A Contribution to the ATLAS Open Data Collaboration at CERN

Candidate Number: 166992

May 20, 2020
Word Count: 11,145
Project Supervisor: Dr. Lily Asquith

Abstract

ATLAS Open Data Collaboration aims at giving open access to proton-proton collision data collected by the ATLAS experiment at the LHC, as well as release relevant tools needed for analysis to the public. A review of available resources from the Open Data educational platform, as well as analysis development and update of various signals has been carried out. This dissertation further contributes to the open access educational tools and helps lift the limited person-power problem and reproducibility issues imposed by the blue-sky research undertaken at CERN. This is done through the replication of *HZZ* and $H\gamma\gamma$ studies from the first research paper on the observation of the Higgs boson, as well as through the development of *ZZ* and heavy *ZZ* resonance analyses. A reproduction of selected results from four different papers published by the ATLAS Collaboration has been achieved. The analyses performed on *HZZ*, $H\gamma\gamma$, and *ZZ* decays isolated the required signal events to a reasonable extent, validating them on the new 13 TeV data set. These are ready to be implemented into the educational platform now. The heavy *ZZ* resonance analysis excluded the mass of the Kaluza–Klein Graviton *G* up to mass of 1000 GeV. With minor adjustments discussed in this dissertation, this analysis can also be implemented into the educational platform.

Contents

Preface

- **Chapter 1:** All work was written by me and includes review of background material from outside sources. All information is cited respectively. All of the Feynman diagrams in this chapter were drawn by me.

- **Chapter 2:** All work was written by me and includes a review of background material from outside sources along with a review of my previous research on ATLAS detector structure. All information is cited respectively.

- **Chapter 3**: All information was written by me and includes a review of Monte Carlo Simulations collected from outside sources. All information is cited respectively.

- **Chapter 4:** The analyses made on the 13 TeV Open Data set discussed in this chapter were performed using the code obtained from Mr. Meirin Evans' work on the 13 TeV data sets. The relevant configuration files and event selection requirements were made by me and are cited respectively where the information came from outside sources. All Feynman diagrams were obtained from outside sources and are cited respectively. All code used for this project is hosted on GitHub and can be found at the following link: https://github.com/StanislavBiryukov11/ATLAS-Open-Data-Project-BSc-

- **Chapter 5:** All the resultant graphs were created by me using Pythons' builtin Matplotlib visualisation module.

- **Chapter 6:** All information was written by me. The relevant sources are cited respectively where the results from previous studies for replication purposes are shown.

- **Chapter 7:** All information was written by me.

- **Chapter 8:** Bibliography shows respectively cited sources in the text.

- **Chapter 9:** The Results of the Appendix section were obtained by me. Where comparisons are made sources are respectively cited.

Chapter 1

1. Introduction

1.1 ATLAS Open Data

1.1.1 Collaboration Motivations

ATLAS, or A Toroidal LHC ApparatuS, is one of four main experiments at the Large Hadron Collider at CERN. Particle physicists at the ATLAS experiment search for evidence of particle physics beyond the Standard Model (SM), the main theory of fundamental forces that also classifies all known elementary particles. [1] This is done by smashing hadrons into each other, where in the case of the LHC the colliding particles are beams of protons. Once data have been collected, ATLAS Open Data "provides open access to proton-proton collision data at the LHC for educational purposes". [2] Some of the main aims of the ATLAS Open Data project are to "generate excitement and enthusiasm for fundamental research, inspiring physicists of the future", as well as "releasing data for education, training, and outreach activities". [3] One of the reasons for the importance of inspiring new physicists to undertake analytical work with the ATLAS Open Data resources is technological innovation. The skills developed through Open Data can then be transferred to other areas to induce further technological progress that is in public interest. One of such examples is the development of the World Wide Web (WWW). In addition to the WWW being one of the most important factors in everyday life for billions of people, this CERN development has contributed a staggering 1,672 billion dollars to the global gross domestic product (GDP). [4] Not only can ATLAS Open Data students induce technological innovation in other areas of science or undertake projects that provide substantial economic return, they can also use their newly formed skills and abilities to share innovative ideas and developments with the public. Since public acceptability highly depends on the trustworthiness of the source, Open Data students can have a high impact on the future of scientific communications. [5]

1.1.2 Complications

The research undertaken at ATLAS and the other three experimental detectors at CERN is sometimes described as blue-sky research. Blue-sky research has been defined as "research without a clear goal" and "curiosity-driven science". Hence, one of the big drawbacks of blue-sky research is lack of funding. Even though funding is not a big problem for ATLAS and LHC experiments, the possibility of not getting enough funding scares a proportion of scientists from engaging with this research, limiting the available person-power in experimental particle physics groups. As not too many scientists want to engage with blue-sky research, lack of transparency and completeness in reporting of data, methods, and analyses, as well as the lack of sufficient replication studies imposes a reproducibility issue. [6] In addition to this, limited person-power reduces the total available time spent researching or analysing, so providing students with more available resources becomes more challenging.

1.1.3 Educational Platform

The complications were attempted to be lifted by the ATLAS Open Data educational platform by providing easy access to the data, software and tools with documentation in the form of step-by-step instructions for users. [2] Providing easy access to data, simulations, tools, and software is ATLAS Open Data's attempt at fixing the reproducibility issue and person-power problem. Having a centralised location for all the necessary information to be able to carry out the analyses of given data allows for reproduction of results obtained by the students. In addition to this, other ATLAS particle physicists or students taking part in ATLAS Open Data that undertake their work in developing, upgrading or studying from the Open Data educational platform can have an easier access to all the previously collected material and results. This aids in making data and analyses reproducible, as well as reducing person-power needed to keep the platform up to date. By allowing the students to teach themselves by following procedures designed to maximise efficiency in learning outcomes, the students can reproduce known results and gain a better understanding of particle physics and analyses techniques with limited reproducibility issues.

1.1.4 Motivation for Dissertation

Even though limited person-power and reproducibility issues are addressed by the creation of the educational platform there is always room for improvement. By reviewing available resources from the ATLAS Collaboration, as well as helping in analysis development and updating for various signals, this dissertation further contributes to the open access educational tools and helps lift the limitations imposed by the blue-sky research undertaken by ATLAS Collaboration. The first three chapters are aimed at giving conceptual information, terminology and definitions to necessary background information needed for data analysis on the 13 TeV run of the LHC. The remaining chapters reproduce selected results from four different papers published by the ATLAS Collaboration.

1.2 Background Information on Particle Physics and Accelerators

1.2.1 Quark Parton Model and Quantum Chromodynamics

During the period of exploration of electron-proton scattering in particle physics, particle physicists came up with a different model to describe the constituents of protons. It was previously thought that the proton consists of three quarks, up, up, and down. It was then proposed that the proton consists not only of the three valence quarks described above, but also from an undetected sea of quarks and antiquarks, strongly interacting through the strong force carriers, the gluons. These three constituents, valence quarks, sea quarks (antiquarks) and gluons are all collectively known as the partons and can be seen in Figure 1.1. With time, particle physicists have discovered further particles; one of these high interest particles was the delta baryon Δ^{++}, which was found to have a quark composition of up, up, and up. According to Pauli's Exclusion principle, the total wavefunction of the system must be antisymmetric under particle exchange. Then, if each up quark has the same flavour state, the total flavour wavefunction will be symmetric under particle exchange. At the same time, it was found that the Δ^{++} baryon was the lowest mass member of its

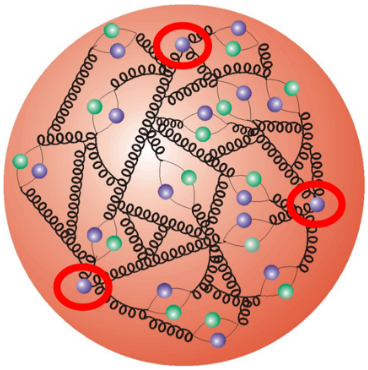


Figure 1.1: The proton structure in terms of partons. The circled points represent the three valence quarks, the black lines represent the gluons, and the non-circled points represented the virtual sea of quarks and antiquarks. [7]

multiplet, so its orbital angular momentum wavefunction is zero as well. In addition, the spin of each of the up quarks is the same and hence the spin wavefunction must also be symmetric under particle exchange. This suggested an overall symmetric system wavefunction, as can be seen in Figure 1.2, so a new quantum number was needed to bring the wavefunction back to antisymmetric. At this point the new quantum number known as colour charge was introduced. [7]

$$\Psi = \underbrace{\psi_{orbital}(\vec{x}) \cdot \xi_{spin}}_{\text{symmetric}} \cdot \underbrace{\chi_{colour}}_{\text{anti-symmetric}}$$

Figure 1.2: General wavefunction of a particle in terms of the product of orbital, spin, and colour wavefunctions. Flavour wavefunction is omitted but is symmetric under particle exchange. [7]

The colour charge wavefunction is comprised of the combination of the three colour charges: red, green and blue. It is important to note that all baryons and mesons have an overall colour charge of white, or neutral. Similar to positive charge cancelling out negative charge, the combinations of the three colour charges are assigned such that the overall colour charge wavefunction is neutral (white) in all baryons and mesons. In baryons, the combination of three colour charges are chosen due to there being three quarks, while for mesons a combination of two colour charges are chosen due to there being two quarks. The study of the strong interaction between quarks and gluons is called quantum chromodynamics (QCD), and it pays special attention to colour charge in particles. [8]

1.2.2 Quark Confinement

One of the interesting effects of colour charge in elementary particles is quark confinement. It is said that real particles always have a neutral or white colour charge, while gluons can also carry a colour charge themselves. As the only way to get multiple colour charges in the wavefunction is through the combination of different quarks' colour charges, the individual colour charges cannot by themselves make a neutral colour charge and hence cannot exist on their own. This prevents us from directly observing isolated quarks. [7]

1.2.3 Hadronization

Considering the proton-proton collisions made at the LHC, it is not actually the protons that collide with protons, but rather the partons colliding with other partons. As an example of the weak interaction, which may change the flavour of quarks, one quark from the proton might change to another. At the same time, some of the virtual quarks can fly out of the sea of virtual particles. However, due to quark confinement they are forced to combine with other quarks and gluons until a combination which gives a neutral colour charge is created. The process of quarks and gluons recombining into baryons and mesons is called hadronization, which is the main process behind the formation of jets. Jets are formed of many hadrons that are created due to hadronization of quarks and take the form of a cone. These jets can be used as the final decay products to infer properties of the quarks due to which the hadronization has occurred. [9]

1.2.4 Signal and Background Events

In particle physics, "events" describe the result of the interaction between subatomic particles. This includes the scattering of incoming particles or their destruction, sometimes followed by the creation of other new particles. Since we generally use very high energies at accelerators, the product particles can be created with a mass higher than that of the parent particles at rest. This is due to Einstein's famous mass energy equivalence relation and the mass-energy-momentum relation seen in Equation 1.1 in natural units:

$$E = \sqrt{(m)^2 + (\boldsymbol{p})^2} \quad . [10] \qquad (1.1)$$

While some product particles created will have high mass and hence decay very fast, others might have a lower mass but then will decay slower. In accelerators and particle detectors we can identify and analyse the heavier fast-decaying particles through their decay products. In addition, when collecting information on these kinds of events, we must be able to categorise them into signal events and background events. Signal events are those events that represent the results of interactions in which we are interested, while background events are all the events that have similar decay products to the signal but are not the result of interactions for which we are looking.

1.2.5 Signal to Background Ratio

Since we encounter around 1.7 billion collisions per second, we have huge data sets in experimental particle physics. [11] To be able to perform analysis on the signal events, we must first differentiate them from the background events. Then, to be able

$$Signal\ to\ Background\ Ratio = \frac{Number\ of\ Signal\ Events}{Number\ of\ Background\ Events}\ , \quad (1.2)$$

to come up with a reasonable result, particle physicists try to increase the signal to background ratio, where the higher the signal to background ratio is, the 'cleaner' the data set is. To achieve the highest signal to background ratios, particle physicists impose certain conditions on the collected data event parameters, but more on this in the Triggers section (2.4) and Event Selection section (4.3). By filtering out more and more unwanted background events, we can increase the ratio until a reasonable level is achieved for analyses. Various parameters can be accessed to increase the signal to background ratio, one set being the tracking parameters.

1.2.6 Tracking Parameters

1.2.6.1 Classical Parameters Representation

The ATLAS detector collects many different values of varying parameters. To understand these parameters, it is first necessary to describe the space in which collisions take place. The most general representation of parameters is the classical representation seen in Figure 1.3. [12] This makes use of five main classical parameters; l_1 and l_2 which are the parameters of local positions transverse to the beam pipe plane (corresponding to *x*-coordinate and *y*-coordinate in Cartesian system, respectively), θ which is the polar angle from the z-axis (in the range $0 \leq \theta \leq \pi$) , ϕ which is an azimuthal angle in the transverse plane (in the range $-\pi \leq \theta \leq \pi$), and q/p , which is the particles' charge to momentum ratio.

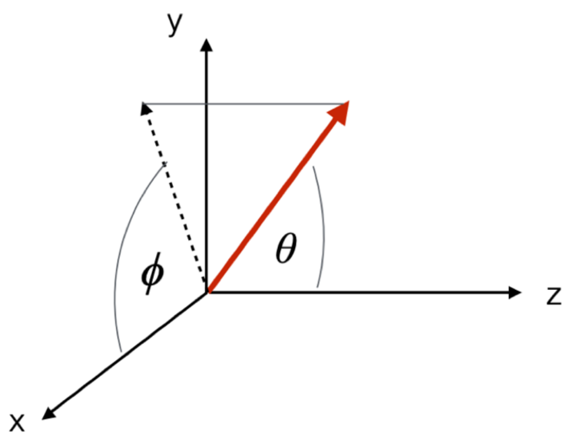


Figure 1.3: Representation of space, centred at the intersection between oppositely travelling beam pipes. Cartesian coordinates form a positional reference plane, while polar angles form a directional reference frame. [12]

1.2.6.2 Pedigree Parameters Representation

Another representation of parameters collected by the ATLAS detector is the Pedigree representation, which is a representation of the closest approach to the reference line (*z*-coordinate in Cartesian system). This makes use of some similar parameters as the classical representation (θ, ϕ, q/p), but with the substitution of d_0 and z_0 for l_1 and l_2.

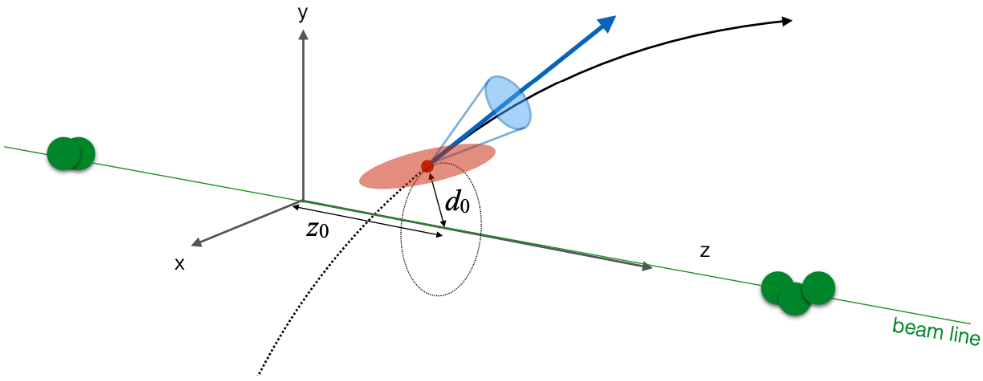


Figure 1.4: The Pedigree representation of closest approach parameters. The curved black line represents the track of the particle, the blue arrow represents its total momentum. Transverse and longitudinal parameters are also shown. [12]

The parameter d_0 is called the transverse impact parameter, while the parameter z_0 is called the longitudinal impact parameter. Both of these give a longitudinal and transverse reference for the particle coming out of a collision and can be seen in Figure 1.4. [12]

1.2.6.3 Rapidity and Pseudorapidity

To aid in tracking of a particle coming out of the collision, particle physicists have come up with another way to describe where the particle is in reference to the beam line. This time they have used a new coordinate in terms of the energy and momentum of the particle in the direction of the beam line, known as rapidity, which is given by

$$y \equiv \frac{1}{2} ln \left(\frac{E+p_L}{E-p_L}\right). \text{ [13]} \qquad (1.3)$$

The longitudinal momentum, or the beam line momentum in this case, is defined as the momentum in the direction of the beam pipe (positive z-direction in Cartesian coordinates). If the highly relativistic particle coming out of the collision is travelling in the plane transverse to the beam pipe, then the longitudinal momentum component will be very small, hence the natural logarithm approaches zero. But when the particle moves in the positive or negative beam pipe direction, then its energy will be roughly equal to its longitudinal momentum if the mass of the particle is much smaller than its momentum. Hence, the value of rapidity approaches either negative infinity for the particle moving in the negative beam pipe direction or approaches positive infinity if it moves in the positive beam pipe direction. It is important to notice that for highly relativistic particles for which we can approximate the energy as roughly being their momentum, it is the same as having massless particles whose energy is their momentum. For these kinds of particles, particle physicists use a different spatial coordinate that can be derived from the rapidity defined above. The new coordinate was named pseudorapidity, which is expressed in terms of the polar angle and can be seen in Equation 1.4 and Figure 1.5 respectively:

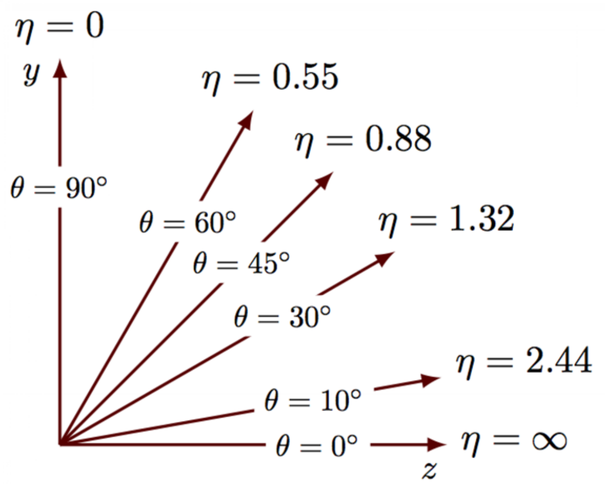


Figure 1.5: Diagram showing some values of pseudorapidity for their associated polar angle θ. The lines represent the direction to the particle from the beam pipe. Here y does not stand for rapidity but rather for the *y*-axis in Cartesian coordinates. [14]

$$\eta \equiv -ln\left[tan\left(\frac{\theta}{2}\right)\right] . [13] \qquad (1.4)$$

1.2.6.4 Angular Separation

Since in particle collision at the LHC we usually see more than one product particle coming out of the collision, it is important to be able to tell the angular separation between particles. The angular separation can be defined in two ways. The first is the angular separation for massive particles, which is expressed in terms of rapidity difference and azimuthal angle difference between two particles, as seen in Equation 1.5,

$$\Delta R \equiv \sqrt{(\Delta y)^2 + (\Delta \phi)^2} . [12] \qquad (1.5)$$

The other definition of angular separation is expressed in terms of the pseudorapidity difference and azimuthal angle difference between the two particles and can be seen in Equation 1.6,

$$\Delta R \equiv \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} . [12] \qquad (1.6)$$

This definition allows events containing particles with undesirable angular separations to be discarded in the analysis stage.

1.2.7 Missing Transverse Momentum

One of the universal conservation laws that applies to all interactions is the conservation of momentum. In a collider the beams of protons are accelerated in opposite directions to each other so the momenta of the colliding beams of protons will be in opposite directions. Since the momentum is equal to zero before the collision, and momentum before and after the collision must be equal, momentum after the collision must also be zero. By tracking and collecting all the transverse momenta of all the particles after the collision and adding them together it is possible to infer the presence of “invisible” particles such as neutrinos if the addition of momenta gives a non-zero value. This non-zero value of combined transverse momentum is called the missing transverse momentum and is a very helpful tool in identifying “invisible” particles that rarely interact with matter. [15]

1.2.8 Centre of Mass Energy

The same universal conservation law of energy-momentum leads to another useful definition. When talking about the available energy for particle creation after collisions, or in other words the available energy for physics experiments, physicists use a definition of the centre of mass energy in terms of the total energy of the system and the total momentum of the system before the collision. This can be seen in Equation 1.7,

$$\sqrt{s} = \sqrt{(E_1 + E_2)^2 - (\vec{p_1} + \vec{p_2})^2} \quad . [16] \qquad (1.7)$$

In the case of accelerators with beams of protons travelling in opposite directions, the momenta of beam 1 and beam 2 are equal in magnitude but opposite in sign, which causes the total momentum of the system to be zero. Hence for two proton beams of the same energy E the centre of mass energy is described by Equation 1.8,

$$\sqrt{s} = \sqrt{(2E)^2} = 2E \ . [16] \qquad (1.8)$$

This centre of mass energy is larger than that for the case of a stationary target, and hence particle colliders that collide two accelerated particles will give higher available energies for experiments and will allow for the formation of heavier and more interesting particles.

1.2.9 Luminosity and Cross-Section

To form heavier and more interesting particles a large number of collisions are needed. Instantaneous luminosity, denoted by $\mathcal{L}$, is a measurement of the number of collisions that can be produced in a detector per centimetre squared per second. This parameter can be described in terms of cross-section, σ, and the rate of events detected per second (time derivative of number of events N) through the relation

$$\mathcal{L}(t) = \frac{1}{\sigma}\frac{dN}{dt} \ . \qquad (1.9)$$

Integrated luminosity is denoted by $\mathcal{L}_{int}$ and is the integral of the instantaneous luminosity over the time the collider was active. This describes the total collected data set size over total run time. The integrated luminosity has the units of inverse femtobarns, fb^{-1}, where one fb^{-1} equals about 10^{12} events, and can be expressed in terms of the instantaneous luminosity and total run time T as

$$\mathcal{L}_{int} = \int_0^T \mathcal{L}(t)dt \ . [17] \qquad (1.10)$$

A particle's cross-section, σ , is a measure of the rate at which a certain process can occur. In other words, the cross-section of a particle gives the probability for an interaction (scattering) to take place and can be expressed in terms of the number of events, integrated luminosity of the experiment, and the detector efficiency factor ϵ. This relation is shown in Equation 1.11,

$$\sigma = \frac{N}{\mathcal{L}_{int}}\epsilon \quad . \tag{1.11}$$

Not only can the cross section be derived from N, $\mathcal{L}_{int}$, and ϵ, but it can also be derived in terms of the transition rate between particles' states. Such a relation is known as Fermi's Golden Rule and relates the transition rate between states to the transition matrix element, and the transition phase space. The transition matrix element amplitude is almost impossible to compute exactly, so particle physicists approximate it by asymptotic series expansion, similar to Taylors expansion. In addition, it is possible to approximate the matrix element amplitude and hence the cross-section of a process through the use of pictorial representations known as Feynman Diagrams, where labelling all vertices with the appropriate coupling constants corresponding to the different types of interactions may aid in approximating the cross-section. [18]

1.2.10 Decay Rates and Branching Ratios

Just like particle physicists use the cross-section observable σ to describe the probability of a scattering process to occur, they also use the decay rate observable Γ to describe the probability of a decay process to occur. The decay rate can also be approximately found from Fermi's Golden Rule and the use of coupling constants in Feynman Diagrams. From the decay rates it is possible to construct the branching ratios of certain processes, for example decays of a particle. The branching ratio gives the probability for the unstable particle to decay to "i-th final state" and can be given by the ratio of a particular decay mode rate Γ_i to the total decay rate (sum of all possible product particles' decay rates). This can be observed in Equation 1.12:

$$BR(i) = \frac{\Gamma_i}{\sum_i \Gamma_i} \; . \; [19] \tag{1.12}$$

1.2.11 Pile Up

At the LHC we do not just accelerate individual protons to collide with each other, but instead we choose to squeeze many protons into beam bunches. To increase the luminosity of the accelerator, and hence increase the chances that two protons will collide, particle physicists at LHC collide these beam bunches of protons at the beam pipe intersection point. A "pile-up" occurs when more than one proton-proton collision has taken place per bunch crossing. Even though one might think that increasing the number of protons per bunch might increase the chances that the protons will collide, and hence should increase the number of events for analysis, in reality more collisions also means more possible background events, so isolating the signal from the background may become more challenging. [20]

1.3 Background Information on Physics for Analysis

1.3.1 The $W^{\pm}$ and Z^0 Bosons

Examples of heavier and more interesting particles include the three force carriers (gauge vector bosons), W^+, W^- and Z^0, which are all mediators of the weak interaction. This interaction is unique because it allows for quarks to swap their flavour for one another. As protons are made up of quarks, in collisions at the LHC some quarks change flavour, some reform into other hadrons, and the weak interaction becomes an essential interaction to be considered while analysing data collected by ATLAS. [21]

The Z^0 boson is neutrally charged and is a spin-1 particle. Since the Z^0 boson is charge neutral, it can only share a vertex, or an interaction point, with two particles whose charge adds up to zero. [22] Hence the Z^0 boson generally shares a vertex with two oppositely charged fermions but can also share a vertex with two other vector bosons whose charge adds up to zero, or the W^+ and W^- bosons mentioned earlier. At the same time, flavour changing neutral currents (FCNC) are not allowed, hence the Z^0 bosons cannot share a vertex with differently flavoured fermions. The possible vertex combinations for the Z^0 boson can be seen in Figure 1.6. Out of the three fermion-antifermion vertices of the Z^0 boson in neutral current reactions, the lepton pair decay mode of the Z^0 boson decay leaves the cleanest signature, so particle physicists generally focus on detecting the leptons from neutral current reactions and using their properties to infer the presence of the parent. There are two main processes for the production of Z^0 bosons at LHC. The Drell Yan Process is the dominant production process for the Z^0 boson in which a particle-antiparticle (parton-antiparton in the case of deep inelastic scattering) annihilate each other to create the Z^0 boson. [23] The other mode of production of the Z^0 boson is quark-gluon scattering, in which the interaction between the valence quark and a parton gluon from the proton beams creates the Z^0 boson. [24] The Drell Yan processes are thought of as the cleanest and simplest channels for analysis as the only product particles are leptons that are easily identifiable in the detector, and this process happens about 65% of the time. The other 35 % of the identified Z^0 bosons come from quark-gluon scattering.

