Exploring the diphoton final state at the LHC at 13 TeV : searches for new particles, and the Higgs boson mass measurement with the ATLAS detector

Kirill Grevtsov

To cite this version:

Kirill Grevtsov. Exploring the diphoton final state at the LHC at 13 TeV : searches for new particles, and the Higgs boson mass measurement with the ATLAS detector. High Energy Physics - Experiment [hep-ex]. Université Grenoble Alpes, 2017. English. NNT : 2017GREAY047 . tel-01717720

HAL Id: tel-01717720
https://theses.hal.science/tel-01717720
Submitted on 26 Feb 2018

Communauté
UNIVERSITÉ Grenoble Alpes

THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTÉ UNIVERSITÉ GRENOBLE ALPES

Spécialité : **Physique Subatomique et Astroparticules**

Arrêté ministériel : 25 mai 2016

Présentée par

KIRILL GREVTSOV

Thèse dirigée par **Isabelle Wingerter-Seez**

préparée au sein du **Laboratoire Laboratoire d'Annecy-le-Vieux de Physique des Particules**
dans **l'École Doctorale Physique**

Exploration du canal diphoton au LHC: recherche de nouvelles particules et mesure de la masse du boson de Higgs avec le détecteur ATLAS

Exploring the diphoton final state at the LHC at 13 TeV: searches for new particles, and the Higgs boson mass measurement with the ATLAS detector.

Thèse soutenue publiquement le **4 juillet 2017**,
devant le jury composé de :

Monsieur Giovanni LAMANNA
LAPP, Président du jury
Monsieur Fabio MALTONI
Université catholique de Louvain, Examinateur
Monsieur Bruno MANSOULIÉ
CEA-IRFU Saclay, Examinateur
Madame Lydia ROOS
LPNHE, Examinatrice
Monsieur Paraskevas SPHICAS
CERN et Université d'Athènes, Rapporteur
Monsieur Guillaume UNAL
CERN, Rapporteur

Abstract

New high-mass states decaying into two photons are predicted in many extensions of the Standard Model (SM). The diphoton final state provides a clean experimental signature with excellent invariant mass resolution and well-known smooth backgrounds.

This document presents a search for new particles with the diphoton final state at the Large Hadron Collider with the ATLAS detector. The *pp* collision data used were collected during 2015 and 2016 runs with a center-of-mass energy of $\sqrt{s}$=13 TeV. The total corresponding luminosity is 37 fb^{-1}.

In this thesis, I show my contribution to the search of scalar particle. The studies of signal modeling for different mass and width hypothesis will be described in details. The estimation of selection efficiencies and statistical interpretations of results are performed. The data are consistent with the Standard Model background-only hypothesis. Limits on the production cross section times branching ratio to two photons of such resonances as a function of the resonance mass and width are presented.

The Liquid Argon electromagnetic calorimeter plays a crucial role in the diphoton analysis. The excellent energy resolution allows to reconstruct objects with high precision. The contribution to operation of LAr calorimeter and its online software will be discussed.

The calibration of the electron and photon energy measurements with the electromagnetic calorimeter is performed. The systematic uncertainties related to energy response are one of the largest contribution limiting the precision measurements of the Standard Model Higgs boson mass. An approach to improve the energy response taking into account the lateral shower shape development is applied in the calibration procedure.

Résumé

De nombreux modèles, au delà du modèle standard, prédisent l'existence de résonances se désintégrant en deux photons. La signature expérimentale très spécifique, associée à l'excellente résolution en masse et à la distribution du bruit de fond bien comprise, en font un canal roi pour la recherche de nouvelle physique.

Ce document présente la recherche de nouvelles particules se désintégrant en une paire de photons effectuée auprès du Grand Collisionneur de Hadrons (LHC) du CERN avec le détecteur ATLAS. L'ensemble des données de collisions de protons, accumulées en 2015 et 2016, à l'énergie dans le centre de masse de 13 TeV, correspond à 37 fb^{-1}.

J'ai développé la méthode et la procédure pour décrire un éventuel signal, en simulant des résonances de masses et largeurs variables. Avec ce modèle j'ai pu calculer l'efficacité de sélection et donner une interprétation statistique des résultats. Cette analyse a permis de montrer que les données collectées sont en accord avec les prédictions du modèle standard. Des limites sur la section efficace de production couplée au rapport d'embranchement en deux photons, en fonction de la masse et de la largeur de la résonance considérée sont présentées.

Le calorimètre électromagnétique de ATLAS joue un rôle clé pour cette analysis di-photon. L'excellente résolution en énergie permet de reconstruire électrons et photons avec une précision de l'ordre de 1% aux énergies considérées. J'ai pris part au fonctionnement du calorimètre lors du démarrage du run-II et décris ici mes contributions.

Finalement, un des points clé pour la mesure de la masse du boson de Higgs est la qualité de la calibration de l'énergie des photons. J'ai développé une approche qui permet de prendre en compte l'impact du développement latéral des gerbes électromagnétiques sur la reconstruction de l'énergie. Cette méthode permettra de réduire l'erreur systématique sur la mesure de la masse du boson de Higgs, mesurée par son mode de désintégration en deux photons.

Acknowledgements

With a great pleasure and respect, I would like to thank my supervisor Isabelle Wingerter-Seez. Thank you for your support in difficult situations, for your wise advice and for the freedom you gave me in making decisions. I have learned a lot from you about specific things, like knowledge about the calorimeter structure and electronics, and as well as about general aspects - you gave me an idea that ”not to know something is not a crime” - and I appreciate that very much!

I would like to thank my colleagues from LAPP ATLAS group for their advices and participation. I would like to acknowledge the group leaders over my thesis: Stéphane Jézéquel and Emmanuel Sauvan, who welcomed me at the group and helped with all organisational issues. I am grateful to Marco Delmastro for the discussions about physics, for sharing his deep knowledge about ”photon-business” and invaluable aid in career planning; to Nicolas Berger for his guidance over analysis strategy, explanations about statistics tricks and advices with coding. Marco and Nicolas, thank you for your help, your support and **patience** in ”hard-times” during bump-pursuing rush! I want to thank Isabelle, Marco and Thibault Guillemin for their contribution in calibration studies, their believing in my idea and help to convince other people.

As well I would like to thank people from other groups with whom I shared this exciting ”diphoton” journey. I am thankful to the conveners Tancredi Carli, Leonardo Carminati, Elisabeth Petit and German Carrillo-Montoya for their organization and support. I give my special thanks to Lydia Roos, Liron Barak, Jan Stark, Hongtao Yang, Yee Yap, Simone Mazza and other colleagues in the analysis team. Thanks to Kerstin Tackmann, Giovanni Marchiori and Chris Mayer for their help with photon-related studies.

I wish to express my gratitude to Guillaume Unal for his outstanding help to understand calorimeter response and sophisticated calibration procedure, to Louis Fayard for his detailed questions on the topic, to Ruggero Turra and Bruno Lenzi for their help with calibration tools. I would like to thank Martin Aleksa for sharing his expertise in calorimeter operation and guidance during my participation in LAr-team.

Many thanks to Alexis Vallier, Sergii Raspopov, Zuzana Barnovska, Angela Burger, Saskia Falke, Olympia Dartsi for their friendship and support. And to Andre Rumler, who shared with me so many hours staying in a traffic jam on the way to CERN.

I would like to thank the ”rapporteurs” for this thesis, Guillaume and Paraskevas Sphicas, for their time of careful reading and commenting my manuscript. I am very thankful to Lydia, Bruno Mansoulié, Fabio Maltoni and Giovanni Lamana to accepting to be in my "jury de these".

As a different type of gratitude, I would like to thank my parents, Alexey and Svetlana, who provide me a chance to choose what I want to do in my life regardless of the circumstances. And my family, my wife Polina and daughter Margarita, for their love and patience - without you there will be no point in doing all this.

Contents

Introduction

The particle physics is the branch of physics that studies the nature of the elementary particles that constitute matter and their interactions. The notion of elementary particles has evolved in time, following progress in experimental techniques; systems that were believed to be "elementary" are in fact composites of smaller constituents. Our understanding of Nature evolves. Currently the dominant theory explaining elementary particles and their dynamics, is called the Standard Model of particle physics.

The Standard Model is an elegant mathematical framework, well tested by experiments. Until the 1960-s, it had an unsolved question: how do the particles gain their mass? The Brout-Englert-Higgs mechanism, was introduced into the theory to accommodate massive $W^{\pm}$ and Z gauge bosons. It predicts the existence of a massive scalar boson - the Higgs particle, it had been searched since 1980s and remained hidden for decades, until July 2012.

The Large Hadron Collider is the world largest particle collider. On July 4th 2012, the ATLAS and CMS collaborations announced the discovery of a new particle with a mass around 126 GeV, compatible with the predicted Higgs boson, the last missing piece of the Standard Model.

Following discovery, the goal has been moved to the study of the properties of the new boson, such as the measurement of its mass and couplings to fermions and bosons. The diphoton decay channel provides a clean experimental signature, with an excellent invariant mass resolution. In my doctoral studies, I have contributed to the photon calibration procedure, attempting to reduce the systematical uncertainty to the Higgs mass measurement.

While the Standard Model is well supported by experimental observations, there are still remaining open questions: it does not explain the gravitational force and it does not provide candidates to invisible energy provided by cosmological observations. Matter-antimatter asymmetry in the universe is unexplained. Neutrinos are massless in the theory, while neutrino oscillation experiments have shown that neutrinos do have mass. To answer those questions, extensions of the Standard Model are presented, denoted as Beyond Standard Model (BSM).

There are BSM models in the extended Higgs sector [1, 2, 3, 4] where new scalar resonances are predicted. During my thesis I have searched for new heavy scalar bosons decaying into two photons using data collected by ATLAS experiment at the centre of mass energy of 13 TeV. Diphoton data collected in 2015 intrigued the physics community by the presence of a small local excess at $m_{\gamma\gamma}$ ~750 GeV. A very strong effort was done to scrutinise carefully the detector response, the reconstruction of the events and the statistical modelling. Data collected in 2016 do not reproduce the excess, which appears to be a statistical fluctuation. This work will be discussed in details in this thesis.

This thesis is organised as follows. The Chapter 1 will provide a brief review of the Standard Model, describing the mechanism of spontaneous symmetry braking, the properties of the Higgs production and decay modes at the LHC. An introduction to BSM models with decays to two photons in the final state will be discussed, including Two Higgs Doublet model and theories introducing gravitational force.

Chapter 2 briefly describes the LHC complex with its main parts and operation conditions. The performance of the LHC is discussed in term of the luminosity. The overview of the ATLAS detector is given together with the summary of the physics performance of each subdetector. The Liquid Argon (LAr) electromagnetic calorimeter is discussed in details, covering the description of the shower development in the calorimeter, the geometrical structure of the detector, the achieved energy resolution and the overview of the read-out electronics. The format of the data taking is summarised, to introduce the software of the LAr Online system, which I was working on. The simulation of *pp* collision events associated to the simulation of particle interactions in the detector is presented. The adjustment of the simulation of the pileup disrtibution to reproduce the one from the data is described, as I was responsible for the procedure during my PhD.

Chapter 3 is devoted to the photon reconstruction in the ATLAS detector. The object reconstruction in the electromagnetic calorimeter associated to information from the tracker detector is presented. Further identification, isolation and energy calibration of the photon candidate is discussed. My personal contribution to the performance studies of the photon conversion is presented.

Chapter 4 introduces my contribution to the calibration of the electrons and photons energy. The linearity of the energy response at the electronics level is presented. The correlation of the observed non-linearity with the shower development in the calorimeter is studied and method for possible improvement of the energy calibration is discussed.

Chapter 5 presents the search for new resonance decaying into two photons. The overview of the analysis is presented starting with the motivations and the selections of diphoton candidates, followed the decomposition of the sample into true and fake photons. The modelling of the signal and the expected backgrounds are presented, to test the compatibility of collected data with the Standard Model expectations.

In the Conclusion chapter, the search results and impact of the energy calibration on the Higgs mass measurement are summarised and further prospects are given.

Personal contribution

In the frame of my doctoral studies, I participated to many activities within the ATLAS collaboration. The work presented in this thesis has been done partly using tools and frameworks developed by ATLAS collaborators. To avoid ambiguities, my personal contributions will be listed below with the links to the corresponding sections. To make a clear difference between my personal commitments and results obtained by collaboration, everywhere in the text, Figures and Tables, which have their caption over grey background correspond to material produced by colleagues. The inputs with **captions without background** correspond to my personal work.

I list below the technical developments and analysis where my contribution is significant:

- Section 2.3 is dedicated to the description of the LAr calorimeter, which plays a crucial role in the analysis with photons in the final state. One of the part of the qualification task to become an ATLAS author was my contribution to the operation of the LAr calorimeter. I did shifts in the ATLAS Control Room and on-call expert shifts for the LAr Online Software. I maintained the software, which reads raw data and translates it to a human readable format (Section 2.3.5).

- Section 2.4 presents the simulation used in ATLAS. As an example, the comparison of the average number of interactions per bunch crossing (pileup) between data and simulation (Section 2.4.2). I was responsible to produce inputs for creating the pileup weights, to implement tool in the analysis framework and to validate the performance of the reweighting, which are used in the entire Higgs group.

- The reconstruction of photons in the ATLAS detector is described in Chapter 3. Using the first collision data at centre of mass energy 13 TeV I analysed the performance of the photon reconstruction (Section 3.1). This results were presented at the 2015 EPS conference, as part of the ATLAS performance in Run-2.

- Chapter 4 is devoted to calibration studies:
  - The study of the energy response of the calorimeter as a function of the electronics gain and its impact of MVA-based calibration is discussed in Section 4.1.1.
  - The study of the correlation of the energy response with the shower shapes in data and MC (Section 4.1.3). The systematic uncertainty related to observed dependency (Section 4.1.4.2).
  - The implementation of the lateral width of the shower into the calibration training and its impact on the precision of the Higgs mass measurement are summarised in Section 4.2.

- Chapter 5 presents the search for new a resonance decaying into two photons, where I contributed in several aspects of the analysis from the beginning of Run-2, in addition I was co-editor of the first note about the 750 GeV excess [107]. I presented the analysis on behalf of the ATLAS collaboration at the international conference HiggsHunting 2016 [108], and was responsible for the ”spin-0” analysis results from May 2016. I was responsible for the following parts of the analysis:

- The study of the cutflows, and the understanding of the dependence of the selection efficiencies with the pileup (Section 5.2.2).
- The signal parametrization: the detector resolution and the implementation of the large width description, and the corresponding systematic uncertainties (Section 5.4).
- The statistical interpretations: the definitions of the compatibility between datasets, the compatibility with background-only hypothesis (p_0 scan), the cross-section limits and the estimation of the global significance (Sections 5.6 and 5.8).

Chapter 1

Theoretical introduction

Contents

The Standard Model of Particle Physics gives a description of the known existing elementary particles and of their interactions. The electroweak symmetry breaking is a very specific mechanism introduced in the SM by Higgs-Brout-Englert-Guralnik-Hagen-Kibble in 1964. This mechanism generates the existence of a spin-0 boson called the Higgs boson. In 2012, the two LHC collaborations CMS and ATLAS announced the discovery of the neutral boson of mass $\simeq 125$ GeV, $H(125)$, which had properties compatible with the Higgs boson. Since then all measurements of the $H(125)$ properties confirmed that it is of the nature of one Higgs boson. The LHC was designed to discover the Higgs boson, to explore physics at TeV scale and possibly discover new phenomena and to make precise measurements to confront the SM to a new domain.

The SM has so far resisted the confrontation to all existing measurements. The LHC experiments recently presented their results at the LHCP conference [18]: no significant deviation of the data with respect to SM prediction are present. This success of the SM needs to be tempered by the fact that the model has limitations and cannot explain the universe. For instance: the SM cannot explain the existence of Dark Matter and Dark Energy, the level of CP-violation in the SM is too small to explain the asymmetry of matter and anti-matter observed in the Universe.

In the following sections, I will try to make connections between the phenomenology of the Standard Model and the physics at the LHC. The introduction to the quantum field theory is presented to describe the interactions between fields in the framework of the Standard Model in the Section 1.1. The Lagrangian for the SM controls the dynamics and kinematics of the theory; the terms responsible for the interactions of the observed particles will be described. The mechanism to introduce the mass terms which are conserved in the gauge invariance of the theory is described; it is denoted as the spontaneous breaking of the symmetry which introduces a scalar boson, later called the Higgs boson. The developed mathematical framework for the Standard Model will be used to show how the model can be extended to provide predictions for particles beyond SM. The description of the proton-proton collisions, the parton interaction and the cross section of the main processes will be introduced, in Section 1.2. The production modes and decay channels of the observed Higgs boson will be discussed in Section 1.3. The theoretical motivations for models beyond SM will be introduced, and several models predicting new phenomena will be discussed in Section 1.4.

1.1 Standard Model phenomenology

The Standard Model (SM) is a common theoretical framework based on quantum field theory (QFT), which was developed to describe the nature and dynamics of the particles and their interactions. The SM includes three out of the four known interactions, namely: the electromagnetic interaction which is responsible for the atomic structure; the weak interaction which is responsible for the radioactive decay (β decay); and the strong interaction which is responsible for the atom nucleus cohesion. The gravity is not included in the theory, while possible extensions of the SM - beyond Standard Model (BSM) - propose solutions and introduce the force and its carrier, as will be discussed in Section 1.4.2.

The Standard Model mathematically describes all the observed elementary particles and force mediators with fields: scalars (spin 0), bi-spinors (spin 1/2) and vectors (spin 1). Besides spin, the particles are characterized by other quantum numbers such as the electric charge, weak isospin and colour. As the SM describes it, the matter consists in point-like fermions of spin 1/2 which interact by exchanging point-like gauge bosons of integer spin (0 or 1). These fermions comprise the charged leptons (e, μ, τ), the corresponding neutral neutrinos $(\nu_e, \nu_\mu, \nu_\tau)$, the quarks (u, d, c, s, t, b), and the antiparticles of each of the leptons and quarks. There are four bosons in the electroweak sector. The massless photon is the mediator of the electromagnetic interaction. The three massive $W^{\pm}$, Z^0 bosons carry the weak interaction. The strong interaction is mediated by eight massless gluons. The existence of these vector bosons is the direct result of the SM gauge symmetries[1]. The Standard Model of elementary particles, with the three generations of matter, gauge bosons in the fourth column, and the Higgs boson in the fifth are shown on Figure 1.1.1.

Gauge invariance is the basic principle of the SM in which each interaction is described in a gauge theory. The Lie algebra of the local gauge transformations group completely determines the nature and properties of the interactions between particles. The strong interaction is described by quantum chromodynamics (QCD) based on $\mathrm{SU}(3)_C$ group, where C corresponds to the colour quantum number, quarks and gluons are coloured objects, called **partons**; it is presented in Section 1.1.1. The electromagnetic interaction and the weak interaction is represented by $\mathrm{U}(1)_Y$ and $\mathrm{SU}(2)_L$ symmetry groups, where Y corresponds to hyper charge quantum number and L indicates that the coupling is only to left-handed fields. The electromagnetic and weak forces are unified to the electroweak theory, which is described in Section 1.1.2. The intensity of interactions between particles is quantified by the free parameters called **couplings**.

The requirement of the local gauge invariance of the Lagrangian forbids the bosonic and fermonic mass terms (will be shown later in the text), therefore all the particles are massless in the SM theory. The interaction with an extra field causes the symmetry to break, allowing mass terms and at the same time preserves the gauge invariance of the theory. The spontaneous symmetry breaking is realized by the Higgs mechanism in the electroweak theory; it will be discussed in Section 1.1.3. The Higgs boson acts as the mediator of a new class of interactions which, at the tree level, are coupled in

---

[1] Invariance under field transformation

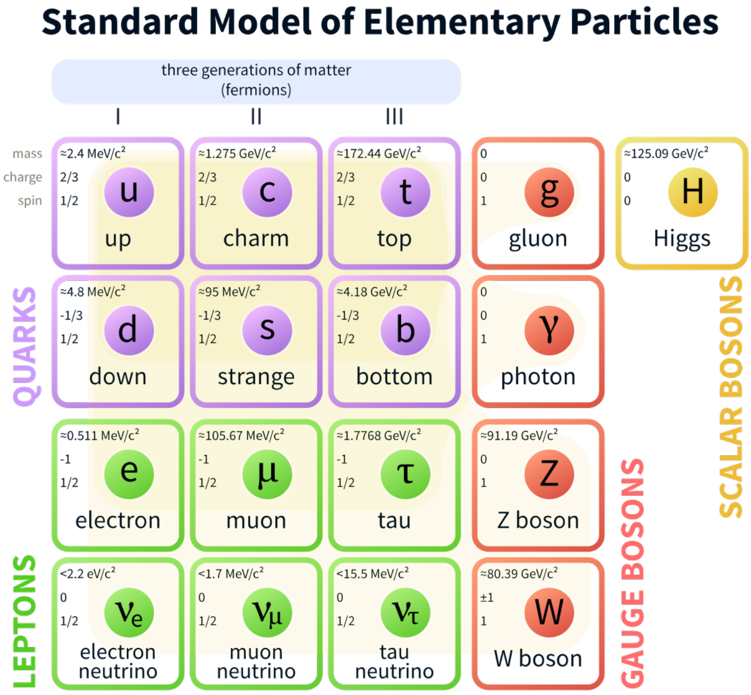


Figure 1.1.1: The particle content of the Standard Model, including the fermions (quarks and leptons), the gauge bosons, and the Higgs boson

proportion to the particle masses. The Higgs particle has now been observed at the LHC with $m_H \sim 125$ GeV, thus making a big step towards completing the experimental verification of the SM.

The SM is a renormalizable field theory, which means that the ultraviolet divergences that appear in loop diagrams can be eliminated by a suitable redefinition of the parameters already appearing in the bare Lagrangian: masses, couplings, and field normalizations.

1.1.1 Quantum chromo dynamics

The strong interaction, quantum chromodynamics (QCD), is a renormalizable gauge theory based on the group SU(3) with colour triplet quark matter fields [9, 10], described by the Lagrangian density:

$$\mathcal{L}_{SU(3)} = -\frac{1}{4}\sum_{A=1}^{8} F^{A\mu\nu}F^{A}_{\mu\nu} + \sum_{j=1}^{n_f} \bar{q}_{j\alpha}\, i\not{D}^{\alpha}_{\beta}\, q^{\beta}_{j}, \tag{1.1}$$

here q_j are the quark fields with n_f different flavours, and D_μ is the covariant derivative of the form:

$$D^\alpha_{\mu\beta} = \partial_\mu \delta_{\alpha\beta} + i g_s G^i_\mu \frac{\lambda^i_{\alpha\beta}}{2}, \tag{1.2}$$

where g_s is the QCD gauge coupling constant, α, β =1,2,3 are colour indices. The quarks transform according to the triplet representation matrices $\lambda^i/2$. The λ's are the Gell-Mann SU(3) matrices, normalised by $\mathrm{Tr}\lambda^i\lambda^j = 2\delta^{ij}$ The gluon field corresponds to G^i_μ, $i = 1, ..., 8$, and it enters the field strength tensor, describing the dynamics of the gluon field:

$$F^i_{\mu\nu} = \partial_\mu G^i_\nu - \partial_\nu G^i_\mu - g_s f_{ijk} G^j_\mu G^k_\nu, \tag{1.3}$$

the last term multiplies two gluon fields allows gluons to interact with each-other; the $f_{ijk}(i, j, k = 1, ..., 8)$ are the structure functions defined by

$$[\lambda^i, \lambda^j] = 2i f_{ijk} \lambda^k \tag{1.4}$$

The first term in 1.1 corresponds to three and four-point gluon self-interactions and the second to the quark interactions. QCD has the property of asymptotic freedom [11]: the coupling becomes weak at high energies or short distances (and accordingly increases at low energy/large distances). The strength of interaction between two partons increases with the relative distance, therefore it is not possible to observe a free parton but only color-neutral objects called **hadrons**. As soon as two coloured objects are pulled apart, the potential energy available will create new $q\bar{q}$ pairs that neutralize the original quarks colours: this property is denoted as **confinement**[12]. Therefore, the decay products of the QCD process with quarks and gluons create flows of hadrons forming **jets**; these are the jets which are reconstructed in the detector. The dynamics of the QCD processes will be discussed in details in Section 1.2.

1.1.2 The electroweak interactions

The electromagnetic interaction describes the dynamics of the charged fermions, it is based on a local gauge transformation of the U(1) symmetry group [13], with the following properties:

- it describes the interaction of a charged particle with the photon;
- the photon is massless (mass term is not gauge invariant);
- the gauge field does not have self-interactions.

The weak interaction is based on the SU(2) symmetry group, and the weak force is characterised by the following properties:

- it is capable of changing the flavour of quarks and leptons ;
- it has massive force-carriers;
- it violates the parity transformation **P**, the charge symmetry **C** and the combined **CP** [1] symmetry [14].

---

[1] Violation of P and C are maximal, while CP violation is much smaller

The maximal violation of C and P implies that the weak interaction to act differently on fermions depending on the projection of the spin onto the momentum direction, and by that property they are split into left- and right-handed fermions, ψ_L and ψ_R correspondingly. Under weak isospin SU(2) transformations the left-handed particles are weak-isospin doublets ($I_3 = \pm\frac{1}{2}$), whereas the right-handed are singlets ($I_3 = 0$) - therefore the weak interaction only acts on left-handed fermion. This property implies that the right-handed neutrino is sterile, since it has no charge for any of the interactions described by the SM: no electric charge, no weak isospin, no colour.

The unification of the electromagnetic and weak interactions is presented by the **electroweak theory**, which is based on the $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ group gauge transformation. The $\mathrm{SU}(2)_L$ group is associated with the weak isospin I_3 and the gauge invariance conditions introduces three gauge bosons $W^i_\mu, i = 1, 2, 3$. The $\mathrm{U}(1)_Y$ is associated with the hypercharge Y and one gauge boson B_μ. The electromagnetic charge can be related to weak isospin and hypercharge through the relation $Q = I_3 + \frac{Y}{2}$. B_μ couples both to the left- and right-handed components of the fermion fields, the W^i_μ fields only couple to the left-handed fermions. The invariance under $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ is obtained by the definition of two covariant derivatives, applied on left-handed and right-handed fermions in the Lagrangian describing the electroweak interaction:

$$\begin{aligned} D_{\mu,L} &= \partial_\mu + ig\frac{\sigma_i}{2}W^i_\mu + ig'\frac{Y}{2}B_\mu \\ D_{\mu,R} &= \partial_\mu + ig'\frac{Y}{2}B_\mu \end{aligned} \tag{1.5}$$

The g and g' coupling constants are the ones for $\mathrm{SU}(2)_L$ and $\mathrm{U}(1)_Y$, respectively and σ_i are the Pauli matrices. The Lagrangian of electroweak interaction can be presented as the following:

$$\mathcal{L}_{EW} = -\frac{1}{4}W^i_{\mu\nu}W^{\mu\nu,i} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} + i\bar{\psi}_L \not{D}_L \psi_L + i\bar{\psi}_R \not{D}_R \psi_R \tag{1.6}$$

where the field strength tensors correspond to the kinematic of the bosons and are defined as:

$$\begin{aligned} B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu \\ W^i_{\mu\nu} &= \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g\epsilon_{ijk}W^j_\mu W^k_\nu, \end{aligned} \tag{1.7}$$

where ϵ_{ijk} are the group structure constants which for SU(2) coincides with the totally antisymmetric Levi-Civita tensor, with $\epsilon_{123} = 1$.

The physical bosons are linear combinations of the $\mathrm{SU}(2)_L$ and $\mathrm{U}(1)_Y$ fields. The W^+ and W^- are responsible for the charged current interactions :

$$W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu) \tag{1.8}$$

while the Z^0 boson and the photon are the physical particles responsible for the neutral current interactions:

$$\begin{aligned} A_\mu &= W^3_\mu \sin\theta_W + B_\mu \cos\theta_W \\ Z^0_\mu &= W^3_\mu \cos\theta_W - B_\mu \sin\theta_W \end{aligned} \tag{1.9}$$

where θ_W is the Weinberg angle which defines how the B_μ and W^3_μ rotate and mix to form the observable A_μ, Z^0_μ fields. The Weinberg angle can be expressed in terms of the g and g' coupling constants $\cos\theta_W = \frac{g}{\sqrt{g^2+g'^2}}$.

Bosonic mass terms ($\frac{1}{2}m_X X_\mu X^\mu$) as well as fermionic mass terms ($m_i\psi_i\psi^i$) would break the local gauge invariance SU(2)$_L$ and U(1)$_Y$. Therefore at this stage of the Standard Model construction all the particles are massless. While the numerous experimental evidences show that both fermions and bosons have a non-zero mass, a mechanism to introduce mass to the particles while preserving the local gauge invariance is provided by the spontaneous breaking of the symmetry.

1.1.3 Spontaneous symmetry breaking

The problem of the origin of the mass of quarks and leptons is solved by introducing the Higgs-Brout-Englert-Guralnik-Hagen-Kibble mechanism [15, 16, 17]. The Higgs boson is a consequence of the electroweak symmetry breaking, which gives other particles mass through a dynamical mechanism - by introducing a new field. To give masses to the particle we define a complex SU(2)$_L$ doublet:

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \tag{1.10}$$

where ϕ^+ is a charged field with I$_3$ = +1/2 and ϕ^0 is a neutral field with I$_3$ = -1/2.

The Lagrangian to describe the dynamics of this field is written as:

$$\mathcal{L}_{EWSB} = (D_\mu\Phi)^\dagger(D^\mu\Phi) - V(\Phi), \tag{1.11}$$

where D_μ, the covariant derivative is given by: $D_{\mu,L} = (\partial_\mu + ig\frac{\sigma_i}{2}W^i_\mu + ig'\frac{Y}{2}B_\mu)$, is used to ensure gauge invariance under U(1)$_Y$⊗ SU(2)$_L$. The symmetry breaking is obtained through the form of the scalar potential V:

$$V(\Phi) = \frac{1}{2}\mu^2\Phi^\dagger\Phi + \frac{1}{4}\lambda(\Phi^\dagger\Phi)^2, \tag{1.12}$$

where it will be assumed that $\lambda > 0$, to ensure potential stability. The value of the potential $V(\Phi)$ in the vacuum (no excitation) is obtained when $\frac{\partial V}{\partial\Phi} = 0$, i.e. for:

$$\mu^2\,|\Phi| + \lambda\,|\Phi|^3 = 0, \tag{1.13}$$

i.e. when its module $|\Phi| = 0$ (if $\mu^2 > 0$) or when $|\Phi| = \pm\sqrt{-\frac{\mu^2}{\lambda}}$ (if $\mu^2 < 0$, the SU(2)$_L$ symmetry is broken only in this case).

Defining $\nu/\sqrt{2} = \sqrt{-\mu^2/\lambda}$ and under a particular gauge transformation, the ground state simply writes as:

$$\Phi = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \nu \end{pmatrix}, \tag{1.14}$$

This field has only a neutral and scalar component. This form of field explicitly breaks the SU$(2)_L$ invariance (electroweak symmetry breaking, or EWSB). The parameter ν is called the *vacuum expectation value.* The excitations from the ground state are given by:

$$\Phi = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \nu + H(x) \end{pmatrix}, \tag{1.15}$$

where H is a physical scalar field, which quantum excitation is called the Higgs boson. This field gives rise to masses of the weak bosons W, Z - put field Φ to term $(D_\mu\Phi)^\dagger(D^\mu\Phi)$, using definition of covariant derivative from 1.5 we obtain:

$$\begin{aligned}
(D_\mu\Phi)^\dagger(D^\mu\Phi) &= \left|\left(\partial_\mu + \frac{i}{2}g\sigma_i W^i_\mu + \frac{i}{2}g' B_\mu\right)\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \nu \end{pmatrix}\right|^2 \\
&= \frac{\nu^2}{8}\left|\left(g\sigma_i W^i_\mu + g' B_\mu\right)\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right|^2 \\
&= \frac{\nu^2}{8}\left|\begin{pmatrix} gW^1_\mu - igW^2_\mu \\ -gW^3_\mu + g' B_\mu \end{pmatrix}\right|^2 \\
&= \frac{\nu^2}{8}\left[g^2\left((W^1_\mu)^2 + (W^2_\mu)^2\right) + (-gW^3_\mu + g' B_\mu)^2\right]
\end{aligned} \tag{1.16}$$

If one use Equation 1.8, the term $W^\pm_\mu W^\mu_\pm$ equal to the first term of Equation 1.16: $(W^1_\mu)^2 + (W^2_\mu)^2$. Using the Z from Equation 1.9 and notion for the Weinberg angle one get the second term of Equation 1.16:

$$Z_\mu = (W^3_\mu \cos\theta_W - B_\mu \sin\theta_W) = \frac{1}{\sqrt{g^2 + g'^2}}(gW^3_\mu - g' B_\mu) \tag{1.17}$$

Therefor, Equation 1.16 provides mass terms[1] for the physical $W^\pm$ and Z^0 bosons. The A_μ field quantum excitation does not gain mass, as no corresponding mass term is found in the Lagrangian. The corresponding masses are:

$$\begin{aligned}
m_\gamma = 0, \quad m_Z &= \frac{\nu}{2}\sqrt{g^2 + g'^2} \\
m_W = \frac{g\nu}{2}, \quad m_H &= \sqrt{-2\mu^2} = \sqrt{2\lambda\nu^2}
\end{aligned} \tag{1.18}$$

the Higgs boson obtains a mass through the same type of couplings. Since λ is a free parameter of the theory, the Higgs mass is not predicted by the SM.

To give masses to fermions we use the Higgs boson field to define Yukawa couplings

---

[1] $\frac{1}{2}m_X X_\mu X^\mu$

y_X (where X is a fermion) into the Lagrangian:

$$\begin{aligned}\mathcal{L}_d &= -y_d(\bar{\psi_L}\Phi\psi_R) + h.c. \\ &= -\frac{1}{\sqrt{2}}y_d\nu\bar{d_L}d_R - \frac{1}{\sqrt{2}}y_d\bar{d_L}Hd_R + h.c.\end{aligned} \tag{1.19}$$

where y_d is the fermion Yukawa coupling which denotes the strength of the scalar field coupling. The first term gives the mass of the down-type quark $m_d = y_d\nu/\sqrt{2}$, while the second shows how fermions couple with strength y_d to the Higgs field H. Since they have right and left components, only the quarks and the charged leptons gain mass. A fermion X mass writes $m_X = y_X\nu/\sqrt{2}$. The neutrinos which are never observed to be right-handed are not given mass in the SM.

The U(1) symmetry and the SU(3) colour symmetry remain unbroken and therefore their carriers, photon and gluons, remain massless.

Standard Model summary

The SM describes the elementary particles and their interactions; it contains 19 free parameters that are necessary to describe the masses of the particles and the different couplings between the particles. Once the mass of the Higgs boson is measured, all the fundamental parameters of the Standard Model can be computed. This requires the precision measurement of the parameters of the Higgs boson, such as its mass and the couplings to fermions and bosons.

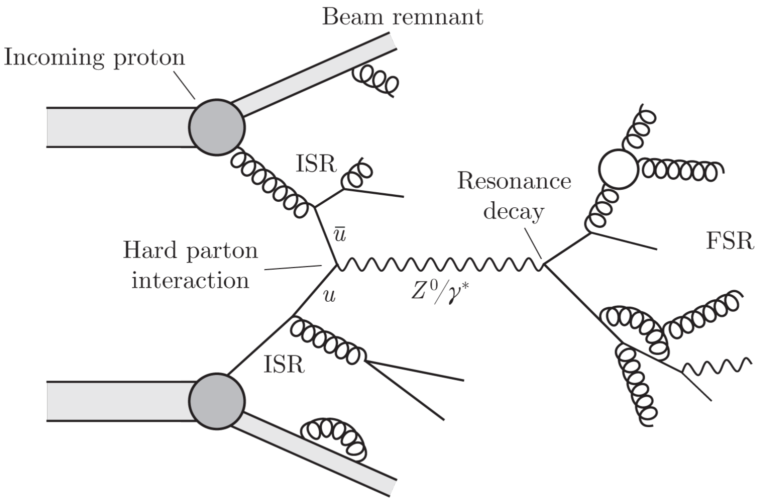


Figure 1.2.1: Schematic representation of a proton-proton hard scattering process [19].

1.2 Proton-proton interactions at LHC

Proton-proton collisions can be described by three types of interactions: elastic (∼25 mb) when the incoming protons remain, diffractive (∼10 mb) where the energy transfer is low, and non-diffractive (∼60 mb) when two partons from the protons interact with a high momentum transfer. I shall describe the hard scattering interactions as it is main process via the Higgs boson or any new object would be produced. The ATLAS trigger selection is defined to select events produced by hard scattering processes. Hard scattering interactions are described by perturbative QCD. A schematic representation of a hard scattering proton-proton collision is given in Figure 1.2.1.

The hard scattering is accompanied by the interactions of the remanent partons from the proton, leading to extra activity around the objects produced by the hard scattering. This is called the *Underlying Event* (UE). Particles from the UE contribute to the triggered event (see Figure 1.2.2).

A second contribution, called pile-up, is superimposed to the hard scattering process at LHC. Because of the very high density of protons at the collision points, more than one proton interact when the two LHC proton bunches cross each other at the center of the experiment. For instance, the average number of interactions per bunch crossing (μ) was 23 during the 2016 data taking campaign. These pile-up events are essentially produced by soft processes and described by non-perturbative QCD effects. Figure 1.2.3 shows an event display of Higgs candidate decaying to $2e2\mu$. The insert at the bottom shows the distribution of the 25 pile-up vertices measured by the ATLAS pixel detector.

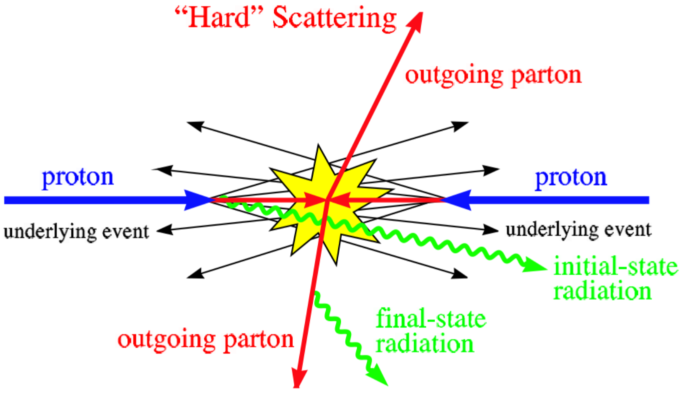


Figure 1.2.2: Schematic representation of a proton-proton hard scattering and of the underlying event [20].

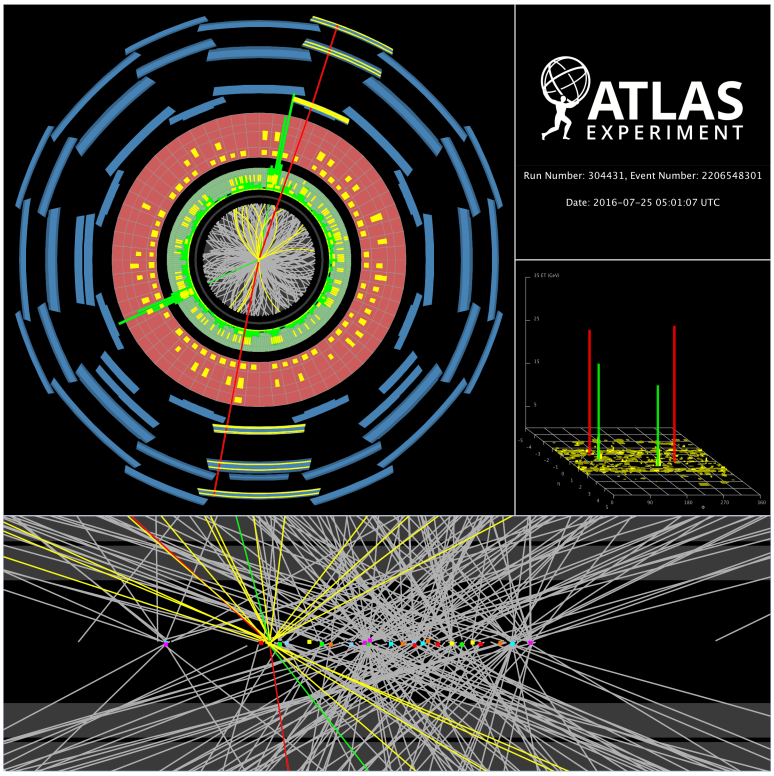


Figure 1.2.3: Event display of Higgs candidate decaying to $2e2\mu$ [31]. The insert at the bottom shows the distribution of the 25 pile-up vertices measured by the ATLAS pixel detector.

1.2.1 Hard process

The LHC provides high energy proton-proton collisions at the energy of 6.5 TeV per beam. The factorisation theorem [21] states that hadron-hadron collision can be split into two parts: one describing structure of the hadron and interaction of the partons. Therefore the cross-section $\sigma_{pp\to X}$ can be written as:

$$\sigma_{pp\to X} = \text{PDFs} \otimes \sigma_{hard\ scatter} \tag{1.20}$$

First term is denoted to the **parton distribution functions** (PDFs), they are universal property of the incoming hadron, they do not depend on the hard scattering process. Parton distributions are not calculable in perturbative QCD and have to be obtained from experiment. The second term, the partonic cross section, is denoted as **hard scattering**, it can be computed as a perturbation series in α_s and does not depend on the type of incoming hadron.

$$\sigma_{pp\to X} = \sum_{q,q'} \int dx_1 dx_2 f_q(x_1, Q^2) f_{q'}(x_2, Q^2) \times \hat{\sigma}_{qq'\to X}(x_1 s, x_2 s) \tag{1.21}$$

where x_1 and x_2 are the momentum fractions of the protons carried by the partons q and q' respectively and f_q $(f_{q'})$ represents the momentum fraction distribution of a parton q (q') which depends also on the four momentum of the process Q^2. The partonic cross-section can be expressed as a power series expansion of the α_s coupling constant:

$$\hat{\sigma}_{q\bar{q}\to X} = \underbrace{\hat{\sigma}_0}_{LO} + \underbrace{\alpha_s \hat{\sigma}_1}_{NLO} + \underbrace{\alpha_s^2 \hat{\sigma}_2}_{NNLO} \tag{1.22}$$

LO refers to the leading order, NLO to the next-to-leading order and NNLO to the next-to-next- to-leading order calculations. Figure 1.2.4 shows the predictions for some important Standard Model cross sections at $p\bar{p}$ and pp colliders, calculated at next-to-leading order in perturbation theory. The cross section of the QCD jet production has dominant contribution to the total cross section. The Higgs boson production cross section is about seven-eight orders of magnitude smaller then QCD jet production, three-four orders smaller than EW production of W and Z bosons and about two orders with respect to the top-quark production.

1.2.2 Parton density function

The parton distribution functions are defined as the probability of finding a parton in a proton with the momentum fraction x, at momentum transfer Q^2. The set of distributions $f_i(x, Q^2)$ describes how the momentum of the proton is shared between the individual partons (f_i = valence quark, sea quarks and gluons). Figure 1.2.5 displays an example of $xf(x, Q^2)$ distributions for the valence quarks u and d, the sea quarks $\bar{u}, \bar{d}, s, \bar{s}, b, \bar{b}$ and the gluon g for two different scales $Q^2 = 10$ GeV^2 and 10^4 GeV^2 . In most of the phase space the gluon PDF dominates.

Perturbative QCD predicts the evolution of the PDF with Q^2 through the Dokshitzer-

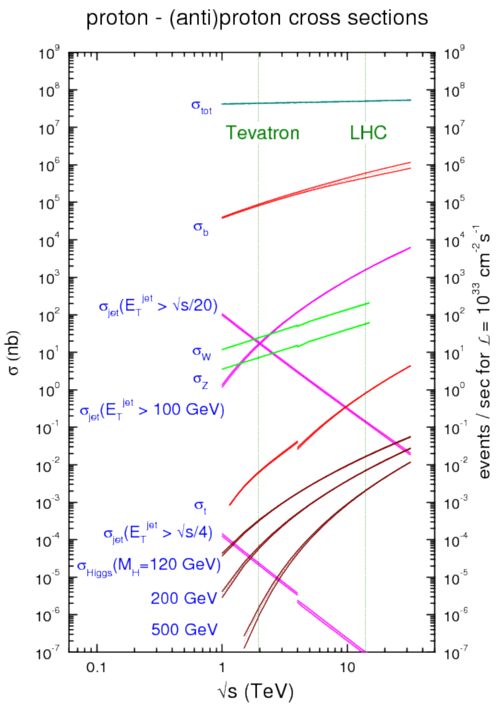


Figure 1.2.4: Main SM processes cross sections in hadronic collisions as a function of $\sqrt{s}$ (center- of-mass energy) [28].

Gribov-Lipatov- Altarelli-Parisi (DGLAP) equations[1]. However, the PDFs themselves are not calculated perturbatively but are derived by fitting the experimental data in fixed target and collider experiments.

1.2.3 Hadronization and fragmentation

The scattering of the proton constituents leading to outgoing partons (quarks and gluons) with large transverse momenta. The partons are coloured objects and due to the confinement can't propagate free, therefore additional $q\bar{q}$ pairs will be created to build colourless hadrons. The flow of the hadrons will constitute "jet" structure via a "fragmentation process". The process involves the production of hadrons and takes place at an energy scale where the QCD coupling constant is large and perturbation theory cannot be used. Fragmentation is therefore described using a QCD-motivated model with parameters that must be determined from experiment.

---

[1] The QCD evolution equation [22, 23].

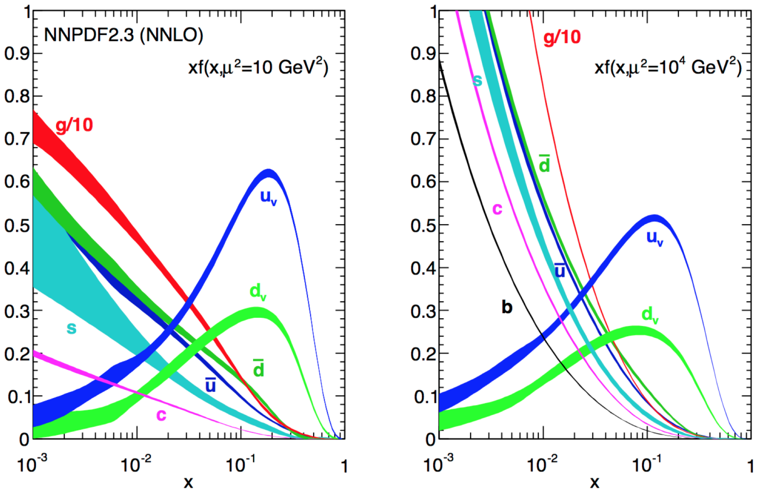


Figure 1.2.5: Parton distribution functions of the proton at next-to-leading order (NLO) for two different scales Q^2 as predicted by the NNPDF collaboration. The band represents the 68% confidence level [29].

1.2.4 Monte Carlo generators

Monte Carlo generators produce complete events starting from a proton-proton initial state. They are used standalone or with specialized generators that improve the description of certain final states. They have many parameters, some of which are related to fundamental parameters such as the QCD coupling constant and electroweak parameters, and some of which describe the models used to parametrize long distance QCD, soft QCD, and electroweak processes.

Sherpa uses an interface to Pythia's hadronization model and produces complete events [24]. It provides approximations for final states with large numbers of isolated jets. Sherpa generates underlying events using a simple multi-parton interaction model based on that of Pythia [25]. It can include high order LO matrix elements and also NLO computations for some processes. Sherpa allows to merge matrix elements with the parton-shower simulation [26], using different prescriptions (e.g. ME+PS@LO [27])

MC@NLO uses fundamental (hard scattering) processes evaluated at next to leading order in QCD perturbation theory [30]. MC@NLO includes one loop corrections, with the consequence that events appear with negative and positive weight which must be taken into account when they are used. MC@NLO has been used for large-scale production of top, W and Z events.

1.3 Higgs physics at LHC

A new particle, with the properties consistent with the SM Higgs boson has been observed by the ATLAS and CMS collaborations in the mass region around 126 GeV [105, 106]. In this section, the Higgs boson production and decay modes will be discussed. The properties of the recently observed Higgs boson will be presented. The review of the parameters of the SM will be presented, comparing predicted and measured values.

1.3.1 Higgs boson production

At the LHC, the SM Higgs boson is produced in proton-proton collisions through four dominant processes: gluon gluon fusion (ggF), vector boson fusion (VBF), associated production with a W or Z boson (VH), or associated production with a pair of top quarks($t\bar{t}H$). The Feynman diagrams of the processes are presented on Figure 1.3.1.

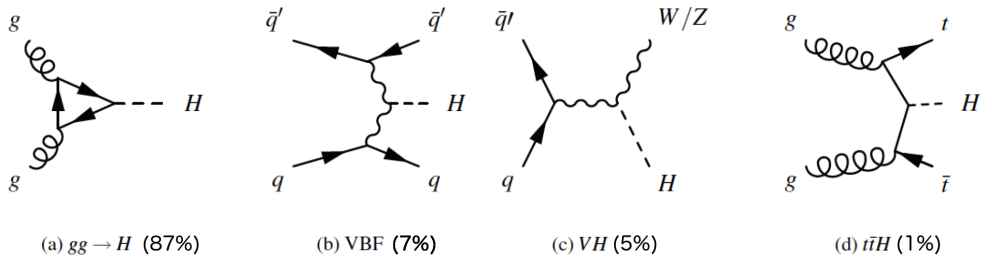
- The Higgs boson cannot couple directly to massless gluons, therefore the process is mediated by triangular loops of heavy quarks. The gluon-gluon interaction is the leading production process in pp collisions since gluons has the largest contribution in the parton density function, therefore the $\boldsymbol{ggF}$ process has largest contribution in the Higgs production. Cross sections of the ggF process are calculated at N3LO QCD and NLO EW accuracies [33].

- The $\boldsymbol{VBF}$ has the second largest cross section at the LHC. It has distinguishable signature in the detector with two high p_T jets which can be exploited at the LHC to discriminate the VBF signal against backgrounds and other signals. The VBF cross sections are calculated at (approx.) NNLO QCD and NLO EW accuracies [34, 35].

- The associated production ($\boldsymbol{VH}$) corresponds to the process when the Higgs boson is radiated from an off-shell $W^{\pm}$ or Z weak boson from quark-antiquark annihilation. The decay products of the associated bosons (presence of lepton(s) and/or missing transverse energy) are used to tag the event. Cross sections of the VH are calculated at NNLO QCD and NLO EW accuracies [36].

- The associated production with a pair of top-quarks ($\boldsymbol{ttH}$) process is similar to the gluon gluon fusion process. It has the smallest cross section, an order of magnitude smaller than ggF due to the presence of two real top-quarks in the final state. The top decays lead to high jet multiplicity in the final state, leptons and transverse missing energy, which makes this production mode challenging to explore due to low signal over background ratio. Cross sections of the ttH are calculated at NLO QCD and NLO EW accuracies [37].

Values of the cross section for each process for the Higgs boson at $\sqrt{s}$ =13 TeV as a function of mass are presented on Figure 1.3.2 (left).



(a) $gg \to H$ (87%) (b) VBF (7%) (c) VH (5%) (d) $t\bar{t}H$ (1%)

Figure 1.3.1: Leading-order Feynman diagrams of the four main SM Higgs boson production processes with the fraction of the contribution to the total cross section. (a) gluon gluon fusion. (b) Vector Boson Fusion. (c) associated production with a vector boson. (d) associated production with a pair of top quarks.

1.3.2 Higgs boson decay channels

The Higgs boson can decay to several channels. The branching fractions are calculated through the computation of all the partial decay widths in the HDecay [41] and Prophecy [42] programs, including all kinematically allowed channels and all relevant higher-order QCD corrections to decays into quark pairs and into gluons. The branching fractions (**BRs**), as a function of the Higgs boson mass, are represented in Figure 1.3.2 (right) for the relevant decay modes.

The Higgs boson decays mainly to pairs of b- and c-quarks, pairs of W and Z bosons, pair of gluons and $\tau\tau$. Those processes has largest production rate, the final states are composed of hadronic jets from quarks and the decay products of the bosons.

The diphoton channel has a considerably lower branching fraction than the other channels, but it is actual final state. For example, the branching fraction to diphoton is 2.27×10^{-3}, while the branching fraction to ZZ is one order magnitude larger, 2.619×10^{-2}. To obtain the rate of $ZZ \to 4e$ one needs to take into account the Z decay fraction to electrons, which is 3.36×10^{-2} for each of the Z. Therefore, comparing the rate of diphoton events to the one of the channel with four electrons one gets 2.27×10^{-3} versus 2.95×10^{-5}. Despite the fact that the γ is mass-less, the $H \to \gamma\gamma$ decays is possible via loop diagrams containing massive charged particles like $W^{\pm}, b$ and t - since the Higgs couplings to fermions is proportional to the mass, contribution of light fermions to the loop are neglected. The Feynman diagrams for the Higgs boson decays to a pair of photons are presented on Figure 1.3.3.

The diphoton channel played a crucial role in the discovery of the Higgs boson. This decay process will be used for search for new scalar resonances. This search is described in details in Chapter 5.

