The Higgs boson and other animals

Prof. William Murray
STFC-RAL / Warwick University

May 17, 2016

This document contains some comments on the paractical methods used for the Higgs boson discovered at CERN in 2012, and consideres their current application to the di-photon resonance.

Chapter 1

The Higgs boson

1.1 LHC as a Higgs factory

The discovery, or falsification, of the Brout-Englert-Higgs theory of spontaneous symmetry breaking was one of the major goals of LHC when it was first proposed in 1983. As is well known, the interactions with the weak bosons, W and Z, are integral to the model, but it is used in the Standard Model to give mass to all the fermions too.

If correct, then the Higgs boson couples to fermions in proportion to their mass, and so will be preferentially created by, and decay to, heavy objects. The production at LHC is not simple, because the valence constituents of a proton, up and down quarks, are very light. The main production mechanisms at LHC are shown in figure 1.1

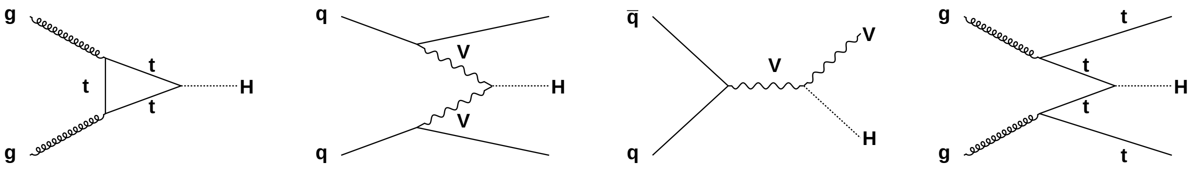

Figure 1.1: Diagrams for Higgs production. There are referred to as ggF, VBF, VH and ttH. V refers to a W or a Z boson.

There are various diagrams; in each case a heavy object, a top quark, W or Z boson, is created and it may or may not survive the production of the Higgs boson. The highest rate is the first, known as gluon fusion. Here the only thing made is a Higgs boson, and therefore we will have to rare decays of the Higgs boson to trigger the readout. In the other modes there is at least the possibility to trigger on the accompanying activity.

Before the LHC started, the LEP experiments had excluded the SM Higgs with mass below 114.4 GeV/c^2, and the consistency of the electroweak fit set upper limits at about 150 GeV/c^2. The relative rates of the main Higgs production and decay modes at 125 GeV can be seen in figure 1.2.

The largest fraction decay to b quarks, but the jets arising from b quarks are difficult to distinguish from all the other jets which dominate LHC interactions. There are no experimental results on the Higgs boson produced in the highest rate, gluon fusion, process and decaying to the highest rate, bb, state.

Second largest is the WW, which is now the best measured channel, at least in ATLAS. However, this has involved painstaking analysis of the decay of both W boson into a lepton and an invisible neutrino. The presence of two neutrinos means that the four-vectors of neither the W bosons not the Higgs can be reconstructed, so there is no sharp peak visible. There are many background sources, each of which had to be calibrated in data control samples. The discovery was primarily through two rare, but cleaner decays: $\gamma\gamma$ and ZZ.

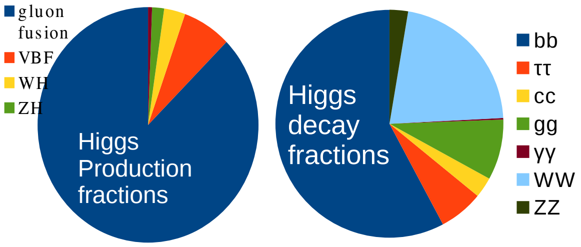

Figure 1.2: The pie-charts above show the fraction of the different production modes (left) and decay modes (right).

1.2 Higgs to $\gamma\gamma$

The Higgs boson decay to $\gamma\gamma$ is not a leading-order process in the SM, because the photon is massless and has no coupling to the Higgs boson. It decays through WW and/or tt loops. These are both massive and charged and so interact with the Higgs and the photon. For Higgs masses from 110 to 135 GeV/c^2 the decay to $\gamma\gamma$ has sufficient rate to be usable, but is it is small - about 0.24%.

The background is fairly featureless, because there are no known 2-body decays producing γ pairs from high mass processes.[1] The primary background is events where quark and antiquark annihilate producing two photons, as for example 1.3 (left). Gluons are neutral and do not couple to photons so gluon interactions can only produce a photon pair if they first creates quarks.