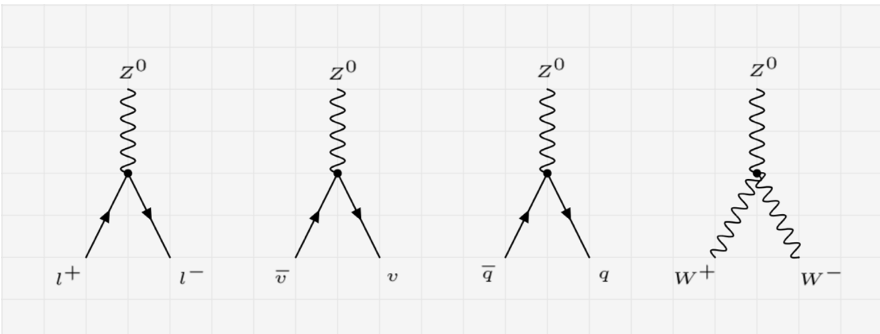


Figure 1.6: Feynman Diagrams of all possible vertices between the Z^0 boson and other SM particles, which includes a lepton-antilepton vertex with BR = (10.099 ± 0.011) %, neutrino-antineutrino vertex with BR = (20.00 ± 0.06) %,quark-antiquark vertex with BR = (69.91 ± 0.06) %, and triple gauge vertex (BR < 0.1 %) respectively left to right. The lepton-antilepton mode can decay either via electron-positron, muon-antimuon, or tau-antitau channels, each having a branching ratio of BR ~ 3.3 %. [25]

The W^+ and W^- bosons have positive and negative charge respectively and are also spin-1 particles. Since the $W^\pm$ bosons are charged, they can have a much higher variety of combinations of vertices than the neutral Z^0 boson. In 1963, Cabibbo proposed that down-type particles participating in weak interactions are not the down and strange quarks, but rather two separate linear superpositions of the down and strange quarks. According to this, vertices with same family quarks are Cabibbo favoured, vertices with cross-one family quarks are Cabibbo suppressed, and cross-two family quarks are doubly Cabibbo suppressed. In this instance suppressed does not necessarily mean it cannot happen. In reality, these processes still occur just at a much smaller rate. The possible vertex combinations for the $W^\pm$ boson can be seen in Figure 1.7.

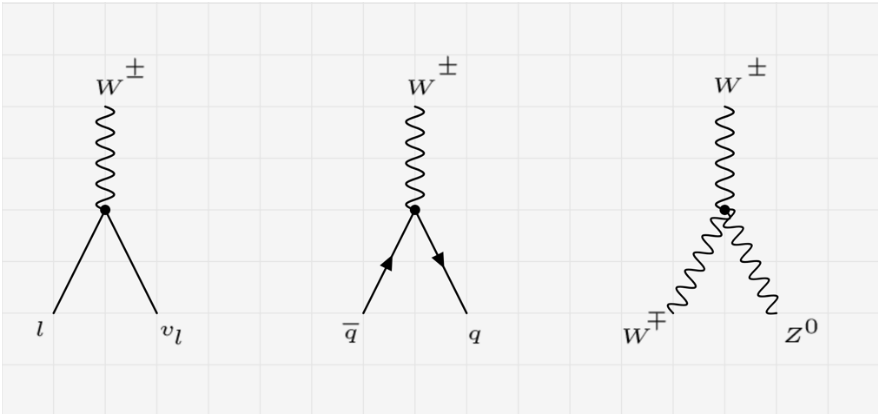


Figure 1.7: Feynman Diagrams of all possible vertices between the $W^\pm$ bosons and other SM particles, which includes lepton-neutrino vertex with BR = (32.72 ± 0.30) %, quark-antiquark vertex with BR = (67.41 ± 0.27) %, and a triple gauge vertex with BR < 0.1 % respectively left to right. The lepton-neutrino mode can decay either via same flavour electron-neutrino, muon-neutrino, or tau-neutrino pairs, each having a branching ratio of BR ~ 10.9 %. [26]

As the $W^\pm$ boson is responsible for transforming one quark flavour to another, or on a larger scale transforming a proton into a neutron or vice versa, the general production mode of the $W^\pm$ boson can be attributed to positive or negative beta-decay. In colliders, $W^\pm$ bosons are mostly produced through gluon induced processes, which are the natural production channels mostly considered as backgrounds to other modes. They can also be produced through top quark loops or direct diboson production, but more information on the diboson production can be found in the Dibosons section (1.3.3). It is important for particle physicists to choose the cleanest channels for analysis. As such, the leptonic decay mode of the W bosons is generally chosen as the cleanest channel since leptons are easily identifiable.

1.3.2 The Higgs Boson

The Higgs Boson (H^0) was first introduced to solve the problem of divergent contributions to the cross-section, or "infinities", from higher order Feynman Diagrams in electroweak processes, whilst keeping the gauge invariance intact. Without the Higgs boson, particles in the SM, including the $W^\pm$ and Z^0 bosons, should not be allowed to have mass, as mass itself is not gauge invariant. Hence, the prediction of a scalar (spin-0) particle by Robert Brout, Francois Englert, and Peter Higgs was necessary to fit the theory of massive particles in the SM. Not only was it necessary to predict the new particle, but it was also necessary to describe the mechanism in which particles in the SM gain mass. The SM particles obtain mass through the coupling to the BEH field vacuum expectation value (*vev*), which is non-zero. The more a particle interacts with the BEH field *vev*, the more massive it becomes. [27]

Since the Higgs boson couples to mass, the strength of the coupling depends on the mass of the particle to which it has coupled. As such, the Higgs boson does not directly couple to massless particles, like photons or gluons. Even though it cannot directly couple to gluons or photons, it can still indirectly couple to them, primarily through a top quark loop. The possible vertices for the Higgs boson can be seen in Figure 1.8, where the Higgs couples stronger to vector bosons than it does to fermions because the relationship of the coupling constant to the mass of the vector boson is quadratic while being linear in the case of the fermions. Since the branching ratio depends on the coupling strength of the interaction in the process considered, the highest branching ratios will come from heavy vector bosons, which means more events to be analysed. Due to this, the diboson processes are essential in the understanding of the Higgs boson and the BEH mechanism.

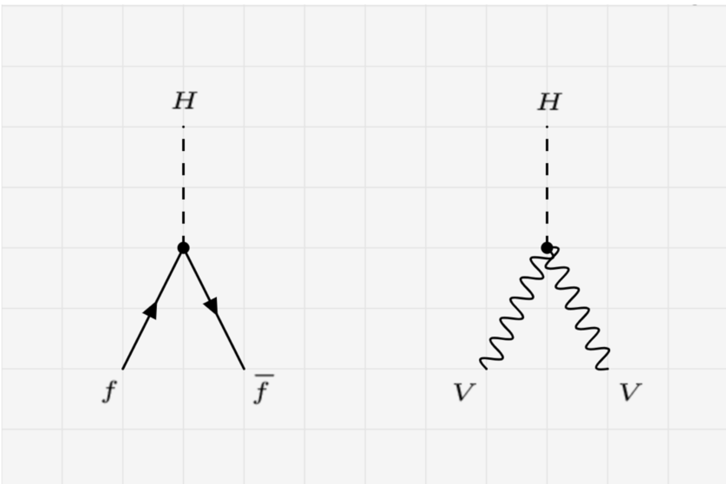


Figure 1.8: Feynman diagrams of all possible vertices for the Higgs boson coupling, including a fermion-antifermion vertex and a vector boson – vector boson vertex. [28]

Since the Higgs boson branching ratios depend on the mass of the coupled particle and the mass of the Higgs boson, of which both observables can only be obtained experimentally, the branching ratios are not fixed and can be seen in Figure 1.9. Even though the highest branching ratio is for the bottom-antibottom decay of the Higgs, this decay mode has the highest background due to the hadronization of the heavy bottom quarks and the jet pollution of the signatures in the detector. Hence, this decay mode will have a lower signal to background ratio than other decay modes, but some interesting analysis can still be made if certain selection requirements are applied to increase the signal to background ratio. On the other hand, the photon-photon (diphoton) decay of the Higgs has a small branching ratio of BR ~ 0.2 %. Yet this decay mode has substantially less background processes. Hence, its signal to background

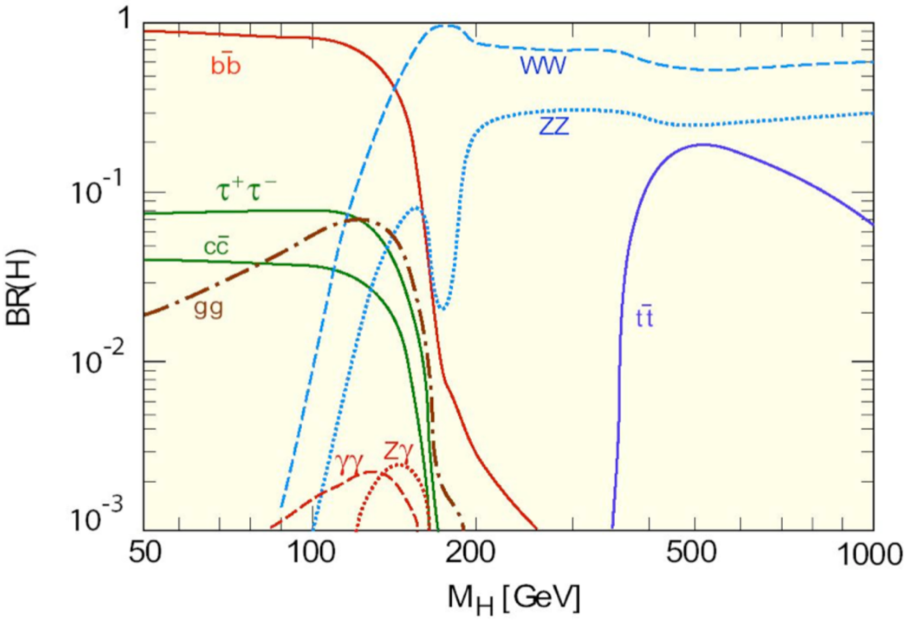


Figure 1.9: Branching Ratios for the decay processes of the Higgs boson are shown. For the Higgs at an observed mass of m_H = (125.1 ± 0.14) GeV, the branching ratios are BR ~ 57.8 % for the bottom-antibottom pair, BR ~ 21.6 % for the $W^{\pm}$ boson pair, BR ~ 8.6 % for the gluon pair, BR ~ 6.4 % for the tau-antitau pair, BR ~ 2.9 % for the charm-anticharm pair, BR ~ 2.7 % for the Z^0 boson pair, BR ~ 0.2 % for the photon pair, and BR < 0.1 % for the top-antitop pair and the Z^0 boson-photon pair. [28]

ratio should be higher and the analysis should be more certain. One of the examples already discussed was the clean signal nature of the Z^0 boson decay to a lepton pair. Since the Higgs can decay via two Z^0 bosons, the *Z*-diboson decay process is a good mode of decay for analysis of the Higgs and the BEH mechanism.

1.3.3 Dibosons

Even though diboson-decaying processes such as the Higgs diphoton decay or Higgs *Z*-diboson decay give particle physicists information on particles in the SM, it is predicted that many other non-SM particles can decay into dibosons. Because of this, diboson processes are essential for exploration of new particle physics beyond the SM. The diboson processes considered in this dissertation are the Higgs decay via two photons, the Higgs leptonic decay via two Z^0 bosons, and the Z^0 boson pair decay to four leptons seen in table 1.1.

Table 1.1: Symbolic representation of diboson processes considered for analysis.

Diboson Processes Considered:	Symbolic:
Higgs Decay via $\boldsymbol{\gamma\gamma}$ pair	$H \rightarrow \gamma\gamma$
Higgs Decay via *ZZ* pair	$H \rightarrow ZZ \rightarrow l^+l^-l^+l^-$
ZZ decay	$ZZ \rightarrow l^+l^-l^+l^-$

1.3.4 Resonance

A resonance is a large peak in cross section when plotted against the energy of the incoming particles. The resonant particle could be understood as the excited level or bound state of two colliding particles. If there is no such state, two particles bounce from each other and the cross-section has no peaks. But if there is such a state, then particles "stick" to each other, which in turn increases the cross-section at a certain peak invariant mass of the product particles. It is therefore possible to find the intrinsic width of the resonance through the peak of the cross-section value, which then allows particle physicists to characterise the interaction, and hence obtain information on the type of resonant particle even if no direct observation can be made. [29] Because of this, heavy diboson resonances can be utilised in search for physics beyond the SM and are especially valuable in search for the hypothetical force carrier of the Gravitational Interaction (considered in this dissertation) that would allow to take a step towards unification of all four fundamental forces of nature. In addition to resonance being utilised in search for physics beyond the SM, it can also be used to verify or revalidate SM particles. The $W^{\pm}$, Z^{0} and Higgs bosons are all examples of resonances that particle physicists have already discovered. The resonant process considered in this dissertation is the heavy *ZZ* resonance seen in Table 1.2.

Table 1.2: Symbolic representation of resonant diboson processes considered for analysis.

Resonant Processes Considered:	Symbolic:
Heavy *ZZ* resonance	$X \rightarrow ZZ \rightarrow l^{+}l^{-}l^{+}l^{-}$

Chapter 2

2. The ATLAS Detector

2.1 Inner Detector

The inner detector is the closest part of the ATLAS detector to the intersection point of the two beampipes. The inner detector is immersed in a 2 T magnetic field and covers the pseudorapidity range of $|\eta|$ < 2.5. [30] It is comprised of pixel detectors, semiconductor trackers and transition radiation trackers, which can all be seen in Figure 2.1. All of these parts give particle physicists information on the tracks of the charged particles coming out of the collisions, which can be used to calculate the momenta of the particles. However, the different parts of the inner detector work on different bases. Pixel detectors and semiconductor trackers use similar techniques to track particles. Pixel detectors track particles by silicone pixel plates that work similar to digital cameras, while semiconductor trackers use semiconductor plates; when the charged particles go through the detector, they hit the semiconducting layer that in turn releases electrons. [31] The electrons start to flow under the presence of a potential in the wire and are pushed towards the readout sensor for analysis. On the other hand, the transition radiation tracker is comprised of tubes filled with gas. While passing through the semiconducting materials with different resistances between the tubes, charged particles produce transition radiation in the form of photons, which is related to the particular energy of the particle. [32] After the collision, strong magnetic fields redirect particles' paths and their tracks can be reconstructed depending on which tubes or plates were hit at what time. From the bend in the tracks of charged particles it is possible to identify the charge to mass ratio, and hence help identify the momenta of the particles involved. The transition radiation gives even further energy information on the particles, which along with their momenta can be used to calculate the particles' mass and hence identify them.

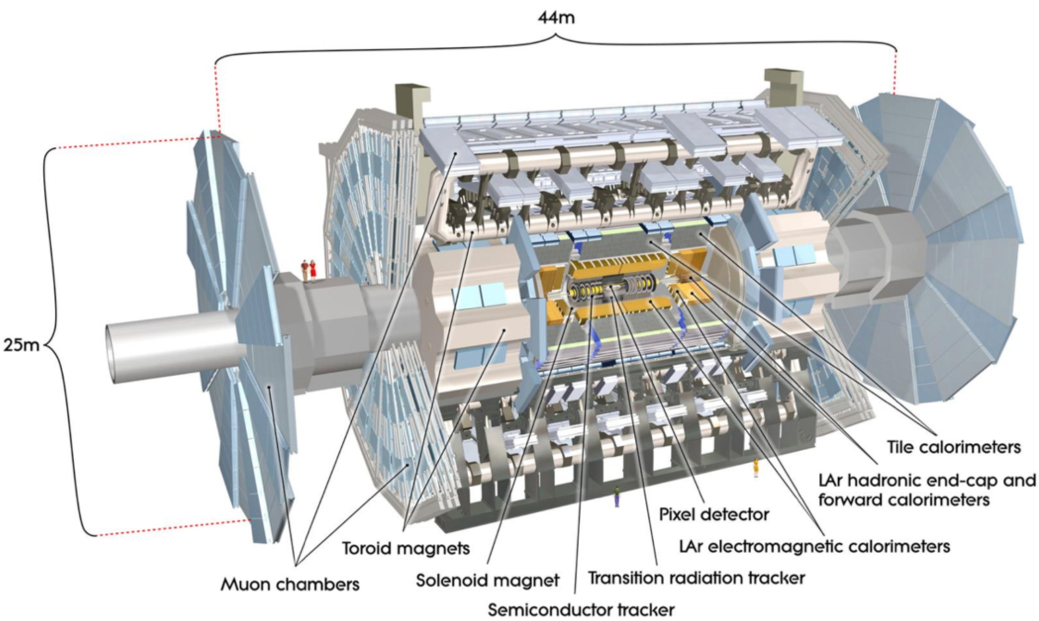


Figure 2.1: Computer generated 3D model of the ATLAS detector and its constituents. [33]

2.2 Calorimeters

The next closest parts of the ATLAS detector to the beam pipes are calorimeters. Calorimeters are used in ATLAS to measure the energy different particles lose as they pass through them, encompassing the pseudorapidity range of $|\eta|$ < 4.9. They are usually made of materials that are good at absorbing energy, such as lead, solid-lead glass or liquid argon. There are two types of calorimeter used in ATLAS; the Hadronic Calorimeter (Hcal) and Electromagnetic Calorimeter (Ecal). The Hcal is used to determine the energy of hadronic particles originating from proton collisions after their interactions with atomic nuclei of the Hcal material. Scintillator Plates are plates made of material that produces light when hit by the charged particles coming through them. These plates are put together to form tiles that are used in the Tile Calorimeter (TileCal, part of Hcal), where the light is collected from the tiles and analysed using photo sensors such as photomultipliers to determine the particles' energies. Electromagnetic calorimeters (Ecal) are used to determine the energies of electromagnetically interacting particles that come out of the collisions. This is primarily done by the Liquid Argon (LAr) Calorimeter, where the energised charged particles go through the sea of liquid argon atoms and ionize them. After ionization, due to an electric field that is present in the calorimeter, the electrons produced are pushed to the readout sensor for analysis. By knowing the number of electrons captured it is possible to back track the energy deposits in the Ecal or Hcal, which in addition to tracking information helps particle physicists identify particles. [34] The hadronic and electromagnetic calorimeters can be seen in Figure 2.1.

2.3 Muon Spectrometer

The outermost part of the ATLAS detector is called the Muon Spectrometer. The Muon Spectrometer encompasses the pseudorapidity range of $|\eta|$ < 2.7 and was included in the ATLAS detector for particles that can pass the calorimeters undetected, such as muons. For this reason, the Muon Spectrometer is located at the ends of the detector as the muons will be the only possible particles reaching that far without being stopped in other parts of the detector. The Muon Spectrometer is composed of tubes, similar to the transition radiation tracker. When a muon passes through the tubes, which are filled with gas and have a gold wire embedded in their centres, the muon rarely interacts with the matter around it but leaves a trail of charged ions and electrons. This trail happens when a particle is travelling faster than light through a given medium and is called the Cherenkov radiation. The trail of radiation drifts towards the wires in the centre. By detecting the starting point of these trails, and by knowing how much time it took them to reach the wire, it is possible to determine the position of the particle as it passed through. In addition, by knowing at which angle the Cherenkov radiation was emitted, it is possible to find the velocity of the particle, hence combining the velocity information with the momentum information, it is possible to discover the mass of the particle and help identify it. [35] The muon spectrometer can be seen in Figure 2.1.

2.4 Triggers

As stated in Signal to Background Ratio section (1.2.5), there are about 1.7 billion collisions happening every second at the centre of the ATLAS detector. It is therefore essential for particle physicists to filter out or cut unwanted events if they have nothing interesting to offer. Triggers are introduced into the detector that allow particle physicists to isolate interesting events by applying conditions on certain parameters such as number of final state particles or energies of particles coming out of proton-proton collisions. For these reasons the three-level trigger system known as The Trigger and Data Acquisition (TDAQ) system was created. [36] Even though applying cuts on certain parameters may increase the signal to background ratio, it is a common problem that some relevant information gets cut along with non-relevant information. This in turn causes concern as evidence of new physics can be cut from analysis altogether. As such it is of great importance in particle physics to choose the correct triggers for the correct experiments.

Chapter 3

3. Monte Carlo Simulations

3.1 Importance of MC Simulations

To be able to perform some kind of reasonable physical analysis on the data collected by ATLAS, particle physicists need to not only maximise the signal to background ratio, but also compare the results of the analysis with simulated data. The simulations used in particle physics are called Monte Carlo (MC) Simulations. These MC simulations can model the results of interactions between particles and their probabilities of occurring. [37] Since we can describe particles in quantum mechanical wavefunction terms, the simulations and the real data both possess probabilistic characteristics and hence can be compared to each other. MC simulations are created independent of luminosity of the experiment, and therefore can be scaled to have the same number of simulated events as the experimentally collected data events. This technique is known as luminosity weighting. [38] It is therefore essential to use these MC simulations to validate theoretical models through direct comparison of simulated and real events.

3.2 Stages of MC Simulation

MC simulation data sets used in particle physics are created in steps, where the first step is the event generation stage. In this stage, programs such as Pythia (next-to-leading order NLO generator), Powheg (NLO), Herwig (NLO), Sherpa (next-to-next-to-leading order NNLO generator) [39] or others generate the resultant product particles from considered interaction events (both signal and background), where all necessary parametrisation information about parent particles is known and fed into the generator. Since probabilities are essential in these simulations, and the most important probabilistic quantity of an event is its cross section, cross sections for simulated events have to be corrected to higher and higher orders to reduce their uncertainties. NNLO generators are corrected to a higher order than NLO generators, and hence give the best certainty for event generation. The most widely used event generators for QCD event simulations for the ATLAS detector are the NLO generators Pythia and Powheg, but NNLO generators such as Sherpa are more frequently used now as they become more accessible. [40] Most event generators simulate different phenomenological processes such as hard processes, parton showers, secondary decays, hadronization, multiple scattering, and hadron decays. [41] Some of these phases can be seen in the sketch of a simulated event in Figure 3.1. Different event generators focus on simulating different phases, where some are better for certain processes than others. It is therefore important for particle physicists to choose MC data sets that are simulated with the appropriate event generator to maximise the analysing capabilities in their experiments.

Detector simulation follows the event generation stage. In this stage, the simulation helps to model what happens to those final state particles simulated in the event

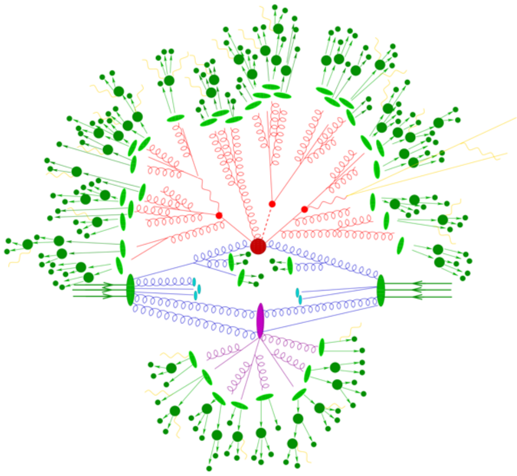


Figure 3.1: A general sketch of the outcome of a proton-proton collision simulated by a Monte Carlo event generator. The centred red blob corresponds to the first hard scattering (between two partons of the protons), the red tree-like structure originating from the centred red blob corresponds to parton showers, the purple blob corresponds to the second hard scattering (between other partons of the original protons), parton-to-hadron transitions are indicated by light green blobs, hadron decays are indicated by dark green blobs, and the yellow lines indicate soft photon radiation. [37]

generation stage as they pass through and interact with different components of the ATLAS Detector. For this reason, particle physicists model not only different parts of the detector, but also model the phase space of the detector including wires, pipes, and even detector limitations and imperfections. Geant4 is the main program used in ATLAS detector simulation; the output consists of a long list of locations of different components of the detector and materials they are made of, values of energy deposits in different components of detector, times over which the final state particles have deposited the energy, etc. This information is useful, but it still requires some tuning to be applicable to the actual experiments. [42]

Event reconstruction is the final stage of the MC simulation. This stage allows particle physicists to simulate the signatures different final state particles from the initial collision leave in different parts of the detector, such as bent tracks in the Inner Detector trackers, energy deposits in electromagnetic and hadronic calorimeters, or bent tracks and energy deposits in the muon spectrometers, along with interactions with imperfections of the detector. This is done by taking all those quantities simulated in the detector simulation stage and turning them into quantities that can be given by the detector itself. These quantities usually involve voltages, times, and currents but are highly dependent on the detector in use. To be able to view the signatures in the detector from these simulations, event displays are generally used. The main event display used at the ATLAS detector is called ATHENA, which on its basic level manages almost all production workflows specified in this section, such as event generation, simulation, and reconstruction. [42]

3.3 Systematic Uncertainties

As the phase space of the detector and the detector limitations are programmable in the detector simulation stage, and as the particle physicists can use multiple event generators specific to their experiment, the difference in the outcomes of the different simulations get assigned to the systematic uncertainties. This gives particle physicists insight into the cause of systematic uncertainties in their experiments depending on the simulation programs used, and also allows for a deeper discussion of analysis. In this dissertation no systematic uncertainties were considered as they are not provided by ATLAS Open Data, hence only graphical observations were made. [43]

3.4 Detector Signatures

The detector signatures from real particles coming out of proton-proton collisions will be very similar to the simulated signatures of the final state particles from the event generation stage. It is therefore important for particle physicists to be able to categorise the type of a particle based on the signature in the detector. Some SM particles' signatures can be observed in Figure 3.2. Muons barely interact with matter around them and usually show signatures in the muon spectrometer but are charged so do show some tracks in the inner detector (small depositions in the Ecal are also possible). Photons are massless and leave no tracks in the inner detector but show themselves by energy deposits in the Ecal. Protons do have tracks in the inner detector, along with some signatures in both the Ecal and Hcal. Neutrons are neutral and hence have no tracks in the inner detector but show themselves by energy deposits in the Hcal. Electrons are charged and will leave tracks in the inner detector, as well as leaving signatures in the Ecal. Neutrinos are almost massless and are neutral hence hardly interact with the detector whatsoever. For this reason, the presence of neutrinos is attributed to the missing transverse momentum discussed in Missing Transverse Momentum section (1.2.7). [44] A more precise description of signatures for SM particles can be seen in Table 3.1.

