February 24th, 2021

Chris Peterson

Sunny Ng(partner)

Measurement of Higgs Boson production via the diphoton decay channel

Abstract:

Number of Higgs boson decay events and Higgs boson mass was measured to confirm the discovery of the Higgs boson. Analysis of the 13TeV ATLAS Open Data set was performed to isolate events in the H-> yy diphoton decay channel. Diphoton decay events were isolated by applying selection criteria for photon pairs present in the ATLAS data. Statistical analysis was performed to demonstrate a peak in diphoton decay events corresponding to the Higgs mass. The Higgs mass was found to be 124.76 ± .86GeV. Using Monte Carlo simulation for Higgs with mass 125.09 ± 0.24GeV, number of Higgs decay events found in data was compared to number of events predicted by MC simulation. The chi^2 value of the Higgs peak found in data compared to MC prediction was found to be 1.92 with probability .983. Thus, validity of MC prediction was demonstrated confirming the discovery of the Higgs boson.

Introduction:

In the beginning of the 1900's a number of fundamental particles were discovered. During the 1960's many physicists were working on theories which proposed new fundamental particles but many approaches, including the yang-mills model and Goldstone's theorem, made predictions of massless particles which did not agree with experimental results. In 1964 three groups, one of which included Peter Higgs, independently developed relativistic models which addressed this problem.

In 1964 Higgs suggested an additional mechanism which would predict the gauge bosons to have finite mass. [1] He suggested that an additional field in Yang-Mills gauge theory could spontaneously break the electro-weak symmetry in Glashow's model for the weak and electro-magnetic interactions. In 1966 Higgs published another paper in which he proposed the Higgs field and the Higgs boson, as well as the decay mechanism of the Higgs boson. [2] His prediction of the Higgs boson decay mechanism introduced a method for experimental confirmation of the theory.

While Higgs provided a rigorous mathematical theory, its predictions have proved difficult to confirm experimentally. In order to produce the Higgs boson two beams of particles must be accelerated to relativistic speeds then the particles must collide. The Higgs will only be produced if the particle colliders operate at a high enough energy; The required energy depends on the mass of the Higgs, which was unknown at the time. Thus, between the 1960's and the year 2000 particle colliders were made at higher and higher energies hoping to discover the Higgs, with the Large Electron-Positron Collider (LEP) being the highest energy at 209GeV. LEP found no evidence for the Higgs boson; it did not operate at high enough energy. By the year 2000 evidence from LEP had determined that if the Higgs boson did exist its energy must be greater than 114.4GeV/c^2. [3]

The highest energy particle collider to date is the Large Hadron Collider (LHC). At the LCH beams of protons are accelerated to speeds up to 99.9999991% of the speed of light before colliding with each other, giving them a collision energy of up to 13TeV. Once the protons collide various other particles predicted by the Standard Model (SM) are produced and measured by several tracking detectors and calorimeters. The calorimeter measures the energy of the particles created in the collision and the tracking detectors measure their position. The LHC can collide extremely large numbers of particles, collecting data on over $3x10^{14}$ proton-proton collisions. A large number of collisions is necessary to collect statistically significant data, since the Higgs boson production is predicted to be roughly 1 per 10 billion collisions and other background SM events must be screened out. Using the LHC CERN was able to experimentally confirm the prediction of the Higgs boson. [4]