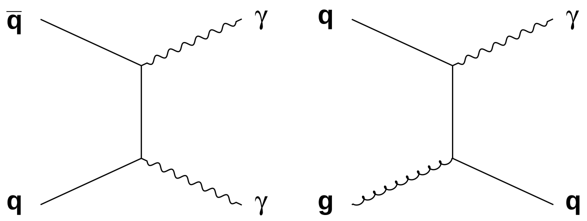

Figure 1.3: Left is the pair production of photons. Right the production of one photon and one jet.

There is also significant background from events with only one real photon, such as figure 1.3 (right). This is a much more common process than the diphoton production, but the detectors are unlikely to mistake a quark for a photon. Still, at some level it happens. For example, as the quarks fragments into a jet it might end up putting nearly all the energy into one π^0. This decays to two photons, but if they are sufficiently close together, as is likely for high p_T, they can be mistaken for just one.

Neither di-photon nor photon plus jet have any particular mass expected. They have not passed through an s channel resonant state. The cross-section of the process decreases as the energy rises, and so each gives a smooth, falling, mass spectrum.

The analysis has two main problems: to efficiently identify photons with low background and to measure the invariant mass of the Higgs candidate as well as possible.

---

[1] Toponium decay to $\gamma\gamma$ is not expected to be measureable at LHC.

1.2.1 Trigger and Photon identification

The trigger required two photon candidates, with transverse momenta repectivey 35 and 25 GeV. This is very efficient for a mass over 100 GeV.

The first starts from requiring that the pattern of energy deposition be EM-like. This uses the spread of the shower laterally (and perhaps in depth) and the lack of energy in the hadron calorimeter to identify that it is an EM shower. Then the absence of a charged particle track is used to distinguish photons from electrons. Unfortunately the amount of material in the tracking detectors, at the very least 20% X_0, means that photons can convert into e^+e^- pairs. It may be that one of the electrons is not reconstructed (e,g, if low momentum), and then the photon would look like an electron, so normally the experiments require that the silicon pixel detector closest to the beam interaction point does not have a hit consistent with the track.

The background from jets where most of the energy has gone into one $\pi^0 \rightarrow \gamma\gamma$ is minimised by requiring 'isolation': that there not be much energy in the calorimeter or momentum in tracks in the tracker, in a cone around the photon direction.

The composition of the selected events is characterised by studying the particles which are nearly accepted as photons. They is used to estimate the composition in terms of $\gamma\gamma$, γ-jet and jet-jet, see figure 1.4. DY refers to 'Drell-Yan' events, e^+e^- production where both electrons are mistaken for photons which is in fact pretty rare.

The selected events are over 75% real pairs of prompt photons, and increasing the purity further would gain little but must surely reduce the efficiency.

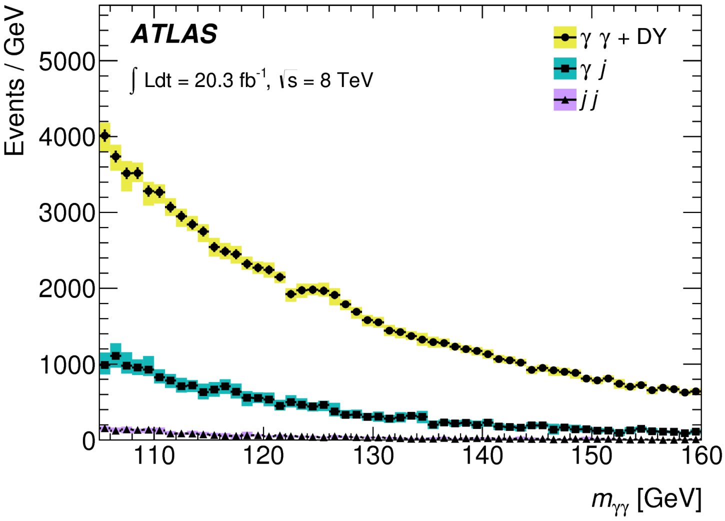

Figure 1.4: The composition of the diphoton sample measured by ATLAS as a function of the mass of the pair.

1.2.2 Mass calculation

Now it is necessary to reconstruct the mass of the pair of photons. As photons are massless, the calculation reduces to

$$m_{\gamma\gamma}c^2 = \sqrt{2E_{\gamma 1}E_{\gamma 2}(1 - cos\theta_{\gamma\gamma})}, \tag{1.1}$$

where $\theta_{\gamma\gamma}$ is the angle between the two photons. Clearly this depends upon accurate measurement of the energies, However the cos $\theta_{\gamma\gamma}$ factor is also important. As mentioned before, this is found by using tracking (and in the case of ATLAS photon pointing - the direction of the photon as measured in the calorimeter) to identify the proton vertex from among the many collisions seen per event. The di-photon mass resolution is however dominated by the energy measurement of the photons, and the resulting resolution is typically 1.5 GeV, i.e. 68% of masses will be in a window 3 GeV wide.