1.3.3 Higgs boson discovery and success of the Standard Model

The Higgs boson, which was predicted in 1964, has been discovered at the LHC in 2012 by the ATLAS and CMS collaborations. The measurements of the Higgs boson

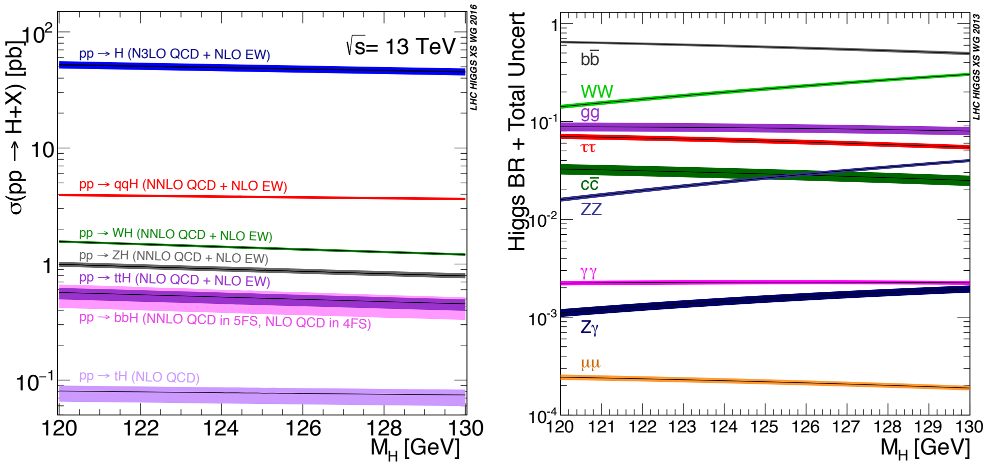


Figure 1.3.2: Standard Model Higgs boson production cross sections at $\sqrt{s} = 13$ TeV (left) and decay branching ratios (right) [39].

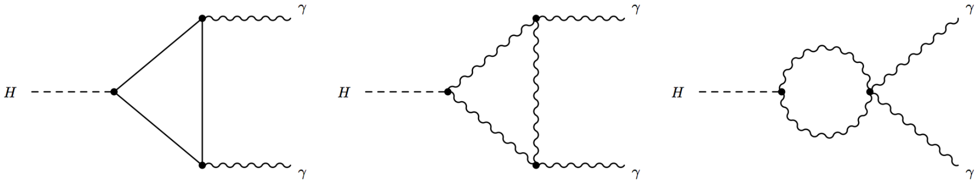


Figure 1.3.3: Examples of leading-order Feynman diagrams for Higgs boson decays to a pair of photons.

mass in the diphoton and ZZ decaying to four leptons channels corresponds to m_H = 125.09±0.21(stat.)±0.11(syst) GeV with the full combined Run-1 dataset of two experiments and the combination presented on Figure 1.3.4 [43].

The SM has been successfully tested at an impressive level of accuracy and provides at present our best fundamental understanding of the phenomenology of particle physics. Various experimental facilities (LEP, Tevatron, LHC) have allowed us to measure the parameters of the Standard Model with incredible precision, as it is summarised in Table 1.3.1. Figure 1.3.5 (left) shows agreement of the parameters with the data. The measured couplings of the Higgs boson to fermions and bosons are in excellent agreement with the behaviour expected in the SM, as shown on Figure 1.3.5 (right).

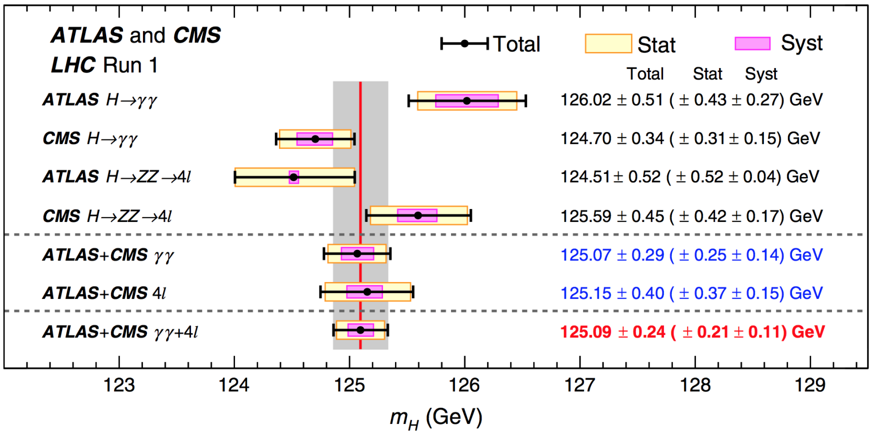


Figure 1.3.4: Summary of Higgs boson mass measurements from the individual analyses of ATLAS and CMS and from the combined analysis $H \to \gamma\gamma$ and $H \to ZZ \to 4l$ [43].

Quantity	Measurement [GeV]	Standard Model Predictions [GeV]	Pull, [σ]
m_t	173.34 ± 0.81	173.76 ± 0.76	-0.5
M_W	80.387 ± 0.016	80.361 ± 0.006	1.6
Γ_W	2.046 ± 0.049	2.089 ± 0.001	-0.9
M_Z	91.1876 ± 0.0021	91.1880 ± 0.0020	-0.2
Γ_Z	2.4952 ± 0.0023	2.4943 ± 0.0008	0.4

Table 1.3.1: The measurements of the Standard Model parameters, compared to the predicted values, where pull is the deviation of the measured values from predicted in terms of standard deviations [44].

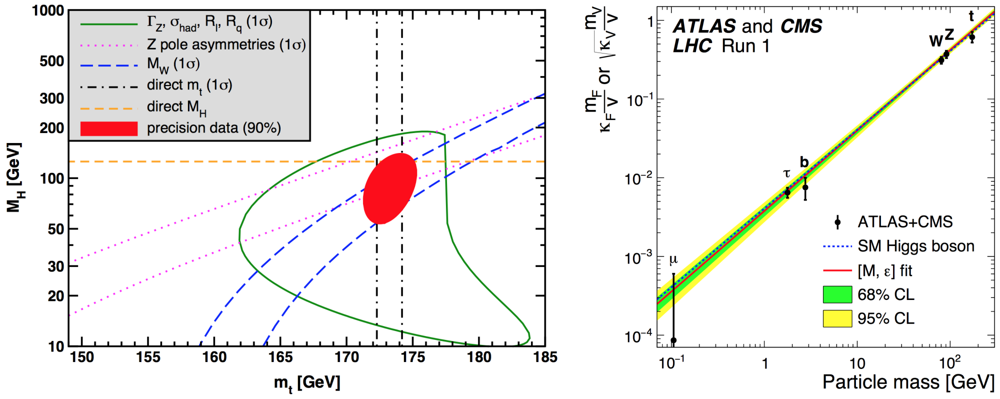


Figure 1.3.5: Success of the Standard Model. Left: summary of the measured parameters of the model (W, Z, H bosons, top quark) and their dependence on the Higgs boson (M_H) and top quark (m_t) masses [44]. Right: Best fit values of the couplings of fermions ($\kappa_F \frac{m_F}{\nu}$) and the weak bosons ($\sqrt{\kappa_V}\frac{m_V}{\nu}$) as a function of particle mass for the combination of ATLAS and CMS data, where $\nu = 246$ GeV is the vacuum expectation value of the Higgs field. 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 $[M, \varepsilon]$ phenomenological model with the corresponding 68% and 95% CL bands [45].

1.4 Beyond Standard Model

Despite the fact that the SM is a very successful theory, describing particle physics phenomena at energies up to the TeV scale, and supported by all experimental observations, it still has open questions in the description of Nature. The SM does not explain the gravitational force and it does not provide candidates for the dark matter and invisible energy inferred by cosmological observations. Matter-antimatter asymmetry in the universe is unexplained.

Several new physics extensions attempt to account for these observations. Examples of extensions are *supersymmetry*, where each SM particle is endowed with a superpartner, or *extra dimensions*, where one or several extra spatial dimensions are added to the already existing three. The Minimal Supersymmetric Standard Model (MSSM) contains two Higgs doublets model and it is one of the simplest possible extensions of the SM. The MSSM predicts additional scalar particles with the similar properties as the Higgs boson. An introduction will be presented in Section 1.4.1. The original proposal for extra dimensions by Kaluza and Klein where attempts to unify electromagnetism with gravitation, will be briefly discussed in Section 1.4.2.

1.4.1 Two Higgs Doublet Model

The Two Higgs Doublet Mode (2HDM) model is a simple extension of Standard Model, where an additional Higgs doublet is added [1, 3, 4]. One of the motivations for the 2HDMs is that it can generate a baryon asymmetry in the Universe, due to the flexibility of their scalar mass spectrum and the existence of additional sources of CP violation. There has been many works on baryogenesis in the 2HDM [46].

In theories with two Higgs doublets, the Yukawa couplings are:

$$V_{Yukawa} = -\sum_{i=1,2} (Q\Phi_i y_i^u \bar{u} + Q\Phi_i y_i^d \bar{d} + L\Phi_i y_i^e \bar{e} + h.c.), \tag{1.23}$$

where y_i^u (y_i^d) corresponds to couplings for up-type (down-type) quarks and y_i^e is couplings for charged leptons.

In the 2HDM the flavour-changing neutral currents (FCNC) are allowed at tree level, which creates a potential problem of the model, since FCNC are not observed [5]. Depending on which type of fermions couples to which doublet Φ (by convention up-type quarks are always taken to couple to Φ_2), one can divide two-Higgs-doublet models into the following tupes:

- Type I: in which $y_1^{u,d,e} = 0$; all fermions couple to one doublet.
- Type II: in which $y_1^u = y_2^d = y_2^e = 0$; the up-type quarks couple to one doublet and the down-type quarks and leptons couple to the other.
- Type III: in which $y_1^u = y_1^d = y_2^e = 0$; quarks couple to one doublet and leptons to the other.

- Type IV: in which $y_1^u = y_1^e = y_2^d = 0$; up-type quarks and leptons couple to one doublet and down-type quarks couple to the other.

In the following, in order to avoid FCNC, only the types I and II are considered. This is achieved by imposing a discrete or continuous symmetry. For instance, for the Type-I 2HDM model, one assumes the symmetry of type $\Phi_1 \to -\Phi_1$, where all fermions with same quantum numbers couples to the same Higgs multiplet; in that case FCNC is absent. Also, it is assumed that CP is conserved to distinguish between scalars and pseudoscalars.

Exploiting the SM mathematical framework described in Section 1.1, we introduce a new potential $V(\Phi)$ similarly to the one used in Equation 1.12. The most general scalar potential for two doublets Φ_1 and Φ_2 with hypercharge +1 is:

$$\begin{aligned} V(\Phi) =& m_{11}^2\Phi_1^\dagger\Phi_1 + m_{22}^2\Phi_2^\dagger\Phi_2 - m_{12}^2(\Phi_1^\dagger\Phi_2 + \Phi_2^\dagger\Phi_1) + \frac{\lambda_1}{2}(\Phi_1^\dagger\Phi_1)^2 + \frac{\lambda_2}{2}(\Phi_2^\dagger\Phi_2)^2 \\ &+ \lambda_3\Phi_1^\dagger\Phi_1\Phi_2^\dagger\Phi_2 + \lambda_4\Phi_1^\dagger\Phi_2\Phi_2^\dagger\Phi_1 + \lambda_5\Big[(\Phi_1^\dagger\Phi_2)^2 + \Phi_2^\dagger\Phi_1)^2\Big], \end{aligned} \tag{1.24}$$

where m_{12}^2 and $\lambda_k, k = 1...5$ (Higgs self-couplings) are free and real parameters.

For a region of parameter space, the minimization of this potential gives

$$\langle\Phi_1\rangle_0 = \begin{pmatrix} 0 \\ \frac{\nu_1}{\sqrt{2}} \end{pmatrix}, \langle\Phi_2\rangle_0 = \begin{pmatrix} 0 \\ \frac{\nu_2}{\sqrt{2}} \end{pmatrix}, \tan\beta \equiv \frac{\nu_2}{\nu_1} \tag{1.25}$$

The angle β is the rotation angle which diagonalizes the mass-squared matrices of the charged scalars and of the pseudoscalars. The angle α is defined to be the rotation angle which diagonalizes the mass-squared matrix of the CP even neutral scalars.

With two complex scalar SU(2) doublets there are eight fields:

$$\Phi_a = \begin{pmatrix} \phi_a^\dagger \\ (\nu_a + \rho_a + i\eta_a)/\sqrt{2} \end{pmatrix}, a = 1, 2. \tag{1.26}$$

Three of those gives mass to the $W^\pm$ and Z^0 gauge bosons; the remaining five are physical scalar ("Higgs") fields. There is:

- the CP even neutral scalars h and H
- the CP odd pseudoscalar A
- and two charged Higgs bosons $H^\pm$

The two parameters α and β determine the interactions of the various Higgs fields with the vector bosons and (given the fermion masses) with the fermions; they are thus crucial in discussing phenomenology. Those parameters are the rotational angles which diagonalize the mass-squared matrices of the scalars.

The vacum expectation values (vev) of the doublets from Equation 1.26 can be expressed as $\nu_1 = \nu\cos\beta$, $\nu_2 = \nu\sin\beta$ and $\sqrt{\nu_1^2 + \nu_2^2} = \nu_{SM} = 246$ GeV. The physical

	2HDM I	2HDM II
hVV	$\sin(\beta-\alpha)$	$\sin(\beta-\alpha)$
hQu	$\cos\alpha/\sin\beta$	$\cos\alpha/\sin\beta$
hQd	$\cos\alpha/\sin\beta$	$-\sin\alpha/\cos\beta$
hLe	$\cos\alpha/\sin\beta$	$-\sin\alpha/\cos\beta$
HVV	$\cos(\beta-\alpha)$	$\cos(\beta-\alpha)$
HQu	$\sin\alpha/\sin\beta$	$\sin\alpha/\sin\beta$
HQd	$\sin\alpha/\sin\beta$	$\cos\alpha/\cos\beta$
HLe	$\sin\alpha/\sin\beta$	$\cos\alpha/\cos\beta$

Table 1.4.1: The tree-level couplings of the neutral Higgs bosons h and H to up- and down-type quarks, leptons, and massive gauge bosons relative to the SM Higgs boson couplings as functions of α and β in the two types of 2HDM [7].

	2HDM I	2HDM II
hVV	1	1
hQu	1	1
hQd	1	1
hLe	1	1
HVV	0	0
HQu	-1/$\tan\beta$	-1/$\tan\beta$
HQd	-1/$\tan\beta$	$\tan\beta$
HLe	-1/$\tan\beta$	$\tan\beta$

Table 1.4.2: The tree-level couplings of the neutral Higgs bosons h and H to up- and down-type quarks, leptons, and massive gauge bosons relative to the SM Higgs boson couplings as functions of α and β in the two types of 2HDM in case of **alignment limit** ($\beta-\alpha = \pi/2$).

scalars are a lighter h and a heavier H, which are orthogonal combinations of ρ_1 and ρ_2:

$$\begin{aligned} h &= \rho_1 \sin\alpha - \rho_2 \cos\alpha, \\ H &= -\rho_1 \cos\alpha - \rho_2 \sin\alpha. \end{aligned} \tag{1.27}$$

the Standard-Model Higgs boson would be:

$$\begin{aligned} h^{SM} &= \rho_1 \cos\alpha + \rho_2 \sin\alpha, \\ &= h \sin(\beta-\alpha) - H \cos(\beta-\alpha). \end{aligned} \tag{1.28}$$

The coupling of the light Higgs, h, to either WW or ZZ is the same as the Standard-Model coupling times $\sin(\beta-\alpha)$ and the coupling of the heavier Higgs, H, is the same as the Standard-Model coupling times $\cos(\beta-\alpha)$, as presented in Table 1.4.1. Therefore, in case when $\cos(\beta-\alpha) = 0$, the so-called *alignment limit*, the lighter CP even h has couplings like the Higgs boson of the Standard Model, as shown in Table 1.4.2.

Diphoton resonances

Searches for heavy scalar (H) in diphoton channel in general are suppressed by large partial decay widths into WW, ZZ and $t\bar{t}$. The coupling values of the second boson to the SM particles are constrained by the resonance already observed at 125 GeV. This allows to probe alignment limit of the 2HDM-type models, where the couplings of the second scalar particle to the VV are suppressed, as shown in Table 1.4.2. Under these assumptions, the second resonance is expected to have sizeable branching fractions in the $\gamma\gamma$ final states through the up-type quarks loops [5]. In the 2HDM Type-II model for low values of $\tan\beta$ parameter the decay modes to bottom-type quarks ($b\bar{b}$) and leptons ($\tau\tau$) are suppressed with respect to $\gamma\gamma$. This enhancement of diphoton decay channel is valid up to $t\bar{t}$ threshold, above which the $t\bar{t}$ decay mode should become dominant.

In 2015 the LHC resumed operation after a long shutdown of two years. The centre of mass energy had increased from 8 to 13 TeV, opening a new phase space to search for new resonances. As part of my PhD, I searched for a new resonance decaying to a pair of photons, as described in Chapter 5.

1.4.2 Extra dimension theory

One of the questions unsolved in the SM, is the so called ”hierarchy problem”: why the gravity is so weak (the weak force is 10^{24} times as strong as gravity)? Extra dimensions are proposing additional space or time dimensions beyond the standard (3 + 1) typical of our observed space-time, to solve the this problem, and were introduced in the Kaluza-Klein theory.

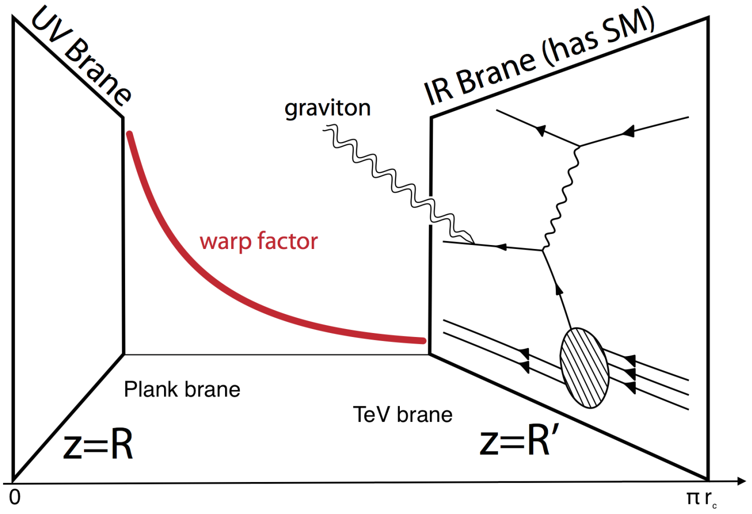


Figure 1.4.1: Cartoon of the RS scenario with a brane-localized SM. The warp factor, $(R/z)^2$, causes energy scales to be scaled down towards the TeV brane [48].

Warped extra dimensions, such as those proposed by the Randall-Sundrum model (RS) [47], introduce separation of two 4-D hypersurfaces (branes) by a small extra (5th) dimension. In their original model, our universe lives on one brane, the ”TeV brane” with negative energy density, the fields of the Standard Model are generally confined to this brane. There exists a second, ”Planck brane” of positive energy density separated by the extra dimension, as shown on Figure 1.4.1. The energy density of the branes warps the space-time in the 5-D bulk. This causes gravity which is a weak force corresponding to a high mass scale on the TeV brane to become a strong force corresponding to a low mass scale on the Planck brane. Given this space-time configuration, TeV scales are naturally generated from the Planck scale due 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 Λ_π defined as:

$$\Lambda_\pi = \overline{M}_{\mathrm{Pl}}\ exp(-k\pi r_c) \tag{1.29}$$

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. This effectively solves the hierarchy problem and predicts that there should be exotic TeV-scale spin 2 states (Kaluza-Klein or KK gravitons). In the minimal RS model, gravitons are the only particles that can propagate in the bulk. 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. A striking signature of the RS model at hadron colliders would be graviton production [49], followed by their decay to pairs of SM fermions or bosons. The decay $G^* \to \gamma\gamma$ is a particularly interesting example, since observation of a resonance in the diphoton final state would rule out some possible interpretations, such as a Z' boson.

The search for the graviton decaying to pair of photons has been performed by the ATLAS collaboration [116], using the same dataset used for the search for scalar resonance presented in Chapter 5. Different kinematic selections and treatment of the background with respect to spin-0 analysis are introduced. No significant deviation from the Standard Model background-only hypothesis is observed. Upper limits on the spin-2 RS graviton cross section times branching ratio to two photons as a function of the mass and coupling $k/\overline{M}_{\mathrm{Pl}}$ are set.

Chapter 2

The LHC and the ATLAS Detector

Contents

The ATLAS experiment is one of the four majour experiments at the Large Hadron Collider (LHC) at CERN. It is a general-purpose particle physics experiment run by an international collaboration involving more than 3000 physicists from 182 institutions of 38 countries. It is designed to exploit the full discovery potential and the huge range of physics opportunities that the LHC provides.

The structure of LHC machine and its performance will be discussed in Section 2.1. The description of ATLAS detector with a brief presentation of the subsystems and their role are presented in Section 2.2. A more detailed overview of the Liquid Argon (LAr) Electromagnetic calorimeter, its performance and read-out electronics are summarised in Sections 2.3.2, 2.3.3 and 2.3.4 correspondingly. The description of the general format of the data stream (Event Format) and my contribution to the tool to access and decode the raw data stream is given in Section 2.3.5. Section 2.4 is devoted to the description of the Monte Carlo simulation of the ATLAS detector and of the proton-proton physics precesses.

2.1 The Large Hadron Collider (LHC)

2.1.1 LHC Machine

The Large Hadron Collider (LHC) is the world's largest instrument for Particle Physics research. It is a two-ring-superconducting-hadron accelerator, designed for proton-proton collisions at the center-of-mass energy $\sqrt{s}$ = 14 TeV with an instantaneous luminosity of 10^{34} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and for heavy ions (Pb) collisions with an energy of 2.8 TeV per nucleon and a luminosity up to 10^{27} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The LHC is located near Geneva in a 27 kilometer circular tunnel, about 100 meters below ground level. The tunnel has eight straight sections, 528 m long each for experimental or utility insertions (denoted as Points), and eight arcs 2.45 km each containing the magnet system.

The LHC has performed very well; in following sections I will briefly describe the key elements of the collider for accelerating and steering the beams. The high performance accelerating cavities are used to push the beam and the magnetic field at the limit of technology - to bend the protons around the ring. Each of the parts of LHC is state of the art, technical details can be found in reference [50, 51].

2.1.2 LHC layout and beam facility

To produce collisions in the center of the LHC detectors, the two beams follow a long path through the accelerating complex, as it is presented on Figure 2.1.1. The protons are first produced by ionising hydrogen; the resulting protons are first accelerated to an energy of 750 keV. They are further accelerated in a linear accelerator, LINAC 2, to an energy of 50 MeV. The protons are split into four parts and injected into the Proton Synchrotron Booster (PSB), a 4-layer ring which accelerates them up to an energy of 1.4 GeV. In the Proton Synchrotron (PS) circular accelerator, the protons are collected to form *bunches* which are brought to an energy to 26 GeV. Bunches are injected to the Super Proton Synchrotron (SPS), where they are accelerated to an energy of 450 GeV. At that point, bunches are prepared for the LHC ring. Each proton gains 485 keV in one turn, and continue to ramp up to operating energy - 6.5 TeV per beam. The two beams at full intensity will consist of 2808 bunches, each of them containing 1.15×10^{11} protons.

The layout of LHC consists of four experimental and four utility insertions, as shown on Figure 2.1.2. The two high luminosity experiments, ATLAS [56] and CMS [57], are located at diametrically opposite straight sections, Point 1 and 5 respectively. Two more experimental insertions, caverns of ALICE [58] and LHCb [59] detectors, are located at Point 2 and Point 8, which also include the injection systems. The beams travel in opposite directions in an ultrahigh vacuum inside two separate beam pipes, Beam 1 is injected clockwise at Point 2 and Beam 2 is injected anti clockwise at Point 8. The two beams, which are share the same vacuum chamber, are squeezed and guided to the interaction point, as shown on Figure 2.1.3.

The collimation systems are important to provide stable, uniform and focused beams. Particles with a large momentum offset are scattered by the primary collimator (sequence of quadrupoles and dipoles) in Point 3, and particles with a large betatron

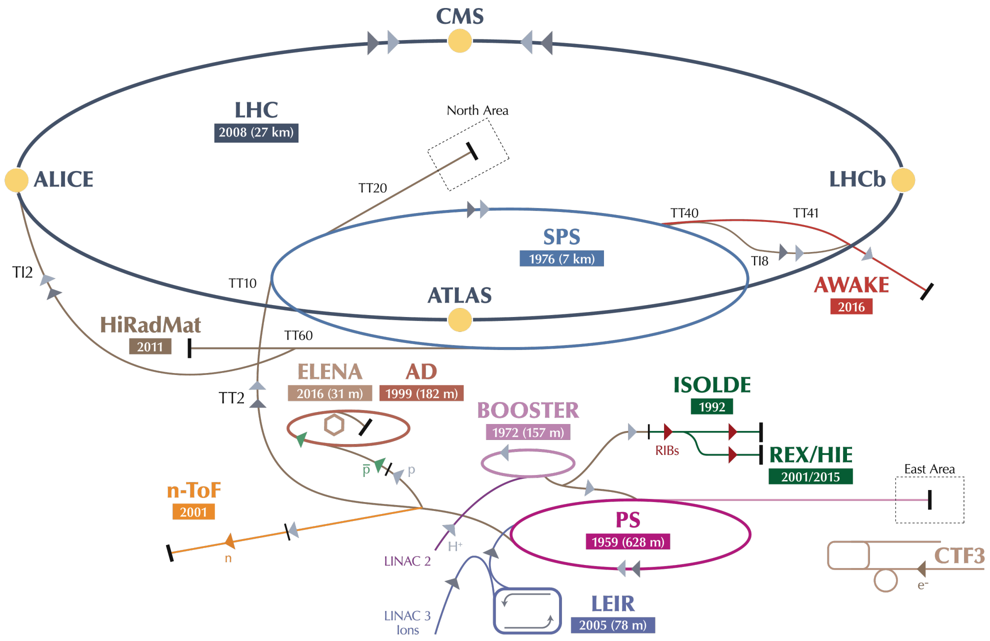


Figure 2.1.1: Overview of the CERN accelerator complex [52].

amplitudes (oscillation around the circular orbit) are scattered by the primary collimator in Point 7.

Two **radio frequency (RF) systems** are key elements to ramp up the beam energy. RF super-conducting cavities are operating at 400 MHz and producing an intense accelerating electric field of 5 MV/m. Cavities are made of Copper and Niobium films cooled to an operating temperature of 4.5 K. They are located at Point 4, one independent system for each LHC beam.

The **beam dump** serves to absorb the beams when they are extracted away from the machine (located in Point 6).

2.1.3 LHC magnets

Beams are guided around the accelerator ring by a strong magnetic field maintained by superconducting electromagnets. Various types of magnets have been designed for different purposes: dipole magnets bend the path of a beam of particles to the curvature radius of the LHC ring; quadrupole magnets act like lenses to focus a beam; sextupole, octupole and decapole magnets are used to correct for small imperfections in the magnetic field at the extremities of the dipoles.

The LHC is using superconducting Niobium-Titanium (Nb-Ti) coils, efficiently conducting electricity without resistance or loss of energy. This requires the magnets to be cooled to a temperature below 2 K, using superfluid helium, to allow to operate at fields above 8 T. The coils are surrounded by non-magnetic 'collars' of austenitic steel, a material that combines the required properties of good thermal contraction and

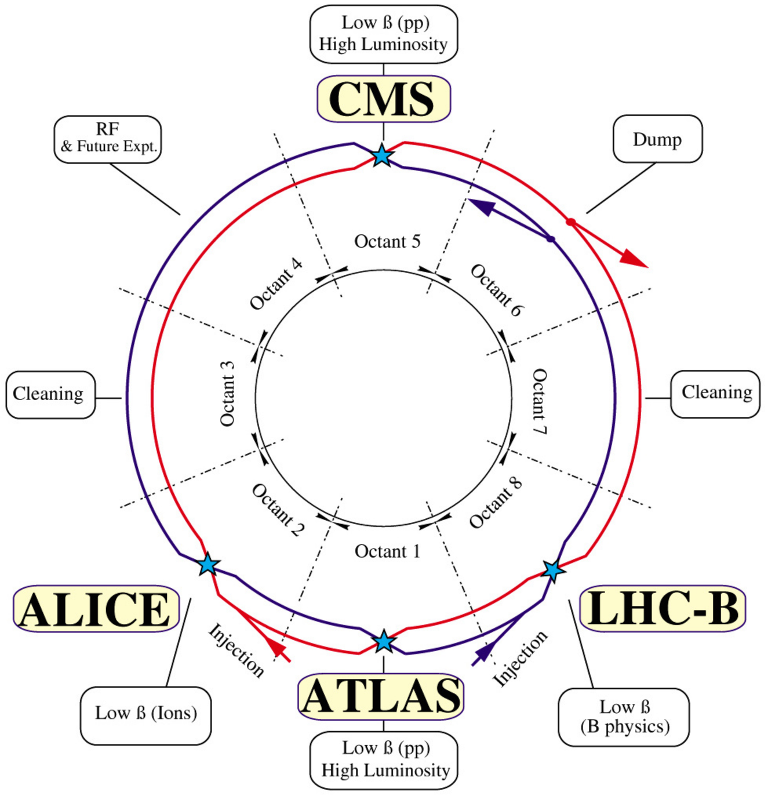


Figure 2.1.2: Schematic layout of the LHC (Beam 1- clockwise, Beam 2 - anti-clockwise) [53].

magnetic permeability. The collars hold the coils in place against the strong magnetic forces that arise when the coils are at full field - the force loading 1 m of dipole is about 400 tonnes. The magnetic system of LHC consists of 1232 dipoles (Fig. 2.1.4), 15 meter long, and 392 quadrupole magnets, each 5-7 meters long.

2.1.4 LHC performance

The luminosity is the quality factor for colliders, because the number of collected events is proportional to the luminosity multiplied by cross section:

$$N_{events} = \mathcal{L} \times \sigma. \tag{2.1}$$

The rate of useful interactions (R_{inel}) is related to the inelastic cross section of the proton-proton collision(σ_{inel}), and luminosity can be expressed as [60]:

$$\mathcal{L} = \frac{R_{inel}}{\sigma_{inel}} \tag{2.2}$$

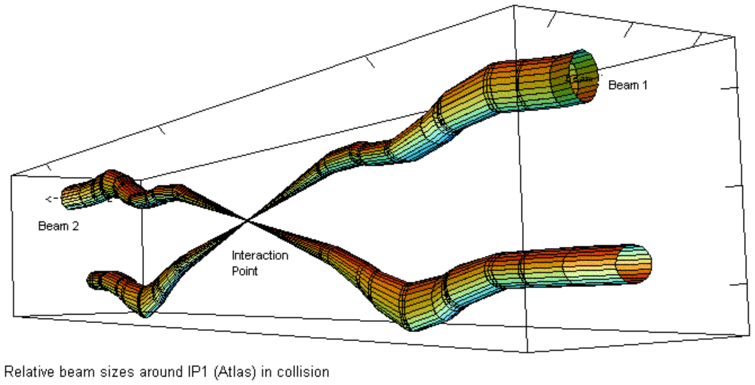


Figure 2.1.3: The beams squeezed down to 64 μm before the interaction point to increase the chances of a collision, providing in average about 23 interactions per bunch crossing.

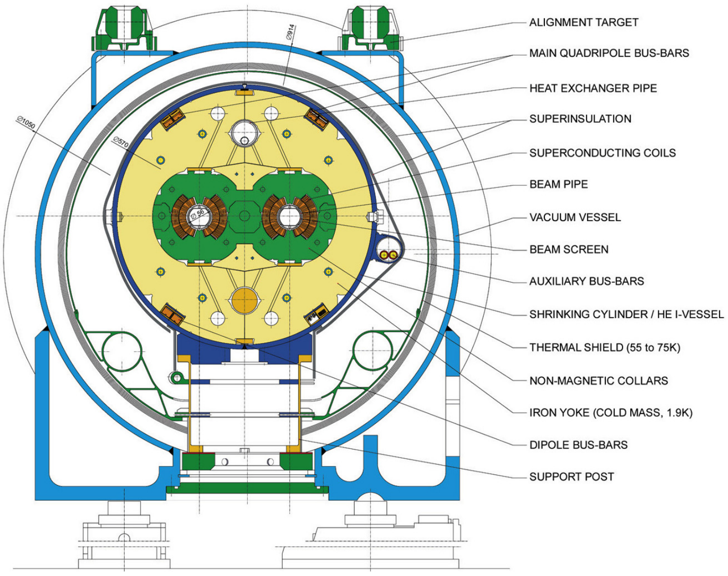


Figure 2.1.4: Cross-section of cryodipole (lengths in mm) [54].

The rate can be related to the average number of inelastic interactions per bunch crossing (μ), the number of bunches per ring (k_b) and the revolution frequency (f_{rev}):

$$\mathcal{L} = \frac{\mu k_b f_{rev}}{\sigma_{inel}} \tag{2.3}$$

To measure the luminosity, special sets of detectors are installed very close to the beam line, to monitor the delivered luminosity by measuring the observed interaction rate per crossing, μ_{vis}. The luminosity can then be written as:

$$\mathcal{L} = \frac{\mu_{vis} k_b f_{rev}}{\sigma_{vis}} \tag{2.4}$$

where $\sigma_{vis} = \varepsilon\sigma_{inel}$ is the total inelastic cross-section multiplied by the efficiency ε of a particular detector and algorithm, and similarly $\mu_{vis} = \varepsilon\mu$ (μ_{vis} denoted as '**pileup**'). Since μ_{vis} is an experimentally observable quantity, the calibration of the luminosity scale for a particular detector and algorithm is equivalent to determining the visible cross-section σ_{vis}.

The calibration of σ_{vis} is performed using dedicated beam-separation scans, also known as van der Meer (νdM) scans, where the absolute luminosity can be inferred from direct measurements of the beam parameters. The delivered luminosity can be written in terms of the accelerator parameters as [50]:

$$\mathcal{L} = \frac{k_b N_b^2 f_{rev}\gamma}{4\pi\varepsilon_n\beta^*}F = \frac{k_b N_b^2 f_{rev}}{4\pi\sigma_x^*\sigma_y^*}F \tag{2.5}$$

where:

- N_b the number of protons per bunch
- γ relativistic factor;
- ε_n the normalised transverse beam emittance - characterizes its spread in coordinate and momentum phase space
- β^* the beta-function at the interaction point, determined by the magnets configuration
- $\sigma_x^*\sigma_y^*$ characterize the widths of the horizontal and vertical beam profiles at the interaction point. It has direct relation with β^* and emittance:

  $$\sigma_x^*\sigma_y^* = \frac{\varepsilon_n\beta^*}{\gamma} \tag{2.6}$$

  The beta function is typically adjusted to have a local minimum at IP in order to minimize the beam size and thus maximize the interaction rate.
- F is the geometric luminosity reduction factor due to the crossing angle at the interaction point

In a νdM scan, the beams are separated by steps of a known distance, which allows a direct measurement of $\sigma_x^*\sigma_y^*$. Combining this scan with an external measurement of the bunch population product (N^2) provides a direct determination of the luminosity when the beams are unseparated.

Parameter	Design	2015	2016
Beam energy (TeV)	7	6.5	
Revolution frequency f_{rev} (kHz)	11.25		
Number of bunches k_b	2808	2232	2208
Number of protons per bunch $N_b(10^{11})$	1.15		
Bunch spacing (ns)	25		
Beta function at IP β^* (m)	0.55	0.8	0.4
Peak Luminosity $\mathcal{L}$ ($10^{33}cm^{-2}s^{-1}$)	10.0	5.0	**13.7**

Table 2.1.1: Parameters of the LHC during Run 2 data taking in 2015 and 2016, compared to the design parameters.

In ATLAS, we have defined a basic time unit called a Luminosity Block (LB). The luminosity is assumed to be stable inside each LB. The typical LB duration is one minute. Data are analysed under the assumption that each luminosity block contains data taken under uniform conditions. The average luminosity and data quality information for each LB are stored in the database. To define a data sample for physics, quality criteria are applied to select LBs where conditions are acceptable, then the average luminosity in that LB is multiplied by the LB duration to provide the integrated luminosity delivered in that LB. Additional corrections can be made for trigger deadtime and trigger prescale factors, which are also recorded on a per-LB basis.

To get the final luminosity estimation for one given dataset, an offline procedure (ATLAS Luminosity Calculator, [121]) retrieves the information about data quality, trigger conditions and duration of LBs from the database. I was in charge of providing the luminosity calculation information for the diphoton analysis. This meant to provide to the collaboration the correct integrated luminosity of the considered data set but also to compute the weights to adjust the pileup distribution in Monte Carlo to the measured one in the data. This will be discussed in Section 2.4.2.

Luminosity in Run 2.

In the Table 2.1.1, the conditions of the machine during the data taking campaign of 2015-2016 of Run 2 at $\sqrt{s} = 13$ TeV are presented together with the design parameters listed for reference. The LHC machine performed excellently in 2016, exceeding the designed luminosity by almost 40%. The collected luminosity over the first two years of Run 2, 2015 and 2016, is presented on Figure 2.1.5. The recorded dataset used for the analysis presented in this thesis will be discussed in details in Section 5.2.2.

2.1.5 The LHC upgrade plan

Following the successful first data-taking period, the LHC will need to shut down to allow maintenance and upgrade work to be carried out, both on the accelerator itself and on the detectors. To sustain and extend its discovery potential, the LHC will need a major upgrade in the 2020s. This long shutdown is part of a long-term agenda for the LHC, which is illustrated in Figure 2.1.6. One objective for the upgrade

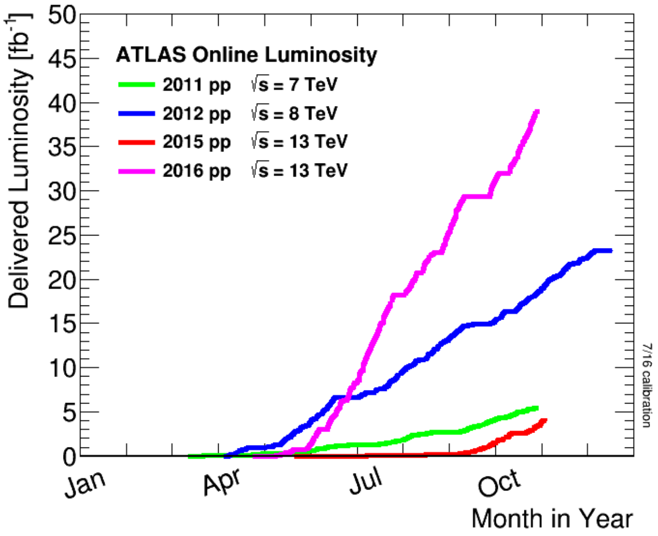


Figure 2.1.5: Overview of collected luminosity in Run 1 and Run 2 [120].

of the accelerator itself is the increase of the instantaneous luminosity provided to the experiments.

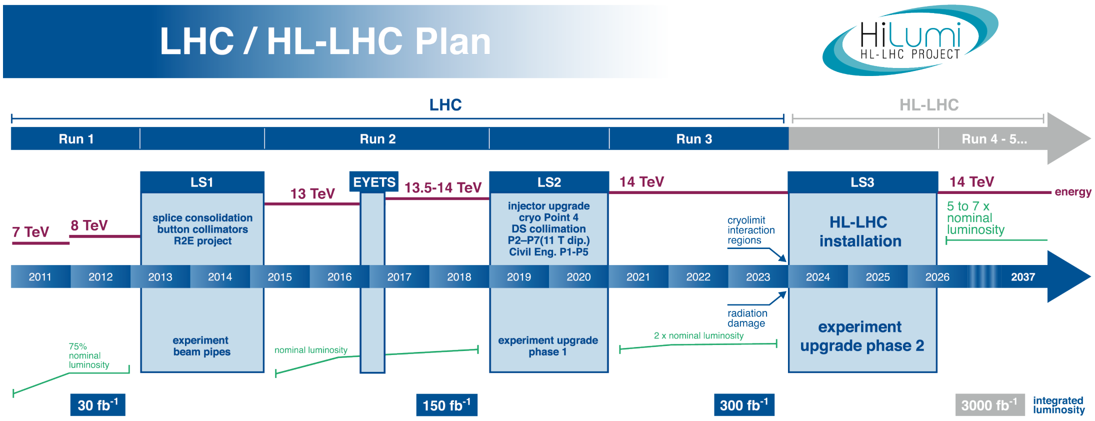


Figure 2.1.6: Timeline of the LHC including the long shutdowns and the phase upgrade [61].

The new configuration, known as High Luminosity LHC (HL-LHC), will rely on a number of key innovations that push the accelerator technology beyond its present limits: 11-12 tesla superconducting magnets; compact superconducting cavities for beam rotation with ultra-precise phase control; new technology and physical processes for beam collimation and 300 metre-long high-power superconducting links with negligible energy dissipation [62]. The main goal of High Luminosity LHC is to reach a peak luminosity of $5(7.5) \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$, which corresponds to an increase by a factor at least five beyond the original design value. This will allow to collect an integrated luminosity of 250 fb^{-1} per year, making possible to reach 3000 (4000) fb^{-1} after the upgrade [61] (estimated for different modes of operation). The amount of analysed data collected by now corresponds to about 1-2% of the full LHC program. The LHC experiments will require upgrade of detector subsystems, triggering and data acquisition techniques, to deal with an average number of interactions per bunch crossing of order 140 (200) and significant increase of the integrated radiation dose.

High-luminosity LHC offers the potential to increase the precision of several key measurements, to uncover rare processes, and to guide and validate the progress in theoretical modeling, thus reducing the systematic uncertainties in the interpretation of the data. The HL-LHC is supported by strong physics program in increase of precision in the Standard Model measurements and study the properties of the Higgs boson: spin, CP properties, and couplings. The increase in luminosity will allow to study rare processes like $H \to \mu\mu$ and to study the di-Higgs processes $HH \to b\bar{b}\gamma\gamma$, providing access to self-coupling (λ_{HHH}) measurements. Another opportunity is the search for new physics, where various scenarios can be performed: constraints on a non-standard mass scaling of the couplings, models with an additional electroweak singlet, an additional electroweak doublet (two-Higgs-doublet model), a composite Higgs boson, a simplified Supersymmetric Standard Model (MSSM), and a Higgs portal to dark matter [63].

2.2 ATLAS Detector

The ATLAS (A Toroidal LHC ApparatuS) detector is a general purpose detector, designed for discovery of new particles and tests of predictions of the Standard Model. It was designed in the 90s to cover the broad physics program of the LHC for more than 20 years and was built between 2000 and 2006. The ATLAS detector has excellent radiation robustness and stability of response with time and with high instantaneous luminosity. A complex of tracking systems within a magnetic field, calorimeters and muon system provide outstanding geometrical acceptance of the detector. It has good hermeticity covering nearly 4π in solid angle. The fine granularity of the sub-detectors provides the excellent resolution required for precision measurements. Custom designed front-end electronics provide exceptional timing which allow the detector to register signal at 40 MHz and to trigger at 100 kHz.

The ATLAS detector is made in a concentric design, which consists of a cylindrical barrel part with the same axis as the beam pipe and two wheel-shape end-caps at each side of the barrel. The inner detector is surrounded by a superconducting solenoid providing magnetic field which bends charged particles trajectories in order to reach precision measurement of the momentum and charge of the particles. It is followed by the calorimeter which measures the energies deposited by the particles and give access to the missing transverse energy. The liquid-argon (LAr) sampling electromagnetic calorimeter provides a measurement of the kinematics of EM objects (γ - photons , $e^{\mp}$ - electrons and positrons). The hadronic calorimeter (HCAL), surrounding the EM calorimeter, provides measurements of the energy deposited by hadrons and jets. The measurement of the shower development (shower shape variables) together with information from inner detector helps to identify the particle type. The Muon spectrometer is used to trigger and to identify and measure muons momentum.

The ATLAS detector uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the center of the detector, and the z-axis along the beam line. The x-axis points from the IP to the center of the LHC ring, and the y-axis points upwards. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the beam line.

In the following Sections I will present the general properties of the sub-detectors: the inner detector (2.2.1), the calorimeter (2.2.2) and the muon system (2.2.3). I shall specify the detection technique and the physics parameters. The detailed description and performance of each system can be found in materials [56, 64, 67, 68]. The description of the ATLAS trigger system is given in Section 2.2.4.

2.2.1 Inner Detectors

The ATLAS Inner Detector (ID), shown on Figure 2.2.1, consists of three subsystems, based on different detection techniques, providing a precise and robust determination of tracks coordinates in $R - \phi$ and z. The pixels, semiconductor tracker (SCT) and transition radiation tracker (TRT) subdetectors are immersed in a 2T magnetic field produced by a superconducting solenoid (5.3 m long, 2.5 m in diameter, situated in the calorimeter cryostat). The trajectory of the charged particles traversing the ID are reconstructed from the signals deposited in typically 4 pixels modules, 8 SCT strips and 30 TRT tubes, leading to a typical transverse momentum resolution of $\sigma_{p_T}/p_T \sim 0.05\% p_T(\text{GeV}) \otimes 1\%$

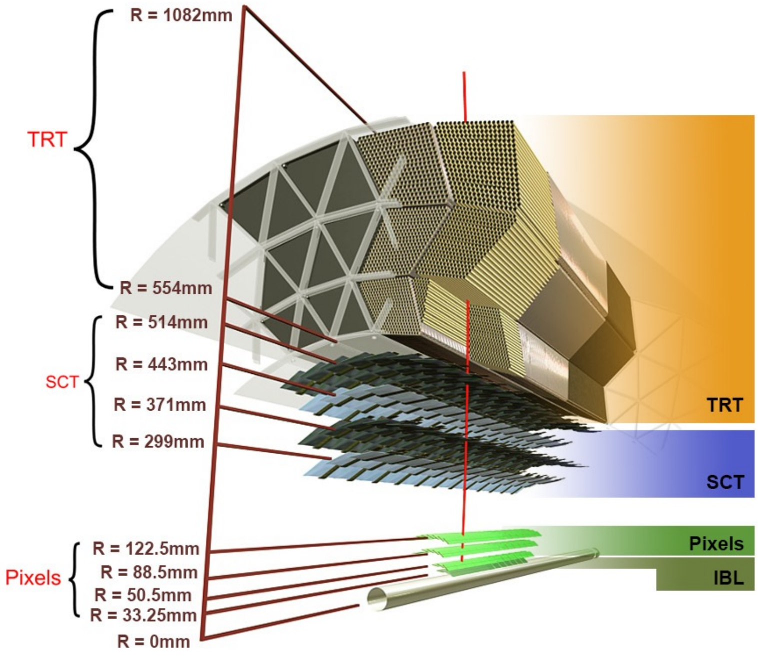


Figure 2.2.1: Different sub detectors of ATLAS Inner Detector [64].

Pixel Detectors

The pixel detector is located very close to the beam pipe (Figure 2.2.1 in green); it allows to reconstruct the vertex position and trajectory of the charged particles. One of the main change in ATLAS between Run 1 and Run 2, is the installation of a new pixel layer (Insertable B-Layer, IBL [65]) at a radius of 33 mm, around a smaller (radius of 25 mm) and thinner beam pipe. It brings improvements on the track and vertex reconstruction and on the b-tagging performance (high efficiency and excellent rejection of jets from light quarks) in the presence of high luminosity pileup. With the IBL, the resolution on transverse and longitudinal impact parameters, $\sigma(d_0)$ and $\sigma(z_0)$, has improved by a factor of two, for low energy tracks [69]. The installation of the fourth pixel layer, at $R = 33\ mm$, has some implication on the reconstruction of converted photons. In particular, one track conversion (where one electron leg is not reconstructed) are identified by requiring the absence of hits in the IBL.

The IBL is composed of 14 lightweight tilted staves 330mm long, arranged in a cylindrical geometry with full coverage in ϕ. Each stave is made of 32/16 (single/double) modules with sensor's dimensions of $50 \times 250\ \mu\text{m}^2$ in the ϕ and z directions, silicon planar sensors in its central region and 3D sensors at the ends. It covers the pseudo rapidity range up to $|\eta| <3.0$.

The IBL is surrounded by three concentric cylindrical silicon-pixel layers at $R = 50.5, 88.5, 122.5$ mm. In the endcaps three silicon-pixel disks perpendicular to the beam pipe are located at $z = \pm495, \pm580, \pm650$ mm. The size of the pixel sensor is 50×400 μm^2 in the ϕ and z directions. The pixels detector covers the $|\eta| <2.5$ region. The intrinsic spatial resolution in the barrel region is 10 μm and 115 μm in $R - \phi$ and z respectively. The reconstruction of the primary vertex plays a crucial role for the diphoton analysis: reconstruction of the mass of the diphoton candidate and the correct identification of the tracks originating from this *pp* collision is necessary to avoid pile-up contributions to the track isolation.

Semiconductor Tracker

The pixel is surrounded by the Semi-Conductor Tracker (SCT) (Figure 2.2.1 in blue). The barrel part of the SCT consists of four concentric cylindrical double-layer silicon microstrips located at $R = 299, 371, 443, 514$ mm and covering $|\eta| <1.1$. One of the layers of each sensor plane is axial (parallel to the beam axis, to measure the $R-\phi$ coordinate) and the other one has a stereo angle of 40 μrad (z coordinate). This stereo angle enables a complete measurement of the particle coordinates. The strips are 126 mm long and 80 μm wide. The endcap, which consists of nine discs perpendicular to the beam pipe located at $z = \pm854$ to ±2720 mm extends the coverage to $|\eta| =2.5$. Each disc is built of radially aligned strips and of stereo strips at an angle of 40 μrad. The length of the strips varies between 55 mm and 120 mm and the width between 57 μm and 90 μm (depending on the disc). The intrinsic spatial resolution in the barrel region is 17 μm and 580 μm in $R - \phi$ and z respectively. Two tracks can be distinguished if they have a ΔR separation of more than 200 μm.

Transition Radiation Tracker

The Transition Radiation Tracker (TRT) is the outermost part of ATLAS tracker (Figure 2.2.1 in orange). It consists of drift tubes, covers the region $|\eta| <2.0$ and have both tracking and identification capabilities. Each polyimide tube (cathode) has a diameter of 4 mm, a length of 144 (39) cm arranged axial (radial) in barrel (endcap); each tube is filled with gas mixture of $70\%Xe$, $27\%CO_2$ and $3\%O_2$. Because of leakage some parts are now filled with liquid argon instead of xenon, due to its cost. The golden wire (anode) of 37 μm collects electrons from ionized gas, when charged particle pass through the straw tube. The intrinsic spatial resolution in the barrel region is 170 μm in $R - \phi$.

To introduce discrimination between charged hadrons and electrons, tubes are surrounded with a radiator fibers (polypropylene/polyethylene). The particles crossing the interface between materials with different indices of refraction, emit a transition radiation (TR). The amount of TR energy depends on the particle's relativistic factor $\gamma = E/m$, therefore depends on its mass. Electrons produce more TR than charged hadrons (pions). Two electronic thresholds have been defined for particle identification. A high threshold (HT) has been defined as a measure of the large energy deposit in the TRT due to absorption of X-rays from transition radiation. Figure 2.2.2 shows the HT fraction for electrons originating from photon conversion and pion candidates in the momentum range 4 GeV $< p <$ 20 GeV, this property used in photon reconstruction and described in Section 3.1.

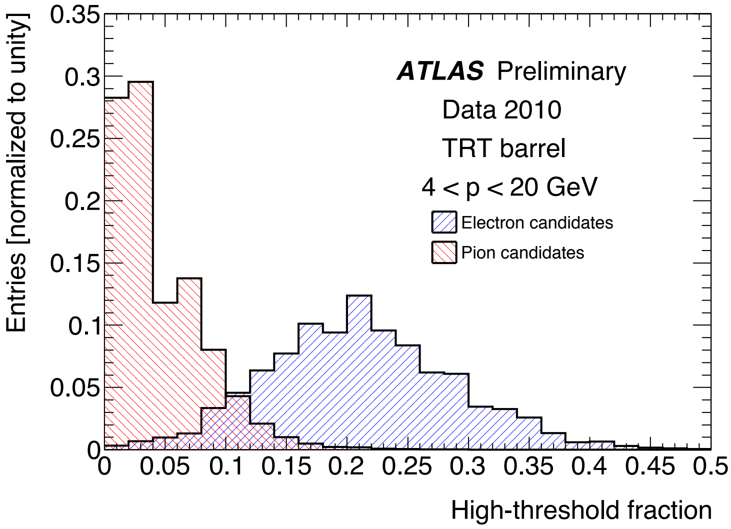


Figure 2.2.2: The HT fraction for electrons from photon conversions and pion candidates in the momentum range 4 GeV $< p <$ 20 GeV, in the barrel region [66].