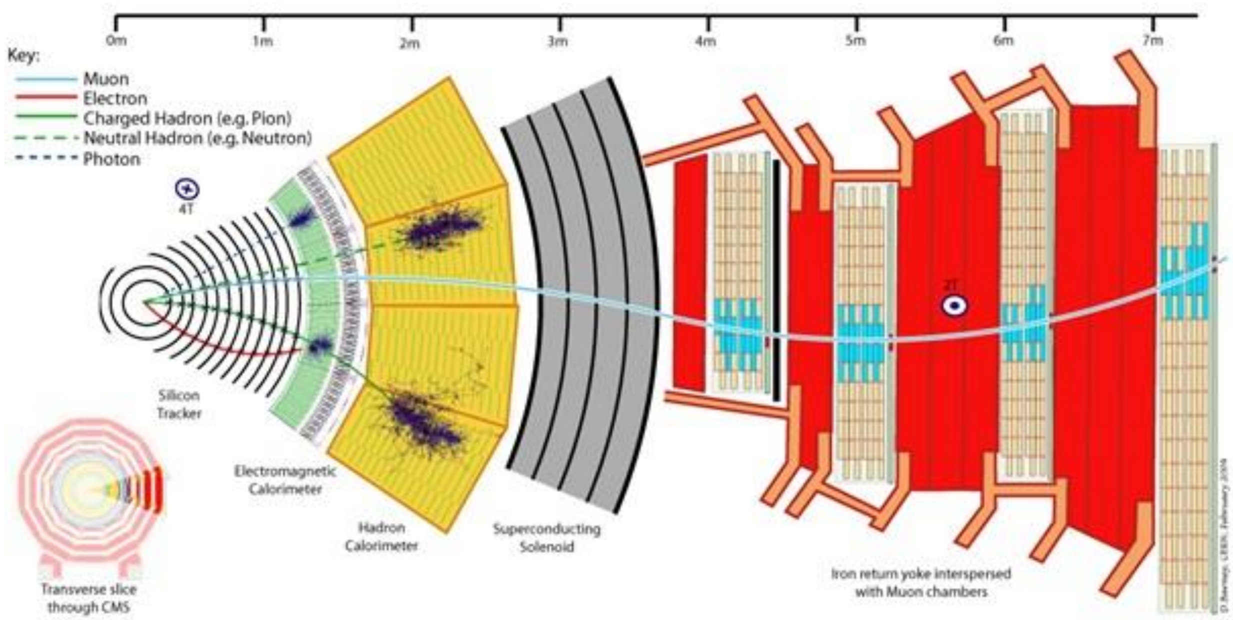

Figure 1: The detector layout in the LHC is shown, with paths for different particles. Photons will be detected by the energy absorbed in the calorimeter. [8]

Measurements are made at LHC by a variety of detectors. The one relevant to our experiment is the calorimeter. When particles detected by the calorimeter cross the detector, they are detected by the energy absorbed by the calorimeter. The calorimeter is made up of layers of high-density material, such as lead, which absorb all the energy in the particles they detect preventing them from passing through. Analyzing the energy absorbed can be used to identify which particle was present.

In this experiment we confirm the discovery of the Higgs boson at the LHC. The Higgs boson is a massive scalar boson with no electric charge and zero spin. When produced it decays too quickly to be directly measured. Instead, particles produced in the Higgs decay mechanism must be measured to determine if the Higgs boson was produced.

When the Higgs boson is produced it decays with a mean half-life of $1.6x10^{-22}$s. There are six decay channels in total, however the three most common account for .852 of the branching ratios. These include the Higgs to bottom-antibottom quark pair, Higgs to pair of W bosons and Higgs to tau-antitau pair with branching ratios .577, .215 and .06, respectively. In this experiment we focus on detecting the Higgs through the H-> yy decay channel, the 6^{th} most probable decay channel, which has a predicted branching ratio of $2.28x10^{-3}$.

Despite the low branching ratio, the high photon reconstruction efficiency of the ATLAS detector enables us to identify a high yield of diphoton decay. In addition, the high resolution of the ATLAS calorimeter enables us to identify the narrow Higgs peak, making the H->yy a cleaner identification process for Higgs detection.

In order to detect the Higgs boson decay, we analyze the 13TeV ATLAS Open Data set, which contains data on 270 million proton-proton collisions measured at the LHC. We apply an analysis technique of implementing strict selection requirements on diphoton candidates for Higgs decay. We then fit the data with a 3^{rd} order polynomial + gaussian fitting function to find the probability of getting our data if the MC prediction is correct. We then compare this with the probability given the null hypothesis, where the null hypothesis is the absence of the Higgs decay contribution to diphoton events in our data.

We show that the Higgs boson mass agrees with the currently accepted value of 125.10+/-.14 by statistical analysis of our data compared with the decay patterns for the Higgs predicted by MC simulation.

Materials and Methods:

In this experiment we investigate the evidence for the Higgs Boson by analysis of diphoton decay events in data released by ATLAS. Data was collected on the decay events present in proton-proton collisions at CERN using the LHC. The methods below describe the techniques used to determine the number of decay events and mass of the Higgs boson.