Table 3.1: Description of SM particles and their signature in the ATLAS detector.

Particle:	Signature:
Electron/Positron	Track in the inner detector + energy deposits in Ecal
Photon	No track in the inner detector + energy deposits in Ecal
Muon/antimuon	Track in the inner detector + small energy deposits in Ecal + muon spectrometer track
Tau/antitau	Multiple tracks in the inner detector + energy deposits in Hcal
Neutrinos	Missing Transverse Momentum
Hadrons	Tracks in the inner detector (if charged) + energy deposits in Hcal
Quarks	Seen as jets of hadrons: multiple tracks in the inner detector + energy deposits in Hcal

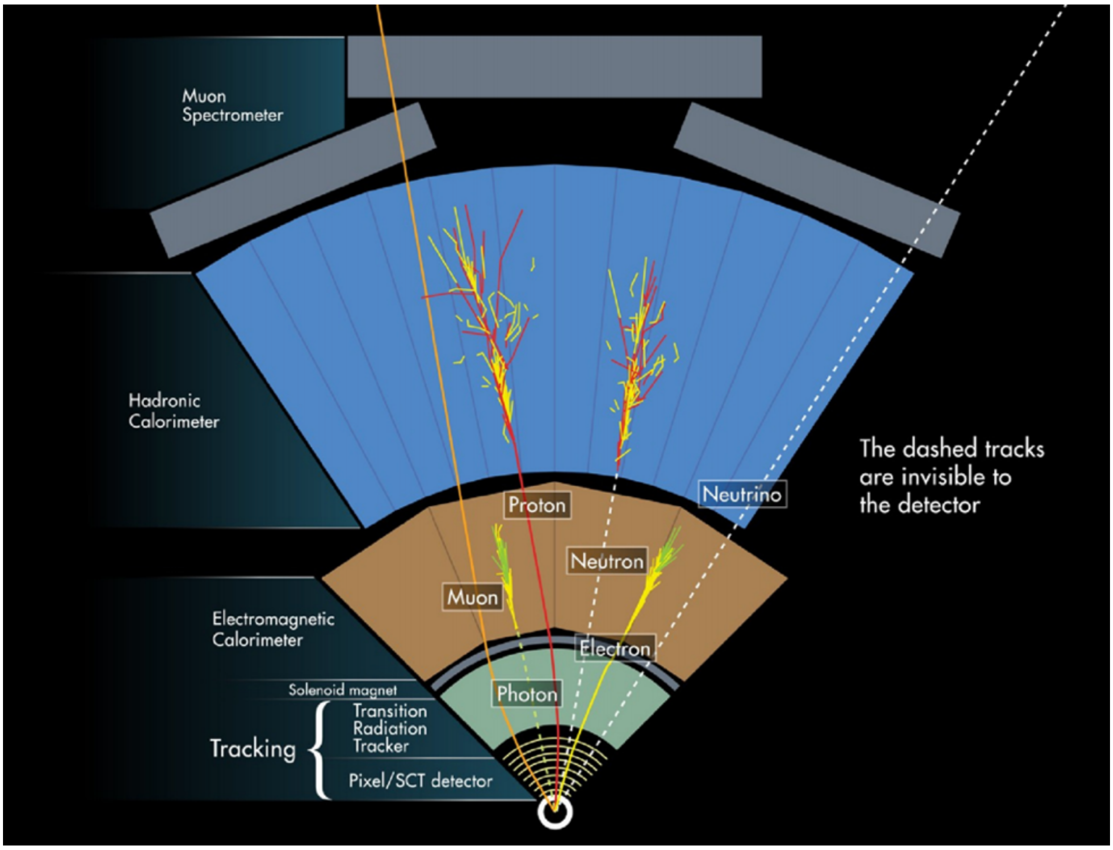


Figure 3.2: Front view of the ATLAS detector, showing Hcal and Ecal, as well as tracking detectors and muon spectrometer. Muon, photon, proton, neutron, electron, and neutrino signatures are shown respectively. [44]

Chapter 4

4. Methodology

4.1 Tools for Analyses

To perform analysis on diboson events described in the Dibosons section (1.3.3) and the resonant events described in the Resonance section (1.3.4) various analysis tools were used. All the analyses were completed in Python by calling upon ROOT functions and commands. Python is an object-oriented programming language that allows users to work quickly and integrate systems more efficiently than other languages such as C++ or JavaScript. [45] ROOT is a modular scientific toolkit which provides scientists with all the functionality needed to perform big data processing, statistical analysis, visualisation and storage. [46] Together Python and ROOT create an analytical software backbone that is essential for the analyses made by particle physicists at CERN and students undertaking part in the ATLAS open data project. The other tools necessary for analysis for new students taking part in the ATLAS open data project are event displays. The two easiest event displays provided to new students by the International Masterclasses group of the ATLAS open data project are Hypatia and Minerva. [47],[48] They work very similarly to the way ATHENA works as described in the Stages of MC Simulation section (3.2) but have a more simplistic interface that allows for easier identification of particles and their signatures in the detector. In addition to the abovementioned tools for analyses, CERN's LXPLUS cluster service was used to obtain necessary data and MC simulation files for the analyses discussed in Dibosons (1.3.3) and Resonance (1.3.4) sections. This Public Login User System is a cluster of computers provided by the IT department of CERN, which allows for interactive work on virtual machines accessible from your own computer. [49] The LXPLUS service can also give access to data collected by different experiments but requires valid CERN member credentials to access this information.

4.2 International Masterclasses Review

The first contribution to the ATLAS open data project originates from the review made on background information and tools provided to new students by International Masterclasses. This work is carried out by review of background material on the Z-path and W-path (Z^0 and $W^{\pm}$ bosons respectively), along with the review, comparison, and features list of Hypatia and Minerva event displays provided to students by ATLAS International Masterclasses. [21]

4.3 Event Selection

As described in the Signal and Background Events section (1.2.4) and Triggers section (2.4), it is of great importance to particle physicists to isolate signal events and increase the signal to background ratio through the use of triggers and application of cuts on unwanted events. This cutting is done by applying event selection requirements to the data sets, both MC simulated data and real data. The preliminary selection requirements that is applied to clear out the biggest backgrounds in most ATLAS open data analyses and this dissertation can be seen in Table 4.1.

Table 4.1: Preliminary selection requirements for standard event-based ATLAS Open Data analyses on 8 TeV data samples. [50]

General Selection:	Requirements:
A single electron or muon trigger has fired	trigE = True trigM = True
The primary vertex has at least 5 tracks	HasGoodVertex = True
There is at least one good lepton with $\boldsymbol{p_T}$ > 25 GeV	lep_IsTightID = True lep_pt > 25 GeV
Leptons are required to be isolated	lep_ptcone20 < 0.15 lep_ptcone30 < 0.15
The event passes the Good Run List (GRL)	passGRL = True

4.3.1 $H\gamma\gamma$ Analysis

The $H\gamma\gamma$ analysis focuses on the diphoton channel for the decay of the Higgs boson via a top quark loop, which can be seen in the Feynman diagram of Figure 4.1. The specific preselection and selection requirements were obtained from the research paper on the "Measurements of Higgs boson properties in the diphoton decay channel with 36 fb^{-1} of pp collision data at $\sqrt{s}$ = 13 TeV with the ATLAS detector". [51] The preselection conditions require that all photons have a transverse energy of E_T > 25 GeV with a maximum pseudorapidity of 2.37, excluding the transition region between the barrel and endcap calorimeters of 1.47 < $|\eta|$ < 1.52. This analysis requires two good reconstructed and well isolated photons that trigger the event. Photon kinematics suggest the first and second reconstructed photons must have transverse momenta p_T > 40 GeV and p_T > 30 GeV, respectively. Converted photons from any part of the detector are excluded, and the diphoton invariant mass window is chosen to be 100 GeV < $m(\gamma\gamma)$ < 160 GeV.

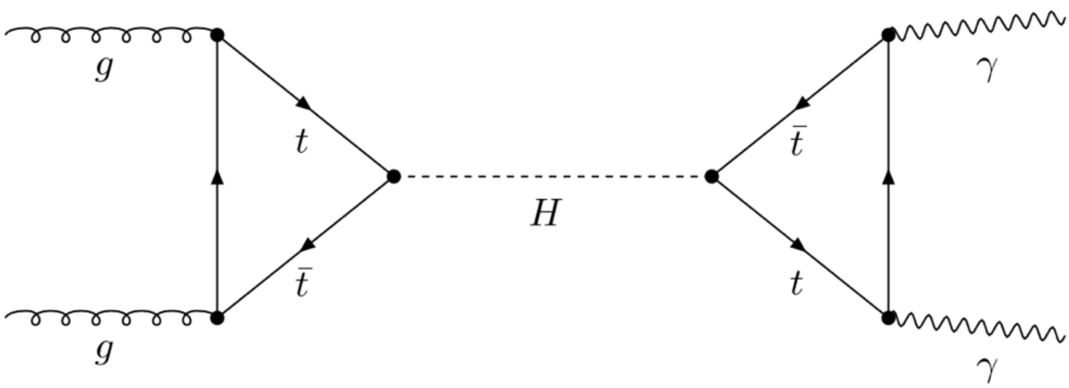


Figure 4.1: Feynman diagram for the ggF production and diphoton decay of the Higgs boson via top quark loop originating from proton-proton collisions at CERN. [52]

Table 4.2: Specific selection requirements for the Higgs diphoton decay analysis. [51]

Specific Selection:	Requirements:
Preselection: pseudorapidity of photons must be $\|\eta\|$ < 2.37, excluding 1.37 < $\|\eta\|$ < 1.52 with transverse energy E_T > 25 GeV	np.absolute(photon_eta) < 2.37 photon_pt > 25 np.absolute(photon_eta) < 1.37 np.absolute(photon_eta) > 1.52
Two photons must be well isolated	photon_n = 2 photon_ptcone20 < 4 GeV for both photon_ptcone30 < 4 GeV for both
No converted photons in any part of the detector	photon_convType = 0 for both
First and second photons must have transverse momentum of p_T > 40 GeV and p_T > 30 GeV	photon_pt[0] > 40 GeV and photon_pt[1] > 30 GeV
Reconstructed diphoton invariant mass must be in the mass window 100 GeV < $m(\gamma\gamma)$ < 160 GeV	100 GeV < $m(\gamma\gamma)$ < 160 GeV
Good reconstructed photons trigger the event	photon_trigMatched = True for both photon_isTightID = True for both

4.3.2 *HZZ* Analysis

The *HZZ* analysis focuses on the *Z*-diboson leptonic decay $H \rightarrow ZZ \rightarrow l^+l^-l^+l^-$, hence fully leptonic decay of the Z^0 bosons is considered. The Feynman diagram for the process considered can be observed in Figure 4.2. The specific cuts applied in this analysis can be seen in Table 4.3. These were obtained from the paper on the "Measurement of inclusive and differential cross sections in the $H \rightarrow ZZ^* \rightarrow 4l$ decay channel in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector". [53] It is important to mention that even though leptonic decay of the *Z*-diboson process is of interest, the Z^0 boson to tau lepton pair decay is not considered as the tau lepton is too massive and too unstable to be quickly detectable on its own, hence events with tau pairs are cut out of both MC and real data sets to increase the signal to background ratio in this analysis. The leptonic channels considered are the $4e$, $2e2\mu$, and 4μ channels. Two individual Same Flavour Opposite Sign (SFOS) lepton pairs are chosen for this analysis. The lepton pair with reconstructed invariant mass closest to that of the Z^0 boson is called the leading pair, and the remaining lepton pair is called the subleading pair. Lepton kinematics suggest first, second, and third highest p_T lepton must have transverse momenta p_T > 20, 15, and 10 GeV respectively. Each good muon (electron) that triggers the detector must also have p_T > 5 GeV (7 GeV) and $|\eta|$ < 2.7 (2.47). The invariant mass of the four leptons must be in the mass window 115 GeV < $m(4l)$ < 130 GeV.

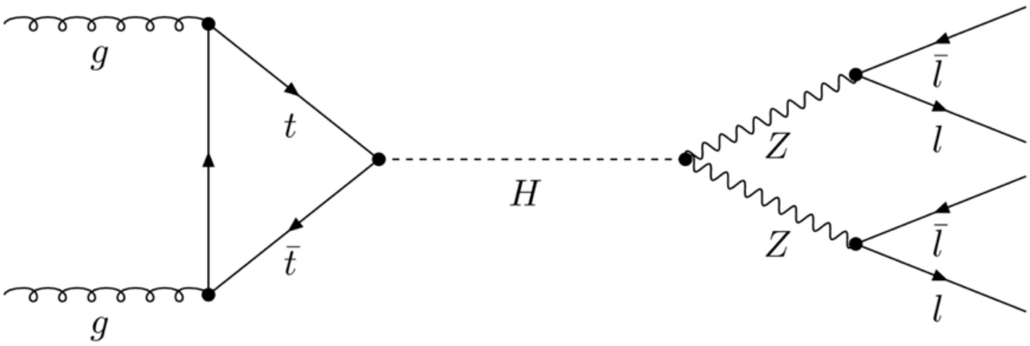


Figure 4.2: Feynman diagram for the gluon-gluon fusion (ggF) production of the Higgs boson and its leptonic decay via *Z*-diboson originating from proton-proton collisions at CERN. [52]

Table 4.3: Specific selection requirements for the Higgs *Z*-diboson decay analysis. [53]

Specific selection:	Requirements:
Two pairs of oppositely charged leptons ($\boldsymbol{4e}$, $\boldsymbol{4\mu}$, and $\boldsymbol{2e2\mu}$ channels are chosen)	lep_n = 4, lep_type = 11 for each lepton ($4e$), or lep_type = 13 for each lepton (4μ), or lep_type = 11 for two leptons and lep_type = 13 for other two leptons ($2e2\mu$)
Each muon (electron) must have $\boldsymbol{p_T}$ > 5 GeV (7 GeV) and $\lvert\boldsymbol{\eta}\rvert$ < 2.7 (2.47)	np.absolute(lep_pt) > 5 GeV (7 GeV) and np.absolute(lep_eta) < 2.7 (2.47) for muons (electrons)
First, second, and third highest $\boldsymbol{p_T}$ leptons must have $\boldsymbol{p_T}$ > 20, 15, and 10 GeV respectively	lep_pt[0] > 20 GeV and lep_pt[1] > 15 GeV and lep_pt[2] > 15 GeV
SFOS lepton pairs with reconstructed invariant mass closest to that of the Z^0 boson (leading pair), must have 50 GeV < $\boldsymbol{m_{12}}$ < 106 GeV	Smallest $\lvert m(Z^0) - m(ll)\rvert$ with 50 GeV < m_{12} < 106 GeV and $m(Z^0)$ = 91.18 GeV
Remaining SFOS lepton pairs with reconstructed invariant mass closest to that of the Z^0 boson (subleading pair), must have 12 GeV < $\boldsymbol{m_{34}}$ < 115 GeV	Remaining smallest $\lvert m(Z^0) - m(ll)\rvert$ with 12 GeV < m_{34} < 115 GeV and $m(Z^0)$ = 91.18 GeV
Invariant mass of the four leptons must be in the mass window 115 GeV < $\boldsymbol{m(4l)}$ < 130 GeV	115 GeV < $m(4l)$ < 130 GeV

4.3.3 *ZZ* Analysis

The *ZZ* analysis focuses on the natural leptonic decay of the *Z*-diboson, $ZZ \to l^+l^-l^+l^-$, whose individual *Z⁰* boson decay Feynman diagrams are represented in Figure 4.3. Understanding the *ZZ* process is important as it is a main background in the *HZZ* and Heavy *ZZ* Resonance analyses. In this analysis the two Z^0 bosons must not have originated from another event such as the decay of the Higgs boson discussed in the *HZZ* Analysis section (4.3.2); the signal events must have two individual Z^0 bosons decaying leptonically. Once again, the tau lepton pair decay of the Z^0 boson decay is not considered similarly to the *HZZ* analysis due to the unstable nature of the tau lepton decay and its difficulty in signal isolation. The specific selection requirements were applied from the research paper on the "Measurement of the *ZZ* Production Cross Section in pp Collisions at $\sqrt{s}$ = 13 TeV with the ATLAS Detector" [54] and can be seen in Table 4.4. The preselection conditions require exactly four good leptons, and that the two reconstructed Z^0 boson candidates' masses from SFOS lepton pairs must have minimum deviation from the accepted Z^0 boson mass of $m(Z^0)$ = (91.1876 ± 0.0021) GeV. [55] The specific event selection for this analysis then requires that all leptons must be energetic with minimum transverse momentum of 20 GeV. All electrons must have a maximum pseudorapidity of 2.47, as well as at least one muon in the 4μ channel having a maximum pseudorapidity of 2.4 corresponding to the acceptance of the muon trigger. All other muons must have a maximum pseudorapidity of 2.7, while the maximum transverse impact parameter significance must be 6.5 for electrons and 3.5 for muons. In addition to this, each reconstructed lepton pair invariant mass is chosen to be in the mass window 66 GeV < $m(ll)$ < 116 GeV. The four-lepton invariant mass window of 120 GeV < $m(4l)$ < 190 was chosen for signal to background ratio calculation.

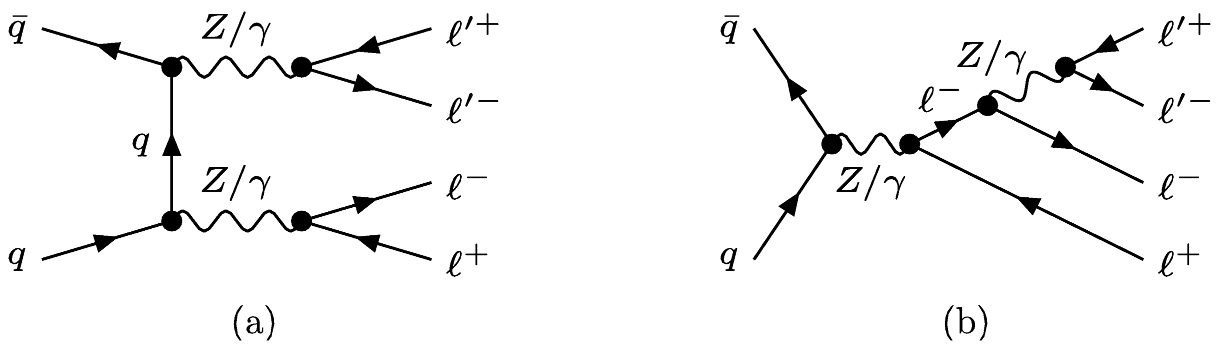


Figure 4.3: Feynman diagrams for the natural leptonic decay of the *Z*-diboson to SFOS lepton pairs. Three such combinations are considered for analysis: $ZZ \to 4e$, $ZZ \to 2e2\mu$, and $ZZ \to 4\mu$. Double-resonant t-channel *ZZ* production is shown in (a) and single-resonant s-channel Drell Yan production process of *Z*-diboson is shown in (b). [56] Branching ratios for these production processes are $BR = 86\%$ and $BR = 10\%$ respectively, excluding small contribution from electroweak vector-boson scattering processes. [57]

Table 4.4: Specific selection requirements for the *ZZ* diboson analysis. [54]

Specific Selection:	Requirements:
Preselection: Exactly four good leptons	lep_n = 4
Preselection: Two Z^0 candidates built from SFOS lepton pairs minimising the total deviation of both candidates from the Z^0 boson mass	$\|m(Z_1) - m(Z^0)\|$ + $\|m(Z_2) - m(Z^0)\|$ < 20 GeV with $m(Z^0)$ = 91.18 GeV
$\mathbf{4e}$, $\mathbf{2e2\mu}$, and $\mathbf{4\mu}$ channels are chosen	lep_type = 11 for each lepton ($4e$), or lep_type = 13 for each lepton (4μ), or lep_type = 11 for two leptons and lep_type = 13 for other two leptons ($2e2\mu$)
Leptons must have transverse momentum of $\boldsymbol{p_T}$ > 20 GeV	lep_pt > 20 GeV for all leptons
Electrons must have pseudorapidity of $\|\boldsymbol{\eta}\|$ < 2.47	np.absolute(lep_eta) < 2.47 for all lep_type = 11
At least one muon in the $\mathbf{4\mu}$ channel must have $\|\boldsymbol{\eta}\|$ < 2.4, corresponding to the acceptance of the muon trigger. The other muons must have $\|\boldsymbol{\eta}\|$ < 2.7	np.absolute(lep_eta[0]) < 2.4 or np.absolute(lep_eta[1]) < 2.4 or lnp.absolute(ep_eta[2]) < 2.4 or np.absolute(lep_eta[3]) < 2.4 if lep_type = 13 and trigM = True for that muon, np.absolute(lep_eta) < 2.7 for other muons
Each lepton pair must have invariant mass in the mass window 66 GeV < $\boldsymbol{m(ll)}$ < 116 GeV	66 GeV < $m(ll)$ < 116 GeV
Transverse impact parameter significance must be $\boldsymbol{sig(d_0)}$ < 6.5 (3.5) for electrons (muons)	lep_tracksigd0pvunbiased < 6.5 (3.5) if lep_type = 11 (13)
Four-lepton invariant mass window chosen to be 120 GeV < $\boldsymbol{m(4l)}$ < 190 GeV for signal to background ratio calculation	120 GeV < $m(4l)$ < 190 GeV

4.3.4 Heavy *ZZ* Resonance Analysis

The heavy *ZZ* resonance analysis focuses on the search for a heavy resonance decaying into two SM Z^0 bosons, which then decay leptonically. This process is symbolically represented by $X \to ZZ \to l^+l^-l^+l^-$ and seen in Figure 4.4. The specific selection requirements were obtained from the research paper on the "Search for heavy *ZZ* resonances in the $l^+l^-l^+l^-$ and $l^+l^-\nu\bar{\nu}$ final states using proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector" and can be seen in Table 4.5. [57]

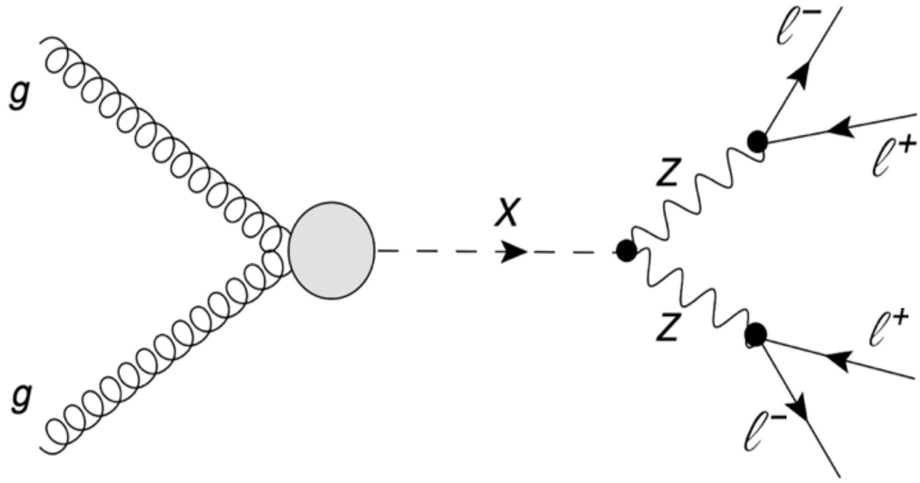


Figure 4.4: Feynman diagram for the ggF production and leptonic heavy *ZZ* resonance decay. [58]

The leptonic channels chosen for this analysis include the $ZZ \rightarrow 4e$, $Z \rightarrow 2e2\mu$, and $ZZ \rightarrow 4\mu$,which are similar to those considered in the *ZZ* analysis and *HZZ* analysis. The Z^0 boson decay to $\tau\tau$ lepton pairs is not considered for the same reasons as described in the *HZZ* Analysis (4.3.2) and *ZZ* Analysis (4.3.3) sections. The selection requirements applied in this analysis are almost identical to the selection requirements applied on the *HZZ* analysis in the *HZZ* Analysis section (4.3.2) as they both treat the leptonic signatures from *Z*-diboson processes with different origins. Two individual SFOS lepton pairs (exactly four leptons) are chosen for this analysis; one of smallest lepton pair invariant mass and Z^0 boson reconstructed mass difference (leading pair) and one of remaining smallest lepton pair invariant mass and Z^0 boson reconstructed mass difference (subleading pair). The considered electron (muon) must satisfy p_T > 7 GeV (5 GeV) and have a pseudorapidity of $|\eta|$ < 2.57 (2.7). Lepton kinematics suggests highest, second highest, and third highest p_T lepton must have p_T > 20, 15, and 10 GeV respectively. The two Z^0 candidates are built from SFOS lepton pairs minimising the total deviation of both candidates from the Z^0 boson mass, with smallest difference attributed to the leading lepton pair and second smallest difference attributed to the subleading lepton pair. The invariant mass of the leading and subleading pair must be in the range 50 GeV < m_{12} < 106 GeV and 12 GeV < m_{34} < 115 GeV respectively. The four-lepton invariant mass was chosen to be in the interval 140 GeV < $m(4l)$ < 190 GeV for signal to background ratio calculation. The choice in invariant mass window to be searched for the presence of a resonance constitutes the only difference between this analysis and the *HZZ* analysis.

