal mass distribution, the bias on the number of fitted signal events is estimated to be less than 1%. This uncertainty is hence neglected.

The most relevant uncertainty in the signal modelling is coming from the photon energy resolution. For this high mass search, the uncertainty in the signal mass resolution is mostly driven by the uncertainty in the constant term of the energy resolution. This mainly impacts the σ_{CB} parameter. The signal samples are generated with upward and downward variations of the photon resolution. The $m_{\gamma\gamma}$ distributions are fitted to extract the value of the σ_{CB} parameter for both upward and downward variations. The width of the DSCB function σ_{CB} varies from $^{+30}_{-20}\%$ to $^{+60}_{-40}\%$ as a function of the mass, in the range from 200 GeV to 1000 GeV, and stays almost constant above 1000 GeV.

The uncertainty in the photon energy scale shifts the resonance peak position. This impacts the precision on the measurement of the resonance mass which is not relevant in this stage of the search, hence it is not included in both analyses.[5]

8.6.2 Signal Yield Uncertainties

The uncertainty on the photon identification efficiency is estimated from the comparison of the efficiencies obtained from simulation before and after applying the shower shape shifts (see Section 3.3) to match data distributions. This gives an uncertainty in signal yield of about ± 2 to 3%.

For the calorimeter-based isolation efficiency, the uncertainty is taken from the difference between the nominal value and that from applying a data-driven shift to the isolation energy in simulation. The uncertainty on the signal yield associated with the calorimeter-based isolation is typically less than 1%.

The track-based isolation efficiency uncertainty also arises from imperfect modelling of the track isolation of photons in MC simulation. The uncertainty on the signal yield from track isolation efficiency is $\pm 1 - 4\%$.

For the spin-0 case, an additional systematic uncertainty is assigned to cover the difference in signal correction factor due to production mode. The C_X factors are evaluated for all five production modes and the uncertainty is taken to be half the difference between the two extremes, giving a final value of $\pm 2.9\%$. For the large-width case, there is an additional 5% uncertainty assigned due to the difference between the C_X factors computed at different widths (see Figure 8.12).

The integrated luminosity of 3.2 fb^{-}1 has an uncertainty of $\pm 5\%$. The uncertainty is derived following a methodology similar to that detailed in Ref. [125] from a preliminary calibration of the luminosity scale using van der Meer beam-separation scans performed in August 2015. Other sources of uncertainties have negligible impact on the signal yield.

8.6.3 Background Modelling Uncertainties

In the spin-0 search, the background modelling uncertainty is estimated from the spurious signal as described in Section 8.3.4. The spurious signal uncertainties vary from 7 (20) events at 200 GeV to 0.006 (0.04) events at 2000 GeV for the case of NWA ($\Gamma_X/m_X = 6\%$). In the functional form approach, the statistical uncertainty is taken from the uncertainty in the determination of the parameters of the function from the fit to data.

[5] If a new resonance is discovered, this would become important.

The uncertainties in the background template for spin-2 resonance search consist of the uncertainties from the shape of the irreducible background, the shape of the reducible background components and the background composition. In the irreducible background, the uncertainties considered are the PDF uncertainty, the choice of the PDF set, the generator-level photon isolation cut in DIPHOX, as well as the fragmentation, factorisation and renormalisation scales. The PDF uncertainty is the dominant theoretical uncertainty, varying from ±2% at a mass of 200 GeV to ±35% at a mass of 3500 GeV. The photon isolation requirement applied at the parton level in DIPHOX gives an uncertainty of about ±10%. In addition, experimental uncertainties in the efficiency of the tight photon identification and isolation requirements are also included.

For the shape of the reducible background, the variations in the fit functions are studied by changing the definition of the control samples. The irreducible background are combined with the reducible background components with the fractions from the 2D isolation fit method. The uncertainties in the background fractions are determined using the background yields from different LoosePrime definitions in the 2D isolation fit. Finally, the bin-by-bin MC statistical uncertainties, which range from ±5 to ±10% in a 5 GeV mass bin are also considered.

Figure 8.13 shows the evolution of the uncertainties in the background prediction for the spin-2 analysis as a function of $m_{\gamma\gamma}$ in the region $m_{\gamma\gamma} > 500$ GeV. At very high masses, the main contribution to the uncertainty originates from the shape of the irreducible background, which in turn mostly arises from the PDF uncertainty.

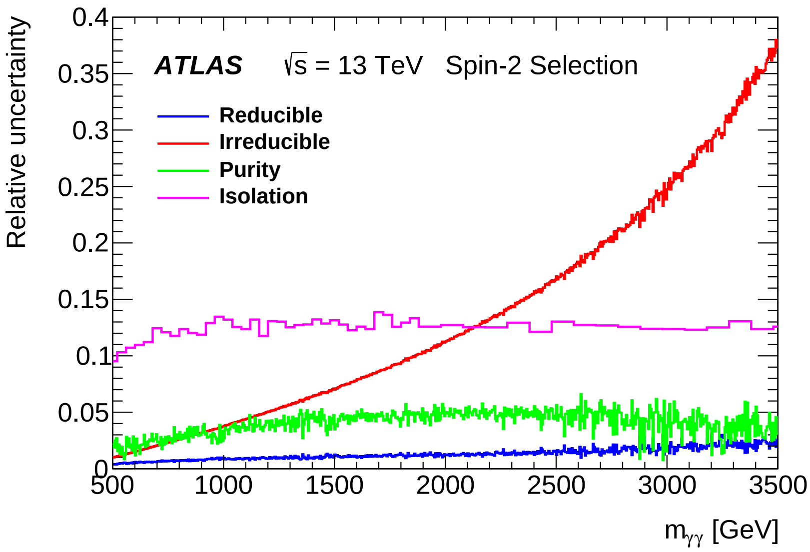


Figure 8.13: Relative uncertainties in the shape of the $m_{\gamma\gamma}$ distribution of the background template used for the spin-2 resonance search prior to fit to data. The reducible and irreducible background uncertainty arise from the uncertainty in the shape of the reducible and irreducible background component respectively. The uncertainty in the purity corresponds to the impact of the relative normalisation of the reducible background compared to the irreducible background while the uncertainty in isolation results from the uncertainty due to the choice of parton-level isolation cut in the DIPHOX NLO computation. [69]

8.7 Statistical Model

The numbers of signal and background events are obtained from maximum-likelihood fits of the $m_{\gamma\gamma}$ distribution of the selected events, for each signal mass and width hypothesis.

In the spin-0 resonance search, the data is described with a per-event likelihood function expressed as

$$\mathcal{L}(m_{\gamma\gamma}; \sigma_{\text{fid}}, m_X, \Gamma_X/m_X, N_{bkg}, \mathbf{a}, \theta) = N_X(\sigma_{\text{fid}}, m_X, \theta_{\mathbf{N_X}}, \theta_{SS}) f_X(m_{\gamma\gamma}, x_X(m_X, \Gamma_X/m_X), \theta_\sigma) + N_{bkg} f_{bkg}(m_{\gamma\gamma}, \mathbf{a}) \tag{8.6}$$

where σ_{fid} is the fiducial production cross-section of the hypothetical resonance of mass m_X; $\mathbf{a}$ denotes the parameters of the background function; $\theta_{\mathbf{N_X}}$ collectively designates the nuisance parameters describing the systematic uncertainties that affect the normalisation of signal yields; θ_{SS} is the nuisance parameter associated to the spurious signal uncertainties; x_X denotes the parameters describing the signal shape; and θ_σ the nuisance parameter of the mass resolution systematic uncertainties

$\theta_{\mathbf{N_X}}$ consist of θ_{lumi} for the uncertainty on the integrated luminosity of the data sample, $\theta_{eff,X}$ for the systematic uncertainty on photon identification efficiency of the resonance, $\theta_{iso,X}$ for the systematic uncertainty on isolation efficiency of the resonance, θ_{ES} for the photon energy resolution systematics and θ_{C_X} for the production-mode uncertainty on the C_X factors.

N_X and N_{bkg} are the number of signal and background events respectively. N_{bkg} is a free parameter in the fit, while N_X is parameterised as

$$N_X(\sigma_{\text{fid}}, m_X, \theta_{\mathbf{N_X}}, \theta_{SS}) = \sigma_{\text{fid}}\, \mathcal{L}_{int}\, C_X(m_X) \prod_{i=1}^{|\theta_{\mathbf{N_X}}|} K_i(\theta_i) + \sigma_{SS}\theta_{SS} \tag{8.7}$$

where $\mathcal{L}_{int}$ is the integrated luminosity of the sample; K_i a function characterising the effect of the normalisation systematics θ_i in $\theta_{\mathbf{N_X}}$; σ_{SS} and θ_{SS} are the spurious signal systematic uncertainty and its associated nuisance parameter. $K_i(\theta_i)$ implements the systematic uncertainties with the expression $K_i(\theta_i) = r_i(m_X)^{\theta_i}$, with

$$r_i(m_X) = \begin{cases} (1+\sigma_i(m_X))N_X & \theta_i > 0 \\ (1-\sigma_i(m_X))N_X & \theta_i < 0 \end{cases} \tag{8.8}$$

where σ_i is the uncertainty on the yield expressed in fraction so that the modifications to the signal yield for $\theta_i = \pm 1$ are exactly equal to the $\pm 1\sigma$ variations used to define the uncertainties. $r_i(m_X)$ is interpolated smoothly between $\theta_i > 0$ and $\theta_i < 0$.

The systematic uncertainties in the signal modelling, spurious signal and on the signal yield are constrained with Gaussian penalty terms.

The signal pdf f_X is the double-sided Crystal Ball function with parameters $x_X = \{\Delta m, \sigma_{CB}, \alpha_{low}, \alpha_{high}, n_{low}, n_{high}\}$, as defined in Section 8.4. The uncertainty on the σ_{CB} parameter is applied using the expression $\sigma_{CB} = \sigma_0 K_\sigma(\theta_\sigma)$, where σ_0 is the nominal value of σ_{CB} and $K_\sigma(\theta_\sigma)$ is defined similarly as for the event yield systematics above. The background pdf f_{bkg} is described by the function $f_{bkg}(m_{\gamma\gamma}; b, a) = (1 - x^{1/3})^b x^a, a = m_{\gamma\gamma}/\sqrt{s}$, with parameters a and b free in the fit, as detailed in Section 8.3.

The overall likelihood is

$$\mathcal{L}(\sigma_{\text{fid}}, m_X, \Gamma_X/m_X, N_{bkg}, \mathbf{a}, \theta) = e^{(N_X+N_{bkg})} \left[\prod_{j=1}^{n} \mathcal{L}(m_{\gamma\gamma,j}; \sigma_{\text{fid}}, m_X, \Gamma_X/m_X, N_{bkg}, \mathbf{a}, \theta) \right] \left[\prod_{k=1}^{|\theta|} \exp(-\frac{1}{2}(\theta_k - \theta_k^{\text{aux}})^2) \right], \tag{8.9}$$

where n is the number of events in the dataset and $\theta^{\mathbf{aux}}$ is the set of auxiliary measurements used to constrain the systematic uncertainties.

For the spin-2 resonance search, the statistical model is similar except for the use of a binned likelihood with 5 GeV bin size due to the template-based background modelling. The uncertainties in background template (the uncertainty in the shape of each background component and in the relative normalisation of the different components) are implemented in the model as three independent nuisance parameters with Gaussian constraints. The effect of the limited MC statistic is also included in the model by means of one nuisance parameter per bin, again with Gaussian constraint.

The determination of local significance, global significance and upper limits follows the procedure described in Chapter 5.

8.8 Results

The search region of the spin-0 resonance search is 200-2000 GeV in mass m_X and 0-10% in relative width Γ_X/m_X. The lower mass range is chosen to minimise the gap between the high mass search and the SM Higgs analysis while allowing sufficient sideband for the fit of the diphoton invariant mass from 150 GeV. Beyond $m_{\gamma\gamma} = 2000$ GeV, there is no event in data to constrain the background shape.

For spin-2 resonance search, the search region is restricted to 500-2000 GeV in mass and 0.01-0.3 in $k/\overline{M}_{\mathrm{Pl}}$. The $200 < m_{\gamma\gamma} < 500$ GeV region is used as a control region to normalise the background template, while the range in $k/\overline{M}_{\mathrm{Pl}}$ is chosen to approximately match the width range of the spin-0 search. The search region for which p-values are computed is up to 2000 GeV for the spin-2 resonance search but limits on the total RS graviton cross-section times branching ratio are computed for graviton mass up to 5 TeV.

8.8.1 Spin-0 Selection

Figure 8.14 shows the diphoton invariant mass distribution for events passing the spin-0 selection together with the background-only fit (signal component being set to zero) using the functional-form approach. A broad excess is observed near a mass of 750 GeV. The local p-values testing the compatibility with the background-only hypothesis are computed as a function of signal mass m_X and relative width Γ_X/m_X in the search range. The local significance as a function of mass and relative width is illustrated in Figure 8.15. The largest deviation from the background-only hypothesis is observed near a mass of 750 GeV, with a local significance of 3.9σ for $\Gamma_X/m_X = 6\%$ corresponding to a width of 45 GeV. The global significance is $2.1 \pm 0.05\sigma$, evaluated using an ensemble of pseudo-experiments with the technique described in Section 5.3 for the defined search region. If a narrow-width signal is assumed, the largest deviation from the background-only hypothesis is found also near 750 GeV, corresponding to a local significance of 2.9σ.

The result of a signal+background fit to 2015 data passing the spin-0 selection with resonance mass and width values floating is shown in Table 8.2.

$N_{\mathrm{background}}$	N_{signal}	m_X	Γ_X/m_X	a	b
7369 ± 86	22 ± 9	749 ± 10 GeV	$6 \pm 3\%$	-3.41 ± 0.30	7.60 ± 2.38

Table 8.2: The number of background events $N_{\mathrm{background}}$, the number of signal events N_{signal}, the resonance mass m_X, the relative width Γ_X/m_X and the parameters a and b of the background function obtained from a signal+background fit to 2015 data passing the spin-0 selection.

In the NWA, the fit pulls the nuisance parameter associated to the photon energy resolution by $\approx 2\sigma$, as the excess is broader than the experimental $m_{\gamma\gamma}$ resolution. The pulls on the resolution uncertainty nuisance parameter are shown in Figure 8.16.

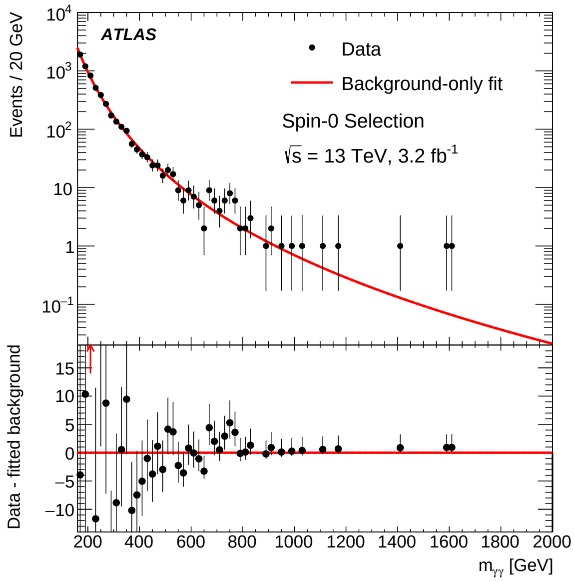


Figure 8.14: Distribution of the diphoton invariant mass in 2015 data with spin-0 selection, with a background-only fit. The difference between the data and the fit is shown in the bottom panel for $m_{\gamma\gamma} > 200$ GeV. The arrow shown in the lower panel indicates a value outside the range with more than one standard deviation. There is no data event with $m_{\gamma\gamma}$ above 2 TeV.

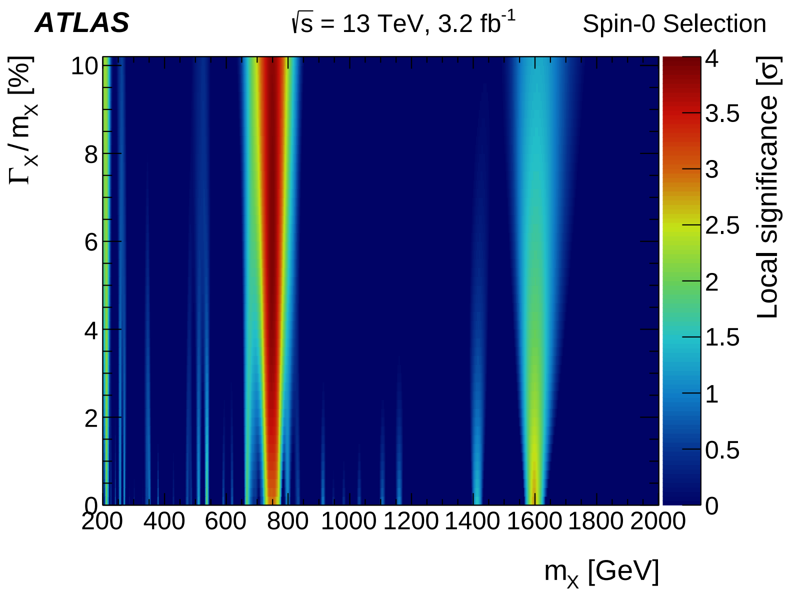


Figure 8.15: Compatibility, in terms of local significance, with the background-only hypothesis as a function of the assumed signal mass m_X and relative width Γ_X/m_X for the spin-0 resonance search using 2015 data.

8.8.2 Spin-2 Selection

Figure 8.17 shows the diphoton invariant mass distribution for events passing the spin-2 selection together with the background-only fit using the MC-based background template. Similar to the spin-0 resonance search, a broad excess is observed near a mass of 750 GeV. The local p-values testing the compatibility with the background-only hypothesis are computed as a function of the signal mass m_{G_*} and $k/\overline{M}_{\mathrm{Pl}}$ in the search range. The local significance as a function of mass and $k/\overline{M}_{\mathrm{Pl}}$ is illustrated in Figure 8.18. The largest deviation from the background-

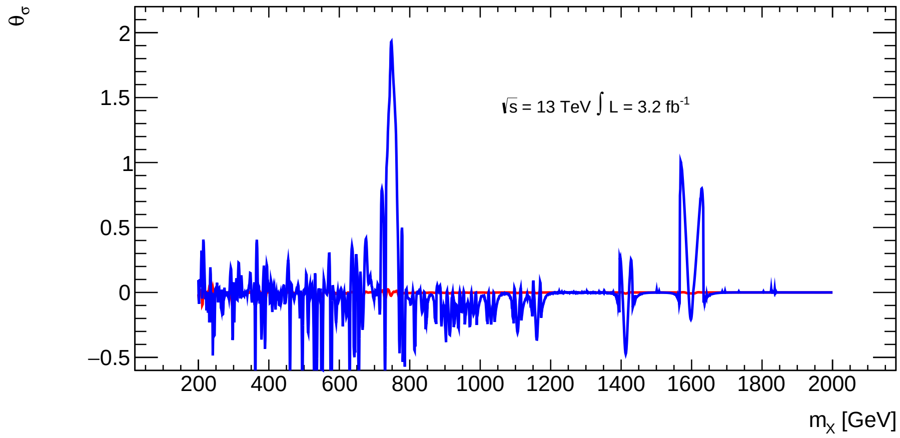


Figure 8.16: Pulls on the nuisance parameter associated to the resolution uncertainty in number of standard deviation as a function of mass assuming a signal of narrow width. The θ_σ is pulled by $\approx 2\sigma$ near 750 GeV due to the presence of a broad excess.

only hypothesis is observed near a mass of 750 GeV, with a local significance of 3.8σ for a $k/\overline{M}_{\rm Pl}$ value of 0.23. The Breit-Wigner width associated with $k/\overline{M}_{\rm Pl} = 0.23$ is 57 GeV. The global significance is $2.1 \pm 0.05\sigma$, evaluated using an ensemble of pseudo-experiments for the search region defined. If a narrow-width signal is assumed with $k/\overline{M}_{\rm Pl} = 0.01$, the largest deviation from the background-only hypothesis is found near 770 GeV, corresponding to a local significance of 3.3σ.

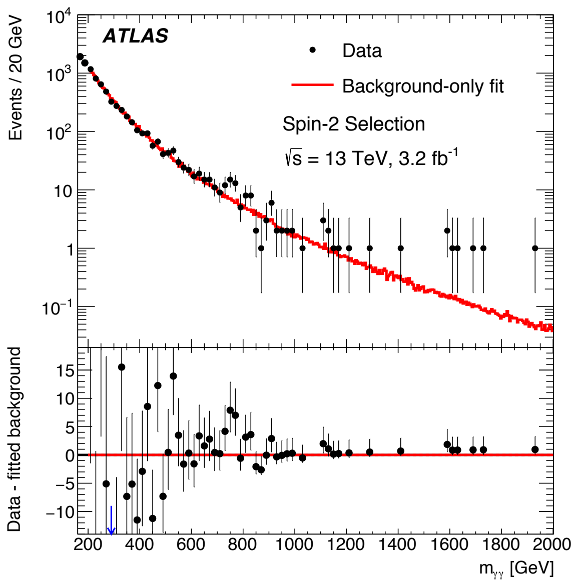


Figure 8.17: Distribution of the diphoton invariant mass in 2015 data with spin-2 selection, with the background-only fit. The difference between the data and the fit is shown in the bottom panel for $m_{\gamma\gamma} > 200$ GeV. The arrow shown in the lower panel indicates a value outside the range with more than one standard deviation. There is no data event with $m_{\gamma\gamma}$ above 2 TeV.