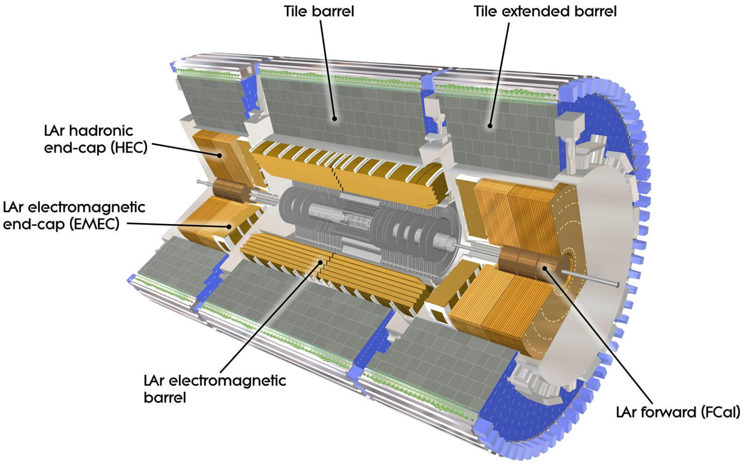


Figure 2.2.3: Overview of the ATLAS calorimetry system [56].

2.2.2 Calorimeters

The ATLAS calorimetry system is composed of an electromagnetic (EM) and an hadronic calorimeters (Figure 2.2.3). The electromagnetic calorimeter with high granularity provides the reconstruction and the identification of EM objects (photons and electrons), while the hadronic calorimeter serves for the reconstruction of hadrons (protons, neutrons, pions, etc.). The two calorimeters which have a "hermetic" coverage down to $|\eta|$ <4.9, ensure the measurement of the missing transverse energy E_T^{miss} [1], defined as the negative vector sum of the p_T of the registered particles (required for studies weakly interacting neutrinos escaping the detector). High energy jets produced in proton-proton collision deposit about half of their energy in the EM calorimeter; the rest of the energy is measured by the hadronic calorimeter. In both calorimeters the sampling technique is used, composed of layers of dense absorbers (lead, iron, copper, tungsten) and active materials (Liquid Argon (LAr), scintillating tiles).

LAr Electromagnetic Calorimeter

The LAr EM calorimeter (Figure 2.2.4) plays a crucial role in the diphoton analysis and calibration studies. It will be described in details in Section 2.3.

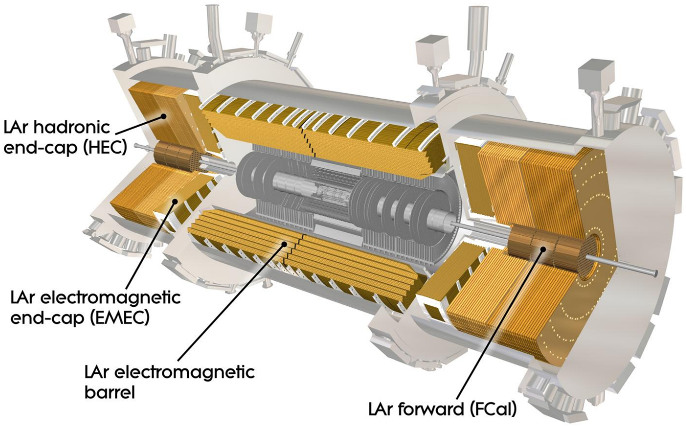


Figure 2.2.4: Overview of ATLAS Liquid Argon (LAr) electromagnetic calorimeter [56].

---

[1] $E_T^{miss} = -\sum_{i=1}^{N} \overrightarrow{p_{T_i}}$, where N indicates the number of final state particles of the event.

Hadronic calorimeter

The hadronic calorimeters complete the measurement of the jets energy. At high energy colliders, quarks and gluons fragment to a beam of particles (mainly hadrons) called jets. The design relative energy resolution of the hadronic calorimeter is:

$$\frac{\sigma_E}{E} \sim \frac{50\%}{\sqrt{E}} \oplus 3\% \tag{2.7}$$

where E is given in GeV, $50\%/\sqrt{E}$ is the sampling term and 3% correspond to the constant term. The energy response is linear at the level of ~1% up to few TeV energies.

The Hadronic sampling calorimeter use steel or copper as absorber. It combines two different techniques used as active material - plastic scintillator and liquid Argon.

Tile calorimeter

The barrel ($|\eta|$ <1.0) and two extended barrels (0.8< $|\eta|$ <1.7) are built of modules of periodic structure steel-scintillator, photomoltiplier tubes (PMT) with wavelength-shifting fiber readout on the tile edge, as shown at Figure 2.2.5 . Fibers collects the UV light emitted by the exited atoms induced by the hadronic shower energy deposits.

Each part of the Tile calorimeter is split into three radial layers with different sampling depths and cell granularity. The first two layers has $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ cell size, and 0.2×0.1 in last layer. The barrel part at η=0 provides 1.5, 4.1 and 1.8 interaction lengths (λ) per layer. The Tile calorimeter is located behind the EM calorimeter; its length is 5.8(2.6 each)m for barrel (extended barrel) and its inner radius is 2.28 m and an outer radius of 4.25 m, as shown on Figure 2.2.3.

LAr Hadronic end-cap calorimeters

The hadronic end-caps(1.5< $|\eta|$ <3.2) use liquid argon as the active material and copper as absorber, combined in flat plates placed in LAr. It consists of two wheels (front and back), sharing one cryostat vessel with the EM end-cap, with an inner radius of 372 to 475 mm and outer radius of 2030mm, as shown on Figure 2.2.6. Each wheel is divided into two layers, with increasing width of absorber from 25 to 50 mm. The ionization current is collected in gaps with the size of the readout cells $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in the region $|\eta| < 2.5$ and 0.2×0.2 for larger values of η.

LAr Forward Calorimeter

The forward detector is a sampling calorimeter that provides both electromagnetic and hadronic energy measurement up to very high η (3.1< $|\eta|$ <4.9), to cover the acceptance as low as possible. It consists of three 45 cm deep modules, Figure 2.2.6: electromagnetic module with copper absorber(FCal1) and two hadronic modules (FCal2 and FCal3) with tungsten to minimize the lateral spread of hadronic showers (high absorption length, about 10λ). The FCal modules are exposed to high particle flux as it is located at a high pseudorapidity. Due to that, the LAr gaps are smaller then in other calorimeters, to avoid the ion build-up problems and to provide high detector density. The signal from the showers is collected by the electrodes placed in rods filled with LAr and oriented parallel to the beam pipe as shown on Figure 2.2.7.

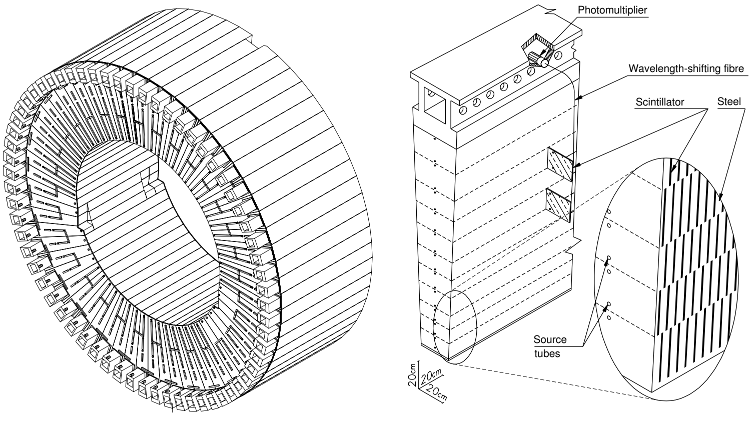


Figure 2.2.5: Schematic view of the Tile Calorimeter. The assembled barrel section is presented (left), together with the elements constituting each module (right) [56].

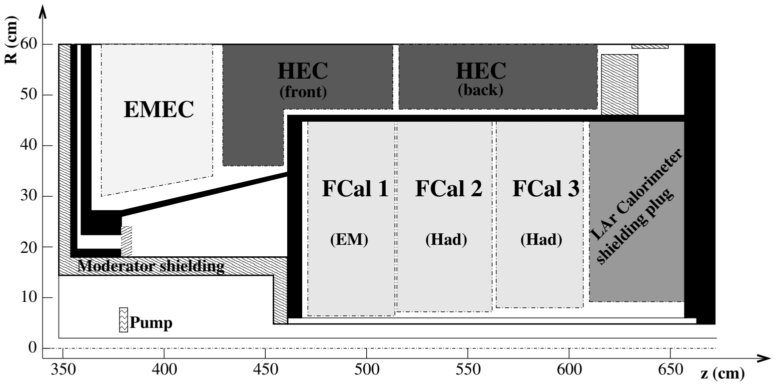


Figure 2.2.6: Schematic diagram showing the three FCal modules located in the end-cap cryostat. The material in front of the FCal and the shielding plug behind it are also shown. The black regions are structural parts of the cryostat. The diagram has a larger vertical scale for clarity [56].

2.2.3 Muon System

The muon spectrometer (MS) is devoted to the identification of muons and to the measurement of their charge and momentum. The measurement is based on the magnetic deflection of tracks in the large superconducting air-core toroid magnets,

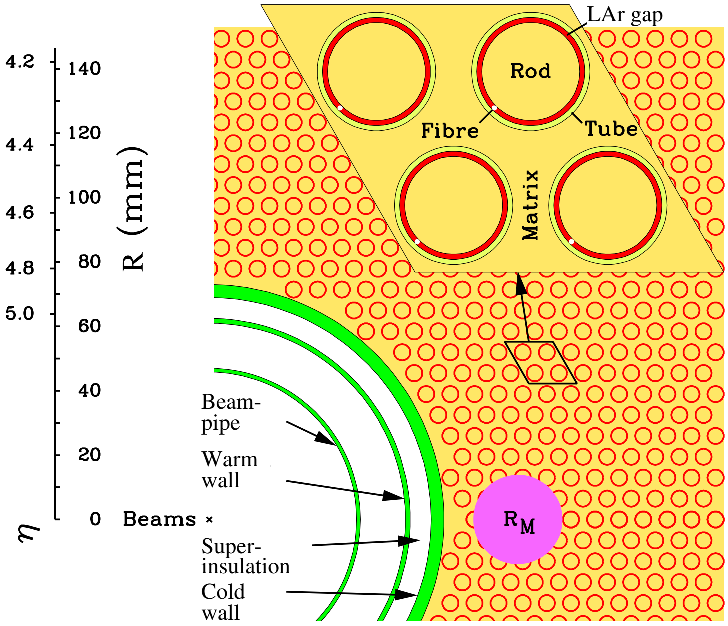


Figure 2.2.7: The electrode structure of FCal1 with the matrix of copper plates, the copper tubes and the rods with the LAr gap for the electrodes. The Moliere radius, R_M, is represented by the solid disk [56].

Figure 2.2.8. The combination of different types of detection techniques enables the ATLAS muon system to trigger and to precisely measure the track coordinates. The MS was designed to provide momentum measurements with a relative resolution of $\sigma(p_T)/p_T$ better than 3% over a wide p_T range and up to 10% at p_T of 1TeV [70].

The **Monitored Drift Tube** chambers (MDT's) provide a measurement of space points with high accuracy. The MDT's cover the pseudorapidity range $|\eta| < 2.7$. The drift tubes are filled with gas under pressure. The chambers are built of several layers, from three to eight (depending on distance to the interaction point). Charged muons ionize the gas when they transverse the tube. The ionisation charge is collected with the anode wire (the cylinder of tube plays the role of cathode) which is at a high potential (See Figure 2.9(a)). The average spatial resolution of a single tube is 80 μm, leading to about 35 μm per chamber.

In the innermost layer of the forward region ($2< |\eta| < 2.7$), MDT's are replaced by **Cathode-Strip Chambers** (CSC). CSC have a higher rate capability and time resolution (safe operation up to counting rates of about 1000 Hz/cm^2). The CSC is a multi-wire proportional chamber with the wires oriented in the radial direction (See Figure 2.9(b)). Strips with fine granularity are used as cathode, one with the strips perpendicular to the wires (providing the precision coordinate) and the other parallel to the wires providing the transverse coordinate. The CSC system is arranged on two discs perpendicular to beam axis and located at $|z|$ of 7.4, 10.8, 14 and 21.5 meters from the interaction point. The achieved resolution is 60μm per CSC plane.

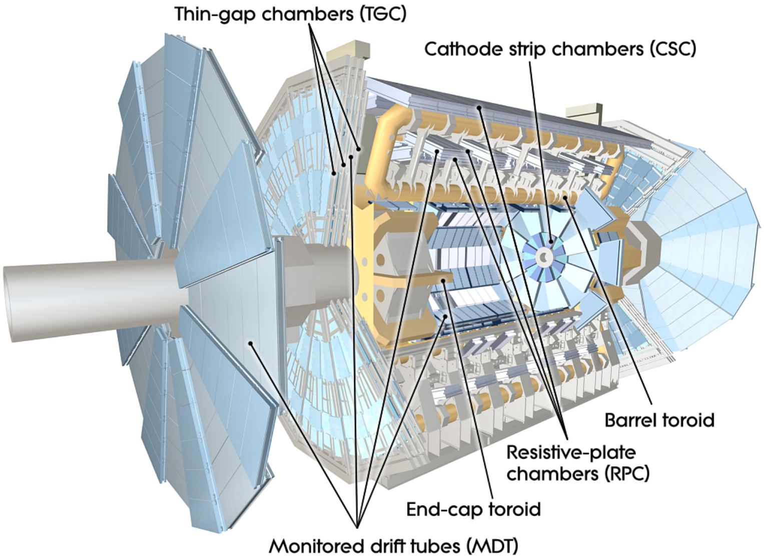


Figure 2.2.8: View of the the ATLAS muon system [56].

The trigger part of Muon system consists of **Resistive Plate Chambers** (RPC's) for the barrel region ($|\eta| \leq$1.05) and of **Thin Gap Chambers** (TGC's) for the end-cap region (1.05$\leq |\eta| \leq$2.4). The RPCs are gaseous parallel electrode-plates, where the signal is collected by strips mounted on the outer faces of the resistive plates. They are arranged in three concentric cylindrical layers around the beam axis at radii of approximately 5, 7.5 and 10 meters, providing three trigger stations. TGC's are multi-wire proportional chambers with an excellent time resolution. The RPCs and TGC provide (η, ϕ) position measurements with a typical spatial resolution of 5-10 mm.

2.2.4 Trigger System

The trigger system is one of the most important part of the data acquisition (DAQ) system of the experiment. It is used to select online potentially interesting physics events to be read out (maximum 1kHz can be handled) and recorded for offline analysis out of 40 MHz pp bunch crossings. The trigger consists of two levels of event selection: hardware-based Level-1 (L1) and a software-based high level trigger (HLT), as shown on Figure 2.2.10.

The L1 trigger is the initial point to identify the interesting collision events to be passed on to the next selection stage, reaching a decision within a latency of less than 2.5μs and a rate of 100 kHz. The signals from the calorimeter (readout with a coarse granularity) and the Muon Trigger System are combined in the Central Trigger Processor which processes the overall L1 Accept (L1A) decision. The Level-1 Trigger identifies event features such as high p_T lepton (electrons and muons), photons, jets and missing transverse energy.

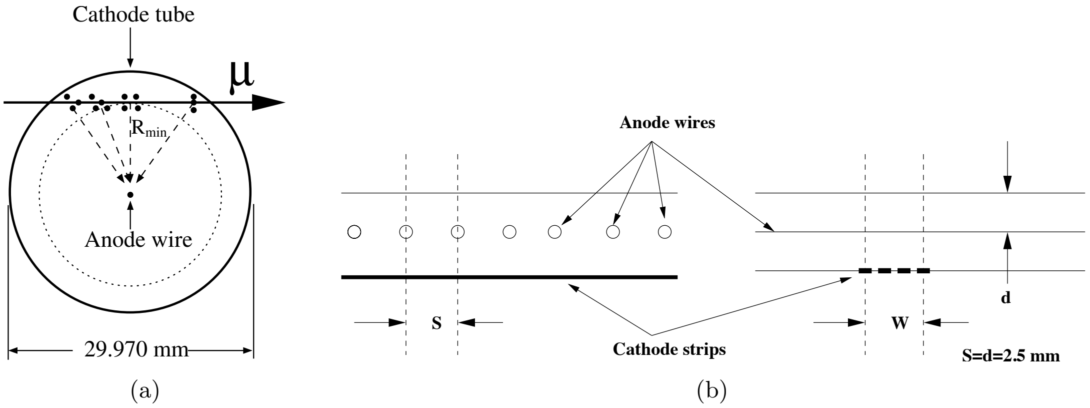


Figure 2.2.9: Precision track position measurement with ATLAS Muon system. Monitored Drift Tube (a) and Cathode-Strip Chambers (b) [56].

After L1 trigger acceptance, the events are buffered in the Read-Out System (ROS) and processed by the HLT using the finer-granularity calorimeter information, precision measurements from the muon spectrometer and the tracking information from the ID. After the events are accepted by the HLT, they are transferred to the local storage space located at the experimental site and then exported to the Tier-0 facility at CERN's computing centre for offline reconstruction.

The algorithms of reconstruction for L1 trigger and details of performance during data taking in Run2 are presented in Reference [71].

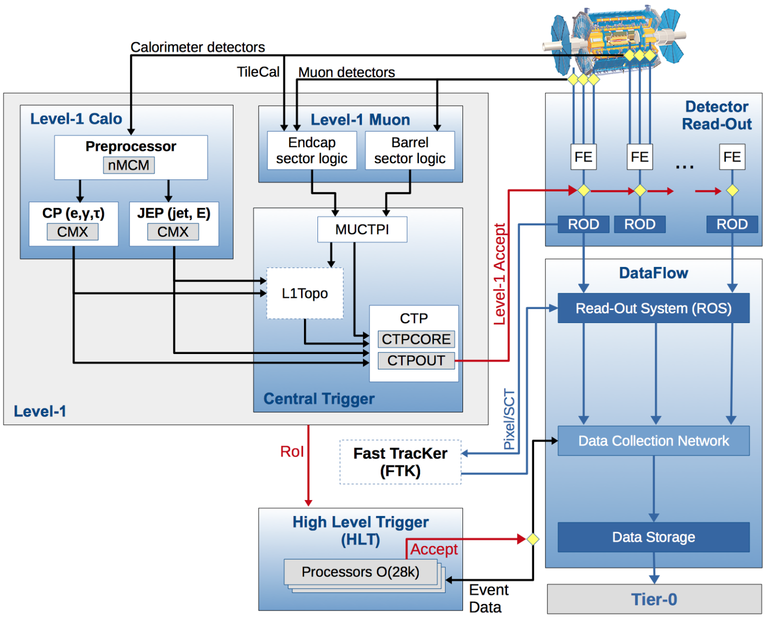


Figure 2.2.10: Schematic layout of the ATLAS trigger and data acquisition system in Run-2 [71].

2.3 The Liquid Argon Calorimeter

The precision electromagnetic calorimeter plays a crucial role in the diphoton analysis. The LAr Calorimeter has an excellent energy resolution, linearity, speed of response and a powerful electron and photon identification capability.

This section is explaining how the ionisation current, created by the particles produced by the interaction of high energy electrons, photons, jets, is proportional to incident particle energy. The detection principle of the calorimeter is described in Section 2.3.1. The detector features are presented in Section 2.3.2, including some details about the cell structure, which allows to reconstruct the shower development. The performance of the calorimeter is discussed in Section 2.3.3. The read-out structure is summarised in Section 2.3.4, the electronics is discussed in details to explain its impact on the energy reconstruction, which is presented in Section 4.1 devoted to the calibration studies. The event format of the data recorded by LAr is discussed in Section 2.3.5. I contributed to the LAr Online Software team activities. I, in particular, updated the tool for decoding the raw data from the detector and shall present this contribution.

2.3.1 Electromagnetic shower development

The principle of calorimeters is based on measuring the energy of the incoming particles, which are losing energy while traveling through matter. The particles cause a shower of many other particles, and the summed ionisation current measured by the calorimeter is proportional to the incident particle energy.

The two dominant processes at high energies ($E > 10$ MeV) by which electromagnetic particles, electrons and photons, interact with matter are the Bremsstrahlung emission (for electrons) and the e^+e^- pair production for photons. The radiation length X_0 is defined as the mean distance over which an electron is left $1/e$ of its incoming energy. It is also defined as 7/9 of the mean free path for high-energy photon e^+e^- pair production. X_0 depends on the material traversed and can be approximated by:

$$X_0 = \frac{716.4\ A}{Z(Z+1)ln\frac{287}{\sqrt{Z}}}\ \mathrm{g\ cm^{-2}} \tag{2.8}$$

where Z and A are respectively the charge and mass of an atom constituting the traversed material. X_0 can also be expressed in cm, dividing by the material density. The choice of an absorber material for electromagnetic particles strongly depends on its atomic charge and structural density. As shown in Figure 2.3.1, electromagnetic showers produced by secondary electrons, positrons and photons, are stopped when the cascade electrons reach energy where the ionization starts to dominate, or when the cascade photons stop to be able to create pairs and mainly interact by Rayleigh or Compton scattering. It is necessary to keep the amount of radiation length in front of the calorimeter as small as possible, to avoid early shower development, which can lead to unaccounted energy loss before the calorimeter.

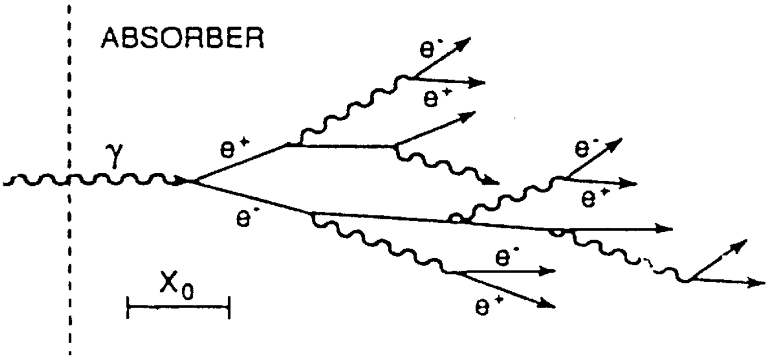


Figure 2.3.1: Schematic development of an electromagnetic shower.

2.3.2 Calorimeter structure

The electromagnetic calorimeter (EM) is divided into a barrel covering $|\eta|$ <1.475 (EMB), and two end-cap sections (EMEC), covering 1.375< $|\eta|$ <3.2. The EM barrel calorimeter consists of two identical half-barrel separated by a small gap (4 mm) at z=0. Each end-cap is divided into two coaxial wheels, an outer wheel covering the region 1.375< $|\eta|$ <2.5 and an inner wheel covering the region 2.5< $|\eta|$ < 3.2. The barrel and end-cap sections are divided into 16 and 8 modules in ϕ, respectively.

The electromagnetic LAr calorimeter is a sampling calorimeter with lead as the absorber and with liquid argon as the detecting medium. The modules are constructed as a continuous sandwich of lead sheets and copper electrodes precisely positioned in the middle of the gap between the two absorbers. The completed structure was then inserted into three cryostats (one barrel and two end-cap cryostats) which have been filled with liquid argon. An accordion geometry has been adopted in order to provide hermiticity and fast signal response. In the barrel, the accordion waves (accordion geometry shown on top of the Figure 2.3.4) are axial and run in ϕ; the folding angles of the waves vary with radius to keep the liquid-argon gap constant. In the end-caps, the waves are parallel to the radial direction and run axially. The granularity in pseudorapidity η is obtained through etching of the copper electrodes prior to insulation and folding. The detector is segmented such that particles traveling from the collision point encounter towers in the $\eta - \phi$ plane. In the η direction, a projective geometry is ensured by a progressive tilt (from the $\eta = 0$ direction) of the etching pattern as shown on Figure 2.3.2.

The accordion-shape allows to have several active layers in depth, three in the precision-measurement region ($0 < |\eta| < 2.5$) and only two in the overlap region between the barrel and the endcap and in the higher$-\eta$ region (2.5< $|\eta|$ < 3.2) with a coarser granularity.

In the region $|\eta| < 2.5$ the EM calorimeter is segmented into three longitudinal sections as presented on Figure 2.3.3. The first layer (L1), also often called strip layer,

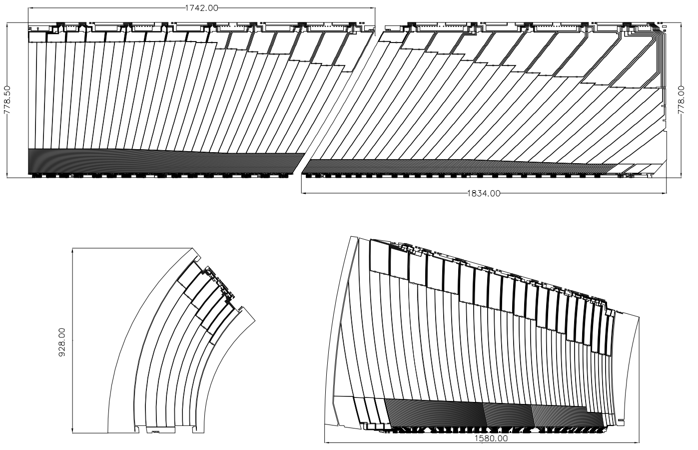


Figure 2.3.2: Layout of the signal layer for the four different types of electrodes before folding. The two top electrodes are for the barrel and the two bottom electrodes are for the end-cap inner (left) and outer (right) wheels. Dimensions are in millimetres. The drawings are all at the same scale. The two or three different layers in depth are clearly visible. [56]

has a depth of $\sim 4.3X_0$ [1]; it has a very fine granularity in η (0.003×0.1 in $\Delta\eta \times \Delta\phi$ for EMB) mainly for the discrimination between single photon showers and two overlapping showers from the decay of neutral hadrons in jets ($\pi^0 - \gamma$ separation Fig. 3.2.1). The second layer (L2), also often called the middle layer has a depth of $\sim 16X_0$; most of the energy of the electrons and photons is deposited in the middle layer; its granularity is 0.025×0.025 in $\Delta\eta \times \Delta\phi$. The third layer (L3), also often called the back layer has a depth which varies from $2X_0$ to about $6X_0$; it measures the energy of the tail of the shower; its granularity is 0.05×0.025 in $\Delta\eta \times \Delta\phi$. The transition region between the barrel and the end-cap, $1.37 < |\eta| < 1.52$, due to the services (cooling system and cables) has a large amount of material in front of the first active calorimeter layer ranging from 5 to almost 10 radiation lengths.

In the region $|\eta| < 1.8$, a detector (presampler, PS) is used to correct for the energy lost upstream of the calorimeter. The presampler consists of an active LAr layer of thickness 1.1 cm (0.5 cm) in the barrel (end-cap) region, with a granularity of 0.025×0.1 in $\Delta\eta \times \Delta\phi$.

The segmentation of the EM calorimeter in $\eta \times \phi$, for each pseudorapidity region, is summarised in Table 2.3.1

The liquid argon fills the gaps between the sampling layers of absorbers made of lead and stainless steel. The readout electrode is located in the middle of a gap, it consists of three planar layers of copper which are separated by insulator layers of kapton. Electrodes and absorbers are separated using honeycomb-shaped spacers as shown on Figure 2.3.4. Two voltage lines per sector ($\Delta\eta \times \Delta\phi = 0.2 \times 0.2$, in the barrel) fed the

---

[1] X_0 - radiation length

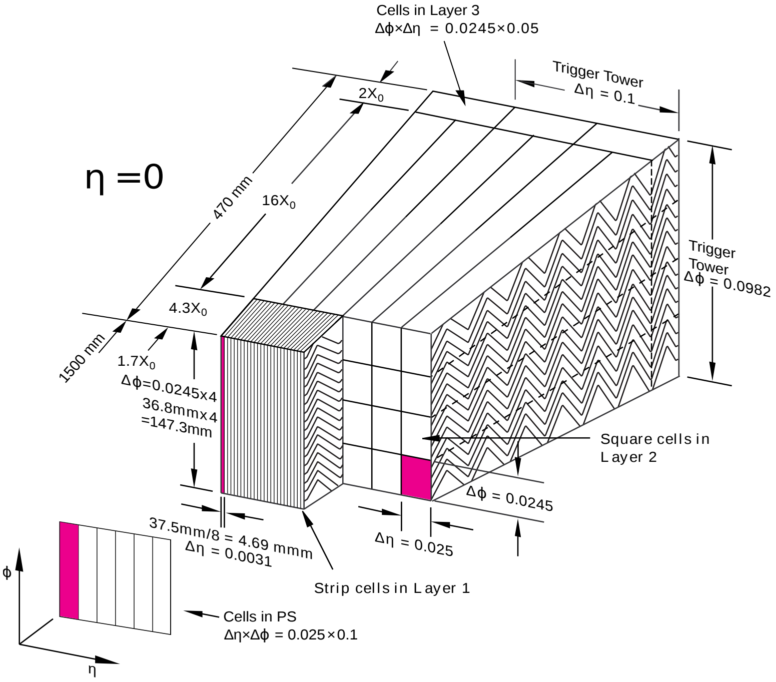


Figure 2.3.3: Sketch of a barrel module with lateral segmentation. The granularity in η and ϕ of the cells of each of the three layers and of the trigger towers is also shown [56] .

copper electrode with a high voltage of 2000V (in the barrel), providing an electric field of 1kV/mm inside the LAr gap. The inner electrode collects the signal from ionization through capacitive effect. A sketch of the triangular current pulse from the calorimeter cell is presented on Figure 2.3.6. The amplitude of that pulse is proportional to the deposited energy by the electromagnetic shower. The energy of the initial particle, can be measured from the reconstructed deposited shower energy. The signals from different longitudinal compartments of the calorimeter are read out at both radius of the electrodes, i.e. Layer 1 is readout from the inner radius and Layer 2 and 3 from the outer radius. The signals are received by summing boards and then sent to the Front-end boards (FEBs).

2.3.3 Energy resolution of the electromagnetic calorimeter

The energy resolution of the calorimeter is described by the equation:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \tag{2.9}$$

where a, b and c are η-dependent parameters. The parameter a is the coefficient of the sampling (”stochastic”) term, representing the statistical fluctuations of the

	Barrel		End-cap	
Presampler	0.025×0.1	$\lvert\eta\rvert < 1.52$	0.025×0.1	$1.5 < \lvert\eta\rvert < 1.8$
First layer	$0.025/8 \times 0.1$	$\lvert\eta\rvert < 1.40$	0.050×0.1	$1.375 < \lvert\eta\rvert < 1.425$
	0.025×0.025	$1.40 < \lvert\eta\rvert < 1.475$	0.025×0.1	$1.425 < \lvert\eta\rvert < 1.5$
			$0.025/8 \times 0.1$	$1.5 < \lvert\eta\rvert < 1.8$
			$0.025/6 \times 0.1$	$1.8 < \lvert\eta\rvert < 2.0$
			$0.025/4 \times 0.1$	$2.0 < \lvert\eta\rvert < 2.4$
			0.025×0.1	$2.4 < \lvert\eta\rvert < 2.5$
			0.1×0.1	$2.5 < \lvert\eta\rvert < 3.2$
Second layer	0.025×0.025	$\lvert\eta\rvert < 1.40$	0.050×0.025	$1.375 < \lvert\eta\rvert < 1.425$
	0.075×0.025	$1.40 < \lvert\eta\rvert < 1.475$	0.025×0.025	$1.425 < \lvert\eta\rvert < 2.5$
			0.1×0.1	$2.5 < \lvert\eta\rvert < 3.2$
Third layer	0.050×0.025	$\lvert\eta\rvert < 1.35$	0.050×0.025	$1.5 < \lvert\eta\rvert < 2.5$

Table 2.3.1: Granularity of the EM calorimeter in $\Delta\eta \times \Delta\phi$ versus pseudorapidity [56].

electromagnetic shower development in the sampling structure of the LAr calorimeter (it is dominant at intermediate energies). The parameter b is the "noise" term due to the electronic and pile-up noise, and c is the "constant" term which takes into account the non-uniformity of the calorimeter and of its response.

The parameters were measured during the test beam campaigns [74]. The energy resolution is obtained from the standard deviation of the Gaussian fit, the noise is subtracted for each energy point to obtain the intrinsic resolution of the calorimeter. The noise term $b \sim 170$ MeV without pileup and is about $350\times \cosh\eta$ MeV for an average number of interactions per bunch crossing of μ=20 [76]. The intrinsic resolution of the calorimeter is measured to be:

$$\left.\frac{\sigma_E}{E}\right|_{measured} = \frac{10.1\%}{\sqrt{E}} \oplus 0.17\% \tag{2.10}$$

The design value of the "sampling" term ($9 - 10\%/\sqrt{E}$ [68]) is fulfilled and measured to be 10.1±0.1%. At high energies the resolution is dominated by the constant term, which was measured to be 0.17±0.04% locally (the designed value being 0.7% for the entire EMB). An additional effective constant term c', that depends on the pseudorapidity η, to account for the difference in energy resolution between data and Monte Carlo is measured on data (Section 3.4.4).

2.3.4 Calorimeter read-out structure

The read-out electronics system of the LAr calorimeter is presented on Figure 2.3.5; it consists of two parts: the *Front End* (FE) system which consists of the boards mounted directly on the cryostat feedthroughs; and the *Back End* (BE) system located

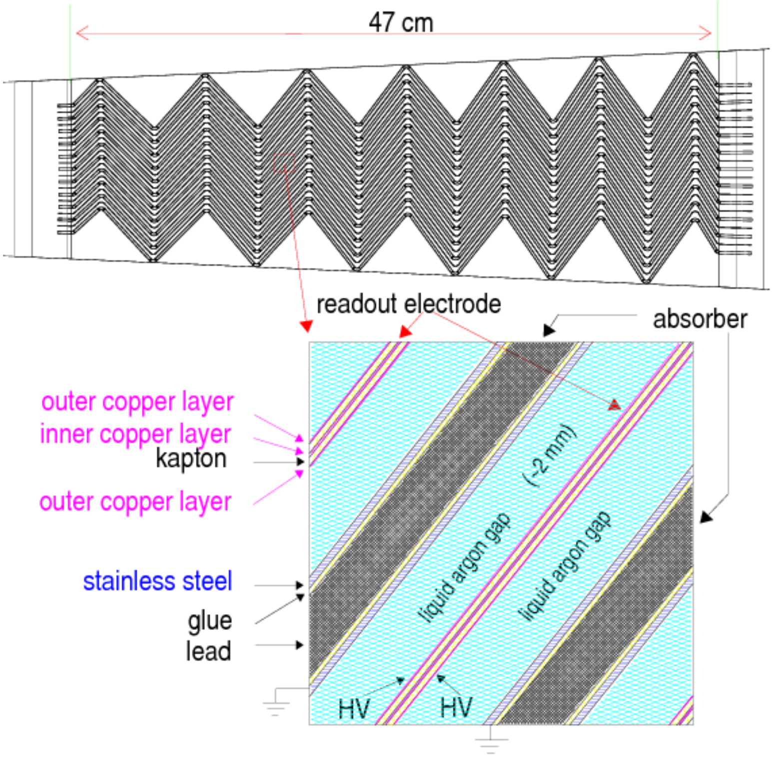


Figure 2.3.4: Accordion structure of the barrel. The top figure is a view of a small sector of the barrel calorimeter in a plane transverse to the LHC beams. [73]

outside the detector cavern, in the experiment counting room (USA15). The FE system is responsible for amplification, shaping and sampling and digitisation of the signals. The *Front-end boards* (FEBs) are built of preamplifiers, shapers, analog pipelines and ADC. In addition, to inject precision calibration signals on the electrodes, *Calibration boards* are also installed on detector. The FE system also consists of *summing boards* to make the analogue sums from adjacent calorimeter cells for the Level 1 trigger and *controller boards* which receive and distribute the 40 MHz LHC clock as well as other configuration and control signals. The BE electronics includes *Readout Driver* (ROD) boards which receive the digitized signals from the FEBs. The RODs perform digital filtering of the signals, formatting and monitoring of the calorimeter signals using *Digital Signal Processor* (DSP), before transmitting the processed data to the ATLAS *Data AcQuisition* system (DAQ) [72].

Front-End Board overview

The raw analog signal from the calorimeter cell is transmitted to the FEBs, where it passes several stages of analog processing. There is a total of 1524 FEBs distributed among 58 FE crates where the signal is amplified, shaped and digitised. Each FEB

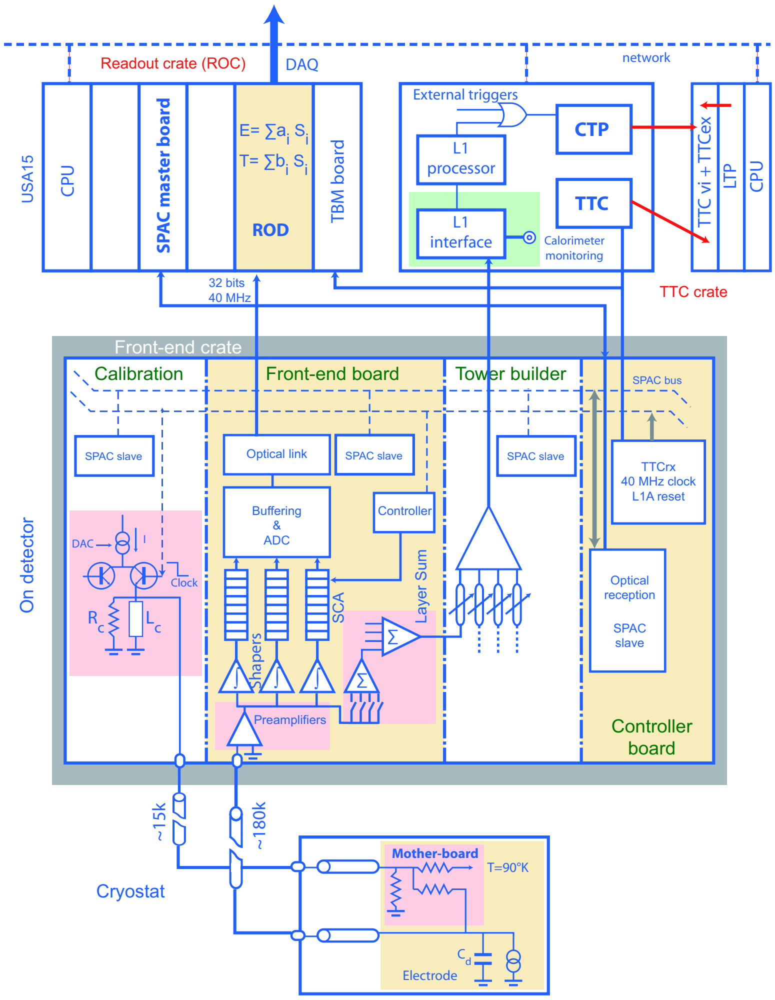


Figure 2.3.5: Block diagram of the LAr readout electronics. The cold electrical circuit is depicted at the bottom, followed above by the on-detector front-end electronics crate and at the top (left) by a schematic view of the readout crate with its ROD boards and TTC modules. Also indicated at the middle and top (right) are the LAr front-end tower builder electronics and the interfaces to the L1 trigger system with its central trigger processor (CTP) [56].

processes up to 128 calorimeter channels. At the exit of the preamplifiers, the signal is split into three overlapping linear gain scales with a ratio of 1/9.3/93 (low, medium and high gain). Data from each gain are stored into pipelines is needed to digitize a 16 bits range of energy (from 40 MeV to 3 TeV) using 12 bits analog-to-digital converters (ADC). The typical energy depositions in the second layer of EM calorimeter readout with different gains is presented:

- low gain: ~300 GeV - 4 TeV
- medium gain: ~30 - 300 GeV

- high gain: up to ∼30 GeV

Then signal is shaped in the form of a bipolar pulse. The pulse of a LAr cell is shown on Figure 2.3.6 before and after the bipolar shaping.

The shaped signals are then sampled at the LHC bunch-crossing frequency of 40 MHz by four-channel switched-capacitor arrays (SCA) analogue pipelines that store the signals during the L1 trigger latency (∼2.5 μs). During LHC data taking, the FEB read out four samples per channel around the signal peak - a compromise between the noise reduction achieved and the amount of data that can be digitized and processed in real time. For events accepted by the L1 trigger, the four stored samples of the chosen gain are digitised by a 12-bit ADC as shown in Figure 2.3.7. To optimize the precision of the energy measurement, the Gain Selector chips (GSEL) chooses for each channel, in each event, which of the three gains to use, based on the value of the peak sample in the medium gain compared to two reference thresholds: to keep the lowest non-saturated gain, in order to use the minimum ADC→MeV conversion factor. The digitized data are formatted, multiplexed, and then transmitted optically out of the detector to the ROD via a single 1.6 Gbps optical output link per FEB. The FEBs also have the task to sum the signal of adjacent calorimeter cells inside each depth layer and to prepare the input for the tower builder boards, which is the interface board between the calorimeter and the L1 trigger system.

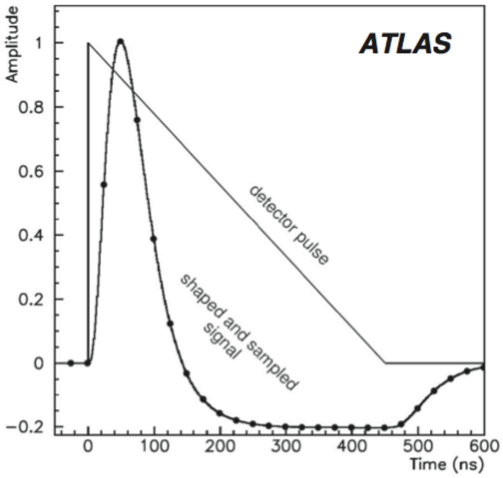


Figure 2.3.6: Ionisation pulse in a LAr cell and FEB output signal after bi-polar shaping. Also indicated are the sampling points every 25 ns. During normal data-taking the signal is sampled four times. A maximum of 32 samples can be attained and is used for calibration purpose [56].

Calibration boards

The calibration boards are used to calibrate the energy response of the LAr calorimeter by sending precise pulses to the electrodes. Non-uniformity of the calibration pulse

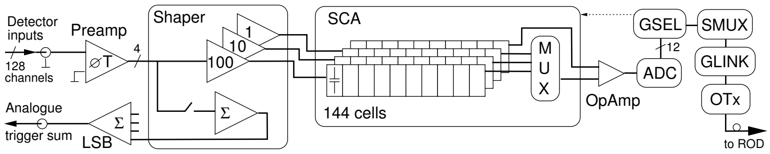


Figure 2.3.7: Block diagram of the FEB architecture, depicting the dataflow for one of the 128 channels [56].

amplitude affects directly the constant term in the energy resolution function and therefore the measurement of high energy electromagnetic showers. The calibration system is designed to limit this contribution to the constant term to less than 0.25% for the EM calorimeter, less than 1% for the HEC and less than 2% for the FCal [56].

Three different types of calibration runs are taken to equalise the response of the LAr calorimeter cells: *pedestals*, *delays* and *ramps*. For the *delays* calibration runs the FEB read out up to 32 samples per channel, but at a lower trigger rate. Sets of calibration runs are acquired for each of the three LAr electronic readout gains (high, medium and low). During the *pedestal* runs no signal is injected into the calorimeter cells thus the electronics baseline level and the electronic noise for each cell are measured. *Delay* runs measure the shape of the pulse as a function of time for each readout cell. *Ramp* runs are used to extract the response of each cell as a function of the injected current in DAC units. The values of the current used varies according to the gain of the electronics being probed [77].

Back-End Board overview

The digital signals from the FEBs are then transmitted via optical links to the back-end electronics which are located in the main cavern, 70 m away from the detector. The Readout Driver system digitally processes the data before transmitting it to the data acquisition system at a L1 trigger rate of 100 kHz. Each ROD module receives data from up to eight FEBs (ie. up to 1024 channels) through eight optical fibres. The ROD motherboard houses four processing units (PU) and each PU is equipped with two Digital Signal Processors (DSP). The Digital Signal Processors synchronise the FEB data with the *Trigger-Time and Control* (TTC) signals, process the data and apply an optimal-filtering method to compute the energy, time and a quality factor of the signal for each cell, prepare and send the data to the ROD motherboard.

The DSP reconstructs the cell energy by applying the calibration constants to the amplitude (A_{ADC}) of the signal pulse reconstructed from the four samples values (in ADC counts) received from the FEB, as well as the time offset of the deposition (τ), by applying a digital filter to the recorded signal samples in ADC counts (s_i) according to the following equations:

$$A_{ADC} = \sum_{i=0}^{N_{samples}-1} a_i(s_i - p) \tag{2.11}$$

and

$$\tau = \frac{1}{A_{ADC}} \sum_{i=0}^{N_{samples}-1} b_i(s_i - p), \tag{2.12}$$

where p is the pedestal, a_i and b_i are the energy and timing *Optimal Filtering Coefficients* (OFCs), $N_{samples}$ is the number of samples used for the computation (while five samples were readout during Run-1, four are now read out, since the beginning of Run-2, due to the increase in L1 trigger rate) [77]. The Optimal Filtering method is a digital filtering technique which allows to compute the peak amplitude of the signal minimising the effect of the electronic and pile-up noise [75]. It requires the knowledge of the shape of the ionisation pulse and the noise auto-correlation matrix for the samples. The shape is predicted from the information contained in the calibration pulses obtained in delay runs.

The energy computation on each DSP is performed with formulae:

$$E = \underbrace{F_{DAC\to\mu A} \times F_{\mu A\to MeV} \times C_{HV} \times \frac{1}{\frac{M_{phys}}{M_{cal}}}}_{f} \times \sum_{j=0}^{N_{ramps}} G_j A^j \tag{2.13}$$

where f correspond to the product of all the conversion and correction factors, listed below:

- $F_{DAC\to\mu A}$ is a factor that converts the current measured in DAC units to μA, and accounts for the values of the local motherboard injection resistor.
- The factor $F_{\mu A\to MeV}$ converts the current to a raw estimate of the energy released in the active and passive part of the calorimeter cell using an average value of the sampling fraction.
- C_{HV} is a correction factor applied when the high voltage is not nominal.
- The factor $\frac{M_{phys}}{M_{cal}}$ corrects for the difference between the calibration and the ionisation pulse shape and is computed from the calibration pulse and from the properties of the readout cells.
- The factors G_j are coefficients of a polynomial of degree N_{ramps} in powers of A that parameterize the cell electronic gain. It is obtained from the ramps calibration run as the slope and offset of the linear fit of the DAC setting versus the reconstructed pulse amplitude in ADC counts.

A quality factor Q is computed to quantify the quality of the measurement:

$$Q = \sum_{i=0}^{N_{samples}-1} [b_i(s_i - p) - A_{ADC}(g_i - g_i^{'}\tau)]^2, \tag{2.14}$$

where g_i is the normalised predicted ionisation pulse shape (after shaping), calculated from calibration pulse shapes and $g_i^{'}$ its derivative.

The computed energy, and time and pulse shape quality factor for cells with energy above a given threshold, are transmitted to the ATLAS DAQ readout part.

2.3.5 Event format and bytestream

During the restart of LHC for Run-2 after the Long Shutdown 1 (LS1) I was a member of the LAr Online Software team. One of the task was to make accessible the information from the detector in convenient way, to be able to control the recorded data within the update for Run-2 ATLAS DAQ software. Therefore, in this section I will introduce the structure of the data format recorded in the ATLAS detector and describe the tool I maintained to access the data stream.

Data from subdetector systems (here I shall only discuss the LAr calorimeter) are stored using the *Event Format* (EF) structure. The EF defines the structure of the data at various readout stages within the *high level trigger* (HLT) and DAQ. The structure is *self descriptive* imposing to include the length of each data block. This allows elements of the Data Flow and HLT processing tasks to access the data without requiring the use of external resources (like Data Bases). EF defines additional data that is added to the detector data, by elements of the TDAQ, allowing the processing tasks to quickly identify the type and origin of each event.

In general, an event is composed as following [78]: Full Event is built from fragments *ReadOut Buffer* (ROB). Each of the ROB is mapped onto a single *ReadOut Driver* (ROD) fragment. Each fragment type, except the ROD fragment, has a header which contains all the event formatting information needed to decode itself. Besides the generic header, ROB and Full Event fragments may contain a single 32 bit word trailer with a checksum of its contents.

Each header has a fixed structure:

1. Start of header marker.
2. Total fragment size.
3. Header size.
4. Format version number.
5. Source identifier: It consists of a sub detector ID, and Module ID. The combination of these fields should allow the Source identifier to be unique across the whole of ATLAS. The Module ID refers to the module which builds and adds the header to the event fragment.
6. Number of status elements.
7. Status element.
8. Check Sum Type.

A ROD fragment contains the data from a given part of the detector - the data is formatted in hardware (not necessarily by programmable devices). Within the ROD fragment header five additional elements are explicitly defined, these are:

(a) Run Number (value is unique during the lifetime of the experiment).

(b) Extended Level 1 ID (generated in the TTC).

(c) Bunch Crossing ID (generated in the TTC).

(d) Level 1 Trigger Type, generated by the Central Trigger Processor and transmitted by the TTC system.

(e) Detector event type: This element allows additional information to be supplied on the type of event, particularly in the case of calibration events. It allows the detectors to specify the exact type of calibration event that they have generated.

I updated the code `DumpeFormat` to run with the version 5 of the `EventFormat`, which is in used since the begining of Run-2. This is a tool, which can access the raw data stream and get information from a particular sub detection system. The key point of this program is that it does not require to wait for reprocessing as it reads directly bytes information. Inside the bytestream, the code searches for a particular header structure, which is shown in Table 2.3.2.

Fragment Type	**Header Marker**
ROD	0xee1234ee
ROB	0xee1234dd
Full Event	0xee1234aa

Table 2.3.2: Start of Header Markers [78].

Then it decodes the source identifiers to recognise Sub-detector ID and Module ID. The collection of LAr identifiers is given in Table 2.3.3.

Detector	EMB A	EMB C	EMEC A	EMEC C	HEC A	HEC C	FCAL A	FCAL C
ID	0x41	0x42	0x43	0x44	0x45	0x46	0x47	0x48

Table 2.3.3: LAr subdetector identifiers [78].

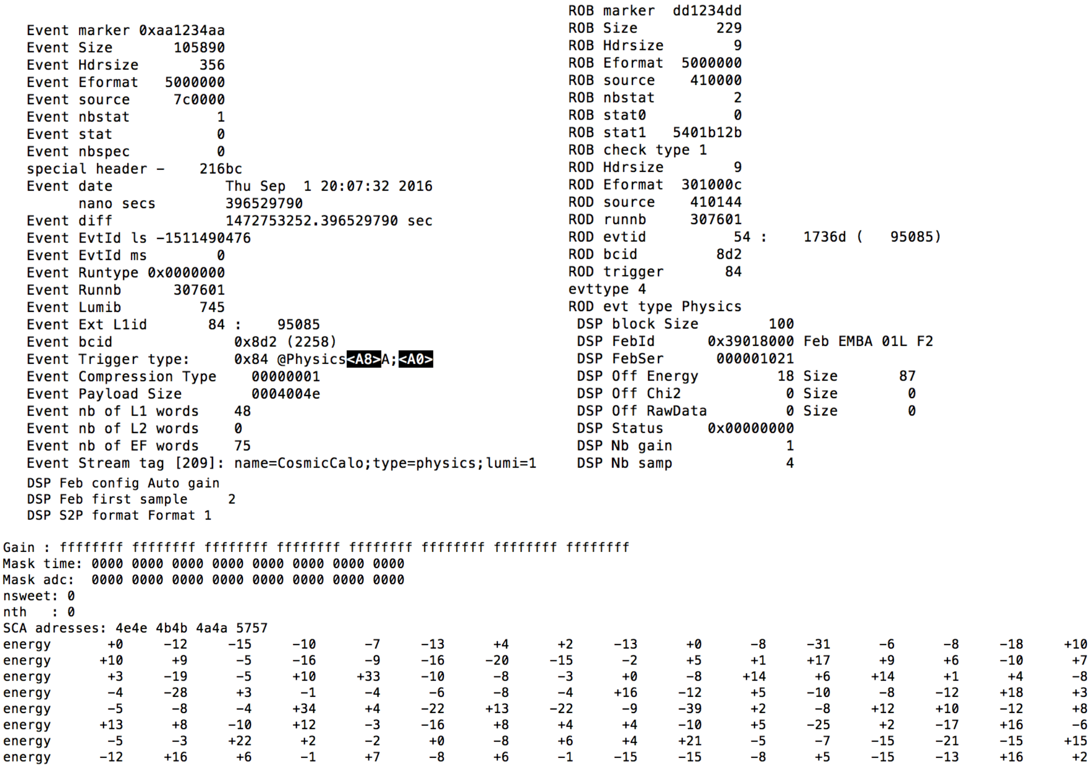
To accomodate multiple readout modes, the LAr readout, and consequently the LAr bytestream, can be configured in three main modes:

- **Transparent**: No energy is calculated in the DSP. The full digitized analog waveform of the FEB is stored, can be up to 32 samples for each channel.
- **Calibration**: Accumulated waveforms average is sent after a give number of events. Intermediate events are empty.
- **Physics**: The energy is calculated for each cell in the DSP and stored in the bytestream. For cells above a given energy threshold, keep also time/quality.

> Another threshold defines whether or not to send samples. All the necessary conditions to perform the optimal filtering processing are loaded into the DSP.

The *Transparent* mode is used in very special conditions, like during the original ATLAS commissioning or when LHC splash events are delivered.