We used the 13TeV ATLAS Open Data set along with the virtual machines and ROOT analysis frameworks provided by ATLAS. [5] Code was written in Python to isolate diphoton decay events in both the data and MC simulation files by implementing strict selection criteria.

The data files were added to ROOT TChains, then an analysis class was defined. Within the analysis class the following selection criteria were applied to the four data files, as well as the five MC simulation files found on the 13TeV ATLAS Open Data website. [5]

We first selected for the diphoton trigger, trigP. We are interested only in events with two photons meeting criteria which indicate they were produced via Higgs decay and this selection cuts out all events which do not contain two photons.

We then selected for photon_ptcone30 < .065 and photon_etcone20 < .065, which are defined as the standard photon selection requirements by the ATLAS technical note. [6] Photon etcone20 is the sum of the energies found by the calorimeter in clusters with within a cone of $\Delta R = 0.2$ around the particle and photon ptcone30 is the sum of the transverse momentum found within a cone of $\Delta R = 0.3$.

These selection criteria are used isolate from detector measurements signals which correspond to photons.

Next a selection was made for exactly two photons, one with photon_pt > 25GeV and one with photon_pt > 35GeV. Photon pt is the transverse momentum of the photon, the energy in the plane transverse to the particle beam. We select for these values of transverse momentum to isolate diphoton candidates of Higgs decay from other events which contain photon pairs missing the required energy to be produced from Higgs decay.

A selection for the absolute value of photon_eta less than 2.37 and excluding values between 1.37 and 1.52 was applied. Eta, the pseudo-rapidity, is the angle between photon and collider beam. This cut is made for technical reasons related to the detector design; We exclude the transition region between the barrel and end cap of the detectors.

Using TLorentz vectors from the ROOT package, we calculated the invariant mass of the diphoton candidates. While photons are massless particles, they have invariant mass due to their momentum. To further select for only diphoton candidates with enough energy to be produced from Higgs decay, we selected for exactly one photon with photon_pt/$m_{\gamma\gamma}$> 35GeV and one photon with photon_pt/$m_{\gamma\gamma}$> 25GeV, where $m_{\gamma\gamma}$ is the invariant mass of the diphoton TLorentz vector.

Finally, we select for a diphoton invariant mass to be 105GeV < $m_{\gamma\gamma}$ < 160GeV. This is to further isolate diphoton candidates produced from Higgs decay, as the Higgs mass was discovered to be 125.09 ± 0.24GeV. [4]

Once these selection criteria identified diphoton candidates, we weighted them to fill TH1F histograms. We used a weight of 1 to scale the histogram made for the data files. The product of mcWeight, PhotonTRIGGER, scaleFactor_PILEUP, and scaleFactor_PHOTON was used to weight the MC files. Because the not every photon produced in the collider is able to be detected by the calorimeter, this weight is used to reduce the weight of the MC prediction by an amount that would be expected to be undetected or unidentified by the detector.

We then scaled the individual MC histograms by a scale factor before stacking them into a single histogram. We scaled them by (luminosity * cross section) divided by the sum of weights. Luminosity is the number of events per area, and cross section is the area corresponding to the area of the colliding beams. Hence their product is the expected number of events, divided by the sum of the weights to account for weighting described in previous step. The MC predictions were made by specifying the number of events, and this scale factor scales the MC predictions to number of events expected from the data collected in the particle collider.

Once the histograms were made a gaussian fit was applied to the MC histogram. We then used a 3^{rd} order polynomial + gaussian to fit the data histogram, which for future reference we call sig+bkg. Parameters obtained from the gaussian fit on the MC histogram were used to set the initial parameters for the sig+bkg fit on the data.

The background was assumed to be smoothly falling, with a peak near 125GeV from the Higgs decay. Thus, the parameters found from the 3^{rd} order polynomial part of the 3^{rd} order polynomial + gaussian fit on the data were used to define a 3^{rd} order polynomial that fit the background diphoton events, which we refer to as bkg.