Table 4.5: Specific selection requirements for the heavy *ZZ* resonance analysis. [57]

Specific selection:	Requirements:
Two pairs of oppositely charged leptons ($\boldsymbol{4e}$, $\boldsymbol{4\mu}$, and $\boldsymbol{2e2\mu}$ channels are chosen)	lep_n = 4, lep_type = 11 for each lepton ($4e$), or lep_type = 13 for each lepton (4μ), or lep_type = 11 for two leptons and lep_type = 13 for other two leptons ($2e2\mu$)
Each muon (electron) must have $\boldsymbol{p_T}$ > 5 GeV (7 GeV) and $|\boldsymbol{\eta}|$ < 2.7 (2.47)	np.absolute(lep_pt) > 5 GeV (7 GeV) and np.absolute(lep_eta) < 2.7 (2.47) for muons (electrons)
First, second, and third highest $\boldsymbol{p_T}$ leptons must have $\boldsymbol{p_T}$ > 20, 15, and 10 GeV respectively	lep_pt[0] > 20 GeV and lep_pt[1] > 15 GeV and lep_pt[2] > 15 GeV
SFOS lepton pairs with reconstructed invariant mass closest to that of the Z^0 boson (leading pair), must have 50 GeV < $\boldsymbol{m_{12}}$ < 106 GeV	Smallest $|m(Z^0) - m(ll)|$ with 50 GeV < m_{12} < 106 GeV and $m(Z^0)$ = 91.18 GeV
Remaining SFOS lepton pairs with reconstructed invariant mass closest to that of the Z^0 boson (subleading pair), must have 12 GeV < $\boldsymbol{m_{34}}$ < 115 GeV	Remaining smallest $|m(Z^0) - m(ll)|$ with 12 GeV < m_{34} < 115 GeV and $m(Z^0)$ = 91.18 GeV
Invariant mass of the four leptons must be in the mass window 140 GeV < $\boldsymbol{m(4l)}$ < 190 GeV	140 GeV < $m(4l)$ < 190 GeV

Chapter 5

5. Results

5.1 Event Displays Features List

The review of material and tools provided to open data students by the ATLAS International Masterclasses group of the ATLAS Open Data project showed inconsistencies between the tools taught and how to utilise them. As a result, a features list was created consisting of a list of inconsistencies between the utilisation methods of event displays such as Minerva and Hypatia, comparison between the two event displays, and discrepancy in descriptions of software tools. The event display features list can be found in the Results of International Masterclasses Review and Event Display Features List section (Appendix A) of this dissertation.

5.2 Results of Analyses

5.2.1 Results of $H\gamma\gamma$ Analysis

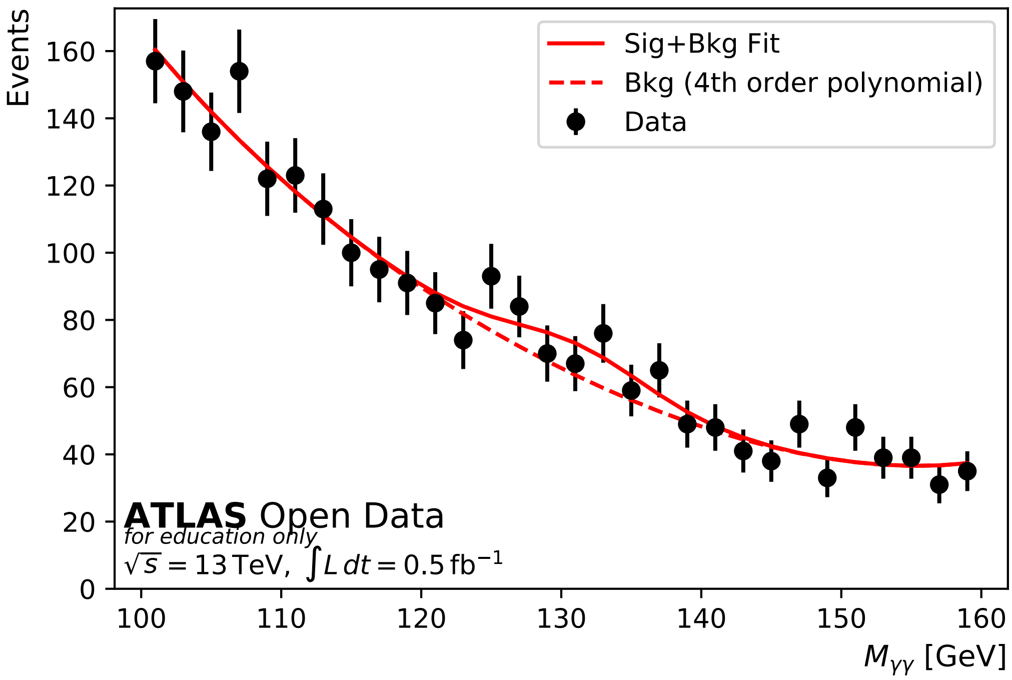


Figure 5.1: The distribution of the invariant mass of diphoton candidates after selection for the 13 TeV Data A sample. The dashed line represents the fourth order Bernstein polynomial background, the solid line represents signal plus background fit, while the black data points represent real data events.

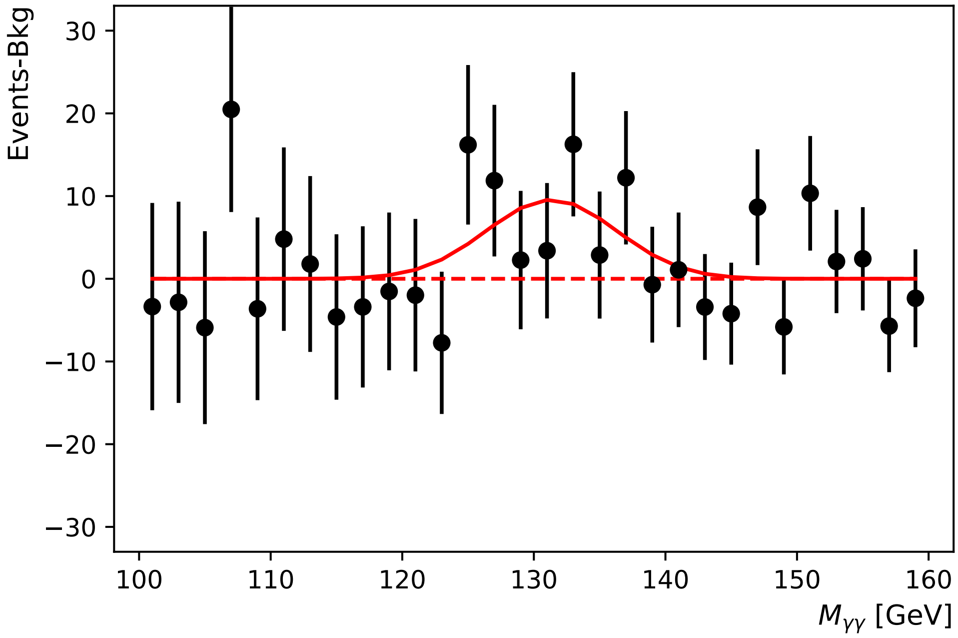


Figure 5.2: The residuals of Data A diphoton events with respect to the respective fitted background component. The Dashed line represents the fourth order Bernstein polynomial background, the solid line represents signal plus background fit, while black data points represent real data events.

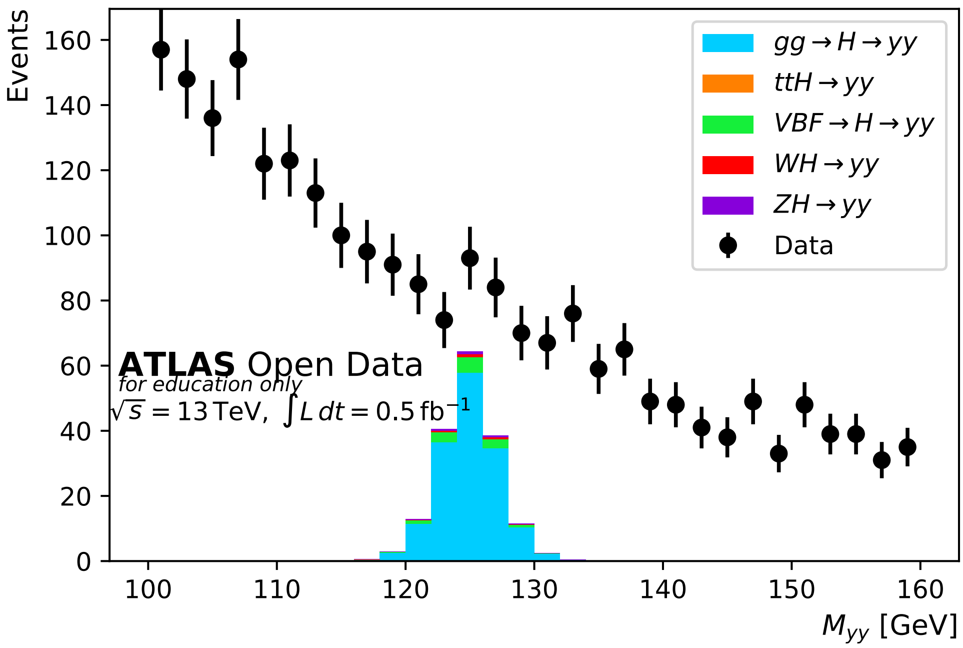


Figure 5.3: Resultant histogram showing number of events against the invariant mass of the diphoton candidate. The coloured bars represent MC data events from different production modes leading to diphoton decay of the Higgs boson seen in the legend. Highest background contribution is from the $gg \to H \to \gamma\gamma$ process in the considered diphoton invariant mass window.

5.2.2 Results of *HZZ* Analysis

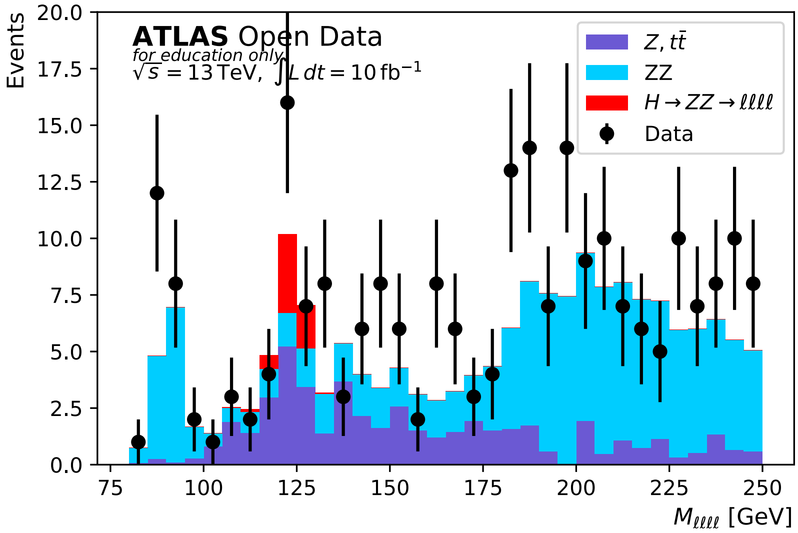


Figure 5.4: Resultant histogram showing number of events against the four-lepton invariant mass (Higgs boson candidate). Real data are represented by black data points while MC simulation data are represented by coloured bars. The histogram shows a number of data events passing the *HZZ* analysis selection candidates are from Higgs decay to the *Z*-diboson. Also included are backgrounds from $Z \rightarrow ll$ processes, $t\bar{t}$ annihilation and hadronization processes, as well as natural diboson processes $ZZ \rightarrow llll$.

5.2.3 Results of *ZZ* Analysis

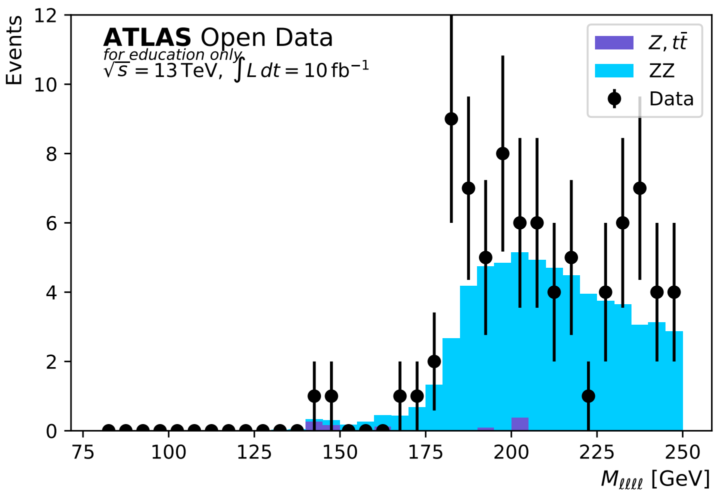


Figure 5.5: Resultant histogram showing number of events against four-lepton invariant mass. Real data are represented by black data points while MC simulation data are represented by coloured bars. The histogram shows majority of data events passing the *ZZ* analysis selection requirements are from simulated *ZZ* processes $ZZ \rightarrow llll$ and background $ZZ \rightarrow lll\nu$ processes ($ZZ \rightarrow ll\nu\nu$ background not considered since data was corrupted). $Z \rightarrow ll$ processes, $t\bar{t}$ annihilation and hadronization processes are also seen to have small background contributions.

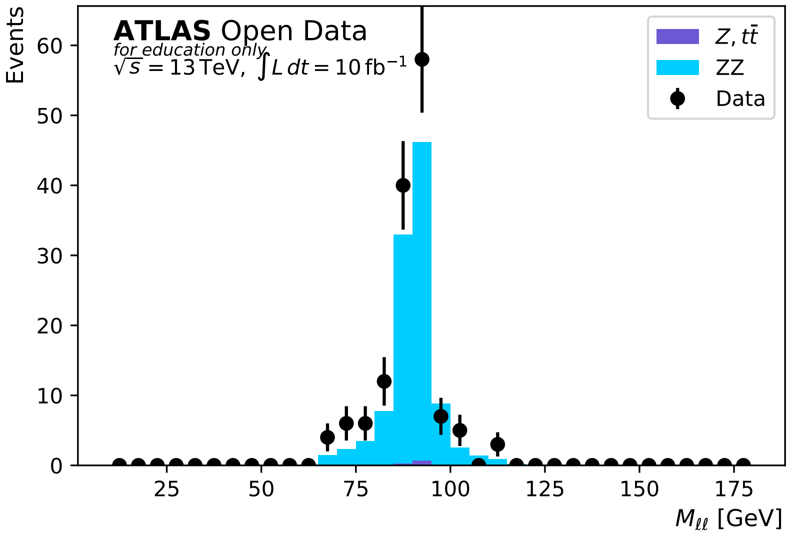


Figure 5.6: Resultant histogram showing number of events against the leading lepton pair invariant mass (m_{12}) distribution for a sample defined by the presence of a Z^0 boson candidate. Background contributions are described in Figure 5.5.

5.2.4 Results of Heavy *ZZ* Resonance Analysis

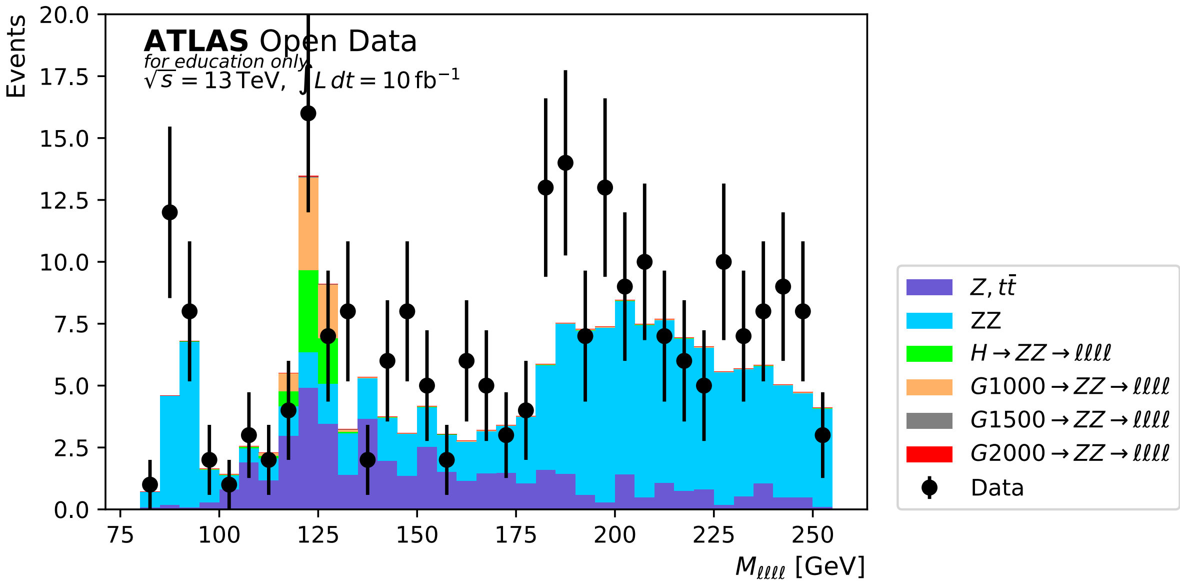


Figure 5.7: Resultant histogram showing number of events against the four-lepton invariant mass (Graviton candidate *G*). Real data are represented by black data points while MC simulation data are represented by the coloured bars. The histogram suggests that a Graviton of mass m_G = 1000 GeV could contribute to the data events. Also visible are large contributions from Higgs decay processes $H \rightarrow ZZ \rightarrow llll$ and natural *Z*-diboson processes $ZZ \rightarrow llll$. Relatively small backgrounds from $Z \rightarrow ll$ processes, $t\bar{t}$ annihilation and hadronization processes can also be seen. Contributions from Gravitons of mass m_G = 1500 GeV and m_G = 2000 GeV may be too small to be seen.

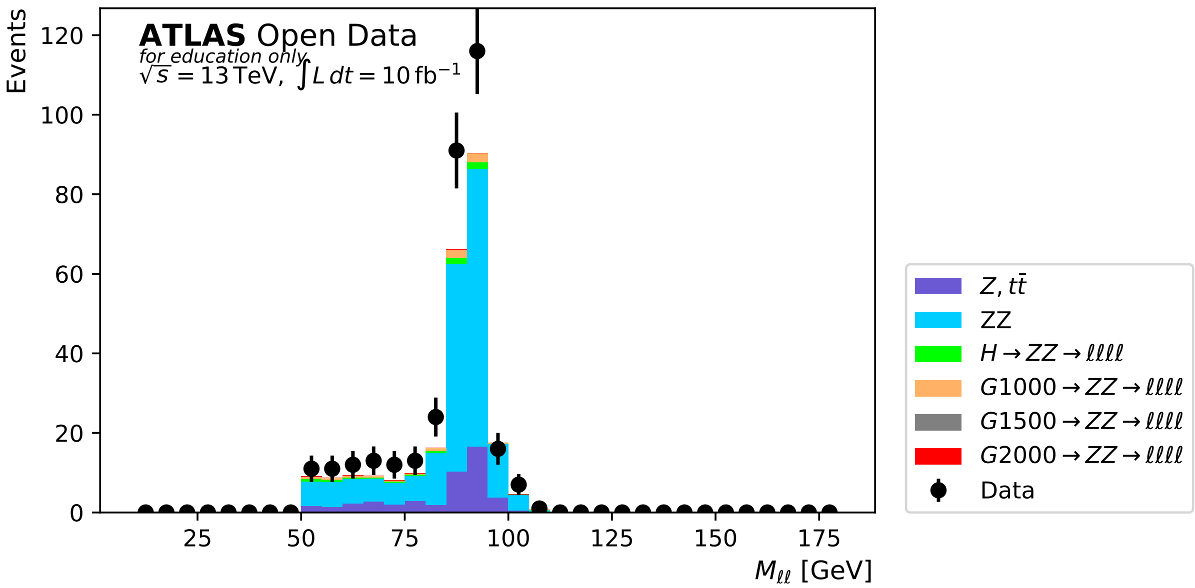


Figure 5.8: Resultant histogram showing number of events against the leading lepton pair invariant mass (m_{12}) distribution for a sample defined by the presence of a Z^0 boson candidate. Background contributions are described in Figure 5.7.

Chapter 6

6. Discussion

6.1 Discussion of Event Display Features List

Through the review of material provided by International Masterclasses, it was found that the Hypatia event display is more experimentally tuned for work with identification of particles and events. The full features list can be found in the Results of International Masterclasses Review and Event Display Features List section (Appendix A). The main difference between provided event displays is that Hypatia can be used for work with the Track Tables and Invariant Mass Tables, as well as inbuilt histograms. On the other hand, when it comes to event identification and more parameters are needed, it was more convenient to switch to Minerva for the same event and note the Isolation, Charge, and Opening Angle, as these are not provided by Hypatia. Following this, it was necessary to return to Hypatia and continue with the event identification process and analysis. The features list also contains information on the inconsistencies between provided material and provided tools. The main inconsistency was observed in the utilisation of Minerva while taking down the Isolation, Charge, and Opening Angle parameter values. It was found that charge and isolation numbers were not shown by the Minerva event display, contradicting the text provided to student learning from International Masterclasses resources. It was also found that contradictory to the text provided, Minerva does not show the missing transverse energy in the 3D lego plot. Instead it shows the missing transverse energy in a 2D lego plot.