1.2.3 Signal significance

The analysis now looks for an excess of photon pairs with a mass compatible with a hypothetical Higgs mass. The expected rate is about 400 events in the 2012 ATLAS data shown above, so if we tried to do the analysis by counting the number of photons in a window of 1σ about the Higgs mass, we would expect $0.68 \times 400 = 272$ signal events. The background level at a mass of 125 GeV/c^2 in the plot is 2000 events/GeV so estimate 6000 background in this window. The statistical error on the background of 6000 is $\sqrt{}6000$=77. So the significance expected is $s/\sqrt{b} = 272/77 = 3.5\sigma$.

However, this assumes we knew the expected background rate. In fact the theoretical calculation of this rate is difficult and has uncertainties around 20%. Thus the total error on the background prediction is σ_{tot}=$\sqrt{}(77^2+1200^2) = 1202$, completely dominated by the systematic error. The significance now becomes 272/1202=0.2σ, i.e. there is no useful statement we can make on the presence of a signal. So how was this channel one of the main contributories to the Higgs discovery?

There are around 100,000 photon pairs in the plot above, and we can use them to measure the background rate. We can count the number excluding the region about 125 GeV we are testing for a signal. This will have a fractional error of about $1/\sqrt{N} = 1/\sqrt{100,000}$ =0.3%. if we assume we know from theory the shape of the distribution (i.e. the fraction of the total falling in the signal window), and we measured the total rate, then in our window of 3 GeV around 125 GeV we predict 6000±18(stat). This 18 events is much less than the statistical error on 77 on the predicted rate, so we can now find a total error of σ_{tot}=$\sqrt{}(77^2+18^2)$=79. This makes the significance of our signal 3.4σ, not very different from the original 3.5σ estimate.

This approach above is a bit naive. I assumed the shape was perfectly known, and this is not justifiable. The actual analysis uses simple functions such as exponentials (or exponentials of quadratics) to model the shape, and the parameters of the exponential functions are fit to the data. This is taking less information about the shape from theory – only that the form chosen has enough flexibility to describe the data, not what the parameters of the function are. The total data seen by the ATLAS experiment can be seen in figure 1.5.

The fraction of events in the signal region is small, but there is a clear region where the smooth background function is unable to explain the data, while adding a few hundred signal events makes for a function which describes the observed data well.

1.2.4 Look-elsewhere effect

Finally, the analysis ignores the look-elsewhere effect, or trials factor. If we are looking for a signal and we do not know where to look, we have to consider multiple possible signal masses. Take a simple example, a search with events categorised into 5 bins in a histogram, call them 1, 2, 3 4 or 5. If we see a signal that has a particular p-value, say 0.01 in the bin 4, can we conclude that this result is a 1 in 100 result? Obviously not, there were 5 bins in which it could have been observed.

The correct calculation is this: The most extreme p-value observed is X. What is the probability of getting X or more in any bin of my distribution? That is most simply expressed through the probability of not seeing X in all the bins, i.e. $\mathrm{X_{global}}$=1-(1-X)$^{\mathrm{n}}$, where n is the number of bins. If X is small then we can use the first order expansion $\mathrm{X_{global}}\approx$1-(1-nX)=nX, and this of course is the same as saying 'I have a probability of X, but there are n chances of getting X so the total probability is nX.'

In the case of the Higgs search above there is perhaps a 30 GeV interesting region 110-140 GeV/c^2, and if it were broken into 10 bins then the look-elsewhere effect would mean the 3.4σ p-value found earlier, with X=0.0005, corresponds to an $\mathrm{X_{global}}$=0.005, which is equivalent to a Gaussian 2.6σ.