The sig+bkg and bkg fits were used to calculate the Chi^2 and probability of Chi^2 for our data. We calculated Chi^2 and probability of Chi^2 for the sig+bkg fit on the data, the bkg fit on the data, and for the gaussian fit of the Higgs peak compared with both MC prediction and null hypothesis. These probabilities represent the probability that our data represent a fluctuation from expected values with the Higgs contribution included, and against a null hypothesis not including Higgs decay contribution. The probability of Chi^2 was done using ROOT in terminal, using values of Chi^2 and NDF obtained from the fitting functions.

By subtracting the bkg fit from the sig+bkg fit, we determined the Higgs peak and width of the peak. The integral of the sig+bkg fit minus bkg fit, divided by bin width, was used to compute the number of events within the Higgs peak. We integrated from 115 to 135GeV, in order to compare the data with the gaussian fit on the MC prediction.

We then repeated the same process varying the selection criteria on transverse momentum. We performed the same analysis on the data files by varying the pt cut by +/-5GeV. This was done to establish that our cuts were justified, as well as to check the validity of our method in finding the Higgs peak.

Next we used another 3rd order polynomial + gaussian fit to find a 2nd peak in addition to the Higgs peak. Finding a 2nd peak was used to compare the probability of this peak in our data with the Higgs peak, in order to further demonstrate the statistical significance of the Higgs peak finding. To avoid finding the Higgs peak again with this fit, we used fit parameters which limit the gaussian part of the fit to an identified peak in the data.

A gaussian fit on the MC histogram and sig+bkg – bkg was used to create a histogram comparing the Higgs peak with MC prediction. Data points with errors were added to the histogram by subtracting bkg from the data histogram.

Results:

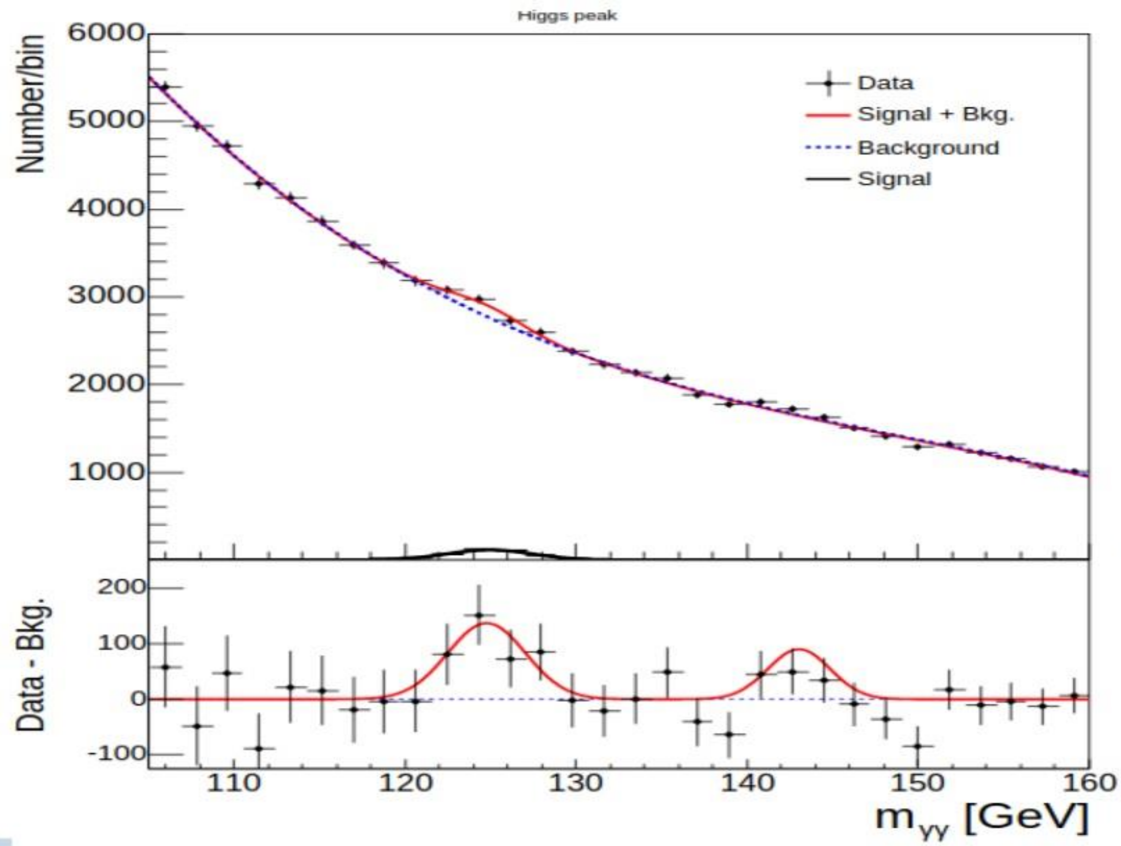

Figure 2: Diphoton invariant mass vs number of events. The top panel includes the data points taken from the data files shown as crosses indicating the top of the bins. The black histogram bins in top panel are taken from MC prediction. Sig+bkg is shown in red, bkg shown in blue. The Higgs peak at 124.76+/-.86GeV is shown in red. The bottom panel shows the difference between sig+bkg – bkg, with the included 2nd peak found at 142.57+/-1.30GeV. Error bars in bottom panel indicate the statistical uncertainty of data – bkg. The small black histogram is the MC prediction for Higgs decay diphoton contribution.

Using the gaussian for sig+bkg – sig we found a μ = 124.76+/-.86GeV, corresponding to the Higgs peak, with σ = 2.27+/-.27GeV, indicating a width of 2σ = 4.48+/-.54Gev for the Higgs peak. Integral of (sig+bkg – sig) divided by 1.83 for bin width, from 115 to 135GeV found a total of 397.81+/-19.45 events within the Higgs peak.

The statistical significance of our data both with and without the prediction of Higgs decay expected contribution was computed. The Chi^2 of the sig+bkg fit was found to be 19.44, with a probability of .675 and the Chi^2 of the bkg fit was found to be 27.82, with a probability of .367. This indicates it is more probable that our data result from fluctuations in diphoton production with the existence of the Higgs and its decay contribution than without.

A 2nd peak was fit to the data, as shown in the bottom panel of figure 2. For the gaussian fit on the 2nd peak, we found a μ = 143.025+/-.91GeV corresponding to the peak and a σ = 1.81+/-.60GeV corresponding to a width of 2 σ = 3.62+/-1.2GeV. Total events above background in the 2nd peak, found by integrating the polynomial + 2nd peak gaussian fit on the data minus bkg from 133 to 153 divided by bin width of 1.83, was found to be 114.96+/-10.72. The Chi^2 for the 2nd was found to be 11.44 with a probability of .179. The low Chi^2 probability found indicates the low statistical significance of this peak in the data.

