Potential for future e^+e^- colliders to measure the decay process: $H \to \gamma\gamma$

Elliott Jones, University of Birmingham, England

September 7, 2017

Abstract

This project aims to determine how precisely future e^+e^- colliders can measure the process $H \to \gamma\gamma$, this process is very important as an input for the global Higgs coupling fit. The project intends to determine the precision with which the Future Circular Collider (FCC) can measure the cross-section x branching ratio of the process. This project begins to reproduce the results from the LEP3 note. The cut efficiencies are reproduced almost exactly and the signal plus background histogram is partially recreated. A study is also carried out on important detector components for the measurement of the decay process, $H \to \gamma\gamma$. It is determined that the resolution of the Electromagnetic Calorimeter (ECAL) is the most important component.

Contents

1 Introduction

The Higgs decay mode, $H \rightarrow \gamma\gamma$, is a historically important decay channel. It played a huge part in the discovery of the 2012 Higgs boson[1][2] due to the signal being much easier to distinguish than most other decay modes in the hadronic collider. This is because it has an electromagnetic signature, which is easy to distinguish from the largely hadronic background. This decay process is important for the global fit of Higgs couplings[3], which in turn is an important test of the Standard Model (SM).
To make these measurements precisely a huge amount of data is required due to the fact that the decay process is so rare - the branching ratio is only 0.2%. It is also very beneficial to have a good understanding of which detector components are important for measuring the decay process. The Future Circular Collider (FCC) plans to collect 5 ab^{-1} of data at $\sqrt{s} = 240$ GeV, which is a huge amount of data. For 5 ab^{-1} of data only 2000 events would be expected, therefore we need this much data to get decent statistics. Measuring the Higgs couplings is the cornerstone of the Higgs physics programme for the FCC and so feasability studies such as this one are of key importance for the upcoming FCC conceptual design report - planned for 2018.

2 Theory

The dominant Higgs production channel at $\sqrt{s} = 240$ GeV is the Higgsstrahlung process as can be seen in figure 1 [3]. This is the production of a Higgs boson along with a Z

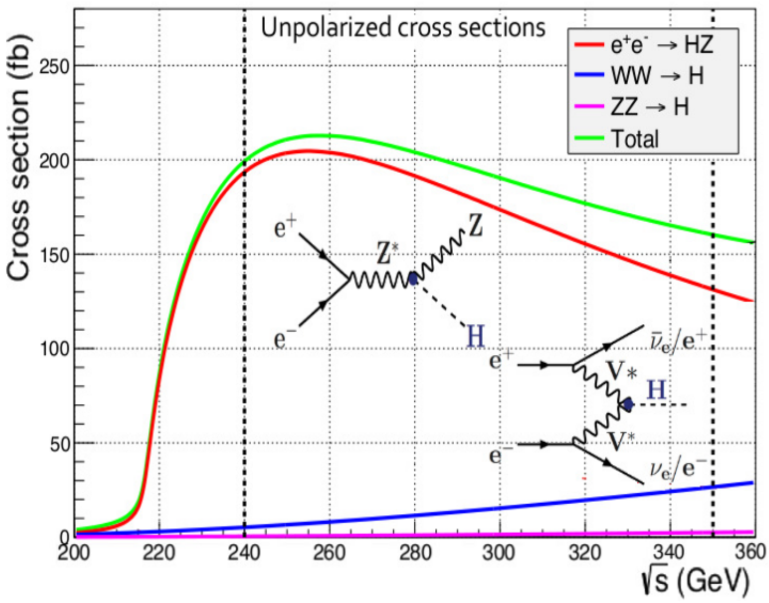


Figure 1: Cross-section of Higgs production process at a range of energies.

boson. A Feynman diagram for the process in an e^+e^- collider can be seen in figure

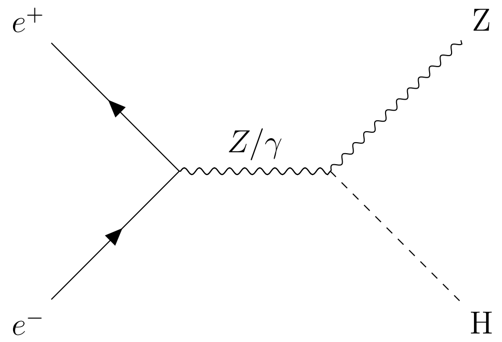


Figure 2: Production of Higgs particles from e^+e^- collisions.