6.2 Discussion of $H\gamma\gamma$ *Analysis*

The results of the $H\gamma\gamma$ Analysis were obtained after application of specific selection requirements mentioned in the $H\gamma\gamma$ Analysis section (4.3.1) on data sample A. Originally it was planned to do one analysis on the combined data samples A, B, C, and D. After this combined data sample analysis was carried out, a distribution of the invariant mass of diphoton candidates along with the residuals of diphoton events with respect to the respective fitted background component were obtained and can be seen in Figure 9.10 and Figure 9.11 respectively. From both those distributions it was observed that the signal, which is modelled by a Gaussian distribution function, drops below the background, which is modelled by the fourth order Bernstein polynomial, in the mass window of 105 GeV < $m(\gamma\gamma)$ < 125 GeV. However, this was not observed in the analysis done on data sample A whose resultant distributions can be seen in Figure 5.1, Figure 5.2 and Figure 5.3 of the Results of $H\gamma\gamma$ Analysis section (5.2.1). This sample showed an excess of signal events in the invariant mass window of 120 GeV < $m(\gamma\gamma)$ < 140 GeV for the diphoton candidate centred around 131 GeV. Even though the invariant mass window of the diphoton candidate encompasses the measured value of the Higgs boson of m_H = (125.1 ± 0.14) GeV, the mean of the Gaussian distribution in this analysis is slightly higher than the accepted value for the mass of the Higgs. In addition, when comparing these results with the previous results on the original observation of the Higgs boson seen in Figure 6.1, it was found that the mean of the Gaussian distribution for the mass of the Higgs was 6 GeV larger in this analysis than in the previously performed analysis. Hence the results of this analysis are more consistent with a Higgs of mass of m_H = 131 GeV than the Higgs of mass

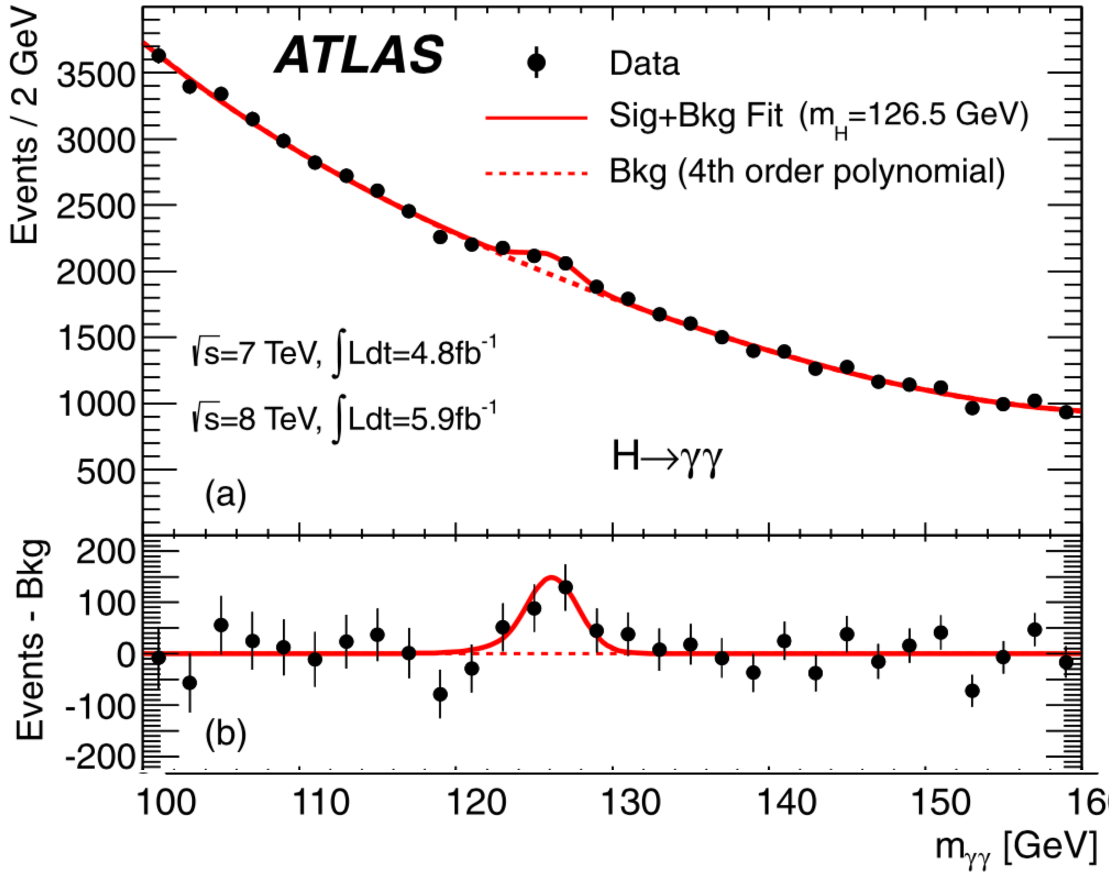


Figure 6.1: The distributions of the invariant mass of diphoton candidates after all selections for the combined 7 TeV and 8 TeV data sample. The inclusive sample is shown in (a), and the residuals of the data with respect to the respective fitted background component is shown in (b). [59]

m_H = 126.5 GeV but do allow for replication of original scientific results obtained at CERN as the residuals signal mass window does encompass the resultant Higgs mass from original findings. By assessing the chi squared value (χ^2) for the Data A sample, it was found to be χ^2 = 1347.42. To be able to tell the goodness of fit, it was first necessary to obtain the reduced chi squared value (χ_v^2), which is expressed in terms of χ^2 divided by the number of degrees of freedom (v). The number of degrees of freedom corresponds to the number of data being fit (30 bins), minus the number of free parameters (three for the Gaussian signal modelling, and four for the background fourth order polynomial modelling). As a result, the reduced chi squared value was found to be χ_v^2 = 58.58, which is still too large to be considered a “good fit”. Hence this problem is further explored in the Appendix B section. The results of the $H\gamma\gamma$ analysis still encompass the accepted value for the mass of the Higgs boson which are replicated and confirmed from the original paper on the observation of the Higgs. As such, a reasonable attempt at isolation of Higgs diphoton decay process was achieved in this analysis.

6.3 Discussion of *HZZ* Analysis

The results for the number of events versus the invariant mass of the Higgs boson candidate can be seen in Figure 5.4. The leading lepton pair invariant mass distribution was also created and can be found in the Resultant Leading Lepton Pair Invariant Mass Distribution of the HZZ Analysis section (Appendix C). It was found that the invariant mass of the Higgs boson in the considered four-lepton invariant mass window of 115 GeV < $m(4l)$ < 130 GeV had a signal to background ratio of 1.46. The fact that the MC simulation of the Higgs leptonic decay via *Z*-diboson process (red bars) shows an excess of simulated events centred around the 125 GeV mark and supported by the peak in real collected events at around the same point on the invariant mass axis suggests the signal has been isolated to a reasonable extent. The simulated prediction fits the real data for the considered selection requirements since the real data match simulated data in the considered mass window. Therefore, the obtained signal to background ratio of 1.46 infers the domination of signal over background and the presence of a boson with a signature that is consistent to that of Higgs with mass m_H = 125 GeV. This result was obtained to replicate the results of the first observation of the Higgs boson back in 2012 updated for the new 13 TeV data set. [59] The results from the original paper on the observation of the Higgs boson can be found in Figure 6.2. Comparing the invariant mass of the four leptons distributions in Figure 5.4 and Figure 6.2 shows similarities in the shape of the peaking Higgs decay simulation. Overall, the results obtained in this analysis are consistent with a measured Higgs of m_H = (125.1 ± 0.14) GeV, backed by the peak in the four-lepton invariant mass, and show a good example for replication of experimental results along with isolation of signal events.

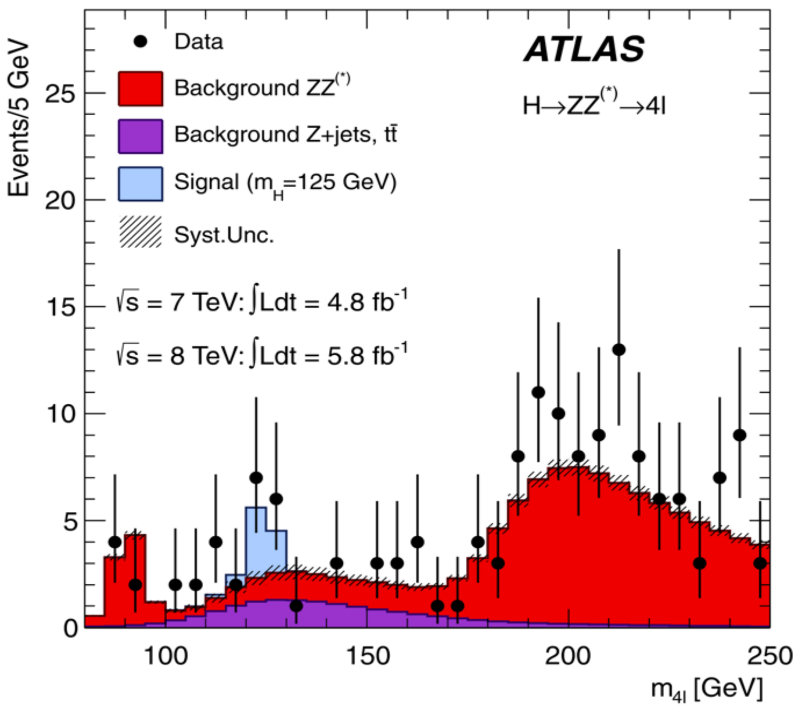


Figure 6.2: The distribution of the four-lepton invariant mass, m_{4l}, for the selected candidates, compared to the background expectation in the 80–250 GeV mass range, for the combination of the $\sqrt{s}$ = 7 TeV and $\sqrt{s}$ = 8 TeV data. The signal expectation for a SM Higgs with m_H = 125 GeV is also shown. [59]

6.4 Discussion of *ZZ* Analysis

The results of the *ZZ* analysis on the 13 TeV data set provided by ATLAS Open Data were obtained and can be seen in Figure 5.5 and Figure 5.6 of the Results of *ZZ* Analysis section (5.2.3). These results show a histogram for the distribution of four-lepton invariant mass and the distribution of the leading lepton pair invariant mass, after specific selection described in the *ZZ* Analysis section (4.3.3). By observing the distribution for the four-lepton invariant mass, no definite sign of the peak in number of events were seen from the MC simulations even though the real data suggests an excess of events with four lepton invariant mass in the window 180 GeV < $m(ll)$ < 185 GeV. Although the accepted mass for two Z^0 bosons (m_{ZZ} = 182.3752 GeV ± 0.0042 GeV) [55] was found to be within the experimentally obtained mass window, due to the lack of similar validation trend from MC simulations it was considered to perform a distribution for the invariant mass of the leading lepton pair as well. This distribution can be seen in Figure 5.6 and gives another representation for the signal of a Z^0 boson originating from *Z*-diboson processes. The reconstructed mass of the leading lepton pair shows reasonable agreement between data and simulation as the trend in both simulation and data appear to be very similar. Even though the trend is similar, it was noticed that the error bars of the real data did not encompass the number of events from the simulated *ZZ* processes. By considering the types of MC simulations used for comparison in this analysis, it was found that the $ZZ \rightarrow ll\nu\nu$ background process was not included since this simulation data was corrupted (the data contained empty lists). It was hence assumed that the discrepancy between the MC and real events in this analysis might have originated from the incomplete set of MC simulations used for comparison in this analysis, but further investigation is needed to improve upon experimental results. Overall this analysis shows that the trend of data is consistent with the presence of a Z^0 boson, supported by a peak in the invariant mass of the leading lepton pair originating from *Z*-diboson processes at the accepted mass of the Z^0 boson (m_Z = 91.1876 GeV ± 0.0021 GeV) with a signal to background ratio of 104 in the leading lepton invariant mass window of 66 GeV < $m(ll)$ < 116 GeV. Since the signal to background ratio for the four-lepton invariant mass window 120 GeV < $m(4l)$ < 190 GeV came out to be 14, the distribution of leading lepton pair invariant mass gave best results for the *ZZ* analysis and confirms the presence of the Z^0 boson centred around 92 GeV originating from a Z^0 boson pair. Hence the *ZZ* analysis isolated the *ZZ* processes to a reasonable extent and both the results can be used to infer the presence of *Z*-diboson events.

6.5 Discussion of Heavy *ZZ* Resonance Analysis

The results of the heavy *ZZ* resonance on the 13 TeV data set provided by ATLAS Open Data were obtained and can be seen in Figure 5.7 and Figure 5.8. These results show a histogram for the number of events versus the four-lepton (Graviton candidate) invariant mass and the distribution of leading lepton pair invariant mass respectively. As there are many different theories that predict the existence of particles beyond the SM, with different masses and different conditions, there exist many simulations for these hypothetical particles depending on their theory. The $X \rightarrow ZZ \rightarrow llll$ decay channel allows for an analysis on the hypothetical spin-2 particle, the Graviton. For this reason, multiple MC simulations for different hypothetical masses of the Kaluza–Klein Graviton (*G*) processes were considered for this analysis. [57] It was first considered to perform the analysis on the combined MC simulation for all the masses of the Graviton accessible. First, the resultant four-lepton and leading lepton invariant mass distributions were obtained and can be seen in Figure 9.13 and Figure 9.14 of the Results for Heavy *ZZ* Resonance Analysis Starting With All Graviton MC Files section (Appendix D) respectively. It was found that too many events have passed the specific selection requirements in simulations, as the considered mass window of 140 GeV < $m(4l)$ < 190 GeV was highly noisy for the simulation of the Graviton with mass of m_G = 200 GeV. For this reason, the simulation of Graviton of mass m_G = 200 GeV was excluded from the MC samples, and the analysis was repeated. Therefore, this analysis suggests that a Graviton of mass m_G = 200 GeV does not exist. As a result, new four-lepton and leading lepton pair invariant mass distributions were obtained, which can be seen in Figure 9.15 and Figure 9.16 of the Appendix D respectively. In the case of Figure 9.15, it was also found that an excess of MC simulated events for the hypothetical Graviton of mass m_G = 500 GeV exists in the four-lepton invariant mass window of 120 GeV < $m(4l)$ < 130 GeV. Due to this noisy region of the distribution the simulation of Graviton of mass m_G = 500 GeV was excluded as well. Therefore, this analysis suggests that a Graviton of mass m_G = 500 GeV does not exist as well. The resultant distributions were assessed on the presence of the Graviton excitation from heavy *ZZ* resonance in Figure 5.7 and Figure 5.8. Since the leftover MC simulation samples gave information on the number of events from Gravitons of mass higher than 1000 GeV, and no peak in invariant mass of the four leptons was observed other than at the accepted mass of the Higgs boson, it was considered that the Kaluza–Klein Graviton is excluded from the invariant mass region of 200 GeV < $m(4l)$ < 1000 GeV. The exclusion of the Graviton excitation from heavy *ZZ* resonance up to a mass of 1000 GeV was found to be consistent with the exclusion of the spin-2 Graviton up to a mass of 1300 GeV obtained from the paper on the search for heavy *ZZ* resonances in the $l^+l^-l^+l^-$ final states mentioned earlier in this section. For this reason, the results of the heavy *ZZ* resonance analysis did not isolate the resonance of the hypothetical particle beyond the SM well, but upon further inspection and addition of new specific selection requirements this analysis can be improved further. A good example of replication of scientific results is demonstrated by the findings of the heavy *ZZ* analysis.

6.6 Discussion of Further Progression

In the *HZZ* analysis discussed in the Discussion of *HZZ* Analysis section (6.3) it was possible to isolate the Higgs *Z*-diboson decay signal through the four-lepton channel. The replication of the original *HZZ* analysis was made, but the paper for comparison was obtained from year 2012. By now there exist many studies of the Higgs mechanism and the Higgs boson analysis becomes more and more reproducible. For this reason, a good further progression of this analysis would be to replicate the results for newer searches on the Higgs decay via the *Z*-diboson. In addition to this, other channels for the decay of Higgs can also be analysed and provided to ATLAS Open Data educational platform. In case of the $H\gamma\gamma$ Analysis, some investigation has already been made into the origin of underfitting described in the Discussion of $H\gamma\gamma$ Analysis section (6.2). To be able to improve upon the certainty of experimental results, it is essential to pinpoint the exact origin of the overcutting, hence further progression is needed in this analysis. In addition to this, since the ATLAS data obtained was already preselected, it would be beneficial to apply the majority of selection requirements one by one rather than using the preselected sets. This might reveal more inconsistencies between code versions described in section 6.2. Considering the *ZZ* analysis, which was able to isolate the *Z*-diboson processes pretty well, further progression should be made to include all relevant background processes. The most important background component that should be added to the *Z*-diboson simulation is considered to be $ZZ \rightarrow llvv$; this was originally excluded due to the corrupt nature of the data file upon download as mentioned in the Discussion of *ZZ* Analysis section (6.4). Regarding the heavy *ZZ* resonance discussed in the Discussion of Heavy *ZZ* Resonance Analysis section (6.5), it was found that the four-lepton channel for the decay of the hypothetical Graviton has little to no sensitivity in the higher mass range. For this reason, it would useful to perform the analysis on heavy *ZZ* resonance which decays via the $X \rightarrow ZZ \rightarrow llvv$ channel for higher mass ranges which could give further information on the Graviton's invariant mass window.

Chapter 7

7. Conclusion

Through this project a contribution to the development and update of the ATLAS Open Data educational platform on the newly provided 13 TeV data sets has been made. A features list was created to help eliminate inconsistencies found within provided text, as well as give feedback on the educational tools provided in the form of comparison and contrast of event displays. Supplying the features list to the International Masterclasses source may aid in updating material provided to students and help with solving limitations imposed by the limited person-power problem. The results for the *HZZ*, $H\gamma\gamma$, and heavy *ZZ* resonance analyses were compared to the original papers on these analyses, and it was concluded that a good replication study has been made due to the consistency in comparison. Updating of previous analyses allows for reproduction of known results, and acts as a reference point for students' attempts at developing their own analyses. Supplying the updated code and results to ATLAS Open Data may increase the amount of available reference material and reduce the effects of limited person-power and reproducibility issues. The analyses performed on *HZZ*, $H\gamma\gamma$, and *ZZ* decays isolated the required signal events to a reasonable extent, whereas the heavy *ZZ* resonance analysis excluded the mass of the Kaluza–Klein Graviton up to a mass of 1000 GeV. The *HZZ* and $H\gamma\gamma$ analyses are ready to be implemented into the ATLAS Open Data educational platform now, while the *ZZ* and heavy *ZZ* resonance analyses can be implemented after minor adjustments as discussed in the Discussion of ZZ Analysis section (6.4) and the Discussion of Heavy ZZ Resonance Analysis section (6.5).

Acknowledgments

I would like to give special thanks to Mr. Meirin Evans for providing necessary information throughout the course of this project and helping out in the most anxious of times. I would also like to give special thanks to Dr. Lily Asquith for pointing me in the right direction throughout this project and providing great feedback. I would also like to thank CERN and ATLAS Open Data Collaboration for providing me with relevant resources and materials to carry out this dissertation.

Bibliography

[1] CERN Accelerating science. (n.d.). Retrieved from https://home.cern/science/physics/standard-model

[2] ATLAS Open Data. (2020, February 28). Retrieved from http://atlas.cern/resources/opendata

[3] The Impact of CERN. (2016). The Impact of CERN. Retrieved from https://cds.cern.ch/record/2256277/files/ CERN - Brochure-2016-005-Eng.pdf

[4] Manyika, J., & Roxburgh, C. (2011, October). The Great Transformer: The impact of the internet on economic growth and prosperity. . Retrieved from http://dese.ade.arkansas.gov/public/userfiles/Legislative_Services/Quality Digital Learning Study/Facts/McKinsey_Global_Institute-Impact_of_Internet_on_economic_growth.pdf

[5] IPSOS MORI Veracity Index 2019. (2019). IPSOS MORI Veracity Index 2019. Retrieved from https://www.ipsos.com/sites/default/files/ct/news/documents/2019-11/trust-in-professions-veracity-index-2019-slides.pdf

[6] Fidler, F., & Wilcox, J. (2018). Reproducibility of scientific results.

[7] Griffith, C. (2020*). Strong Interactions and Deep Inelastic Scattering*. Personal Collection of Dr. Clark Griffith, University of Sussex, Brighton, East Sussex.

[8] Marciano, W., & Pagels, H. (1978). Quantum chromodynamics. *Physics Reports*, 36(3), 137-276.

[9] Webber, B. R. (1999). Fragmentation and hadronization. *arXiv preprint hep-ph/9912292*.

[10] Halliday, D., Walker, J., & Resnick, R. (2011). *Halliday, Resnick and Walkers Principles of physics: international student version* (10th ed.). Singapore: John Wiley & Sons Singapore Pte. Ltd. , 1033

[11] Trigger and Data Acquisition System. (2019, March 25). Retrieved from https://atlas.cern/discover/detector/trigger-daq

[12] Tracking Lecture 1, Indico CERN, (2018). Retrieved from https://indico.fnal.gov/event/8769/session/12/material/0/0.pdf

[13] Liu, F. H., Chen, Y. H., Gao, Y. Q., & Wang, E. Q. (2013). On current conversion between particle rapidity and pseudorapidity distributions in high energy collisions. *Advances in High Energy Physics*, 2013.

[14] How to draw diagrams in LaTeX with TikZ. (2018, September 10). Retrieved from https://wiki.physik.uzh.ch/cms/latex:tikz

[15] Griffith, C. (2020*). Particle production and detection*. Personal Collection of Dr. Clark Griffith, University of Sussex, Brighton, East Sussex.

[16] Griffith, C. (2020*). Special relativity: recalling relevant concepts*. Personal Collection of Dr. Clark Griffith, University of Sussex, Brighton, East Sussex.

[17] Herr, W., & Muratori, B. (2006). Concept of luminosity.

[18] Griffith, C. (2020*). Observables and Feynman Diagrams*. Personal Collection of Dr. Clark Griffith, University of Sussex, Brighton, East Sussex.

[19] Griffith, C. (2020*). Observables and Feynman Diagrams*. Personal Collection of Dr. Clark Griffith, University of Sussex, Brighton, East Sussex.

[20] Reconstructing a multitude of particle tracks within CMS. (n.d.). Retrieved from https://cms.cern/news/reconstructing-multitude-particle-tracks-within-cms

[21] Jende, K. (n.d.). International Masterclasses at CERN. Retrieved from https://atlas.physicsmasterclasses.org/en/index.htm

[22] CERN Accelerating science. (n.d.). Retrieved from https://writing-guidelines.web.cern.ch/entries/vertex

[23] Kenyon, I. R. (1982). The drell-yan process. *Reports on Progress in Physics*, 45(11), 1261.

[24] Schott, M. (2007). Study of the z Boson Production at the ATLAS Experiment with First Data (No. CERN-THESIS-2007-177).

[25] J. Beringer et al. (Particle Data Group), PR D86, 010001 (2012), Retrieved from http://pdg.lbl.gov/2012/listings/rpp2012-list-z-boson.pdf

[26] M. Tanabashi et al. (Particle Data Group), Phys. Rev. D 98, 030001 (2018), Retrieved from http://pdg.lbl.gov/2018/listings/rpp2018-list-w-boson.pdf

[27] Englert, F. (2014). The BEH mechanism and its scalar boson. *Annalen der Physik,* 526(5-6), 201-210.

[28] M. Tanabashi et al. (Particle Data Group), Phys. Rev. D 98, 030001 (2018) and 2019 update, Retrieved from http://pdg.lbl.gov/2019/listings/rpp2019-list-higgs-boson.pdf

[29] Ceci, S., Hadžimehmedović, M., Osmanović, H., Percan, A., & Zauner, B. (2017). Fundamental properties of resonances. *Scientific reports*, 7, 45246.

[30] Aad, G., Abbott, B., Abdallah, J., Khalek, S. A., Abdinov, O., Aben, R., ... & Abreu, H. (2015). Observation of top-quark pair production in association with a photon and measurement of the t t̄ γ production cross section in p p collisions at s= 7 TeV using the ATLAS detector. *Physical Review D*, 91(7), 072007.

[31] First collisions with the pixel detector. (2018, June 11). Retrieved from https://atlas.cern/updates/atlas-blog/first-collisions-pixel-detector

[32] CERN Accelerating science. (n.d.). Retrieved from https://home.cern/science/experiments/how-detector-works

[33] Detector & Technology. (2017, August 2). Retrieved from https://atlas.cern/discover/detector

[34] Calorimeters. (2019, December 10). Retrieved from https://atlas.cern/discover/detector/calorimeter

[35] Detecting Muons. (n.d.). Retrieved from https://cms.cern/detector/detecting-muons

[36] Aaboud, M., Aad, G., Abbott, B., Abdallah, J., Abdinov, O., Abeloos, B., ... & Abreu, H. (2017). Performance of the ATLAS Trigger System in 2015. *The European Physical Journal C*, 77(5), 317.

[37] Höche, S. (2016). Introduction to parton-shower event generators. In Journeys Through the Precision Frontier: Amplitudes for Colliders: *TASI 2014 Proceedings of the 2014 Theoretical Advanced Study Institute in Elementary Particle Physics* (pp. 235-295).

[38] Rizvi , E. (2010, November). London . Retrieved from http://pprc.qmul.ac.uk/~rizvi/Talks/hep-analysis.pdf

[39] Sjostrand, T. (2012, July). *Cern Summer Student Lecture*. CERN Summer Student Lecture. Lund . Retrieved from http://home.thep.lu.se/~torbjorn/talks/cern12a.pdf

[40] Gütschow, C. (2016, November). *Intercollegiate Postgraduate Course*. Intercollegiate Postgraduate Course. London . Retrieved from https://www.hep.ucl.ac.uk/postgrad/teaching/lhc/MC_2017.pdf

[41] Richardson , P. (2018). Durham . Retrieved from https://slideplayer.com/slide/12955540/

[42] Defending Your Life (Part 2). (2017, May 22). Retrieved from https://atlas.cern/updates/atlas-blog/defending-your-life-part-2

[43] Aaboud, M., Aad, G., Abbott, B., Abdinov, O., Abidi, S. H., AbouZeid, O. S., ... & Acharya, B. S. (2018). Measurement of differential cross-sections of a single top quark produced in association with a W boson at s = 13 TeV with ATLAS. *The European physical journal. C*, Particles and fields, 78(3).

[44] Potter , T. (n.d.). Lecture 3 . *Lecture 3* . Cambridge . Retrieved from https://www.hep.phy.cam.ac.uk/~chpotter/particleandnuclearphysics/Lecture_03_CollidersDetectors.pdf

[45] Welcome to Python.org. (n.d.). Retrieved from https://www.python.org/

[46] ROOT a Data analysis Framework: ROOT a Data analysis Framework. (n.d.). Retrieved from https://root.cern.ch/

[47] Jende, K. (n.d.). Visualization with HYPATIA. Retrieved from https://atlas.physicsmasterclasses.org/en/zpath_teilchenid2.htm

[48] Jende, K. (n.d.). The Event Display MINERVA. Retrieved from https://atlas.physicsmasterclasses.org/en/wpath_teilchenid2.htm

[49] CERN Accelerating science. (n.d.). Retrieved from http://information-technology.web.cern.ch/services/lxplus-service

[50] Event selection. (n.d.). Retrieved from http://opendata.atlas.cern/books/current/openatlasdatatools/_book/event_selection.html

[51] Aaboud, M., Aad, G., Abbott, B., Abdinov, O., Abeloos, B., Abidi, S. H., ... & Abulaiti, Y. (2018). Measurements of Higgs boson properties in the diphoton decay channel with 36 fb− 1 of p p collision data at s= 13 TeV with the ATLAS detector. *Physical review D*, 98(5), 052005.