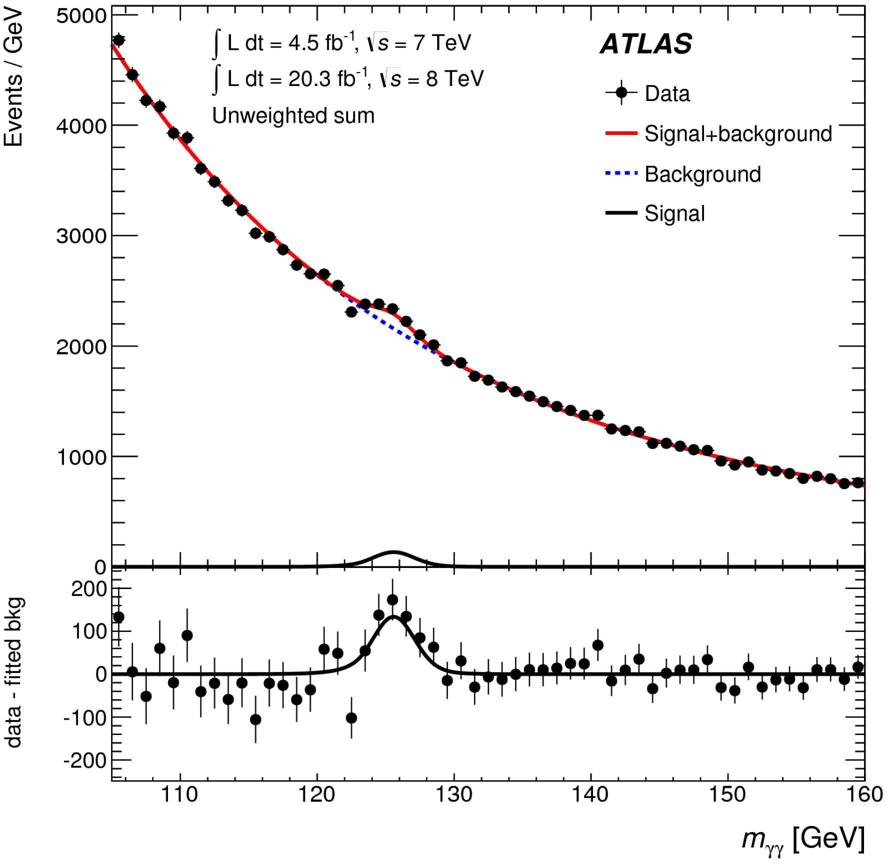

Figure 1.5: The di-photon mass distribution as seen in ATLAS in run 1

However there is a philosophical problem here – how did I decide there was a 30 GeV interesting region? This was roughly taken from the region where the H→ $\gamma\gamma$ search is expected to be useful in the SM. If we had seen a peak outside that region it would still have been interesting. So was the 30 GeV region really the only search region? In truth it is often hard to define the look-elsewhere-effect. For this reason the LHC experiments, when claiming to have discovered the Higgs boson, highlighted the *local* significance.

1.3 Higgs to four leptons

The second of the two discovery channels was Higgs discovery was H→ZZ→llll. This decay mode has a problem: the Higgs boson weighs 125 GeV/c^2, while the Z boson has a mass of 91 GeV/c^2. Thus for the decay to proceed one or both of the Z bosons has to be off mass shell. The normal solution is ZZ^*, with one of them on shell and the other with 30 GeV/c^2 or less.

The decay rate is 3%. This is 10 times the Higgs branching ratio into photons and at first sight is very promising. However, only 6.6% of Zs decay to electron or muons pairs, which can be cleanly identified and accurately measured. Both Zs must decay this way if we are to measure the mass. The end result is that only about 66 such decays are expected from the SM Higgs from Run 1 of LHC - an order of magnitude less than the diphoton. It is then essential that the efficiency of reconstructing leptons is maximised.

The off-shell Z presents a real challenge. The LHC detectors were designed assuming that leptons would

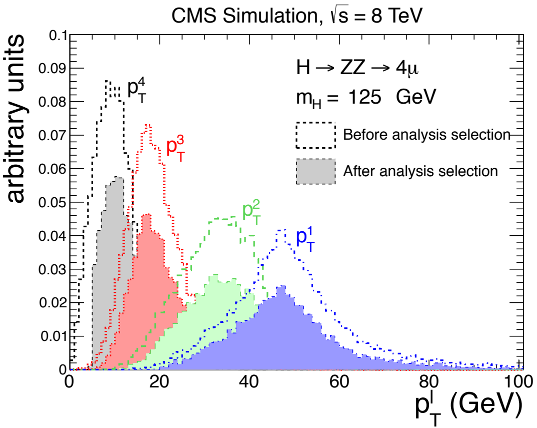

Figure 1.6: The p_T distribution of the four leptons from a $H \to ZZ \to llll$ decay, ordered by p_T.

have momenta around 50 GeV/c. But the lowest momentum lepton, see plot, has typically a $\mathrm{p_T}$ of about 10 GeV/c. For this measurement the detectors accept a minimum $\mathrm{p_T}$ of about 6 GeV/c. The lower the p_T the more slowly the dominant π^+ will travel through the detector, and the higher the chance that they decay to μ^+. In addition the calorimetric energy measurement is getting worse as the momentum drops, which makes the electron measurement more difficult.