The code successfully was implemented in the LAr online system and was used in operational tasks. The documentation of the tool can be found in Reference [79]. Snapshot of the decoded bytestream using the `DumpeFormat` tool is shown on Figure 2.3.8. The description starts with the Event Information: the physics data collected on 1st of September 2016 in the CosmicCalo stream. The ROB identifiers show that the Event Format 5 was used. The ROD presents the DSP information: data from the electromagnetic barrel, side A, the signal is digitised using four samples. In the bottom lines, the cell energies for each channel are listed.



```
Event marker 0xaa1234aa
Event Size       105890
Event Hdrsize       356
Event Eformat   5000000
Event source     7c0000
Event nbstat          1
Event stat            0
Event nbspec          0
special header -    216bc
Event date             Thu Sep  1 20:07:32 2016
      nano secs        396529790
Event diff             1472753252.396529790 sec
Event EvtId ls -1511490476
Event EvtId ms        0
Event Runtype 0x0000000
Event Runnb      307601
Event Lumib         745
Event Ext L1id       84 :    95085
Event bcid              0x8d2 (2258)
Event Trigger type:     0x84 @Physics<A8>A;<A0>
Event Compression Type    00000001
Event Payload Size        0004004e
Event nb of L1 words    48
Event nb of L2 words    0
Event nb of EF words    75
Event Stream tag [209]: name=CosmicCalo;type=physics;lumi=1
DSP Feb config Auto gain
DSP Feb first sample     2
DSP S2P format Format 1
```

```
ROB marker  dd1234dd
ROB Size         229
ROB Hdrsize        9
ROB Eformat  5000000
ROB source    410000
ROB nbstat         2
ROB stat0          0
ROB stat1   5401b12b
ROB check type 1
ROD Hdrsize        9
ROD Eformat  301000c
ROD source    410144
ROD runnb     307601
ROD evtid          54 :    1736d (   95085)
ROD bcid         8d2
ROD trigger       84
evttype 4
ROD evt type Physics
 DSP block Size        100
 DSP FebId      0x39018000 Feb EMBA 01L F2
 DSP FebSer      000001021
 DSP Off Energy         18 Size       87
 DSP Off Chi2            0 Size        0
 DSP Off RawData         0 Size        0
 DSP Status     0x00000000
 DSP Nb gain             1
 DSP Nb samp             4
```

```
Gain : ffffffff ffffffff ffffffff ffffffff ffffffff ffffffff ffffffff ffffffff
Mask time: 0000 0000 0000 0000 0000 0000 0000 0000
Mask adc:  0000 0000 0000 0000 0000 0000 0000 0000
nsweet: 0
nth   : 0
SCA adresses: 4e4e 4b4b 4a4a 5757
energy       +0    -12    -15    -10     -7    -13     +4     +2    -13     +0     -8    -31     -6     -8    -18    +10
energy      +10     +9     -5    -16     -9    -16    -20    -15     -2     +5     +1    +17     +9     +6    -10     +7
energy       +3    -19     -5    +10    +33    -10     -8     -3     +0     -8    +14     +6    +14     +1     +4     -8
energy       -4    -28     +3     -1     -4     -6     -8     -4    +16    -12     +5    -10     -8    -12    +18     +3
energy       -5     -8     -4    +34     +4    -22    +13    -22     -9    -39     +2     -8    +12    +10    -12     +8
energy      +13     +8    -10    +12     -3    -16     +8     +4     +4    -10     +5    -25     +2    -17    +16     -6
energy       -5     -3    +22     +2     -2     +0     -8     +6     +4    +21     -5     -7    -15    -21    -15    +15
energy      -12    +16     +6     -1     +7     -8     +6     -1    -15    -15     -8     +5    -15    -13    +16     +2
```

Figure 2.3.8: Snapshot of the decoded bytestream using the `DumpeFormat` tool.

2.4 Monte Carlo Simulation

ATLAS has developed a very detailed description of the detector geometry, implemented a precise simulation of the detector response and interfaced both to physics processes generators. Simulated events are reconstructed with the same software as recorded data. Having a precise simulation of proton-proton collisions at $\sqrt{s}$ =13 TeV is a key ingredient to the physics analysis as it permits to confront any observation to expectations.

The simulation starts with the generation of events and immediate decays, using various sets of generators (Pythia, Sherpa, MC@NLO, etc.). The interactions of the decay products with the detector is described using GEANT4 simulation toolkit [80, 81]. And the last step is the digitisation of the energy deposited in the detector into voltages and currents for comparison to data readout of the ATLAS detector. Both the simulated events and data from the detector then run through the same ATLAS trigger and reconstruction. This allows to make a detailed comparison of the two.

In this section I will briefly introduce the simulation steps in Section 2.4.1. The role of Monte Carlo simulation is to describe, as accurately as possible, the experimental characteristics of physics processes. Therefore in ATLAS we study data and Monte Carlo response, and introduce corrections and reweighting to remove residual discrepancy. As an example of the excellent performance of the simulation after corrections, I will show my studies of the impact of the reweighting on the distributions of the average interaction per bunch crossings in Section 2.4.2.

2.4.1 Simulation

The generators produce complete events starting from a proton-proton initial state. Each generated event contains the particles from a single interaction of the two partons inside the colliding hadrons; this process is denoted as **hard scattering**. **Parton distribution functions** (PDFs) are used to describe the substructure of the proton and are used by all the event generators as external inputs. The matrix element of the hard scattering process is calculated at *leading order* (LO) in QCD; some generators provide next-to-leading (NLO) order computations. The generator adds additional QCD and QED radiations from the parton, which produce **parton shower**. The shower approximation is most accurate when the radiation is emitted at small angle and is poorest in those cases with a large number of widely separated emissions of comparable energy. The generator simulates the **underlying event** (UE) using the multiple-parton interaction model; UE describes the soft interaction from partons not involved in the hard scattering. At the end of the shower, a phenomenological model is used to combine the quarks and gluons into hadrons, denoted as **hadronization**. Many of these hadrons are unstable, and subsequently decay. The combination of the processes listed above represent the **generated event**.

After the event is generated the **simulation of the detector response** to the particles from the event is performed. The detector geometry, the material budget and magnetic field is described using GEANT4 simulation toolkit [80, 81]. It performs the accurate modelling of the interactions of particles with the detector, providing Monte Carlo *truth* record. The ATLAS digitization software converts the hits produced by the simulation into detector responses - the voltage or current on a particular readout channel. After last step, data events and simulation events share the same format (simulated events contains extra information from the *truth* level).

The comparison of the objects properties predicted from simulation to data can reveal problems or confirm proper understanding. In most cases the agreement is fairly good, but for some properties, like shower shapes, an additional correction to the simulation is required, in order for the simulation to describe the data. An example of such correction is presented on Figure 3.2.3, where the distribution of the shower shapes of photons from $Z \to l^+l^-\gamma$ events is compared between data and simulation: a clear shift is observed. The origin of this shift is not yet understood. Various methods exist to account for such a difference: applying shifts (explained in Section 3.2), introducing smearing (example in Section 3.4.4) or applying reweighting on MC. An example of a reweighting procedure is presented in the next section.

2.4.2 Pileup Reweighting

To simulate the correct impact from pileup in the simulated events, the number of interactions per bunch crossing, μ, was generated, prior to data taking, to cover the expected experimental conditions. The simulated values of pileup form a discrete set, unlike data where the amount of pileup is a continuous variable. Using event-weights applied on simulation we correct the difference between MC and collected data. The instantaneous luminosity distributions and the trigger prescale configurations are used to define the weights.

A **PileupWeight** is defined [136], which can be applied to Monte Carlo samples to correct for the difference between the instantaneous luminosity distribution used to produce the sample, and the instantaneous luminosity distribution measured in data, as described in Section 2.1.4.

In MC, μ should be scaled to recover the same fraction of visible cross section as in data. First, MC is corrected for difference in longitudinal beamspot size to take into account merged vertex effects. Then, the distribution of the vertex multiplicity, the number of primary vertices (NPV), is compared between data and MC as a function of the instantaneous luminosity. In data μ is determined by independent luminosity detectors sensitive to inclusive number of interactions, and NPV is sensitive to the number of interactions which give significant activity within the ID acceptance. The difference is observed in correlation of NPV with μ between data and MC. The agreement can be improved by scaling the μ of MC by the appropriate factor. The best fit value for a scaling of pileup in MC to match distribution in data is estimated to be: 1.16 ± 0.03 [82]. Due to the discrete nature of the values of μ used in MC, it is more practical to scale the value of μ in data (which is a continuous variable) by the inverse scale factor (1/1.16=0.862069). The correlation between μ and NPV is presented on Figure 2.4.1.

The `PileupReweighting` tool was implemented in the analysis framework to calculate the weights for MC and to correct the pileup in the data. The `iLumiCalc` tool [121] is used to calculate the instantaneous and integrated luminosities for data files. The average number of interactions per bunch crossing for data and Monte Carlo before and after applying the tool are presented in Figure 2.4.2 (left).

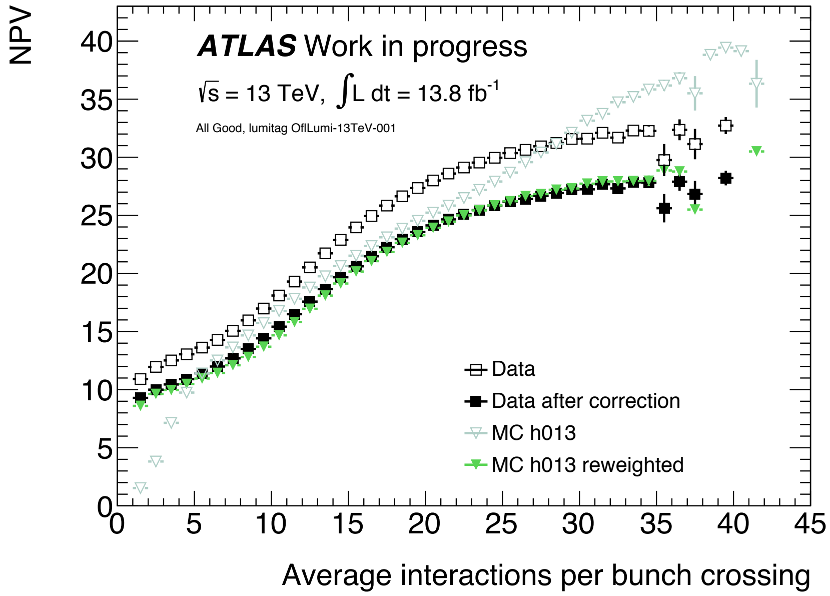


Figure 2.4.1: Number of primary vertices (NPV) versus the average number of interactions per bunch crossing. The points from Monte Carlo (from xAOD) are shown with grey triangles, weighted MC with green triangles, data before scaling with open boxes and filled boxes correspond to data after correction.

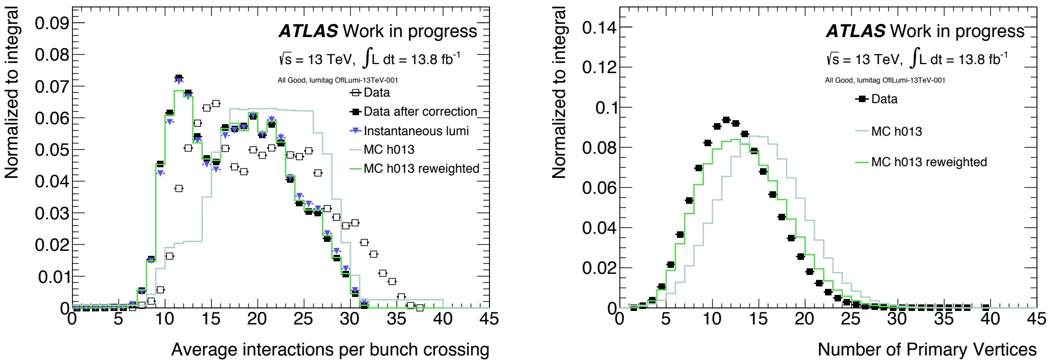


Figure 2.4.2: On the left, the average number of interactions per bunch crossing (from xAOD) for Monte Carlo in grey and data in open boxes. Data after applying the PileupReweighting tool correction corresponds to the filled boxes, the weighted MC in green and the instantaneous luminosity from `iLumiCalc` are shown with blue triangles. Values of the pileup for the data and corrected data (scaled to 1/1.16) are presented . On the right, the number of primary vertices (NPV) for Monte Carlo (from xAOD) is shown in grey, for the weighted MC with green and for the data with boxes.

The impact of the pileup reweighting procedure is crosschecked with the number of primary vertices (NPV) on Figure 2.4.2 (right). The Monte Carlo events have slightly more PVs than the data events, for a given μ due to the "screening" effect. The beamspot is wider in MC than in data. This means that PVs are more separated in MC than in data. In data, we will hence more often have PVs that end up on top of each other, which means we will more often loose a PV since another PV "screens" it. This effect likely explains the small difference between the black squares and the green solid line on Figure 2.4.2 (right).

I was responsible to produce inputs for the creating the pileup weights, to implement tool in the analysis framework and to validate the performance of the reweighting, which are used in the entire Higgs group, including analysis which will be discussed in Chapter 5.

Chapter 3

Photons in the ATLAS Detector

Contents

The reconstruction of invariant mass of the pair of photons is the crucial ingredient when searching for resonances decaying to a pair of photons. The energy and the direction of the reconstructed photon need to be measured with high precision. This Chapter is devoted to description of the methods used to reconstruct and identify photons and to measure their energy.

The photons are identified by an electromagnetic (EM) cluster without associated tracks in the inner detector when the photon has not interacted in the inner detector, or with one or two tracks in case of photon conversion. The reconstructed object can be genuine photon - electromagnetic particle produced in proton-proton (*pp*) collisions, so called *prompt* photon, or the decay product of the initial hadron from hard scattering, or fragmentation photons, denoted as *fake* photons. Hadron misidentified as photon during reconstruction are categorized as fake photons. To separate this two types of objects several criteria are applied to the photon candidates, based on the shape and the properties of the associated electromagnetic showers and by requiring them to be isolated from other particles in the event.

The production of the prompt photons described by Standard Model (SM) mechanisms is presented on Figure 3.0.1, like parton scattering, QED radiation, more complex processes beyond LO and decays of the Higgs boson ($H \to \gamma\gamma$). The diagrams for the production of fake photons are presented on Figure 3.0.2.

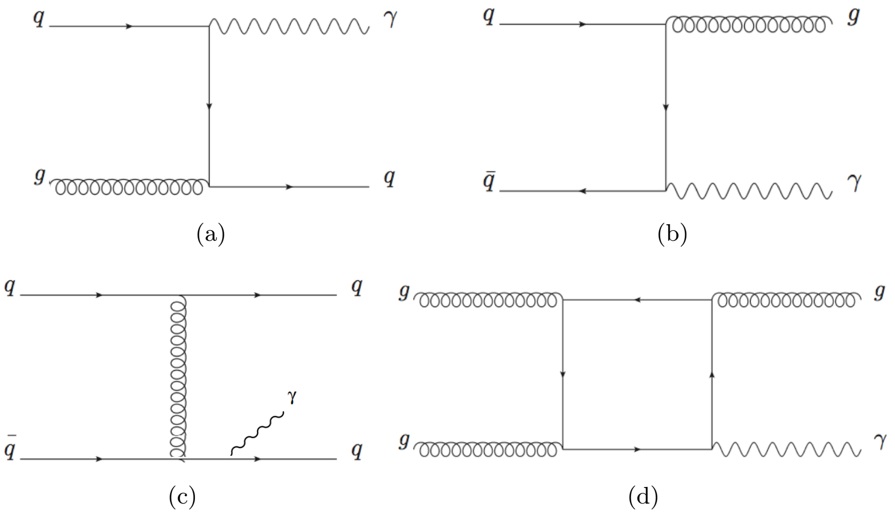
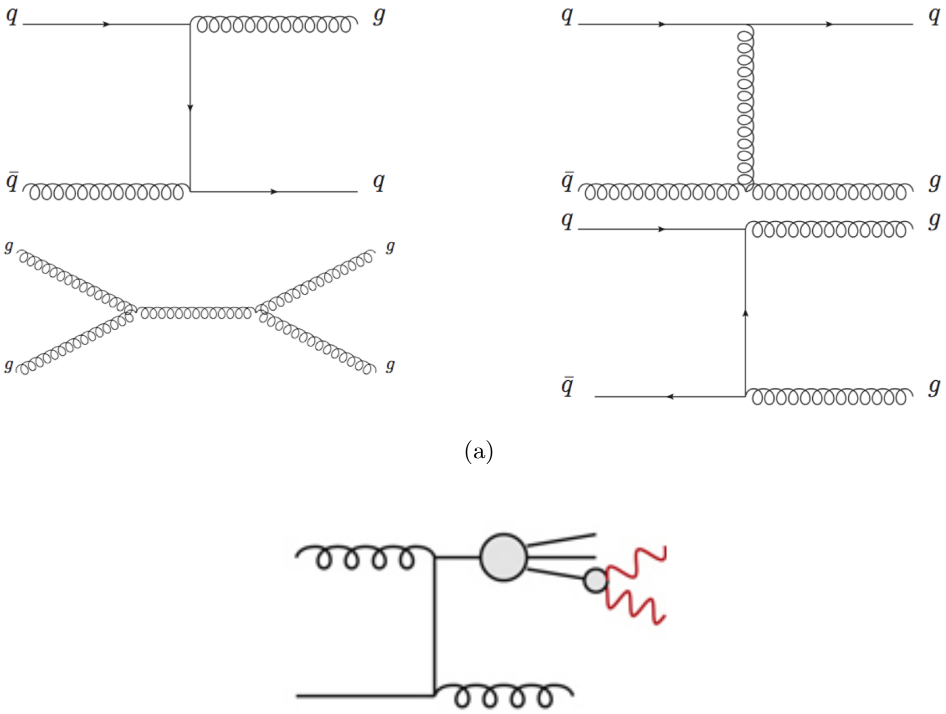


Figure 3.0.1: Feynman diagrams of prompt photon production by a) quark-gluon Compton scattering, b) quark-anti-quark annihilation, c) bremsstrahlung radiation of an outgoing quark and d) NLO process of gluon-gluon annihilation.

In the following, the methods to select prompt photons and to reject fakes are presented. The photon reconstruction and my study of the performance of photon reconstruction in Run-2 are discussed in the Section 3.1. Due to certain ambiguities between the converted photon and the electron candidates, the reconstruction of electrons is also briefly discussed. The different levels of identification of photons are presented in details in Section 3.2. Another property of the object, helping to distinguish between prompt and fake photons is isolation; it is discussed in Section 3.3. The overview of the calibration procedure of the energy response of electromagnetic particles is given in Section 3.4. This section will describe the procedure in details, as an introduction to my contribution to the calibration procedure, which will be discussed in Chapter 4.



(a)

(b)

Figure 3.0.2: Feynman diagrams presenting *fake* photons. a) Diagrams of various processes where quark or gluon initiated jet can be misidentified as a photon, b) photons coming from the decay of the light hadrons.

3.1 Photon reconstruction

Photons and electrons are reconstructed in the ATLAS detector (Figure 3.1.1) by collecting depositing energy into a cluster of cells from the electromagnetic calorimeter, called an EM cluster. The signal reconstruction for each cell is discussed in detail in Section 2.3.4. EM clusters have a fixed size; they are built from contiguous EM cells assembled from the grid of $N_\eta \times N_\phi$ cells from the LAr calorimeter middle layer of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. Inside each of these elements, the energy of all cells in the four longitudinal layers is summed into the cluster energy.

A sliding window algorithm [83] is used to find the *seed* cluster, applying a cut on the transverse momentum above 2.5 GeV. The position of the window of size $\Delta\eta\times\Delta\phi=$ 0.075×0.125 is adjusted to find a local maximum of the transverse energy. The initial cluster reconstruction efficiency for photons with transverse energy E_T >25 GeV is estimated from simulation to be greater than 99%.

After defining the seed cluster, reconstructed tracks in the inner detector are associated to the cluster. Tracks with a loose angular match with the EM cluster are considered for the identification and reconstruction of the electron and the photon conversions. The matching requires the angular distance between the cluster barycenter and the extrapolated track's intersection to the middle layer of EM calorimeter to be less than 0.05 radian along ϕ and η. Tracks with hits in the silicon detectors are extrapolated from the point of closest approach to the primary vertex, while tracks without hits in the silicon detectors are extrapolated from the last measured point.

In case the EM cluster is matched to a pair of oppositely charged tracks in the inner detector, a search for a "double-track" conversion vertex is performed. Collinear track pairs passing the requirement are considered as originating from "converted photon" candidate. Track pairs are classified into three categories, whether both tracks (Si-Si), none (TRT-TRT) or only one of them (Si-TRT) have hits in the silicon detectors.

The profile of signal from the TRT (see Figure 2.2.2 in Section 2.2.1) is used to check the compatibility of the track to be an electron originating from a photon conversion, against the probability to be a charged pion. Tracks without hits in the IBL are considered as photon candidates with a "single-track" conversion vertex . There are different possibilities to loose the second track of the conversion: in case one of the two produced electron tracks is not reconstructed either because it is very soft (asymmetric conversions where one of the two tracks has $p_T < 0.5$ GeV), or because the two tracks are very close to each other and cannot be adequately separated.

Clusters to which neither a conversion vertex candidate nor any track has been matched during the electron reconstruction are considered "unconverted photon" candidates.

The sketch of reconstruction of EM objects presented on Figure 3.1.1. The electron and double-track conversion photon are shown, with tracks associated to the corresponding seed cluster in the EM calorimeter. The unconverted photon has no activity in the inner detector. After the seed cluster has been identified as one of the EM objects (electron, converted or unconverted photon), the final cluster is rebuilt. The cluster size is set to 3×7 and 5×5 calorimeter cells in the second layer (where $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$) in the barrel and endcap calorimeter respectively. These lateral cluster sizes were optimized to take into account the different energy distributions in the barrel and endcap calorimeters while at the same time minimising the impact of pile-up and electronic noise on the total energy. The barrel cluster is broader in ϕ to account for the spread due to opening of electron pair from the converted photon at low E_{T} under magnetic field.

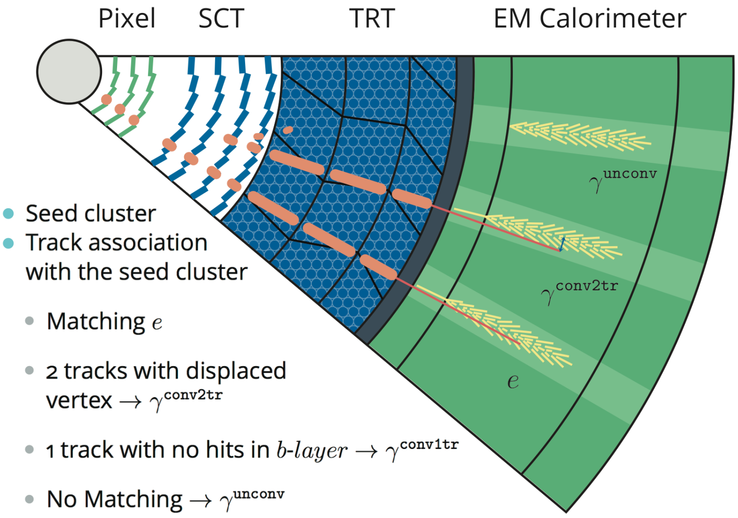


Figure 3.1.1: The sketch of reconstructed EM objects in a sector of the inner detector followed by the electromagnetic calorimeter. Examples of the electron, converted and unconverted photons are shown.

In the beginning of Run-2, during the data preparation effort, I was in charge of the validation of the photon conversion reconstruction with first set of data (called 50ns data, as the protons bunches were still separated by 50 ns), and later update studies with 25 ns datasets. I presented a poster describing these studies at the European Physical Society Conference on High Energy Physics (EPS-HEP) 2015, as part of the ATLAS performance after the start of LHC Run-2 [87].

In the analysis the photon candidates are required to pass the photon identification criteria which are based on the shapes of the associated electromagnetic cluster and to be isolated. Converted photon candidates are also split into single-track and double-track conversions. The fraction of photon candidates reconstructed as unconverted or converted photons as a function of the candidate transverse momentum and average number of interactions per bunch crossing are presented on Figure 3.1.2. The studies shows the excellent stability of the reconstruction over the large range of photon's energy and different pileup conditions. The data sample was collected by ATLAS at the beginning of the 2015 data taking period; it corresponds to about 19pb^{-1} of integrated luminosity. The results are presented in Reference [86]. The results have been updated with a luminosity of 1.7 fb^{-1} (which are presented on Figure 3.1.2) in order to reduce statistical uncertainty. The contamination of background photons from the decays of neutral hadrons in jets is estimated to be smaller than 5% (details can be found in Reference [85]).

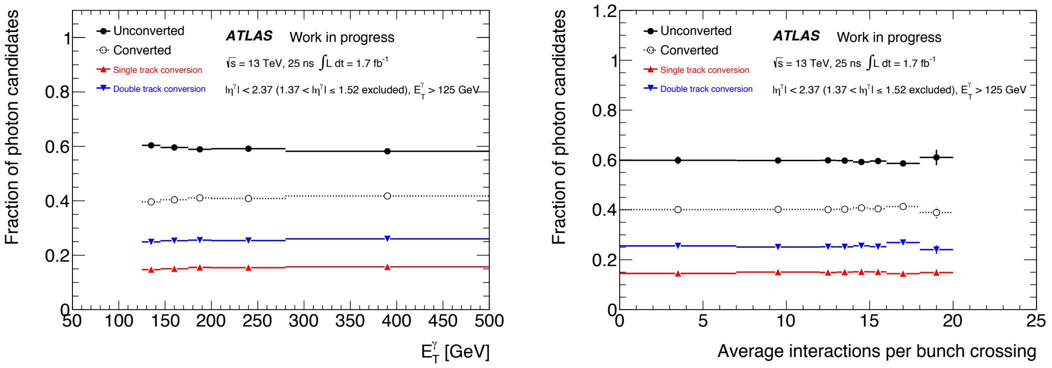


Figure 3.1.2: Fraction of photon candidates reconstructed as unconverted or converted photons in pp collisions at $\sqrt{s} = 13$ TeV, as a function of the candidate transverse momentum (left) and pileup conditions (right).

3.2 Photon identification

The photon identification (ID) performs the selection to distinguish *prompt* photons from *fakes* candidates. The procedure is based on cuts applied on *discriminating variables* (DV), aimed to provide a high signal efficiency and a powerful background rejection. The list of DV is summarised in Table 3.2.1. The high granularity of the first layer (*strips*) of the electromagnetic calorimeter provides detailed information on the lateral shower development. The energy distribution of the shower in the first and second layers is a precious indicator of the longitudinal shower development which can discriminate between electrons and photons vs hadrons. The information from the hadronic calorimeter is used to estimate the leakage outside the EM calorimeter. The fake photons release a substantial portion of the energy in the hadronic calorimeter and produce a broader transverse energy deposit in the EM calorimeter. The prompt photons typically produce narrower energy deposits in the EM calorimeter and have smaller leakage to the hadronic one. The photons from the hadron decays are typically pairs of genuine photons close to each other ($\pi^0 \to \gamma\gamma$) and two separate local energy maxima in the finely segmented strips of the first layer can be recognized on Figure 3.2.1 (right) and the profile of energy deposition is different from the single prompt photon shown on Figure 3.2.1 (left).

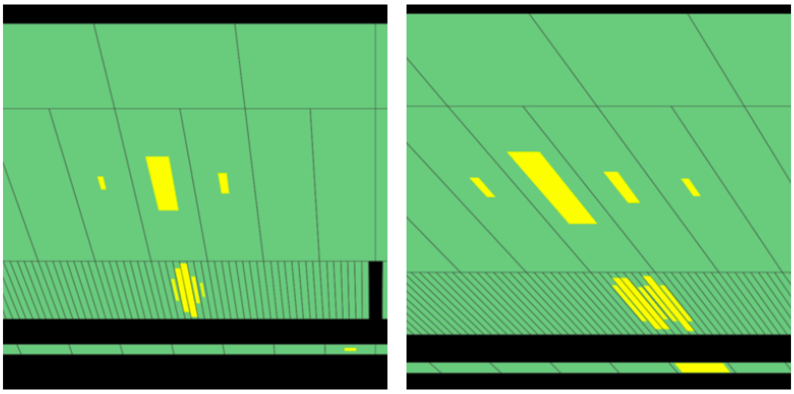


Figure 3.2.1: Shower shapes for a photon candidate (left) and a candidate for a jet decaying to pair of genuine photons $\pi^0 \to \gamma\gamma$ (right), in data recorded in proton-proton collisions [88].

Two reference selections are defined to identify photon candidates; a set of *loose* criteria is used for the preselection of photon candidates and a *tight* - including additional selections, set of criteria is applied to reach a higher purity of the selected photons.

The ***loose*** selection exploits the DVs only in the hadronic calorimeter and in the second layer of EM calorimeter; it is used for trigger and background studies. The *loose* requirements provide a highly efficient selection of 97% for E_T^γ=20 GeV to 99% for E_T^γ >40 GeV, with a fair background rejection factor of order 1000. The rejection is defined as the ratio of the number of initial jets with $p_T > 40$ GeV in the acceptance of the calorimeter to the number of reconstructed background photon candidates satisfying the identification criteria [84].

The ***tight*** ID requirement exploits information from the full granularity of the EM calorimeter - the finely segmented strip layer of the calorimeter - and applies tighter requirements also on the DVs used by the *loose* ID. Optimization of the requirements on the DVs are done separately for unconverted and converted photons, separately in different pseudorapidity regions. The *tight* criteria provides a photon identification efficiency of about 85% (90%) for unconverted (converted) photon candidates with transverse energy E_{T} >40 GeV and reaches 93% (99%) for candidates with transverse energy E_{T} >100 GeV, as shown on Figure 3.2.4. The *tight* ID requirement corresponds to a background rejection factor of about 5000.

The list of the DVs used by the *loose* and *tight* photon identification algorithms is given in Table 3.2.1. The illustration for DVs are summarised on Figure 3.2.2.

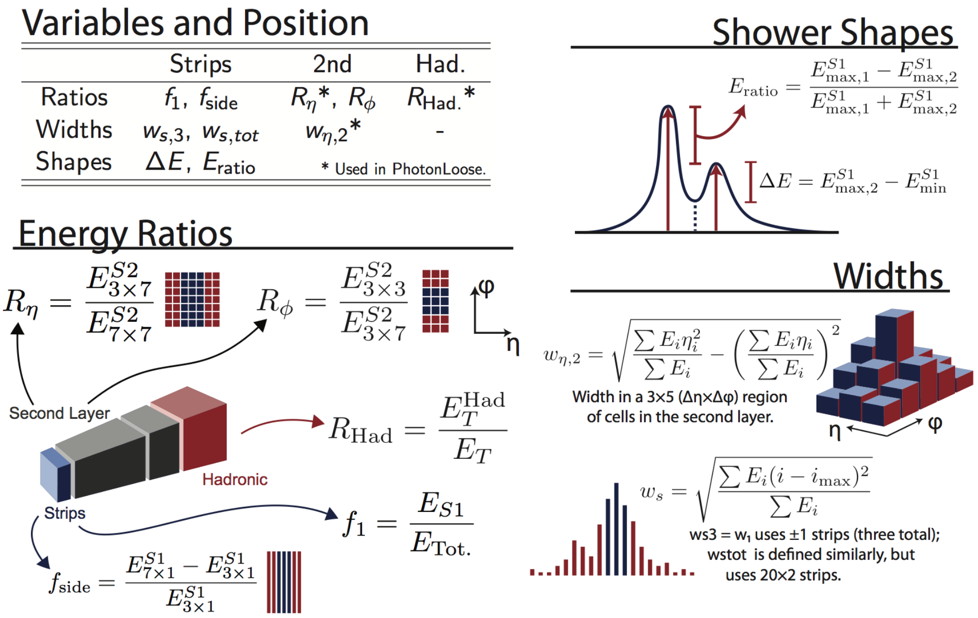


Figure 3.2.2: Sketch of the calorimetric discriminating variables used for photon identification in ATLAS [89].

The optimization of the selections applied on the DV are derived from Monte Carlo simulation, but as was previously discussed in Section 2.4.1, we know that sometimes MC does not perfectly describes the data. A systematic shift in the distributions of the shower shapes is observed, while the shapes of distributions are in agreement. In order to improve the description of the photon DVs, corrections are applied to the simulated values (Figure 3.2.3). These shifts are calculated by minimizing the χ^2 between the data and the shifted MC distributions of photon candidates, for different pseudorapidity and transverse momentum regions. The typical size of the correction factors is 10% of the RMS of the distribution of the corresponding variable in data.

Category	Description	Name	*loose*
Acceptance	$\|\eta\| < 2.37$, with $1.37 < \|\eta\| < 1.52$ excluded	–	●
Hadronic leakage	Ratio of E_{T} in the first sampling layer of the hadronic calorimeter to E_{T} of the EM cluster (used over the range $\|\eta\| < 0.8$ or $\|\eta\| > 1.37$)	R_{had_1}	●
	Ratio of E_{T} in the hadronic calorimeter to E_{T} of the EM cluster (used over the range $0.8 < \|\eta\| < 1.37$)	R_{had}	●
EM second layer	Ratio of 3×7 $\eta \times \phi$ to 7×7 cell energies	R_η	●
	Lateral width of the shower (as a function of η)	w_{η_2}	●
	Ratio of $3{\times}3$ $\eta \times \phi$ to $3{\times}7$ cell energies	R_ϕ	
EM first layer	Shower width (as a function of η) calculated from three strips around the strip with maximum energy deposit	$w_{s\,3}$	
	Total lateral shower width (as a function of η)	$w_{s\,\mathrm{tot}}$	
	Energy outside the core of the three central strips but within seven strips divided by energy within the three central strips	F_{side}	
	Difference between the energy associated with the second maximum in the strip layer and the energy reconstructed in the strip with the minimum value found between the first and second maxima	ΔE	
	Ratio of the energy difference associated with the largest and second largest energy deposits to the sum of these energies	E_{ratio}	

Table 3.2.1: Discriminating variables used for the photon identification [90]. All variables are used for *tight* crateria, while *loose* use only the shower shape in the EM layer 2 and the energy deposited in the hadronic calorimeter, mentioned with bullets in the table.

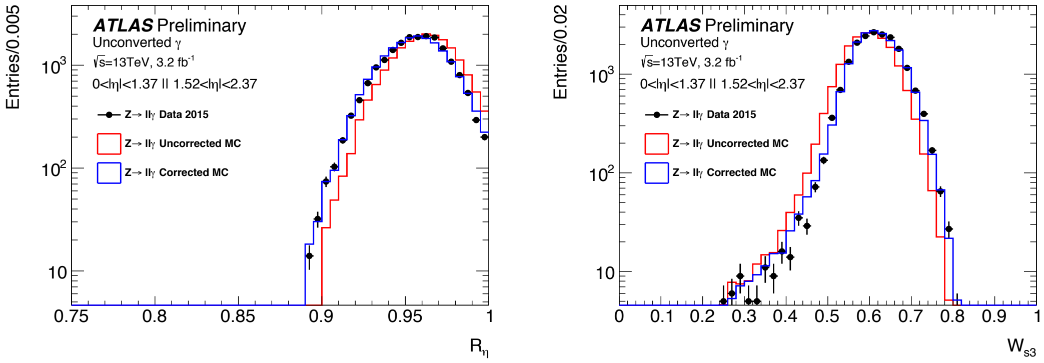


Figure 3.2.3: Two examples of shower-shape variables for unconverted photons: the black points are for the photons from $Z \to l^+l^-\gamma$ events from data, the red histogram for the photons from $Z \to l^+l^-\gamma$ event from simulation before applying the shift, and the blue histogram after applying the shift [91].

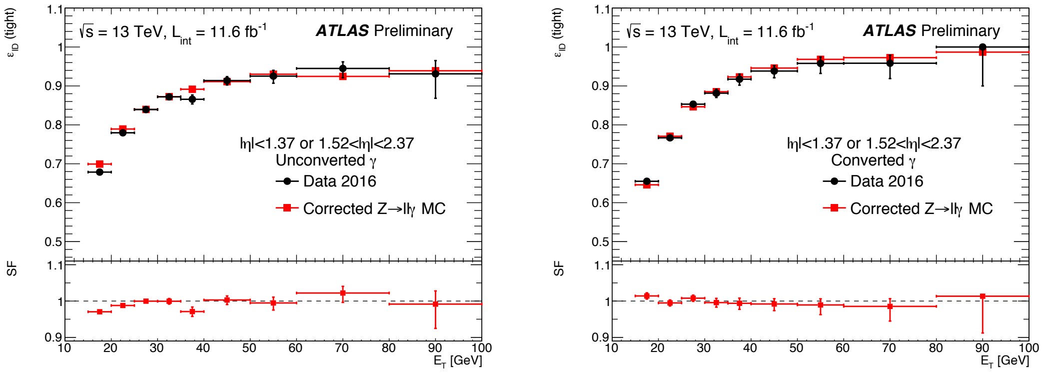


Figure 3.2.4: Comparison of the radiative Z boson data-driven tight identification efficiency measurements of unconverted (left) and converted (right) photons to the $Z \to l^+l^-\gamma$ simulation as function of E_T of the photon in the region 15< E_T <100 GeV, for the entire pseudorapidity region (excluding the transition region between the barrel and end-cap calorimeters) [92]. The bottom panel shows the ratio of the data-driven results to the MC predictions (scale factor, SF).

3.3 Photon isolation

The amount of the energy deposited around the reconstructed object provides information of its isolation, and can be used as a discriminant variable to further separate prompt photons from fake photons. Typically, fake photons are surrounded by a sizeable hadronic activity, and is therefore poorly isolated. The isolation is computed both in the calorimeter using the deposited energy and in the inner detector using tracks.

The calorimetric isolation variable $E_{\mathrm{T}}^{\mathrm{iso}}$ is defined as the sum of the transverse energy E_{T} of topological clusters inside a cone around photon candidate, within an angular distance $\Delta R = \sqrt{(\eta - \eta^{\gamma})^2 + (\phi - \phi^{\gamma})^2} = 0.4$, subtracting the area $\Delta\eta \times \Delta\phi = 0.125 \times 0.175$ corresponding to the energy of the photon candidate, as shown on Figure 3.3.1. The energy collected in the cluster does not correspond to the total energy deposited by the incoming photon. A fraction of its energy (typically a few percents) is deposited outside the cluster and therefore enters the isolation cone. A correction, function of the photon transverse energy, and estimated from simulation, is therefore applied. The contribution of the underlying event and of the pileup is subtracted from the energy inside the isolation cone, event-by-event [93, 94]. An example of the distributions of the calorimeter isolation variable is presented on Figure 3.3.2.

The track isolation variable $p_{\mathrm{T}}^{\mathrm{iso}}$ is defined as scalar sum of the transverse momenta of all the tracks with transverse momentum above 1 GeV, within a cone of $\Delta R = 0.2$ around the cluster direction. Only tracks having a distance of closest approach to the primary vertex along the beam axis ($|z_0 \times \sin\theta|$) of less than 3 mm considered. The tracks associated to photon conversions are excluded from the sum.

The efficiency and background rejection of the isolation requirement will be presented on the example of the diphoton high mass search selections, which will be discussed in details in Section 5.2.1. The efficiency of the combined isolation requirement for photons passing these selections in signal diphoton MC samples is 90% to 96% in the E_{T} range 100 GeV to 500 GeV, with an uncertainty between 1% and 2% [109]. The isolation requirement reduces the rate at which jets are misidentified as photons by about one order of magnitude.

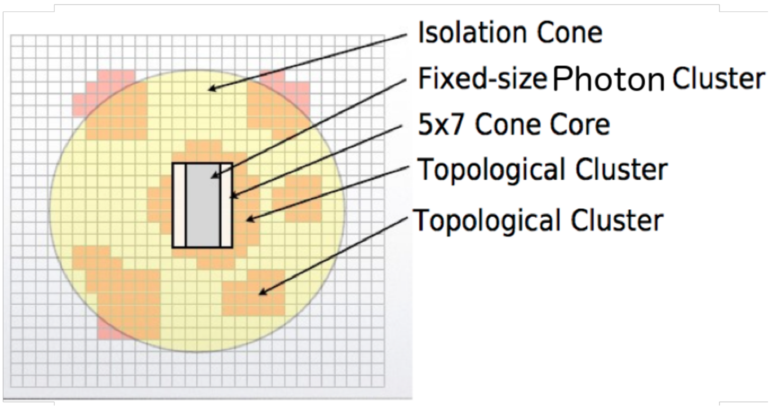


Figure 3.3.1: Sketch of the isolation cone around the photon candidate in the calorimeter. The photon shower is excluded from computation of the isolation transverse energy, and it is corrected for the leakage of the photon shower into the isolation cone.

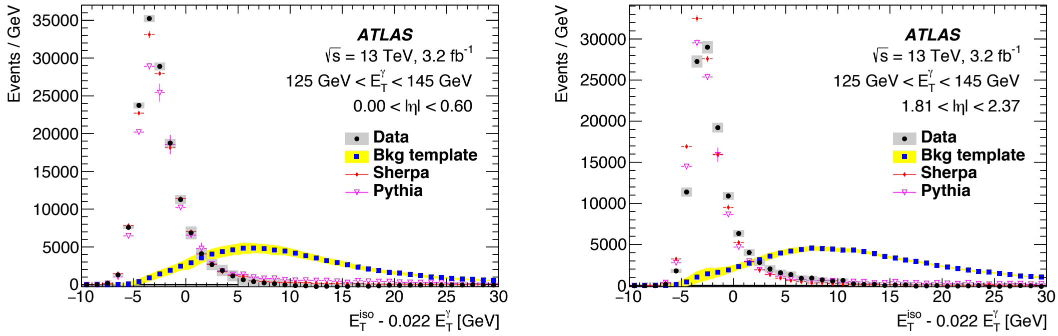


Figure 3.3.2: Distributions of the calorimeter isolation variable ($E_\mathrm{T}^\mathrm{iso} - 0.022E_\mathrm{T}(\gamma)$) for photon candidates fulfilling the tight identification criteria for 125 GeV< E_T <145 GeV for barrel (left) and endcap (right) regions. The background contribution to the data, 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 $E_\mathrm{T}^\mathrm{iso} - 0.022E_\mathrm{T}(\gamma) > 12$ GeV. The 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. The calorimeter isolation requirement corresponds to a cut at 2.45 GeV on this variable [109].

3.4 Photon energy calibration

An excellent performance of the electron and the photon energy calibration is an important ingredient in most physics analyses, in particular for precision measurements of the parameters of the Standard Model processes with electrons and photons in the final state. Electromagnetic particles deposit most of their energy inside the calorimeter, but a small fraction may be lost in the material in front of the calorimeter. Longitudinal or lateral leakages and local effects lead to a loss of reconstructed energy inside the EM cluster. Therefore, the calibration of reconstructed object is required to recover for any energy loss and to reconstruct the particle's initial energy.

The calibration of the photon and electron energies starts with the electronic calibration, done at the cell level, which was described in Section 2.3.4. This calibration converts the observed signal in ADC counts to a measured energy in GeV. The calibration of the reconstructed cluster energy of electromagnetic particles is based on a multivariate regression algorithm correcting for the energy loss, originating from the unaccounted deposition of the energy in front of the calorimeter, from the longitudinal and lateral leakage and for other local effects. The steps of the energy calibration for data and simulation are summarized on the sketch presented on Figure 3.4.1. The details of the MC-based calibration algorithm will be discussed in Section 3.4.1, the methods to account for inter calibration of the calorimeter layers are presented in Section 3.4.2, and the corrections applied on the collision data are discussed in Section 3.4.3. After applying the steps mentioned above, the calibrated energy between data and simulation does not perfectly match. To account for the residual disagreement, η-dependent correction factors are extracted using $Z \to ee$ events through an in-situ procedure, as discussed in Section 3.4.4. The energy scale of electrons applied on data and on Monte Carlo simulation is smeared to match the energy resolution in data. The detailed description of the calibration procedure can be found in Reference [95, 96].

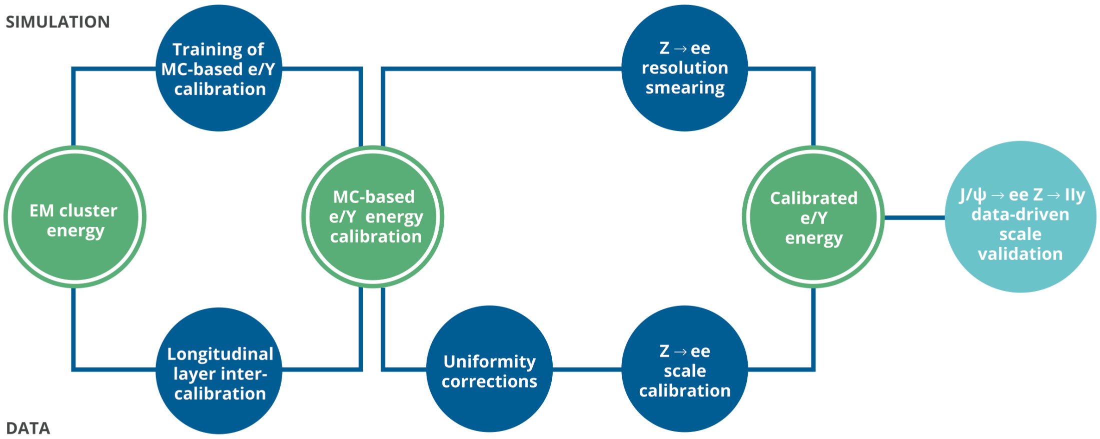


Figure 3.4.1: Schematic overview of the procedure used to calibrate the energy response of electrons and photons in ATLAS.

3.4.1 Training of MC-based calibration

The goal of the calibration is to reconstruct the initial particle's energy from the deposited energy in the EM cluster, corresponding to the sum of the energy collected in the three layers within calorimeter tower as described in Section 3.1. Also, one should recover for energy loss in front of calorimeter, and the longitudinal segmentation of the EM calorimeter and the presempler (PS) helps to provide necessary information. The large energy deposition in the presampler and large ratio of the energies deposited in first and second sampling layers is a hint, that the particle starts showering by interacting with the material before EM calorimeter. The absence of signal in PS with large deposition in the sampling calorimeter indicates that most probably the particle is an unconverted photon, as illustrated on Figure 3.4.2 (left). The measured track properties from the inner detector provides additional information for the converted photons, to estimate the energy loss before the EM calorimeter.

To take into account the correlation between the parameters describing the shower development, a multivariate regression algorithm is implemented in the calibration procedure [97]. This MC-based calibration relies on an accurate description of detector geometry and interactions of particles with matter in the MC simulation. The cluster-level calibration constants are extracted from single particle simulations without pileup, based on a multivariate boosted decision tree [98]. It is performed separately for electrons, converted and unconverted photons in bins of transverse momentum and pseudorapidity. The binning was chosen to match approximately the non-uniformities of the detector and significant changes in the response.

The variables used in the training of the MVA for electrons and photons are listed below:

- $E_{\text{calo}} = \sum\limits_{i=1}^{3} E_i$, the total energy deposited in the three layers of LAr calorimeter
- E_0/E_{calo}, the ratio of the presampler energy to the calorimeter energy, used up to $|\eta| < 1.8$ (available PS coverage)
- E_1/E_2, the ratio of the values of the energy measured in the first two layers of the calorimeter
- η_{cluster}, the cluster barycentre pseudorapidity in the ATLAS coordinate system [1]
- η_{calo} and ϕ_{calo}, the cluster barycentre in the calorimeter frame
- $E^{\text{TileGap}}_{\text{E4scintillators}}$, the energy measured by the E4 scintillators installed in the transition region between the barrel and endcap compartments of the EM calorimeter.

For converted photons additional parameters are included in the training. Photons are considered converted if the conversion radius R_{conv} is smaller than 800 mm. The empty squares, in the list below, correspond to conversions with both tracks containing at least one hit in either the pixel or SCT detector.

---

[1] The ATLAS coordinate system has its origin at the nominal interaction point, with respect to which the calorimeter is displaced by a few millimeters in the barrel and up to ~5 cm in the endcaps, while all calorimeter cells are in their nominal position in the EM calorimeter frame.

- ■ R_{conv}, the conversion radius (only if the vectorial sum of the conversion track momenta, p_{T}^{conv}, is above 3 GeV)

- ☐ $p_{\mathrm{T}}^{\mathrm{conv}}/E_{\mathrm{calo}}$, the ratio of the vectorial sum of the conversion track momenta to he calorimeter energy

- ☐ $p_{\mathrm{T}}^{\mathrm{max}}/p_{\mathrm{T}}^{\mathrm{conv}}$, the fraction of the conversion momentum carried by the highest-pT track

The output of the MVA simulation-based calibration, is a set of calibration constants for numerous bins in $E_{\mathrm{T}} \times \eta$ fine segmented grid. The same constants are applied on Monte Carlo simulation and collision data. A relative calibration is applied to the layers of EM calorimeter in the data, before applying the calibration weights. This correction is discussed in the next Section 3.4.2.

3.4.2 Intercalibration of the LAr calorimeter longitudinal layers

The response of the layers in the calorimeter needs to be corrected in data to adjust for residual effects not perfectly accounted by the cell electronic calibration, in order to ensure the correct extrapolation of the response in the full p_{T} range used in the various analyses.

The energy scale between the first and the second layers has been studied using muons from $Z \to \mu\mu$ decays. Most of muons from this decay are minimal ionizing particles, which are insensitive to material upstream the calorimeter. They deposit a constant amount of energy over the full depth of the calorimeter and constitute a direct probe of the energy response. To perform the relative calibration between two layers, the average ratio of energy deposited in the strips to the middle layer $\langle E_{1/2}\rangle$ in data is compared to Monte Carlo simulation. The intercalibration is done, by applying an η-dependent correction $\alpha_{1/2} = \langle E_{1/2}\rangle^{\mathrm{data}}/\langle E_{1/2}\rangle^{\mathrm{MC}}$ to the second layer energy measured in data: $E_2^{\mathrm{corr}} = E_2 \times \alpha_{1/2}$. The results of intercalibration are verified using electrons from $Z \to ee$ decays.

The energy scale of Presampler is estimated from the ratio of the PS energy deposited by electrons from W and Z decays in data and MC, denoted $\alpha_{\mathrm{PS}} = E_0^{\mathrm{data}}/E_0^{\mathrm{MC}}$. The simulated energy, E_0^{MC}, is corrected to account for any material mis-modeling with respect to data. The correction is derived using the correlation between E_0 and $E_{1/2}$ variables. Different assumptions of additional material budget in the detector geometry description were made. Increased material upstream the calorimeter in simulation leads to earlier shower development, which results to the increase of the PS energy E_0 and to a higher fraction of the strips energy (increase of $E_{1/2}$). While any extra material between the PS and strips will not change E_0 but will impact on $E_{1/2}$ - introduce showering which leads to increase E_1. Unconverted photons are less affected by upstream material mismodeling than electrons (Fig. 3.4.2), and therefore can be used to determine a possible bias in the simulation: compare $E_{1/2}^{\mathrm{data}}/E_{1/2}^{\mathrm{MC}}$ requiring low presampler energy.

The contribution of the third layer to the total energy scale for Z decays is neglected - contribution of E_3 to full energy is of the order of 1%.

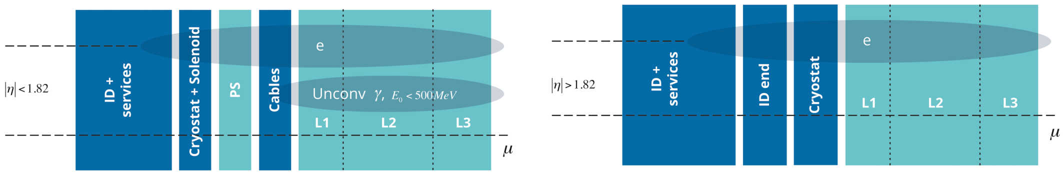


Figure 3.4.2: Sketch of EM shower development for the different particle categories, for $|\eta| < 1.82$ (left) and $|\eta| > 1.82$ (right).

3.4.3 Detector non-uniformities

The detector has some features in specific regions, which are not described in the detector geometry. The high voltage in a few sectors of the calorimeter is set to a non-nominal values due to short circuits occurring in specific LAr gaps. The gravity-induced widening of the gaps between modules in ϕ, following the structure described in Section 2.3.2 leads to energy loss between the barrel calorimeter modules; this effect is not described in the simulation. Therefore, data-driven corrections have been optimised to mitigate the non-uniformity in the detector response and are applied to data.

Another correction was associated with the different gains of the calorimeter cells readout electronics (High/Medium Gain), but additional studies described in Chapter 4, revisited the impact of this correction on MVA-based calibration, and justify that correction is not needed. The correlation with the shower development has been observed, and a corresponding systematic uncertainty has been derived to cover the difference of the energy response in data and MC simulation.

3.4.4 In-situ calibration

After applying the layer inter calibration on data, the energy response is calibrated using the MVA-based procedure both on data and MC, and then the detector uniformity corrections applied on data. After those steps, a residual disagreement in the energy scale and resolution may be present between data and simulation. In-situ corrections using $Z \to ee$ events are derived in order to correct this residual difference.

The energy mis-calibration is defined as the difference in response between data and simulation, and is parametrised as follows:

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

where E^{data} and E^{MC} are the electron energy in data and simulation, and α_i represents the departure from optimal calibration, in a given pseudorapidity bin labelled i. Electron resolution corrections are applied as additional effective constant term for a given η region:

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

The measured values of the energy scale correction factor α_i and the additional constant term c' for the energy resolution from $Z \to ee$ events are measured using data collected in 2015 and 2016. The distributions of α_i and c' are shown on Figure 3.4.3.