8.8.3 Limits on Cross-section

Limits on the cross-section times branching ratio to diphoton are derived using the asymptotic approximation (see Section 5.5 and 5.6). For the spin-0 search, the cross-section limits are

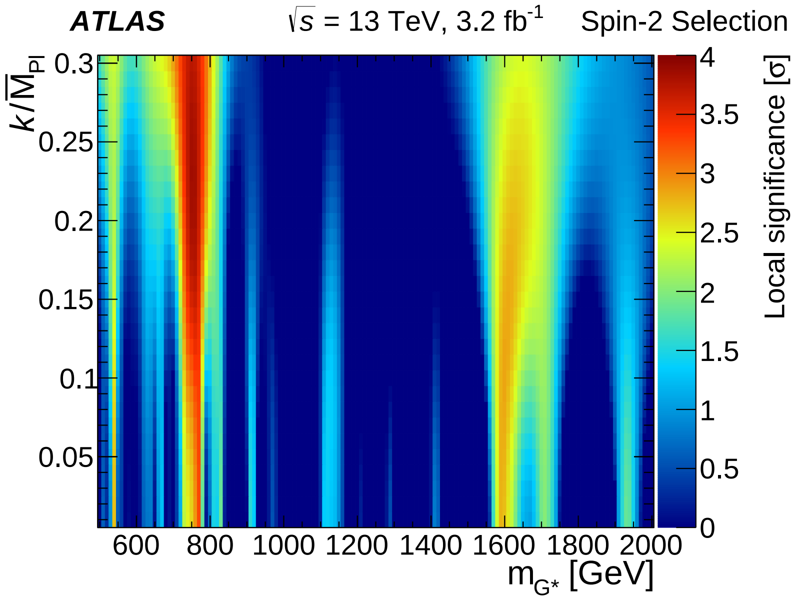


Figure 8.18: Compatibility, in terms of local significance, with the background-only hypothesis as a function of the assumed signal mass and $k/\overline{M}_{\rm Pl}$ for the spin-2 resonance search using 2015 data.

expressed within the fiducial acceptance defined in Section 8.5.1. This allows the limits to be interpreted in a nearly model-independent manner. Figure 22 shows the limits on the signal fiducial cross-section times branching ratio to two photons for a spin-0 resonance as a function of the hypothesised mass and for four different values of signal width. The observed limits are within the $\pm 1, 2\sigma$ bands of the expected limits consistent with a background-only hypothesis except around 750 GeV where the observed limit is less stringent due to the presence of an excess around that mass. Better limits are set for the narrow-width hypothesis compared to large widths due to the higher sensitivity to search for narrow signals.

For the spin-2 case, the analysis is performed for a specific benchmark model of a spin-2 graviton, hence limits are given on the total cross-section times branching ratio to diphoton. Figure 8.20 shows the limits as a function of mass for different $k/\overline{M}_{\rm Pl}$ values from 0.01 to 0.3. The limits are compared to the LO cross-sections predicted by the benchmark model in Pythia8. The uncertainty band shows the PDF uncertainty estimated from the variations of the NNPDF23LO PDF set.

8.8.4 Validity of NWA Spin-0 Limit for Resonances with Larger Width

The spin-0 analysis initially presented in December 2015 [119] was restricted to the narrow-width hypothesis since the large-width signal modelling was not fully developed and integrated in the statistical framework. Instead, a range of natural widths that can be excluded with the NWA limit with a bias on the yield of less than 10 and 20% was presented.

The allowed width range is computed by calculating the fit bias from fitting a non-narrow-width signal+background sample with the narrow-width model. Asimov datasets consisting of a background component modelled by

$$f_{bkg}(m_{\gamma\gamma}) = (1 - x^{1/3})^{4.09} x^{-3.8}, x = m_{\gamma\gamma}/\sqrt{s} \tag{8.10}$$

and a signal component with varying width are generated for each mass point in the range $200 < m_X < 1600$ GeV. [6] For each Asimov dataset, 7353 background events were generated,

[6] The parameters of the background function are obtained from a background-only fit to the preliminary 2015

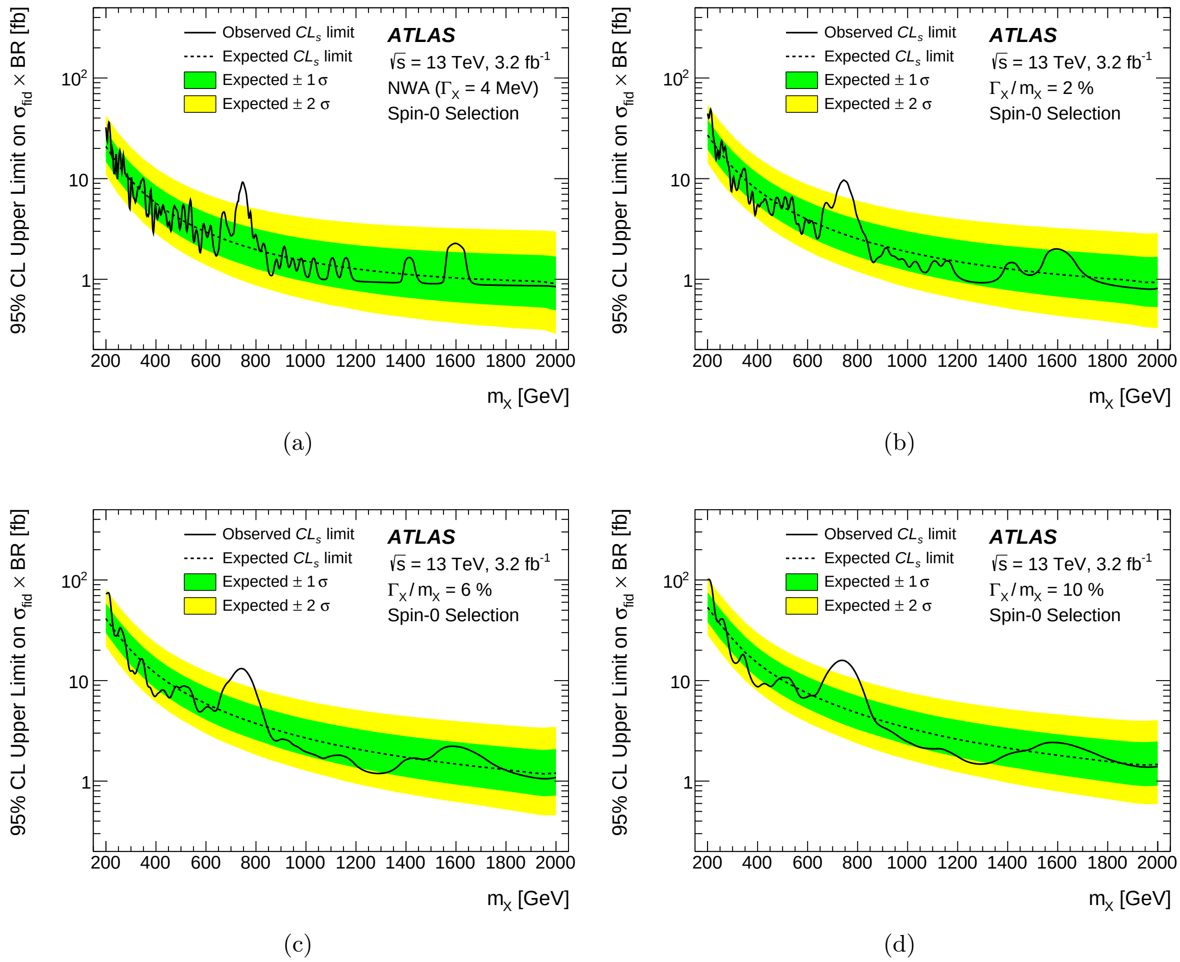


Figure 8.19: 95% CL upper limits on the fiducial cross-section of a spin-0 resonance as a function of the assumed mass, for different values of the decay width divided by the mass.

and the number of signal events is the expected number of events for a resonance having a cross-section equal to twice the expected NWA limit.

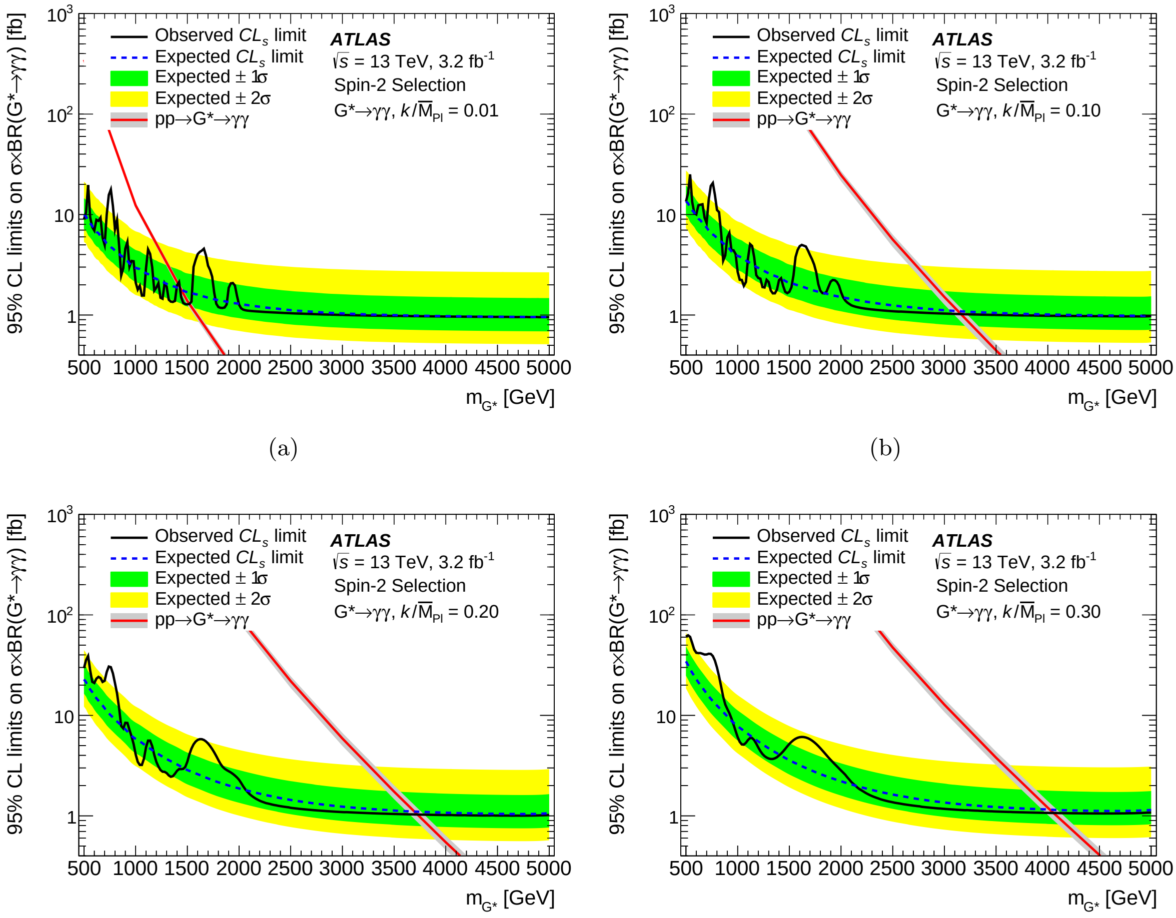
Figures 8.21 and 8.22 show respectively the fitted nuisance parameter associated to the uncertainty on σ_{CB}, θ_σ, and the bias on the number of fitted signal events for each mass point as a function of the injected resonance width. The range of acceptable widths for the signal model is defined such that the bias on the fitted signal yield is kept below 10 or 20%. The acceptable width ranges extracted from this plot are shown in Figure 8.23 as a function of the resonance mass m_X.

The bias on the number of fitted signal events using the NWA signal parameterisation is found to be smaller than 10% (20%) for a natural width between $\Gamma_X/m_X = 0.4 - 0.7\%$ ($\Gamma_X/m_X = 1.0 - 1.7\%$).

8.8.5 Checks on Asymptotic Approximation of p-value

For the spin-0 analysis, the smallest p-value at the assumed mass and width corresponding to the largest deviation from background-only hypothesis is checked using an ensemble of pseudo-experiments. The observed test statistic t_0^{uncap} is 15.3. Using asymptotic formula, the local

data as presented in December 2015. As such, the background parameters, the number of background events and the search range are slightly different to what presented in Section 8.8.1.



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

Figure 8.20: 95% CL upper limits on the production cross-section of an RS graviton as a function of the assumed mass, for different values of $k/\overline{M}_{\mathrm{Pl}}$. The predicted cross-sections times branching ratio to two photons for the RS graviton model, computed at LO, are also shown. The uncertainty in the cross-section values represent the PDF uncertainty.

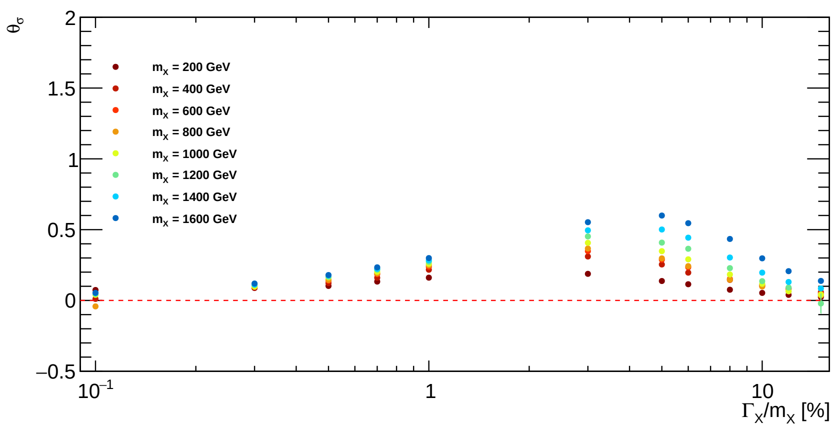


Figure 8.21: Fitted nuisance parameter θ_σ as a function of the injected signal width, for $200 < m_X < 1600$ GeV.

p-value is 4.6×10^{-5} or in terms of significance, 3.91σ.

4.5 million background-only pseudo-experiments are generated to test the $m_X = 749$ GeV,

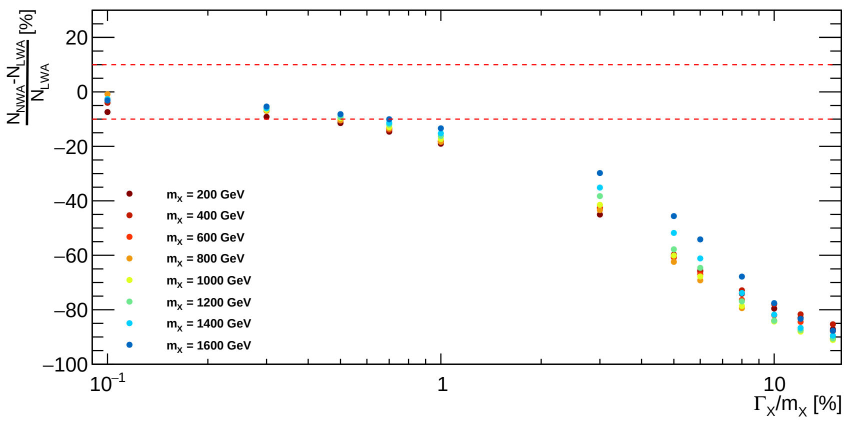
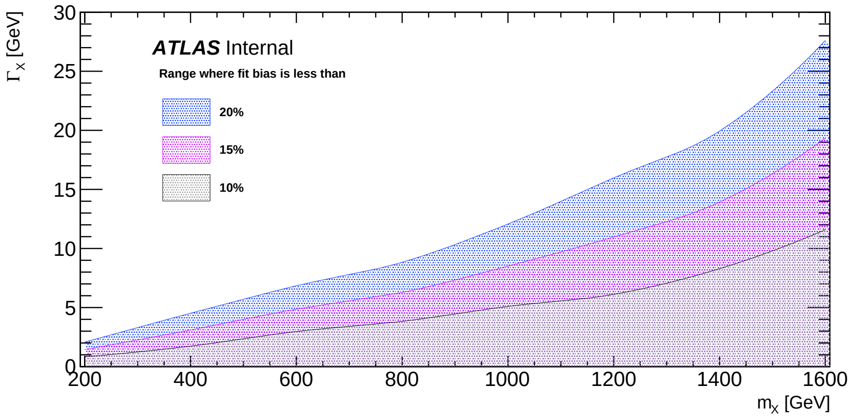
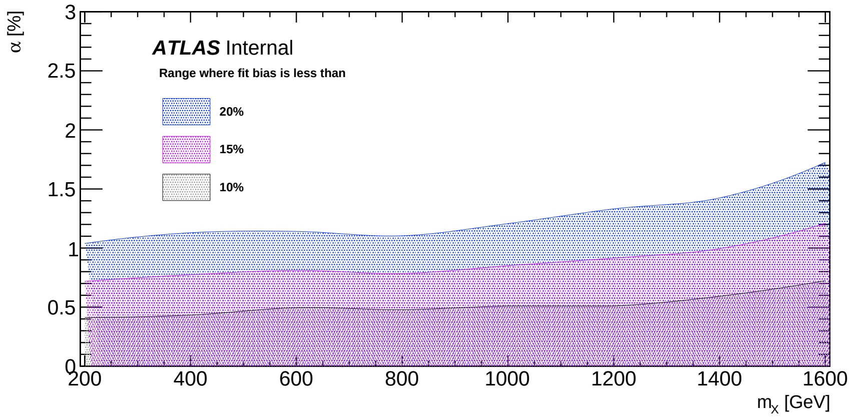


Figure 8.22: Bias on the number of fitted signal events as a function of the injected signal width, for $200 < m_X < 1600$ GeV. The dashed line shows the acceptable limit of 10% for the fit bias.



(a)



(b)

Figure 8.23: Range of natural signal widths compatible with the fiducial limits presented in the December 2015 CONF note [119], as a function of the resonance mass m_X.

$\Gamma_X/m_X = 6\%$ hypothesis. The distribution of t_0^{uncap} is shown in Figure 8.24. The distribution can be approximated with a χ^2 distribution with one degree of freedom. Counting the fraction of events above the observed t_0^{uncap}, the local p-value is $3.9 \pm 0.29 \times 10^{-5}$. The uncertainty on the p-value is statistical and arises from the limited number of pseudo-experiments. This corresponds to local significance of $3.95 \pm 0.02\sigma$. The p-value obtained using asymptotic formula is thus a good approximation.

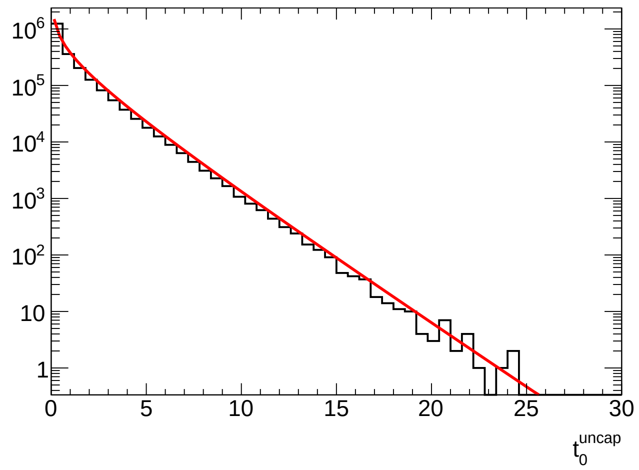


Figure 8.24: Distribution of the test statistic t_0^{uncap} obtained from pseudo-experiments (histogram). Only positive values of the test statistic are shown. A χ^2 distribution with one degree of freedom (red curve) has been overlaid to show the validity of asymptotic approximation.

8.9 8 TeV Re-analysis

Searches for diphoton resonances with 8 TeV data have been reported in [126] and [127]. Given the presence of an excess around 750 GeV in 2015, the compatibility with 8 TeV results is checked.

The 8 TeV data recorded in 2012 with an integrated luminosity of 20.3 fb^{-1} is re-analysed with the final Run 1 calibration as described in Ref. [61]. This is to ensure the calibration is similar to ease comparison of the two datasets. The selections are unchanged with respect to those in the original publication ([126,127]). In particular, in the spin-2 analysis, both photons are required to have a transverse momentum greater than 50 GeV. The isolation requirements are also different in both selections.

In the original publication, the spin-0 search at 8 TeV was performed only up to a diphoton invariant mass of 600 GeV due to the sliding window background modelling method. In this method, the fit is performed in a window of $m_{\gamma\gamma}$ around the tested mass. For each mass hypothesis, an upper sideband is needed to constrain the background fit, hence the search was only performed up to 600 GeV. In order to compute the compatibility with 13 TeV data, the search is extended up to 2000 GeV.

The signal and background modelling follow those used in the 13 TeV analysis. For the large-width part of spin-0 search, no MC samples were available. Therefore, the line shape was not fully validated and the fiducial C_X factors are inferred from the narrow-width C_X factor and the line shape assumption. This is described in some details in Section 8.9.1. The results of the 8 TeV re-analysis are presented in Section 8.9.2.

8.9.1 Large-Width Scalar Signal and Fiducial Correction Factor

Large-width spin-0 signal modelling at $\sqrt{s} = 8$ TeV follows those from 13 TeV analysis. The same line shape parameterisation as in Eq. 7.30 is used where the gluon-gluon luminosity is computed at a centre-of-mass energy of 8 TeV. The theoretical line shape is validated with truth-level samples generated with PowHeg. It is assumed that the convolution of the detector resolution obtained from narrow-width samples with the theoretical line shape would yield a good description of the reconstructed line shape, as it has been demonstrated on the 13 TeV MC samples.

For the large-width scalar model, the fiducial mass volume of $m_X - 2\Gamma_X < m_{\gamma\gamma}^{\mathrm{true}} < m_X + 2\Gamma_X$ is applied on top of the NWA case, the same as for the 13 TeV analysis. The numerator of the C_X factor is determined from a signal + background fit of the large-width sample combined with an Asimov background component, as done in Section 8.5.2. The number of injected signal events is chosen to be twice the expected upper limit on the event yield.

As there is no large-width sample available for this study, the extraction of C_X factors is therefore based on pseudo-samples. Those pseudo-samples containing information on the true and reconstructed masses are generated from the parameterisation of the theoretical line shape as implemented in PowHeg and its convolution with the detector resolution respectively. The resolution function is extracted from the narrow-width sample at $\sqrt{s} = 8$ TeV.

To use pseudo-samples to determine the large-width C_X factors, the C_X factor as defined in Section 8.5.1 is factorised for the large-width case:

$$C_X = \frac{\mathrm{N_{selection}}}{\mathrm{N_{acceptance}}} \times \frac{\mathrm{f_{reco}}}{\mathrm{f_{fiducial}}} \tag{8.11}$$

where $\mathrm{N_{selection}}$ and $\mathrm{N_{acceptance}}$ are as defined in Section 8.5.1, $\mathrm{f_{reco}}$ is the fraction of reconstructed events fitted by our large-width signal parameterisation from a signal plus background fit, and $\mathrm{f_{fiducial}}$ is the fraction of signal events within the fiducial mass volume.

$\mathrm{N_{selection}}/\mathrm{N_{acceptance}}$ is the definition of narrow-width C_X factor which neither include a fiducial mass volume nor taken from the number of fitted signal events. $\mathrm{f_{reco}}/\mathrm{f_{fiducial}}$ represents the correction on top of narrow-width C_X factor definition for the large-width case. It corrects the $\mathrm{N_{selection}}$ to the number of fitted signal events and adding a fiducial mass volume to $\mathrm{N_{acceptance}}$.