2. This process has a cross-section of 200 fb at $\sqrt{s} = 240$ GeV at leading order and was the production process used in the study. The Z boson was allowed to decay via any process and the Higgs boson restricted to just the desired decay process to two photons. This can occur via either a fermionic or bosonic loop as can be seen in figure 3. The decay into two photons has a very small branching fraction of only 0.2%. The last

(a) Higgs decay to 2 photons via fermionic loop.

(b) Higgs decay to 2 photons via bosonic loop.

Figure 3: Decays of a Higgs boson to 2 photons.

process considered in the project is for the background - the process is $e^+e^- \rightarrow Z\gamma$ and a Feynman diagram for this can be seen in figure 4. For the background an extra photon is required for the Higgs reconstruction, this comes from the radiation of a photon from one of the initial state particles.
The Higgs mass can be reconstructed from the measured properties of the two photons it decays into and is given by:

$$M_H = \sqrt{2E_1E_2(1 - cos\theta)}, \quad (1)$$

where E_1 and E_2 are the energies of the two photons and θ is the angle between them. Based on this equation, one can expect the Higgs mass resolution, and hence the Higgs cross-section measurement, depends crucially on the granularity and the ECAL resolution. By increasing the granularity, better measurements of the angle between the two photons can be made which in turn produces more precise measurements of the Higgs

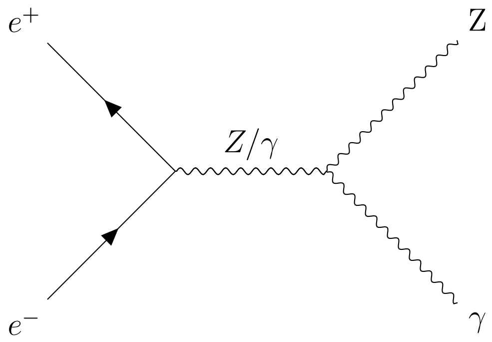


Figure 4: Production of Z and γ from e^+e^- collisions.

mass. By increasing the ECAL resolution the energy measurements for the two photons can be improved - this also results in an improved Higgs mass measurement. Both of these detector components are important for getting a good measurement of the decay process considered. The HCAL resolution may help to improve the measurements of particles that contribute to the reconstruction of the Z boson and therefore also improves the measurement of the Higgs mass.

3 Method

3.1 Simulation

The first stage of the project was to produce some events in PYTHIA8[4]. Only 1000 events were produced for the process, $H \to \gamma\gamma$, so that initial tests for the analysis code could be run quickly. The events were produced by allowing only the Higgsstrahlung production process, $e^+e^- \to HZ$, and then allowing only the Higgs decay to two photons and any decay mode for the Z boson. Feynman diagrams for the processes involved can be seen in figures 2 and 3. Once the analysis code was fully functioning with 1000 events, 10000 events could be used to run analysis on the pure signal. Eventually 40000 events were used for the pure signal as this gave much better statistics without increasing the processing time too significantly.

The other events which needed to be produced were for the background. The background events came from the process $e^+e^- \to Z\gamma$. Initially 10000 events were produced to check the efficiency when passing the events through the analysis cuts. Only 0.2% were found to pass the cuts so 1M events were produced for the main measurements with the hope of giving good statistics for the plots.