[52] Jende, K. (n.d.). Higgs Events . Retrieved from https://atlas.physicsmasterclasses.org/en/zpath_hevents.htm

[53] Aaboud, M., Aad, G., Abbott, B., Abdinov, O., Abeloos, B., Abidi, S. H., ... & Abreu, R. (2017). Measurement of inclusive and differential cross sections in the H→ ZZ*→ 4 ℓ decay channel in pp collisions at $$\sqrt {s}= 13$$ TeV with the ATLAS detector. *Journal of High Energy Physics*, 2017(10), 132.

[54] Aad, G., Abbott, B., Abdallah, J., Abdinov, O., Abeloos, B., Aben, R., ... & Abreu, R. (2016). Measurement of the Z Z Production Cross Section in p p Collisions at s= 13 TeV with the ATLAS Detector. *Physical review letters*, 116(10), 101801.

[55] M. Tanabashi et al. (Particle Data Group), *Phys. Rev*. D 98, 030001 (2018) and 2019 update, Retrieved from http://pdglive.lbl.gov/Particle.action?node=S044&init=0

[56] Kallweit, S., & Wiesemann, M. (2018). ZZ production at the LHC: NNLO predictions for 2ℓ2v and 4ℓ signatures. *Physics Letters B*, 786, 382-389.

[57] Aaboud, M., Aad, G., Abbott, B., Abdinov, O. B. O., Abeloos, B., Abidi, S. H., ... & Buanes, T. (2018). Search for heavy ZZ resonances in the ℓ+ ℓ− ℓ+ ℓ− and ℓ+ ℓ− νν¯ final states using proton–proton collisions at s√= 13 TeV with the ATLAS detector. *European Physical Journal C*, 78(293), 1-34.

[58] Sirunyan, A. M. (2018). Search for ZZ resonances in the 2ℓ2v final state in proton-proton collisions at 13 TeV. *Journal of High Energy Physics*.

[59] Aad, G., Abajyan, T., Abbott, B., Abdallah, J., Khalek, S. A., Abdelalim, A. A., ... & AbouZeid, O. S. (2012). Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1), 1-29.

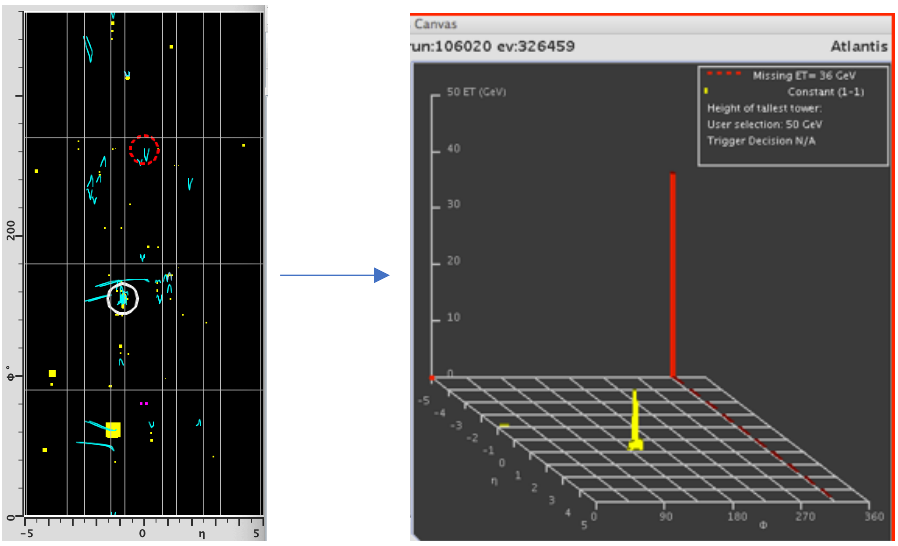
Appendix

A. Results of International Masterclasses Review and Event Display Features List

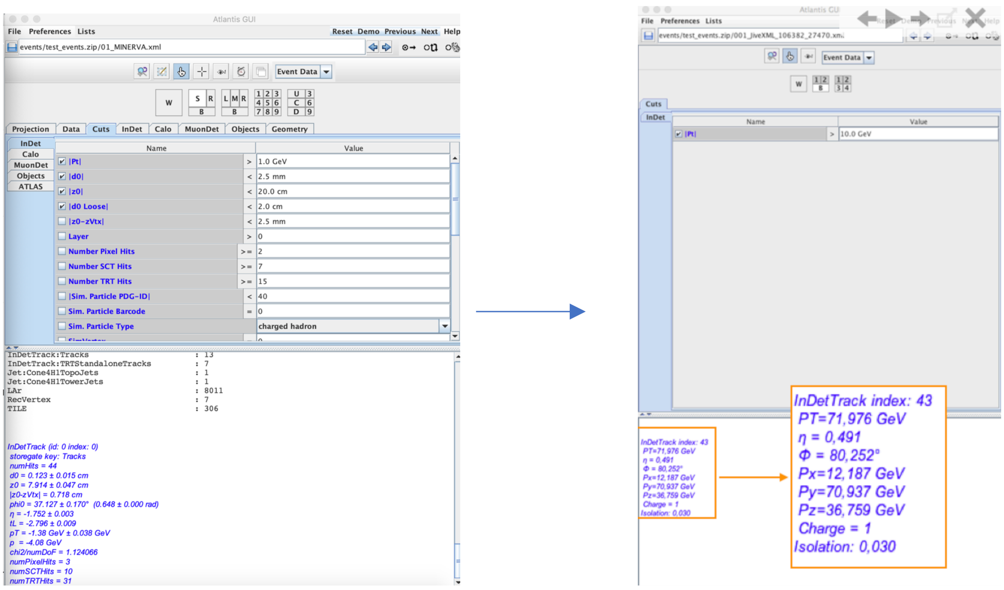
If the arrow is pointing from a picture, that picture is what the program actually shows. If arrow is pointing to a picture, that picture is what International Masterclasses says the program should show.

<u>Minerva inconsistencies:</u>

- Lego Plot only in 2-D
- Lego Plot doesn't show Missing ET box

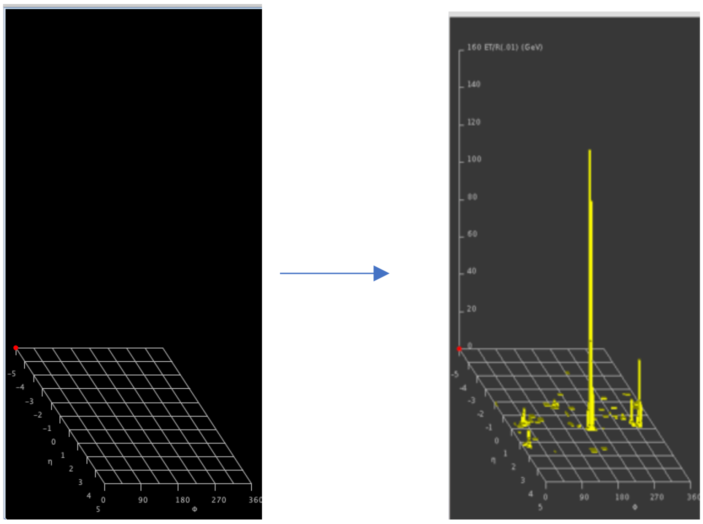


- Charge and Isolation numbers not shown in the information box.

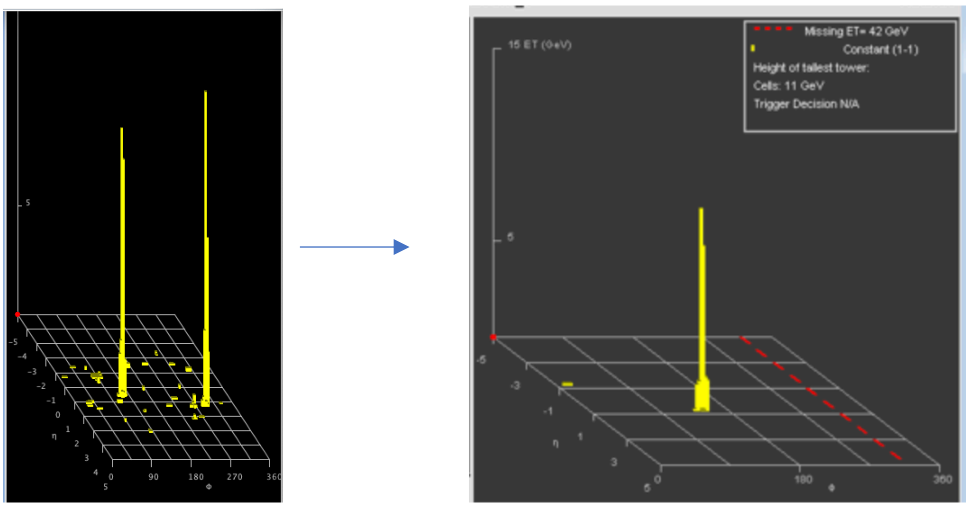


Hypatia Inconsistencies:

- Lego Plot Inactive after initial launch of the program. Need to load an exercise file first to activate the plot.



- Lego plot doesn't show missing ET box



Event Display Comparison:

- Both Hypatia and Minerva have essentially the same side view and cross-sectional view of the detector.
- Both Show vertices of collisions.
- Both have the same Control Window.
- Both have generally the same colour scheme.
- Both can load the same format of files (data).

Event Display Contrast:

- Only Minerva can show angles between two tracks.
- Only Minerva has a 2-D Lego Plot.
- Only Minerva can show Charge and Isolation numbers after selecting a track. (Bugged: doesn't show)
- Only Minerva shows jets as white cones in the Canvas Window.
- While in Minerva the Missing ET is shown with a red arrow, in Hypatia it is shown with a red dashed line.
- While in Minerva one has to click on the missing ET's red track to see the missing ET, in Hypatia one can see it automatically in a separate window (underneath the previous event button). (Bugged: both should show missing ET box in lego plot)
- Only Hypatia has a Track Window where one can see some physical information on the selected track, as well as identify the track and move it to the invariant mass table for further analysis.
- Only Hypatia has an Objects Table which shows tracks of strong signal objects from the event.
- Only Hypatia has an Invariant Mass Table Window.
- Only Hypatia has a builtin histogram function for the tracks in the invariant mass table.

Why is Hypatia Better?

Generally, it can be said that Hypatia event display is more useful in identification of particles than Minerva due to its mathematical and analytical nature. Here are a couple of reasons why Hypatia is better. The first reason is that Hypatia comes with a Track Window and Invariant Mass Window, as well as a builtin Histogram Function. Being able to view the selected tracks and compare them in the Track Window makes identification process easier. In addition, the histogram function can map the invariant mass tracks for further analysis, hence providing a broader understanding of the events taking place. The second reason why Hypatia is better is the existence of the Objects Window. This is very useful as it automatically shows strong signal objects once the selected track is clicked, making it easier to analyse and identify.

Why is Minerva Better?

Even though Hypatia is more mathematical in its nature and allows to work with more physical quantities, Minerva has some very important tools that can only be accessed through it. Hence Minerva is the better event display when it comes to the tools provided. The extra quantities that cannot be obtained from Hypatia are the Isolation and Charge numbers, as well as the 2-D version of the Lego plot. Some events require the isolation numbers to confirm the type of decay happening, while it is also sometimes more convenient to look at the 2-D version of the Lego plot rather than 3-D. In addition to this, one very important tool of Minerva that cannot be utilised in Hypatia is the opening angle calculation between two tracks. This can easily be done in Minerva and is very useful when identifying particles and events or imposing selection requirements on analyses.

B. Discussion of the Cause of Underfitting in the Results of $H\gamma\gamma$ Analysis

The dip in signal discussed in the Discussion of $H\gamma\gamma$ Analysis section (6.2) suggests the domination of the background events over signal events, and hence it was considered to separate the data samples and perform individual analysis on each one. The results of analysis on data sample B, C, and D can be found in the Results of $H\gamma\gamma$ Analysis on Data B, C, D, and Their Combined Data Sample section (Appendix E) (upon download each data sample already had some preselection requirements applied). From the above-mentioned samples, the residuals distributions for samples B, C, and D showed no promise in Higgs diphoton signal identification as each distribution either showed no deviation of signal from background or showed the domination of background over signal events. Due to the encountered variation in the fitting of the signal it was considered that there exists a problem with the fitting function in use. To obtain information on the origin of the underfitting, it was first necessary to exclude the cuts on converted photons in any part of the detector, cuts on first isolation of photons, and cuts on photon triggering to bring the code to the original form that was adapted from Mr. Meirin Evans. The analysis on original code and the updated version of the same code were run simultaneously, hence two cut flows on the same data were obtained and compared to each other at each stage of event selection. An output of a distribution of the invariant mass of diphoton candidates along with the residuals of diphoton events with respect to the respective fitted background component was created for each stage in the selection process for each code version and can be seen in form of a cut flow in the Cut flow for $H\gamma\gamma$ Analysis on Data A Sample section (Appendix F). By comparing the two cut flows, it was found that the two codes give different outputs on selected data events at stage (4.) of the cut flow, which performs the cut on pseudorapidity inside barrel/end-cap transition region. By comparing the definitions of these cuts in both versions of the code, they were found to be identical, therefore a question was posed regarding the reason for the variation in amount of cut data events. To investigate this further, a look into the transverse momentum distributions of subleading and leading photons was made in a similar way to the cut flow of Appendix F to see how the cuts from stage (3.) to stage (4.) differ in the amount of cut events. The results of this cut flow comparison can be seen in Figure 9.18. It was found that the number of events in a maximally filled bin drops from around 275,000 in the leading photon distribution of Figure 9.18 (a) to around 225,000 in Figure 9.18 (b), while dropping from around 310,000 in the subleading photon distribution of Figure 9.18 (c) to around 290,000 in Figure 9.18 (d). A quick estimation for the number of dropped events shows a total cut of roughly 70,000 data events (20,000 from leading photons and 50,000 from subleading photons) going from stage (3.) to (4.) of the cut flow. Comparing this result with the resultant output of the cut flow at stage (4.) found in Appendix F between the two versions of the code, it was found that about 24,000 events were taken out in Mr. Meirin Evans' analysis code at this stage compared to 70,000 events taken out in the updated version of the code. Hence it was found that roughly three times more events were cut from the updated version of the code than the original version. A slight hint of this factor of three in the cutting process was additionally observed in the residual distributions for the two versions of the code seen in stage (4.) of the cut flow in Appendix F. In this comparison the trend of the curve encompassing all residuals was found to be identical in both cases, differing only by the scale of the number of events; the factor of three. Hence it was confirmed that the updated code overcuts events going from stage (3.) in the cut flow to stage (4.), however further work is needed to pinpoint the origin and correct the overcutting to improve the fitting of the data.

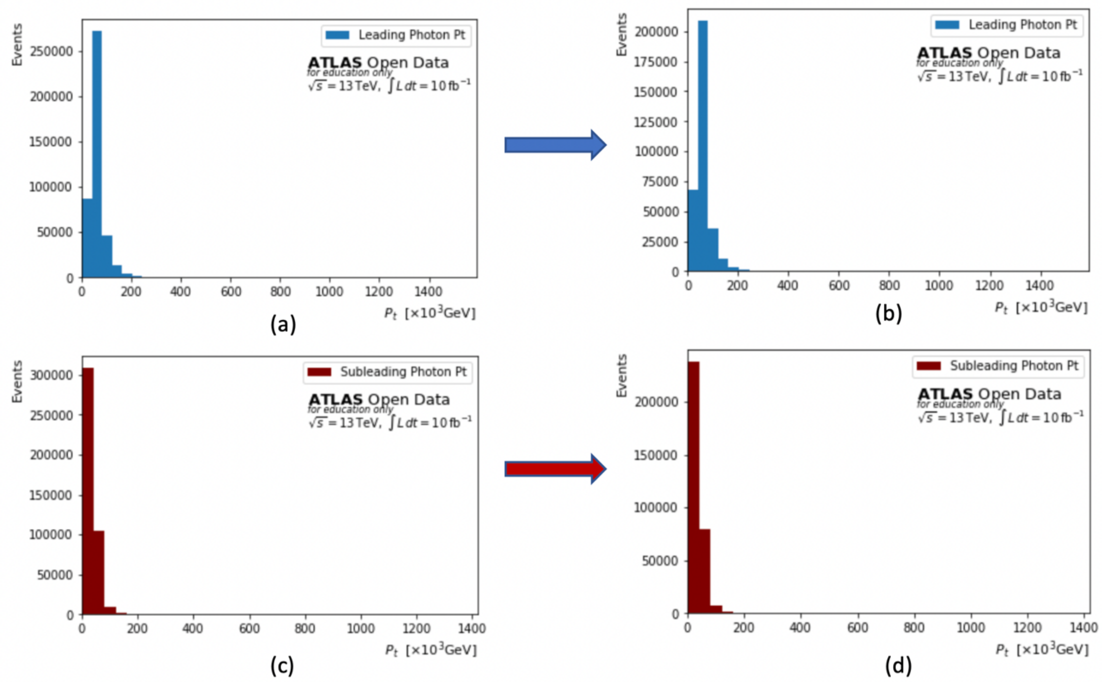


Figure 9.18: Resultant transverse momentum distributions of leading and subleading photons at stage (3.) of the cut flow can be seen in (a) and (c) respectively, while transverse momentum distributions of leading and subleading photons at stage (4.) of the cut flow can be seen in (b) and (d) respectively. Note: The integrated luminosity for these results is actually 0.5 fb^{-1}, not 10 fb^{-1} as shown.

C. Resultant Leading Lepton Pair Invariant Mass Distribution of the *HZZ* Analysis

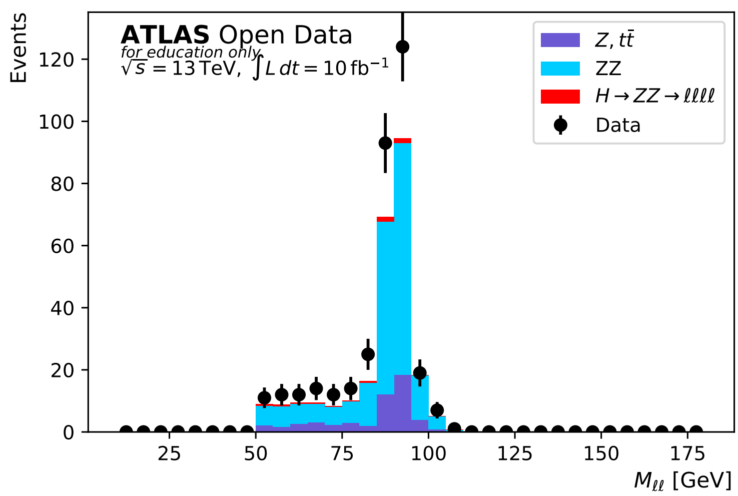


Figure 9.17: Resultant histogram showing number of events against the leading lepton pair invariant mass (m_{12}) distribution for a sample defined by the presence of a Z^0 boson candidate. Background contributions are described in Figure 5.4.

D. Results for Heavy *ZZ* Resonance Analysis starting with all Graviton MC files

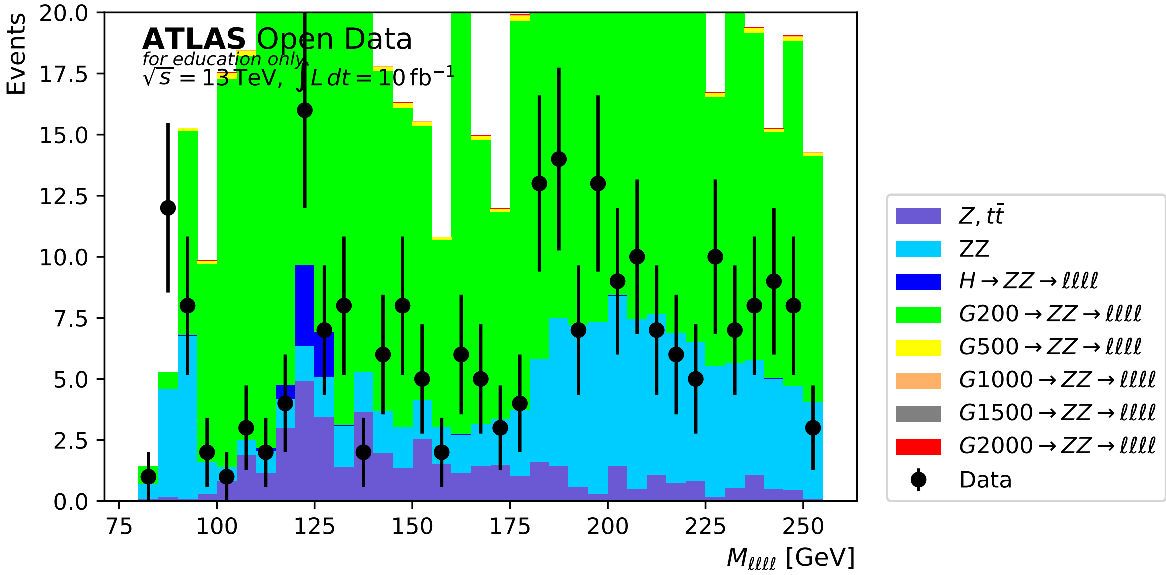


Figure 9.13: Resultant histogram showing number of events against the four-lepton invariant mass (Graviton candidate). Real data are represented by black data points while MC simulation data are represented by coloured bars. The histogram shows majority of simulated events passing the heavy *ZZ* analysis selection requirements are from the hypothetical Graviton of mass m_G = 200 GeV decay process. Also visible are large contributions from Higgs decay processes $H \rightarrow ZZ \rightarrow llll$ and natural *Z*-diboson processes $ZZ \rightarrow llll$. Relatively small backgrounds from $Z \rightarrow ll$ processes, $t\bar{t}$ annihilation and hadronization processes can also be seen. Contributions from Gravitons of mass m_G = 1500 GeV and m_G = 2000 GeV may be too small to be seen.

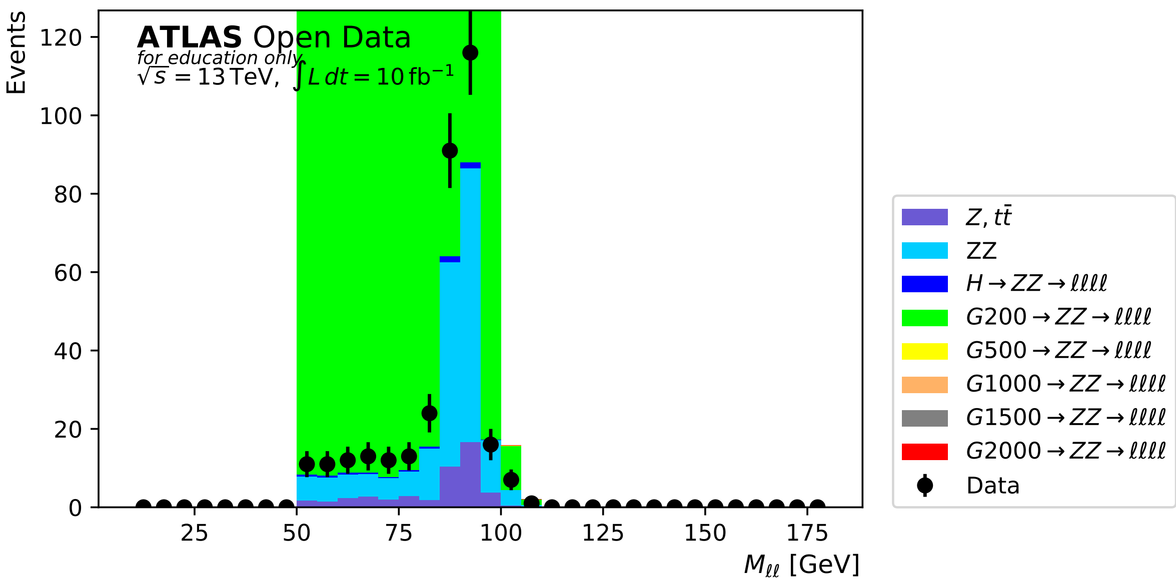


Figure 9.14: Resultant histogram showing number of events against the leading lepton pair invariant mass (m_{12}) distribution for a sample defined by the presence of a Graviton candidate. Background contributions are described in Figure 9.13.