$\mathrm{f_{reco}}$ and $\mathrm{f_{fiducial}}$ are determined in the exact same way as with fully-simulated MC samples, but using pseudo-samples instead. $\mathrm{f_{reco}}$ is obtained from the fraction of reconstructed signal events from signal+background fit of a sample combining an Asimov background component with the signal events from pseudo-sample at given mass and width weighted such that the total number of signal events is twice the expected upper limit on the event yield. $\mathrm{f_{fiducial}}$ is computed from the fraction of signal events within the fiducial mass volume from the true mass distribution of the pseudo-sample at a given mass and width.

From the 13 TeV samples, it is observed that $\mathrm{N_{selection}}/\mathrm{N_{acceptance}}$ from narrow-width MC samples is very similar to those from the large-width MC samples. To estimate the C_X factors for a large-width scalar model, the $\mathrm{N_{selection}}/\mathrm{N_{acceptance}}$ from narrow-width samples is multiplied by $\mathrm{f_{reco}}/\mathrm{f_{fiducial}}$ calculated from pseudo-samples with resonance mass of 500, 750, 800 and 1000 GeV with few width scenarios (5, 7 and 10 % except 750 GeV where widths of 3, 6 and 10 % are used).

This estimation of C_X factors using pseudo-samples has been checked at 13 TeV and yields consistent results to the C_X factors calculated using fully-simulated MC samples. Thus, the estimation of large-width C_X factors using pseudo-samples at 8 TeV is justified. Figure 8.25 illustrates the parameterisation of the C_X factors for large-width scalar signals for the 8 TeV re-analysis. Very little width dependence is seen and the same parameterisation of C_X factors for widths up to 10 % is used.

8.9.2 Compatibility between 8 TeV and 13 TeV Data

The invariant mass distributions of the 8 TeV data recorded in 2012 is shown in Figure 8.26 for the two selections. In the search optimised for a spin-0 resonance, checking for excess at the mass and width of the largest deviation in 13 TeV data, a small excess is seen in the 8 TeV data, corresponding to 1.9σ local significance. For the spin-2 search, no excess is observed at the mass and width of the largest deviation in 13 TeV data.

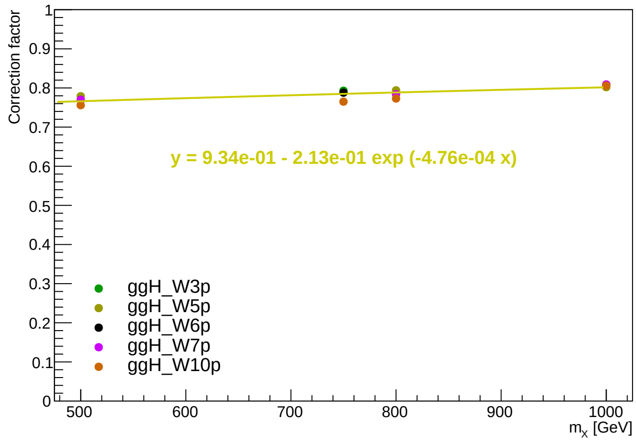
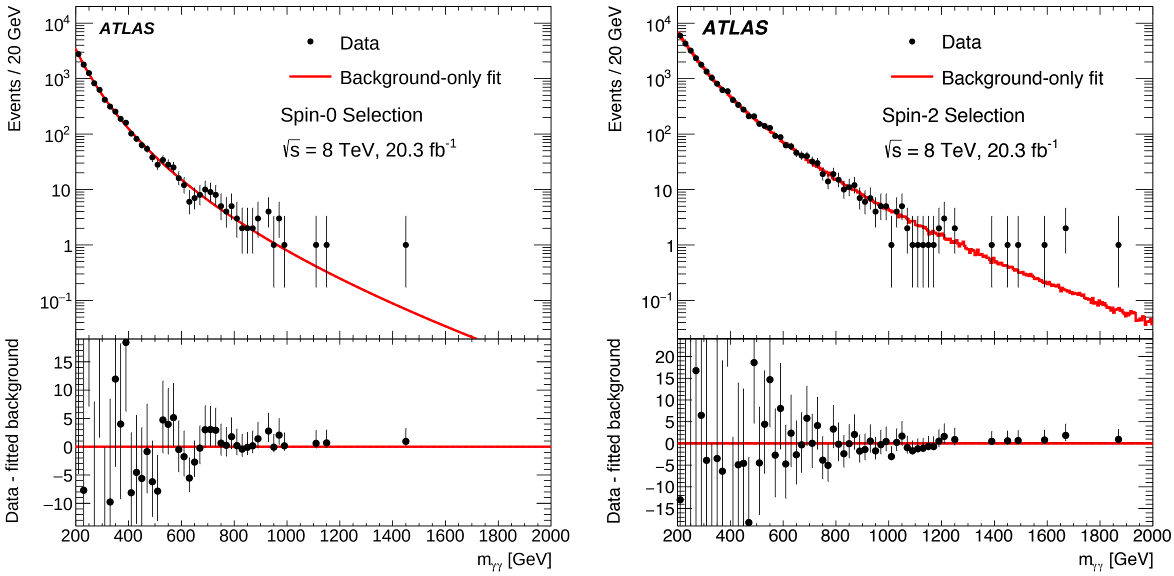


Figure 8.25: C_X factor as a function of m_X for the large-width scalar models.



(a) (b)

Figure 8.26: Distribution of the diphoton invariant mass in 8 TeV data passing the (a) spin-0 and (b) spin-2 selections defined in the 8 TeV publication. Background-only fit to the data is shown in red. There is no data event with $m_{\gamma\gamma}$ above 2 TeV.

To evaluate the compatibility of the excess near mass of 750 GeV between the 8 TeV and 13 TeV data, a common signal model is assumed. For a particle of mass 750 GeV produced in an s-channel process, the cross-section is expected to increase by the ratio of the parton luminosity at $\sqrt{s} = 13$ TeV to 8 TeV. The increase in cross-section is estimated with the MSTW2008NLO PDF set to be 4.7 for a gluon-gluon initial state and 2.7 for a light quark-antiquark initial state. These numbers are consistent with the increase predicted by the PDF4LHC15 NNLO PDF set.

The compatibility of the results at 8 TeV and 13 TeV is estimated from a combined analysis of both datasets, accounting for correlations in systematic uncertainties in particular for the photon energy scale. The compatibility between the two datasets is quantified using $\sigma_{8\,\mathrm{TeV}}/\sigma_{13\,\mathrm{TeV}}$ as parameter of interest.

Assuming a signal with mass $m_X = 750$ GeV and $\Gamma_X/m_X = 6\%$, the difference between the 8 TeV and the 13 TeV results passing the spin-0 selection corresponds to 1.2σ if gluon-gluon production is assumed, and 2.1σ for quark-antiquark production. For the analysis optimised for a spin-2 resonance, assuming signal mass of 750 GeV and $k/\overline{M}_{\rm Pl} = 0.2$, the difference corresponds to 2.7(3.3)σ for gluon-gluon (quark-antiquark) production.

8.10 Studies on Categorisation

Categorisation is a way to increase the sensitivity of the analysis by splitting the samples into two or more categories with at least one having better mass resolution or higher ratio of signal over background. Aside from increasing the sensitivity, categorisation also provides a measurement of the properties of the excess, if any, as a higher significance in a category indicates a more prominent contribution of the excess in that category.

Some studies on categorisation were performed before 2016 data was taken in order to enhance the discovery potential. If the excess was due to new physics, categorisation would also provide some information on its properties, such as spin, production mode, etc.

In the following, the categorisation studies are performed assuming a signal at 750 GeV. Only the spin-0 resonance is considered and the resonance is assumed to have narrow width for simplicity and to test the potential of resolution-based categories.

8.10.1 Resolution-Based Categories

The SM $H \to \gamma\gamma$ analysis profits from resolution-based categories. At 125 GeV, the signal-over-background ratio is very small; an improvement in the resolution in certain categories boosts the sensitivity of the search.

In the context of the high mass search, I tried to define categories that classify events in terms of the mass resolution. One of the simplest criterion to define the categorisation is the pseudorapidity of the photon candidates. In general, photon candidates in the central part of the detector are better reconstructed than those in the forward region due to the smaller amount of material and finer segmentation in the calorimeter.

One could define categorisation based on simple cut on the photon pseudorapidity. To define a categorisation that have the best separation of the mass resolution, multivariate analysis (MVA) techniques were employed to identify regions of detector with different resolution. The input variables are the absolute pseudorapidity and transverse momentum of both photons, while the target variable is the difference between the reconstructed diphoton invariant mass and the true invariant mass. The 750 GeV NWA MC sample is used in the MVA training. Both BDT and multilayer perceptron (MLP) methods are tested. It is found that the photon transverse momenta are not discriminant, likely due to the hard spectrum of photon p_T for an assumed 750 GeV scalar resonance. For a given pseudorapidity, the photon energy resolution is dominated by the constant term at high p_T. Hence, the transverse momenta are dropped and only the pseudorapidity of the two photon candidates are retained in the MVA. Both BDT and MLP perform similarly well. MLP was chosen ultimately as the regions in photon pseudorapidity distribution with good and poor resolution are identified in such a way that the categorisation based on MLP score can be replaced with a simple cut on the photon pseudorapidity instead (see Figure 8.27).

The sample is divided into three categories based on the MLP scores, with equal amount of events in each category. The lower the MLP score of an event, the better the mass resolution is in that event. Three categories are thus defined: one with low MLP score and excellent resolution, one with medium MLP score and moderate resolution, and the other one with high MLP score and poor resolution. Figure 8.27 shows the distribution of low, medium and high MLP scores against the absolute pseudorapidity of the leading and subleading photons.

The signal shape in each category is parameterised with a DSCB function. The Crystal Ball width and the fraction of signal events obtained from simulation are presented in Table 8.3.

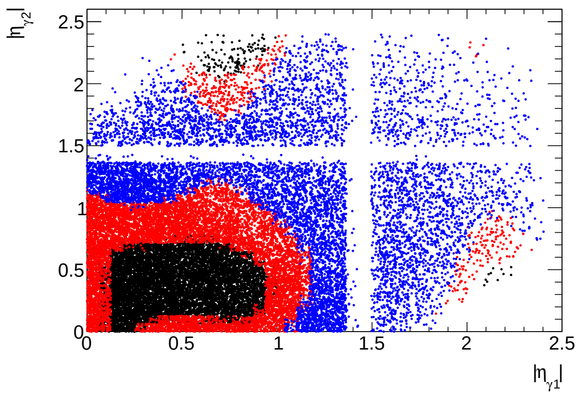


Figure 8.27: Distribution of MLP scores which is an estimate of the mass resolution as a function of the absolute pseudorapidity of the leading photon, $|\eta_{\gamma 1}|$ and of the subleading photon, $|\eta_{\gamma 2}|$. The category with the lowest MLP score and having the best resolution (MLP1) is shown in black, the category with medium resolution (MLP2) shown in red and the one with the worst resolution (MLP3) in blue.

8.10.2 Signal-over-Background Ratio-Based Categories

Categories based on the ratio of the number of signal events to the number of background events can also improve sensitivity. This relies on an assumption of the signal model and properties of the resonance. Examples of categories sensitive to the properties of a potential signal are the number of jets[7], sensitive to the production process and the decay angle of the photons in the resonance rest frame, $\cos\theta^*$, sensitive to the spin.

Given the generic search and lack of motivation for any particular signal properties, these types of categorisation are not studied here. In the case of discovery, such a categorisation can be pursued.

8.10.3 Method

To study the improvement in sensitivity, one needs the signal and background shape, as well as the signal and background fraction in each category. Once known, an Asimov dataset can be generated and the expected significance is computed per category. The improvement in sensitivity is evaluated by comparing the combined significance to the significance in the inclusive case.

The signal shape and fraction in each category are obtained from signal MC. A dedicated signal shape parameterisation is performed for each category. The background shape and normalisation can be estimated from either MC or data. In the inclusive analysis, the background modelling is data-driven. However, a data-driven background estimate for the categorisation study suffers from low statistics while an MC-based estimation relies on a reasonably accurate description of the data by the MC. In both cases, the functional form approach described in Section 8.3.4 is used to model the background. For each category, the chosen function is validated on LO Sherpa $\gamma\gamma$ samples alone. For simplicity, the background template the function is validated on is built with only irreducible background and without correcting for NLO effects. The spurious signals are evaluated per category using the expected number of background events.

In the data-driven estimation, the fraction of background events in each category is obtained from data by counting the events passing all selection cuts and with invariant mass greater than 150 GeV. A signal+background fit is performed on data simultaneously in all categories, and

[7]If the resonance is produced preferentially via vector boson fusion, the signal-over-background ratio in the two-jet category could be high.

the background prediction in each category is taken from the background component of the fit. In the MC-based background estimation method, the fraction of background events in each category is taken from LO SHERPA simulation and the background shape is also fitted on the simulated $\gamma\gamma$ background sample.

In both background estimation methods, the total number of background events is set to 7391, corresponding to the total number of events of the events passing the spin-0 selection in 2015 data. The Asimov dataset is generated from the background estimation (either MC-based or data-driven) along with some amount of signal events at 750 GeV assuming a narrow-width signal. Using the Asimov dataset, the p-value is computed at 750 GeV with the narrow-width hypothesis. The combined local p_0 significance from all the categories is compared to the local p_0 significance in the inclusive case.

8.10.4 Results

In parallel to the MLP-based categorisation, another categorisation inherited from the SM $H \to \gamma\gamma$ analysis which divides the sample into Central-Central and Rest category is tested. The Central-Central category contains events where both photons are in the central region defined by $|\eta| < 0.75$. All other events are grouped in the Rest category.

Table 8.3 shows the mass resolution in terms of the width of the DSCB function and the fraction of signal and background events in each category from MC for the two tested categorisations. Using the MC-based background, both the MLP-based and the Central-Central/Rest categorisation do not improve the sensitivity of the search as the combined p_0 significance is comparable to the inclusive case. Since the more sophisticated method, MLP-based categorisation, does not bring any visible improvement compared to a simpler category definition based on the photon pseudorapidity, only the latter is used in the following.

Category	σ_{CB} / GeV	$fSig$	$fBkg$ (150-3000 GeV)	$fBkg$ (700-800 GeV)	$fData$ (150-3000 GeV)
MLP1	4.60	33.3%	21.6%	22.9%	
MLP2	5.21	33.3%	26.9%	28.8%	
MLP3	6.57	33.3%	55.5%	48.3%	
Central-Central	4.66	36.1%	21.5%	22.7%	20.9%
Rest	5.86	63.9%	78.5%	77.3%	79.1%

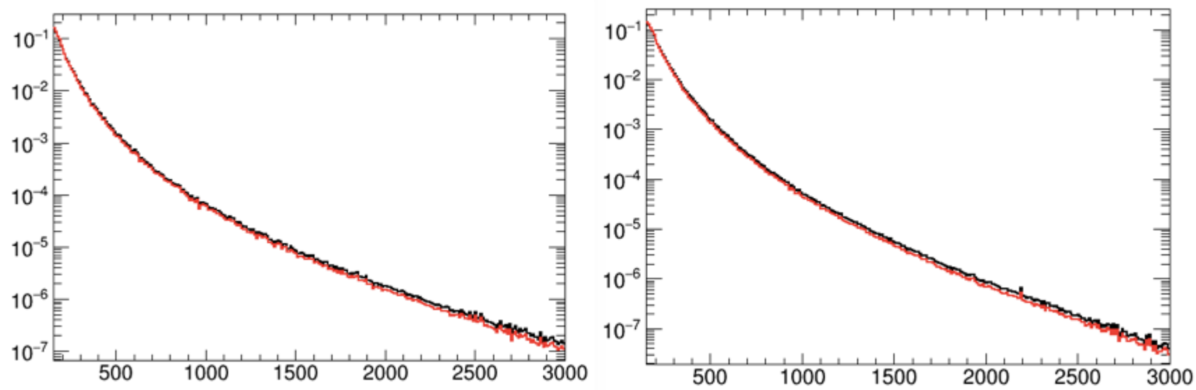
Table 8.3: Width of the double-sided Crystal Ball function σ_{CB} describing the signal shape in each category, fraction of signal events $fSig$ and fraction of background events $fBkg$, all obtained from simulations. For the background, the LO SHERPA $\gamma\gamma$ prediction (without reducible background components) is used. The fraction of events in each category as obtained from data, $fData$, is also shown.

The Central-Central/Rest categorisation is further tested using data-driven background estimation and a sizeable improvement is observed. To understand the difference, the background predictions from both methods are investigated further. The fraction of background events in the Central-Central and Rest categories predicted by LO SHERPA $\gamma\gamma$ simulation and that by data with $m_{\gamma\gamma} \in [150, 3000]$ GeV is found to be similar. However, the background prediction in $m_{\gamma\gamma} \in [700, 800]$ GeV in the Central-Central category is found to be lower when estimated from data compared to simulation, leading to an increase in sensitivity.

Leading-order diphoton MC alone is not expected to describe the data well since the reducible background contribution is sizeable and NLO corrections are not negligible. On the other hand, the data-driven estimation suffers from low statistics especially in the Central-Central category which only accounts for $\sim 20\%$ of the total statistics. The amount of background near the mass of 750 GeV determines the sensitivity of the search. Since there is only a handful of events above 750 GeV to constrain the fit to data, this results in an unreliable background estimate in that region.

Comparing NLO Diphox prediction to LO Sherpa, Diphox is found to predict a harder $m_{\gamma\gamma}$ spectrum. This actually worsens the agreement between the background shape obtained from data compared to MC when the LO Sherpa $\gamma\gamma$ MC is reweighted with NLO k-factor from Diphox. The fraction of background events in the Central-Central category with $m_{\gamma\gamma} \in [150, 3000]$ GeV is found to be 18.5% as Diphox predicts more diphoton events in the forward pseudorapidity than LO Sherpa which predicts 21.5%.

Adding the reducible background component obtained from data control regions to the Diphox reweighted $\gamma\gamma$ MC with the same purity in both categories as obtained for the inclusive case, the overall background shape is found to be quite similar to using LO Sherpa prediction alone. The $m_{\gamma\gamma}$ spectrum is slightly harder in both categories compared to leading-order only prediction as observed in Figure 8.28. The background prediction near 750 GeV in the Central-Central category is still different from the data-driven method when the reducible background component and NLO corrections are added.



(a) (b)

Figure 8.28: $m_{\gamma\gamma}$ spectrum in the (a) Central-Central and (b) Rest category as predicted by LO Sherpa $\gamma\gamma$ MC is shown in red. Black line shows the prediction from Diphox along with reducible background.

Finally, the background shapes in the two categories from the fit to data or from MC-based prediction using LO Sherpa $\gamma\gamma$ prediction are compared. The result is shown in Figure 8.29. In the Rest category, the agreement of the background prediction from both methods is fair while in the Central-Central category, a noticeable difference is observed. The MC-based background estimation fails to describe the $m_{\gamma\gamma}$ distribution in data well while the background function fitted on data has insufficient upper side band to constrain the fit. The excess around 750 GeV is particularly prominent in this category and can change the background fit assuming either the presence or absence of a signal at 750 GeV. However, given the low statistics, it is inconclusive whether MC or function fitted on data provides the most accurate description.

If the data is not divided into categories, LO Sherpa MC manages to describe the data fairly well. Nonetheless, discrepancies emerge when the data is split into η categories or into low and high mass range. Figures 8.30 shows the good agreement in photon η distribution in data compared to LO Sherpa MC passing the spin-0 selection and with $m_{\gamma\gamma} > 150$ GeV. When the diphoton invariant mass is restricted to higher mass range, the events in data tend to populate the more forward region in η. An invariant mass cut of 400-700 GeV is chosen in Figure 8.31 so as not to include any bias from a possible resonance at 750 GeV.

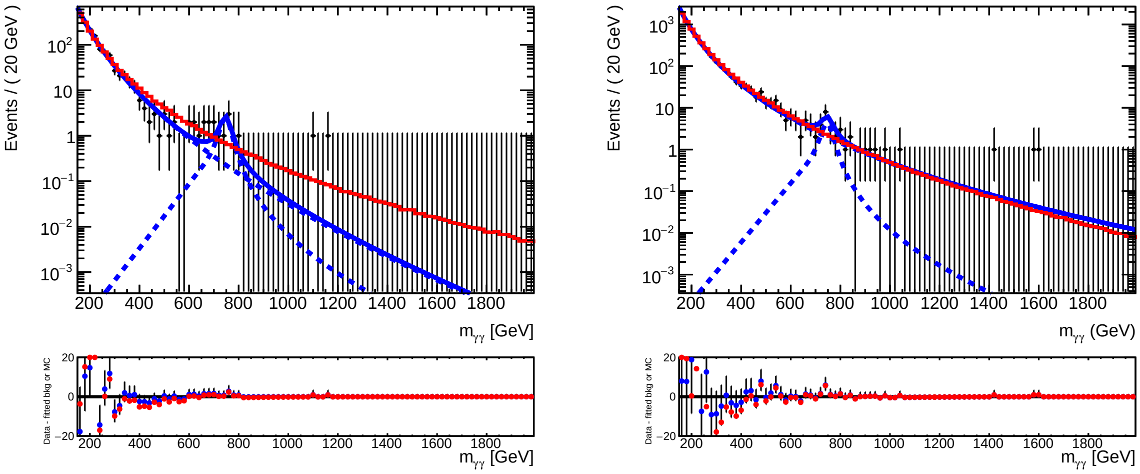


Figure 8.29: $m_{\gamma\gamma}$ distribution in data overlaid with the background template obtained from LO Sherpa $\gamma\gamma$ and the signal+background fit with functional form in blue. The left plot shows the $m_{\gamma\gamma}$ distribution in the Central-Central category whereas the distribution in the Rest category is shown on the right. The bottom panel shows the difference between the data and the background template (red) and between the data and the fit (blue). The uncertainties on the difference between the data and the background template are not plotted.

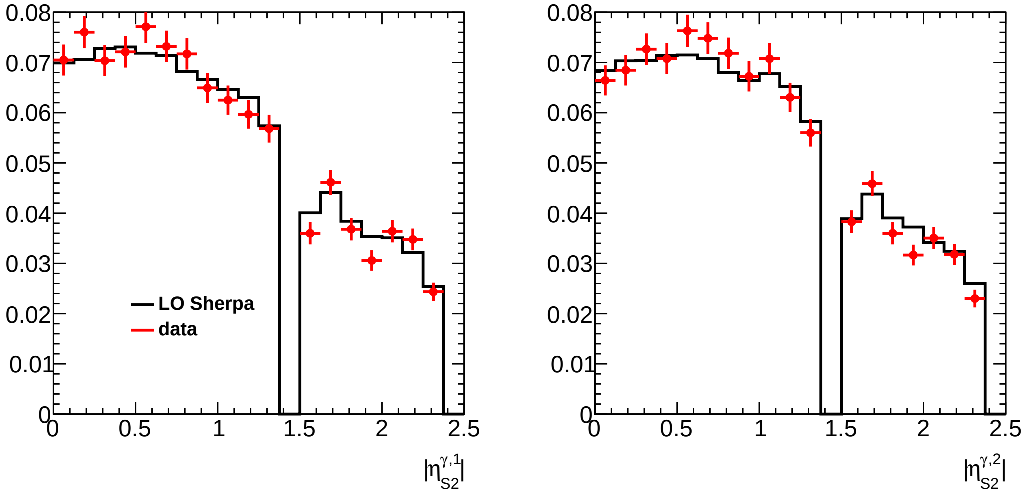


Figure 8.30: Distribution of pseudorapidity of the leading (left) and subleading (right) photon in simulation (black) compared to data (red). The events are required to pass the spin-0 selection and have invariant mass greater than 150 GeV. Both plots are normalised to unity.