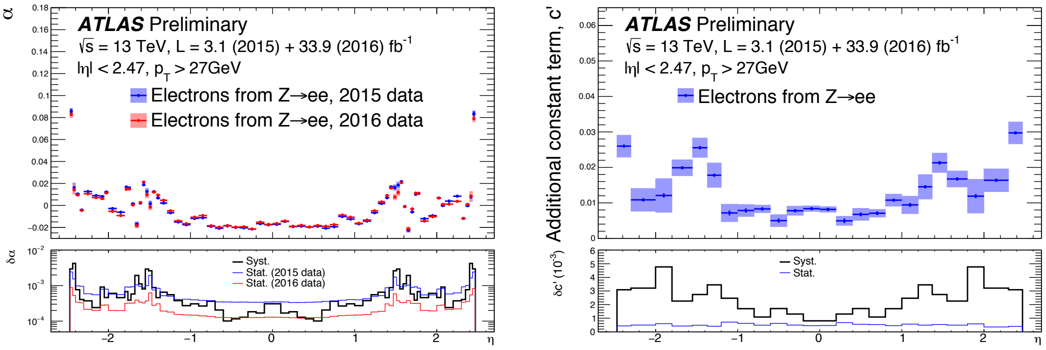


Figure 3.4.3: Energy scale correction factor α_i (left) and additional constant term c' for the energy resolution (right) from $Z \rightarrow ee$ events as a function of η are measured using data collected in 2015 and 2016. The total systematic uncertainty on this quantity is represented as the blue band on the top plot and as the thick black line on the bottom plot. In the bottom inset, the thin blue line represents the 2015 + 2016 statistical uncertainty [99].

After all corrections, the accuracy of the whole calibration procedure is checked by comparing the dielectron (m_{ee}) mass distribution in data and simulation. The response in data and Monte Carlo agrees at the level of 2% in the mass range 80< m_{ee} <100 GeV, as shown on Figure 3.4.4. The instantaneous luminosity delivered by the LHC has largely increased during 2015 and 2016, and the average number of simultaneous interactions per bunch crossing (μ) has increased from 13.6 for 2015 to 25 for the 2016 dataset. The stability of the calibration has been studied as a function of μ and of time, estimated to be stable at the 0.05% level, presented on Figure 3.4.5.

3.4.5 Systematic uncertainties

The main sources of systematic uncertainties (with largest contribution) are discussed below. The full list can be found in Reference [96].

- **Presampler:** the uncertainty on the calibration of the thin presampler layer α_{PS}.
- **Layer intercalibration:** the uncertainty on the scale factors $\alpha_{1/2}$. An additional systematic uncertainty related to the observed discrepancy between Run-1 and Run-2 intercalibrations (~1.5%) has been included in the uncertainty model.
- **LAr electronic response:** this source of uncertainty has been reevaluated, and a detailed description can be found in Chapter 4.
- **Material:** the uncertainties relative to the description of the material before the calorimeter for $|\eta|$ <2.5. An additional source of uncertainty for the mismodeling

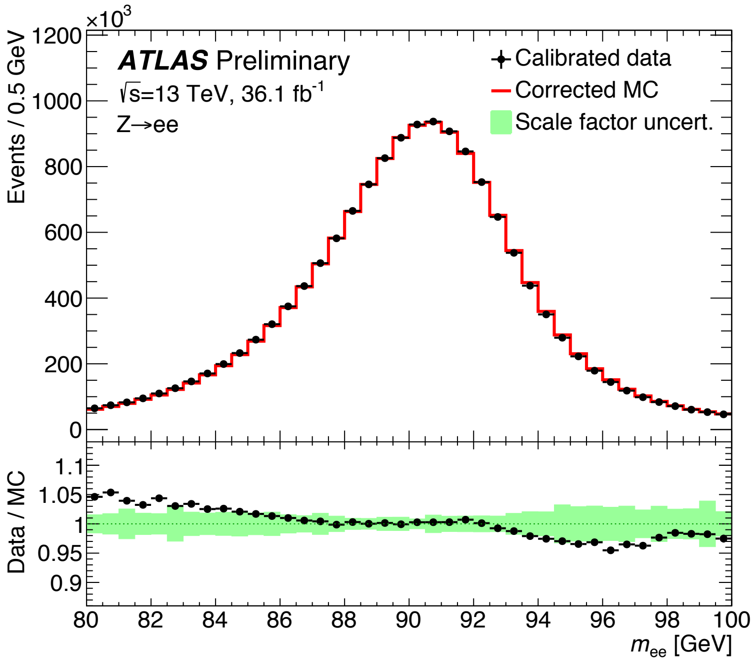


Figure 3.4.4: Inclusive dielectron invariant mass distribution from $Z \to ee$ decays in data compared to MC after applying the full calibration. No subtraction of the background (expected to be at the level of 0.5% and with a non-peaking m_{ee} distribution) is applied, and the simulation is normalised to data. The lower panel shows the data to simulation ratio, together with the total scale factor uncertainty [100].

of the IBL material in the simulation has been considered for Run-2. It includes the $PP0$ area for the inner detector which is different between Run-1 and Run-2.

- **In-situ calibration ($Z \to ee$):** The uncertainty on the scale and additional constant term have been re-evaluated using Run-2 datasets,
- **Pileup and temperature:** The uncertainty related to difference between pileup conditions and temperature difference of the liquid argon in between 2015 and 2016 datasets in Run-2.

The main sources of uncertainty on the energy scale and the ones that have been updated are shown in Figure 3.4.6 for $|\eta|$ <0.6 and 5 GeV< E_T <1TeV for electrons, unconverted and converted photons.

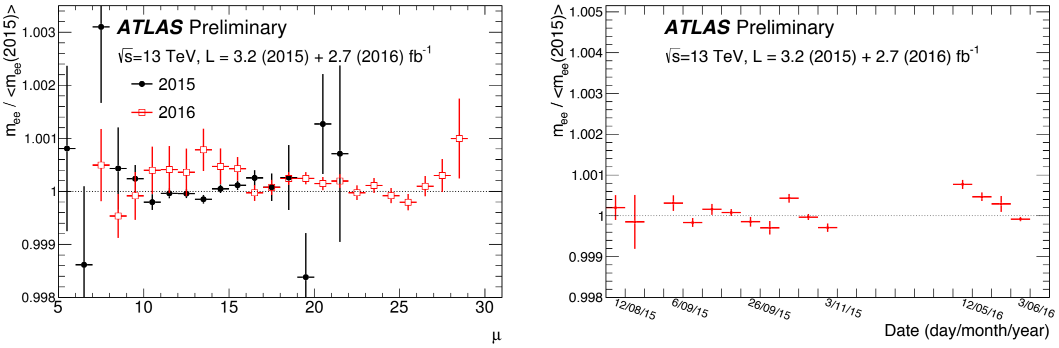


Figure 3.4.5: Stability of the average value of the electron pair invariant mass for $Z \to ee$ events as a function of the average number of interactions per bunch crossing, μ, (left) and as a function of the time of data taking using data collected in 2015 and 2016 (right). Each bin shows the average invariant mass divided by the average invariant mass measured in 2015.

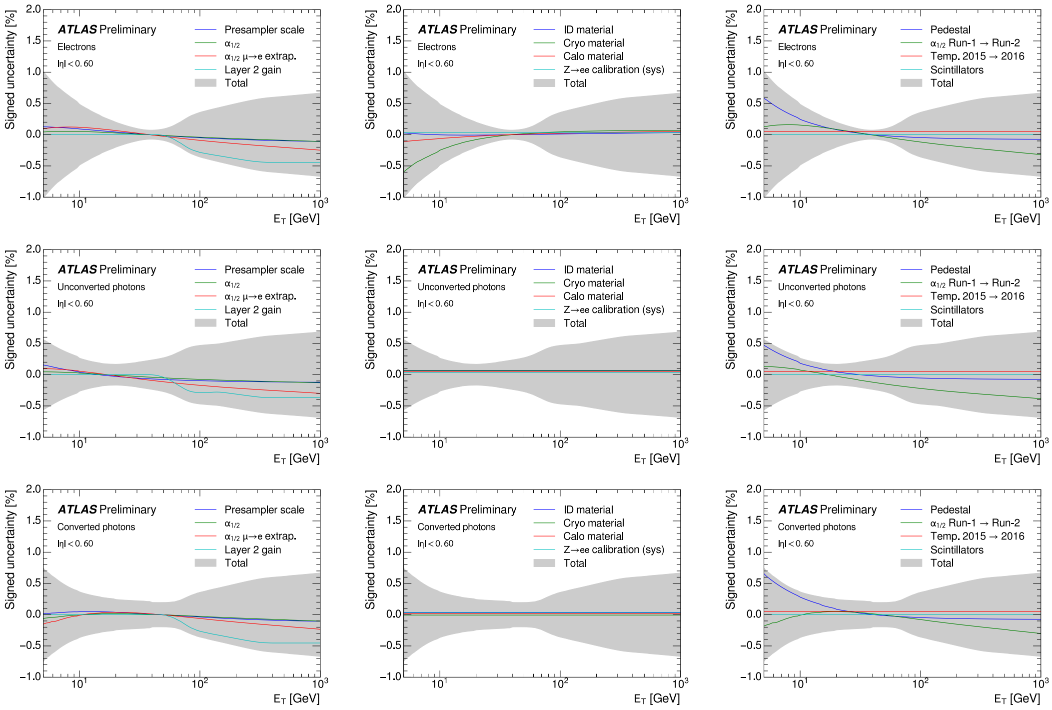


Figure 3.4.6: Main contributions to the systematic uncertainty on the energy scale after calibration corrections, shown as a function of E_T and for $|\eta|$ <0.6, in the case of electrons (top row), unconverted photons (middle row), and converted photons (bottom row). The total uncertainty bands represent the quadratic sum of all systematic uncertainty sources described in the text and represented in the various columns. The last column in particular shows the additional systematic uncertainties introduced from Run-1 to Run-2 and the temperature uncertainty added when comparing 2015 data to 2016 data.

Chapter 4

Improvement of energy calibration in Electromagnetic calorimeter

Contents

The precision of the measurements of the Standard Model parameters in the processes with electrons and photons in the final state, strongly depend on the systematic uncertainty on the energy measurement. For example, the measurement of the Higgs boson mass in the $H \to \gamma\gamma$ decay channel is presented using Run-1 data collected at center-of-mass energies of 7 TeV and 8 TeV, corresponding to an integrated luminosity of 25 fb^{-1} [102]. The measured value of the Higgs boson mass is :

$$\begin{aligned} m_H &= 125.98 \pm 0.42(\text{stat}) \pm 0.28(\text{syst})\ \text{GeV} \\ &= 125.98 \pm 0.50\ \text{GeV} \end{aligned}$$

where the first error represents the statistical uncertainty and the second the systematic uncertainty. This measurement is dominated by the statistical uncertainty, but in the Run-2 with the increased center-of-mass energy (larger cross section) and expected integral luminosity of the order 100 fb^{-1}, the statistical uncertainty will decrease down to order of $\sim 150 - 180$ MeV. Therefore, the key point is to reduce the systematical uncertainty related to the photon energy measurement, to be able to improve the measurement of the mass.

The dominant contribution to the total systematics uncertainty originates from the photon energy scale calibration and corresponds to 0.21% uncertainty. One of the largest source is denoted as "LAr cell nonlinearity" - uncertainty on the nonlinearity of the energy measurement at the cell level: this arises mostly from the relative calibration of the different gains used in the calorimeter readout, which was discussed in details in Section 2.3.4. The uncertainty on the nonlinearity of the cell energy calibration contributes an uncertainty of about 0.1% to the energy scale of photons from Higgs boson decays (up to 1% for $1.5 < |\eta| < 1.7$).

This Chapter will be devoted to understand the linearity of the response at the electronics level, find the way to address its associated systematic uncertainty. The study of the observed effect of the non-linearity of the energy response will be revisited using an improved calibration (discussed in the previous Chapter 3 in the Section 3.4) using the Run-1 dataset. Its impact is summarised in Section 4.1. The same procedure is performed for the Run-2 dataset and a similar effect is observed; this is summarised in Section 4.1.1.2. The correlation of the effect is observed between the gain switch and the lateral development of the shower in the calorimeter as discussed in Section 4.1.3. Therefore, the study of the initial "LAr cell nonlinearity" contribution is split into the two questions:

1. Is there an intrinsic non-linearity in the calorimeters electronics readout?
2. Is the observed difference of the energy response coming from the calibration procedure, failing to treat different types of shower?

The studies related to the intrinsic electronics response (1) have been studied by collaborators, and I provide a short summary in Section 4.1.2. The investigation of the response of the calibration procedure (2) is discussed in Sections 4.1.1 and 4.1.3, and a possible solution accounting for the shower development is discussed in Section 4.2.

4.1 Energy response in high and medium gain

With the Run-1 dataset, a difference in the energy response between the high and medium gain of the ATLAS electromagnetic calorimeter was observed [101]. The average Z boson mass difference between data and simulation events (Data-MC) is of opposite sign between the events reconstructed in different electronic gain (High or Medium gain): $\Delta m_{\mathrm{HG}}<0-\Delta m_{\mathrm{MG}}>0$. The double difference (Data-MC, HG-MG) reaches 1 GeV in pseudorapidity region 1.5< $|\eta|$ <1.8 (denoted below as "bad region"). An ad-hoc correction was derived to correct for this effect and a systematic uncertainty was assigned, corresponding to size of the difference in each η region. This correction denoted as "LArCell non-linearity", will be named "Gain Correction" in the following. While studying further the "LArCell non-linearity", after the publication of the result of the Higgs boson masse measurement, it was understood that the "Gain Correction" had been derived from a dataset calibrated with an older calibration procedure (calibration hits) and not with the procedure applied to the dataset used for the publication. It turned out that the older calibration procedure indeed had a bias on the energy reconstruction. It became obvious that the effect of the Gain Correction had to be revisited using the proper calibration procedure (MVA) [96].

The studies described below, have been performed on reprocessed data and MC from Run-1 to estimate impact of new calibration on observed effect. Then, the checks of the energy response for different gain chains will be performed for Run-2 dataset. Discussion of the other possible sources of this discrepancy will be performed in Section 4.1.3.

4.1.1 Impact of MVA calibration on effect of the electronic gain

The difference in the energy response between data and MC is observed. The checks performed to analyse the impact of the updated calibration procedure on the energy response, for different electronic gains will be discussed, both for Run-1 and Run-2 datasets in following Section.

Methodology

In the studies presented below, the electrons from Z decay will be categorised in terms of the electronic gain in which they were readout. Event will be denoted as "Medium" gain electron (MG), if at least one cell of the cluster of the second layer of EM calorimeter (3×7 for Barrel and 5×5 for EC)[1] was readout using "Medium" gain. If all cells of the cluster of the second layer were readout in High gain, the electron is categorised as High gain (HG). The events with at least one cell readout in Low gain are not used in the analysis. As the most energetic cell usually carries one third of the cluster's energy, this split intrinsically introduces a kinematical bias on the selected events.

The gain switching in MC is done at a given and constant ADC hits count. In data the threshold of switching between gains are slightly different for different regions of

---

[1] Detector cells used: 97 (EMB2) or 289 (EMEC2 OW)

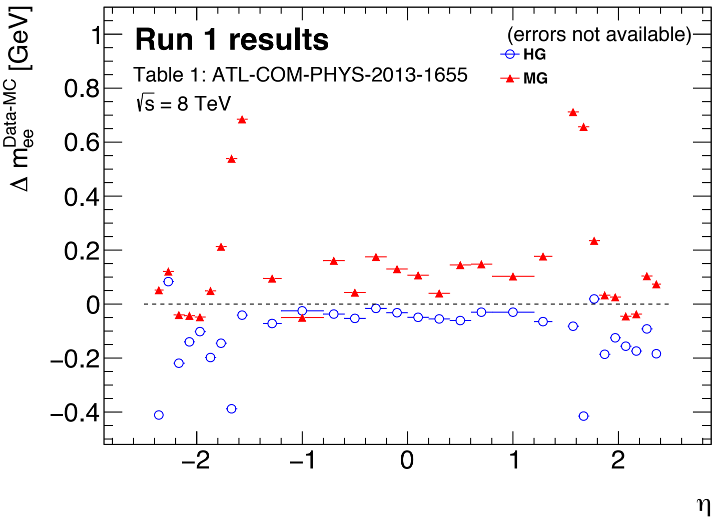


Figure 4.1.1: The average impact of the gain in η bins observed in Run-1 [101]. The plot is produced using values from Table 1 from the note.

calorimeter. Thus, in the region of p_T close to transition between High and Medium gain, the gain assigned to event could be different between data and MC. To avoid such bias in the computation of the difference in energy response (data-MC), the events will be split in small bins of energy, and the difference will be taken bin-by-bin.

One m_Z value enters two categories or twice the same category, depending on the gain and η of each electron from the Z.

4.1.1.1 Impact of MVA calibration on Run-1 data

The size of effect observed in Run-1 of difference in the mass reconstruction between events recorded using High and Medium electronics gain is presented on Figure 4.1.1. A correction was derived to remove the difference between two categories. The size of the correction was assigned as a systematic uncertainty of the electronics non linearity. The Run-1 dataset will be used to estimate improvements, which MVA calibration introduces to the energy reconstruction of the Z boson mass. The simulated samples and data from Run-1 collected at the centre of mass energy $\sqrt{s}$=8 TeV had been reprocessed using the latest reconstruction and calibration (MVA-based) procedure. Further in this section, the "gain correction" is switched off to study the initial observed effect.

Impact of MVA on gain effect in MC

To estimate the impact of the modified procedure of the calibration (i.e. MVA calibration without applying the Gain correction), the comparison of Z events as a function of the electrons transverse energy in MC for different gain is performed for Hit-based (old) and MVA (new) calibrations, as presented on Figure 4.1.2. In the barrel region, the agreement between two calibration procedures is excellent for the energy reconstructed in different gain categories, while the endcap region in Hit-based calibration shows a clear distortion at high energy. For MVA-based calibration, there is no significant difference over the entire pseudo rapidity acceptance. Figure 4.1.3 shows the difference of the reconstructed mass in MC between gain categories (MG-HG) taken in bins of energy, comparing the two calibration procedure methods. It shows that in MC, where gain switching is modelled by definition linear, the Hit-based calibration generate a significant discrepancy in the energy reconstruction between gain categories. This can be a sign, that the observed (Data-MC) difference can be related to the calibration procedure and not to the gain switching linearity.

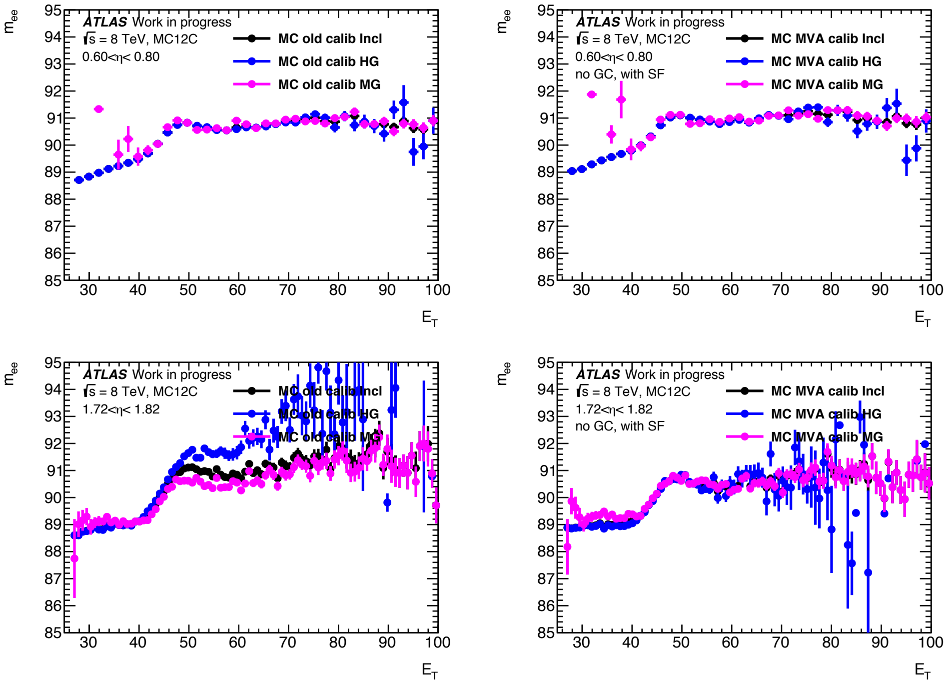


Figure 4.1.2: Invariant mass distribution m_{ee} in MC as a function of the electrons transverse energy for Hit-based on the left, and MVA-calibration on the right. The inclusive distribution presented in black, MG events in pink and HG in blue. The top plots present the barrel region, 0.6< η <0.8, and bottom plots the End-Cap "bad-region", 1.72< η <1.82.

Impact of MVA on gain effect in data

The comparison of the energy response for different gain categories was performed in data for two calibration procedures, and the difference in MG-HG as a function of transverse energy is shown on Figure 4.1.4. An energy dependence is observed for endcap region for both methods, while the size of dependency is slightly smaller in

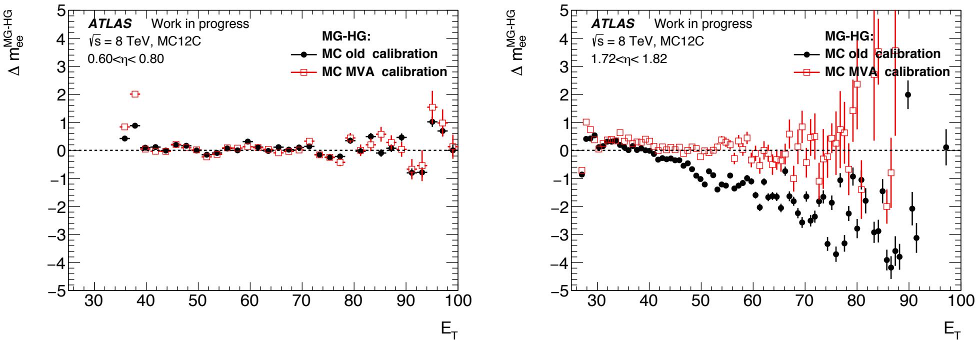


Figure 4.1.3: Difference MG-HG of the invariant mass distribution Δm_{ee} in MC as a function of transverse energy of the electron. The Hit-based calibration is presented in black, and the MVA-calibration calibration in red. The left plot presents the barrel region, 0.6< η <0.8, and the right plot the End-Cap "bad-region", 1.72< η <1.82.

MVA-based calibration. The energy reconstruction in the barrel region is stable with respect to transverse energy between gain categories for both calibrations, while for the Hit-based method an offset of the order of 200 MeV is observed. For the Hit-based calibration, there is no difference between the gain categories for the barrel region in MC but an offset is observed in data, as visible on Figure 4.1.1, where the difference data-MC is presented as a function of η for the two categories. In the barrel region the double difference (data-MC)(MG-HG) corresponds to 200 MeV. This offset is not present in MVA-based calibration, where inter-layer calibration (E_1/E_2) is performed, as discussed in Section 3.4.2. To check this, the LayerRecalibration option was switched off in the MVA-based calibration and compared to the invariant mass difference between gain categories reconstructed with Hit-based calibration. The result is presented on Figure 4.1.5, where the gain difference in the barrel region agrees between two methods.

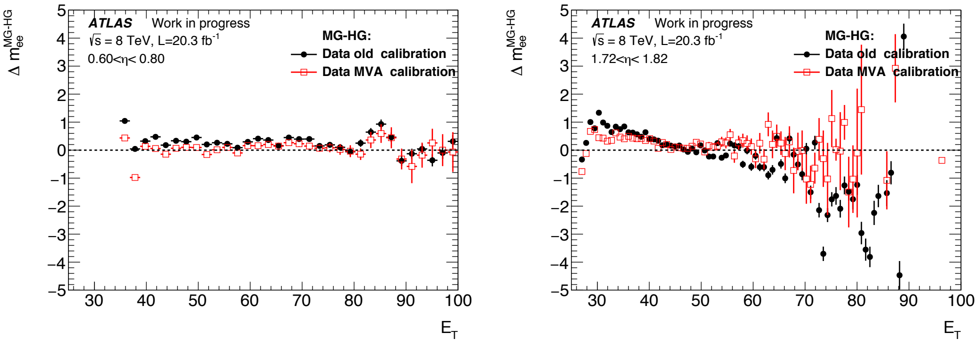


Figure 4.1.4: Difference MG-HG of invariant mass distribution Δm_{ee} in Data as a function of transverse energy of the electron. The Hit-based calibration presented in black, and MVA-calibration in red. The left plot for barrel region, 0.6< η <0.8, and the right for End-Cap "bad-region", 1.72< η <1.82.

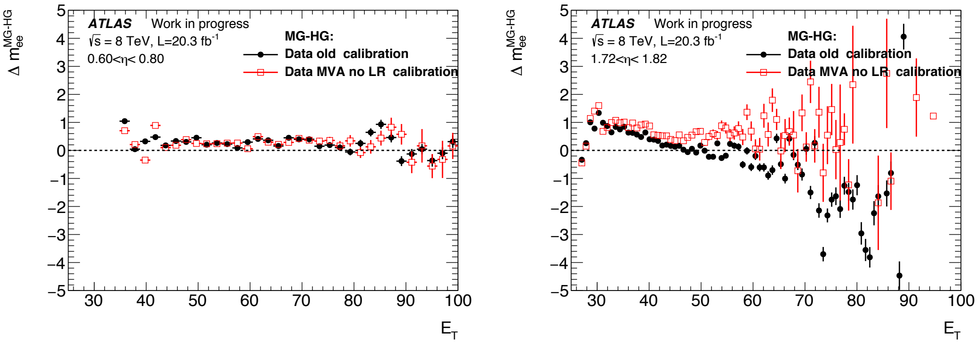


Figure 4.1.5: Check of impact of layer recalibration on difference MG-HG of invariant mass distribution Δm_{ee} in Data as a function of transverse energy of the electron. The Hit-based calibration presented in black, and MVA-calibration without LayerRecalibration in red. The left plot for barrel region, 0.6< η <0.8, and the right for End-Cap "bad-region", 1.72< η <1.82.

The impact of the MVA-based calibration on the final difference in the invariant mass distribution Data-MC ($\Delta m_{ee}^{Data-MC}$) for the two gain categories over all pseudo-rapidity region is presented on Figure 4.1.6 (to be compared with Figure 4.1.1). Each point of the plot is obtained by averaging over electrons energy, taking the weighted average. The size of difference in "bad" region in the endcap remains significant. As discussed above, applying the layer recalibration, removes the systematic offset in the barrel region. Therefore, the impact of the double difference (data-MC)(MG-HG) exist only in particular region of the calorimeter. In the next Section, the check of the effect will be discussed using Run-2 data collected at $\sqrt{s} = 13$ TeV.

4.1.1.2 The difference in energy response between Data and MC in Run-2

The same procedure discussed above was performed for Run-2, comparing the difference between Data and MC at $\sqrt{s} = 13$ TeV. The impact of the different gain categories on the reconstructed Z mass in data and MC is compared. To avoid any kinematic bias, in each η bin difference is taken in bins of transverse energy of the leading electron. The summary of data-MC difference for energy response in gain categories is presented on Figure 4.1.7. The results are similar to the ones obtained with the Run-1 dataset. Both data and simulation sets have been reprocessed with the latest reconstruction and calibration (MVA-based).

The fact that the observed effect strongly depends on the way the electrons are calibrated, and that the effect is localised to a particular region of the calorimeter, triggers the question about the origin of the effect. The study of intrinsic impact of the gain switching will be discussed in the next Section 4.1.2. The correlation of the energy response with the properties of the electron shower will be discussed in Section 4.1.3.

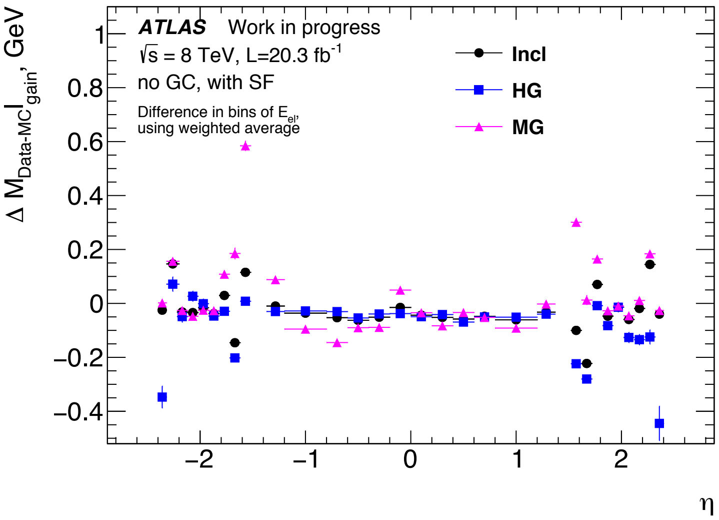


Figure 4.1.6: Distribution of the difference of invariant mass reconstructed with MVA-based calibration, $\Delta m_{ee}^{Data-MC}$ between Data and MC as a function of the pseudorapidity of the electron, for the two gain categories. The Medium gain category is presented with pink points, the High gain category in blue, and the inclusive distributionis in black.

4.1.2 Study of the impact of the intrinsic electronics using special data

In order to study the impact of the electronics gain switch, a special run was taken in 2015[1], in which the electronics was configured such that cells would be readout in MG, down to a few GeV. This run was taken to try to factorize the impact from the electronics chain from the shower shape development. The integrated luminosity collected with these settings was 12pb^{-1}.

The special run corresponds to data taking conditions, where the threshold of switching between Medium Gain and High gain was lowered in layer 2 of electromagnetic Barrel and Endcap, (60 ADC instead of 300 ADC) such that almost all electrons from Z decays had a least the highest energy cell in layer 2 recorded using the medium gain readout. The invariant mass distribution of $Z \to ee$ events is compared between ”special” and ”standard” runs, corresponding to nominal switching thresholds. The advantage of this ”special” run with respect to standard data taking conditions, is that most cells from the electrons clusters are readout in a single gain, avoiding the bias when studying the shower shape. This measurement can then be interpreted as a constraint on the relative calibration of the medium and high gain readouts and the resulting uncertainty can be propagated to the uncertainty on the energy linearity for electrons and photons in standard physics run.

The relative energy scale calibration between the special and the standard runs

---

[1] lower HG/MG switching threshold runs: 276073 , 276147, 276161, 276176 , 276181, 276183, 276189, 276212, 276245 (periods D1+D2)

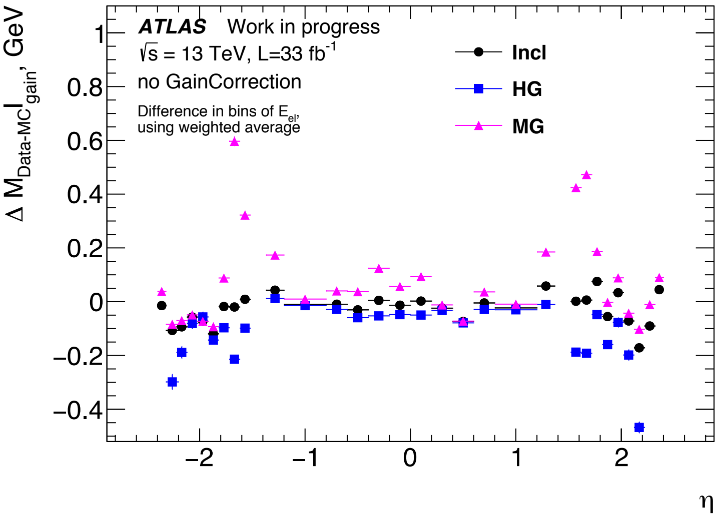


Figure 4.1.7: Difference of invariant mass distribution $\Delta m_{ee}^{Data-MC}$ between Data and MC at $\sqrt{s}$ = 13 TeV, as a function of the pseudorapidity of the electron, for the two gain categories. The Medium gain category is presented with pink points, the High gain category in blue, and the inclusive distributions in black.

is derived by comparing the invariant mass distributions from $Z \rightarrow ee$ events. The α energy scale corrections are computed in 5 η regions following the two steps procedure as described in Section 3.4.4 for the data/MC energy scale. The derived scale corresponds to the correction to apply on the standard standard run energy scale to match the special run energy scale. The obtained values for different η bins are consistent with 0 within two standard deviations, meaning that given the available statistics no significant energy scale difference is found between the two datasets. Those results are documented in [103].

4.1.3 Shower Shapes

The dependence of the energy response with the electronics gain readout has been questioned in the previous section. The study presented in this section shows that the energy response depends on the development of the shower, which is tightly correlated to the gain readout. For example, very broad showers will deposit less energy per cell in the second layer of the calorimeter, and will therefore have a lower probability to be made of cells which exceed the threshold and be recorded in MG. For very narrow showers it is the opposite, the energy deposit being collimated in fewer cells: it is then more probable that one cell goes above the threshold and be readout in MG. The lateral size of the shower can we expressed as its width measured in the first Layer of the calorimeter. It is determined in a window $\Delta\eta \times \Delta\phi = 0.0625 \times 0.2$, corresponding typically to 20 strips in η:

$$\omega_{\text{stot}} = \sqrt{\frac{\sum E_i(i - i_{max})^2}{\sum E_i}}, \tag{4.1}$$

where i is the strip number and i_{max} the strip number of the first local maximum. The illustration of the ω_{stot} in the calorimeter and an example of the ω_{stot} distribution are shown on Figure 4.1.8, left and right correspondingly.

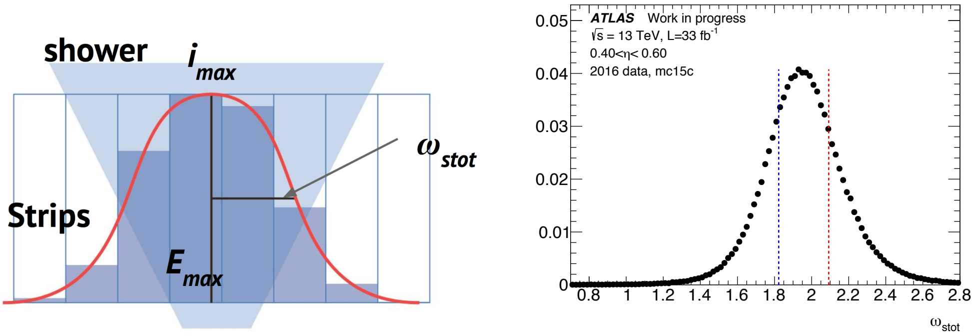


Figure 4.1.8: Left: The sketch of the shower development in the first layer of EM calorimeter and illustration of the definition of the shower width (ω_{stot}) Right: The distribution of ω_{stot} variable in the pseudo rapidity region 0.4< η <0.6.

The calibration procedure includes the information of the longitudinal shower development, assuming that it correlates with the lateral development. The residual energy variation with shower width are possible if longitudinal shower shapes don't account for all fluctuations related to shower shapes. Therefore, the energy response should be uniform for various showers widths. To check the impact of the lateral width of the shower on the energy calibration, the reconstructed Z mass as a function of the ω_{stot} is presented in each pseudorapidity bin as defined in previous Section. Figure 4.1.9 shows comparison of dependency of m_Z with shower width between data and MC simulation in different pseudorapidity bins. The top plots correspond to the barrel region (0.6< η <0.8) and for the endcap, in the region 2.12< η <2.22, a good agreement between data and MC is observed. In the contrary, a dependence with the shower width is present in some of the endcap regions. The bottom plots correspond to the "bad" region (1.52< η <1.62 and 1.62< η <1.72), where a significant difference in energy response is observed between data and MC.

The qualitative correlation of the lateral size of the shower with gain categories was discussed in the beginning of this Section. To provide a quantitative comparison of the observed effect of the difference of energy response between gain and shower shape, we split the data in terms of different shower size. Three regions of ω_{stot} distribution are defined, using interquartile ranges: "Narrow" showers correspond to the first quartile of ω_{stot} distribution, "Normal" showers make the core of distribution (IQR) and "Wide" correspond to the right side of the distribution (third quartile). The example of the division is presented on the right of Figure 4.1.8, where the blue line corresponds to

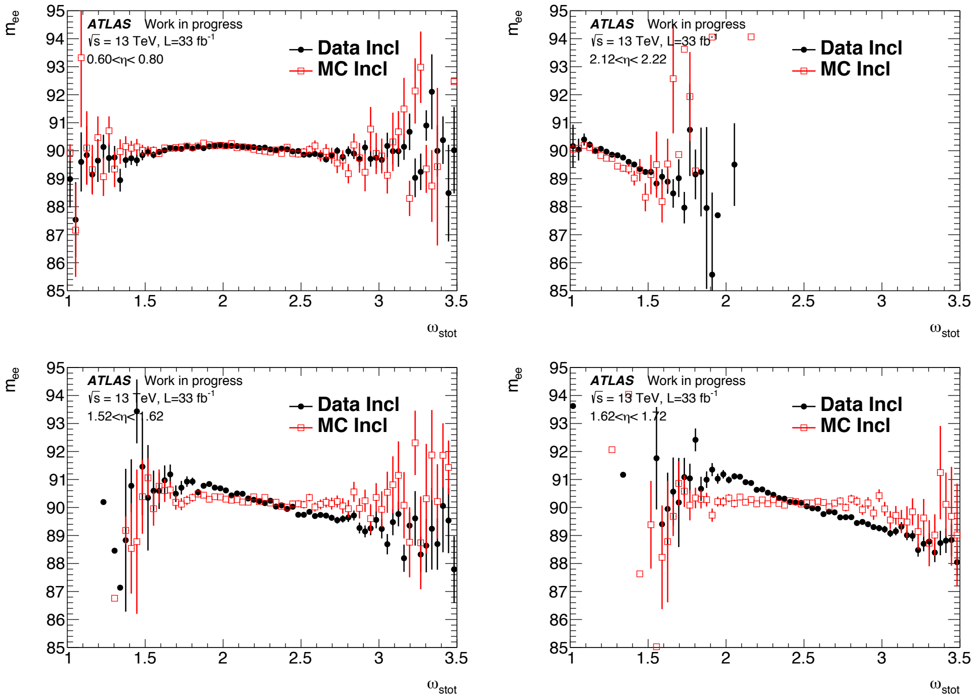


Figure 4.1.9: Dependence of invariant mass distribution for Data and MC as a function of lateral shower width (ω_{stot}). Top plots present barrel (0.6< η <0.8) and far endcap (2.12< η <2.22) bins, where behaviour of data is well reproduced by MC. Bottom plots correspond to "bad" region, where significant discrepancy is observed.

the first quartile and the red line to third quartile; the events between the two lines are categorised as "Normal" showers. Since there is some minor discrepancy between the modelling of the ω_{stot} in data and MC (small shift but similar shape), the categorisation is done separately for data and MC in each η bin. The reconstructed Z mass as a function of energy of the leading electron is performed in categories of the shower width for data and MC. The bin by bin difference data-MC, for four different pseudo rapidity regions, is presented on Figure 4.1.10. The top plots show a good agreement between the response in different categories, while in the bottom plots, corresponding to the bad region, a clear difference is observed for Narrow and Wide showers with respect to the Nominal (and inclusive) case.

The energy dependence is stable, the weighted average is taken in each η bin for each category. The effect of discrepancy between different shower sizes over the overall pseudo rapidity range is presented on Figure 4.1.11. A similar structure is observed, as for the categories made with the gain readout, as was presented on Figure 4.1.7.

This study shows, that the difference in the energy response may come from the miscalibration of the different types of the shower of the particle. Since this is the property of the shower, we may improve the calibration procedure, by introducing additional parameter in the training. This should equalise response between data and MC in the regions where behaviour is not uniform now. The studies of this approach will be

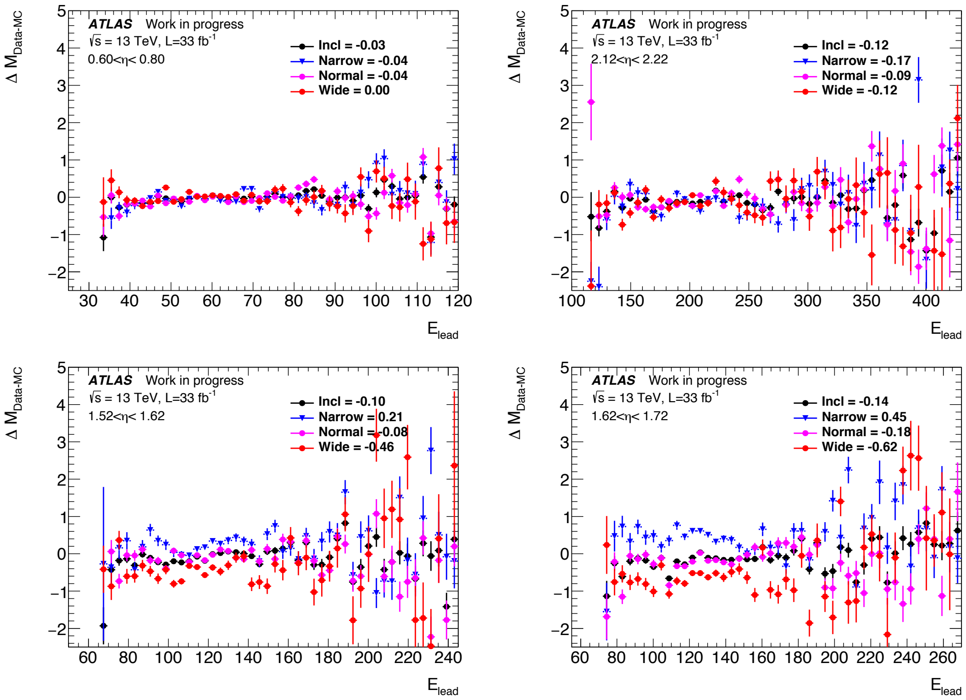


Figure 4.1.10: The difference of invariant mass distribution for Data and MC as a function of energy for categories of the lateral shower width (ω_{stot}). Top plots presents barrel (0.6< η <0.8) and far endcap (2.12< η <2.22) bins, the bottom plots correspond to "bad" region.

presented in Section 4.2. Meanwhile, the observed difference between shower categories should be covered by systematic uncertainties, which is presented in Section 4.1.4.2.

4.1.4 Systematics uncertainty

The list of systematic uncertainties, used for early Run-2 analysis contains the "LAr-cell non-linearity" contribution. As it was shown in Section 4.1.1, the observed Data to MC difference was deduced after applying the MVA-based calibration. The effect observed in categories of electronic gain has a strong correlation with the lateral shower shape categories, which can explain the difference in energy reconstruction coming from the lateral leakage of the shower, discussed in Section 4.1.3. The study of the impact of the intrinsic electronics gain switching was presented in Section 4.1.2. Therefore, the correction related to the gain non-linearity from Run-1 and the corresponding systematic uncertainty were removed. This requires to set a new systematic uncertainty, which will cover the difference presented on Figure 4.1.11. The systematic uncertainty will be split into two parts. The corresponding intrinsic gain non linearity is summarised in Section 4.1.4.1; and the second systematic uncertainty, related to the shower shape, is presented in Section 4.1.4.2.

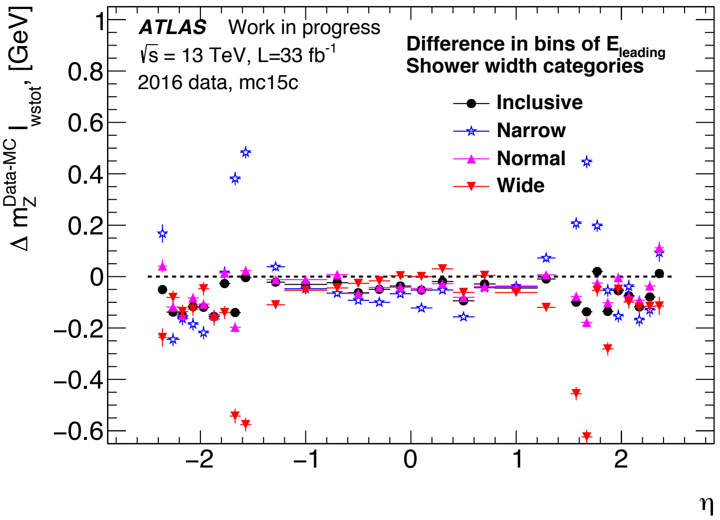


Figure 4.1.11: Difference of invariant mass distribution $\Delta m_{ee}^{Data-MC}$ between Data and MC at $\sqrt{s} = 13$ TeV, as a function of pseudorapidity of the electron for lateral shower width categories.

4.1.4.1 Systematic uncertainty on the special run analysis

An uncertainty on the scale difference between the two datasets (special and standard run) has been derived, and is documented in the note [103]. The final uncertainty related to the intrinsic gain miscalibration is presented as a sum of contributions from: α uncertainty, the special run sensitivity and the impact on electron and photon energies. The typical uncertainty for 60 GeV $E_{\rm T}$ unconverted photons is about 0.10-0.15% in the barrel region and 0.3-0.5% in the endcap region.

4.1.4.2 Systematics related to shower width

Different behaviour of data and MC for dependency of reconstructed invariant mass as a function of lateral shower width in particular regions of the calorimeter is observed, as shown on Figure 4.1.9. Corresponding systematic uncertainty should cover this difference, before necessary improvements will be introduced in the training procedure, to account for lateral shower development.

The origin of the discrepancy may come from difference of shower shape ($\omega_{\rm stot}$) in Data and MC, this implies a change in lateral leakage. The correlation between leakage and width, different in data and MC, can drive the uncertainty. A systematic deviation of the electrons energies translates in a deviation of m_Z following the formula: $\Delta m_Z = \frac{m_Z}{2}\left(\frac{\Delta E_1}{E_1} + \frac{\Delta E_2}{E_2}\right)$. For electrons from Z events, in particular at $p_{\rm T}$ =40 GeV, the difference in data to MC comparison should be absorbed in in-situ scales. Therefore, the systematic uncertainty can be estimated as:

$$\frac{\Delta E}{E}(E_{\mathrm{T}}, p) = \frac{2}{m_Z}\Big(A*[\omega_{\mathrm{stot}}^{\mathrm{Data}}(E_{\mathrm{T}}, p) - <\omega_{\mathrm{stot}}^{\mathrm{Data}}(40, e)>] - B*[\omega_{\mathrm{stot}}^{\mathrm{MC}}(E_{\mathrm{T}}, p) - <\omega_{\mathrm{stot}}^{\mathrm{MC}}(40, e)>]\Big), \tag{4.2}$$

where A is a slope of dependency on the electron m_Z as a function of ω_{stot} in data for the particular η bin and B is the same slope on MC. These slopes are presented on Figure 4.1.9. The variable p is the particle type: electron(e), converted or unconverted photon (γ^{conv}, γ^{unconv}).

The parameters A and B are obtained from linear approximation of the dependency $m_Z(\omega_{\mathrm{stot}})$ in each η bin for data and MC, in the region where most of the statistics lie[1]. The correlation with energy of parameters A and B are checked, splitting data and simulation in bins of p_{T} ([27,40,50,100] GeV), and no strong dependency is observed, following Figure 4.1.12.

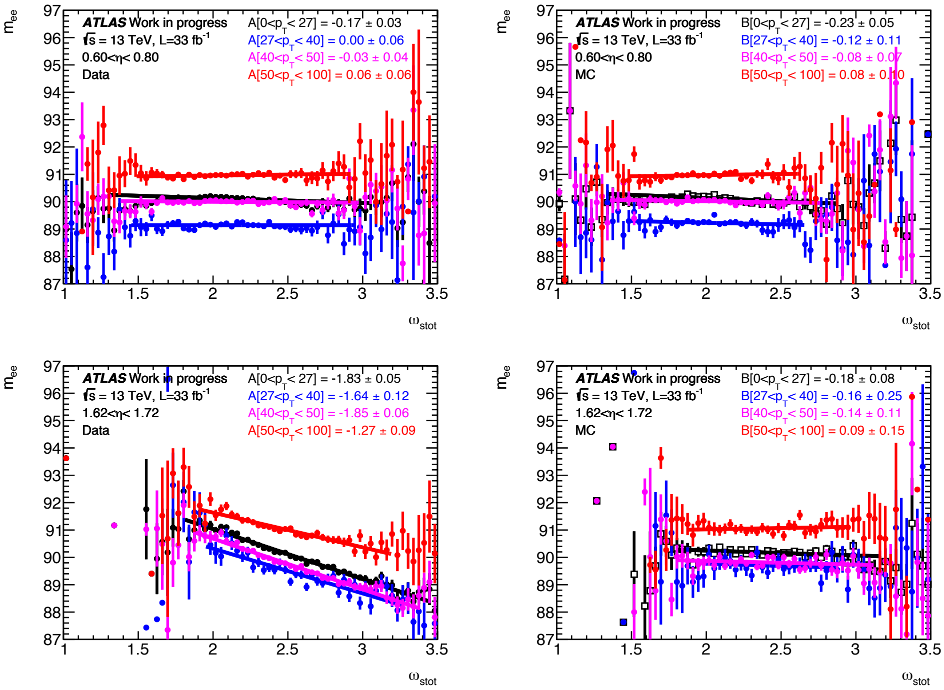


Figure 4.1.12: Dependence of invariant mass distribution for data and MC as a function lateral shower width (ω_{stot}) for different energy categories for Data and MC.

The estimation of shower shape related systematic uncertainties is performed using the definition of Equation 4.2 and parametrising the dependency of the lateral shower width with the transverse energy of the particle for data and MC. For electrons, the dependency $\omega_{\mathrm{stot}}(p_T)$ in data is compared with MC from $Z \to ee$. Since this process has limited kinematic region, the shower width for MC events passing dielectron selection

---

[1] Require at least 10 successive points with relative statistical uncertainty less then 5 permille.

from Z sample is compared to single electron simulation [1], and found to be in good agreement in overlapping region. Therefore, the function to describe the behaviour over the full p_T spectrum was optimised on single electron MC simulation:

$$\omega_{\text{stot}}(E_{\text{T}}) = a + \frac{b}{\sqrt{E_{\text{T}}}} \tag{4.3}$$

The parametrisation for electrons from data is obtained by a fit to the spectrum in the range [27-200] GeV, and interpolating over the full spectrum [10 - 1000] GeV as presented on Figure 4.1.13. The MC from $Z \to ee$ events is fitted in the same region ([27-200] GeV) and extrapolated to full range ([10 - 1000] GeV) to be compared with single electron MC, which is obtained by fit in range [10-1000] GeV. In the most bins in pseudorapidity the two MC parametrisation are in good agreement, nevertheless, the parametrisation obtained from fit of the single particle simulation is used in estimation of the systematic uncertainties.

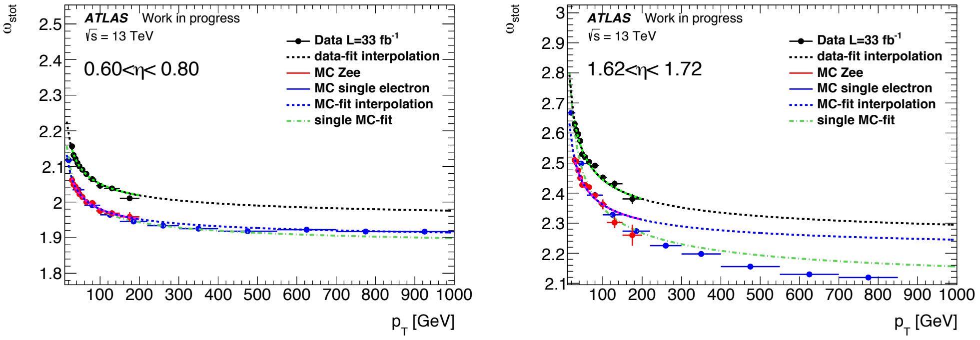


Figure 4.1.13: Dependence of lateral shower width (ω_{stot}) as a function of transverse energy for Data and MC from $Z \to ee$ process and MC from single electrons generation. The parametrisation obtained with a function from Equation 4.3. The data are fitted in the range [27-200] GeV and extrapolated to the full p_T spectrum. The single electron MC is fitted over full range and compared to MC from Z fitted in same range as the data.

A similar procedure was done for photons. The data sample was obtained from the high-mass diphoton analysis[2] , requiring pass preselection, Isolation, ID cuts and requiring the diphoton invariant mass to be larger then 100 GeV[3]. The single photon MC sample[4] is used to estimate dependency of lateral shower width with transverse energy. The same function (Equation 4.3) is used for converted and unconverted photons. The data fitted in the range from 10 GeV to last bin with non-zero error and extrapolated to full p_T spectrum, and single photon MC fitted over full range [10-1000] GeV are presented on Figure 4.1.14. For some η bins, the behaviour of the shower

---

[1] mc15_13TeV.423000.ParticleGun_single_electron_egammaET.merge.AOD.e3566_s2726_r7728_r7676

[2] `MxAOD` samples h014 for full 2016 dataset.

[3] HGamEventInfoAuxDyn.isPassedPreselection, HGamEventInfoAuxDyn.isPassedIsolation, HGamEventInfoAuxDyn.isPassedPID, HGamEventInfoAuxDyn.m_yy>100 GeV.

[4] mc15_13TeV.423001.ParticleGun_single_photon_egammaET.merge.AOD.e3566_s2726_r7728_r7676

width at high energy for unconverted photons has unphysical growth due to effect from cross-talk between Layer 2 and Layer 1 of the EM calorimeter.