The primary background is the production of ZZ^* with no Higgs involved. Backgrounds from other processes exist, e.g. Zbb→llbb and tt→WbWb→lνblνb. The b quark is problematic because it decays to b→lνc 10% of the time for each lepton species, and these leptons can easily have momenta of order 10 GeV/c. However, with careful choices of isolation it is still possible to ensure that the primary background is genuine ZZ^* production, while still keeping the efficiency high. As with the $\gamma\gamma$ analysis the fakes rates are measured in data by studying candidates which just fail to pass the selections.

1.3.1 Mass calculation

The tracks of the four leptons are all measured, and so their four-vectors can be summed and the $\sqrt{s}/c$ extracted. The mass resolution is slightly better than the two-photon case. The final mass distribution, as measured by ATLAS, is in figure 1.7

The red background shape labelled ZZ^* has two major structures. Above 190 GeV two real on-shell Z bosons can be produced, and here is the bulk of the background. However, the llll mass distribution also shows a spike at 91 GeV, corresponding to the very rare case where a single Z decays into four leptons, not the usual two. Between these two the expected rate is quite low, although there is a small rise around 120-140 GeV, partly from the misidentified background.

The peak from the Higgs boson at 125 GeV is rather striking. 37 events are observed in 120-130 GeV, where the expected background rate was about 10. If we approximate the Poisson with a Gaussian we might assign an error of $\sqrt{10}$=3.2. The statistical significance is then 37-10/3.2=8.4σ. However, we should mistrust the use of Gaussian errors for such low numbers.

There is a look-elsewhere effect in this analysis too, but again the range is unclear. Was it 120 to 150, or 120 to 600 GeV/c^2? Both possibilities were put into the discovery paper, but the ‘5 sigma’ significance was held to apply to the local significance, as that is well defined.

Systematic errors appear in this analysis too, some 10% for the ZZ^* background and 20% for the 'fake' backgrounds. However, because the background level is so low the statistical error dominates everything. As a much larger dataset becomes available the background will be measured using the sidebands in a similar way to the $\gamma\gamma$.

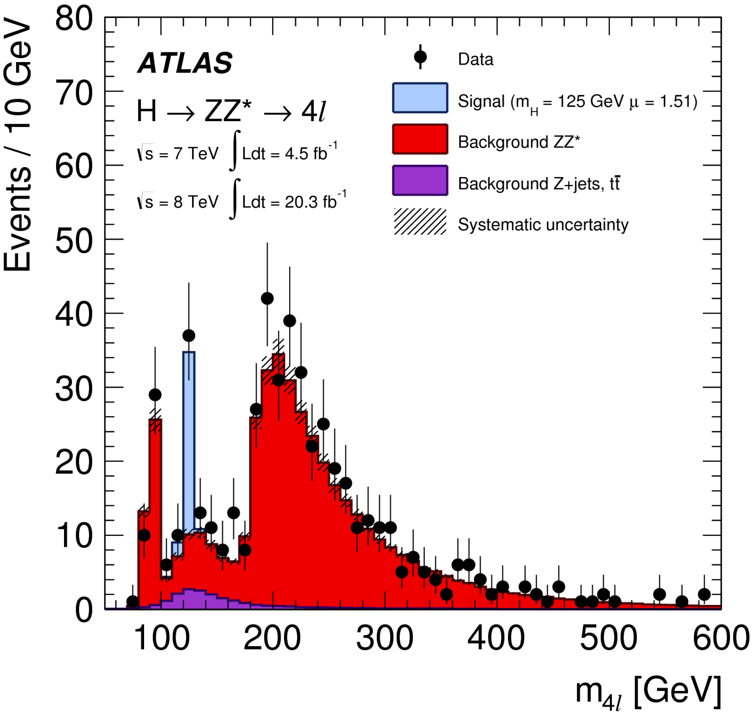

Figure 1.7: The four-lepton mass distribution measured by ATLAS.

1.4 Spin and Coupling studies

The conclusion is that the Higgs boson has appeared in two very different analyses but each time with a rate similar to expectations. Each of the analyses sees more than 5σ evidence for a peak, and the discovery is therefore made twice over. In addition both ATLAS and CMS strongly observe the same two channels. This is certainly a new particle.

The mass has been measured carefully and the spin has been studied - one feature of the Higgs is that is the first fundamental spinless particle, and this new object appears to have this property.

Since 2012 the decays to WW and $\tau\tau$ have been seen, though the last requires both ATLAS and CMS combined data and only preliminary results exist so far. There is some evidence for $H \to b\bar{b}$, mostly from production in association with a Z boson which is used for the trigger. Limits have been set on other decays, for example $\mu\mu$ and ee which should not be sensitive with this dataset, and do not see anything.

LHC measures the product of production and decay, so couplings cannot be extracted without some assumption. A rather model-independent analysis is to assume one particle is produced, and that production and decay factorise. This then allows to measure the ration of production and decay to a reference channel, and this is shown in figure 1.8. , although there is some tension in ttH production and bb decay.