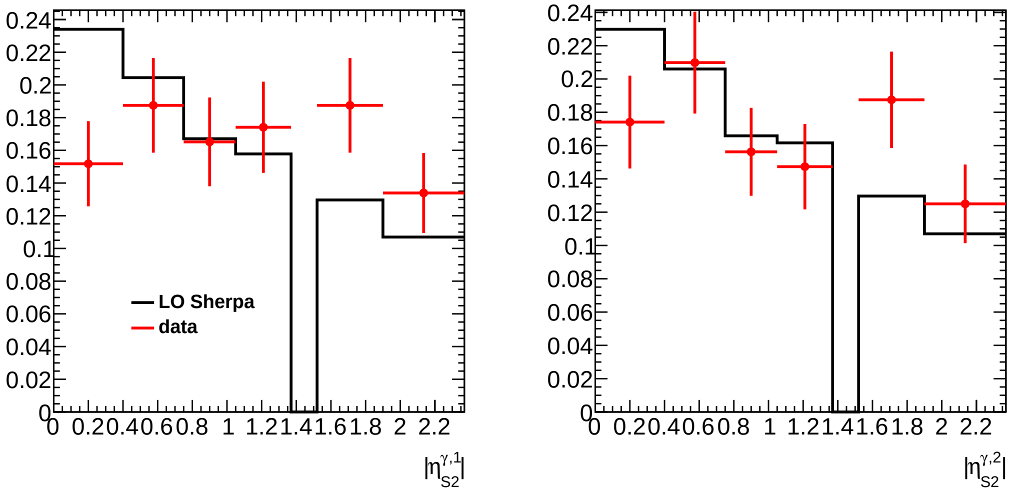


Figure 8.31: Distribution of pseudorapidity of the leading (left) and subleading (right) photon in simulation (black) compared to data (red). The events are required to pass the spin-0 selection and have invariant mass between 400 and 700 GeV. A wider binning than that in Figure 8.30 is chosen due to the low statistics in data. Both plots are normalised to unity.

8.11 Conclusion

Based on the first 3.2 fb^{-1} of 13 TeV data recorded in 2015, the search for high mass diphoton resonances is carried out. Two searches are performed: one optimised for spin-0 resonances and the other for spin-2 RS graviton. An excess is observed near a mass of 750 GeV with a best-fit width of around Γ_X/m_X=6%, corresponding to a local significance of 3.9σ and 3.8σ in the spin-0 and spin-2 searches respectively. The global significances are estimated to be 2.1σ in both searches.

Given the apparent large width of the excess, the resonance shape studies performed in this thesis are essential to extend the search to large-width resonances. A similar search for high mass diphoton resonances using the data recorded in 2015 by the CMS experiment also showed an excess in the same mass range [128]. Nonetheless, the excess seen in the CMS dataset is more compatible with the narrow-width hypothesis. The excess could well be a statistical fluctuation of the data, hence the analysis of 2016 data in the next chapter is indispensable to check if the 750 GeV excess remains.

Chapter 9

Search for Diphoton Resonances at $\sqrt{s} = 13$ TeV Updated with 2016 data

As described in the previous chapters, in the search for high mass diphoton resonances at $\sqrt{s} =$ 13 TeV with 2015 data, an excess around 750 GeV was observed. This sparked a huge interest in the high energy physics community. Such an excess could be a much sought-after sign of new physics, or simply a statistical fluctuation. In order to establish the origin of this excess, the search is updated with 2016 data which has much higher integrated luminosity. In this chapter, an update of the high mass diphoton resonance search is performed using both the 2015 and the first half of 2016 datasets as presented in the *38th International Conference on High Energy Physics* in August 2016 [129].

My main contributions to the analysis are investigating the new scalar signal model and modelling of the large-width line shape for scalar and graviton model. In the same manner as in the analysis using 2015 data, I was also responsible for the statistical analysis on the spin-0 resonance search. In addition, I also performed several checks on the functional form-based background modelling, kinematic distributions and statistical results.

The updated analysis uses similar techniques and the selection requirements are unchanged with respect to the previous analysis using 2015 data presented in Chapter 8. The changes with respect to the 2015 analysis are reported in Section 9.1. The results are discussed in Section 9.2. Checks on photon pseudorapidity distributions in data compared to simulation and the comparison of the background shapes between the two years are shown in Section 9.3.1 and 9.3.2 respectively. Finally, the upper limits computed with asymptotic formulae are checked against the results using toy Monte Carlo in Section 9.3.3.

9.1 Changes with Respect to 2015 Analysis

The analysis uses 3.2 fb^{-1} and 12.2 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV data collected by ATLAS in 2015 and first half of 2016 respectively. The measurement of the integrated luminosity has an uncertainty of $\pm 2.1\%$ for the 2015 data and a preliminary uncertainty of $\pm 3.7\%$ for the 2016 data. The preliminary uncertainty on the combined 2015+2016 integrated luminosity is $\pm 2.9\%$. These are based on the beam-separation scans performed in August 2015 and May 2016. The luminosity uncertainty for the 2015 data is improved with respect to the preliminary uncertainty quoted in Section 8.6.

Due to the higher pile-up in the 2016 data, more work is needed to complete the analysis in the extended acceptance of the spin-2 selection. The spin-2 search using 2016 data is thus not presented.

9.1.1 Changes in Reconstruction Software Release

One of the major updates with respect to the 2015 analysis is in the reconstruction software. The changes include small improvements in the reconstruction and selection of converted photons and in the photon energy calibration. The 2015 data and simulated samples presented in the previous chapter have been reprocessed with the same reconstruction software as used for the 2016 data processing.

For the converted photons, all the tracks associated to the reconstructed photon conversion vertices are excluded from the p_T^{iso} computation. This gives a better track isolation variable definition for photons which undergo significant interactions in the early part of the inner detector. It improves the efficiency for genuinely converted photons by a few percent and makes the efficiency less p_T-dependent. The same requirement on the track isolation variable ($p_T^{\text{iso}} < 0.05 \times E_T$) as used in 2015 is retained in the analysis.

The cuts used for photon conversion reconstruction were retuned following a change in the TRT particle identification tool to use the local occupancy in the detector to correct for pile-up effects. In addition, the GSF track momentum fit was changed to correct for a small bug present in the previous release. The training of the MVA energy calibration was redone to take into account these changes. The MVA energy calibration was also retrained in the crack region ($1.425 < |\eta| < 1.55$) to use consistently the crack scintillator in the calibration.

9.1.2 Pile-up

The average number of pp interactions per bunch crossing was 13.7 in 2015 but it has increased to 23.2 in 2016. The peak instantaneous luminosity in this dataset reaches 1.2×10^{34} cm^{-2}s^{-1}, which is higher than the design instantaneous luminosity, as shown in Figure 9.1.

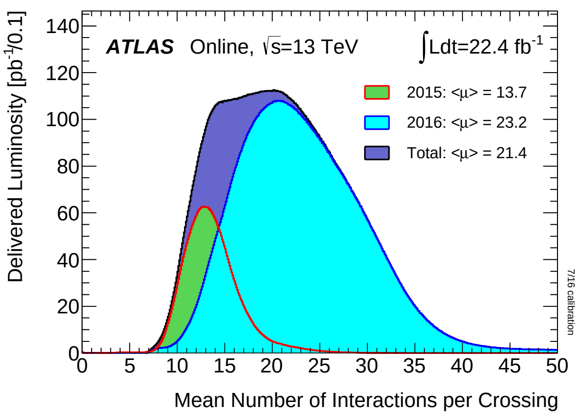


Figure 9.1: Luminosity-weighted distribution of the mean number of interactions per bunch crossing for the total 2015 pp collision data and the 2016 data recorded from 16 April - 25 July (ICHEP dataset). All data delivered to ATLAS during stable beams is shown.

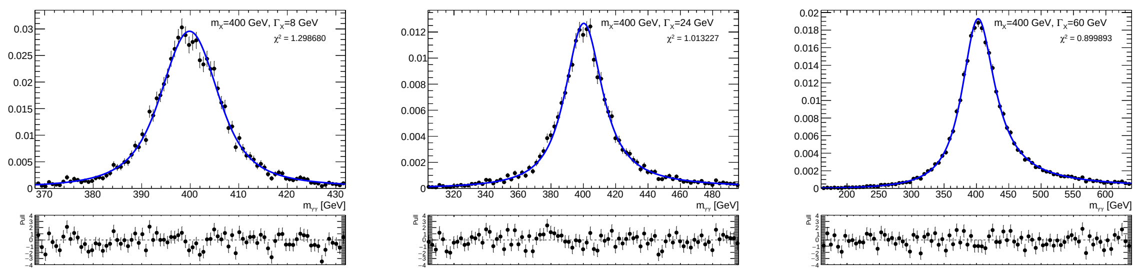
Both the photon identification and isolation requirements are slightly sensitive to the pile-up activity, with a +4% (+1.5%) expected difference for the combined identification and isolation efficiency for 100 GeV (300 GeV) E_T photons for the data taking conditions in 2015 compared to those in 2016 dataset.

9.1.3 Signal Modelling

The PowHeg Higgs-like signal model has been replaced by a scalar singlet model using an effective field theory (EFT) approach in MG5_aMC@NLO, in order to have a more model-independent description of the large-width resonance shape, as mentioned in Section 7.2.3 and

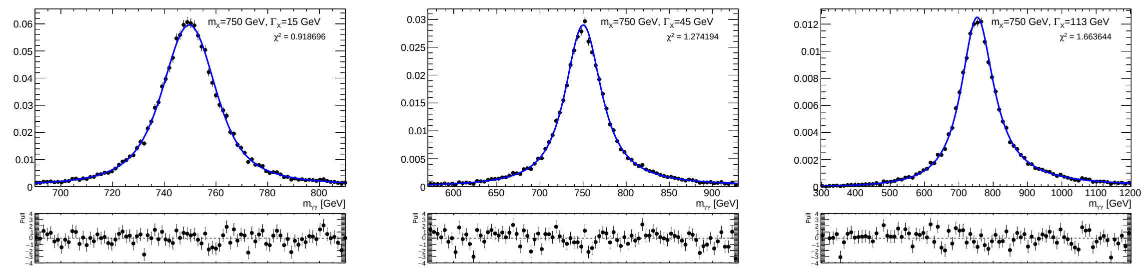
7.3.2. The comparisons of the two line shapes are shown in Figure 7.17 for $\Gamma_X/m_X = 10\%$. The asymmetry of the mass tail distribution as well as the mass peak position has changed, resulting in a slight modification to the statistical results.

The DSCB function is once again used to model the detector resolution from the NWA MG5_aMC@NLO samples. Using the same 2016 reconstruction software to process both PowHeg and MG5_aMC@NLO samples, the NWA mass distributions are very similar between the two models, as expected. Convolving the detector resolution to the MG5_aMC@NLO theoretical line shape as derived in Section 7.3.2, good agreement with the reconstructed mass distributions in the MC samples is found. The validation of the convolution on the MG5_aMC@NLO samples using the MG5_aMC@NLO EFT theoretical line shape is shown in Figures 9.2, 9.3 and 9.4 for resonance mass values of 400, 750 and 1600 GeV.



(a) (b) (c)

Figure 9.2: Reconstructed $m_{\gamma\gamma}$ distributions of a resonance with m_X=400 GeV and width of 2% (left), 6% (centre) and 15% (right) of the m_X value. The solid blue line which shows the convolution of the theoretical line shape with the detector resolution is overlaid with the reconstructed mass of the MG5_aMC@NLO sample passing spin-0 selection.

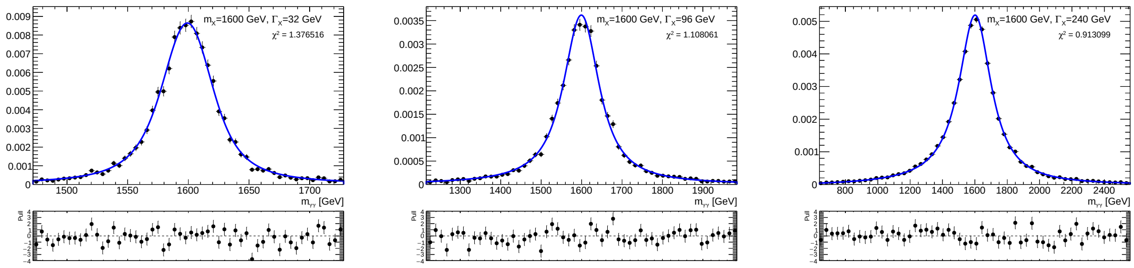


(a) (b) (c)

Figure 9.3: Reconstructed $m_{\gamma\gamma}$ distributions of a resonance with m_X=750 GeV and width of (2% (left), 6% (centre) and 15% (right) of the m_X value. The solid blue line which shows the convolution of the theoretical line shape with the detector resolution is overlaid with the reconstructed mass of the MG5_aMC@NLO sample passing spin-0 selection.

The large-width signal shape is modelled with the convolution of a DSCB function describing the detector resolution and the theoretical line shape in Eq. 7.42. The full line shape including the tail distribution is used in the statistical model. This is in contrast to the DSCB function used to model the resonant part of the PowHeg signal shape in the previous iteration of the spin-0 analysis.

Although the spin-2 analysis has not been finalised, the signal modelling was studied and changed to using the line shape parameterisation in Section 7.5. The MC samples have been reprocessed with the reconstruction software used in 2016 but no replacement to the samples



(a) (b) (c)

Figure 9.4: Reconstructed $m_{\gamma\gamma}$ distributions of a resonance with m_X=1600 GeV and width of 2% (left), 6% (centre) and 15% (right) of the m_X value. The solid blue line which shows the convolution of the theoretical line shape with the detector resolution is overlaid with the reconstructed mass of the MG5_aMC@NLO sample passing spin-0 selection.

generated with Pythia version 8.186 which has an erroneous $\hat{s}/m^2_{G^*}$ factor in $q\bar{q} \to G^*$ cross-section was made, since a flat mass sample is used in most cases.

9.1.4 Signal Efficiency

The signal efficiency has been recomputed with the MG5_aMC@NLO samples. The fiducial volume is unchanged except for the removal of the mass window for the large-width fiducial definition.

Due to the different photon isolation and identification efficiency in 2016 compared to 2015 as a result of the higher pile-up in 2016, separate C_X factors for 2015 and 2016 are computed (see Section 8.5.1 for definition). This is achieved by reweighting the pile-up profile in MC to match that in the 2015 or the 2016 data. The C_X factors are computed using all available MC samples. They are found to be similar between the narrow-width and the large-width signals. Therefore, the narrow-width C_X factors are used to derive a parameterisation as a function of the mass, as shown in Figure 9.5 for the 2016 dataset.

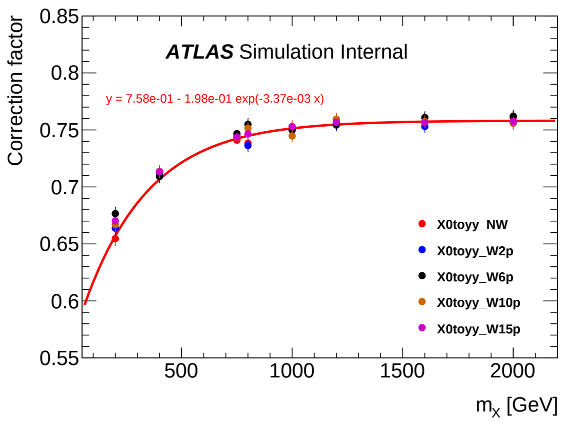


Figure 9.5: C_X factors computed using the MG5_aMC@NLO scalar singlet samples as a function of m_X for 2016 dataset. In the samples, the pile-up profile is reweighted to match that of 2016 data. The C_X factors are shown for NWA (red), $\Gamma_X/m_X = 2\%$ (blue), $\Gamma_X/m_X = 6\%$ (black), $\Gamma_X/m_X = 10\%$ (brown) and $\Gamma_X/m_X = 15\%$ (purple).

The production mode uncertainty on the C_X factors is $\pm$2.1%. For the large-width case, the small difference between the C_X factors computed at different widths is taken into account as an additional uncertainty, which has a value of ±1.9%. This uncertainty is combined in quadrature with the production mode uncertainty to obtain a theory uncertainty of ±2.8% on the C_X factor in the large-width case.

The experimental uncertainties in C_X consist of the uncertainties in the photon identification efficiency (±2 to 3% depending on the assumed mass), the photon isolation efficiency (±1 to 4% depending on the mass), the trigger efficiency (±0.6%) and the luminosity uncertainty.

9.1.5 Miscellanea

The data sample composition has been measured in the combined 2015 and 2016 dataset with similar techniques as described in Section 8.3.3. The average purity of the spin-0 selection is 90^{+2}_{-10}%. The reducible background shape has been updated with the combined 2015 and 2016 data with higher statistics resulting in a more accurate and reliable determination of the background shape.

The same functional form with two parameters described in Eq. 8.2 is chosen to model the total background. The *spurious signals* have been re-evaluated using a background template built with the updated purity, reducible and irreducible backgrounds, corresponding to an integrated luminosity of 16 fb^{-1} expected for the combined dataset. In order to satisfy the spurious signal criteria (less than 30% of the statistical uncertainty in the fitted signal yield over the full mass range), the function is fitted to data in the range $m_{\gamma\gamma} > 180$ GeV. For the analysis of 2015 data alone, the fit range starts at 150 GeV, as in the analysis described in Chapter 8, in order to increase the number of events in the low invariant mass sideband and improve the precision of the fit results. For the analysis of the combined data, the 2015 and 2016 data are fitted simultaneously from 180 GeV.

The search range is defined to be 200-2400 GeV in mass m_X and 0-10% in relative width Γ_X/m_X. The upper search range is extended as higher diphoton mass events are expected from the higher integrated luminosity.

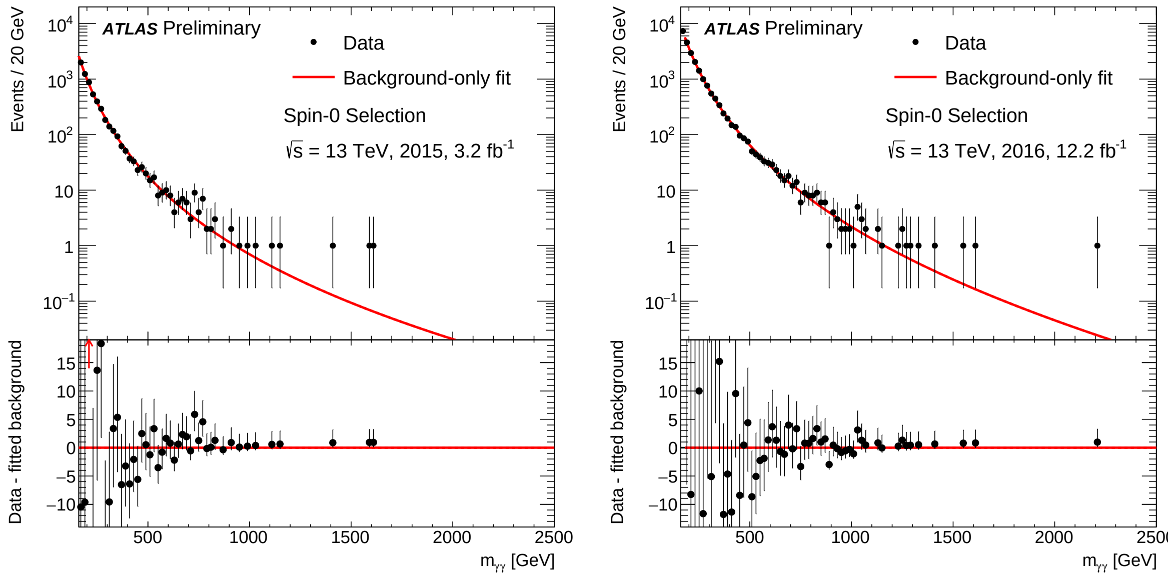
9.2 Results

The 2015 data reprocessed with the improved photon reconstruction algorithm have been re-analysed with the above changes. This results in the local significance of the largest excess decreasing from 3.9σ to 3.4σ. Due to the change in signal model and the migrations of events in the invariant mass distribution, the corresponding signal mass and width with largest deviation from background-only hypothesis have also changed, from a mass of 750 GeV and a relative width of 6% to a mass of 730 GeV and a relative width of 8%.

The decrease in the local significance of the excess is mainly due to two events affected by the new reconstruction algorithms. In one of these two selected event, one of the two photon candidates is at $|\eta| = 1.53$, where the improved calibration of photon candidates near the transition region of the electromagnetic calorimeter changes the diphoton invariant mass from 757 GeV to 722 GeV. In the second event, one track previously associated to one of the two selected photon candidates is not considered as originating from the photon conversion due to the change in the photon conversion reconstruction. Therefore, the track enters the calculation of the track isolation. This photon candidate thus fails the track isolation requirement and the event no longer passes the selection.

The diphoton invariant mass distribution together with the background-only fit is shown in Figure 9.6a for events selected in the reprocessed 2015 dataset and in Figure 9.6b for the 2016 dataset. The 2015 and 2016 datasets are combined, and the diphoton invariant mass distribution of the selected events is shown in Figure 9.7.

In the 2016 data, no significant deviation from the background-only hypothesis is observed near 730 GeV, the mass of the most significant excess in the 2015 data. The diphoton mass



(a) (b)

Figure 9.6: Invariant-mass distribution of the selected diphoton candidates, with the background-only fit overlaid, for (a) 2015 data and (b) 2016 data. The difference between the data and this fit is shown in the bottom panel. The arrow shown in the lower panel indicates a value outside the range with more than one standard deviation. There is no data event with $m_{\gamma\gamma} > 2500$ GeV.