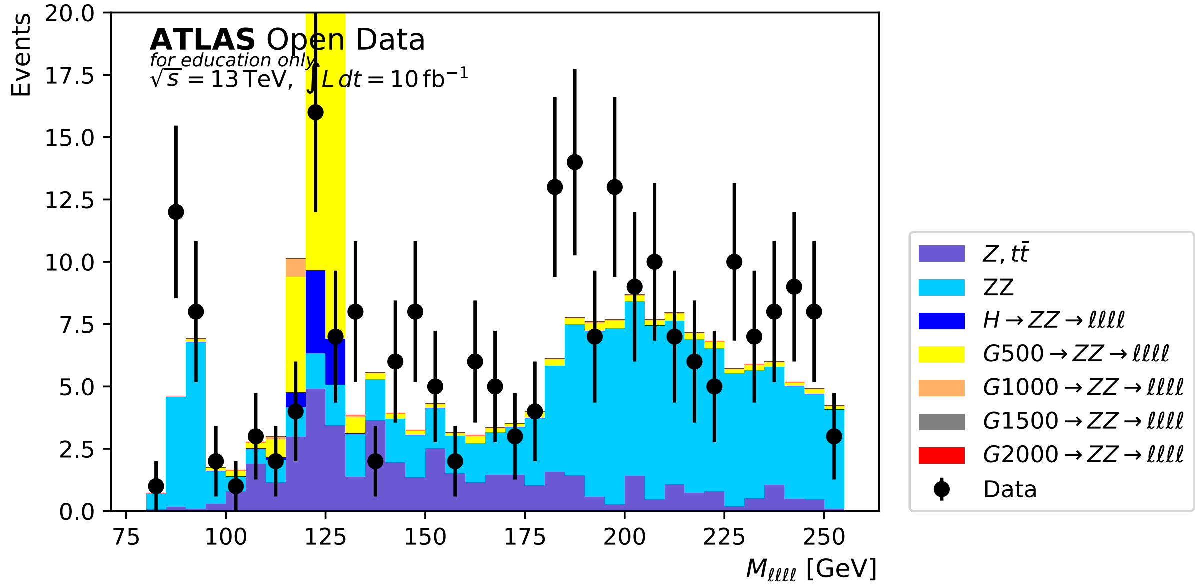


Figure 9.15: Resultant histogram showing number of events against the four-lepton invariant mass (Graviton candidate). Real data are represented by black data points while MC simulation data are represented by coloured bars. The histogram shows majority of simulated events passing the heavy *ZZ* analysis selection requirements are from the hypothetical Graviton of mass m_G = 500 GeV decay process. Also visible are large contributions from Higgs decay processes $H \rightarrow ZZ \rightarrow llll$ and natural *Z*-diboson processes $ZZ \rightarrow llll$. Relatively small backgrounds from $Z \rightarrow ll$ processes, $t\bar{t}$ annihilation and hadronization processes can also be seen. Contributions from Gravitons of mass m_G = 1500 GeV and m_G = 2000 GeV may be too small to be seen.

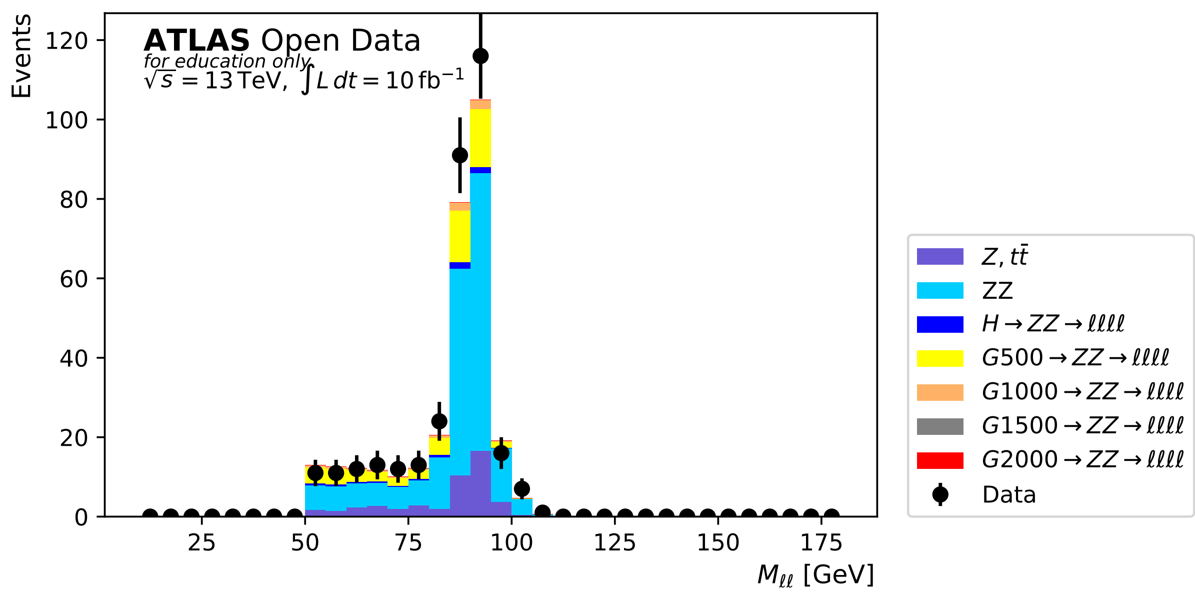


Figure 9.16: Resultant histogram showing number of events against the leading lepton pair invariant mass (m_{12}) distribution for a sample defined by the presence of a Graviton candidate. Background contributions are described in Figure 9.15.

E. Results of $H\gamma\gamma$ Analysis on Data B, C, D, and Their Combined Data Sample.

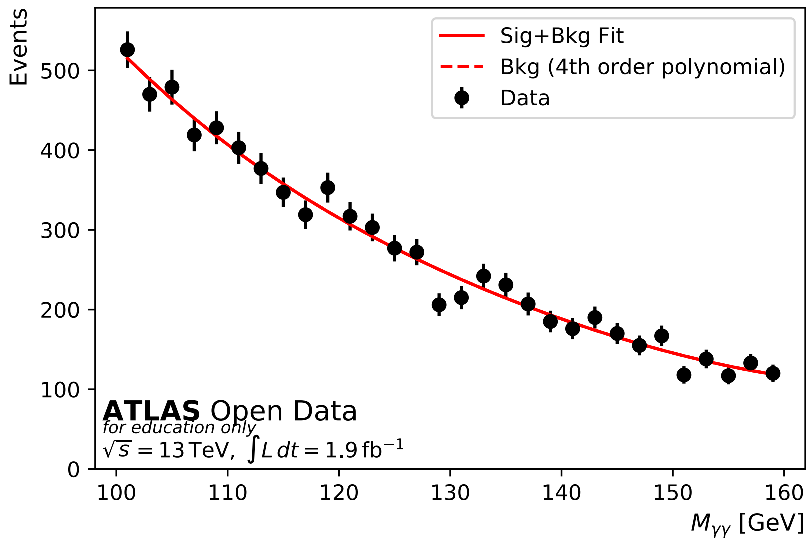


Figure 9.1: The distribution of the invariant mass of diphoton candidates after selection for the 13 TeV Data B sample. The dashed line represents the fourth order Bernstein polynomial background, the solid line represents signal plus background fit, while black data points represent real data events.

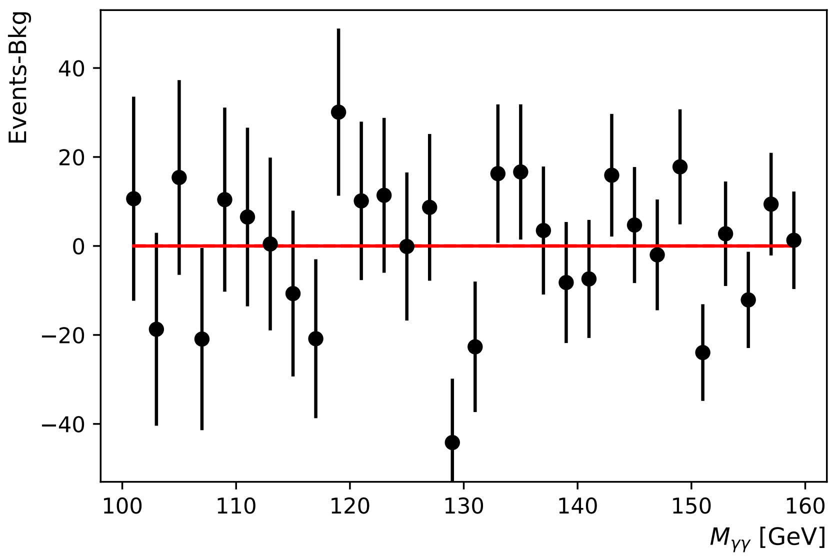


Figure 9.2: The residuals of Data B diphoton events with respect to the respective fitted background component. The dashed line represents the fourth order Bernstein polynomial background, the solid line represents signal plus background fit, while black data points represent real data events.

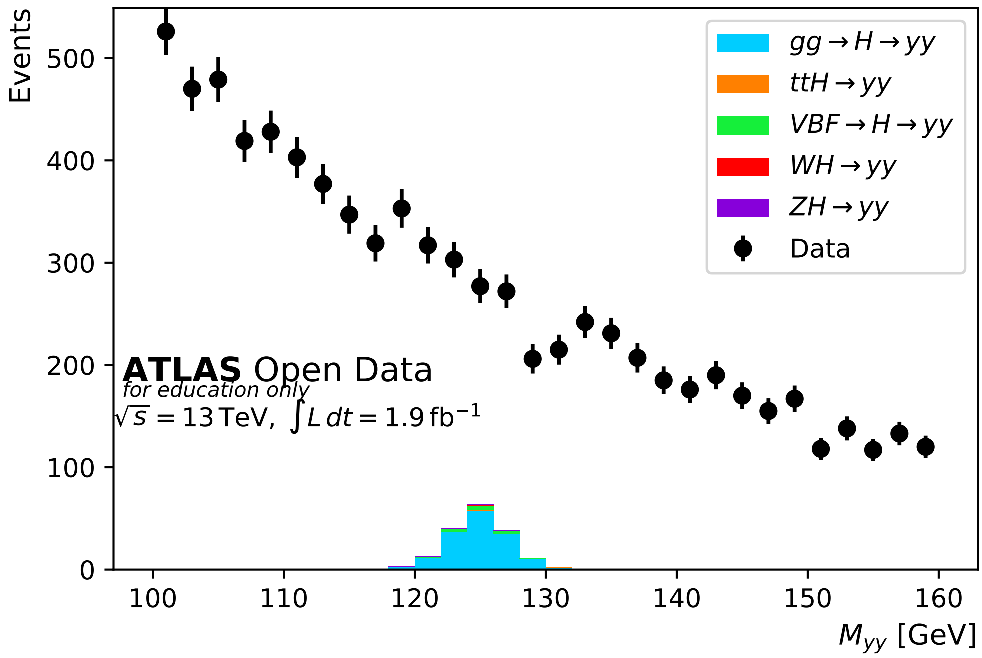


Figure 9.3: Resultant histogram showing number of events against the invariant mass of the diphoton candidate from Data B sample. The coloured bars represent MC data events from different production modes leading to diphoton decay of the Higgs boson seen in the legend. Highest background contribution is from the $gg \to H \to \gamma\gamma$ process in the considered diphoton invariant mass window.

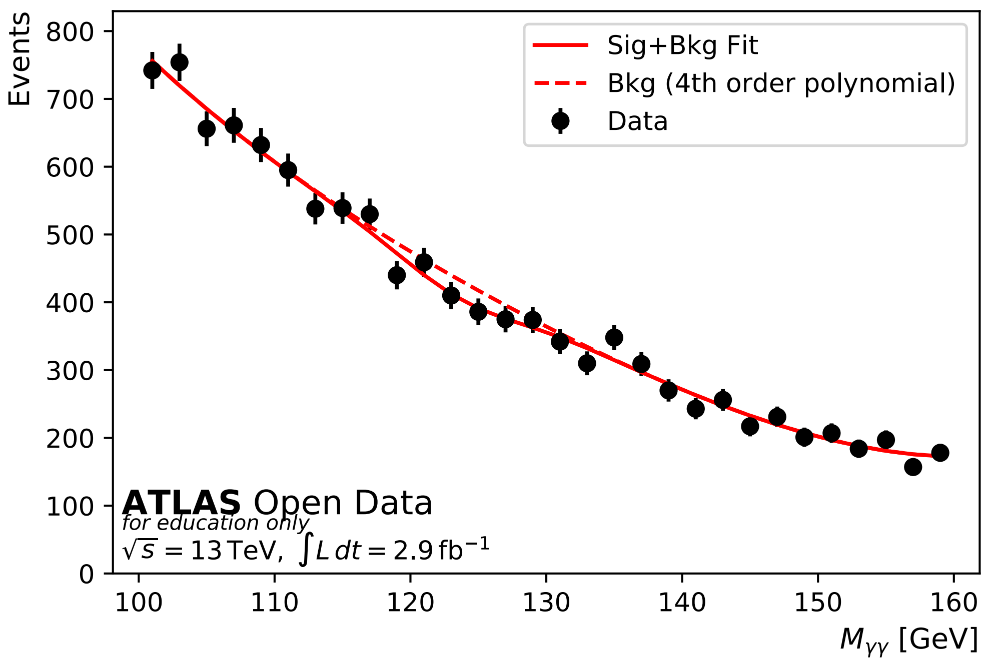


Figure 9.4: The distribution of the invariant mass of diphoton candidates after selection for the 13 TeV Data C sample. The dashed line represents the fourth order Bernstein polynomial background, the solid line represents signal plus background fit, while black data points represent real data events.

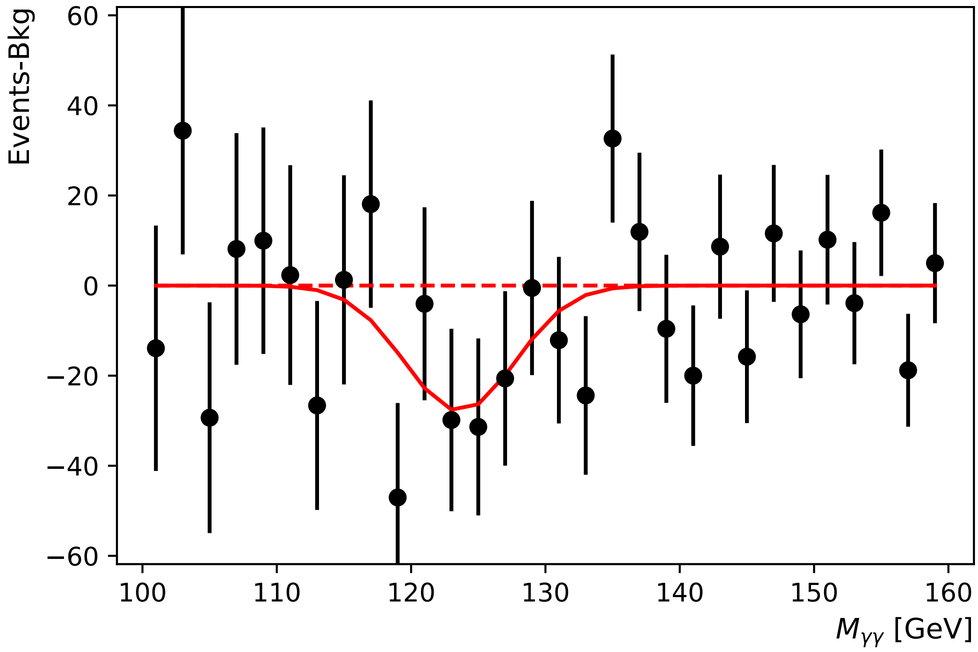


Figure 9.5: The residuals of Data C diphoton events with respect to the respective fitted background component. The dashed line represents the fourth order Bernstein polynomial background, the solid line represents signal plus background fit, while black data points represent real data events.

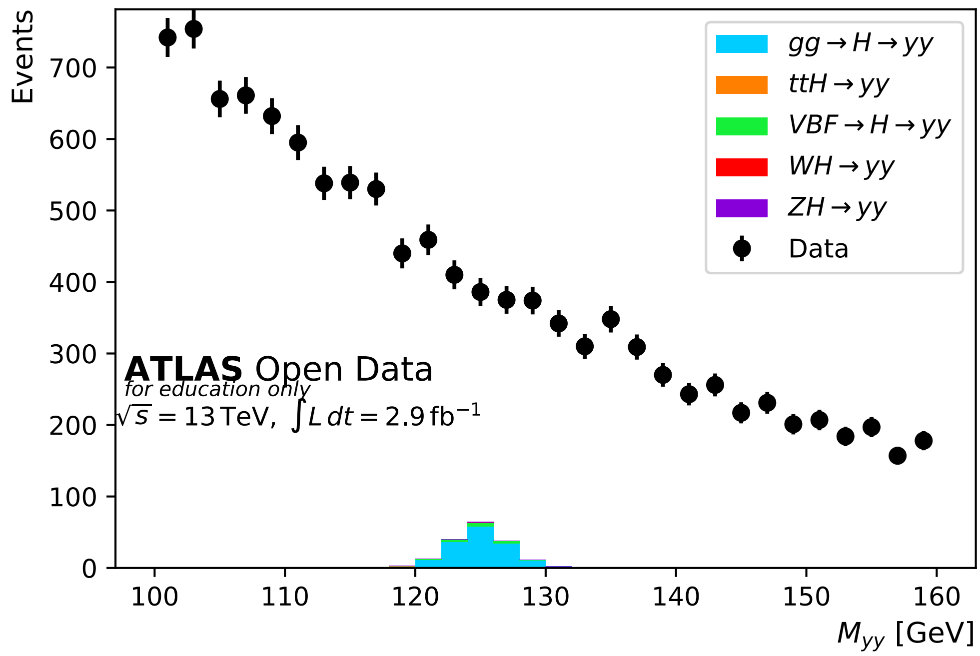


Figure 9.6: Resultant histogram showing number of events against the invariant mass of the diphoton candidate from Data C sample. The coloured bars represent MC data events from different production modes leading to diphoton decay of the Higgs boson seen in the legend. Highest background contribution is from the $gg \to H \to \gamma\gamma$ process in the considered diphoton invariant mass window.

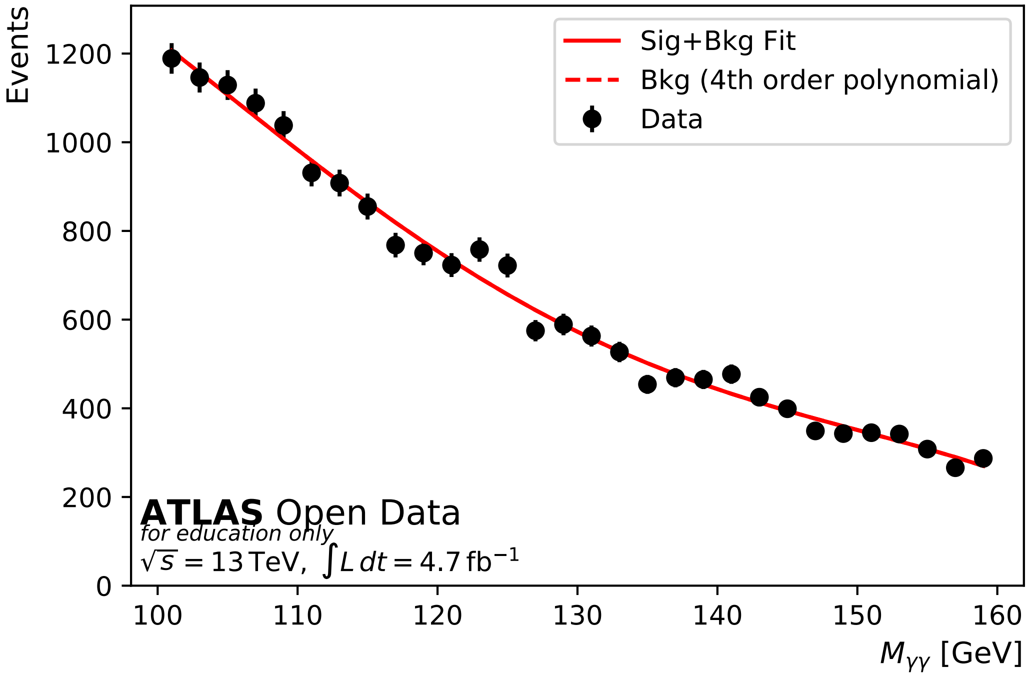


Figure 9.7: The distribution of the invariant mass of diphoton candidates after selection for the 13 TeV Data D sample. The dashed line represents the fourth order Bernstein polynomial background, the solid line represents signal plus background fit, while black data points represent real data events.

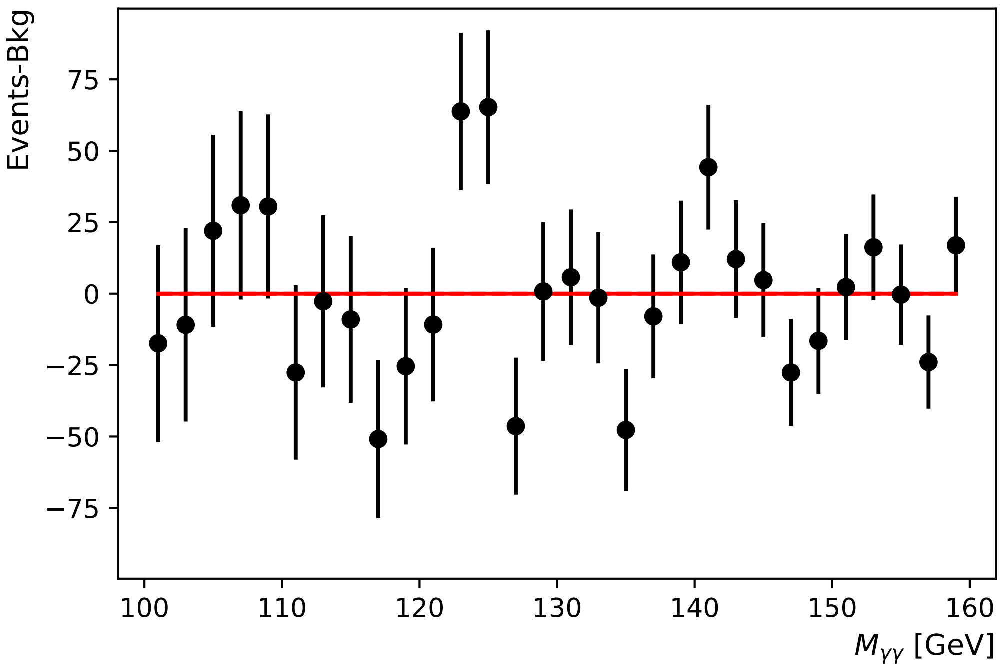


Figure 9.8: The residuals of Data D diphoton events with respect to the respective fitted background component. The dashed line represents the fourth order Bernstein polynomial background, the solid line represents signal plus background fit, while black data points represent real data events.

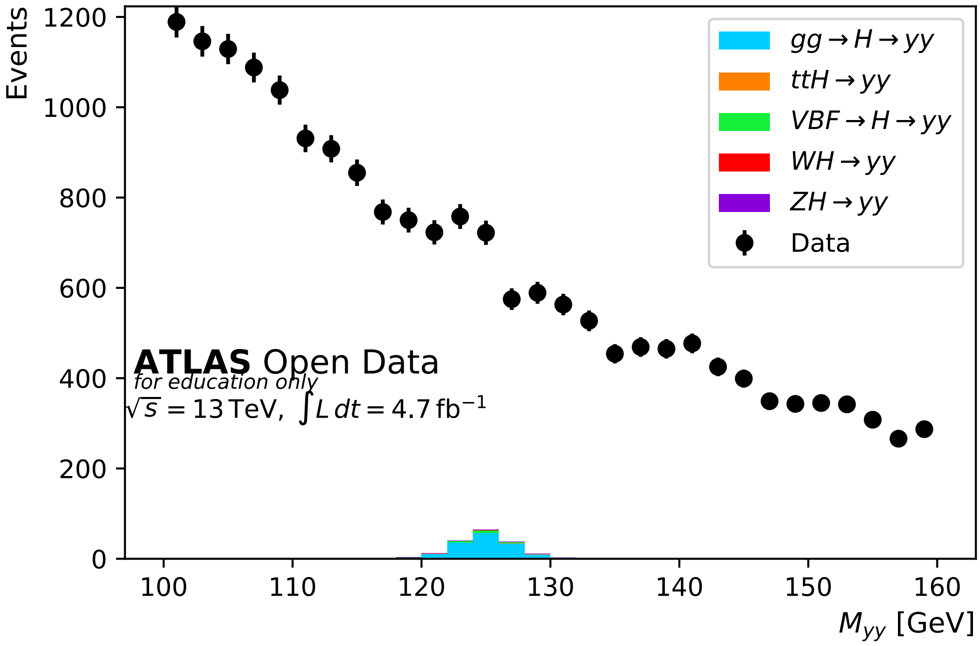


Figure 9.9: Resultant histogram showing number of events against the invariant mass of the diphoton candidate from Data D sample. The coloured bars represent MC data events from different production modes leading to diphoton decay of the Higgs boson seen in the legend. Highest background contribution is from the $gg \rightarrow H \rightarrow \gamma\gamma$ process in the considered diphoton invariant mass window.

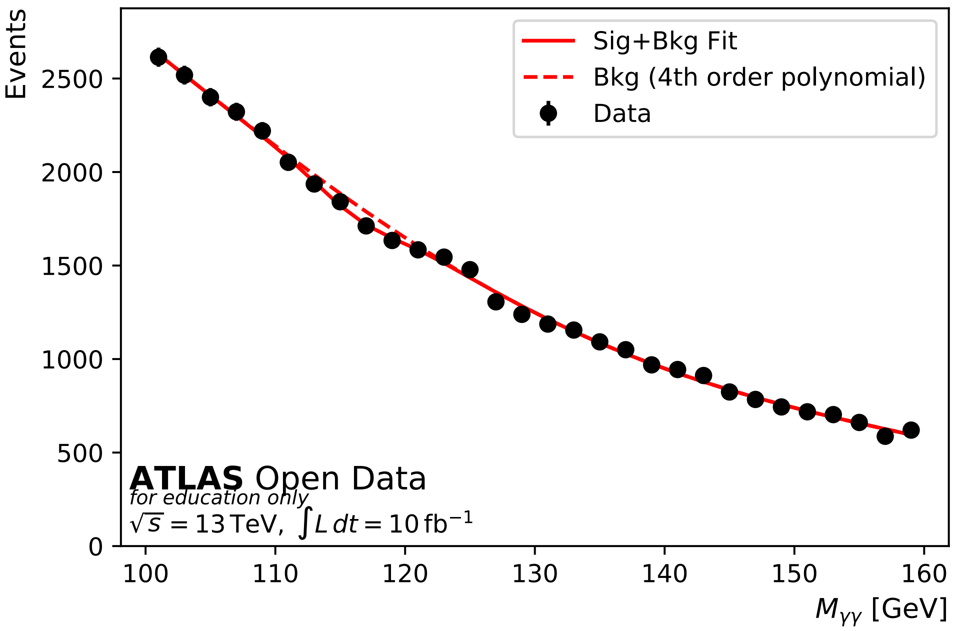


Figure 9.10: The distribution of the invariant mass of diphoton candidates after selection for the combined 13 TeV Data A, Data B, Data C, and Data D samples. The dashed line represents the fourth order Bernstein polynomial background, the solid line represents signal plus background fit, while black data points represent real data events.