Figure 1.8 (right) shows a much more model-dependent analysis, where the particles couplings are scaled in a leading-order way (the 'κ-framework'). The dependence of the coupling on mass is nicely illustrated. All the coupling studies are consistent with the standard model Higgs. All in all, we have good evidence we are living in a sea of Brout-Englert-Higgs field.

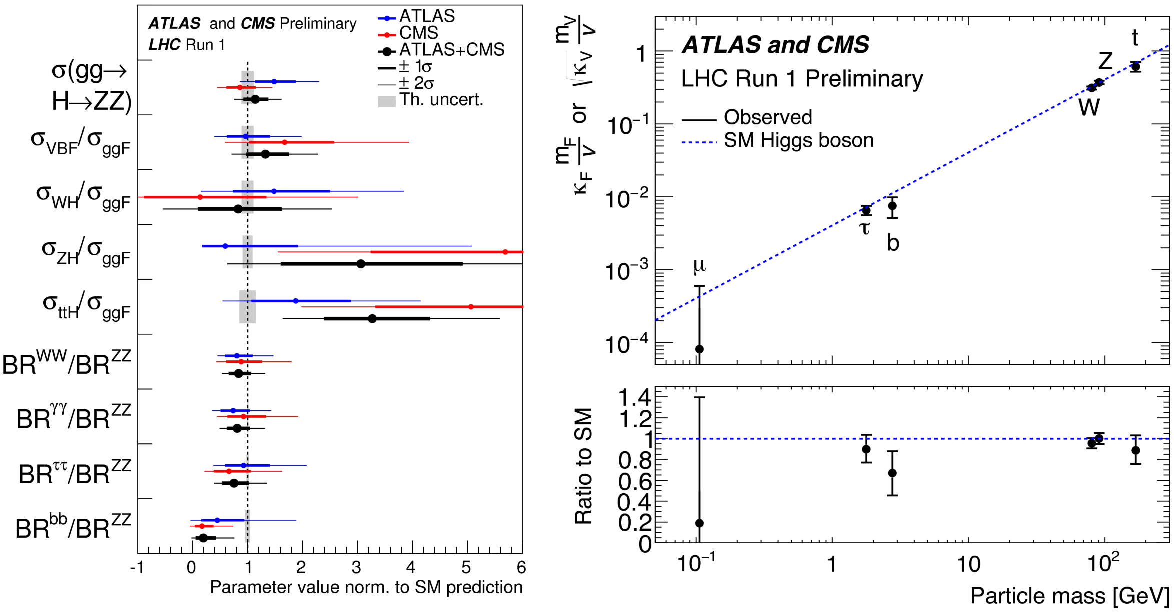

Figure 1.8: Left: The production rates and decay rates of the Higgs boson, as measured by a combination of ATLAS and CMS data. Right: The same data used to extract the coupling strength.

Chapter 2

Beyond the Higgs boson

2.1 LHC today

The LHC beam energy in the 2010-2012 run which discovered the Higgs boson was 3.5 or 4 TeV, (i.e. $\sqrt{s}/c = 7/8$ TeV) and data-taking restarted in 2015 at 6.5 TeV. The data collected or expected in various years is shown in table 2.1. The first 0.5 fb^{-1} of 2016 data has been delivered. There is a lot of room for new discoveries in 2016 and beyond.

Year	$\sqrt{s}$/c	$\int Ldt$, fb^{-1}
2010	7	0.05
2011	7	5
2012	8	20
2015	13	3
2016	13	25 est.
by end 2018	13	100 total
by end 2023	14	300 total

Table 2.1: The collision energy and amount of data recorded by the general purpose LHC detectors in various years.

There was a meeting in December 2015 where the collaborations reported first results from the data take that year. Some 30 different searches and measurements were shown by each collaboration; for example the production rates of W, Z, $t\bar{t}$ and H were all tested. No unexpected results were observed. However, there are always statistical fluctuations, and so in some places the agreement is less perfect than in others. One measurement attracted interest because a deviation was seen in the same place by both ATLAS and CMS - the $\gamma\gamma$ resonance search. The results shown here are from the scalar search as shown at Moriond 2016.

Figures 2.1 and 2.2 shows the diphoton mass spectra seen by ATLAS and CMS. This spectrum was analysed in much the same way as the Higgs boson discovery, as seen in figure 1.5. The CMS data is split into barrel-barrel and barrel-endcap events, and there is a separate analysis for 20% of the data where the magnet was not operating.

The background is fitted to arbitrary functions, tested in simulation for lack of bias. The CMS plot shows the error on these background functions.