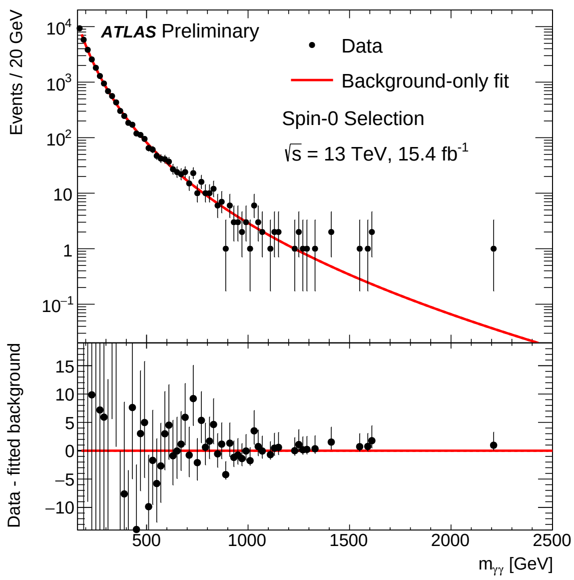


Figure 9.7: Invariant-mass distribution of the selected diphoton candidates, with the background-only fit overlaid for the combined 2015 and 2016 dataset totalling 15.4 fb^{-1}. The difference between the data and this fit is shown in the bottom panel. There is no data event with $m_{\gamma\gamma} > 2500$ GeV.

spectrum for the events selected in the 2016 data is shown in Figure 9.6b.

The compatibility between the 2016 data and the best fit signal associated to the largest excess in the 2015 data is evaluated by assuming a signal mass and width determined from the smallest local p-value in the 2015 data. The compatibility is determined using the difference between the signal cross-sections extracted from the 2015 and 2016 datasets $\Delta\sigma = \sigma_{2015} - \sigma_{2016}$ as parameter of interest. The two datasets are compatible at the level of 2.7 σ.

The local p-values testing the compatibility of the combined dataset with the background-only hypothesis are computed as a function of the signal mass m_X and relative width Γ_X/m_X in the search range. The local significance as a function of mass and relative width is illustrated in Figure 9.8. Figure 9.9 shows the local p-value as a function of the assumed mass for NWA, $\Gamma_X/m_X = 2, 6$ and 10% hypotheses, comparing the results observed with the 2015 dataset only, the 2016 dataset only and the combined dataset.

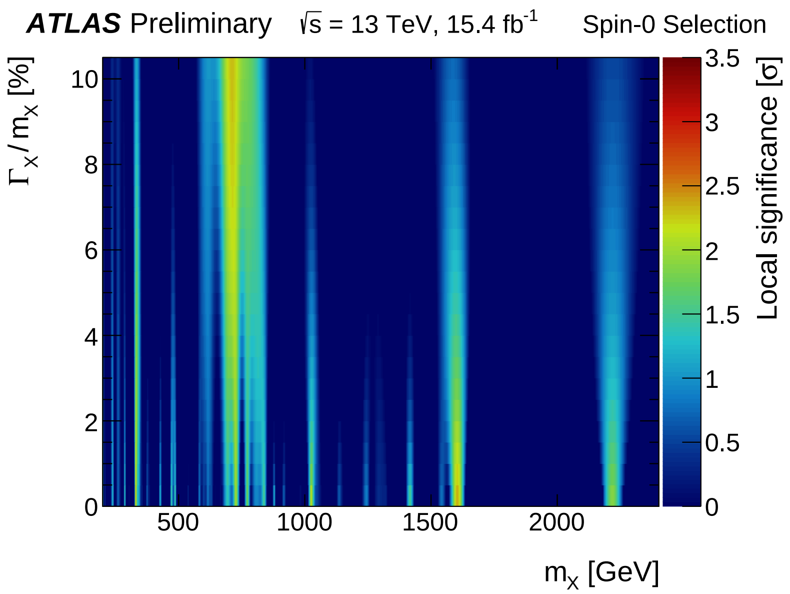
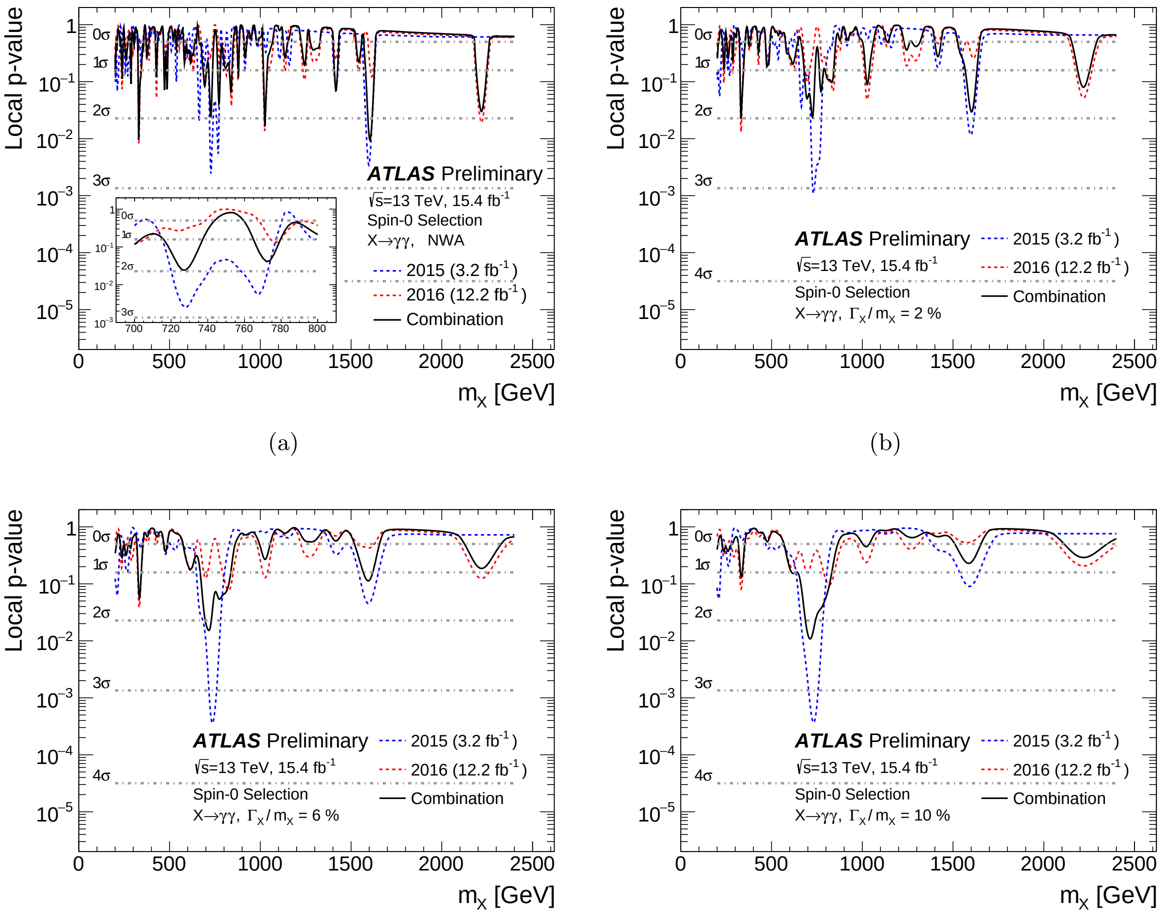


Figure 9.8: Compatibility, in terms of local significance with the background-only hypothesis as a function of the assumed signal mass m_X and relative width Γ_X/m_X for the spin-0 resonance search using combined data.

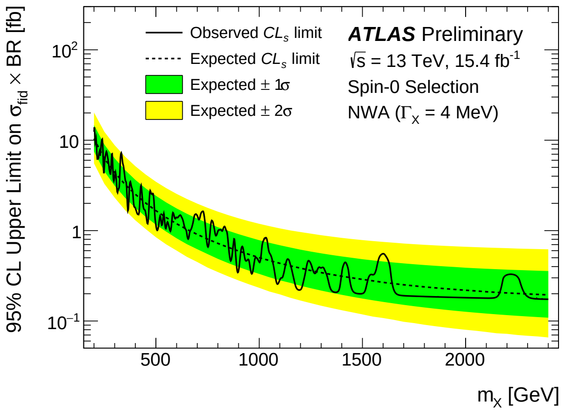
The largest deviation over the background-only hypothesis in the combined dataset is observed at a mass of 1600 GeV for the narrow-width hypothesis. This corresponds to a local significance of 2.4 σ, or a global significance of less than 1 σ. In the 700-800 GeV mass range where the excess was seen in the 2015 data, the largest local significance is 2.3 σ for a mass near 710 GeV and $\Gamma_X/m_X = 10\%$.

Figure 9.10 shows the limits on the signal fiducial cross-section times branching ratio to two photons for a spin-0 resonance as a function of the hypothesised mass and for four different values of signal width using the combined dataset. Two-dimensional plots of observed and expected upper limits as a function of the resonance mass m_X and the relative width Γ_X/m_X can be found in Appendix D. In the same appendix, the comparison of the observed limits as a function of mass using the 2015, 2016 and combined dataset is also shown. The asymptotic approximation is used to derive the limits (see Section 5.5).

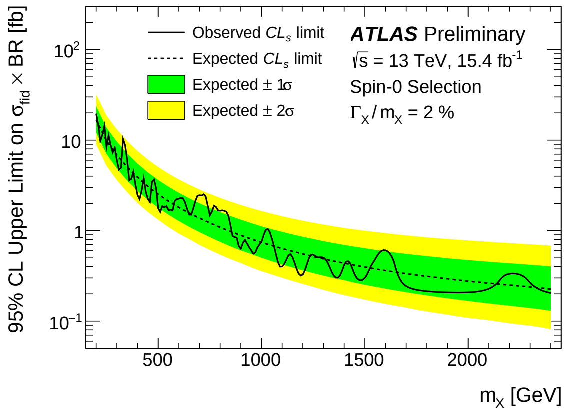


(a) (b) (c) (d)

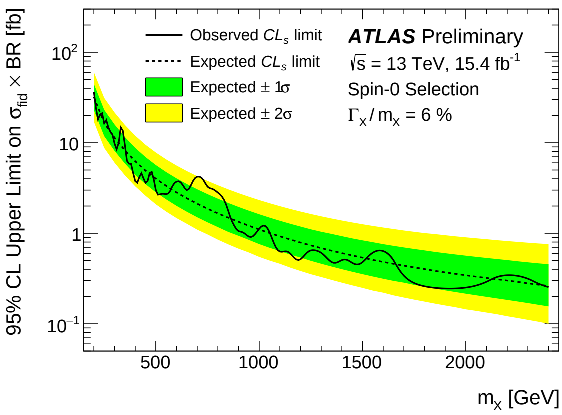
Figure 9.9: Compatibility with the background-only hypothesis as a function of the assumed signal mass for different values of relative width for the spin-0 resonance search. The p-value using 2015 (2016) data is shown in dotted blue (red) line and the p-value using combined 2015 and 2016 dataset is shown in black line. In (a) a narrow-width signal, with $\Gamma_X = 4$ MeV, is assumed and the inset shows the p-values for m_X between 700 and 800 GeV.



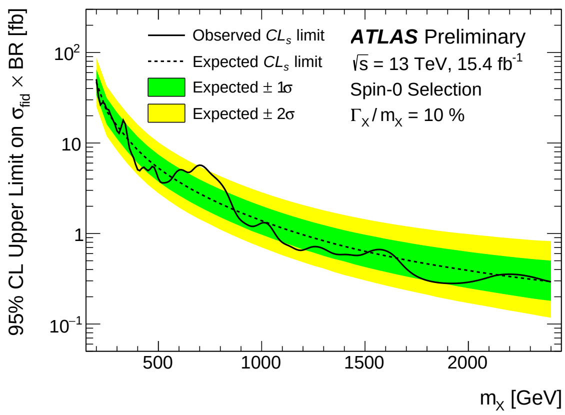
(a)



(b)



(c)



(d)

Figure 9.10: 95% CL upper limits on the fiducial cross-section times branching ratio to two photons of a spin-0 resonance as a function of the assumed mass, for different values of the decay width divided by the mass, using the combined dataset.

9.3 Additional Checks

9.3.1 Checks on the Photon Pseudorapidity

The photon pseudorapidity distribution in 2015 data shows a discrepancy with the LO Sherpa prediction, as shown in Section 8.10.4. The check on photon pseudorapidity is repeated with the higher statistics of the 2016 run. Figure 9.11 shows the comparison between the distribution of the leading and the subleading photon pseudorapidity in 2016 data compared to 2015 data and simulation. The distributions in 2015 data and simulation are similar to what shown in Figure 8.31 except for the update in reconstruction software. The discrepancy at large $|\eta|$ is much less evident in the 2016 data.

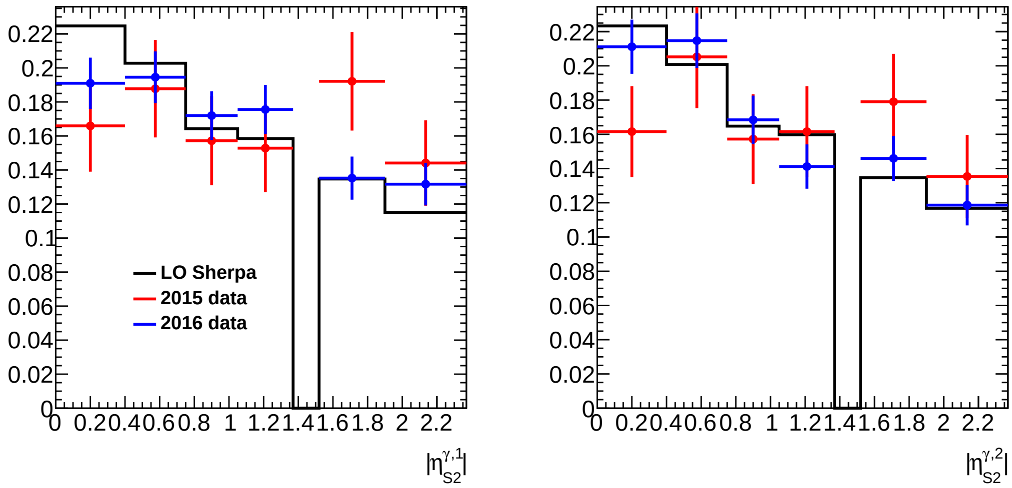


Figure 9.11: Distribution of pseudorapidity of the leading (left) and subleading (right) photon in simulation (black) compared to 2015 data (red) and 2016 data (blue). The events are required to pass the spin-0 selection and have invariant mass between 400 and 700 GeV. Both plots are normalised to unity.

Figure 9.12 shows the comparison of the $m_{\gamma\gamma}$ distribution in 2016 data to LO SHERPA $\gamma\gamma$ prediction in the Central-Central and Rest category (see Section 8.10.4). The agreement between the MC prediction and the data is much better than what was observed in Figure 8.29.

9.3.2 Checks on the Background Modelling

The background functional forms obtained from 2015 and 2016 data are compared. The background shape should be similar between the two years as the same reconstruction algorithm is used to process the two datasets. The higher pile-up in 2016 could induce some slight reduction in the efficiency for lower p_T photons.

The background shapes from the functional form fit to the two datasets are compared. Figure 9.13 shows the comparison of the background function as fitted on the 2015 data and on the 2016 data. A background-only fit is performed on the 2016 data since there is no significant excess in the dataset. The background shape obtained from 2015 data differs whether a background-only fit or a signal+background fit is performed, due to the presence of an excess near 750 GeV. Both types of fit are hence shown. The statistical uncertainty on the background shape originates from the uncertainty in the determination of the parameters of the function from the fit to the data. The statistical uncertainty band is obtained from an ensemble of toy experiments. For each toy, the background normalisation and shape parameters are drawn randomly from a multivariate Gaussian distribution with the central values defined by the best-

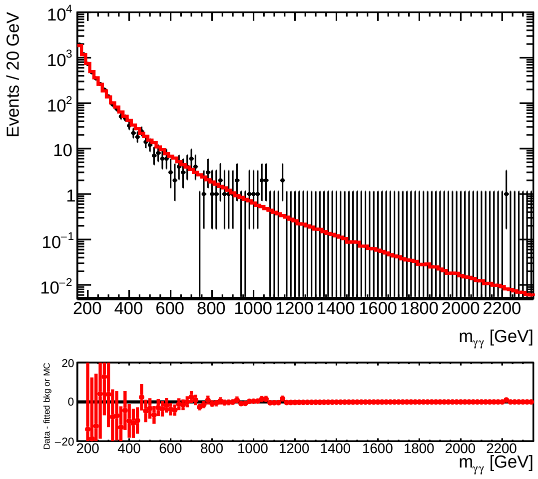
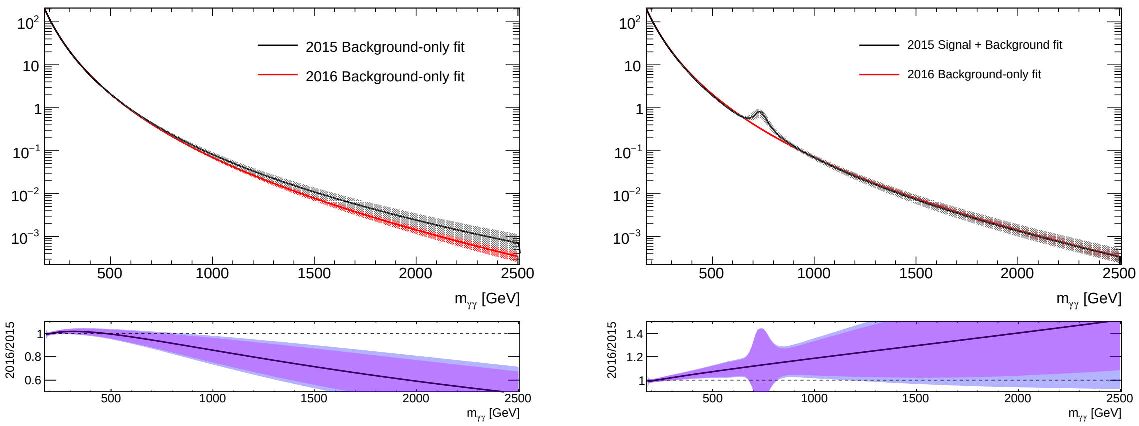


Figure 9.12: $m_{\gamma\gamma}$ distribution in 2016 data overlaid with the background template obtained from LO SHERPA $\gamma\gamma$ (red histogram). The left plot shows the $m_{\gamma\gamma}$ distribution in the Central-Central category whereas the distribution in the Rest category is shown on the right. The bottom panel shows the difference between the data and the background template.

fit values in data, and the widths defined by the covariance matrix of the fit. In each mass bin, the uncertainty band is then defined as an interval containing 68% of the toy experiments.



(a) (b)

Figure 9.13: Comparison of the background function as fitted on the 2015 data (black) and on the 2016 data (red). In (a), a background-only fit is performed on 2015 data and in (b), a signal+background fit is performed on the 2015 data. The fit range for 2015 data is $m_{\gamma\gamma} >$ 150 GeV while it is $m_{\gamma\gamma} >$ 180 GeV for 2016 data. The 2016 background function is normalised to the background function of 2015 data in the range $m_{\gamma\gamma} >$ 180 GeV in order to compare the two background shapes. The bottom panel shows the ratio of the background function fitted on 2016 data to the 2015 background function. The uncertainty band on the ratio is calculated from the quadratic sum of the statistical uncertainties on the two background functions.

From the figure, one observes that the 2015 background shape as estimated from the background-only fit does not agree well with the 2016 data-driven background shape due to the bias from the 750 GeV excess. When performing a signal+background fit, the background component in 2015 then agrees better with the background shape in 2016 data.

9.3.3 Checks on the Asymptotic Limits

Upper limits are set on the fiducial cross-section times branching ratio to two photons as a function of mass up to 2.4 TeV, at which point there is no observed event. Above 1.5 TeV, the diphoton mass distribution is sparsely populated and the asymptotic approximation breaks down. This results in better limits obtained using the asymptotic formulae than expected in the high mass range, especially for the NWA case where the limits are more stringent than for the larger widths.

When there is zero number of observed events, one expects to set an upper limit at about 3 events, based on Poisson distribution:

$$\mathcal{P}(n|\mu) = \frac{\mu^n e^{-\mu}}{n!}, \tag{9.1}$$

with n=0. Solving $\mathcal{P}(n = 0|\mu) = 5\%$ for μ gives $\mu = -\ln(0.05) = 2.99$.

With the asymptotic formulae, the expected (observed) upper limit on the fiducial cross-section times branching ratio to diphoton for the combined 2015+2016 dataset is 1.66 (1.14) fb at 500 GeV, 0.50 (0.51) fb at 1000 GeV and 0.22 (0.18) fb at 2000 GeV for the case of NWA. This is equivalent to a limit on the number of signal events of 18.49 (12.71) at 500 GeV, 5.80 (5.92) at 1000 GeV and 2.56 (2.12) at 2000 GeV. The upper limits obtained at $m_X = 2000$ GeV appear to be better than what is expected from Poisson statistics.

When the number of observed events is low, one should compute the upper limits using toy Monte Carlo experiment instead of using the asymptotic formulae. This is however computationally expensive and not feasible for a fine scan in mass and width such as that performed in this analysis. Thus, the upper limits at three mass points (500 GeV, 1000 GeV and 2000 GeV) for a narrow-width hypothesis are checked using toy Monte Carlo to determine how much bias is induced by using the asymptotic approximation.

The following procedure is used to compute the expected and observed upper limits using toy Monte Carlo experiment for a given resonance mass and width:

- For each μ (or fiducial cross-section times branching ratio σ_{fid}, which is the parameter of interest in the analysis) hypothesis:
  - Evaluate the observed test statistic $\tilde{q}_{\mu,obs}$.
  - Construct $f(\tilde{q}_\mu|\mu)$ (or preferably $f(\tilde{q}_\mu|\mu, \hat{\hat{\theta}}(\mu, obs)))$, the pdf of observing $\tilde{q}_\mu$ by generating signal+background pseudo-experiments with signal strength μ and calculating $\tilde{q}_\mu$.[1]
  - Construct $f(\tilde{q}_\mu|0)$ (or preferably $f(\tilde{q}_\mu|0, \hat{\hat{\theta}}(0, obs)))$, the pdf of observing $\tilde{q}_\mu$ by generating background-only pseudo-experiments and calculating $\tilde{q}_\mu$.
  - From the constructed distributions, calculate

$$p_\mu = \int_{\tilde{q}_{\mu,obs/exp}}^{\infty} f(\tilde{q}_\mu|\mu) d\tilde{q}_\mu \tag{9.2}$$

  and

$$p_b = \int_{\tilde{q}_{\mu,obs/exp}}^{\infty} f(\tilde{q}_\mu|0) d\tilde{q}_\mu. \tag{9.3}$$

  where $\tilde{q}_{\mu,exp}$ is defined as the median of $f(\tilde{q}_\mu|0)$.