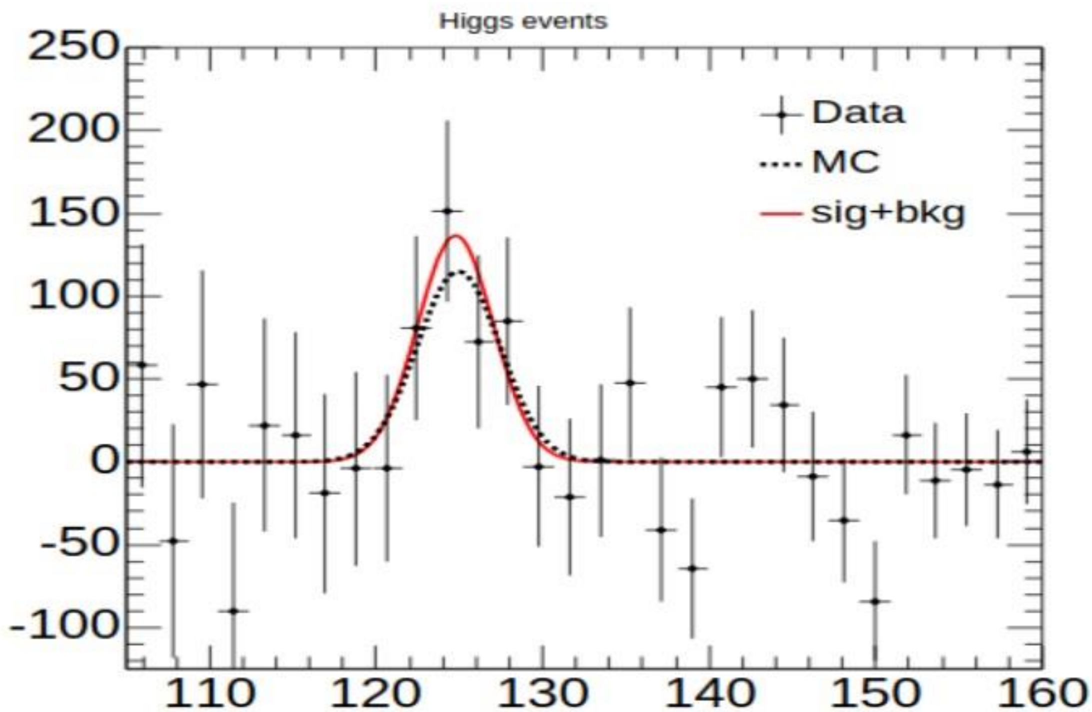

Figure 3: Sig+bkg – sig is shown in red. Gaussian fit of the MC data is shown in dotted black. Data shown with error bars calculated from data – bkg.

A comparison between sig+bkg – sig and the gaussian fit on the MC data is shown in figure 3. The Chi^2 for the sig+bkg – bkg fit between 115 and 135GeV was 1.92. The probability of this Chi^2 was found to be .983, indicating a high probability that our data result from random fluctuations from expected values given the hypothesis of the Higgs decay contribution to diphoton decay events. A second fit was imposed on data – bkg, fixing the parameters to be the constant 0 function expected if there were no Higgs contribution. The Chi^2 for this fit was found to be 14.87, with a probability of .136 indicating low probability that our data result from random fluctuations in decay events without Higgs decay contribution.

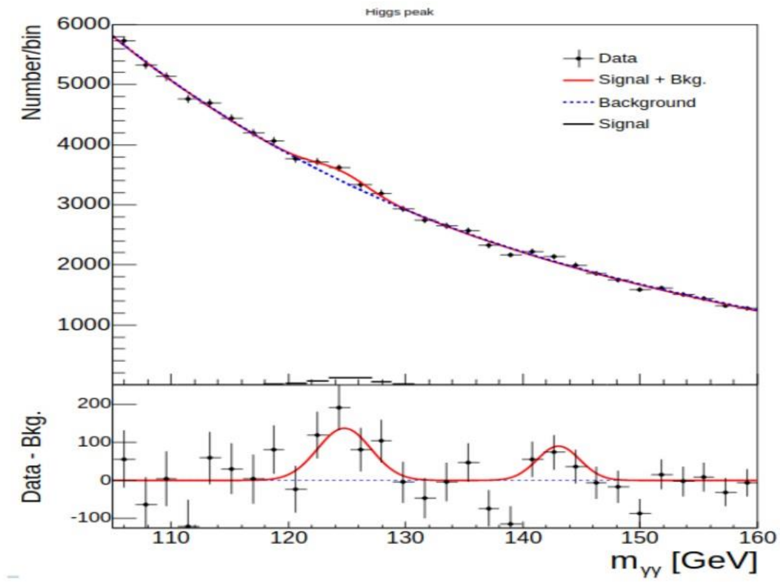

Figure 4: Diphoton invariant mass vs number of events shown for loosening the pt cut by -5GeV. The top panel includes the data points taken from the data files shown as crosses indicating the top of the bins. The black histogram bins in top panel are taken from MC prediction. Sig+bkg is shown in red, bkg shown in blue. The Higgs peak at 124.52+/-.76GeV is shown in red. The bottom panel shows the difference between sig+bkg – bkg, with the included 2nd peak found at 143.46+/-.86GeV. Error bars in bottom panel indicate the statistical uncertainty of data – bkg. The small black histogram in top panel is the MC prediction.

A second analysis was done by loosening the photon transverse momentum cut by -5GeV, shown in figure 4. Using the gaussian for sig+bkg – sig on the loose cut we found a μ = 124.52+/-.76GeV, corresponding to the Higgs peak, with σ = 2.31+/-.25GeV, indicating a width of 2σ = 4.62+/-.50Gev for the Higgs peak. Integral of (sig+bkg – sig) divided by 1.83 for bin width, from 115 to 135GeV found a total of 499.53+/-22.35 events within the Higgs peak. Chi^2 for this fit was found to be 5.19, with probability of .737.