The numerical interpretation is done by hypothesizing a new particle of a particular mass and then fitting the data to see how many such particles best describes the data seen, and the p-value that measures the compatibility of the actual result with the background only model. Then this is repeated for all possible masses

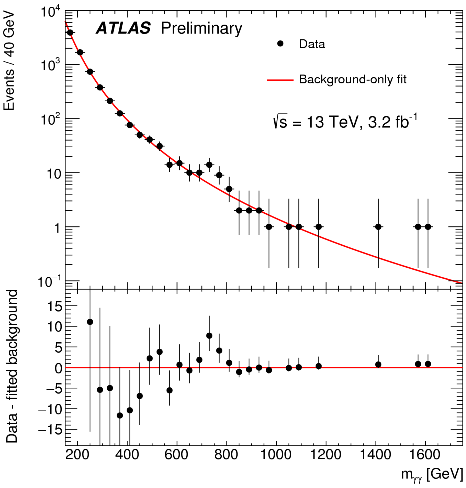

Figure 2.1: The di-photon mass spectrum recorded by ATLAS in 2015.

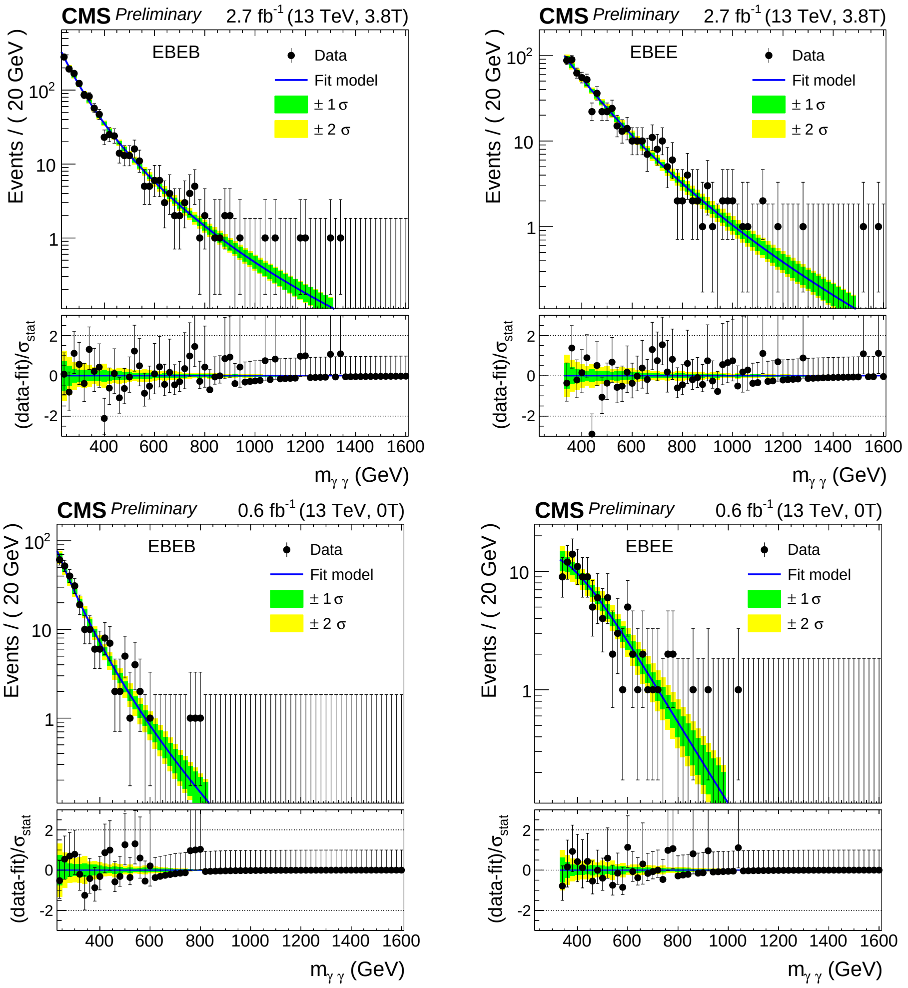

Figure 2.2: The di-photon mass spectra recorded by CMS in 2015. Top row is the field on data, bottom row field-off. The left hand plots have both photons in the barrel, on the right one is in an endcap.

in the range to be tested. The p-values are shown in Figures 2.3 and 2.4.