- For the observed limit, find the value of $\mu_{up}(obs)$ which satisfies $p^{CL_s}_{\mu_{up}} = p_\mu/p_b = 5\%$.

[1] In this study, the pseudo-experiments are generated with the pre-fit nuisance parameters. Ideally, they should be with the conditional MLE of the nuisance parameter based on the observed data $\hat{\hat{\theta}}(\mu, obs)$ including correlations.

- For the expected limit, generate background-only pseudo-experiments and find the value of $\mu_{up}(exp)$ which satisfies $p^{CL_s}_{\mu_{up}} = p_\mu / p_b = 5\%$. The median and the $\pm1\sigma$ and $\pm2\sigma$ bands can be evaluated from the distribution of $\mu_{up}(exp)$.

- Alternatively for the expected limit, find the value of $\mu_{up}(exp)$ and $\mu_{up}(exp \pm 1, 2\sigma)$ which satisfies $p^{CL_s}_{\mu_{up}} = p_\mu / p_b = 5\%$ with the $\tilde{q}_{\mu,exp}$ and $\tilde{q}_{\mu,\pm1,2\sigma}$ obtained from the median and quantiles of the $f(\tilde{q}_\mu|0)$ distribution.

Figure 9.14 shows the $f(\tilde{q}_\mu|\mu)$ and $f(\tilde{q}_\mu|0)$ distributions testing $\mu = \sigma_{fid} = 1.5$ fb. 20,000 pseudo-experiments are generated for each distribution.

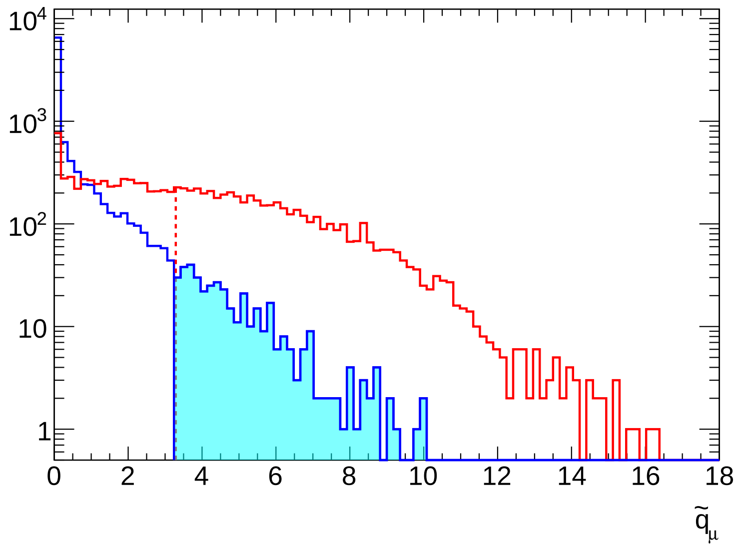


Figure 9.14: Distribution of test statistic $f(\tilde{q}_\mu|\mu)$ for $\mu = \sigma_{fid} = 1.5$ fb in blue and $f(\tilde{q}_\mu|0)$ in red. The median of $f(\tilde{q}_\mu|0)$ is shown at $\tilde{q}_\mu = 3.3$. The shaded area corresponds to p_μ.

The upper limits obtained using the sampling distributions ($f(\tilde{q}_\mu|\mu)$ and $f(\tilde{q}_\mu|0)$) generated with 20,000 pseudo-experiments are compared to those computed using the asymptotic approximation in Table 9.1. For each mass value, 10 μ hypotheses are tested in the evaluation of the upper limits.

m_X [GeV]	Asymptotic/Toy	$\sigma^{obs}_{95\%}$ [fb]	$\sigma^{exp}_{95\%}$ [fb]	$\pm1\sigma$	$\pm2\sigma$
500	Asymptotic	1.14	1.66	1.16, 2.42	0.84, 3.47
500	Toy	1.23	1.67	1.21, 2.40	1.05, 3.35
1000	Asymptotic	0.45	0.50	0.33, 0.78	0.23, 1.19
1000	Toy	0.46	0.51	0.37, 0.74	0.34, 1.07
2000	Asymptotic	0.18	0.22	0.13, 0.39	0.08, 0.66
2000	Toy	0.25	0.26	0.25, 0.30	0.25, 0.43

Table 9.1: Comparison of the observed and expected 95% CL upper limits on the fiducial cross-section times branching ratio to two photons using asymptotic approximation and that calculated with toy Monte Carlo method. The comparison is shown for $m_X = 500, 1000$ and 2000 GeV with the narrow-width hypothesis. The $\pm1\sigma$ and $\pm2\sigma$ entries correspond to the fluctuations in the expected limit, the first (second) value corresponding to the lower (higher) value defining the $\pm1\sigma$ or $\pm2\sigma$ interval.

The asymptotic approximation is valid for the m_X=500 and 1000 GeV hypotheses where the observed and median expected limits agree with those obtained from toy Monte Carlo. The

$\pm 1\sigma$ and $\pm 2\sigma$ error bands on the expected limits using the asymptotic approximation differ slightly from those obtained with sampling distributions. For the $m_X = 2000$ GeV hypothesis where the asymptotic approximation is not expected to be applicable due to the low statistics, large variations between the two results are seen. The asymptotic approximation is biased as it yields an upper limit that is too strict. This bias is negligible at low mass but increases at high mass. In particular, the limits obtained with toy Monte Carlo do not go below 0.25 fb, which is the value expected with no event, while the observed limit using the asymptotic formulae is at 0.18 fb. The upper limits obtained using the asymptotic approximation are biased by up to 30% at high mass.

9.4 Conclusion

The search for high mass diphoton resonances is updated using 12.2 fb^{-1} of data collected up until July 2016. The result was presented at the ICHEP conference [129]. The 2016 dataset is combined with the 2015 data, yielding 15.5 fb^{-1} in total. Only the spin-0 resonance search is presented. The excess observed in the 2015 data near 750 GeV has a lower significance after the dataset is reprocessed. No excess is observed in the 2016 dataset near this mass value, implying that the excess seen in the 2015 data is likely to be a statistical fluctuation.

A comparison of the photon pseudorapidity distributions has been performed and the distribution in the 2016 data does not appear to be very different from the MC prediction. The background $m_{\gamma\gamma}$ shape as estimated from data is compatible between the 2015 and the 2016 dataset. Upper limits on the fiducial cross-section are checked for the NWA hypothesis and a bias up to 30% is found at high mass from the use of asymptotic approximation where the approximation is not valid.

Conclusion

After the impressive success of the LHC Run 1 leading to the discovery of a new boson consistent with the Higgs boson predicted by the Standard Model, the shutdown period presents an opportunity for consolidation. During this period, upgrade and maintenance work on the detector were carried out, the software and reconstruction of physics objects were improved, and many studies were performed in preparation of Run 2 and beyond in light of the new challenges expected in the next runs. In this thesis, studies on an alternative method to classify photon conversions are performed and its performance with respect to the current method in a high pile-up environment assessed. The alternative method based on the electromagnetic shower of the photon candidates is found to be pile-up robust. It may potentially replace or complement the current track-based method in the coming years when the pile-up rate is expected to be high. The performance of the track-based conversion reconstruction in 2012 is also measured with a similar classification method. It yields a more precise measurement on the efficiency than previously due to the increased separation of true converted and unconverted photons provided by the method.

In parallel, a study of the feasibility of the search for a pseudoscalar A boson as predicted in two-Higgs-doublet models decaying to Zh in the $jj\gamma\gamma$ final state is performed using the data collected in 2012 in order to assess its sensitivity compared to the other search channels in ATLAS. In this search, h is assumed to be the 125 GeV Higgs boson discovered in 2012. The study concludes that the $jj\gamma\gamma$ channel is competitive with the $ll\tau\tau$ final states but less sensitive to searches with bb final states. As such, the search is not pursued.

The start of the LHC data taking at a centre-of-mass energy of 13 TeV in 2015 signified a new frontier for beyond the SM physics searches. Making use of the unprecedented collision energy, the search for high mass diphoton resonances is carried out. The preliminary result using the full 2015 dataset was presented in a seminar at CERN in December 2015, then updated with a publication in June 2016. In the 2015 data, an excess is observed near a mass of 750 GeV with a best-fit width of around 45 GeV. Similarly, CMS also reported an excess in the same mass range, albeit with a smaller best-fit width. This prompted considerable interest in the scientific community and over 500 papers with possible explanations appeared in the course of a few months after the seminar. The 2015 data statistics is equivalent to only a quarter of the data collected in 2016. The high mass diphoton resonances search is updated using the data collected up until July 2016 and the result is presented at the ICHEP conference in August 2016. No excess is observed in the 2016 dataset near 750 GeV, suggesting that the diphoton excess seen in 2015 data is a statistical fluctuation. 95% CL exclusion limits are derived on the production cross-section times branching ratio to two photons as a function of the resonance mass and width. For a spin-0 resonance, the fiducial cross-section calculated with the asymptotic approximation ranges from 12 fb (51 fb) at 200 GeV to 0.2 fb (0.3 fb) at 2400 GeV, with the narrow (10% of the mass) width hypothesis.

The Standard Model is known to be incomplete. While no significant sign of new physics has shown up, the LHC Run 2 is still ongoing. The data analysed for ICHEP is only a third of the total delivered data in 2016, and the remainder of Run 2 will yield even more — over 100 fb^{-1} is expected. There is still a vast uncharted territory to be explored at the LHC.

Bibliography

[1] G. 't Hooft, "Why Do We Need Local Gauge Invariance in Theories With Vector Particles? An Introduction," *NATO Sci. Ser. B* **59** (1980) 101–115.

[2] A. Pich, "Aspects of quantum chromodynamics," in *Proceedings, Summer School in Particle Physics: Trieste, Italy, June 21-July 9, 1999*, pp. 53–102. 1999. `arXiv:hep-ph/0001118 [hep-ph]`.

[3] S. L. Glashow, "Partial Symmetries of Weak Interactions," *Nucl. Phys.* **22** (1961) 579–588.

[4] S. Weinberg, "A Model of Leptons," *Phys. Rev. Lett.* **19** (1967) 1264–1266.

[5] A. Salam, "Weak and Electromagnetic Interactions," *Conf. Proc.* **C680519** (1968) 367–377.

[6] A. Pich, "The Standard model of electroweak interactions," in *High-energy physics. Proceedings, European School, Aronsborg, Sweden, June 18-July 1, 2006*, pp. 1–49. 2007. `arXiv:0705.4264 [hep-ph]`. [,1(2007)].

[7] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons," *Phys. Rev. Lett.* **13** (1964) 321–323.

[8] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons," *Phys. Rev. Lett.* **13** (1964) 508–509.

[9] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global Conservation Laws and Massless Particles," *Phys. Rev. Lett.* **13** (1964) 585–587.

[10] **MuLan**, D. B. Chitwood *et al.*, "Improved measurement of the positive muon lifetime and determination of the Fermi constant," *Phys. Rev. Lett.* **99** (2007) 032001, `arXiv:0704.1981 [hep-ex]`.

[11] P. Langacker, "Introduction to the Standard Model and Electroweak Physics," in *Proceedings of Theoretical Advanced Study Institute in Elementary Particle Physics on The dawn of the LHC era (TASI 2008): Boulder, USA, June 2-27, 2008*, pp. 3–48. 2010. `arXiv:0901.0241 [hep-ph]`.

[12] **ATLAS, CMS**, G. Aad *et al.*, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV," *JHEP* **08** (2016) 045, `arXiv:1606.02266 [hep-ex]`.

[13] J. Ellis and T. You, "Updated Global Analysis of Higgs Couplings," *JHEP* **06** (2013) 103, `arXiv:1303.3879 [hep-ph]`.

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

[15] **LHC Higgs Cross Section Working Group**, S. Heinemeyer, C. Mariotti, G. Passarino, and R. Tanaka (Eds.), "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties," *CERN-2013-004* (CERN, Geneva, 2013) , `arXiv:1307.1347 [hep-ph]`.

[16] **LHC Higgs Cross Section Working Group**, D. de Florian *et al.*, "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector," `arXiv:1610.07922 [hep-ph]`.

[17] **ATLAS, CMS**, G. Aad *et al.*, "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) 191803, `arXiv:1503.07589 [hep-ex]`.

[18] **Super-Kamiokande**, Y. Fukuda *et al.*, "Evidence for oscillation of atmospheric neutrinos," *Phys. Rev. Lett.* **81** (1998) 1562–1567, `arXiv:hep-ex/9807003 [hep-ex]`.

[19] S. Boyd, "Neutrino Mass and Direct Measurements." `https://www2.warwick.ac.uk/fac/sci/physics/staff/academic/boyd/stuff/lec_neutrinomass_writeup.pdf`.

[20] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, M. Sher, and J. P. Silva, "Theory and phenomenology of two-Higgs-doublet models," *Phys. Rept.* **516** (2012) 1–102, `arXiv:1106.0034 [hep-ph]`.

[21] S. L. Glashow and S. Weinberg, "Natural Conservation Laws for Neutral Currents," *Phys. Rev.* **D15** (1977) 1958.

[22] E. A. Paschos, "Diagonal Neutral Currents," *Phys. Rev.* **D15** (1977) 1966.

[23] **ATLAS Collaboration**, "Beyond Standard Model Higgs boson searches at a High-Luminosity LHC with ATLAS,". `https://cds.cern.ch/record/1611190`.

[24] C.-N. Yang, "Selection Rules for the Dematerialization of a Particle Into Two Photons," *Phys. Rev.* **77** (1950) 242–245.

[25] L. D. Landau, "On the angular momentum of a system of two photons," *Dokl. Akad. Nauk Ser. Fiz.* **60** no. 2, (1948) 207–209.

[26] A. Hill and J. J. van der Bij, "STRONGLY INTERACTING SINGLET - DOUBLET HIGGS MODEL," *Phys. Rev.* **D36** (1987) 3463–3473.

[27] M. J. G. Veltman and F. J. Yndurain, "RADIATIVE CORRECTIONS TO W W SCATTERING," *Nucl. Phys.* **B325** (1989) 1–17.

[28] T. Binoth and J. J. van der Bij, "Influence of strongly coupled, hidden scalars on Higgs signals," *Z. Phys.* **C75** (1997) 17–25, `arXiv:hep-ph/9608245 [hep-ph]`.

[29] R. M. Schabinger and J. D. Wells, "A Minimal spontaneously broken hidden sector and its impact on Higgs boson physics at the large hadron collider," *Phys. Rev.* **D72** (2005) 093007, `arXiv:hep-ph/0509209 [hep-ph]`.

[30] B. Patt and F. Wilczek, "Higgs-field portal into hidden sectors," `arXiv:hep-ph/0605188 [hep-ph]`.

[31] G. M. Pruna and T. Robens, "Higgs singlet extension parameter space in the light of the LHC discovery," *Phys. Rev.* **D88** no. 11, (2013) 115012, `arXiv:1303.1150 [hep-ph]`.

[32] T. D. Lee, "A Theory of Spontaneous T Violation," *Phys. Rev.* **D8** (1973) 1226–1239. [,516(1973)].

[33] R. Barbieri and R. Torre, "Signals of single particle production at the earliest LHC," *Phys. Lett.* **B695** (2011) 259–263, `arXiv:1008.5302 [hep-ph]`.

[34] L. Randall and R. Sundrum, "A Large mass hierarchy from a small extra dimension," *Phys. Rev. Lett.* **83** (1999) 3370–3373, `arXiv:hep-ph/9905221 [hep-ph]`.

[35] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali, "The Hierarchy problem and new dimensions at a millimeter," *Phys. Lett.* **B429** (1998) 263–272, `arXiv:hep-ph/9803315 [hep-ph]`.

[36] A. Strumia, "Interpreting the 750 GeV digamma excess: a review," 2016. `arXiv:1605.09401 [hep-ph]`.

[37] L. R. Evans and P. Bryant, "LHC Machine," *J. Instrum.* **3** (2008) S08001. 164 p. `https://cds.cern.ch/record/1129806`. This report is an abridged version of the LHC Design Report (CERN-2004-003).

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

[39] **CMS Collaboration**, S. Chatrchyan *et al.*, "The CMS experiment at the CERN LHC," *JINST* **3** (2008) S08004.

[40] **LHCb Collaboration**, A. A. Alves, Jr. *et al.*, "The LHCb Detector at the LHC," *JINST* **3** (2008) S08005.

[41] **ALICE Collaboration**, K. Aamodt *et al.*, "The ALICE experiment at the CERN LHC," *JINST* **3** (2008) S08002.

[42] W. J. Stirling. Private communication.

[43] **ATLAS Collaboration**, M. Capeans, G. Darbo, K. Einsweiller, M. Elsing, T. Flick, M. Garcia-Sciveres, C. Gemme, H. Pernegger, O. Rohne, and R. Vuillermet, "ATLAS Insertable B-Layer Technical Design Report,". `https://cds.cern.ch/record/1291633`.

[44] **ATLAS Collaboration**, G. Aad *et al.*, "The ATLAS Inner Detector commissioning and calibration," *Eur. Phys. J.* **C70** (2010) 787–821, `arXiv:1004.5293 [physics.ins-det]`.

[45] **ATLAS Collaboration**, ATLAS Collaboration, "Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s} = 13$ TeV,". `https://cds.cern.ch/record/2037683`.

[46] **ATLAS Collaboration**, "ATLAS liquid argon calorimeter: Technical design report,". `https://cds.cern.ch/record/331061`.

[47] **ATLAS Collaboration**, "ATLAS tile calorimeter: Technical Design Report,". `https://cds.cern.ch/record/331062`.

[48] **ATLAS Collaboration**, "ATLAS muon spectrometer: Technical design report,". `https://cds.cern.ch/record/331068`.

[49] **ATLAS Collaboration**, J. P. Badiou, J. Beltramelli, J. M. Baze, and J. Belorgey, "ATLAS barrel toroid: Technical Design Report,". `https://cds.cern.ch/record/331065`.

[50] **ATLAS Collaboration**, "ATLAS endcap toroids: Technical design report,". `https://cds.cern.ch/record/331066`.

[51] **ATLAS TDAQ**, "The ATLAS Data Acquisition and High Level Trigger system," *JINST* **11** no. 06, (2016) P06008.

[52] **ATLAS Collaboration**, G. Aad *et al.*, "Performance of the ATLAS Trigger System in 2010," *Eur. Phys. J.* **C72** (2012) 1849, `arXiv:1110.1530 [hep-ex]`.

[53] **ATLAS Collaboration**, "ATLAS level-1 trigger: Technical Design Report,". `https://cds.cern.ch/record/381429`.

[54] **ATLAS Collaboration**, P. Jenni, M. Nessi, M. Nordberg, and K. Smith, "ATLAS high-level trigger, data-acquisition and controls: Technical Design Report,". `https://cds.cern.ch/record/616089`.

[55] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, "Calorimeter Clustering Algorithms: Description and Performance,". `http://cds.cern.ch/record/1099735`.

[56] T. Cornelissen, M. Elsing, S. Fleischmann, W. Liebig, and E. Moyse, "Concepts, Design and Implementation of the ATLAS New Tracking (NEWT),". `http://cds.cern.ch/record/1020106`.

[57] T. Cornelissen, M. Elsing, I. Gavrilenko, W. Liebig, E. Moyse, and A. Salzburger, "The new ATLAS track reconstruction (NEWT)," *J. Phys. Conf. Ser.* **119** (2008) 032014.

[58] **ATLAS Collaboration**, "Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung,". `https://cds.cern.ch/record/1449796`.

[59] **ATLAS Collaboration**, M. Aaboud *et al.*, "Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-1 data," `arXiv:1606.01813 [hep-ex]`.

[60] **ATLAS Collaboration**, "Reconstruction of Photon Conversions,". `https://cds.cern.ch/record/1171942`. CSC PUB Note.

[61] **ATLAS Collaboration**, "Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data," *Eur. Phys. J. C* **74** (2014) 3071, `arXiv:1407.5063 [hep-ex]`.

[62] **ATLAS Collaboration**, "Electron and photon energy calibration with the ATLAS detector using data collected in 2015 at $\sqrt{s} = 13$ TeV,". `https://cds.cern.ch/record/2203514`.

[63] **ATLAS Collaboration**, "Expected photon performance in the ATLAS experiment,". `https://cds.cern.ch/record/1345329`.

[64] **ATLAS Collaboration**, "Photon identification in 2015 ATLAS data,". `https://cds.cern.ch/record/2203125`.

[65] **ATLAS Collaboration**, "Photon Identification Efficiencies Using 2016 Data ," Aug, 2016. `https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/EGAM-2016-003/index.html`.

[66] **ATLAS Collaboration**, "Calorimeter isolation with TopoClusters versus pileup," March, 2012. `https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/EGAMMA/PublicPlots/20120404//TopoIsolation/ATL-COM-PHYS-2012-362/ATL-COM-PHYS-2012-362.pdf`.

[67] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, "Event generation with SHERPA 1.1," *JHEP* **02** (2009) 007, arXiv:0811.4622 [hep-ph].

[68] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1," *Comput. Phys. Commun.* **178** (2008) 852–867, arXiv:0710.3820 [hep-ph].

[69] **ATLAS Collaboration**, M. Aaboud *et al.*, "Search for resonances in diphoton events at $\sqrt{s}$=13 TeV with the ATLAS detector," *JHEP* **09** (2016) 001, arXiv:1606.03833 [hep-ex].

[70] **ATLAS Collaboration**, Y. C. Yap, L. Roos, and B. Lenzi, "Photon conversion classification and measurement using information from the electromagnetic calorimeter,". https://cds.cern.ch/record/2221661.

[71] **GEANT4**, S. Agostinelli *et al.*, "GEANT4: A Simulation toolkit," *Nucl. Instrum. Meth.* **A506** (2003) 250–303.

[72] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, "TMVA: Toolkit for Multivariate Data Analysis," *PoS* **ACAT** (2007) 040, arXiv:physics/0703039.

[73] K. Liu, Y. Liu, G. Marchiori, and E. Soldatov, "Measurement of the identification efficiency of isolated prompt photons using radiative $Z \to \ell\ell\gamma$ decays in 20.7 fb^{-1} of pp collisions at $\sqrt{s} = 8$ TeV,". https://cds.cern.ch/record/1636404.

[74] **ROOT Collaboration**, K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke, "HistFactory: A tool for creating statistical models for use with RooFit and RooStats,". https://cds.cern.ch/record/1456844.

[75] B. Lenzi, "Fraction of converted photons in data and MC," Oct, 2014. https://indico.cern.ch/event/310874/session/6/contribution/52/material/slides/0.pdf.

[76] E. Gross and O. Vitells, "Trial factors or the look elsewhere effect in high energy physics," *Eur. Phys. J.* **C70** (2010) 525–530, arXiv:1005.1891 [physics.data-an].