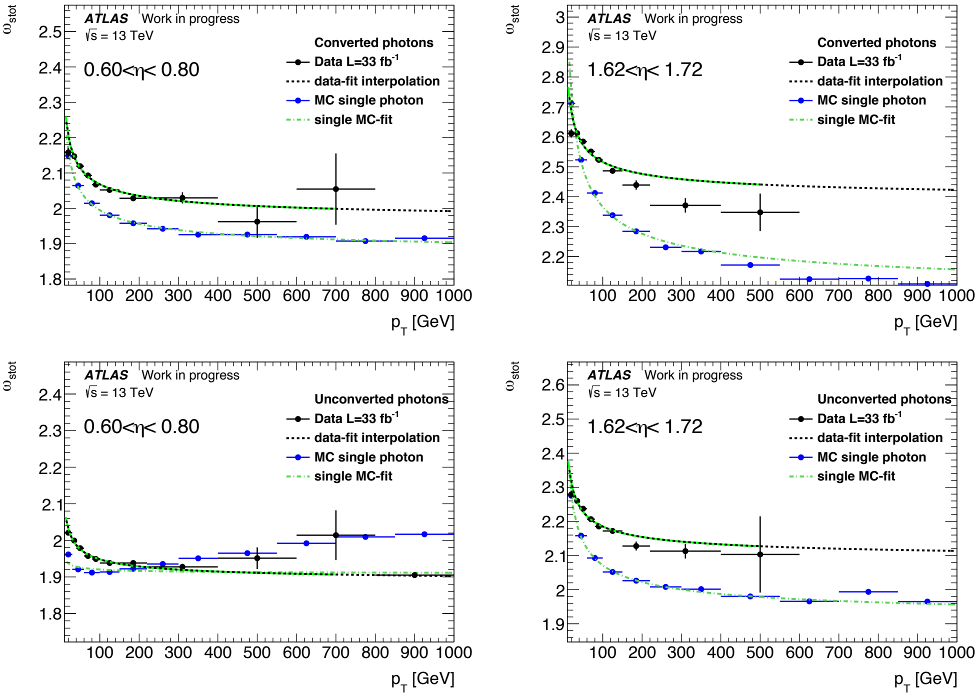


Figure 4.1.14: Dependence of the lateral shower width (ω_{stot}) as a function of the transverse energy for Data and MC for converted (top) and unconverted (bottom) photons. The parametrisation is obtained with a function from Equation 4.3. The data are fitted in the range from 10 GeV to last bin with non-zero error and extrapolated to full p_T spectrum. The single photon MC is fitted over full range.

The systematic uncertainty for electrons, converted and unconverted photons is summarised on Figure 4.1.15. Largest contribution of this uncertainty corresponds to "bad" region η = [1.52 - 1.82]:

- up to 1% for electrons with p_T >500 GeV
- up to 1.5% for p_T >400 GeV for unconverted photons
- within 0.5% everywhere for converted photons

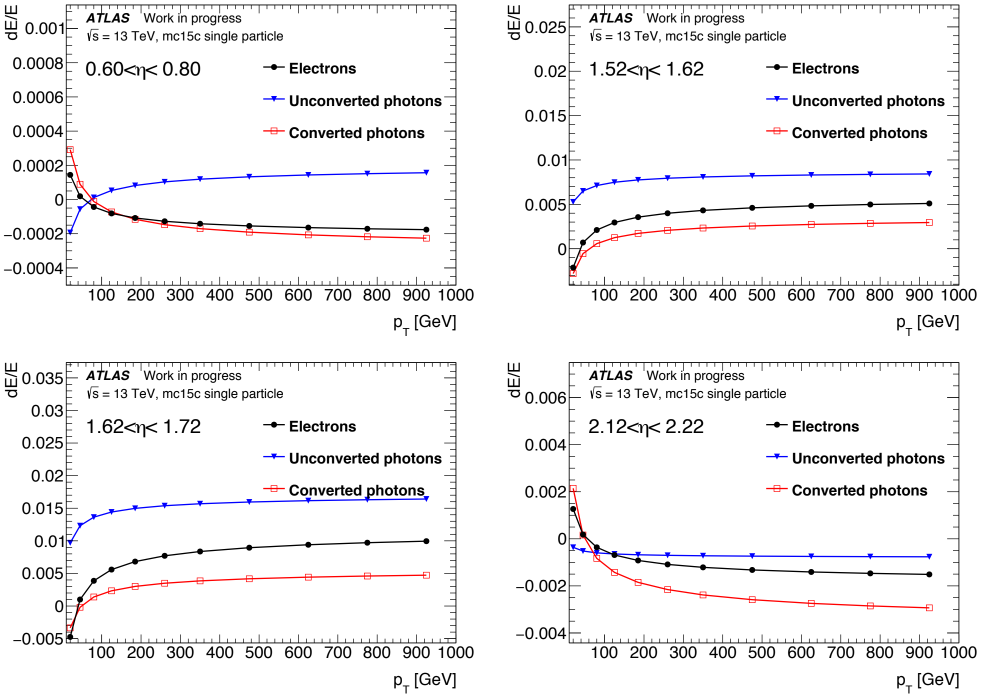


Figure 4.1.15: The systematic uncertainty related to the lateral shower properties. Contribution for electrons (black), converted (red) and unconverted(blue) photons presented.

4.2 Calibration including shower width

In this Section, the procedure to include the lateral shower shape development into the calibration process will be discussed. As shown in the previous Section 4.1.3, a dependence of the reconstructed energy with the shape of the shower is observed, and the response is different between data and simulation. The training procedure and its performance will be presented in Section 4.2.1. The impact on the new calibration on the energy reconstruction of electrons from the $Z \to ee$ events will be discussed in Section 4.2.2.

4.2.1 Training procedure

The procedure of the training of the calibration weights is discussed in details in Section 3.4.1. For the case of the electrons there are nine variables which are used as inputs to the MVA regression: $\eta_{\mathrm{cluster}}, \phi_{\mathrm{cluster}}, \eta_{\mathrm{calo}}, \phi_{\mathrm{calo}}, E^{\mathrm{TileGap}}_{\mathrm{E4scintillators}}$ and E_i, where $i = 0 - 4$ is the index of the calorimeter layer. This set of inputs will be denoted as *"Nominal"* further in the text.

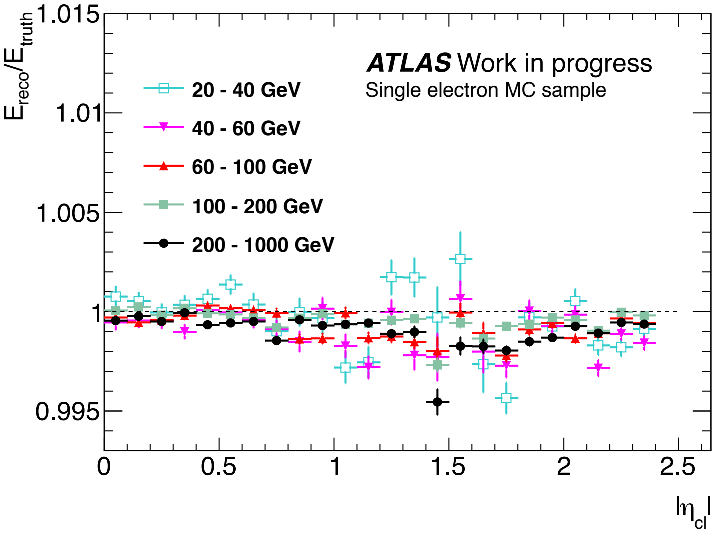


Figure 4.2.1: $E^{\mathrm{reco}}/E^{\mathrm{truth}}$ distribution as a function η_{cluster} for different transverse energies of electrons.

The optimisation of the response is performed in different regions of phase space in bins of $\eta_{\mathrm{cluster}}, E_{\mathrm{T}}^{\mathrm{calo}}$. The binning is chosen to follow the known detector geometry variations and significant changes in the energy response:

- $|\eta_{\mathrm{cluster}}|$:0 - 0.05 - 0.65 - 0.8 - 1.0 - 1.2 - 1.37 - 1.52 - 1.55 - 1.74 - 1.82 - 2.0 - 2.2 - 2.47.
- $E_{\mathrm{T}}^{\mathrm{calo}}$:0 - 10 - 20 - 40 - 60 - 80 - 120 - 500 - 1000 and 5000 GeV.

An independent optimisation is performed for each bin. Multivariate algorithms aim at optimising the energy response and minimising the root mean square (RMS) resolution.

The response of the MVA calibration, $E^{\text{reco}}/E^{\text{truth}}$, as a function of η_{cluster} is illustrated in Figure 4.2.1 for different transverse energies of the electrons.

To address the problem of the dependency of the energy response with the shower shapes, in particular shower width, I introduced in the training additional parameter to the nominal set: ω_{stot}. This configuration will be denoted ad *"Nominal+ω_{stot}"* in the following. The initial problem was observed in dependence of the shower width in "bad" region (1.6<$|\eta|$<1.8). The energy response as a function of ω_{stot} in the barrel and in the endcap is presented on Figure 4.2.3 for two sets of calibration weights. The barrel regions (top two plots) don't show strong dependence with the shower depth, while the "bad"-region and endcap has significant miscalibration in case of wide shower (high ω_{stot} values) for Nominal calibration (in cyan). The improved calibration (in red) removes dependence of energy response for all shower widths.

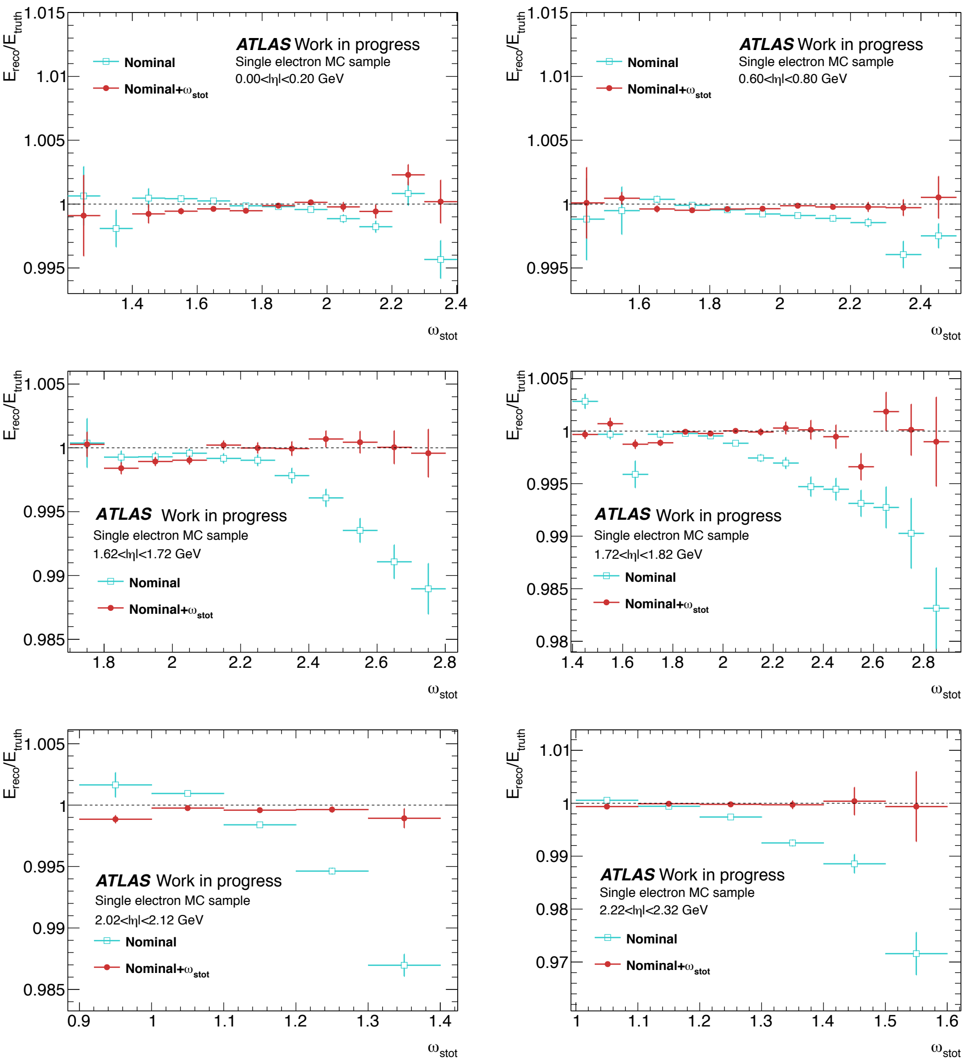


Figure 4.2.2: Comparison of the energy response for two sets of calibration weights: "Nominal" (cyan) and "Nominal+ω_{stot}"(red). $E^{\text{reco}}/E^{\text{truth}}$ distribution as a function ω_{stot} for the different regions of the pseudorapidity.

The comparison of the response of the two calibration sets are presented on Figure 4.2.3, where $E^{\mathrm{reco}}/E^{\mathrm{truth}}$ as a function of η_{cluster} is presented for two kinematic regions, most relevant to electrons from $Z \to ee$ process. In average, the response between two sets is comparable within uncertainties over full region in pseudorapidity. But the impact on dependence with the shower width, should remove the discrepancy observed in Z mass reconstruction, as discussed in Section 4.1.3.

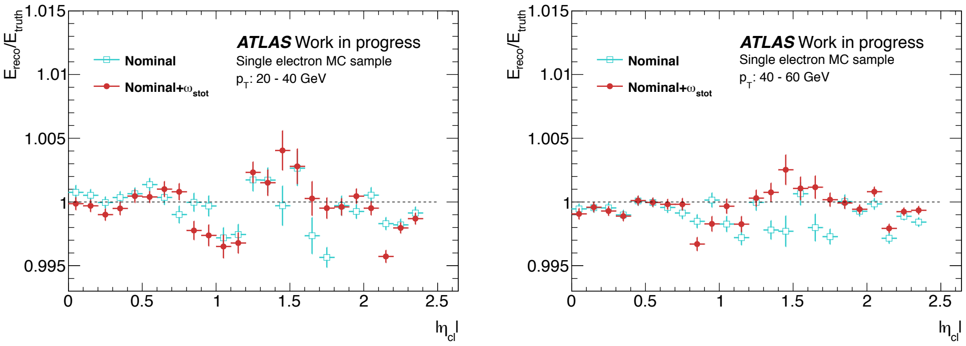


Figure 4.2.3: Comparison of the energy response for two sets of calibration weights: "Nominal" (cyan) and "Nominal+ω_{stot}"(red). $E^{\mathrm{reco}}/E^{\mathrm{truth}}$ distribution as a function η_{cluster} for transverse energies of electrons 20-40 GeV (left) and 40-60 GeV (right).

4.2.2 The energy reconstruction using improved calibration

The impact of the improved calibration (Nominal+ω_{stot}) on the energy response was shown on single particle MC simulation sample, using $E^{\mathrm{reco}}/E^{\mathrm{truth}}$. This procedure will be propagated to the energy response with using the di-electron invariant mass m_{ee} distribution ($Z \to ee$ MC sample), to estimate the difference of invariant mass distribution $\Delta m_{ee}^{Data-MC}$ for the lateral shower width categories and reduce the associated systematic error (derived in Section 4.1.4.2), possibly remove it.

Chapter 5

Search for diphoton resonance

Contents

In this Chapter the search for new resonances decaying into two photons will be discussed. The overview of the analysis will be presented starting with the motivations and the selection of diphoton candidates, followed by the estimation of the sample composition obtained after applying the selection criteria. The modelling of the signal and of the expected backgrounds are presented to test the compatibility of the selected data with the Standard Model expectations. This analysis is the state of the art and is the fruit of the collection of work of many people from different groups. Since this manuscript is the summary of my contribution, the parts where I made a dominant contribution will be revealed in details. To make it clear, the Figures and Tables, which have their caption over grey background correspond to the material coming from the analysis documentation and which was obtained by colleagues from the group. The inputs with captions without background correspond to my personal contributions to this analysis for internal use or for published results. The study of the selection efficiency and its dependence with the pileup is discussed in Section 5.2.2. The description of the signal modelling and the related systematic uncertainties is presented in Section 5.4. The statistical techniques and interpretations of the search are given in the Sections 5.6 and 5.8 correspondingly.

Before entering the description of the analysis, I would like to present the diphoton invariant mass distribution of the selected events which were collected in 2015; it is shown on Figure 5.0.1. Even if those points around 750 GeV are not *statistically* significant, they managed to create significant *resonance* in the physics community. About **200** papers trying to interpret the *bump* were written in the **few months** following the presentation of this result, and reaching 600 after the 2016 dataset did not confirm the excess. The dependence of the number of related papers with time is presented on Figure 5.0.2. The change of the slope is seen around July-August 2016, when the first result with the freshly collected data was presented. I am happy to have been the part of the team of people who pursuit the bump from the beginning and did their best to scrutinise the data, even if in the end it turned to be just a statistical fluctuation.

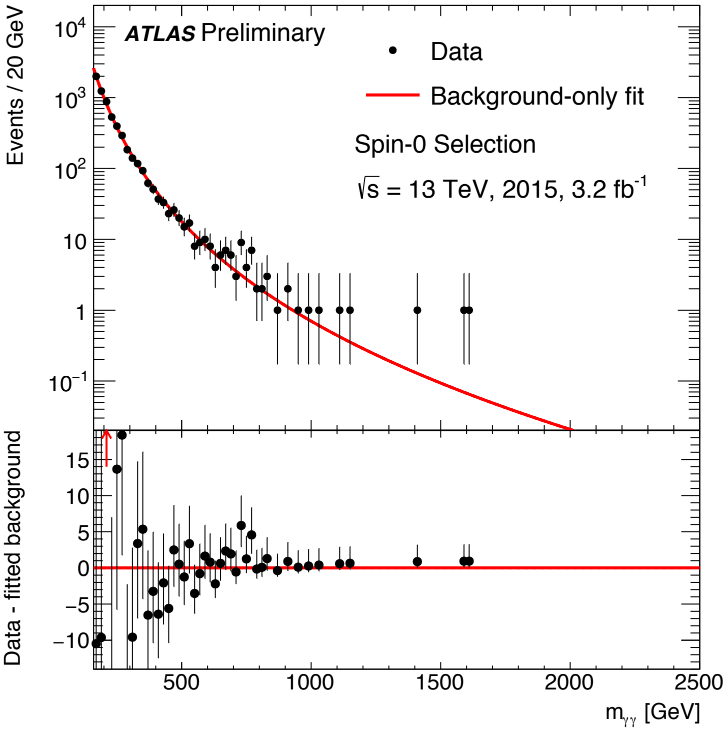


Figure 5.0.1: Invariant-mass distribution of the selected diphoton candidates, with the background-only fit overlaid, for the 2015 dataset. The difference between the data and this fit is shown in the bottom panel. The arrow shown in the lower panel indicates a value outside the range with more than one standard deviation [109].

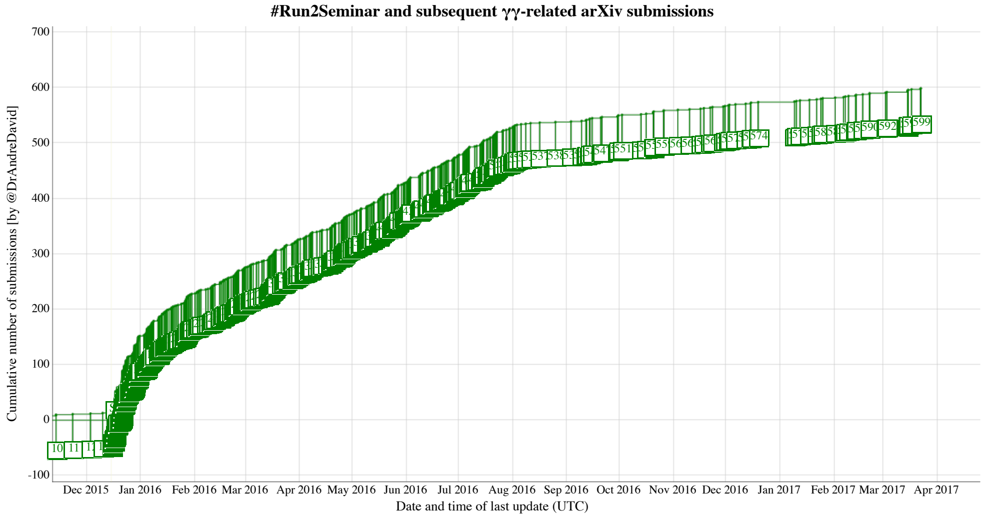


Figure 5.0.2: Cumulative number of arXiv submissions related to the diphoton excess around 750 GeV [104].

5.1 Introduction

The observation of the Higgs-like scalar particle during Run 1 was a successful confirmation of the Standard Model (described in Section 1.1). The diphoton final state provides a clean experimental signature with excellent invariant mass resolution and smoothly falling known backgrounds. It played a crucial role in the discovery of the Higgs boson [105]. The ATLAS collaboration has provided extensive studies of the properties of the new boson. Several models predict the existence of an extended Higgs sector with additional states, and some of them were discussed in Section 1.4. Some models, Beyond the Standard Model (BSM) require a second scalar particle at higher mass to fulfil the unitarity conditions in the WW and ZZ scattering amplitudes at high-energy [117]. The observed $H(125)$ Higgs boson and its measured couplings being compatible with the SM, make strong constrains for couplings to any second boson. The particular scenario of the two-Higgs-doublet extensions of the Standard Model (Section 1.4.1) provides the case where the couplings of the second scalar particle to the W and Z bosons can satisfy this requirement [118]. With these criteria the second resonance is expected to be narrow, and can have sizeable branching fractions in the $\gamma\gamma$ final states.

Searches for high-mass resonances decaying to diphoton final states are highly motivated from the theoretical and the experimental points of view. Such searches in LHC Run-1 data have been reported by the ATLAS and CMS collaborations [110, 111, 112], where no new physics was observed, and upper limits on the production cross section times branching fraction were set.

In this Chapter, a search for new high-mass spin-0 resonances decaying into two photons in the ATLAS experiment at the LHC is described. The analysis is based on the proton-proton collision data collected in 2015 and 2016, corresponding to an integrated luminosity of $36.7\,\text{fb}^{-1}$. The search uses a restricted kinematic range for the photon selection, taking advantage of the isotropic distribution of the decay products in the centre-of-mass frame of the new particle. The signal and background shapes are described by analytical functions. The signal is parametrised in ranges of the mass between 200 and 2400 GeV for different resonance width (Γ_X) hypothesis, from a narrow intrinsic width ($\sim$4MeV) to a width of 10% of the resonance mass m_X; the parametrisation and the corresponding systematic uncertainties are discussed in Section 5.4. The background is estimated by fitting the data distribution with a function; the choice and validation of the function are presented in Section 5.5. The results of the analysis are presented in Section 5.8, in terms of a limit on the production cross-section times branching fraction for the production of a new resonance followed by its decay to $\gamma\gamma$. The cross-section is restricted to a fiducial volume where the detector effects are well understood and the event selection is as model-independent as possible.

5.2 Data Sample and Event Selection

5.2.1 Overview of the diphoton selections

The definitions and procedures of object reconstruction, identification and isolation of photon candidates are described in Sections 3.1, 3.2 and 3.3 correspondingly. Therefore, in this Section the definition will be skipped and only the values of the selection criteria will be mentioned.

Event selection

The analysis requires two photon candidates satisfying *tight* identification criteria based on the EM showers shapes. Both photon candidates are required to be isolated, using a criteria based on the inner tracker and the calorimeter. The track isolation energy in the cone $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\varphi)^2} = 0.2$ around each photon candidate is required to be less than $0.05 \times E_{\mathrm{T}}$ GeV. Only tracks consistent with originating from the diphoton production vertex are used, and the tracks associated to converted photon candidates are excluded. The calorimeter isolation in a cone of radius $\Delta R = 0.4$ is required to be smaller than $0.022 \times E_{\mathrm{T}} + 2.45$ GeV (after subtracting the contributions from the photon itself, and correcting for the leakage of the photon energy and using an event-by-event energy subtraction based on the jet area method).

The precise location of the diphoton production vertex is required for precise computation of track-based quantities. The vertex position along the beam axis is obtained by combining the trajectories of both photons, measured using the longitudinal segmentation of the calorimeter, with constraints from the average beam spot position (*photon pointing* method [119]). For converted photons, the position of the conversion vertex is also used if the conversion tracks have hits in the silicon detectors. The diphoton production vertex is selected among all the reconstructed primary vertices using a Neural Network algorithm. The scalar sum of transverse momenta of the tracks associated to the vertex (p_T) and the spread of the position of the primary vertex ($(z_{\mathrm{com}}$ [1] $-z_{\mathrm{vertex}})/\sigma_{\mathrm{vertex}}$) are used as the inputs for the training procedure. The expected efficiency of finding a primary vertex within 0.3 mm of the true one is on average better than 80%.

The energy of the photon is calibrated using the algorithm described in Section 3.4. It is optimised for the layout of the ATLAS detector used during the 2015 data taking period, including the presence of the Insertable B-Layer module (IBL). The scale adjustment for data is based on $Z \to ee$ sample collected in 2015 and 2016.

Kinematics selection

The photon candidates are required to be within the fiducial calorimeter region of $|\eta| < 2.37$, excluding the transition region at $1.37 < |\eta| < 1.52$ between the barrel and the end-cap calorimeters. In the excluded transition region, the calorimeter granularity

---

[1] z_{com} is the weighted mean of the intersections of the extrapolated photon trajectories given by the calorimeter "pointing" with a constraint from the beam spot position.

is reduced, and the presence of significant additional inactive material degrades the photon identification capabilities and energy resolution.

The decay products of a spin-0 resonance in its center-of-mass reference frame have an isotropic distribution, therefore the transverse energies of the two photons are expected to be higher than those of photons from background processes of the same invariant mass. To reduce the background at high mass, the transverse energy is required to be $E_{\mathrm{T}} > 0.4 m_{\gamma\gamma}$ for the photon with the highest E_{T} and $E_{\mathrm{T}} > 0.3 m_{\gamma\gamma}$ for the photon with the second-highest E_{T}.

The diphoton invariant mass is required to be larger than $m_{\gamma\gamma} > 150$ GeV.

5.2.2 Data Samples

The data used in this analysis were recorded in 2015 and 2016 using *pp* collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV with a minimum bunch spacing of 25 ns. The peak instantaneous luminosity exceeded the LHC designed value of 1×10^{34} and reached $1.37 \times 10^{34} cm^{-2} s^{-1}$.

Events from *pp* collisions were recorded using a diphoton trigger `HLT_g35_loose_g25_loose`. The E_{T} thresholds are 35 GeV and 25 GeV for the leading and sub-leading photons correspondingly. The associated electromagnetic clusters are required to match *loose* criteria according to the properties of showers initiated by photons. The trigger has a signal efficiency larger than $99.9\%^{+0.027}_{-0.26}$(stat) $\pm$ 0.05(syst) for events fulfilling the final event selection [126].

Data Quality requirements are applied, such as requiring that the events were collected during *stable beam* conditions, and that the detector subsystems were fully operational. During the 2016 data-taking, there was period when the Toroid Magnet was switched off; since this system does not impact on the diphoton analysis, data recorded during this period are used. The data-taking efficiency of the ATLAS detector is 92%. The fraction of the full dataset satisfying the data quality requirements for all physics analyses is 92% and the fraction reaches 94% for analysis not vetoing the period when the Toroid Magnet was off. The following Good Run Lists (`GRL`) were applied for the 2015 and 2016 datasets correspondingly:

- `data15_13TeV.periodAllYear_DetStatus-v79-repro20-02_DQDefects-00-02-02_PHYS_StandardGRL_All_Good_25ns.xml`
- `data16_13TeV.periodAllYear_DetStatus-v88-pro20-21_DQDefects-00-02-04_PHYS_StandardGRL_All_Good_25ns_ignore_TOROID_STATUS.xml`

The values of the corresponding integrated luminosities for each dataset and the measurement of the integrated luminosity uncertainty are listed in Table 5.2.1.

The average number of *pp* interactions per bunch crossing, pileup, is presented for 2015, 2016 and the combination of the two datasets on Figure 5.2.1.

Cutflow comparison

One expect that the ratio of selected events between the two datasets should be equal to the ratio of their luminosities. The number of events after selection cut requirements

Year	Integrated Luminosity, fb^{-1}	Uncertainty, %
2015	3.2	2.1
2016	33.5	3.4
$2015 + 2016$	36.7	3.2

Table 5.2.1: Summary table of integrated luminosity per dataset and its uncertainty [122].

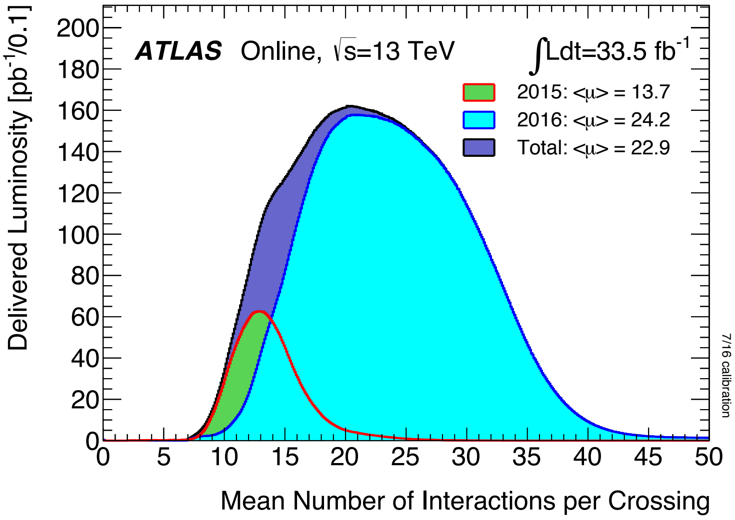


Figure 5.2.1: The luminosity-weighted distribution of the mean number of interactions per crossing ($< \mu >$) for the combined 13 TeV data from 2015 and 2016 [120].

Sample	Spin-0 NWA $m_X = 750$ GeV	Data 2015 3.2 fb^{-1}	Data 2016 33.5 fb^{-1}
Selection	Spin-0 sel. efficiency	Spin-0 sel. event yield	
Trigger	89.53%	8649862	93085472
Detector DQ	89.53%	8649084	93082776
Primary vertex	89.53%	8649050	93082768
2 loose photons	77.96%	2662080	28497832
e-γ ambiguity	77.96%	2662080	28497832
Trigger match	77.83%	2645287	28430104
Tight ID	69.45%	414418	4197665
Isolation	62.15%	124862	1182494
$m_{\gamma\gamma} > 150$ GeV	62.43%	19341	187633
E_T cuts	45.49%	7794	76395

Table 5.2.2: Effect of event selections on a NWA spin-0 MC sample generated for $m_X = 750$ GeV and in the data for 2015 and 2016 datasets. For the MC samples, the efficiency is shown relative to the total event yield after applying event weights. For data, the absolute yields are shown for the spin-0 selections.

(described in the previous section) for MC and data samples of 2015 and 2016 are presented in Table 5.2.2. A discrepancy of yields is observed, and a study of the impact of the cuts, pursued to understand the reason for the difference observed between the two years, is presented on Figure 5.2.2.

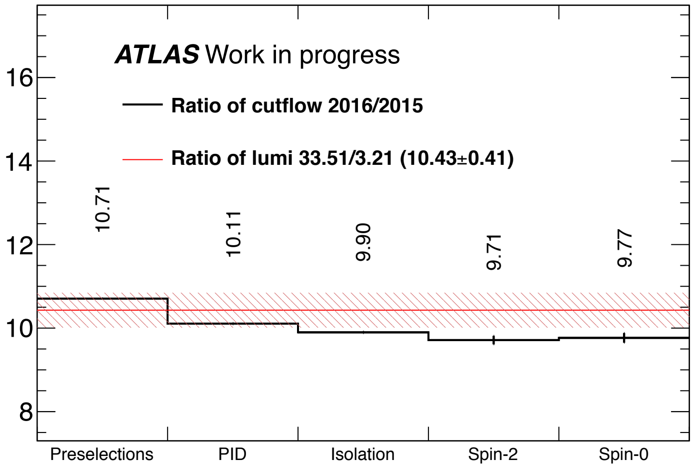


Figure 5.2.2: Distribution of the ratio of selected events for each selection cut: the numerator is the number of selected events in 2016 and the denominator the number of events from the 2015 data. The ratio of luminosities is 10.43 ± 0.41 is presented as the horizontal line with its error represented as the red band.

The comparison shows, that the main difference appears after applying ID and Isolation selections. The main difference in data-taking conditions between the two years is the average number of interactions per bunch crossing (Figure 5.2.1). A dependence of the fraction of selected events for the tight photon identification selection and for the tight isolation with the number of pile-up events (μ) in the event is presented on Figure 5.2.3 on the left for the 2015 dataset and on the right for the 2016 dataset. A clear dependency with μ is observed. The relative loss of selected events depending on the average pileup covers the deviation of the number of events with respect to the ratio of integrated luminosities across the two years. In addition, it has been checked that the dependency of these two selection cuts is also present in SHERPA Monte-Carlo $\gamma\gamma$ events, Figure 5.2.4 .

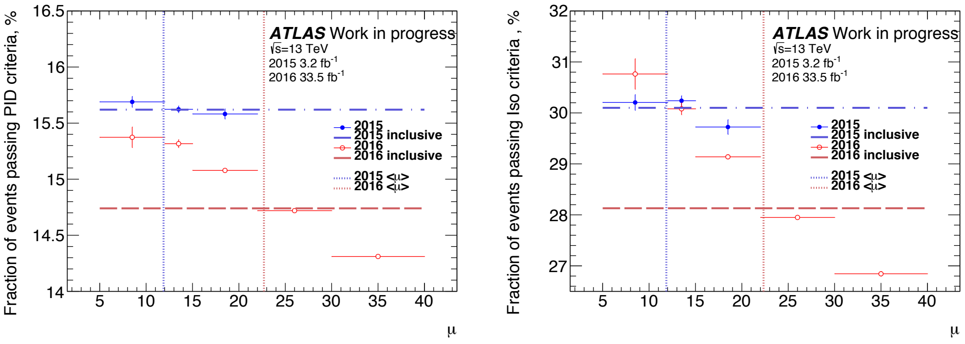


Figure 5.2.3: Distribution of the fraction of events passing the tight photon identification selection cut (left) and for the tight isolation cut (right) as a function of the average number of pile-up events, in blue for 2015 data and in red for 2016 data. The horizontal lines represent the inclusive fractions of selected events and the vertical line represent the average number of pile-up events for each year.

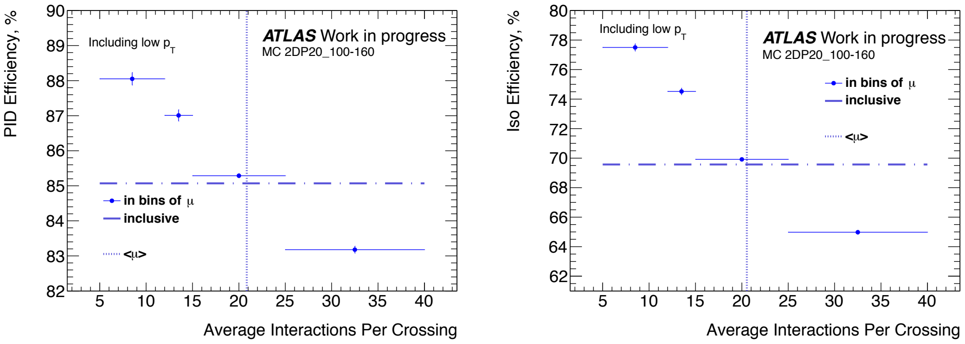


Figure 5.2.4: Distribution of the ratio of selected events for the tight photon identification selection cut (left) and for the tight isolation cut (right) as a function of the average number of pile-up events in SHERPA Monte-Carlo $\gamma\gamma$ events. The horizontal lines represent the inclusive fractions of selected events and the vertical line represent the average number of pile-up events.

The conclusion of this study is that the lower fraction of selected events in 2016 data with respect to 2015 data is explained by the increased average number of pile-up interactions per event: the tight photon identification and the tight isolation cuts are less efficient when the level of pile-up increases (about 4% loss of efficiency for each cut, going from $< \mu_{2015} > \simeq 14$ to $< \mu_{2016} > \simeq 24$).

5.3 Sample Composition and Acceptance

5.3.1 Sample composition

The diphoton spectrum, after applying the selections, consists of events from three contributions in order of fraction:

- Genuine diphoton production, where both photons are prompt photons from signal or background events ($\sim 90\%$). This category is denoted as $\gamma\gamma$.
- Events, where one hadronic jet is misidentified as a photon, "fake photon" ($\sim 10\%$). This contribution is named γj or $j\gamma$ whether the transverse momentum of the reconstructed prompt photon is larger than that of the fake photon.
- Events with two "fake photons" ($> 1\%$). Such events are denoted as jj.

The first category is called "irreducible" background, and the last two - "reducible". Isolated electrons from Drell-Yan, $W\gamma$ or $Z\gamma$ production misidentified as photon has a negligible contribution. A possible resonant signal is considered indistinguishable from the irreducible background, and is included to the first category as the SM prompt diphoton events. The determination of the background composition provides a direct information that the excess, if any, does not come from fluctuation of the reducible components. The relative fraction and shapes of the reducible components are presented on Figure 5.3.1. The fractions are then used for the background modelling, discussed in Section 5.5.

The method to estimate the background decomposition is based on the use of control regions (CR), where one or two reconstructed photons fail the Tight identification and/or the isolation requirements. The events in CRs correspond to the fake photon component, therefore the information on the hadronic background can be derived and propagated to the signal region. In the signal region, where both photons pass the Tight identification (T) and isolation criteria (I) - TITI , the contribution of fake rate is subtracted. Three independent methods based on the data-driven background measurement are presented to perform the measurement of the di-photon purity of events selected in a search. The detailed description and comparison of the methods can be found in [127]. The purities are measured either inclusively or as a function of the di-photon invariant mass and other kinematic variables. The results of the different methods are in agreement within the uncertainty. The inclusive purity results is $(91^{+3}_{-8})\%$.

5.3.2 Signal acceptance and efficiency

The signal yield can be expressed as the product of three terms: the production cross section times branching ratio to two photons, the acceptance (A) of the kinematic requirements, and the reconstruction and identification efficiency (C).

- The acceptance is expressed as the fraction of events satisfying the fiducial ac-

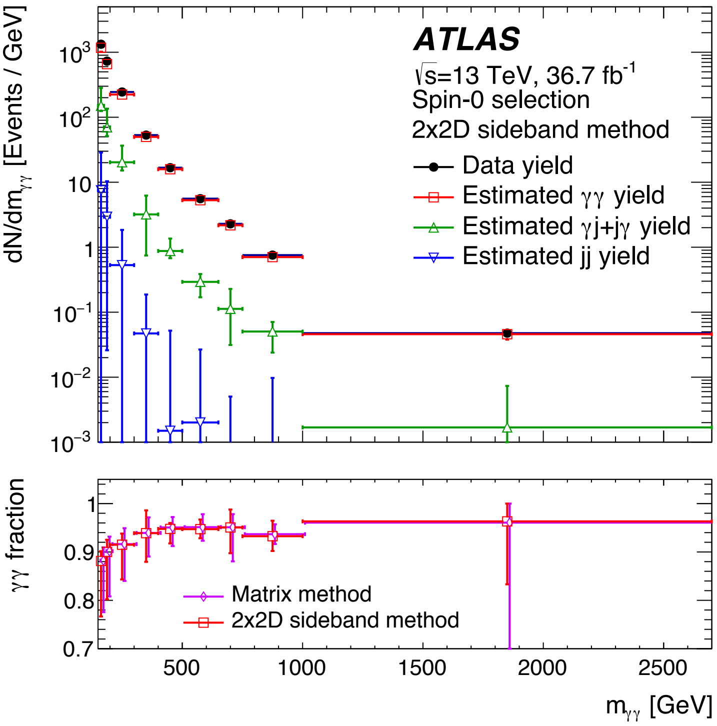


Figure 5.3.1: The diphoton invariant mass distribution (upper panel) of the data and its decomposition in contributions from genuine diphoton, photon+jet and dijet events as determined using of the methods, the 2x2D sidebands method. The bottom panel shows the purity of the diphoton events as determined from the matrix and the 2x2D sidebands methods. The total uncertainties are shown, including statistical and systematic components [116].

ceptance at the generator level:

$$A_X = \frac{N_{\text{fiducial}}}{N_{\text{total}}}, \tag{5.1}$$

where N_{total} is the total number of events corresponding to given values of m_X and α_X=Γ_X/m_X and N_{fiducial} is the number of signal events at the particle-level generated within the fiducial volume.

- The factor C is defined as the ratio of the number of events fulfilling all the selections placed on reconstructed quantities to the number of events in the fiducial acceptance:

  $$C_X = \frac{N_{\text{selection}}}{N_{\text{fiducial}}}, \tag{5.2}$$

  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 particle-level includes all generated particles with a mean lifetime of at least 10 ps.

In order to minimize the model-dependence of the result, the fiducial acceptance closely follows the selection criteria applied to the reconstructed data: $|\eta| < 2.37$ (the transition region between the barrel and the end-cap calorimeters $1.37 < |\eta| < 1.52$ is **not** removed); $E_\text{T}^{\gamma_1} > 0.4\ m_{\gamma\gamma}$and $E_\text{T}^{\gamma_2} > 0.3\ m_{\gamma\gamma}$. Photons reconstructed in events where the resonance is produced in association with many high p_T jets (such as ZH, WH and $t\bar{t}H$ events) have a larger calorimeter isolation energy. Therefore, the isolation requirement E_T^{iso}< $0.05E_\text{T}^{\gamma}$+ 6 GeV is applied, where E_T^{iso} is computed using all particles with lifetime greater than 10 ps at the generator level in a cone $\Delta R = 0.4$ around the photon direction. The value of the isolation requirement applied at the particle level is adjusted to mimic the ones used at the reconstruction level.

The C_X factor is computed on samples produced by gluon fusion in the mass range m_X = 200 GeV to 2 TeV and presented on Figure 5.3.2 [123]. To cover continuous description over m_X the mass-dependence is fitted using the functional form $C_X(m_X) = a_0 + a_1 \exp(a_2 m_X)$, with the values $a_0 = 0.749$, $a_1 = -0.220$, and $a_2 = -3.66 \times 10^{-3}$. The acceptance factor A ranges from 35% to 45% in the mass range from 200 GeV to 700 GeVfor a particle produced by gluon fusion and is almost constant above 700 GeV.

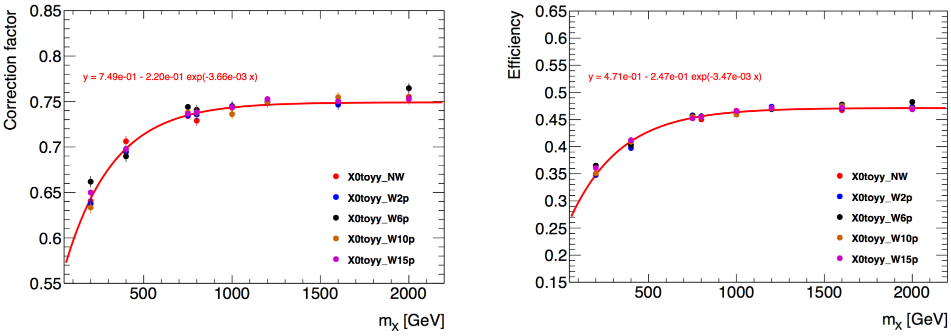


Figure 5.3.2: The estimation of Correction factor C_X (left), and acceptance A_X (right), computed for the gluon fusion Higgs production modes, as a function of m_X [123].

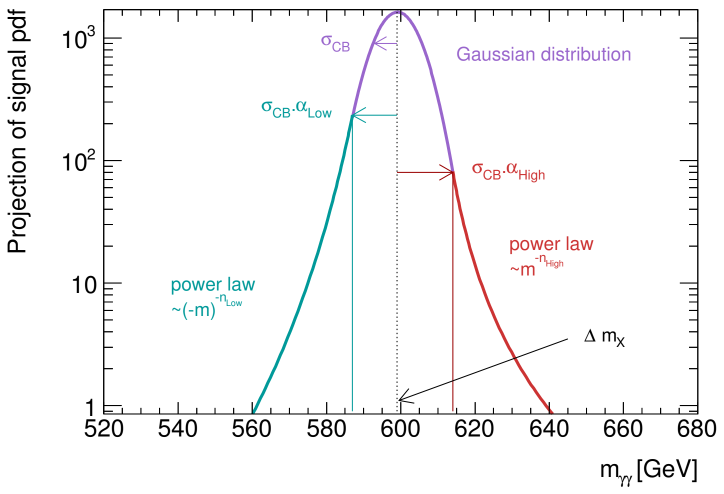


Figure 5.4.1: Description of the double-sided Crystal Ball function parameters, for a signal mass $m_X = 600$ GeV. The different parameters are described in the text.

5.4 Signal Modeling

The invariant mass of the diphoton pair for the signal is expected to peak near the assumed mass of the new particle, with a spread given by the convolution of its intrinsic decay width with the experimental resolution. In this section the description of the signal model, its validation with simulated samples and the related systematics uncertainties will be discussed.

5.4.1 Detector resolution

The signal samples simulated with the gluon-fusion (ggF) production mode are used for the parametriaation of the detector resolution. The width Γ_X of the Higgs-like resonance is set to Γ_X= 4.07 MeV(corresponding to the width of a 125 GeV SM Higgs boson), defined as the Narrow Width Approximation (NWA) - the intrinsic width is negligible with respect to experimental resolution. In order to perform a model-independent search, the properties of the signal shape are also studied for other Higgs-like production modes: vector boson fusion (VBF), associated production with a vector boson (WH, ZH) and associated production with a top quark pair ($t\bar{t}H$). The shapes of the detector resolution for different production modes are not significantly different and the impact of the production mode on the number of fitted signal will be discussed in Section 5.4.2.

The experimental resolution is modelled with a double-sided Crystal Ball (DSCB)

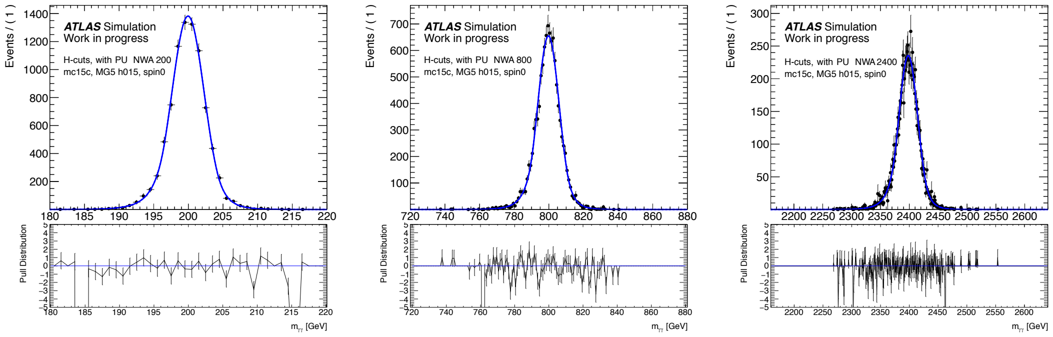


Figure 5.4.2: Fit of the $m_{\gamma\gamma}$ distributions for ggF samples at $m_X = 200$ (left), 800 (middle) and 2400 GeV(right) to a double-sided Crystal Ball function. The bottom insets show the pulls in each bin (the difference between the data point and the fit, divided by the statistical uncertainty in each bin).

function.

$$N \cdot \begin{cases} e^{-t^2/2} & \text{if } -\alpha_{\text{low}} \leq t \leq \alpha_{\text{high}} \\ \dfrac{e^{-0.5\alpha_{\text{low}}^2}}{\left[\frac{\alpha_{\text{low}}}{n_{\text{low}}}\left(\frac{n_{\text{low}}}{\alpha_{\text{low}}} - \alpha_{\text{low}} - t\right)\right]^{n_{\text{low}}}} & \text{if } t < -\alpha_{\text{low}} \\ \dfrac{e^{-0.5\alpha_{\text{high}}^2}}{\left[\frac{\alpha_{\text{high}}}{n_{\text{high}}}\left(\frac{n_{\text{high}}}{\alpha_{\text{high}}} - \alpha_{\text{high}} + t\right)\right]^{n_{\text{high}}}} & \text{if } t > \alpha_{\text{high}}, \end{cases} \tag{5.3}$$

where $t = (m_{\gamma\gamma} - \mu_{CB})/\sigma_{CB}$ with μ_{CB} the peak of the Gaussian distribution and σ_{CB} represents the width of the Gaussian part of the function; N is a normalization parameter; α_{low} (α_{high}) is the position of the junction between the Gaussian and power law 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 peak of the Gaussian and the reference mass value. An illustrative drawing of the double-sided Crystal Ball function is provided in Figure 5.4.1

The parameters of the DSCB function are expressed as analytical functions of the mass of the resonance, providing continuous description of detector resolution over the search range in mass m_X.

The parameterisation is determined in two steps. In a first step, a fit of the $m_{\gamma\gamma}$ distribution of all the events passing the selection cuts is performed using the DSCB function, separately for each mass point m_X (*single mass* fits), yielding a set of DSCB parameters at each point. Analytical functions are selected to describe the behaviour of these DSCB parameters with m_X. The evolution of the width of the Gaussian part (σ_{CB}) is well characterised by a linear function. The difference between the peak of the Gaussian and the reference mass value (Δm_X) and the positions of the junction between the Gaussian and power law ($\alpha_{\text{low/high}}$) are represented with quadratic functions. The exponents of power law ($n_{\text{low/high}}$) are set to constants, because no particular trend is visible with m_X in the single mass point fits. The functional forms are listed in Table 5.4.1 , where a, b and c are free parameters in the fit. Examples of single fit of Narrow Width sample are presented at Figure 5.4.2 for mass points 200, 800 and 2400 GeV .

Parameter	Parameterization	a	b	c
Δm_X	$a + bm_{nX} + cm_{nX}^2$	-0.0151	-0.048	$2.94\,10^{-4}$
σ_{CB}	$a + bm_{nX}$	1.699	0.644	
α_{low}	$a + bm_{nX} + cm_{nX}^2$	1.477	$-7.15\,10^{-3}$	$-8.80\,10^{-5}$
n_{low}	a	12.1		
α_{high}	$a + bm_{nX} + cm_{nX}^2$	1.903	$-2.33\,10^{-3}$	$9.23\,10^{-4}$
n_{high}	a	11.6		

Table 5.4.1: Parameterizations of the DSCB parameters describing the scalar signal shape, as a function of $m_{nX} = \frac{m_X - 100\text{ GeV}}{100\text{ GeV}}$. The results are extracted from the multiple mass point fit.

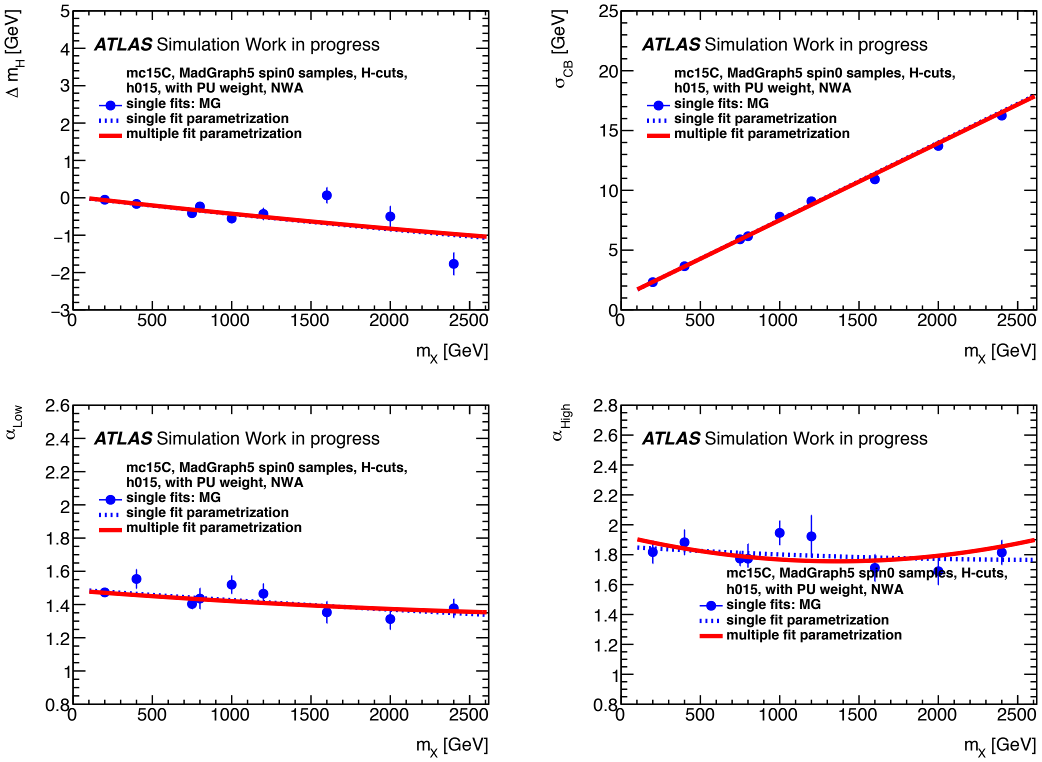


Figure 5.4.3: Comparison of the multiple mass fit parameterisation (red line) to the output parameters of the single mass point fits (blue points) and to the parameterisation of the single mass point fit parameters (blue dashed line) for the scalar model. The parameters are given as a function of m_X for the ggF samples: $\Delta m_X = \mu_{CB} - m_X$ (top left), σ_{CB} (top right), α_{low} (bottom left), α_{high} (bottom right). For Δm_X few points deviate from the fit but that has no effect. The label H-cuts denotes the cuts used to select the spin-0 signal.