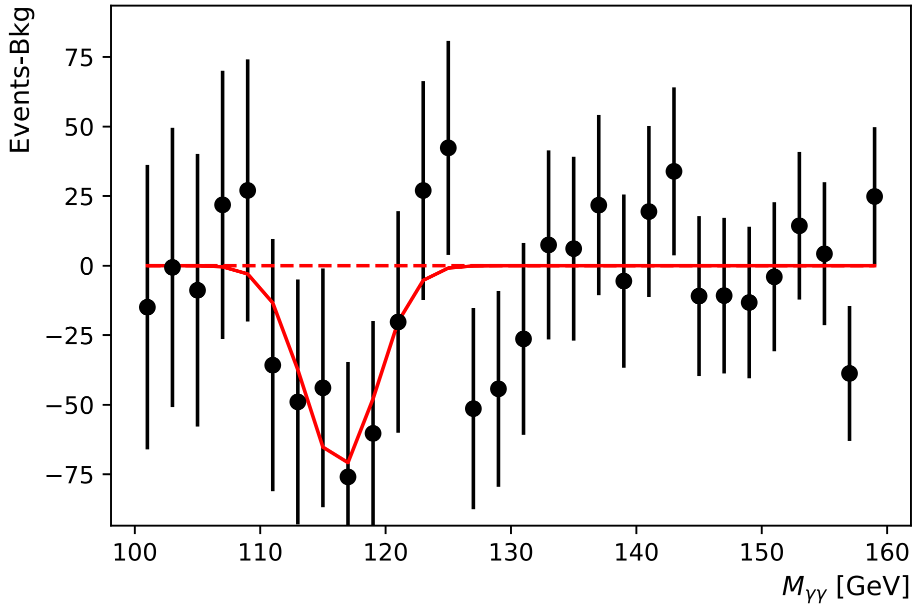


Figure 9.11: The residuals of diphoton events with respect to the respective fitted background component. The dashed line represents the fourth order Bernstein polynomial background, the solid line represents signal plus background fit, while black data points represent real data events. Analysis sample is combined from Data A, B, C and D samples.

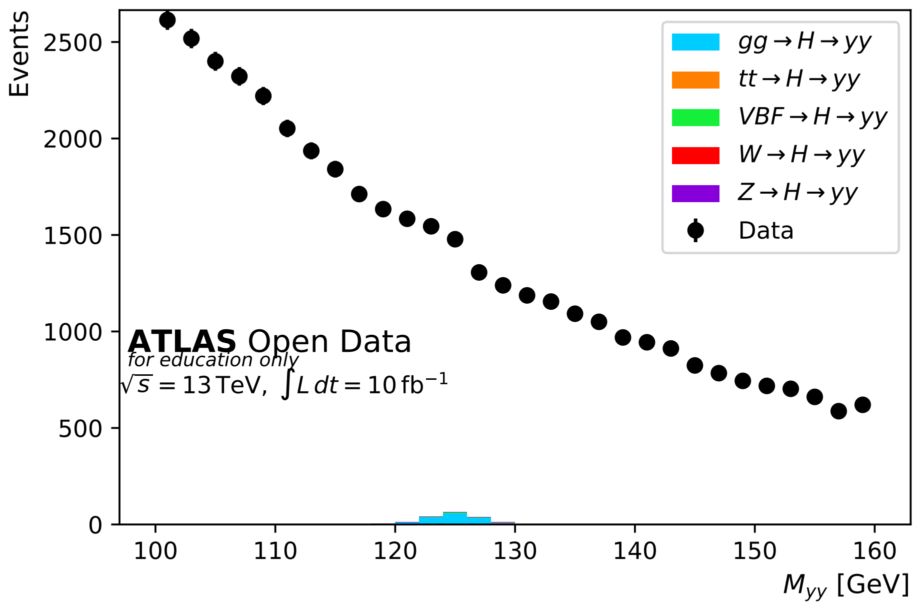
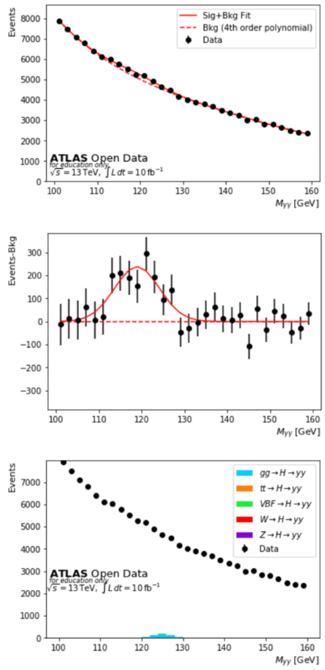


Figure 9.12: Resultant histogram showing number of events against the invariant mass of the diphoton candidates from combined samples Data A, B, C, and D. The coloured bars represent MC data events from different production modes leading to diphoton decay of the Higgs boson seen in the legend. Highest background contribution is from the $gg \to H \to \gamma\gamma$ process in the considered diphoton invariant mass window.

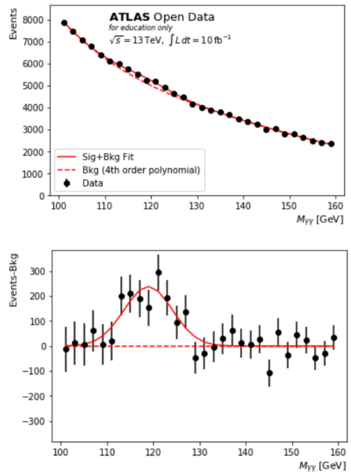
F. Cut flow for *Hγγ* Analysis on Data A Sample

```
data:
Photon_Input/data/data_A.GamGam.root
        Processing: data_A
                Time taken: 3.993748903274536, nIn: 430344, nOut: 430344
chi^2 = 69970.37014523176
Time taken: 29.36070418357849
Signal/sqrt(Background) for 105<mllll<160 1.9574518031359467
```



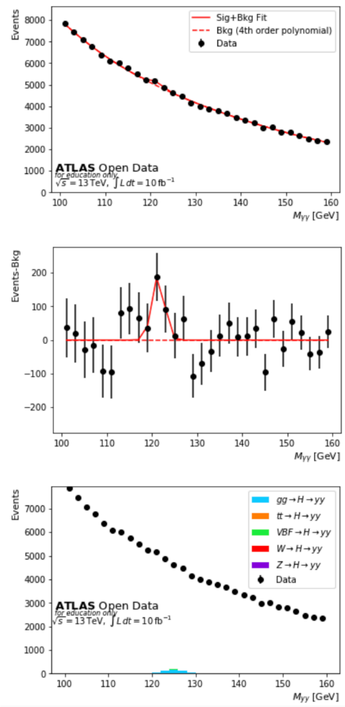
(1.S)

```
data:
Photon_Input/Data/data_A.GamGam.root
        Processing: data_A
                Time taken: 3.335598945617676, nIn: 430344, nOut: 430344
chi^2 = 69970.37014523176
Time taken: 4.198407888412476
```



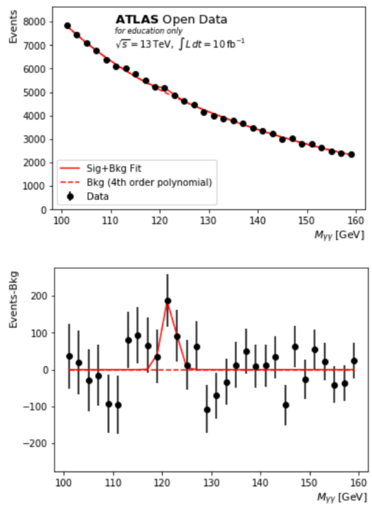
(1.M)

```
data:
Photon_Input/data/data_A.GamGam.root
        Processing: data_A
                Time taken: 4.169710874557495, nIn: 430344, nOut: 428881
chi^2 = 86665.17142942124
Time taken: 30.988357067108154
Signal/sqrt(Background) for 105<mllll<160 1.945273499264401
```



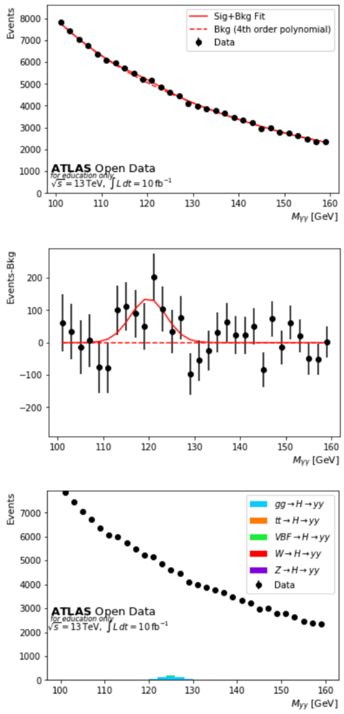
(2.S)

```
data:
Photon_Input/Data/data_A.GamGam.root
        Processing: data_A
                Time taken: 3.4003896713256836, nIn: 430344, nOut: 428881
chi^2 = 86665.17142942124
Time taken: 4.090299129486084
```



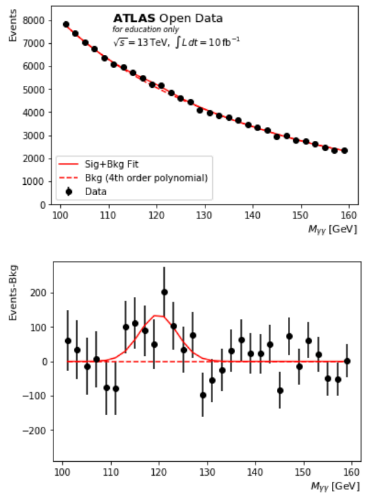
(2.M)

```
data:
Photon_Input/data/data_A.GamGam.root
        Processing: data_A
                Time taken: 6.369947195053101, nIn: 430344, nOut: 426923
chi^2 = 88184.00940731863
Time taken: 46.95648527145386
Signal/sqrt(Background) for 105<mllll<160 1.9319346333249983
```



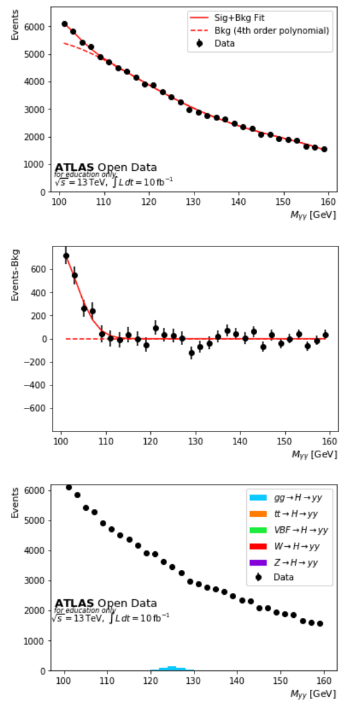
(3.S)

```
data:
Photon_Input/Data/data_A.GamGam.root
        Processing: data_A
                Time taken: 5.941542148590088, nIn: 430344, nOut: 426923
chi^2 = 88184.00940731863
Time taken: 8.823787927627563
```



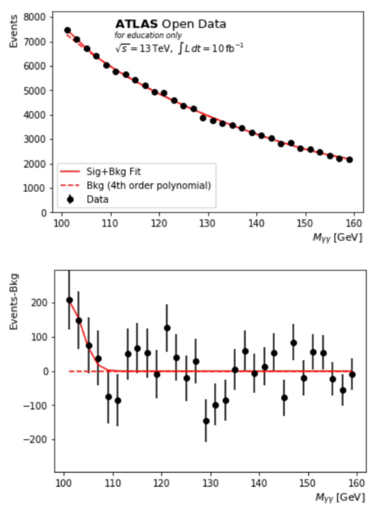
(3.M)

```
data:
Photon_Input/data/data_A.GamGam.root
        Processing: data_A
                Time taken: 9.987621068954468, nIn: 430344, nOut: 328558
chi^2 = 72623.51384703729
Time taken: 71.99606990814209
Signal/sqrt(Background) for 105<mllll<160 1.73052127370774
```



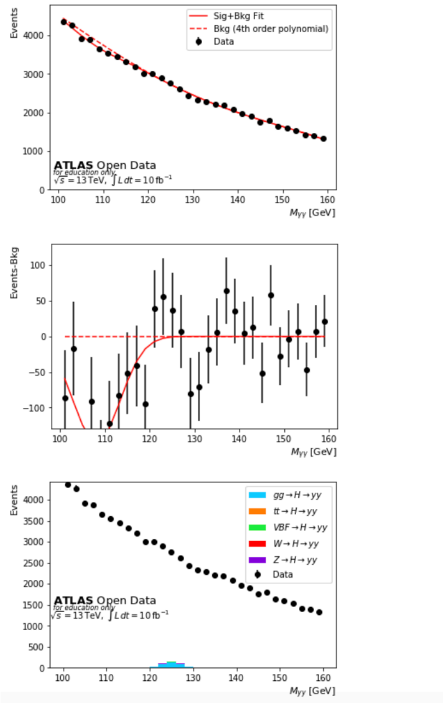
(4.S)

```
data:
Photon_Input/Data/data_A.GamGam.root
        Processing: data_A
                Time taken: 9.517445087432861, nIn: 430344, nOut: 405908
chi^2 = 110717.1626599769
Time taken: 10.442492961883545
```



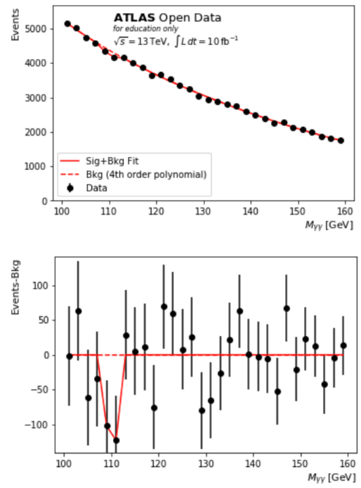
(4.M)

```
data:
Photon_Input/data/data_A.GamGam.root
        Processing: data_A
                Time taken: 10.983433961868286, nIn: 430344, nOut: 200612
chi^2 = 57412.81836223739
Time taken: 77.41481709480286
Signal/sqrt(Background) for 105<mllll<160 1.6468046811243877
```



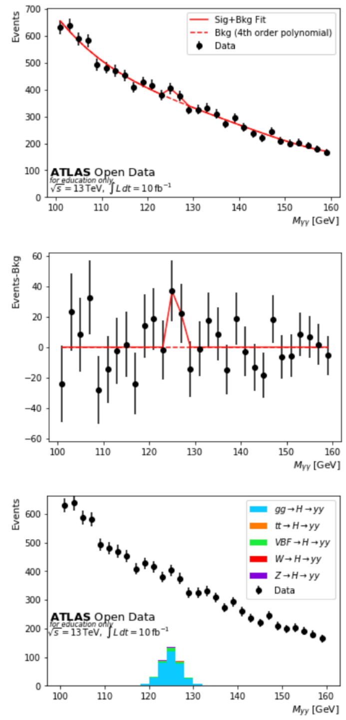
(5.S)

```
data:
Photon_Input/Data/data_A.GamGam.root
        Processing: data_A
                Time taken: 10.386703729629517, nIn: 430344, nOut: 248261
chi^2 = 50769.88916856576
Time taken: 10.984143257141113
```



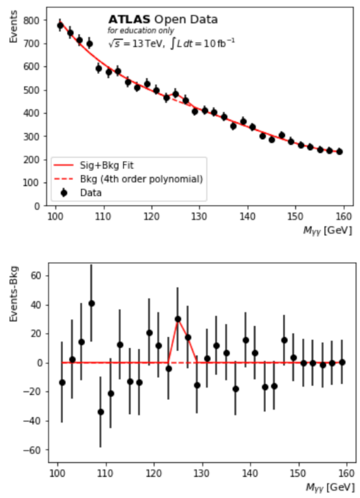
(5.M)

```
data:
Photon_Input/data/data_A.GamGam.root
        Processing: data_A
                Time taken: 11.784507036209106, nIn: 430344, nOut: 27986
chi^2 = 6610.947668555991
Time taken: 82.94891619682312
Signal/sqrt(Background) for 105<mllll<160 1.5409323722368171
```



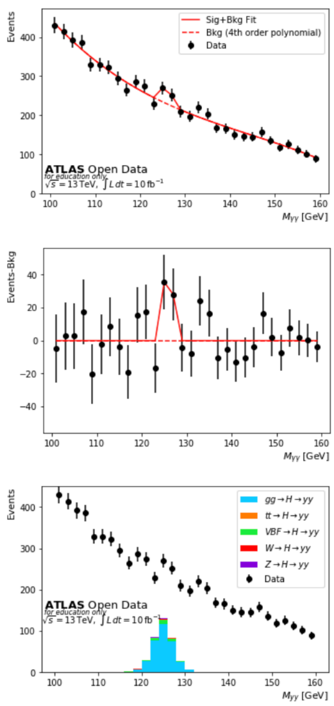
(6.S)

```
data:
Photon_Input/Data/data_A.GamGam.root
        Processing: data_A
                Time taken: 10.998628854751587, nIn: 430344, nOut: 34397
chi^2 = 6587.620305545649
Time taken: 11.535936117172241
```



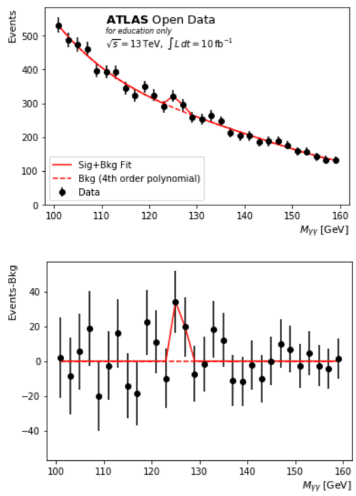
(6.M)

```
data:
Photon_Input/data/data_A.GamGam.root
        Processing: data_A
                Time taken: 12.715303182601929, nIn: 430344, nOut: 18281
chi^2 = 3796.4459699110153
Time taken: 96.82769298553467
Signal/sqrt(Background) for 105<mllll<160 1.5147434620877822
```



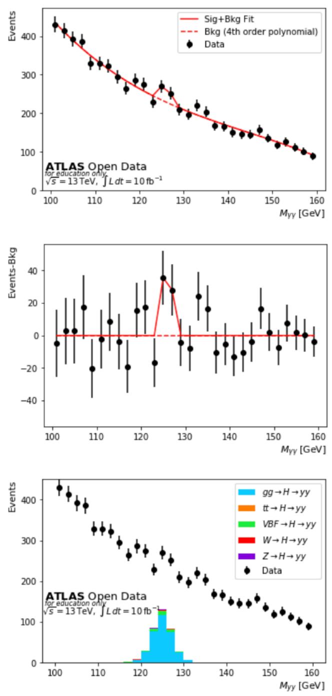
(7.S)

```
data:
Photon_Input/Data/data_A.GamGam.root
        Processing: data_A
                Time taken: 11.572202920913696, nIn: 430344, nOut: 22253
chi^2 = 3517.9367932658706
Time taken: 12.112674951553345
```



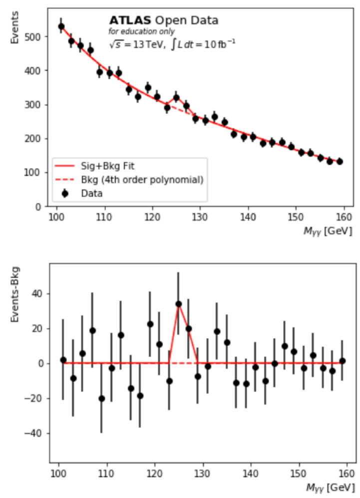
(7.M)

```
data:
Photon_Input/data/data_A.GamGam.root
        Processing: data_A
                Time taken: 12.93391728401184, nIn: 430344, nOut: 9386
chi^2 = 3796.4459699110153
Time taken: 98.20303082466125
Signal/sqrt(Background) for 105<mllll<160 1.5147434620877822
```

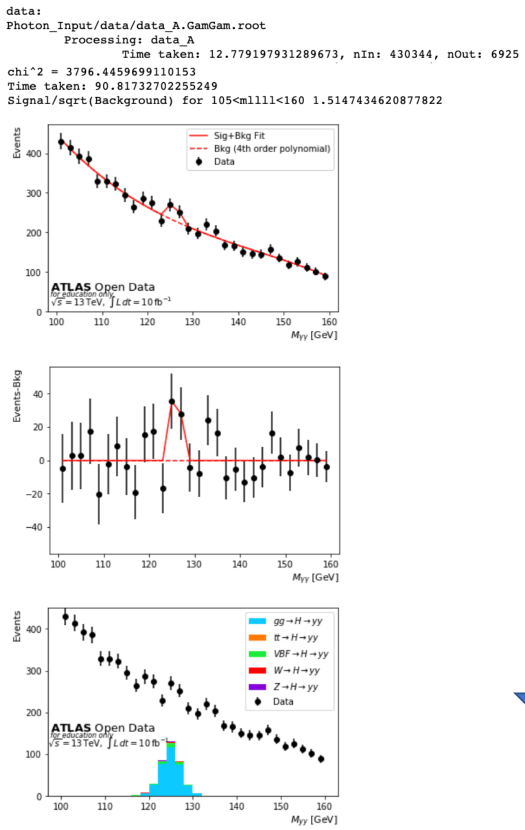


(8.S)

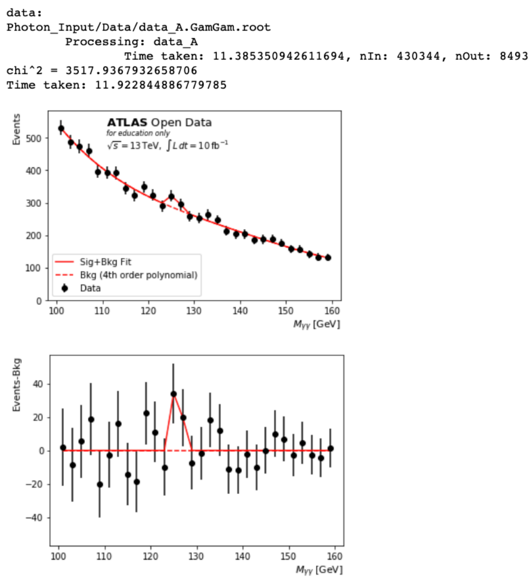
```
data:
Photon_Input/Data/data_A.GamGam.root
        Processing: data_A
                Time taken: 10.967936992645264, nIn: 430344, nOut: 12245
chi^2 = 3517.9367932658706
Time taken: 11.484228134155273
```



(8.M)



(9.S)



(9.M)

Table 9.1: Legend of cut flow stages showing number of stage/analysis output along with respective cuts performed at that stage/analysis output for $H\gamma\gamma$ Analysis on Data A Sample. The suffix ".S" stands for the analysis output obtained through my version of the code adapted from Mr. Meirin Evans' version , while ".M" suffix stands for the analysis output obtained through Mr. Meirin Evans' version of the code.

Analysis Output /Stage	Cuts Performed
(1.)	Cut on reconstructed diphoton invariant mass: data['myy'] = np.vectorize(calc_myy)(data.photon_pt,data.photon_eta,data.photon_phi)
(2.)	Cut (1.) + Cut on number of photons: fail =data[np.vectorize(cut_photon_n)(data.photon_n)].index
(3.)	Cut (2.) + Cut on pseudorapidity 1: fail =data[np.vectorize(cut_photon_eta)(data.photon_eta)].index
(4.)	Cut (3.) + Cut on pseudorapidity 2: fail =data[np.vectorize(cut_photon_eta2)(data.photon_eta)].index

(5.)	Cut (4.) + Cut on transverse momentum of the photons: fail =data[np.vectorize(cut_photon_pt)(data.photon_pt)].index
(6.)	Cut (5.) + Cut on photon reconstruction: fail =data[np.vectorize(cut_reconstruct)(data.photon_isTightID)].index
(7.)	Cut (6.) + Cut on isolation: fail =data[np.vectorize(cut_isolation2)(data.photon_etcone20)].index
(8.)	Cut (7.) + Cut on lower limit on reconstructed invariant mass: fail =data[np.vectorize(cut_mm_r)(data.photon_pt,data.photon_eta,data.photon_p hi)].index
(9.)	Cut (8.) + Cut on upper limit on reconstructed invariant mass: fail =data[np.vectorize(cut_mu_r)(data.photon_pt,data.photon_eta,data.photon_p hi)].index