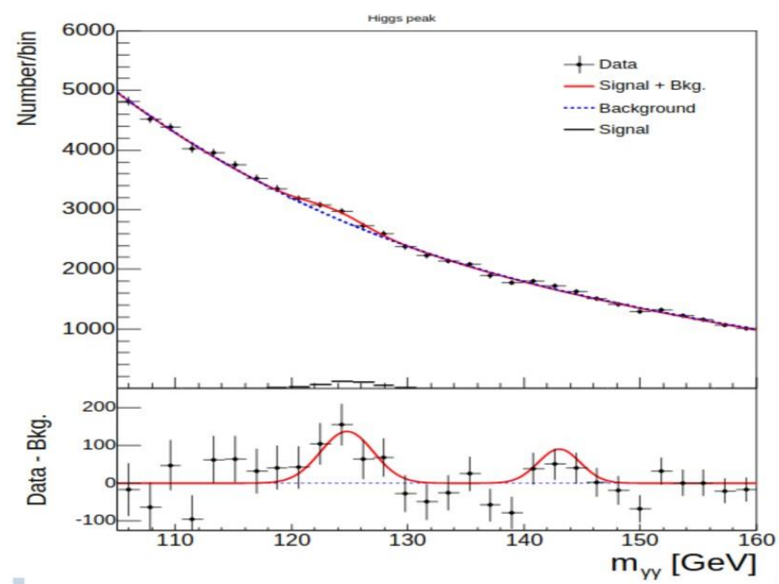

Figure 5: Diphoton invariant mass vs number of events shown tightening the pt cut by +5GeV. The top panel includes the data points taken from the data files shown as crosses indicating the top of the bins. The black histogram bins in top panel are taken from MC prediction. Sig+bkg is shown in red, bkg shown in blue. The Higgs peak at 123.86+/-.92GeV is shown in red. The bottom panel shows the difference between sig+bkg – bkg, with the included 2nd peak found at 142.92+/-1.23GeV. Error bars in bottom panel indicate the statistical uncertainty of data – bkg. The small black histogram in top panel is the MC prediction.

A third analysis was done by tightening the photon transverse momentum cut by +5GeV, shown in figure 5. Using the gaussian for sig+bkg – sig on the tight cut we found a μ = 123.86+/-.92GeV, corresponding to the Higgs peak, with σ = 1.46+/-.68GeV, indicating a width of 2σ = 2.92+/-1.36Gev for the Higgs peak. Integral of (sig+bkg – sig) divided by 1.83 for bin width, from 115 to 135GeV found a total of 445.53+/-21.11 events within the Higgs peak. Chi^2 for this fit was found to be 4.35, with probability of .824.

Conclusion:

We found the value of the Higgs mass to be 124.76+/-.86GeV, which is in agreement with the currently accepted value of 125.10+/-.14. [7] We found a p value of .983 for the Higgs peak, confirming the validity of the model for Higgs diphoton decay mechanism used to generate the MC prediction. The discovery of the Higgs boson found at the LHC in 2012 was confirmed. [4]

Bibliography:

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

[2] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. **145**, 1156 (1966).

[3] J. A. Heister, S. Schael, R. Barate, *Search for the Standard Model Higgs Boson at LEP*, Phys. Lett. Sect. B Nucl. Elem. Part. High-Energy Phys. **565**, 61 (2003).

[4] G. Aad, T. Abajyan, B. Abbott, *Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC*, Phys. Lett. Sect. B Nucl. Elem. Part. High-Energy Phys. **716**, 1 (2012).

[5] ATLAS Open Data 13TeV Documentation. https://atlas-opendata.web.cern.ch/atlas-opendata/release/2020/documentation/atlas/lhc.html [Accessed 14 Dec. 2020]

[6] Review of the 13TeV ATLAS Open Data set, technical documentation. https://cds.cern.ch/record/2707171/files/ANA-OTRC-2019-01-PUB-updated.pdf

[7] P.A. Zyla; et al. (Particle Data Group) (2020). https://pdg.lbl.gov/2020/listings/rpp2020-list-higgs-boson.pdf (PDF). *Progress of Theoretical and Experimental Physics*: 1

[8] LCH detecors, https://www.lhc-closer.es/taking_a_closer_look_at_lhc/1.detectors