In a second step, a simultaneous fit to the signal samples at various masses is performed to obtain the coefficients (a, b, c) of the functional form selected in the first step. The procedure is approximately equivalent to fitting the evolution of the DSCB parameters obtained in the first step, but should be more reliable since the fit is performed in a single step directly from the signal $m_{\gamma\gamma}$ distributions.

Figure 5.4.3 shows the resulting parameterisations of the DSCB parameters from the multiple mass fit (red line). For comparison, the result of the fit of the single-mass parameters to the same functional forms is also shown (blue points fitted with the blue dashed line); in this case, the n_{low} and n_{high} values are fixed to the values obtained from the multiple-mass fit. Very good agreement is found between the two methods.

5.4.2 Systematic uncertainties on the mass resolution

Bias on production mode

The signal shape parameters extracted from ggF samples are compared to those from the VBF, WH, ZH and $t\bar{t}H$ production modes. Because of differences in the kinematics of different Higgs production modes, the photon energy resolution is different in each of the signal samples for the same mass point, and hence the $m_{\gamma\gamma}$ mass distribution. Figure 5.4.4 compares the signal parameterisation functions extracted from a multiple mass-point fit of each of the five signal samples. Significant variations can be observed in the tails. Since this analysis makes no assumption about the production mode of the resonance being searched for, the differences with respect to the ggF parametrisation is included as a systematic uncertainty.

The model-dependence uncertainty is estimated by measuring the bias in the number of fitted signal events induced by using the parametrisation obtained on different production mechanism. To do so, toy Monte Carlo datasets are generated from the $t\bar{t}H$ parametrisation plus a background function $f(x; p_1, p_2) = (1 - x^{1/3})^{p_1} x^{-p_2}$, where $x = \frac{m_{\gamma\gamma}}{\sqrt{s}}$, for each mass point in range $200 < m_X < 3000$ GeV. Each dataset is then fitted with the ggF DSCB function plus a background function.

The DBCB resolution is allowed to vary within $\epsilon = \pm 40\%$ through a constrained nuisance parameter θ, to account for the photon energy scale uncertainty. Figure 5.4.5 shows the evolution of the the bias on the number of fitted signal events. The effect is always smaller then 1%, and this uncertainty is thus neglected.

Photon energy resolution and scale variations

The resolution of the reconstructed photon energy is parameterised as [95]:

$$\frac{\sigma(E)}{E} = \frac{A}{\sqrt{E}} \oplus \frac{B}{E} \oplus C \simeq \frac{A}{\sqrt{E}} \oplus C \tag{5.4}$$

where A, B and C refer to the sampling, noise and constant terms respectively and were discussed in details in Section 2.3.3. The noise term is negligible for photons in the energy range considered for this analysis.

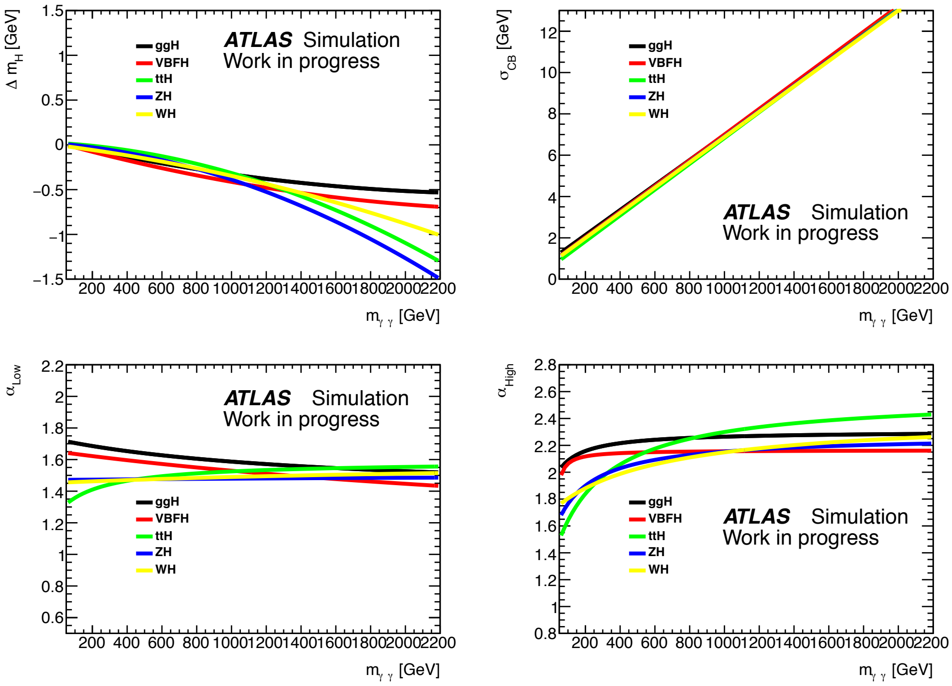


Figure 5.4.4: Multiple mass point fits of the five different Higgs processes as a function of m_X.

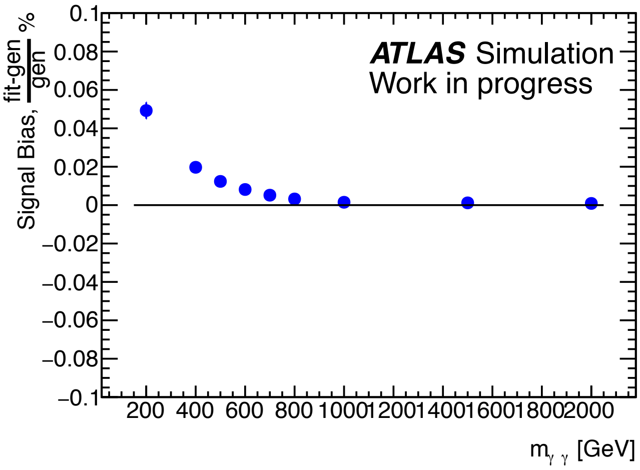


Figure 5.4.5: Bias on the number of fitted signal events induced by fitting the ggF parametrisation on an toy MC dataset generated with the $t\bar{t}H$ parametrisation

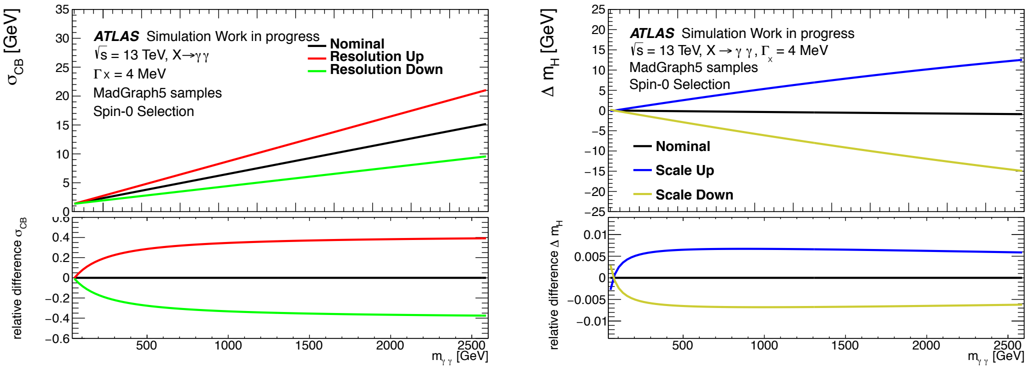


Figure 5.4.6: Variation of the σ_{CB} parameter due to up/down shift of the calorimeter resolution parameters, as a function of m_X, (left) and the energy scale variations (right)for the nominal NW scalar samples . The bottom insets show the relative difference between the shifted samples and the nominal sample.

A single set of variations (`AllUp` and `AllDown`) is used to modify the smearing procedure, associated to a single nuisance parameter in the model. The modified smearing function is then applied to the signal MC samples. The resulting $m_{\gamma\gamma}$ distributions are then fitted using the multiple mass point fit method to extract new signal parameterisation parameters for both the up and down variations.

The change in the resolution parameter values mainly impact the σ_{CB} parameter: Figure 5.4.6 (left) shows the fitted σ_{CB} values for the up and down shifts, compared to the nominal value as a function of m_X. The bottom plot shows the relative difference between the shifted samples and the nominal samples. The systematic effect on σ_{CB} is parameterised by interpolating between the nominal and shifted values of σ_{CB} using a nuisance parameter in the model, as described in detail in Section 5.6.

In order to compute the systematic uncertainty associated to the photon energy scale, the energy of the photon candidates selected in the signal MC samples is shifted up and down by the corresponding uncertainties and selecting the simplified correlation model (`1NP_v1`), where all the components are summed in quadrature. The modified $m_{\gamma\gamma}$ distributions are then fitted using the multiple mass point fit method to extract the new signal parametrisation parameters. The impact of scale on peak position as a function of m_X presented on Figure 5.4.6 (right) . Effect of scale variation is about 0.5%, and can be neglected.

5.4.3 Large-width signal shapes

The parametrisation of the signal shapes over mass and width is obtained by the convolution of the detector resolution described in Section 5.4.1 with the theoretical lineshape of the studied resonance. The convolution itself is implemented in `roofit` [133] using the `RooFFTConvPdf` class, which makes uses the fast Fourier transform (FFT) method. The `MG5_aMC@NLO` lineshape is used in the analysis. The Madgraph effective field theory (EFT) signal lineshape consists of the product of a Breit-Wigner distribution for a given mass and width, squared matrix element and the parton luminosity function. The detailed description of the theoretical lineshape is discussed in [125].

The differential cross-section for $gg \to X \to \gamma\gamma$process is proportional to:

$$d\sigma_{EFT} \propto \int d\hat{s}\ \mathcal{L}_{gg}(\hat{s})\ \frac{1}{\hat{s}}|\mathcal{A}|^2, \tag{5.5}$$

where$\hat{s}$ is the usual Mandelstam variable ($\hat{s} = m^2_{\gamma\gamma}$) and $\mathcal{A}$ is the signal amplitude. It can be expressed as:

$$\mathcal{A} = a_{gg\to X}(\hat{s})\Delta(\hat{s}, m_X)a_{X\to\gamma\gamma}(\hat{s}), \tag{5.6}$$

where m_X is the mass of the resonance and $\Delta(\hat{s}, m_X)$ is the propagator. A relativistic Breit-Wigner distribution is used as propagator in `CalcHEP`, it is given by:

$$f_{BW} = \frac{1}{(\hat{s} - m^2_X)^2 + (m_X\Gamma_X)^2}, \tag{5.7}$$

where Γ_X is the width of the resonance. The squared matrix element $|\mathcal{A}|^2$ is calculated at LO using `Feynrules`, `CalcHEP` and `Mathematica`, where the term $|a_{gg\to X}(\hat{s})|^2|a_{X\to\gamma\gamma}(\hat{s})|^2$ is computed to be $\hat{s}^4$:

$$|\mathcal{A}|^2 \propto \hat{s}^4 f_{BW}. \tag{5.8}$$

Use the squared matrix element to represent the differential cross-section:

$$\frac{d\sigma_{EFT} \propto}{d\hat{s}}\ \mathcal{L}_{gg}\ \frac{1}{\hat{s}}\ \hat{s}^4 f_{BW}, \tag{5.9}$$

taking into account $d\hat{s} = 2m_{\gamma\gamma}dm_{\gamma\gamma}$, one obtain:

$$\frac{d\sigma_{EFT}}{dm_{\gamma\gamma}} \propto \mathcal{L}_{gg}\ f_{BW}\ m^7_{\gamma\gamma}. \tag{5.10}$$

The gluon luminosity $\mathcal{L}_{gg}$ function is obtained with the NNPDF 3.0 NLO parton distribution function, see Figure 5.4.8. The APFEL Web is used to plot the luminosities [134, 135].

The resulting lineshape compared to the truth mass distribution of simulated samples shows an excellent agreement as shown in Figure 5.4.8.

The validation of the convolution of the detector resolution with the line shape is performed using simulated samples for particular masses and widths. An example of comparison for m_X = 800 GeV for widths 2, 6 and 10% is presented on Figure 5.4.9. Excellent agreement is found for all the tested samples.

To summarise, the analytical description of the signal model allows to probe 2D parameter space in search for new resonances. Examples of various mass and width hypothesis are presented on Figures 5.4.10 and 5.4.11.

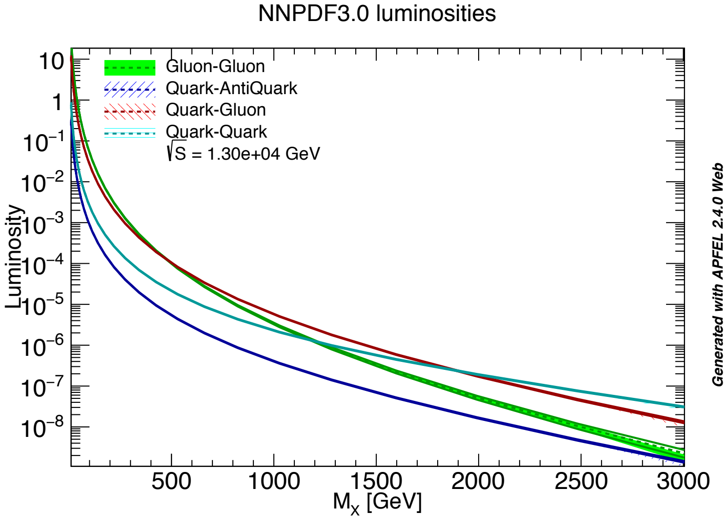


Figure 5.4.7: The parton luminosities as a function of $\sqrt{\hat{s}}$ obtained with the NNPDF 3.0 NLO parton distribution function. The APFEL Web is used to plot the luminosities [134, 135].

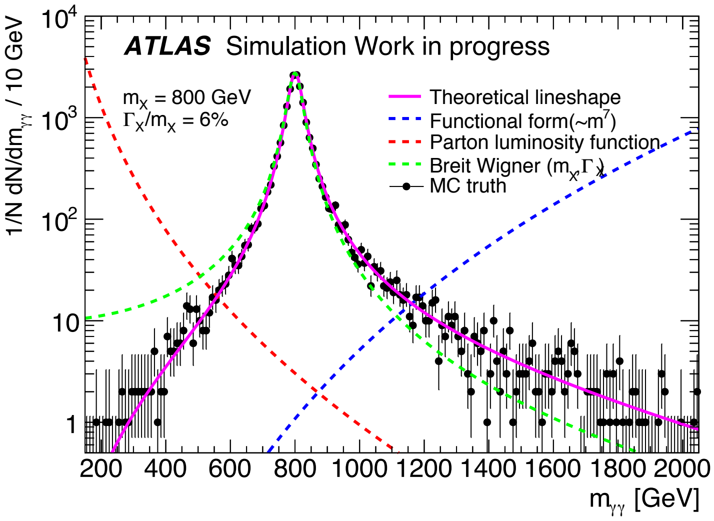


Figure 5.4.8: The true $m_{\gamma\gamma}$ distributions of the resonance with M_X=800 GeV and width of 6% of the m_X value. The dashed red line is the gluon-gluon luminosity, the dashed blue line is the functional form $m^7_{\gamma\gamma}$ and the dashed green line is the Breit-Wigner distribution with 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, shown as the black histogram. No selection cuts have been applied to the true photons.

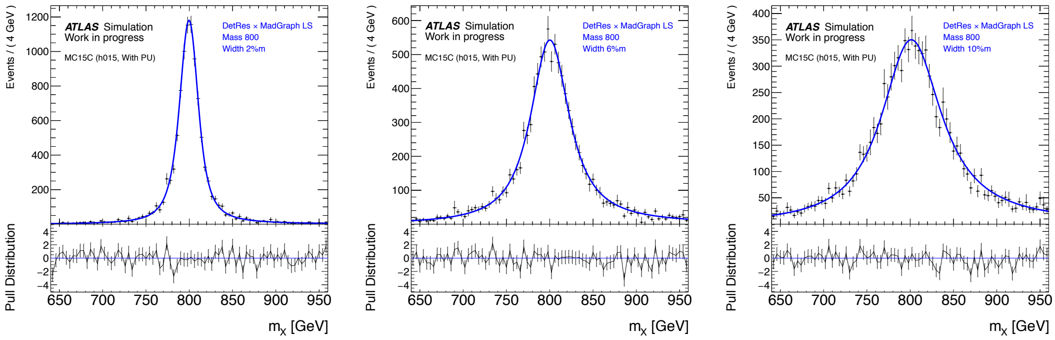


Figure 5.4.9: The reconstructed $m_{\gamma\gamma}$ distributions of the resonance with $m_X =$ 800 GeV and width of (left) 2%, (centre) 6% and (right) 10% of the m_X value. The solid blue lines which shows the convolution of detector resolution and theoretical line shape is overlaid with the reconstructed mass of the `Madgraph` sample passing spin-0 selection.

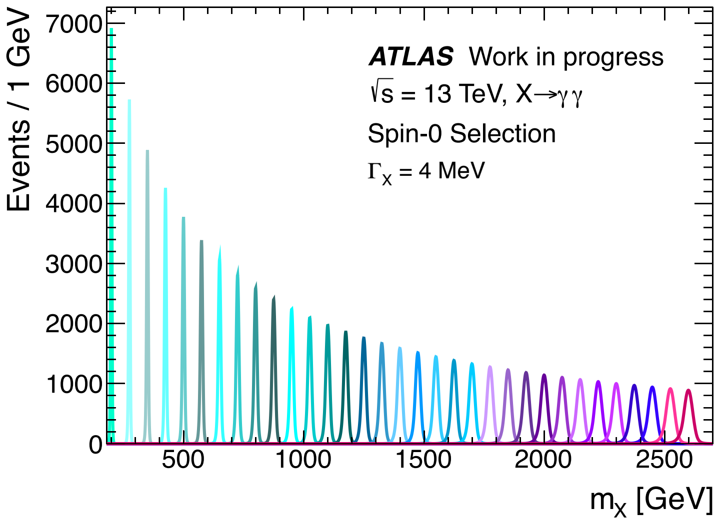


Figure 5.4.10: Analytical description of signal modelling for detector resolution for various mass hypothesis.

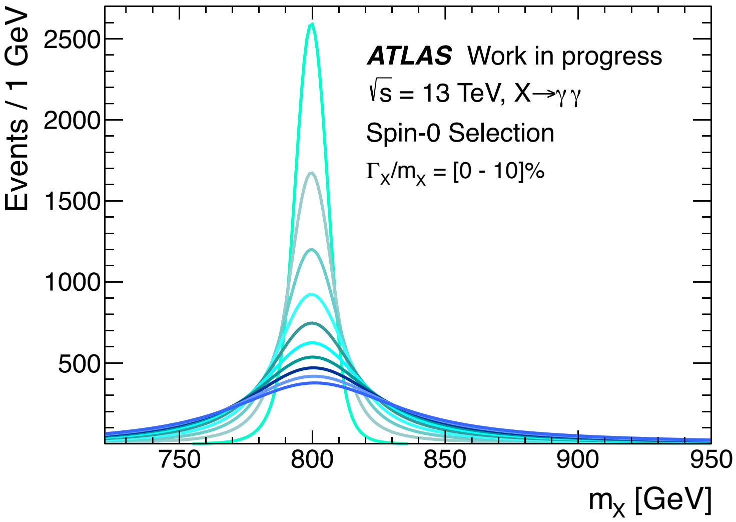


Figure 5.4.11: Analytical description of signal modelling for detector resolution for various width hypothesis [0-10]% for a resonance mass of 800 GeV.

5.5 Background Modeling

The analysis searches for a resonant signal on top of a smooth and rapidly falling spectrum. The shape of the background can be described by an analytical function, and the estimate of the background $m_{\gamma\gamma}$ contribution in the selected sample is based on a fit of the data with this functional form. A family of functional forms, adapted from one of those used by searches for new physics signatures in multi-jet final states [129], is chosen:

$$f_{(k)}(x; b, d, \{a_k\}) = N(1 - x^d)^b x^{\sum_{j=0}^{k} a_j (\log x)^j} , \tag{5.11}$$

where $x = \frac{m_{\gamma\gamma}}{\sqrt{s}}$, d, b and a_i are free parameters ($i = 0 - k$), and N is a normalisation factor. The number of free parameters describing the normalised mass distribution is thus $k + 3$.

This method allows to perform a stable signal-plus-background fit up to the highest $m_{\gamma\gamma}$ values, because the function will be constrained by the data statistics on the lower end of the invariant mass spectrum. The background functional forms are validated for $m_{\gamma\gamma} < 3$ TeV as described in Section 5.5.1.

5.5.1 Validation of the function

To validate the choice of this functional form and to derive the corresponding uncertainties, the so called *Spurious Signal* method is applied to check that the functional form is flexible enough to accommodate different physics-motivated underlying distributions [130]. The idea of the method is to perform a signal-plus-background fit of background-only samples, and the resulting number of signal events from the fit, denoted as $N_{\text{spurious}}(m_X)$ for a given mass hypothesis, is taken as an estimate of the bias in a particular background model under test. A function can be accepted, if it satisfies the requirement that the number of spurious signal is much smaller than the expected signal rate or much smaller than the statistical uncertainty on the number of background events in the fitted signal peak, for cases where the number of expected signal events is very small.

To create a background-only sample we combine the irreducible (diphoton contribution from simulation) and reducible (the photon+jet and jet+photon, and dijet background from data-driven estimates) components, normalised to their yields estimated in the data according to sample composition derived in Section 5.3.1.

Irreducible component

The shape of the mass distribution for the irreducible diphoton background is obtained from the simulated SHERPA diphoton samples. High statistic sample is required, such that the statistic uncertainties of this MC sample are negligible with respect to local modulation of the predicted $m_{\gamma\gamma}$ spectrum.

Reducible component

The reducible background is coming from events in which one or both of the reconstructed photon candidates result from a different physics object, usually a jet, faking the photon signature. The shapes of the mass distribution for the reducible photon-jet, jet-photon and dijet background are estimated from background-enriched data control samples, defined by varying identification selection criteria. The kinematic selection and isolation requirement applied on control samples are the same as for the signal region selections.

The control sample to model the photon-jet component of reducible background is denoted as Tight-AntiTight. It is selected by requiring the leading (most energetic) photon-candidate to pass the photon-ID requirement applied for the signal selection (Tight-ID, Section 3.2). The subleading photon candidate should pass "relaxed" identification criteria, but fail the Tight-ID requirement. The tem *relaxed* means, that in Tight selection several shower shapes variables have a looser cut value (Table 3.2.1); the set of such ID menus denoted as LoosePrime. The cuts are looser for $w_{\eta,1}$, F_{side} and additionally for ΔE, E_{ratio} and $w_{s,\text{tot}}$. The jet-photon sample denoted as AntiTight-Tight, is defined by a selection where the leading candidate pass LoosePrime and fails Tight identification, and the subleading candidate passes Tight-ID. The dijet events selected by requiring both candidates to pass LosePrime and fail Tight identification criteria.

As the limited number of data events does not directly allow a precise estimate of the mass distribution for masses above 500 GeV, the invariant mass distributions of these samples are fitted with various smooth functions providing an adequate fit to the data. The fit use following functional forms:

$$f_{k_1,k_2} = x^{k_1(1-\log(x))} \times x^{k_2 \times \log(x)} \tag{5.12}$$

$$f_{k_1,k_2,k_3,k_4} = x^{k_1(1-\log(x))} \times x^{k_2 \times \log(x)} \times \left(1 - \frac{1}{1+e^{(x-k_3)/k_4}}\right) \tag{5.13}$$

$$f_{k_1,k_2,k_3,k_4,k_5,k_6} = x^{k_1(1-\log(x))} \times x^{k_2 \times 1-\log(x)} \times \left(1 - \frac{1}{1+e^{(x-k_3)/k_4}}\right) \times \left(1 - \frac{1}{1+e^{(x-k_5)/k_6}}\right) \tag{5.14}$$

where the k_i are the free parameters determined by the fit. Unbinned negative log-likelihood (NLL) fits are performed in the mass range [150-3000] GeV. The final choice of the functional form is driven by the maximum likelihood values for the different candidate functions. If the difference of the maximum likelihood value is within 0.5, the function with the lowers number of degrees of freedom is chosen.

The reducible components have limited statistics, therefore lead to bias in fitted mass spectrum. Toy studies performed to estimate it, injecting fitted results and number of events in selected data sample. The background shape obtained on data is compared to shape extracted from toy samples. The statistical uncertainty on the bias is the one obtained using the RMS of all toys.

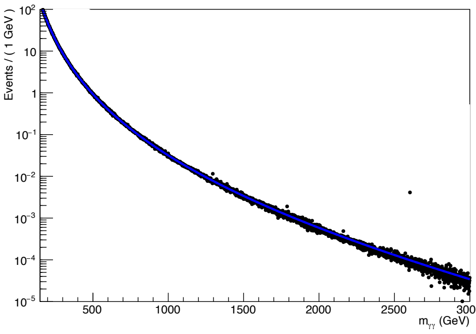


Figure 5.5.1: Background template composed of mixed irreducible ($\gamma\gamma$) and reducible (γj and $j\gamma$) components [123].

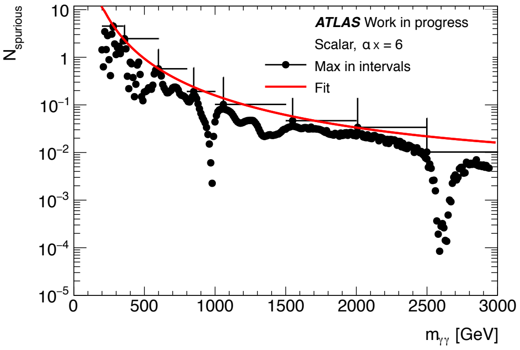


Figure 5.5.2: Result of the spurious signal test for $\alpha_X = 6\%$. The N_{spurious} parametrised as a function of diphoton mass, using maximum value in interval over $m_{\gamma\gamma}$.

Spurious Signal test

The mixed dataset presented on Figure 5.5.1 is used to perform a spurious signal test varying the mass hypothesis in the range of interest in step of one GeV. Different signal width values ($\Gamma_X/m_X = \alpha_X = 0, 0.02, 0.04, 0.06, 0.08$ and 0.10) are considered for the signal-plus-background fit. N_{spurious} is obtained for each m_X value see for instance Figure 5.5.2. To have a continuous coverage over full mass range it is then parametrised with the function:

$$f(x) = N a_0 x^{-a_2} (1 - x^{1/3})^{-a_1} \, , \tag{5.15}$$

where $x = \frac{m_{\gamma\gamma}}{\sqrt{s}}$, N - luminosity normalisation factor. The N_{spurious} distributions are split into several bins with the variable size to get similar amount of the entries within each bin. The maximum of each bin being used for the fit in order to obtain the parameters a_i, $i = 0, 1, 2$. The preparation of the spurious signal study was carried by colleagues from the analysis team, while I contributed to the optimisation of the parametrisation of N_{spurious} as a function of $m_{\gamma\gamma}$ as shown on Figure 5.5.2. For each width hypothesis $N_{\text{spurious}}(m_X, \alpha_X)$ is defined, and a linear interpolation is used to obtain the uncertainty for arbitrary values of α_X in the range $0 < \alpha_X < 0.1$.

The parameterised value of $N^{\text{max}}_{\text{spurious}}(m_X)$ for the chosen function enters the data statistical modelling a spurious nuisance component to be added to fit background and signal functional form. It is described as having a Gaussian distribution with zero mean and a width equal to $N^{\text{max}}_{\text{spurious}}(m_X)$.

A function passes the spurious signal test if in all the mass range of interest it does not generate a value of $N_{max}(m_X)$ larger than a 50% of the expected statistical uncertainty of the background. The selected function satisfying the spurious signal criteria as shown on Figure 5.5.3, with the smallest number of parameters (details can

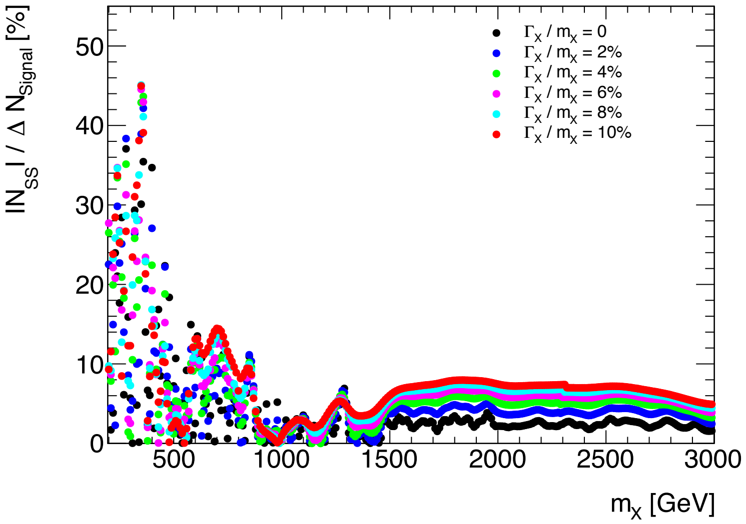


Figure 5.5.3: Result of the test showing relative spurious signal to the background uncertainty performed for different width hypotheses assuming an integrated luminosity of $36.7\,\mathrm{fb}^{-1}$. Several curves corresponding to different signal width are shown [123].

be found in [124]) considered for analysis is:

$$f(x; b, a) = N(1 - x^{1/2})^b x^a. \tag{5.16}$$

where $x = \frac{m_{\gamma\gamma}}{\sqrt{s}}$, a, b are free parameters, and N is a normalization factor.

5.6 Statistical procedure

5.6.1 Statistical model

The diphoton spectrum of the selected events can be described as the sum of signal and background events:

$$N_X(\sigma_X)f_X(m_{\gamma\gamma}) + N_{bkg}f_{bkg}(m_{\gamma\gamma}), \tag{5.17}$$

where N_X is the fitted number of signal events from the new resonance, $f_X(m_{\gamma\gamma})$ is the normalised invariant mass distribution for a given signal hypothesis (Section 5.4), N_{bkg} is the fitted number of the non-resonant background component and $f_{bkg}(m_{\gamma\gamma})$ is the normalized invariant mass distribution of the background events described by functional form (Section 5.5). The fitted number of signal events is related to the assumed signal cross section times branching ratio to two photons (σ_S) in the fiducial acceptance via the integrated luminosity and the detector efficiency correction factors (Section 5.3.2).

N_X and N_{bkg} are obtained from maximum-likelihood fits of the $m_{\gamma\gamma}$ distribution of the selected events. Uncertainties in the signal parameterisation, in the detector efficiency correction factors for the signal and in the description of the background shape are included in the fit via nuisance parameters. Taken this into account, Equation 5.17 can be written in a way:

$$\begin{aligned}\mathcal{L}(m_{\gamma\gamma};\sigma_{fid},m_X,\alpha_X,N_{bkg},\boldsymbol{a},\boldsymbol{\theta}) &= N_X(\sigma_{fid},m_X,\boldsymbol{\theta}_{N_X},\theta_{SS})f_X(m_{\gamma\gamma},\boldsymbol{x}_X(m_X,\alpha_X),\theta_\sigma)\\ &+ N_{bkg}f_{bkg}(m_{\gamma\gamma},\boldsymbol{a})\end{aligned} \tag{5.18}$$

where σ_{fid} is the fiducial production cross-section of the new resonance of mass m_X. The $\boldsymbol{x}_X = \{m_{CB} = m_X + \Delta m, \sigma_{CB}, \alpha_{low}, \alpha_{high}, n_{low}, n_{high}; \alpha_X\}$ is the double-sided Crystal Ball shape description as a function of the resonance mass m_X and the resonance width α_X. The $\boldsymbol{\theta}_{N_X}$ collectively designates the nuisance parameters used to describe the systematic uncertainties, as listed below:

- θ_{lumi}: uncertainty on the integrated luminosity of the data sample;
- $\theta_{eff,X}$, $\theta_{isol,X}$: systematic uncertainties on photon ID and isolation efficiencies for the new resonance;
- θ_{SS}: spurious signal systematic;
- θ_{ER}: photon energy resolution systematics;
- θ_{ES}: photon energy scale systematics.
- θ_{c_X}: production-mode uncertainty on the C_X factor.

The uncertainties related to the signal modelling are constrained with Gaussian or log-normal penalty terms. The nuisance parameters $\boldsymbol{a} = \{a, b\}$ (Equasion 5.16) are the background shape parameters , included without penalty terms. The number of background N_{bkg} is a free parameter in the fit, while the number of signal N_X is parameterized as:

$$N_X(\sigma_{fid}, m_X, \boldsymbol{\theta}_{N_X}, \theta_{SS}) = \sigma_{fid}\,\mathcal{L}\,C_X(m_X) \prod_{n=1}^{|\boldsymbol{\theta}_{N_X}|} K_i(\theta_i) + \sigma_{SS}\theta_{SS} \tag{5.19}$$

where $\mathcal{L}$ is the integrated luminosity of the sample; $C_X(m_X)$ the value of the C_X factor for the considered mass m_X, K_X a function characterizing the effect of the normalization systematics. The second term of the sum corresponds to the background modelling uncertainty ("spurious signal") and its associated nuisance parameter - σ_{SS} and θ_{SS} correspondingly. This "spurious" signal uncertainty is considered separately for each (m_X, α_X) hypothesis without any correlation between the different investigated mass ranges.

The $K_i(\theta_i)$ factors implement each of the systematic uncertainties on the number of signal events listed above. The expression $K_i(\theta_i) = r_i(m_X)^{\theta_i}$ is used, where $r_i(m_X)$ is approximately given by $N_{X,+i}(m_X)/N_X(m_X)$ for $\theta_i > 0$ and $N_X(m_X)/N_{X,-i}(m_X)$ for $\theta_i < 0$, i.e. in both cases the relative variations on N_X for the systematic uncertainty i.This expression ensures that the modifications to the signal event yield for $\theta_i = \pm 1$ are exactly equal to the $\pm 1\sigma$ variations used to define the uncertainties. The actual expression for r_i interpolated smoothly between the cases $\theta_i > 0$ and $\theta_i < 0$ to avoid numerical problems at $\theta_i = 0$. It is implemented using the `RooStats::HistFactory::FlexibleInterpVar` class, modified to allow mass-dependent values for the uncertainties.

5.6.2 Significance

The search for a new resonance considers a signal of unknown mass m_X and width α_X, and the cross-section σ_X, on top of a smooth background distribution described by nuisance parameters ν. The test of the background-only hypothesis against the signal-plus-background hypothesis with specified m, α_X is quantified with the profile likelihood ratio test statistic:

$$q_0(m, \alpha_X) = -2\log\frac{L(\sigma_X = 0, m_X, \alpha_X, \hat{\hat{\nu}})}{L(\hat{\sigma}_X, m_X, \alpha_X, \hat{\nu})} \tag{5.20}$$

where the values of the parameters marked with the hat superscript are chosen to unconditionally (signal-plus-backgound) maximise the likelihood $\mathcal{L}$, while the value with a double hat is chosen to maximize the likelihood in a background-only fit, and ν represents the nuisance parameters which are varied in the fit. The p_0 probability was used to quantify the level of agreement between the observed data and the background-only hypothesis. It was calculated by integrating the normalized distribution of q_0 generated under the background-only hypothesis from the observed value of q_0 to infinity:

$$p_0 = \int_{q_{0,\mathrm{obs}}}^{\infty} f(q_0|0)dq_0 \tag{5.21}$$

where $q_{0,\mathrm{obs}}$ is the test statistic observed in data and $f(q_0|0)$ denotes the pdf of the test statistic. This p_0-value is calculated using the asymptotic approximation to the test statistic distribution [137].

Local significance values Z_0^{local} are computed for each $q_0(m, \alpha_X)$:

$$Z_0 = \Phi^{-1}(1 - p_0(q_0)), \tag{5.22}$$

where Φ^{-1} is the inverse of the cumulative distribution of the standard Gaussian. For a given search, a $Z_{0,\,\mathrm{max}}^{\mathrm{local}}$ value is found for the (m,α_X) hypothesis leading to the largest deviation from the background-only hypothesis, corresponding to a value q_{max}. The distribution of q_{max} does not follow the familiar chi-squared distribution, but has a longer tail.

Fluctuations of data leading to large Z_0^{local} are more common than the corresponding p_0 probability would suggest, since the two additional degrees of freedom from the scan of mass and width provide a larger parameter space for the signal-plus-background to explore. This larger parameter space has a possibility of containing a deeper fit minimum, and hence a lower p_0 value (higher Z_0^{local}). This reduction of the true "global" significance, Z_0^{global}, below $Z_{0,\,\mathrm{max}}^{\mathrm{local}}$, is referred to as the look-elsewhere effect. The look-elsewhere effect can be described in terms of a trial factor, which is the ratio between the probability of observing the excess at some fixed mass point, to the probability of observing it anywhere in the range. The asymptotic distribution in the case of a 1D mass scan [1] was studied in [131]:

$$p_{\mathrm{global}}(c) \approx p_{\mathrm{local}}(c) + \langle N(c_0)\rangle e^{-(c^2-c_0^2)/2}, \tag{5.23}$$

where $N(c_0)$ is the number of 'upcrossings' of the reference level c_0, c - level of observed local significance.

5.6.3 Exclusion limits

The expected and observed 95% confidence level ($C.L.$) exclusion limits on the cross section times branching ratio to two photons are computed using a modified frequentist approach CL_s [138] with the asymptotic approximation to the test statistic distribution [137]. Cross-checks with sampling distributions generated using pseudo-experiments are performed for a few signal mass points. The largest differences are of the order of 10-30% on the cross-section limit for a high-mass, narrow resonance.

5.6.4 Combining two dataset

The 2015 and 2016 datasets are analysed as independent datasets. The combined statistic is used to derive the shapes of the reducible component of the background. The signal model for the two datasets is the same. The same background function is applied to both datasets, while the background normalisation in each dataset is

---

[1] The approach to test n-D parameter space is described in [132].

parameterised by independent free parameters. The luminosity and its uncertainty are separate parameters in the statistical model of each dataset.

The compatibility between the 2015 and 2016 datasets is estimated used a modified version of the statistical model. The cross-section parameters used in the 2015 and 2016 are taken to be different (instead of a single parameter for both the 2015 and 2016 category, as described above). For 2015, a free parameter σ_{2015} is used, while for 2016 the cross-section is expressed as $\sigma_{2016} = \Delta\sigma + \sigma_{2015}$, where $\Delta\sigma$ is a free parameter. A profile likelihood scan is then performed for $\Delta\sigma$. The compatibility significance is then computed as $\sqrt{\lambda(\Delta\sigma = 0)}$, where λ is the profile $-2\log L$ value.

Source	Uncertainty
Signal yield [123]	
Luminosity	±2.1% (2015), ±3.4% (2016)
Trigger	±0.5%
Photon identification	±1.5 – 3%, mass-dependent
Isolation efficiency	±1 – 1.5%
Scalar production process	±2.8%
Photon energy scale	negligible
Photon energy resolution	negligible
Signal modeling, Section 5.4	
Photon energy resolution	$^{+17\%}_{-17\%}$ (at $m_X = 200$ GeV) – $^{+38\%}_{-36\%}$ (at $m_X = 2$ TeV)
Photon energy scale	±0.5%–±0.6%
Production mode	negligible
Background, Section 5.5	
Spurious Signal for 2015	3 – 0.0024 events, NWA
at 200 and 2000 GeV	19.7 – 0.018 events, LWA α= 10%
Spurious Signal for 2016	47.1 – 0.033 events, NWA
at 200 and 2000 GeV	150 – 0.216 events, LWA α= 10%
Spurious Signal for 2015 and 2016	51.5 – 0.036 events, NWA
at 200 and 2000 GeV	164 – 0.236 events, LWA α= 10%

Table 5.7.1: Summary of the main sources of systematic uncertainty for the measurement of σ_{fid} for the Spin-0 analysis. Unless written otherwise, numbers are mass independent. The effect of those systematic uncertainties on the results are shown in Section 5.8.2.

5.7 Systematic uncertainties

Various sources of systematic uncertainties contribute to the analysis. The complete list is summarised in Table 5.7.1. The uncertainties on the luminosity and trigger are mentioned in Section 5.2.2. Other uncertainties related to the signal yield are discussed in detail in Reference [123]. The uncertainties related to the Signal modelling were discussed in Section 5.4.2. The spurious signal estimation is assigned to the background uncertainties, and was presented in Section 5.5.1. The effect of those systematic uncertainties on the results are shown in Section 5.8.2.

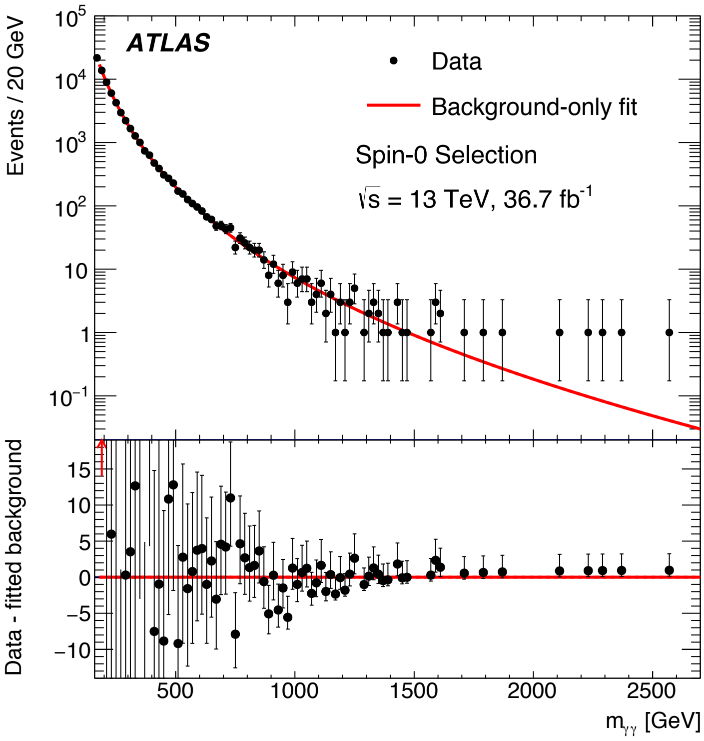


Figure 5.8.1: Distribution of the diphoton invariant mass of the selected events, with the background-only fit. The difference between the data and this fit is shown in the bottom panel. The arrow shown in the lower panel indicates a values outside the range with more than one standard deviation. There is no data event with $m_{\gamma\gamma} > 2700$ GeV [116].

5.8 Results

A search for new spin-0 resonances decaying into two photons have been performed using inputs described above. The data are consistent with the Standard Model background-only hypothesis, and 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. The global significance of these excesses is less than one standard deviation.

5.8.1 Compatibility with background-only hypothesis

Figure 5.8.1 shows the diphoton invariant mass distribution for events passing the Spin-0 selections together with the background-only fit.

The interpretations of the collected data were performed as the compatibility with the background-only hypothesis for different mass and width hypothesis, as shown on Figure 5.8.2. The largest deviation corresponds to a local significance of 2.5σ at 730 GeV preferring narrow width signal hypothesis. The comparison of the local compatibility with the background hypothesis as a function of the assumed mass for different widths (NWA, α_X=2,6,10%) for the 2015 dataset only, the 2016 dataset only and the combined

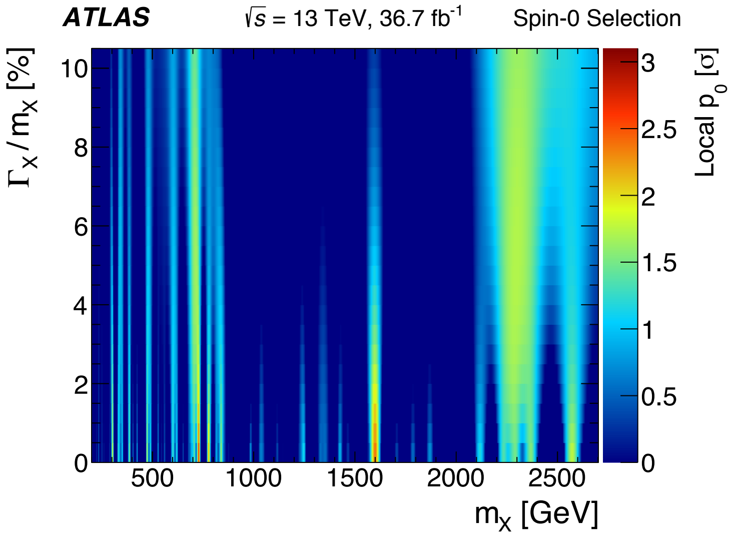


Figure 5.8.2: Observed 2-dimensional scan of the $p-$value for the background-only hypothesis p_0, as a function of the probed resonance mass m_X and relative natural width α_X for 2015 and 2016 combined data [116].

dataset are presented on Figure 5.8.3.

As one can see at Figure 5.8.3 in the 2016 data set (red dashed line), no significant deviation from the background-only hypothesis is observed at the value of the mass corresponding to the most significant excess in 2015 data (blue dashed line), seen by eye on Figure 5.0.1 [109]. The compatibility between the results obtained with 2015 data and 2016 data is estimated to be 2.9 σ following the method described in Section 5.6.4. The profile likelihood ratio is shown in Figure 5.8.4.

The global significance is estimated using formula 5.23 described in Section 5.6.2. Since the maximum observed local significance is rather small (2.5σ), in order to save computational time, it was decided to estimate global significance in simplified way, using only one width hypothesis (where the deviation from the background hypothesis is maximum) - 1D scan over the resonance mass. This means, that if on takes into account the full search phase space (2D-scan over m_X,α_X), the value of the global significance going to be even smaller then in the simplified version. In order to estimate number of uncrossings $< N >$, 80 background-only pseudo-datasets were generated, and for each a scan of $q_0(m)$ for NWA hypothesis was created. The number of uncrossings $< N >$ counted for different tested crossing significances is shown in Figure 5.8.5. The corresponding global significance is smaller than one standard deviation.

An alternative method counting toy experiments in which the local significance exceeds the observed local maximum (2.5σ) in the search range, provide an estimation of the global significance to be ($-0.13^{+0.23}_{-0.24}$). The negative value can be interpreted as the fact that the current dataset expects background fluctuation with higher local significance than observed.

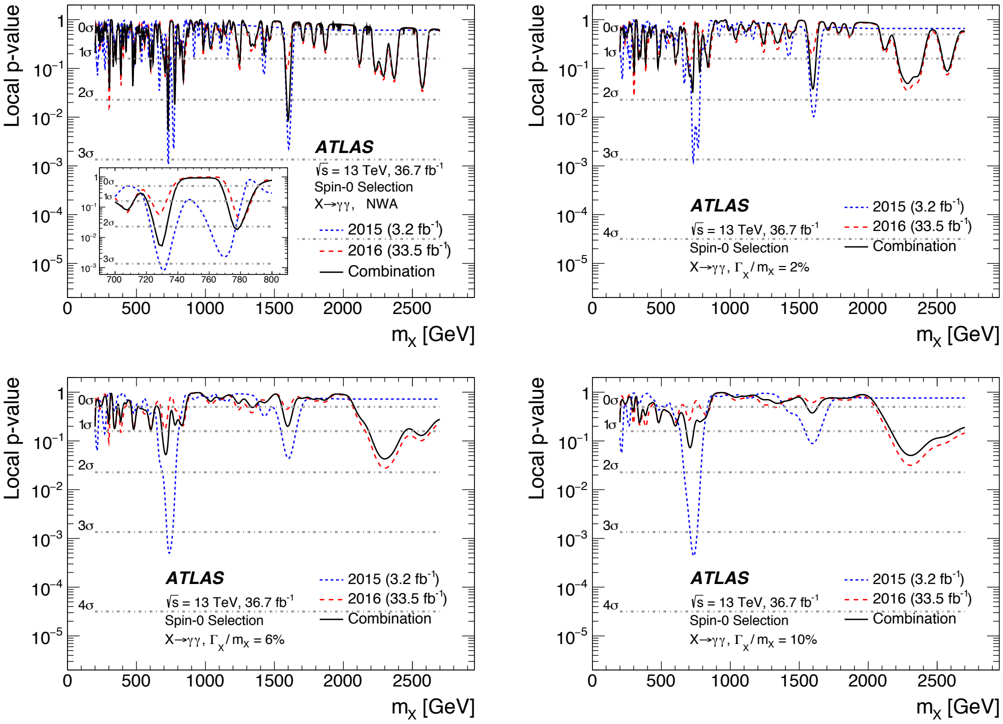


Figure 5.8.3: Observed scan of the p-value for the background-only hypothesis p_0 as a function of the assumed mass, for various values of α_X. Shown are 2015 and 2016 data separately and the 2015 and 2016 combined [116].

5.8.2 Cross-section limits

The data are used to provide 95% CL exclusion limits on the cross section times branching ratio for new scalars decaying to two photons. To be model-independent limits are based on the cross-section within the fiducial acceptance, defined in Section 5.3.2. Figure 5.8.6 shows scans of the expected and observed limits as a function of hypothetical resonance mass for different values of the signal width, including $\Gamma_X =$ 4 MeV, $\alpha_X = 2$, 6 and 10%. The observed limit (solid line, computed with 2 GeV step in the range 200 GeV to 2.7 TeV) is in agreement with the expected limit (dashed line, computed every 20 GeV, accompanied with ±1 and 2 standard deviation band, green and yellow correspondingly) assuming the background-only hypothesis. Figure 5.8.7 provides the expected and observed limits in a two- dimensional plane of hypothetical signal masses and relative widths.

Impact of systematic uncertainties on the cross-section limit

The impact of the systematic uncertainties discussed in Section 5.7 on the expected limit is estimated and presented on Figure 5.8.8. The dominant systematic uncertainty impacting on the limit is the number of the spurious signal, it has the largest contribution in the mass range 200-700 GeV. The impact of the systematic uncertainties on

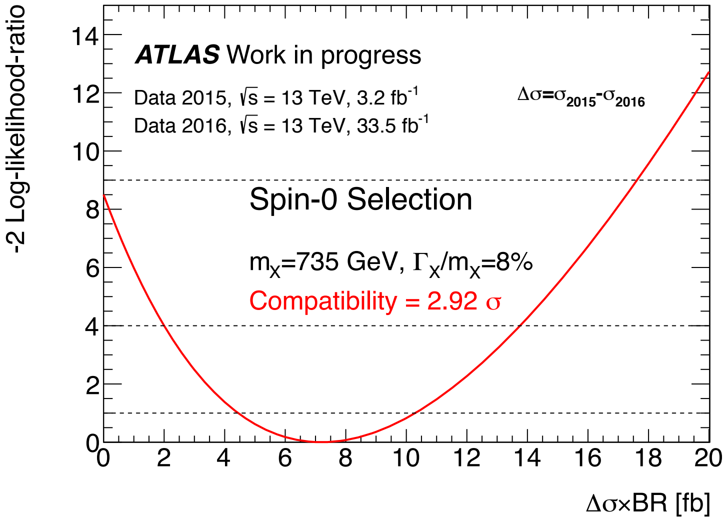


Figure 5.8.4: Profile likelihood ratio as a function of $\Delta\sigma = \sigma_{2016} - \sigma_{2015}$ [123].

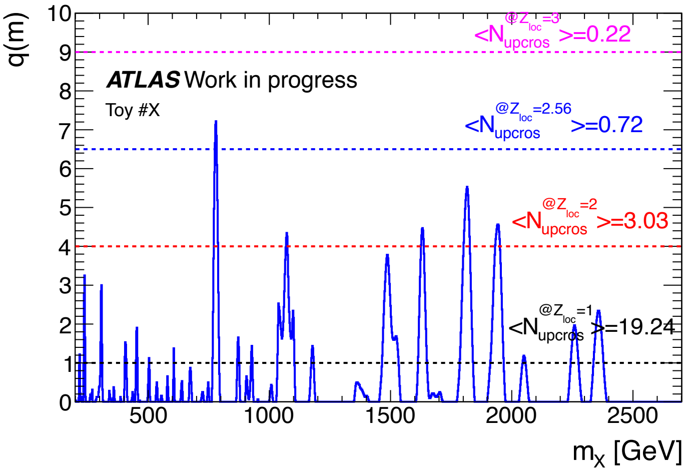


Figure 5.8.5: Example of a scan of $q_0(m)$ for one of the toy, showing the different tested crossing significances level, with averaged number of upcrossings over all toy experiments.