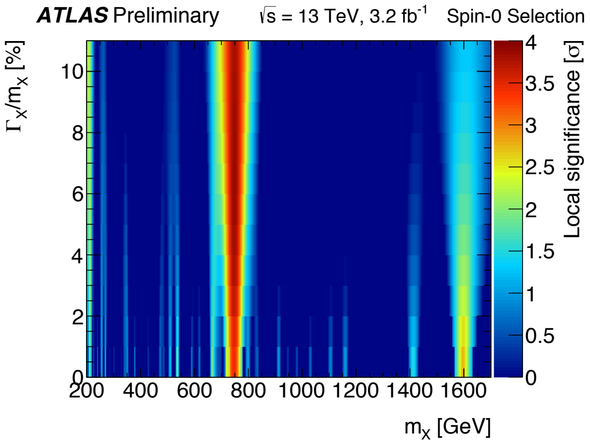

Figure 2.3: The compatibility with the background model estimated by the two experiments when hypothesizing a new particle decaying to $\gamma\gamma$ with a range of possible masses.

ATLAS present a 2D scan of mass and width. The most-extreme p-value in the ATLAS result is at 750 GeV/c^2 for a width of 45 GeV/c^2. It is about 3.9 σ. This is well short of the 5 sigma at which people traditionally claim a discovery - but sufficiently unusual to be interesting.

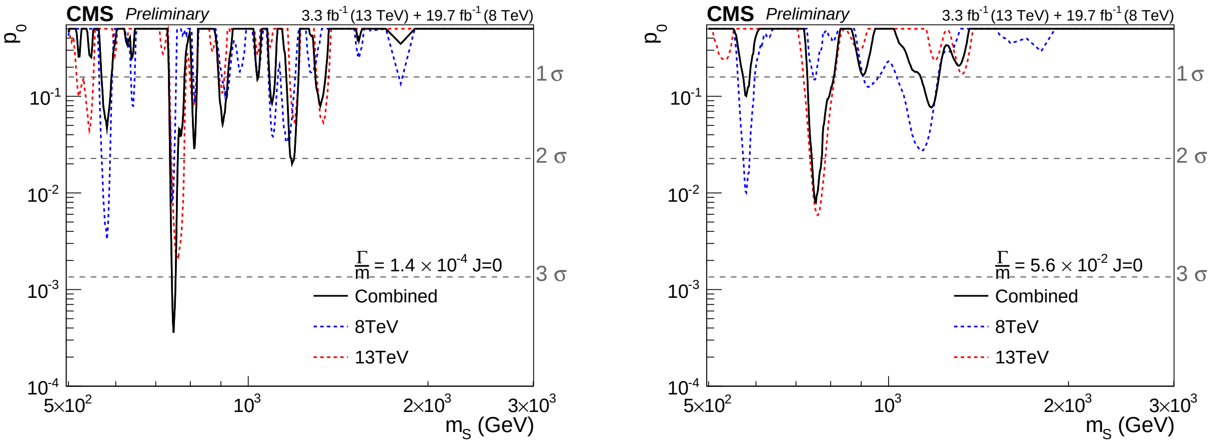

Figure 2.4: The compatibility with the background model estimated by the CMS when hypothesizing a new particle decaying to $\gamma\gamma$ with a range of possible masses and for narrow (left) and wide (right) resonances.

What makes it much more interesting is that the CMS data shows its most extreme p-value at the same mass. CMS report a 3.4 σ excess from combining 2012 and 2015 data, for a narrow width. This falls to 2.4σ if the width is fixed to the 5.6% of the mass - roughly the best fit in ATLAS.

We should not jump to the conclusion that a new particle has been found. There are many dips in the p-value plots, because the good mass resolution means there are many independent places to test - nearly all of them, if not all, will be statistical fluctuations. There was nothing that told us 750 GeV was a special energy. So if we ask the question ‘what is the probability of getting a bump like that anywhere’ we will get a less-unlikely answer. This is called the look-elsewhere effect by particle physicists, or the trials-factor by statisticians. In the ATLAS case the ‘global’ p-value, correcting for the look-elsewhere effect, is 2 σ, or about 5%. The experiments presents 20 or 30 searches, one had a 1 in 20 result..that seems normal variability.

Still the fact remains that both experiments had a peak at the same place. This does make it more unusual, and has sparked a torrent of papers proposing theories to explain the data - hundreds of them. But theories will not tell us whether or not there is something there - that will take more data. If there really is a new particle at about the 3.5 sigma level, then as we record more data both the amount of signal and the amount of background will rise. As the significance is roughly $s/\sqrt{b}$ this rises with the square-root of the luminosity. With 10 fb^{-1} of data both experiments should see a 5 sigma signal - or the significance dropped to 2 σ. if LHC works well we should have this much data by ICHEP conference in August. So we should know the answer by then.

Personally I am sceptical, but each time a more data is seen it seems to get stronger.