[77] O. Vitells and E. Gross, "Estimating the significance of a signal in a multi-dimensional search," *Astroparticle Physics* **35** no. 5, (2011) 230 – 234. http://www.sciencedirect.com/science/article/pii/S0927650511001630.

[78] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics," *Eur. Phys. J.* **C71** (2011) 1554, arXiv:1007.1727 [physics.data-an]. [Erratum: Eur. Phys. J.C73,2501(2013)].

[79] A. L. Read, "Presentation of search results: The CL(s) technique," *J. Phys.* **G28** (2002) 2693–2704.

[80] A. Wald, "Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large," *Transactions of the American Mathematical Society* **54** no. 3, (1943) 426–482. http://www.jstor.org/stable/1990256.

[81] S. S. Wilks, "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses," *Annals Math. Statist.* **9** no. 1, (1938) 60–62.

[82] **ATLAS Collaboration**, "Search for a CP-odd Higgs boson decaying to Zh in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector," *Phys. Lett. B* **744** no. arXiv:1502.04478. CERN-PH-EP-2015-013, (Feb, 2015) 163–183. 30 p. https://cds.cern.ch/record/1989793.

[83] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, "MadGraph 5 : Going Beyond," *JHEP* **06** (2011) 128, `arXiv:1106.0522 [hep-ph]`.

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

[85] D. Eriksson, J. Rathsman, and O. Stal, "2HDMC: Two-Higgs-Doublet Model Calculator Physics and Manual," *Comput. Phys. Commun.* **181** (2010) 189–205, `arXiv:0902.0851 [hep-ph]`.

[86] **ATLAS Collaboration**, "Measurement of the Higgs boson mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector," *Phys. Rev. D* **90** (Sep, 2014) 052004. `http://link.aps.org/doi/10.1103/PhysRevD.90.052004`.

[87] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k(t) jet clustering algorithm," *JHEP* **04** (2008) 063, `arXiv:0802.1189 [hep-ph]`.

[88] **ATLAS Collaboration**, "Pile-up subtraction and suppression for jets in ATLAS,". `https://cds.cern.ch/record/1570994`.

[89] T. Skwarnicki, "Ph.D. thesis," *DESY-F31-86-02* (1986) Appendix E.

[90] M. Oreglia, "Ph.D. thesis," *SLAC-R-0236* (1980) Appendix D.

[91] J. Gaiser, "Ph.D. thesis," *SLAC-R-0255* (1982) Appendix F.

[92] **ATLAS Collaboration**, M. Aaboud *et al.*, "Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC," `arXiv:1608.03953 [hep-ex]`.

[93] **Particle Data Group**, K. A. Olive *et al.*, "Review of Particle Physics," *Chin. Phys.* **C38** (2014) 090001.

[94] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory.* 1995.

[95] S. Frixione, F. Stoeckli, P. Torrielli, B. R. Webber, and C. D. White, "The MC@NLO 4.10 Event Generator.".

[96] S. Goria, G. Passarino, and D. Rosco, "The Higgs Boson Lineshape," *Nucl. Phys.* **B864** (2012) 530–579, `arXiv:1112.5517 [hep-ph]`.

[97] D. Y. Bardin, S. M. Bilenky, W. Beenakker, F. A. Berends, W. L. van Neerven, S. C. Van der Marck, G. Burgers, W. F. L. Hollik, T. Riemann, and M. Sachwitz, "Z line shape,". `https://cds.cern.ch/record/199969`.

[98] **LHC Higgs Cross Section Working Group**, S. Dittmaier *et al.*, "Handbook of LHC Higgs Cross Sections: 2. Differential Distributions," `arXiv:1201.3084 [hep-ph]`.

[99] G. Valencia and S. Willenbrock, "The Heavy Higgs resonance," *Phys. Rev.* **D46** (1992) 2247–2251.

[100] A. Ghinculov and T. Binoth, "On the position of a heavy Higgs pole," *Phys. Lett.* **B394** (1997) 139–146, `arXiv:hep-ph/9611357 [hep-ph]`.

[101] N. Kauer and G. Passarino, "Inadequacy of zero-width approximation for a light Higgs boson signal," *JHEP* **08** (2012) 116, `arXiv:1206.4803 [hep-ph]`.

[102] **ATLAS Collaboration**, G. Aad *et al.*, "Constraints on the off-shell Higgs boson signal strength in the high-mass ZZ and WW final states with the ATLAS detector," *Eur. Phys. J.* **C75** no. 7, (2015) 335, `arXiv:1503.01060 [hep-ex]`.

[103] **CMS Collaboration**, V. Khachatryan *et al.*, "Constraints on the Higgs boson width from off-shell production and decay to Z-boson pairs," *Phys. Lett.* **B736** (2014) 64–85, `arXiv:1405.3455 [hep-ex]`.

[104] **ATLAS Collaboration**, G. Aad *et al.*, "Search for a high-mass Higgs boson decaying to a W boson pair in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector," *JHEP* **01** (2016) 032, `arXiv:1509.00389 [hep-ex]`.

[105] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO Higgs boson production via gluon fusion matched with shower in POWHEG," *JHEP* **04** (2009) 002, `arXiv:0812.0578 [hep-ph]`.

[106] **ATLAS Collaboration**, G. Aad *et al.*, "Measurements of fiducial and differential cross sections for Higgs boson production in the diphoton decay channel at $\sqrt{s} = 8$ TeV with ATLAS," *JHEP* **09** (2014) 112, `arXiv:1407.4222 [hep-ex]`.

[107] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C. P. Yuan, "New parton distributions for collider physics," *Phys. Rev.* **D82** (2010) 074024, `arXiv:1007.2241 [hep-ph]`.

[108] V. Bertone, S. Carrazza, and J. Rojo, "APFEL: A PDF Evolution Library with QED corrections," *Comput. Phys. Commun.* **185** (2014) 1647–1668, `arXiv:1310.1394 [hep-ph]`.

[109] S. Carrazza, A. Ferrara, D. Palazzo, and J. Rojo, "APFEL Web," *J. Phys.* **G42** no. 5, (2015) 057001, `arXiv:1410.5456 [hep-ph]`.

[110] P. Artoisenet *et al.*, "A framework for Higgs characterisation," *JHEP* **11** (2013) 043, `arXiv:1306.6464 [hep-ph]`.

[111] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr, and B. Fuks, "FeynRules 2.0 - A complete toolbox for tree-level phenomenology," *Comput. Phys. Commun.* **185** (2014) 2250–2300, `arXiv:1310.1921 [hep-ph]`.

[112] **NNPDF**, R. D. Ball *et al.*, "Parton distributions for the LHC Run II," *JHEP* **04** (2015) 040, `arXiv:1410.8849 [hep-ph]`.

[113] A. Pukhov, "CalcHEP 2.3: MSSM, structure functions, event generation, batchs, and generation of matrix elements for other packages," `arXiv:hep-ph/0412191 [hep-ph]`.

[114] S. Jung, J. Song, and Y. W. Yoon, "How Resonance-Continuum Interference Changes 750 GeV Diphoton Excess: Signal Enhancement and Peak Shift," *JHEP* **05** (2016) 009, `arXiv:1601.00006 [hep-ph]`.

[115] S. Jung, J. Song, and Y. W. Yoon, "Dip or nothingness of a Higgs resonance from the interference with a complex phase," *Phys. Rev.* **D92** no. 5, (2015) 055009, `arXiv:1505.00291 [hep-ph]`.

[116] S. Jung, Y. W. Yoon, and J. Song, "Interference effect on a heavy Higgs resonance signal in the $\gamma\gamma$ and ZZ channels," *Phys. Rev.* **D93** no. 5, (2016) 055035, `arXiv:1510.03450 [hep-ph]`.

[117] A. Djouadi, J. Ellis, and J. Quevillon, "Interference effects in the decays of spin-zero resonances into $\gamma\gamma$ and $t\bar{t}$," *JHEP* **07** (2016) 105, `arXiv:1605.00542 [hep-ph]`.

[118] J. Bijnens, P. Eerola, M. Maul, A. Mansson, and T. Sjostrand, "QCD signatures of narrow graviton resonances in hadron colliders," *Phys. Lett.* **B503** (2001) 341–348, `arXiv:hep-ph/0101316 [hep-ph]`.

[119] **ATLAS Collaboration**, "Search for resonances decaying to photon pairs in 3.2 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector,". `https://cds.cern.ch/record/2114853`.

[120] **ATLAS Collaboration**, "Search for resonances in diphoton events with the ATLAS detector at $\sqrt{s} = 13$ TeV,". `https://cds.cern.ch/record/2141568`.

[121] P. Nason and C. Oleari, "NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG," *JHEP* **02** (2010) 037, `arXiv:0911.5299 [hep-ph]`.

[122] S. Schumann and F. Krauss, "A Parton shower algorithm based on Catani-Seymour dipole factorisation," *JHEP* **03** (2008) 038, `arXiv:0709.1027 [hep-ph]`.

[123] S. Hoeche, F. Krauss, S. Schumann, and F. Siegert, "QCD matrix elements and truncated showers," *JHEP* **05** (2009) 053, `arXiv:0903.1219 [hep-ph]`.

[124] T. Binoth, J. P. Guillet, E. Pilon, and M. Werlen, "A Full next-to-leading order study of direct photon pair production in hadronic collisions," *Eur. Phys. J.* **C16** (2000) 311–330, `arXiv:hep-ph/9911340 [hep-ph]`.

[125] **ATLAS Collaboration**, G. Aad *et al.*, "Improved luminosity determination in pp collisions at sqrt(s) = 7 TeV using the ATLAS detector at the LHC," *Eur. Phys. J.* **C73** no. 8, (2013) 2518, `arXiv:1302.4393 [hep-ex]`.

[126] **ATLAS Collaboration**, G. Aad *et al.*, "Search for Scalar Diphoton Resonances in the Mass Range $65 - 600$ GeV with the ATLAS Detector in pp Collision Data at $\sqrt{s} = 8$ TeV," *Phys. Rev. Lett.* **113** no. 17, (2014) 171801, `arXiv:1407.6583 [hep-ex]`.

[127] **ATLAS Collaboration**, G. Aad *et al.*, "Search for high-mass diphoton resonances in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector," *Phys. Rev.* **D92** no. 3, (2015) 032004, `arXiv:1504.05511 [hep-ex]`.

[128] **CMS Collaboration**, V. Khachatryan *et al.*, "Search for Resonant Production of High-Mass Photon Pairs in Proton-Proton Collisions at $\sqrt{s}$ =8 and 13 TeV," *Phys. Rev. Lett.* **117** no. 5, (2016) 051802, `arXiv:1606.04093 [hep-ex]`.

[129] **ATLAS Collaboration**, "Search for scalar diphoton resonances with 15.4 fb^{-1} of data collected at $\sqrt{s}$=13 TeV in 2015 and 2016 with the ATLAS detector,". `http://cds.cern.ch/record/2206154`.

Appendix

A Photon Conversion Studies

A.1 Classification performance as a function of pseudorapidity, transverse momentum, radius of conversion and number of primary vertices

The default track-based conversion tagging shows a dip around $|\eta| = 1$ due to the gap at the transition from the barrel to the end-cap TRT and at $|\eta| > 2$ as the limit of the TRT pseudorapidity range is approached, as shown earlier in Figure 3.1. It is also less efficient in reconstructing low p_T photon conversions. On the other hand, since the efficiency of the multivariate classifiers are fixed at 90% across $|\eta|$ categories, the efficiencies are rather flat across the p_T spectrum as well. However, the multivariate classifiers still do not perform very well on low p_T photons as fake conversion rates are higher. In terms of the radius of conversion, the multivariate classifiers are more able to reconstruct conversions at large radii.

In these plots, we see that the default conversion tagging method works very well at very low pile-up, except for the drop in efficiency at the η regions mentioned above and at larger radii. At higher pile-up, the multivariate classifiers perform much better. Since the default conversion reconstruction in use here is not tuned to the high pile-up environment, some tunings and improvements in the track-based reconstruction could result in better performance.

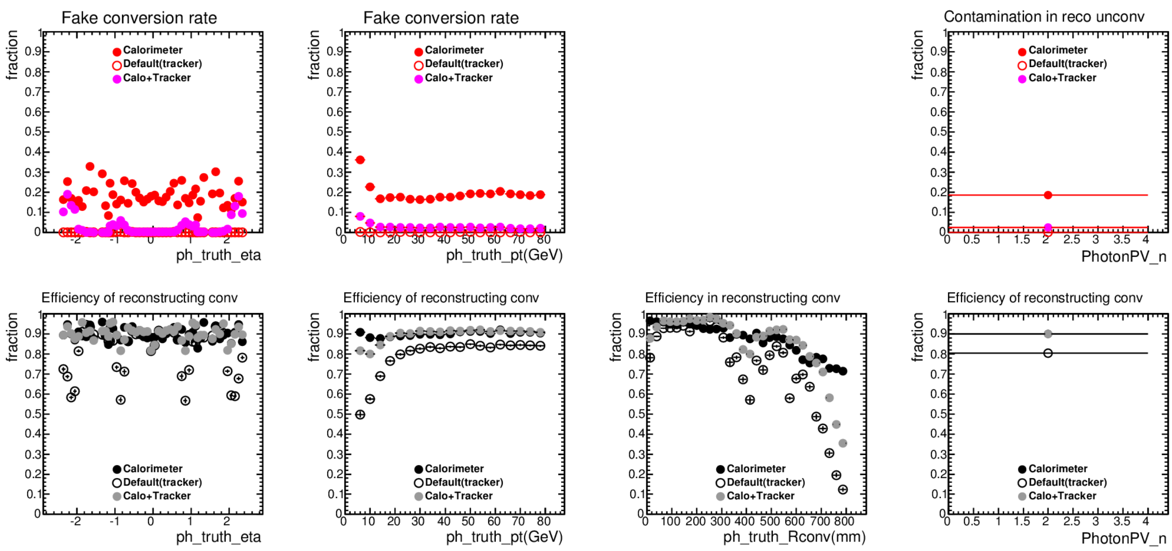


Figure 15: The conversion reconstruction efficiency and fake conversion rate of the calorimeter-based classification, the default track-based reconstruction and the classification using both calorimeter and tracking information. The performance on the single photon Monte Carlo sample with $\mu = 0$ is shown as a function of η, transverse momentum, radius of conversion and number of primary vertices.

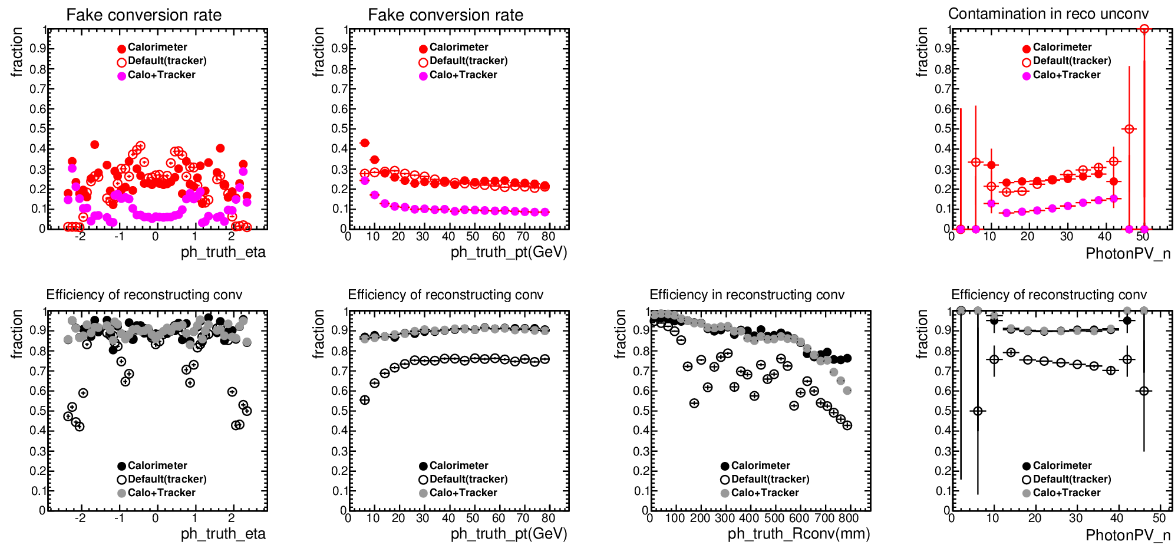


Figure 16: The conversion reconstruction efficiency and fake conversion rate of the calorimeter-based classification, the default track-based reconstruction and the classification using both calorimeter and tracking information. The performance on the single photon Monte Carlo sample with $\mu = 40$ is shown as a function of η, transverse momentum, radius of conversion and number of primary vertices.

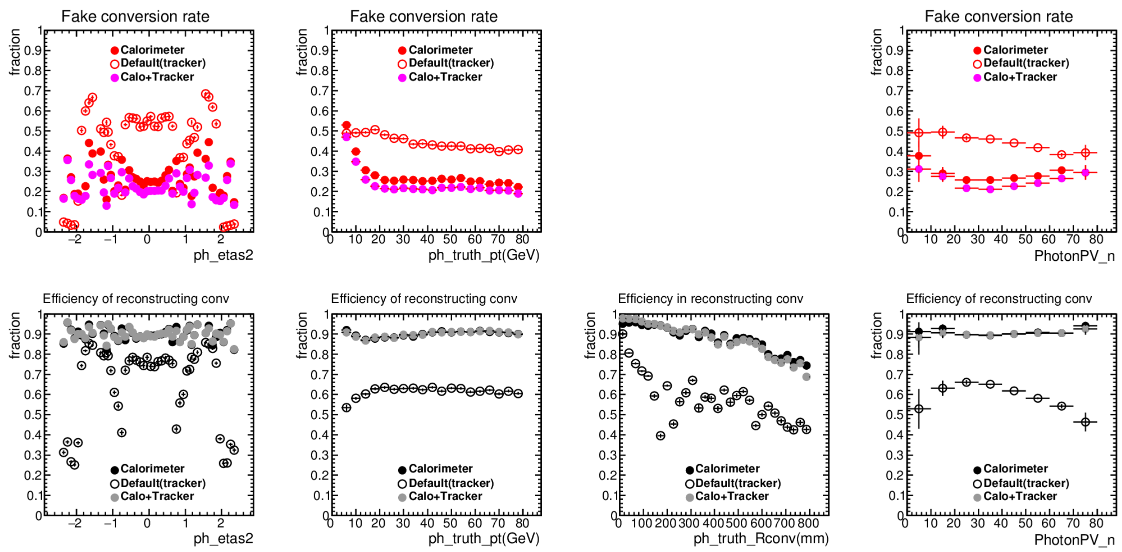


Figure 17: The conversion reconstruction efficiency and fake conversion rate of the calorimeter-based classification, the default track-based reconstruction and the classification using both calorimeter and tracking information. The performance on the single photon Monte Carlo sample with $\mu = 80$ is shown as a function of η, transverse momentum, radius of conversion and number of primary vertices.

A.2 Effect of having no stand-alone TRT tracks

In the case of high pile-up at the end of Run 2 or beyond, it might be difficult or even impossible to construct stand-alone TRT tracks, i.e. tracks that have only TRT hits and no silicon hits. In order to study the effect of such a scenario on the conversion classification, the number of TRT hits is reset to zero for the case when the photon conversion is reconstructed using stand-alone TRT tracks. Nevertheless, there is a small caveat as losing stand-alone TRT tracks means that those photons would be reconstructed as unconverted and hence have smaller cluster size in the barrel. This results in changes in some shower shape variables, which are not taken into account.

Using the multivariate classification method with both calorimeter and tracking variables, the effect of losing stand-alone TRT tracks at high pile-up is not large since the calorimeter variables provide the most discrimination between converted and unconverted photons. For a moderate pile-up of $\mu = 40$, a 9% increase in fake conversion rate is seen (at 90% efficiency). At higher μ, the effect is even smaller (6% and 3% respectively for $\mu = 60$ and $\mu = 80$).

These numbers represent the overall increase in the fake conversion rate (inclusive), rather than a relative increase for stand-alone TRT tracks. Stand-alone TRT tracks type conversion reconstruction is highly unreliable at high pile-up condition, particularly since this reconstruction has not been adapted to high pile-up. Hence, this information has little significance in BDT. To better demonstrate the improvement that the calorimeter-based tagging can bring, we look at both the purity and efficiency of this tagging on the photons reconstructed as converted via stand-alone TRT tracks.

The ability of the BDT to reconstruct true conversion if stand-alone TRT tracks are lost and to restore the fake conversion as unconverted photon is quantified in Table 2. In the $\mu = 80$ sample, only 41% of the photons reconstructed as converted via stand-alone TRT tracks are true converted photons. Of these, 88% can be tagged as converted using the BDT with a working point for overall 90% efficiency in reconstructing conversion. 77% of fake conversions from stand-alone TRT tracks are restored in the unconverted photon collection using BDT. This attests the robustness of a calorimeter-based conversion classification in a high pile-up environment.

μ	Fraction of true converted photons	Conversion reconstruction efficiency of BDT	Fraction of fake conversions that are recovered as unconverted using BDT
40	0.515	0.845	0.803
60	0.417	0.864	0.787
80	0.411	0.878	0.765

Table 2: Breakdown of the photons that would be reconstructed as converted via stand-alone TRT tracks. The second column shows the fraction of photons reconstructed as converted via stand-alone TRT tracks that are true converted photons. The third column shows the efficiency in reconstructing true converted photons (that would have stand-alone TRT tracks) as converted while the last column shows the fraction of fake conversions reconstructed via stand-alone TRT tracks that are restored as unconverted using BDT.

A.3 $p_T > 15$ GeV photons

From the study in Ref. [59], the photon purity for $p_T > 15$ GeV is found to be about 98%, higher than that for $p_T > 10$ GeV. Therefore, the photon transverse momentum cut is raised to 15 GeV. The true values of f_{conv} and f_{reco} extracted from the MC samples increase due to the higher transverse momenta, as shown in Table 3.

η	f_{conv}	f_{fake}	f_{reco}
$\lvert\eta\rvert < 0.6$	0.178	0.023	0.909
$0.6 < \lvert\eta\rvert < 1.37$	0.322	0.034	0.832
$1.52 < \lvert\eta\rvert < 1.8$	0.453	0.017	0.950
$1.8 < \lvert\eta\rvert < 2.37$	0.562	0.008	0.628

Table 3: The values of f_{conv}, f_{fake} and f_{reco} obtained from the nominal Monte Carlo with photon $p_T > 15$ GeV.