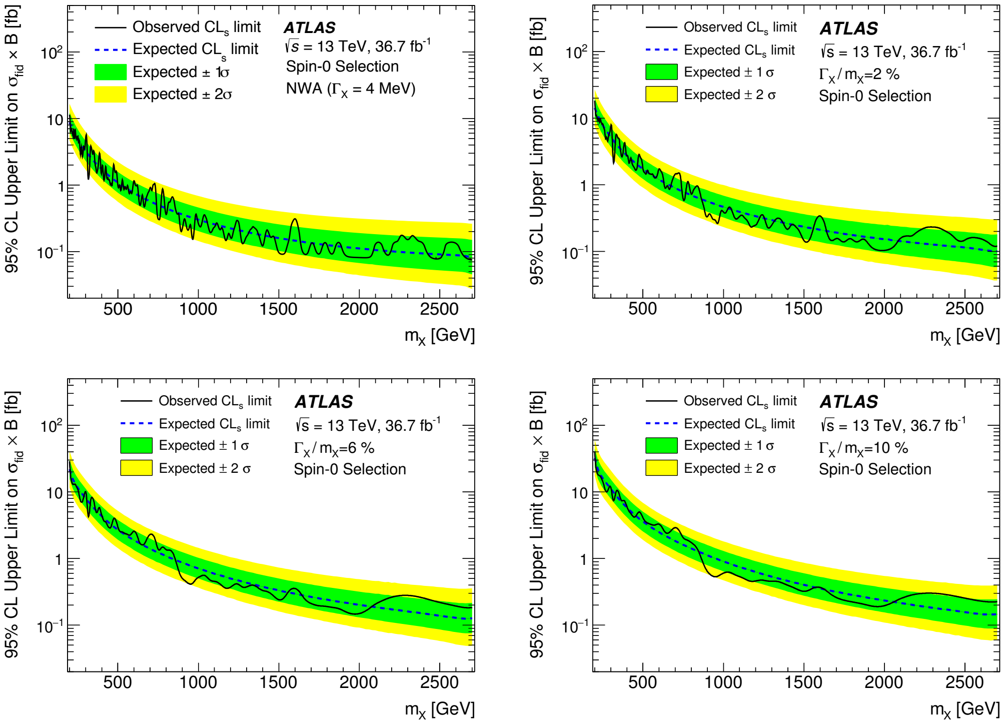


Figure 5.8.6: Expected and observed limit on the fiducial production cross-section limit $\sigma_{\text{fiducial}} \times \text{BR}(X \to \gamma\gamma)$ as a function of the resonance mass m_X, for various values of the signal width α_X for the combined 2015 and 2016 data-set [116].

the limit decreases versus the resonance mass from 29% at 200 GeV to 5% at 700 GeV. Above 700 GeV the impact is typically 2-3%.

5.8.3 Limits on specific models: interpretations in terms of 2HDM

The upper limits showed in Section 5.8.2 are given in a model independent way. They can be translated into an exclusion contour in the 2HDM for the phase space where the narrow width assumption is valid. The different types of the 2HDM have been discussed in Section 1.4.1 of the theoretical introduction.

Decay of a heavy scalar to diphoton may be important at high mass if the 2HDM is very close to the alignment limit, where the contributions from $H \to VV$ vanish from the total width. Contours of the inclusive $\sigma \cdot \text{Br}(H \to \gamma\gamma)$ relative to the Standard Model σ·Br are taken from the Reference [5], which are calculated for 8 TeV center of mass. The benchmark is defined by setting $m_H = 300$ GeV. The fiducial limit obtained in the search at 300 GeV corresponds to 3.23 fb, as presented on Figure 5.8.6 top right. It is converted to the total cross section times branching fraction to the diphoton final state, using the acceptance factor from Figure 5.3.2. The change of inclusive parton luminosity (gg and $q\bar{q}$) between 13 TeV and 8 TeV is taken into account. The overlay exclusion region on the contours of the inclusive σ·BR$(H \to \gamma\gamma)$ for 2HDM models Type-I and Type-II are presented on Figure 5.8.9.

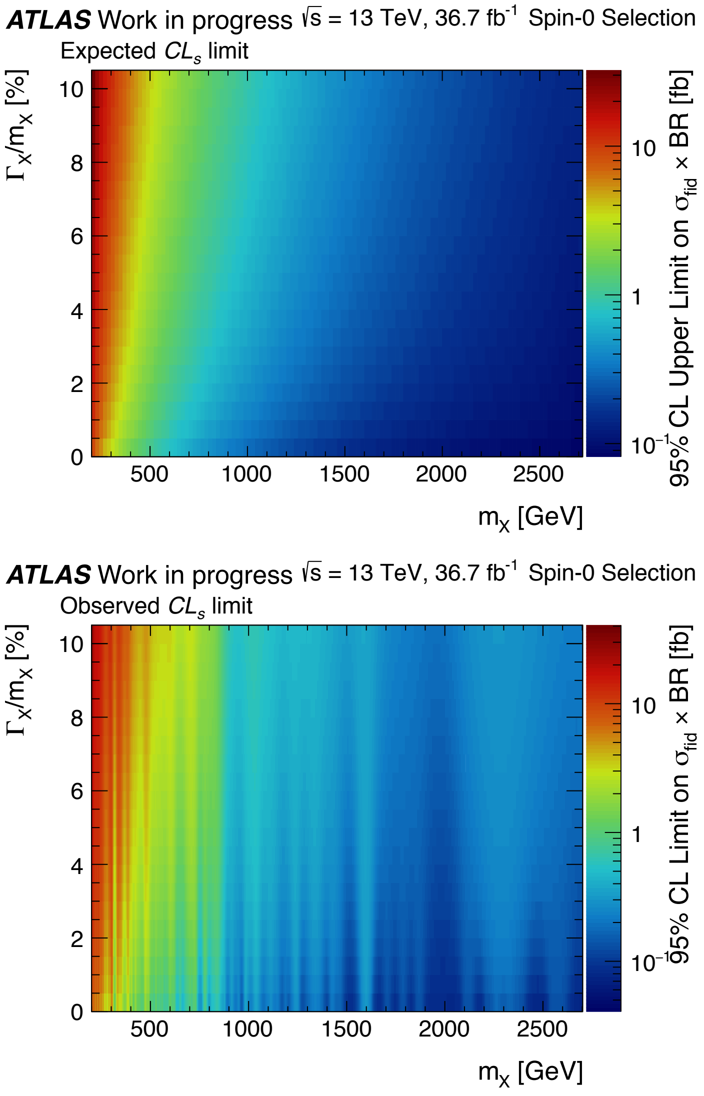


Figure 5.8.7: Expected and observed 2-dimensional limit on the fiducial production cross-section σ_{fid} as a function of the resonance mass m_X and relative natural width α_X [123].

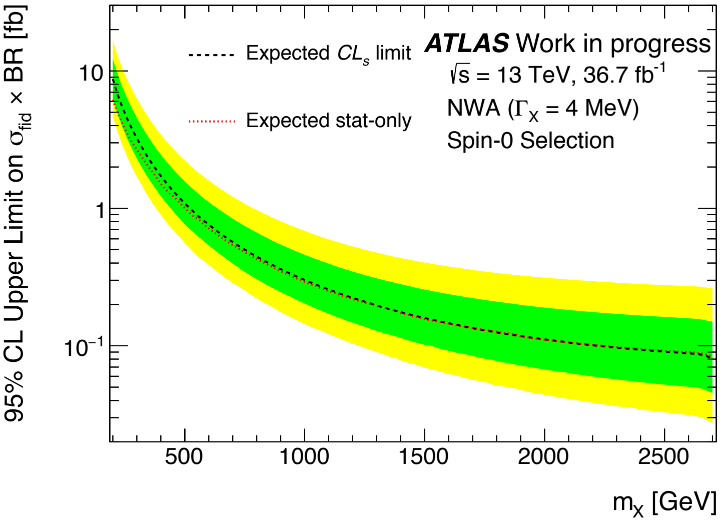


Figure 5.8.8: Expected limit on the fiducial production cross-section σ_{fid} as a function of the resonance mass m_X for NWA signal presented in black line. Expected limit assuming only statistical uncertainty presented in red [123].

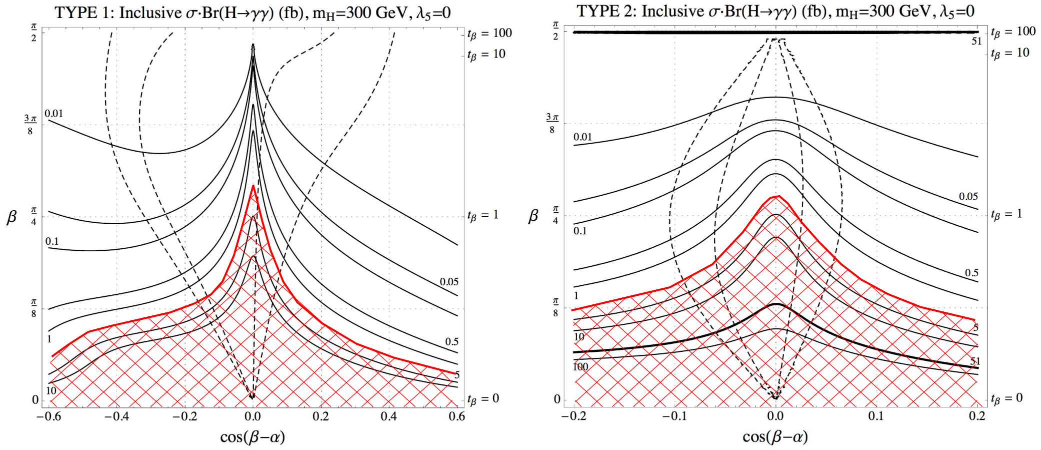


Figure 5.8.9: The exclusion contours of the inclusive $\sigma{\cdot}\mathrm{BR}(H \to \gamma\gamma)$ in units of fb for the non-SM scalar Higgs boson with m_H = 300 GeV, shown in the plane of $\cos(\beta-\alpha)$ and $\tan\beta$ for Type-I (left) and Type-II (right). The red area corresponds to the limit set by the search discussed in this Chapter.

5.9 Conclusions

A search for a high mass diphoton resonance was performed using the data sample collected by the ATLAS detector at the LHC during the 2015 and 2016 data taking campaigns. The complete dataset corresponds to an integrated luminosity of 36.7 fb^{-1} of pp collisions at $\sqrt{s}$ =13 TeV. The data are consistent with the background-only hypothesis over the explored mass range from 200 to 2700 GeV. Limits on the fiducial cross section at 95% CL are derived as a function of the mass and the width of the resonance.

I have contributed to this search performing following studies:

- The difference between the ratio of the event yields and the ratio of the luminosities between 2016 and 2015 datasets has been understood, and the origin of the effect found to be the dependence of the selection efficiency of identification and isolation criterias to the average number of interactions per bunch crossing, which differ by a factor of about two between two data-takings.
- The detector resolution has been extracted, and the signal model has been performed as an analytical description for the 2D parameter space of mass and the width of the predicted resonance. The bias of choice of the production mode and the systematic uncertainties of the photon energy scale and resolution were estimated.
- The statistical interpretation has been performed, and the impact of the systematic uncertainties on the cross section has been estimated.

In the events collected in 2015 (corresponding to ~10% of the total available data sample) a deviation from the background-only hypothesis was observed in a broad region of ±25 GeV near the mass of 750 GeV with local significances of 3.9 standard deviations. This local excess was an interesting case, because a similar search for high mass diphoton resonances had been performed by the CMS collaboration, which also showed an excess in the same mass range. A brief discussion and comparison of results are given in Section 5.9.1.

The prospects of the analysis with the full Run-2 expected dataset and next steps in the strategy are summarised in Section 5.9.2.

5.9.1 Comparison with CMS

In high energy physics, the standard convention of the level of significance where a discovery can be claimed is 5σ-level significance, corresponding to a p-value of 2.87×10^{-7}. So, why the 3.Xσ observed on data 2015 made such interest in the community? Because the two, CMS and ATLAS, collaborations observed a similar deviation in the same mass range. In the following, the comparison of the two analysis is presented, using data collected in 2015 and a second stage with the 2016 data where no deviation from the background-only hypothesis were found.

The two collaborations performed similar analysis, the comparison of the strategies and of the selections are presented in Table 5.9.1. ATLAS analysis uses a restricted

	ATLAS	CMS
Signal hypothesis	Spin-0 and Spin-2	
Kinematic selections	Spin-0: ${E_{\mathrm{T}}}^{\gamma_1(\gamma_2)}/m_{\gamma\gamma}$>0.4 (0.3) Spin-2: ${E_{\mathrm{T}}}^{\gamma_1,\gamma_2}$ >55 GeV	${E_{\mathrm{T}}}^{\gamma_1,\gamma_2}$ >75 GeV
Width hypothesis	Spin-0: [0-10]% m_X Spin-2: [0.01-11]% m_{G^*}	$[1.4{\times}10^{-2}, 1.4, 5.6]\% m_X$
Mass range, TeV	Spin-0: [0.2-2] Spin-2: [0.5-3.5]	[0.5-4.5]
Categories	-	both γ in the barrel $\gamma_{1(2)}$ in the barrel, $\gamma_{2(1)}$ in the endcap

Table 5.9.1: Comparison of the analysis strategies and kinematical selections of the ATLAS and the CMS collaborations.

kinematic range for the photon selection for the search of the Spin-0 resonances (relative cuts $E_{\mathrm{T}}^{\gamma_1(\gamma_2)}/m_{\gamma\gamma} > 0.4(0.3)$), taking advantage of the isotropic distribution of the decay products in the centre-of-mass frame of the new particle. The kinematic selection requirements is looser for the spin-2 resonance search (absolute cuts: $E_{\mathrm{T}}^{\gamma_1,\gamma_2}$ >55 GeV). CMS has same kinematic selections for both Spin-0 and Spin-2 searches, but the use of categories helps to increase the sensitivity to Spin-0 search. Two categories are presented:

- both photons are in the ECAL barrel detector (EBEB)
- one photon in the ECAL barrel detector and the other in an ECAL endcap detector (EBEE).

ATLAS analysis performs the results of the search as a function of the resonance mass and the width, covering the range from narrow width, where intrinsic width of the signal (4 MeV) is negligible with respect to the detector resolution, up to 10% of the resonance mass. CMS analysis interpret the results in three relative width hypothesis: $1.4{\times}10^{-2}\% m_X$ (corresponds to NWA used in ATLAS), $1.4\% m_X$ and $5.6\% m_X$.

2015 dataset

A search for resonant production of high-mass photon pairs was presented by the CMS experiment [113], using 3.3 fb^{-1} of pp collisions collected in 2015 at $\sqrt{s}$ =13 TeV. The search performed for scalar and Randall-Sundrum gravitons for resonance masses $0.5 < m_X < 4.5$ TeV and relative widths of $1.4{\times}10^{-4} < \Gamma_{\mathrm{X}}/m_X < 5.6 \times 10^{-2}$ A modest excess of events over the background-only hypothesis is observed for m_X around 750 GeV. The invariant mass spectrum of selected diphoton events are shown on Figure 5.9.1 (left).

The largest excess observed in data has a p_0 value corresponding to 2.85 (2.9) standard deviations for spin-0 (spin-2) hypothesis at the mass 760 GeV and relative width of $\Gamma_{\mathrm{X}}/m_X = 1.4 \times 10^{-4}$ (the narrow-width hypothesis), as shown on Figure 5.9.1 (right). The global significance of the excess is estimated to be less than one standard deviation. Additionally, the CMS collaboration performed a combined analysis

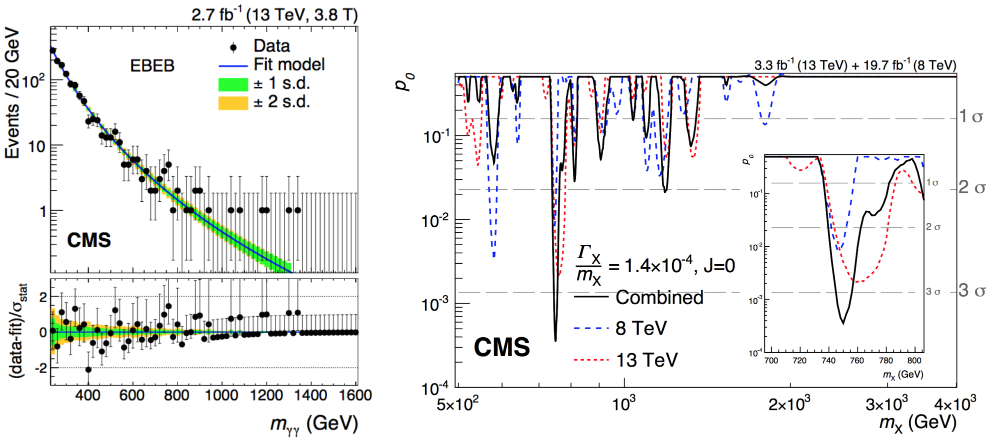


Figure 5.9.1: Left: Observed diphoton invariant mass $m_{\gamma\gamma}$ spectra used in the analysis of the 13 TeV 2015 dataset (magnetic field strength B = 3.8 T; both photons in the ECAL barrel detector). Right: Observed background-only p-values for narrow-width scalar resonances ($\Gamma_X/m_X = 1.4 \times 10^{-2}$) as a function of the resonance mass m_X, the separate 8 and 13 TeV data sets and for the combined analysis of the 8 and 13 TeV data [113].

including ~20fb^{-1} from the 8 TeV dataset and 3.3fb^{-1} of 13 TeV dataset. The result is presented on Figure 5.9.1: the largest excess is observed for mass 750 GeV and $\Gamma_X/m_X = 1.4 \times 10^{-4}$. The local p-value corresponds to approximately 3.4 standard deviations. The global p-value, taking into account the effect of searching for several signal hypotheses is approximately 1.6 standard deviations.

Searches for new resonances decaying into two photons in the ATLAS experiment, were performed on pp collision data corresponding to an integrated luminosity of 3.2 fb^{-1} at $\sqrt{s} = 13$ TeV recorded in 2015 [115]. The largest deviation from the background-only hypothesis is observed in a broad region ±25 GeV near a mass of 750 GeV, with local significances of 3.8 and 3.9 standard deviations in the searches optimized for the spin-2 and spin-0 resonances, respectively. The invariant mass distribution and the local significance for various masses and the decay widths hypothesis are presented on Figure 5.9.2 left and right. The global significances are estimated to be 2.1 standard deviations for both searches. Assuming a scaling of the production cross section for an s-channel resonance produced by gluon fusion (light quark-antiquark annihilation), the consistency between the 13 TeV data and the data collected at 8 TeV is found to be at the level of 2.7 (3.3) standard deviations using results from the searches optimized for a spin-2 particle and at the level of 1.2 (2.1) standard deviations using results from the searches optimized for a spin-0 particle.

The comparison of the most significant results obtained by the ATLAS and the CMS collaborations based on 2015 data is summarised in Table 5.9.2. The detailed view on the results shows that there is a significant discrepancy in the preferable width of the resonance, and a sizeable difference of the local significance of the excess between

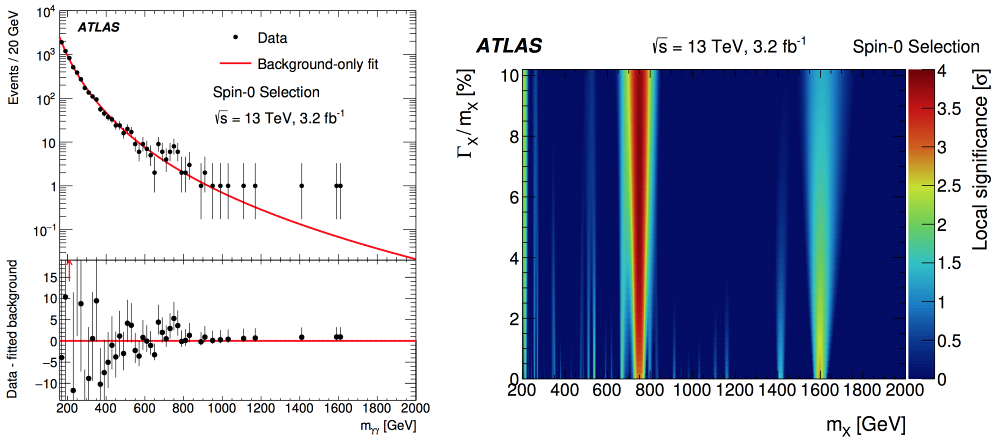


Figure 5.9.2: Left: Distribution of the invariant mass of the diphoton candidates for the selection used in the search for a spin-0 resonance with the best background-only fit. Right: Compatibility, in terms of local significance σ, with the background-only hypothesis as a function of the assumed signal mass m_X and relative width Γ_X/m_X for the analysis optimized for a spin-0 resonance search [115].

	ATLAS	CMS	
$\sqrt{s}$, TeV	13	13	13+8
$\mathcal{L}$, fb^{-1}	3.2	3.3	3.3+19.7
Mass, GeV	750	760	750
Hypothesis	Spin-0	Equal for Spin-0 and Spin-2	
Local Z, σ	3.9	2.9	3.4
Width	6%m_X	1.4%m_X	1.4×10^{-2}%m_X
Global Z, σ	2.1	-	1.6

Table 5.9.2: Comparison of the most significant deviation from the background-only hypothesis observed in data collected in 2015 by the ATLAS and CMS collaborations.

the results of two collaborations.

Including 2016 dataset

The analysis has been updated by both collaborations, using the data collected in 2016. The results of the ATLAS collaboration based on the combined 36.7 fb^{-1} collected at $\sqrt{s}$=13 TeV are summarised in Section 5.8, and the paper is in preparation. The CMS collaboration provided updates of the analysis including 12.9 fb^{-1} collected in 2016 [114] to the results reported in [113]. No significant deviation from the background-only hypothesis have been observed. The summary of the statistical results is presented in Table 5.9.3. Comparing Tables 5.9.2 and 5.9.3 one concludes, that the excess seen in the 2015 data is likely to be a statistical fluctuation.

	ATLAS	CMS
$\sqrt{s}$, TeV	13	13+8
$\mathcal{L}$, fb^{-1}	36.7	16.2+19.7
Mass, GeV	730	~1300
Hypothesis	Spin-0	
Local Z, σ	2.5	2.2
Width	NWA	
Global Z, σ	~0	<1

Table 5.9.3: Comparison of the most significant deviation from the background-only hypothesis observed by the ATLAS and CMS collaborations in Spin-0 hypothesis.

5.9.2 Analysis prospects

Current analysis

As of today, the search for new resonances did not find any deviation from the SM expectations, and has therefore set limits on the fiducial cross section times branching ratio. The available dataset allows to set an exclusion limit of 11.4 fb for a resonance mass of 200 GeV to 0.1 fb for a resonance mass of 2700 GeV. What sensitivity can we reach by the end of Run-2, assuming an integrated luminosity of 100 fb^{-1}?

Here is provided a brief estimation of the limit evolution with increased integral luminosity. The scaled number of selected diphoton events assuming Spin-0 selections, as a function of the integrated luminosity is presented on Figure 5.9.3 (left). The functional form to describe the background is assumed to be the same as discussed in Section 5.5. The number of the spurious signal is scaled according to the increase of luminosity. This is simplified solution with very conservative systematic uncertainty. In general, the increase of the dataset leads to reduction of the statistical uncertainty, it creates the problem of finding the function which is flexible enough to fit the data spectrum and to avoid generating significant bias of the fitted signal (large number of spurious signal) to the search. One of the possible solutions to this problem, is to use several overlapping regions where the function can be optimised such, that it will satisfy the requirement on the number of the spurious signal, but increase the statistical error.

The expected limit has been estimated for dataset of 100 fb^{-1} and is presented on Figure 5.9.3 (right). This simplified projection to an integrated luminosity of 100 fb^{-1}estimates an expected limit on the fiducial cross section times branching ratio ranging from 6.6 fb at m_X=200 GeV to 0.03 fb at m_X=2700 GeV.

Extension of the analysis in the low mass region

Several extensions to the Standard Model predict new scalar resonances also below the Higgs mass which may decay to two photons []. With this motivation, a low-mass analysis was done in Run-1, searching for scalar diphoton resonances in the mass range 65-600 GeV at $\sqrt{s}$= 8 TeV [111].

The analysis of the Run-2 at $\sqrt{s}$= 13 TeV data will be extended to the low mass

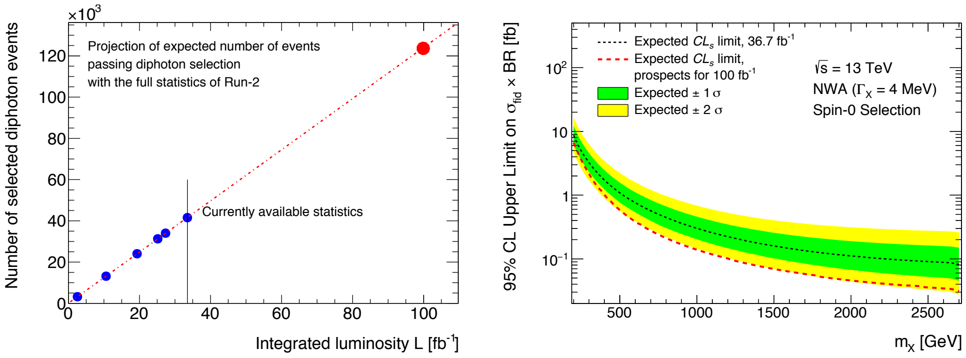


Figure 5.9.3: Left: The scaled number of selected diphoton events as a function of the integrated luminosity. Right: Expected limit on the fiducial production cross-section σ_{fid} as a function of the resonance mass m_X for NWA signal presented in black line for current available statistics, and in red presented projection to 100fb^{-1}.

region. Such a low-mass analysis requires a precise modeling of the Drell-Yan (DY) background, dominated by the Z boson decays to an electron-positron pair, where both electron and positron are misidentified as photons, and mostly classified as converted photons. The exploitation of categories with different signal-to-background ratios, according to the conversion status of the photon candidates can mitigate the loss of the signal sensitivity. In the low-mass region for $m_{\gamma\gamma} < 80$ GeV, the analysis also needs to account for the "turn-on" effect induced on the $m_{\gamma\gamma}$ distribution by the photon looser E_{T} cuts. Using a diphoton trigger with E_{T} threshold of 20 GeV on each photon candidate would allow to have offline cuts of ∼22 GeV, which would move the turn-on outside the analysis range.

Recently theorists (B. Bellazzini et al.) have provided a new model which motivates diphoton searches below the ATLAS threshold of 65 GeV [139], providing a mass range not yet excluded neither by LEP nor LHC. The analysis team is working now on the possibility to reduce the thresholds on the photon E_{T} in the trigger definition. This would allow the analysis to explore the 50-65 GeV region, and to target the region defined in Ref. [139].

Conclusions

The successful restart of the LHC in 2015 with an increased centre of mass energy of $\sqrt{s}$=13 TeV, opened a new page in the history of high energy physics. The increased energy has two main advantages: unexplored phase space became accessible for searches of phenomena Beyond the Standard Model (BSM); and the increase of the signal cross-sections lead to more precise measurements and to increased sensitivity to new phenomena. The ATLAS detector has been operated with an outstanding performance and high quality data have been recorded.

This thesis presents studies of the energy response of photons and electrons with the ATLAS detector and the search for new resonances in the diphoton channel. The reconstruction of the photon energy and the associated systematic uncertainty is one of the critical elements for the precision measurement of the Higgs boson mass in the diphoton channel.

A search for high-mass scalar resonances in the diphoton channel was conducted with 36.7 fb^{-1} of data collected in 2015 and 2016. An intriguing excess, at an invariant mass of about 750 GeV was observed with the 2015 data sample, corresponding to 3.2 fb^{-1}. This excess was not confirmed with the full available dataset, where the data were consistent with the background-only hypothesis. Exclusion limits at 95% CL were set on the production cross-section of a resonance decaying to two photons, as a function of the resonance mass and width. The observed limits for a narrow-width signal range from 11.4 fb for m_X=200 GeV to about 0.1 fb for m_X=2700 GeV.

The dominant contribution to the total systematics uncertainty on the Higgs boson mass in the diphoton decay channel originated from the photon energy scale calibration. I studied the origin of this uncertainty and showed, that the origin of the effect was coming from the shape of the EM shower. This implied that the method used for Run-1 data and the corresponding systematic uncertainty needed to be re-evaluated. The ATLAS simulation gives a good description of the electromagnetic shower shapes for most cases. The lateral shower description is one exception in particular for some regions of the detector. A systematic uncertainty has been derived to cover this difference. The method proposed to improve the response of the calibration procedure, introduces the information of the lateral shower while deriving the energy calibration of EM showers.

The next steps in the particle physics is to study with high precision the properties of the Higgs boson: its mass, spin, CP properties and couplings to existing bosons and fermions. The available integrated luminosity of the Run-2 will allow an accurate determination of the Yukawa coupling to the top quark using the Higgs boson production associated with a $t\bar{t}$ pair. These couplings are particularly sensitive to new physics phenomena, especially when they are generated by quantum loops, as in the case of production via gluon-gluon fusion or in the decay into two photons.

The collected data are in agreement with the SM predictions; searches for supersymmetry, dark matter, and other extensions to the Standard Model are being carried out.

On a longer timescale, the high luminosity LHC (HL-LHC) and the upgraded detectors will extend the era of high precision measurements. With the expected integrated luminosity at the end of LHC, 3000 fb^{-1} the statistical precision will increase by a factor of 10. It will provide the accurate measurement of the 3- and 4-Higgs self-couplings, and the possibility to observe the rare SM process $H \to \mu\mu$.

References

[1] Hill A., Bij B. . **Strongly interacting singlet - doublet Higgs model** PhysRevD.36.3463, 1987 1, 24

[2] Binoth B.. **Influence of strongly coupled, hidden scalars on Higgs signals** hep-ph 9608245 1997 1

[3] Schabinger S.,. **A Minimal spontaneously broken hidden sector and its impact on Higgs boson physics at the large hadron collider** hep-ph 0509209, 2005 1, 24

[4] Branco G., et al.. **Theory and phenomenology of two-Higgs-doublet models** hep-ph 1106.0034 1, 24

[5] Craig, N., Galloway, J., Thomas, S.. **Searching for Signs of the Second Higgs Doublet** hep-ph 1305.2424 24, 26, 144

[6] S. L. Glashow and S. Weinberg. **Natural Conservation Laws for Neutral Currents** Phys.Rev.D15 (1977) 1958.

[7] Craig, N., Thomas, S.. **Exclusive Signals of an Extended Higgs Sector** hep-ph 1207.4835 26

[8] M. E. Peskin and D. V. Schroeder. *An Introduction to quantum field theory.* Addison-Wesley, Reading, 1995.

[9] M. Gell-Mann. **Symmetries of baryons and mesons** Phys.Rev. 125 (1962) 10671084. 8

[10] G. Zweig. **An SU(3) model for strong interaction symmetry and its breaking** CERN-TH-401, Jan 17, 1964 8

[11] D. Gross and F. Wilczek. **Asymptotically free gauge theories.** Phys.Rev. D9 (1974) 980993. 9

[12] K. G. Wilson. **Confinement of Quarks** Phys.Rev. D10 (1974) 24452459 9

[13] R. P. Feynman and M. Gell-Mann. **Theory of Fermi interaction** Phys.Rev. 109 (1958) 193198 9

[14] C. Wu, E. Ambler, R. Hayward, D. Hoppes, and R. Hudson. **Experimental Test of Parity Conservation in Beta Decay** Phys.Rev. 105 (1957) 14131414 9

[15] P. W. Higgs. **Broken Symmetries and the Masses of Gauge Bosons** Phys.Rev.Lett. 13 (1964) 508 509 11

[16] F. Englert and R. Brout. **Broken Symmetry and the Mass of Gauge Vector Mesons** Phys.Rev.Lett. 13 (1964) 321323 11

[17] G. Guralnik, C. Hagen, and T. Kibble. **Global Conservation Laws and Massless Particles** Phys.Rev.Lett. 13 (1964) 585587 11

[18] **The fifth Annual Large Hadron Collider Physics conference (LHCP2017)** https://indico.cern.ch/event/517784/ 6

[19] **Illustration of an LHC proton-proton collision.** http://imperialhep.blogspot.fr/2011/08/strangeness-at-lhcb.html 14

[20] **Introduction to the LHC Experiments** https://inspirehep.net/record/853601 15

[21] S.D. Drell and Tung-Mow Yan. **Partons and their applications at high energies** Ann. Phys. 66:578, 1971. 16

[22] G. Altarelli and G. Parisi. **ASYMPTOTIC FREEDOM IN PARTON LANGUAGE** Nucl.Phys. B126:298 (1977) 17

[23] Yu.L. Dokshitzer. **Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.** Sov.Phys. JETP 46:641 (1977) 17

[24] T. Gleisberg et al.,. **SHERPA 1.alpha, a proof-of-concept version** JHEP 02 (2004) 056, arXiv:0311263 [hep-ph] 18

[25] ATLAS Collaboration **ATLAS Pythia 8 tunes to 7 TeV data** ATL-PHYS-PUB-2014-021 http://cds.cern.ch/record/1966419 18

[26] S. Schumann and F. Krauss, **A parton shower algorithm based on Catani-Seymour dipole factorisation** JHEP 03 (2008) 038, arXiv: 0709.1027 [hep-ph]. 18

[27] S. Hoche et al., **QCD matrix elements and truncated showers** JHEP 05 (2009) 053, arXiv: 0903.1219 [hep-ph]. 18

[28] W.James Stirling. **Progress in Parton Distribution Functions and implications for LHC** https://inspirehep.net/record/805147 17

[29] R. D. Ball et al.. **Parton distributions with LHC data** arXiv:1207.1303 [hep-ph] 18

[30] S. Frixione et al.,. **Matching NLO QCD and parton showers in heavy flavour production** JHEP 08 (2003) 007, arXiv:0305252 [hep-ph] 18

[31] **Event display** https://twiki.cern.ch/twiki/pub/AtlasPublic/EventDisplayRun2Physics/run3044 evt2206548301-05GeV.png 15

[32] The ATLAS Collaboration. **Combined measurements of the mass and signal strength of the Higgs-like boson with the ATLAS detector using up to 25 fb^{-1} of proton-proton collision data** ATLAS-CONF-2013-014 http://cds.cern.ch/record/1523727

[33] Anastasiou, C., Duhr, C., Dulat, F. et al. . **High precision determination of the gluon fusion Higgs boson cross-section at the LHC** J. High Energ. Phys. (2016) 2016: 58 19

[34] M. Ciccolini, A. Denner, and S. Dittmaier,. **Strong and electroweak corrections to the production of Higgs + 2-jets via weak interactions at the LHC** Phys. Rev. Lett. 99 (2007) 161803 arXiv:0707.0381 [hep-ph]. 19

[35] A. Denner, S. Dittmaier, and A. Muck. **HAWK: A Monte Carlo generator for the production of Higgs bosons Attached to WeaK bosons at hadron colliders** http://omnibus.uni-freiburg.de/ sd565/programs/hawk/hawk.html, 2010. 19

[36] O. Brein, A. Djouadi, and R. Harlander,. **NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders** Phys. Lett. B 579 (2004) 149156 arXiv:hep-ph/0307206. 19

[37] W. Beenakker et al.. **NLO QCD corrections to ttH production in hadron collisions** Nucl. Phys. B 653 (2003) 151203 arXiv:hep-ph/0211352. 19

[38] LHC Higgs Cross Section WG. **SM Higgs production cross sections at $\sqrt{s} = 13$ TeV** https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt13TeV

[39] LHC Higgs Cross Section WG. **LHC Higgs Cross Section WG Picture Gallery** https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWGCrossSectionsFigures 21

[40] LHC Higgs Cross Section WG. **SM Higgs Branching Ratios and Total Decay Widths** https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageBR

[41] A. Djouadi et al.. **HDECAY: A Program for Higgs boson decays in the standard model and its supersymmetric extension** Comput.Phys.Commun. 108 (1998) 5674, arXiv:9704448 [hep-ph]. 20

[42] A. Bredenstein et al.. **Prophecy4f: A Monte Carlo generator for a proper description of the Higgs decay into 4 fermions** http://omnibus.uni-freiburg.de/ sd565/programs/prophecy4f/prophecy4f.html 20

[43] ATLAS, CMS Collaboration. **Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments**. Phys. Rev. Lett. 114, 191803, 2015 21, 22

[44] Particle Data Group. **10. Electroweak model and constraints on new physics** http://pdg.lbl.gov/2016/reviews/rpp2016-rev-standard-model.pdf 22, 23

[45] ATLAS, CMS Collaboration. **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**. JHEP08(2016)045, 2016 23

[46] L. Fromme, et al.. **Baryogenesis in the Two-Higgs Doublet Model** hep-ph 0605.242 24

[47] L. Randall, R. Sundrum. **A Large Mass Hierarchy from a Small Extra Dimension** hep-ph 9905221 27

[48] L. Csaki, P. Tanedo. **Beyond the Standard Model** hep-ph 1602.04228 27

[49] H. Davoudiasl and J.Hewett. **Phenomenology of the Randall-Sundrum Gauge Hierarchy Model** PhysRevLett.84.2080 hep-ph 9909255 28

[50] P. Bryant L. Evans. **LHC Machine**. *JINST*, **3**(S08001), 2008. 31, 35

[51] P. Lebrun S. Myers R. Ostojic J. Poole O.S. Brüning, P. Collier and P. Proudlock. *LHC Design Report: The LHC Main Ring.* CERN, Geneva, 2004. 31

[52] **The CERN accelerator complex**. https://cds.cern.ch/record/2197559, 2016. 32

[53] **LHC Layout.**. https://cds.cern.ch/record/841573, 1997. 33

[54] **Cross section of LHC dipole.**. https://cds.cern.ch/record/842530, 1999. 34

[55] **LEP design report**. Technical report, CERN-LEP-84-01, 1984.

[56] The ATLAS Collaboration. **The ATLAS Experiment at the CERN Large Hadron Collider**. *JINST*, **3**(S08003), 2008. 31, 39, 42, 43, 45, 46, 47, 48, 52, 53, 54, 56, 57, 58

[57] The CMS Collaboration. **The CMS experiment at the CERN LHC**. *JINST*, **3**(S08004), 2008. 31

[58] The ALICE Collaboration. **The ALICE experiment at the CERN LHC**. *JINST*, **3**(S08002), 2008. 31

[59] The LHCb Collaboration. **The LHCb Detector at the LHC**. *JINST*, **3**(S08005), 2008. 31

[60] The ATLAS Collaboration. **Luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC**. *Eur. Phys. J. C*, **71**(1630), 2011. 33

[61] L. Rossi et. al. **High-Luminosity Large Hadron Collider; a description of the European strategy preparatory group**. Technical report, CERN, 2012. 37, 38

[62] G. Apollinari et. al. **High-Luminosity Large Hadron Collider (HL-LHC). Preliminary Design Report,**. Technical report, CERN-2015-005, 2015 38

[63] The ATLAS Collaboration. **Prospects for New Physics in Higgs Couplings Studies with the ATLAS Detector at the HL-LHC**. ATL-PHYS-PUB-2014-017, 2014 38

[64] The ATLAS Collaboration. **ATLAS inner detector : Technical Design Report**. Technical report, CERN-LHCC-97-016, 1997. 39, 40

[65] The ATLAS Collaboration. **ATLAS Insertable B-Layer Technical Design Report**. Technical report, CERN-LHCC-2010-013, 2010 . 40

[66] The ATLAS Collaboration. **Particle Identification Performance of the ATLAS Transition Radiation Tracker**. ATLAS-CONF-2011-128, 2011. 42

[67] The ATLAS Collaboration. **ATLAS muon spectrometer : Technical Design Report**. Technical report, CERN-LHCC-97-022, 1997. 39

[68] The ATLAS Collaboration. **ATLAS liquid-argon calorimeter : Technical Design Report**. Technical report, CERN-LHCC-96-041, 1996. 39, 54

[69] The ATLAS Collaboration. **Impact Parameter Resolution**. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2015-007/. 40

[70] The ATLAS Collaboration. **Muon reconstruction performance of the ATLAS detector in protonproton collision data at $\sqrt{s} = 13\ TeV$**. Eur. Phys. J. C (2016) 76: 292. 46

[71] The ATLAS Collaboration. **Performance of the ATLAS Trigger System in 2015.**. CERN-EP-2016-241, 2016. arXiv: 1611.09661 [hep-ex]. 48, 49

[72] N.J. Buchanan et al,. **Design and implementation of the Front End Board for the readout of the ATLAS liquid argon calorimeters**. *JINST*, **3**(S09003), 2008. 55

[73] The ATLAS Collaboration. **ATLAS Calorimeter Performance**. CERN/LHCC/96-40, 1996. 55

[74] M. Aharrouche et al.. **Energy Linearity and Resolution of the ATLAS Electromagnetic Barrel Calorimeter in an Electron Test-Beam**. Nucl.Instrum.Meth.A 568:601-623,2006. 54

[75] M. Aleksa et al.. **ATLAS Combined Testbeam: Computation and Validation of the Electronic Calibration Constants for the Electromagnetic Calorimeter**. ATL-LARG-PUB-2006-003, 2006. 59

[76] The ATLAS Collaboration. **Readiness of the ATLAS Liquid Argon Calorimeter for LHC Collisions**. Eur.Phys.J.C70:723-753, 2010 54

[77] L. Courneyea et al.. **Computation and validation of the electronic calibration constants for the ATLAS Liquid Argon Calorimeters**. ATL-LARG-INT-2010-007, 2010. 58, 59

[78] W. Vandelli et al.. **The raw event format in the ATLAS Trigger and DAQ**. ATL-D-ES-0019, 2014. 60, 61

[79] **LAr bytestream**. https://atlasop.cern.ch/twiki/bin/view/LAr/LArCheckBytestreamType. 62

[80] GEANT4 Collaboration, S. Agostinelli et al.. **GEANT4: A simulation toolkit**. Nucl. Instrum. Meth. A506, 2003. 63

[81] J. Allison et al.. **Geant4 developments and applications**. IEEE Trans. Nucl. Sci. 53-270, 2006. 63

[82] S.P. Griso. **Mu-Rescaling: material for discussion on recommendation**. https://indico.cern.ch/event/437993/contributions/1925644/attachments/1138739/1630981/spagan_MuRescaling.pdf, 2015, Internal documentation. 64

[83] W. Lampl et al.. **Calorimeter Clustering Algorithms: Description and Performance**. ATL-LARG-PUB-2008-002, 2008. 70

[84] The ATLAS Collaboration. **Expected photon performance in the ATLAS experiment**. ATL-PHYS-PUB-2011-007, 2011, http://cdsweb.cern.ch/record/1345329. 74

[85] The ATLAS Collaboration. **Study of inclusive isolated-photon production in pp collisions at $\sqrt{s} = 13\ TeV$ with the ATLAS detector**. ATL-COM-PHYS-2015-603, 2015, Internal documentation. https://cds.cern.ch/record/2030168. 71

[86] K. Grevtsov et al.. **Photon reconstruction using 18.6 pb^1 of pp collisions collected by ATLAS at $\sqrt{s} = 13\ TeV$ in 2015**. ATL-COM-PHYS-2015-725, 2015, Internal documentation. https://cds.cern.ch/record/2034889. 71

[87] Kirill Grevtsov. **Electron and Photon performance in ATLAS at the LHC**. EPS-HEP conference 22-29 July 2015, http://indico.cern.ch/event/356420/contributions/1764221. 71

[88] The ATLAS Collaboration. **Performance of the ATLAS Liquid Argon Calorimeter after three years of LHC operation and plans for a future upgrade**. Advancements in Nuclear Instrumentation Measurement Methods and their Applications, 2013, arXiv: 1306.6756 [physics.ins-det] . 73

[89] J. Saxon and H. Williams. **Neural Networks for Photon Identification in $H \rightarrow \gamma\gamma$**. ATL-COM-PHYS-2013-305, 2013, Internal documentation. https://cds.cern.ch/record/1524081. 74

[90] The ATLAS Collaboration. **Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-1 data**. Eur. Phys. J. C (2016) 76: 666. , 2016 75

[91] The ATLAS Collaboration. **Photon identification in 2015 ATLAS data**. ATL-PHYS-PUB-2016-014, 2015 https://cds.cern.ch/record/2203125. 76

[92] The ATLAS Collaboration. **Photon Identification Efficiencies using 2016 Data with radiative Z boson decays**. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/EGAM-2016-003/index.html, 2016. 76

[93] M. Cacciari, G.P. Salam and G. Soyez. **The Catchment Area of Jets**. J. High Energy Phys. 04, 2008, arXiv: 0802.1188 [hep-ph]. 77

[94] M. Cacciari, G.P. Salam and S. Sapeta. **On the characterisation of the underlying event**. J. High Energy Phys. 04, 2010, arXiv: 0912.4926 [hep-ph]. 77

[95] The ATLAS Collaboration. **Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data**. *Eur. Phys. J. C*, **74**(3071), 2014. arXiv: 1407.5063 [hep-ex]. 79, 124

[96] The ATLAS Collaboration. **Electron and photon energy calibration with the ATLAS detector using data collected in 2015 at $\sqrt{s}$ = 13 TeV**. ATL-PHYS-PUB-2016-015. 79, 83, 89

[97] A. Hoecker et al.,. **Toolkit for Multivariate Data Analysis**. PoS ACAT, 040, 2007. arXiv: 0703039[physics] http://tmva.sourceforge.net/ 80

[98] B. Lenzi, R. Turra. **Monte Carlo calibration update for electrons and photons using multivariate techniques**. ATL-COM-PHYS-2013-1426, 2014, Internal documentation. 80

[99] The ATLAS Collaboration. **Central electrons energy scale factors and resolution constant term using Z to ee events from 2015 and 2016 data**. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/EGAM-2017-002, 2017. 83

[100] The ATLAS Collaboration. **Measurement of the Higgs boson mass in the $H \rightarrow ZZ^{*} \rightarrow 4l$ and $H \rightarrow \gamma\gamma$ channels with $\sqrt{s}$=13TeV pp collisions using the ATLAS detector**. ATLAS-CONF-2017-046, 2017. 84

[101] C. Becot et al.. **Difference of energy response between the high and medium gain of the ATLAS electromagnetic calorimeter**. ATL-COM-PHYS-2013-1655, 2014, Internal documentation. https://cds.cern.ch/record/1637736 89, 90

[102] The ATLAS Collaboration. **Measurement of the Higgs boson mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4l$ channels in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector** Phys. Rev. D 90, 052004, 2014. 87

[103] The ATLAS Collaboration. **Calibration stability, uniformity and linearity measurements of electrons and photons** ATL-COM-PHYS-2017-758, 2017, Internal documentation. https://cds.cern.ch/record/2268802 95, 99

[104] Andre David. **Run 2 Seminar and subsequent $\gamma\gamma$-related arXiv submissions**. http://jsfiddle.net/adavid/bk2tmc2m/show/. 109

[105] The ATLAS Collaboration. **Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC**. Phys.Lett. B716 1-29, 2012. arXiv: 1207.7214 [hep-ex]. 19, 110

[106] The CMS Collaboration. **Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC**. Phys. Lett. B 716 30, 2012. arXiv: 1207.7235 [hep-ex]. 19

[107] The 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**. ATLAS-CONF-2015-081, 2015. http://cds.cern.ch/record/2114853. 3

[108] K.Grevtsov, on behalf of the ATLAS Collaboration. **Search for high-mass diphoton resonances with the ATLAS detector**. https://indico.lal.in2p3.fr/event/3116/session/9/contribution/24/1/material/slides/0.pdf. 4

[109] The ATLAS Collaboration, Aaboud, M., Aad, G. et al.. **Search for resonances in diphoton events at $\sqrt{s} = 13\ TeV$ with the ATLAS detector**. J. High Energ. Phys., 2016. 10.1007/JHEP09(2016)001. 77, 78, 109, 141

[110] The ATLAS Collaboration. **Search for high-mass diphoton resonances in pp collisions at $\sqrt{s} = 8\ TeV$ with the ATLAS detector**. Phys. Rev. D 92, 2015. arXiv: 1504.05511 [hep-ex]. 110

[111] The ATLAS Collaboration. **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, 2014. arXiv: 1407.6583 [hep-ex]. 110, 151

[112] The CMS Collaboration. **Search for diphoton resonances in the mass range from 150 to 850 GeV in pp collisions at $\sqrt{s} = 8\ TeV$**. Phys. Lett. B 750, 2015. arXiv: 1506.02301 [hep-ex]. 110

[113] The CMS Collaboration. **CMS collaboration, Search for resonant production of high-mass photon pairs in proton-proton collisions at $\sqrt{s} = 8\ TeV$ and 13 TeV**. Phys. Rev. Lett. 117 , 2016. arXiv: 1606.04093 [hep-ex]. 148, 149, 150

[114] The CMS Collaboration. **Search for high-mass diphoton resonances in proton-proton collisions at 13 TeV and combination with 8 TeV search**. Phys. Lett. B 767 (2017) 147, arXiv: 1609.02507 [hep-ex]. 150

[115] The ATLAS Collaboration. **Search for resonances in diphoton events at $\sqrt{s}$ = 13 TeV with the ATLAS detector**. JHEP09(2016)001, arXiv: 1606.03833 [hep-ex]. 149, 150

[116] The ATLAS Collaboration. **Search for new phenomena in high-mass diphoton final states using 37 fb^{-1} of proton–proton collisions collected at $\sqrt{s}$ = 13 TeV with the ATLAS detector**. arXiv: 1707.04147 [hep-ex]. 28, 118, 140, 141, 142, 144

[117] LHC Higgs Cross Section Working Group. **Handbook of LHC Higgs Cross Sections: 3. Higgs Properties**. CERN-2013-004, 2013. arXiv: 1307.1347 [hep-ph]. 110

[118] T.D. Lee. **A Theory of Spontaneous T Violation**. Phys.Rev. D8, 1226-1239, 1973. 110

[119] B. Lenzi, D. Delgove. **Selection of the diphoton production vertex using multivariate techniques for $H \to \gamma\gamma$ and other analyses**. ATL-COM-PHYS-2015-1321, 2015, Internal documentation. https://cds.cern.ch/record/2062239 111

[120] The ATLAS Collaboration,. **Luminosity Public Results Run2**. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2 37, 113

[121] **ATLAS Luminosity Calculator**. https://atlas-lumicalc.cern.ch 36, 64

[122] The ATLAS Collaboration, Luminosity Group. **Luminosity for Physics (consulted on Feb. 16th, 2017)**. https://twiki.cern.ch/twiki/bin/view/Atlas/LuminosityForPhysics 113

[123] L. Barak, N. Berger et al.. **Search for new phenomena in diphoton events with the ATLAS detector at $\sqrt{s}$ = 13 TeV: Supporting documentation for the publication based on the 2016 data** . ATL-COM-PHYS-2017-148, 2017, Internal documentation. https://cds.cern.ch/record/2252642 119, 120, 133, 134, 139, 143, 145, 146

[124] L. Barak, N. Berger et al.. **$\sqrt{s}$ = 13 TeV in the mass range from 200 GeV to 3 TeV - Supporting documentation** . ATL-COM-PHYS-2015-1246, 2015, Internal documentation. https://cds.cern.ch/record/2056957 134

[125] Y. C. Yap. **Search for New Physics with Two Photons in the Final State with the ATLAS Detector**. CERN-THESIS-2016-259, 2017 126

[126] The ATLAS Collaboration, ATLAS HGamma Working Group. **Selection and performance for the $H \to \gamma\gamma$ and $H \to Z\gamma$**

analyses. ATL-COM-PHYS-2015-1326, 2015, Internal documentation. https://cds.cern.ch/record/2056957 112

[127] M. Fanti et al.. **Purity studies: Search for new phenomena in diphoton events with the ATLAS detector at $\sqrt{s} = \mathbf{13}\ TeV$**. ATL-COM-PHYS-2017-108, 2017, Internal documentation. https://cds.cern.ch/record/2245447 117

[128] CDF collaboration. **Search for new particles decaying into dijets in proton-antiproton collisions at $\sqrt{s} = \mathbf{1.96}\ TeV$** . Phys. Rev. D 79 112002, 2009 arXiv: 0812.4036 131

[129] CDF collaboration. **Search for new particles decaying into dijets in proton-antiproton collisions at $\sqrt{s} = \mathbf{1.96}\ TeV$** . Phys. Rev. D 79 112002, 2009 arXiv: 0812.4036 131

[130] ATLAS collaboration. **Measurement of Higgs boson production in the diphoton decay channel in pp collisions at center-of-mass energies of 7 and 8TeV with the ATLAS detector** . Phys. Rev. D 90 112015, 2014 arXiv: 1408.7084 131

[131] E. Gross, O. Vitells. **Trial factors for the look elsewhere effect in high energy physics**. Eur. Phys. J C70 (2010) 525. 137

[132] E. Gross, O. Vitells. **Estimating the significance of a signal in a multi-dimensional search**. arXiv:1105.4355 [astro-ph.IM] 137

[133] W. Verkerke, D. Kirkby. **The RooFit toolkit for data modeling**. Statistical problems in particle physics, astrophysics and cosmology. Proceedings, Conference, PHYSTAT05, Oxford, UK, September 12-15, 2005 arXiv: 0306116[physics] 126

[134] V. Bertone, S. Carrazza, J. Rojo. **APFEL: A PDF Evolution Library with QED corrections**. Comput. Phys. Commun., 2014. arXiv: 1310.1394 [hep-ph] 127, 128

[135] S. Carrazza, A. Ferrara, D. Palazzo, J. Rojo. **APFEL Web**. J. Phys. G42, 5, 2015. arXiv: 1410.5456 [hep-ph] 127, 128

[136] W. Buttinger. **Using Event Weights to account for differences in Instantaneous Luminosity and Trigger Prescale in Monte Carlo and Data**. ATL-COM-SOFT-2015-119, 2015, Internal documentation. https://cds.cern.ch/record/2014726 64

[137] G. Cowan et al.. **Asymptotic formulae for likelihood-based tests of new physics**. Eur. Phys. J. C 71 (2011) 1554, arXiv: 1007.1727 [physics.data-an] 137

[138] A. L. Read. **Presentation of search results: The CL_s technique**. J. Phys.G28 (2002) 26932704 137

[139] B. Bellazzini et al.. **R-axion at colliders**. arXiv: 1702.02152 [hep-ph], 2017 152