By raising the cut on the photon transverse momentum, the number of candidate photons is reduced to 56% of the total. This results in increased error sizes on the fit values. Table 4 lists the results of the fit to data. While some discrepancies remain, f_{reco} from the two fits agree much better in $|\eta| < 0.6$ and $1.52 < |\eta| < 1.8$. This suggests again that the background is responsible for the discrepancies observed between the two fits and by removing the background entirely, one may obtain similar fit values for the two fits.

	E_1/E_2		BDTG	
η	f_{fake}	f_{reco}	f_{fake}	f_{reco}
$\lvert\eta\rvert < 0.6$	0.018 ± 0.005	0.926 ± 0.031	0.021 ± 0.004	0.894 ± 0.011
$0.6 < \lvert\eta\rvert < 1.37$	0.034 ± 0.006	0.828 ± 0.018	0.064 ± 0.009	0.775 ± 0.013
$1.52 < \lvert\eta\rvert < 1.8$	0.023 ± 0.019	0.936 ± 0.028	0.026 ± 0.018	0.937 ± 0.017
$1.8 < \lvert\eta\rvert < 2.37$	0.014 ± 0.008	0.601 ± 0.014	0.018 ± 0.009	0.606 ± 0.017

Table 4: The results of the fit to data with photon $p_T > 15$ GeV using E_1/E_2 (left) and the BDTG (right) variable, neglecting background contribution.

A.4 MC test of background impact on template fit

In an attempt to understand the impact of background on the template fit, a test using only Monte Carlo is performed. The background templates are obtained directly from the $Z \to \mu\mu$ MC sample with the same selection as signal except for a requirement on the photon candidate to be fake while signal templates are the usual ones. From the cross-sections and acceptances in Monte Carlo, the purity in each channel (each $|\eta|$ bin and reconstructed conversion status) is estimated. The background templates are then normalised to the number of background events expected in data from the purity calculations and are shown in Figure 4.19.

In this test, a pseudo-data sample is used that is constructed from the sum of signal and background templates, where the contribution of each corresponds to the number expected in the data from purity calculations. Thus, the pseudo-data contains the same number of events as the actual data. The fractions of conversion, efficiencies and fake rates in the pseudo-data differ from those in signal Monte Carlo. This is because the signal contribution in the reconstructed converted channel are scaled by different amount to that in the reconstructed unconverted channel since the purities are calculated separately for the photons that are reconstructed as converted and those reconstructed as unconverted. The input values of f_{conv}, f_{fake} and f_{reco} are presented in Table 5.

η	f_{conv}	f_{fake}	f_{reco}
$\lvert\eta\rvert < 0.6$	0.164	0.021	0.888
$0.6 < \lvert\eta\rvert < 1.37$	0.306	0.032	0.799
$1.52 < \lvert\eta\rvert < 1.8$	0.427	0.015	0.934
$1.8 < \lvert\eta\rvert < 2.37$	0.541	0.007	0.584

Table 5: f_{conv}, f_{fake} and f_{reco} in the signal component of the pseudo-data being tested.

When the pseudo-data is fitted with both signal and background components, f_{fake} and f_{reco} obtained are the same as the input values, as indicated in Table 6.

	E_1/E_2		BDTG	
η	f_{fake}	f_{reco}	f_{fake}	f_{reco}
$\lvert\eta\rvert < 0.6$	0.020 ± 0.002	0.885 ± 0.013	0.020 ± 0.002	0.888 ± 0.010
$0.6 < \lvert\eta\rvert < 1.37$	0.032 ± 0.003	0.797 ± 0.009	0.033 ± 0.002	0.798 ± 0.007
$1.52 < \lvert\eta\rvert < 1.8$	0.013 ± 0.012	0.938 ± 0.016	0.015 ± 0.008	0.935 ± 0.011
$1.8 < \lvert\eta\rvert < 2.37$	0.006 ± 0.004	0.584 ± 0.008	0.006 ± 0.010	0.585 ± 0.010

Table 6: The results of the fit to pseudo-data (sum of MC signal and background) with signal and background components.

However, if we ignore background by fitting only with the signal components of converted and unconverted photons, the fit using the BDTG output yields the wrong values on f_{reco} and f_{fake}, as shown in Table 7. The measured f_{reco} is significantly lower and f_{fake} higher when the background is neglected. On the other hand, the fit using E_1/E_2 is insensitive to background as f_{reco} and f_{fake} obtained are similar to the correct values.

The insensitivity of the fit using E_1/E_2 to background is due to the smaller separation of converted photons and unconverted photons in E_1/E_2. The background, which comes from jets faking photons, have R_ϕ distributions that is very similar to those of the true converted photons, whereas this is less clear in E_1/E_2. Hence, in the fit using E_1/E_2, by ignoring background, the contributions of converted and unconverted photons are scaled to absorb the background in a way that the fractions (f_{reco} and f_{fake}) are not affected significantly. On the other hand, in the fit

	E_1/E_2		BDTG	
η	f_{fake}	f_{reco}	f_{fake}	f_{reco}
$\lvert\eta\rvert < 0.6$	0.023 ± 0.002	0.880 ± 0.013	0.027 ± 0.002	0.816 ± 0.012
$0.6 < \lvert\eta\rvert < 1.37$	0.035 ± 0.005	0.794 ± 0.014	0.051 ± 0.004	0.746 ± 0.008
$1.52 < \lvert\eta\rvert < 1.8$	0.034 ± 0.024	0.928 ± 0.042	0.038 ± 0.022	0.899 ± 0.019
$1.8 < \lvert\eta\rvert < 2.37$	0.009 ± 0.011	0.582 ± 0.022	0.011 ± 0.009	0.575 ± 0.014

Table 7: The results of the fit to pseudo-data (sum of MC signal and background) with only signal components.

using the BDTG output, where the background templates clearly resemble the templates from true converted photons, the converted distribution is the one absorbing most of the background contribution, resulting in lower conversion reconstruction efficiency being measured.

This is consistent with the results in Section 4.4.6 and 4.4.7. It also confirms that in the fit using the BDTG variable, even a small amount of background needs to be taken into account.

A.5 Check on the impact of ignoring statistical errors on background templates

The uncertainties on the background templates were not taken into account due to the reasons mentioned in Section 4.4.4.1. We try a workaround by increasing the errors on the signal templates so that they include also the errors on the background templates. This is done by scaling the errors in the signal true converted and unconverted templates by a factor equal to the ratio of the total errors to the total signal errors in Figure 4.20.

Similar results are obtained, except that the errors on f_{reco} and f_{fake} are larger by no more than 20% with the exception at f_{reco} at $0.6 < |\eta| < 1.37$ where it is considerably higher. The results are presented in Table 8. The increase in the errors is also more apparent in the E_1/E_2 fit than it is in the BDTG fit, which strengthens our claim that the fit with more variables is better-performing. At the same time, the constraints on the nuisance parameters are also slightly reduced (see Figure 18).

The conclusions are unchanged, BDTG still gives a more precise measurement on the conversion reconstruction efficiency compared to E_1/E_2 and some constraints on the material-related and R_ϕ fudge factor nuisance parameters are still seen.

	E_1/E_2		BDTG	
η	f_{fake}	f_{reco}	f_{fake}	f_{reco}
$\lvert\eta\rvert < 0.6$	0.013 ± 0.006	0.921 ± 0.047	0.015 ± 0.003	0.890 ± 0.013
$0.6 < \lvert\eta\rvert < 1.37$	0.029 ± 0.005	0.801 ± 0.023	0.027 ± 0.005	0.804 ± 0.009
$1.52 < \lvert\eta\rvert < 1.8$	0.009 ± 0.015	0.918 ± 0.030	0.015 ± 0.018	0.906 ± 0.019
$1.8 < \lvert\eta\rvert < 2.37$	0.006 ± 0.007	0.568 ± 0.014	0.013 ± 0.009	0.565 ± 0.013

Table 8: The results of the fit to data with the background subtracted using E_1/E_2 (left) and BDTG (right) where the errors on the signal templates have been scaled higher to absorb the errors on the background templates.

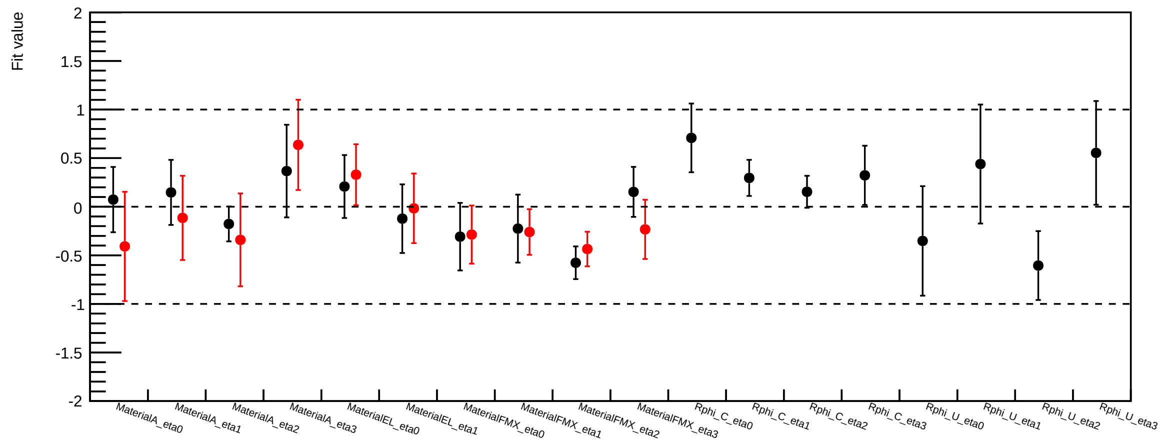


Figure 18: Fitted values of the nuisance parameters from the fit using E_1/E_2 (in red) and BDTG (in black) to the data, with the background contribution subtracted. The errors on the signal templates have been scaled to include the errors on the background templates.

B Higgs-like resonance line shape assuming infinite mass particle in production and decay loop using MC@NLO prescription

Using the MC@NLO prescription, the line shape of a heavy Higgs-like resonance is

$$\frac{d\sigma_{\text{MC@NLO}}}{dm_{\gamma\gamma}} \propto \mathcal{B}_{gg} \cdot \mathcal{L}_{gg} \cdot \frac{\sqrt{\hat{s}}\Gamma_{X\to\gamma\gamma}(\hat{s})}{(\hat{s}-m_X^2)^2 + m_X^2\Gamma_X^2} \cdot m_{\gamma\gamma}. \tag{4}$$

Assuming infinite top quark and W boson mass, the $m_{\gamma\gamma}$-dependence of the production process squared matrix element (see Eq. 7.31) and the partial width of 'Higgs' to two photons [14] can be approximated as:

$$\mathcal{B}_{gg} \propto m_{\gamma\gamma}^2 \tag{5}$$

and

$$\Gamma_{X\to\gamma\gamma} \propto m_{\gamma\gamma}^3. \tag{6}$$

Substituting in the differential cross-section:

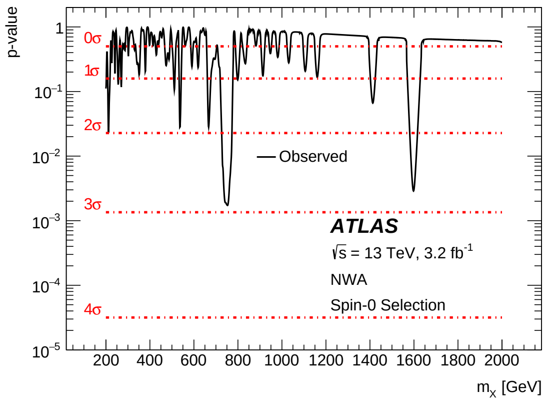
$$\frac{d\sigma_{\text{MC@NLO}}}{dm_{\gamma\gamma}} \propto m_{\gamma\gamma}^2 \cdot \mathcal{L}_{gg} \cdot m_{\gamma\gamma} \frac{1}{(\hat{s}-m_X^2)^2 + m_X^2\Gamma_X^2} m_{\gamma\gamma}^3 \cdot m_{\gamma\gamma}. \tag{7}$$

Finally, one obtains

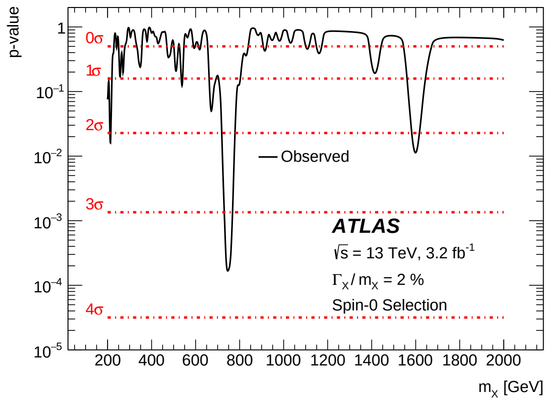
$$\frac{d\sigma_{\text{MC@NLO}}}{dm_{\gamma\gamma}} \propto m_{\gamma\gamma}^7 \cdot \mathcal{L}_{gg} \cdot \frac{1}{(\hat{s}-m_X^2)^2 + m_X^2\Gamma_X^2}, \tag{8}$$

which is equivalent to the EFT line shape implemented in MG5_aMC@NLO (see Eq. 7.42).

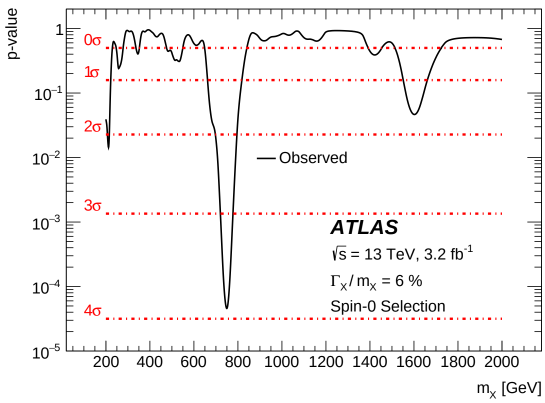
C p-value as a function of mass for different width hypothesis in the diphoton resonance search at $\sqrt{s}$ = 13 TeV with 3.2 fb^{-1} recorded in 2015



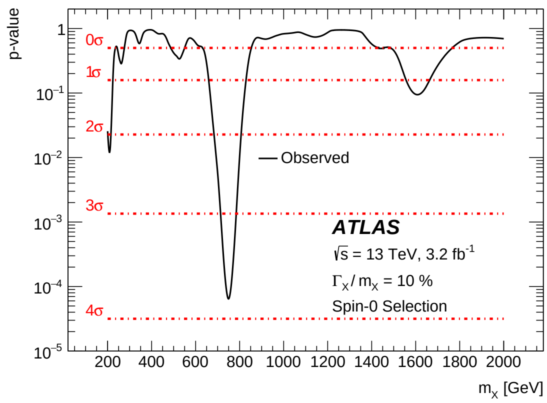
(a)



(b)

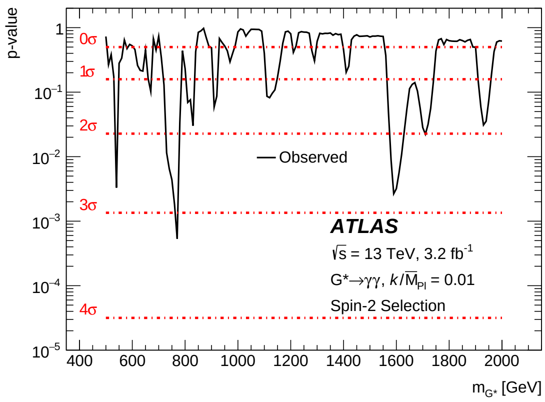


(c)

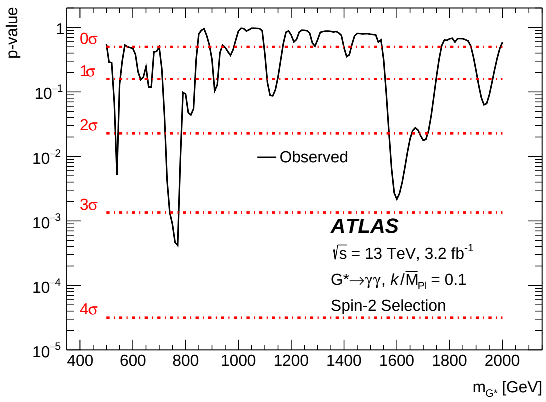


(d)

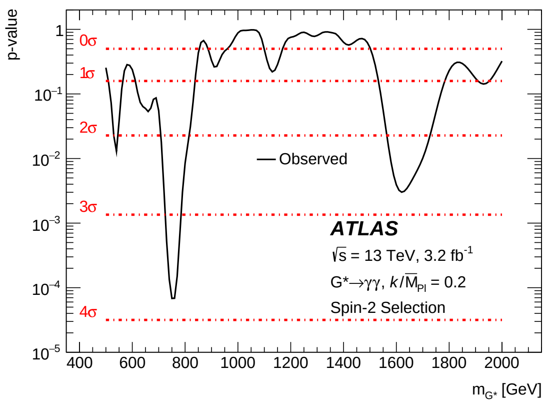
Figure 19: Compatibility with the background-only hypothesis as a function of the assumed signal mass for different values of relative width for the spin-0 resonance search.



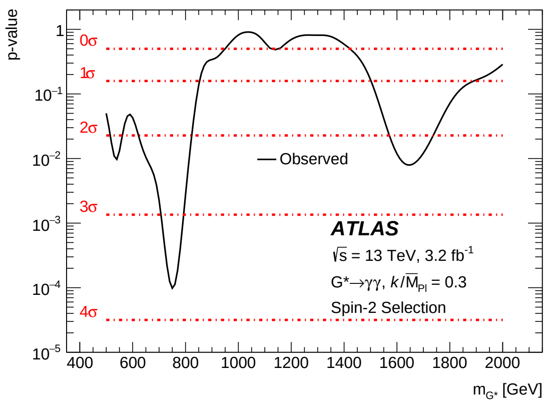
(a)



(b)



(c)



(d)

Figure 20: Compatibility with the background-only hypothesis as a function of the assumed signal mass for different values of $k/\bar{M}_{\mathrm{Pl}}$ for the spin-2 resonance search.

D Upper limits on fiducial cross-section times branching ratio to two photons at $\sqrt{s}$ = 13 TeV with 15.4 fb^{-1} recorded in 2015 and 2016

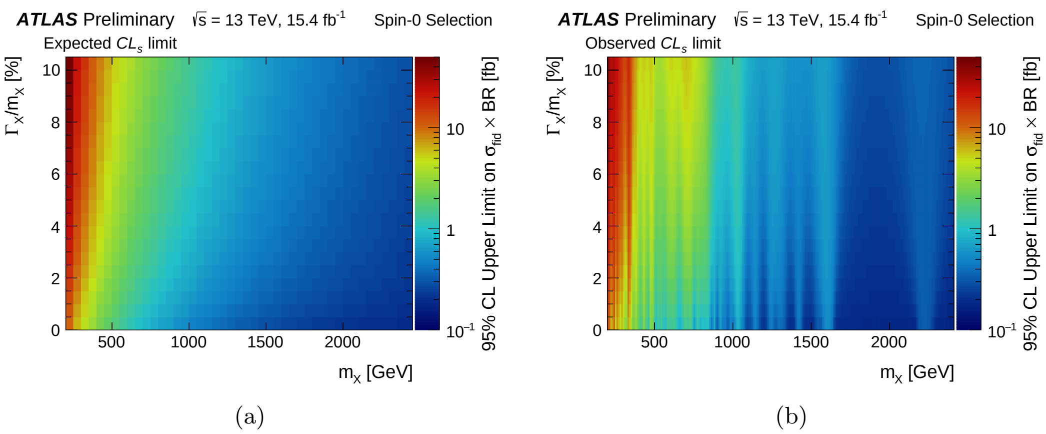


Figure 21: Two-dimensional (a) expected and (b) observed 95% CL upper limits on the fiducial production cross-section times branching ratio to two photons at $\sqrt{s}$ = 13 TeV of a spin-0 resonance as a function of the resonance mass m_X and the relative width Γ_X/m_X.

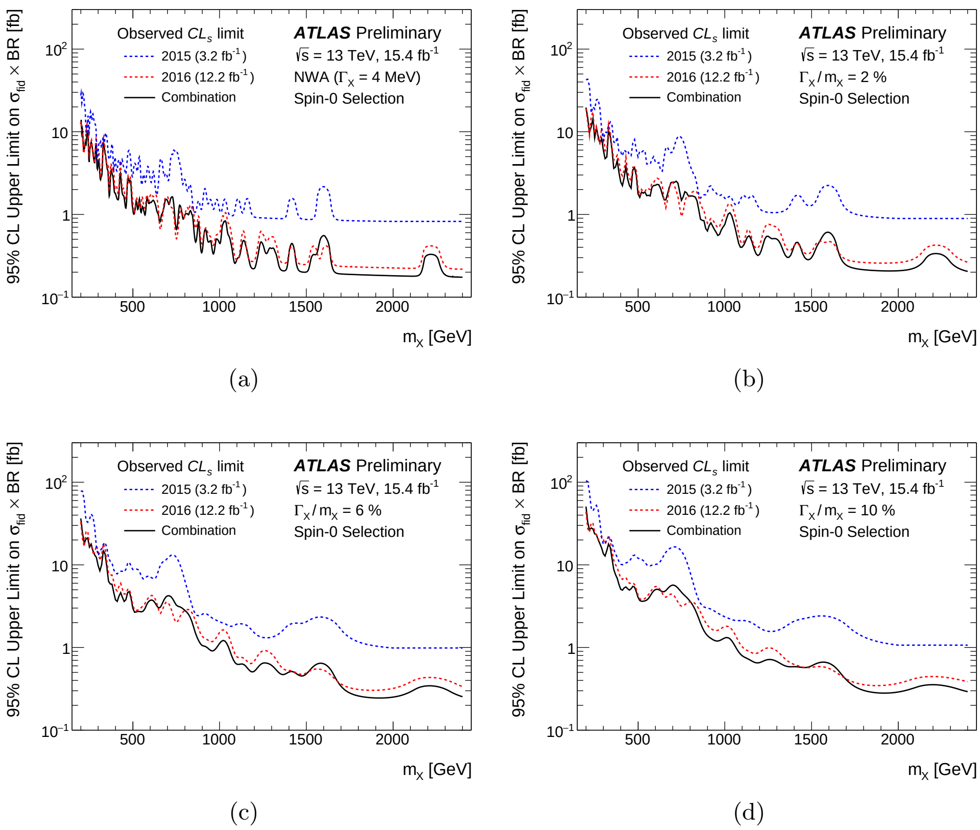


Figure 22: 95% CL observed upper limits on the fiducial cross-section of a spin-0 resonance as a function of the assumed mass, for different values of the decay width divided by the mass, using the 2015, 2016, and combined dataset.