After the events were produced in PYTHIA8 they were passed through a parametrised particle detector simulation (PAPAS). The simulation takes an input detector file and then uses this to simulate a real detector by producing realistic measurements for the observable quantities of the particles produced in the events. The detector used for the main measurements in the project is the CMS detector.

3.2 Analysis

The analysis code made use of the High Energy Physics with Python (HEPPy)[5] framework. The analysis code consisted of a series of selection requirements for the events such that only desirable Higgs candidates passed and as many of the background events as possible were removed. The first selection requirement was to allow only events with at least two photons as these are required for the reconstruction of the Higgs particle. The second requirement was to only allow events where at least two of the photons had an energy of at least 40 GeV. This is motivated by the Higgs mass being known to be 125 GeV and so the photons used to reconstruct it need to have energies large enough to produce a particle of mass 125 GeV. The mass of the Higgs reconstructed from two photons is given by equation 1 and can be seen to depend on the energies of the two photons.

After this the Higgs was reconstructed. For events with more than two photons, the Higgs was reconstructed using each pair separately and then all other particles used to reconstruct the Z. The pair of photons used for the Higgs reconstruction was then chosen by minimising the quantity:

$$\chi^2 = |M_H - M_{H,nominal}| + |M_Z - M_{Z,nominal}|, \tag{2}$$

where M_H is the mass of the reconstructed Higgs; $M_{H,nominal}$ is the nominal mass of the Higgs (125 GeV) and similarly for the Z boson with a mass of 91 GeV.

Once the photon pair and Higgs had been chosen they could then be used to make further selections. The first of these was made to restrict the isolation of the photons. The isolation is defined to be the energy sum of the particles reconstructed in a cone around the photon divided by the energy of the photon itself. The cone, in this case was, defined to have a radius of 0.4. The sum of the isolations of the two photons was to be lower than 0.4. This rejected hadronic events with two highly energetic neutral pions. Another selection was made to the photons related to their pseudo-rapidity difference. This was to reject some events coming from the main background source, $e^+e^- \to Z\gamma$. This process preferentially produces photons close to the beam axis, so the pseudo-rapidity difference of the photons was required to be less than 1.8. The final requirement was for the Higgs candidate to have an angle of more than 25 °from the beam axis. This was also to reject events from the main background source. After all of these have been made a tree is produced in ROOT with the mass of the Higgs stored from each event.

After the analysis code was written it could then be run using the 1k test events to get an idea of the efficiencies of the cuts. The efficiencies can be seen in table 1. The individual efficiency is the efficiency of the selection when no other selection requirements are enforced. The marginal efficiency is the efficiency when the given selection requirement is not enforced. The efficiencies achieved by these selection requirements can be compared to the efficiencies achieved in the LEP3 note[3]. After the photon isolation requirement, the LEP3 note achieved an efficiency of $\approx 85\%$ and after all of the selection requirements they had achieved an efficiency of $\approx 59\%$. These values are very similar to those that I achieved with my cuts, which is an important confirmation for the efficiencies. From

Table 1: Selection requirements and their efficiencies for the $H \to \gamma\gamma$ signal only.

Cut	Continuous	Individual	Marginal
> 2 photons	97%	97%	61%
Photon Energy	93%	96%	61%
Photon Isolation	86%	92%	64%
Photon Pseudo-rapidity	64%	71%	84%
Higgs angle	59%	97%	61%

the individual and marginal efficiencies it can be seen that the the most effective was the pseudo-rapidity selection requirement.
With the efficiencies of the cuts known the number of events required for a good signal was chosen to be 40k. The background events could then be tested for their efficiencies as well. It was then decided that 1M events should be enough for the background. The analysis could then be run on the 40k signal events and 1M background events and trees produced with data for the photons and Higgs candidate.

4 Results

The data from the trees was then used to create some histograms for the reconstructed Higgs mass. Firstly a histogram was produced using the CMS detector for the signal alone and a Gaussian fit to it. This histogram can be seen in figure 5. The width of the

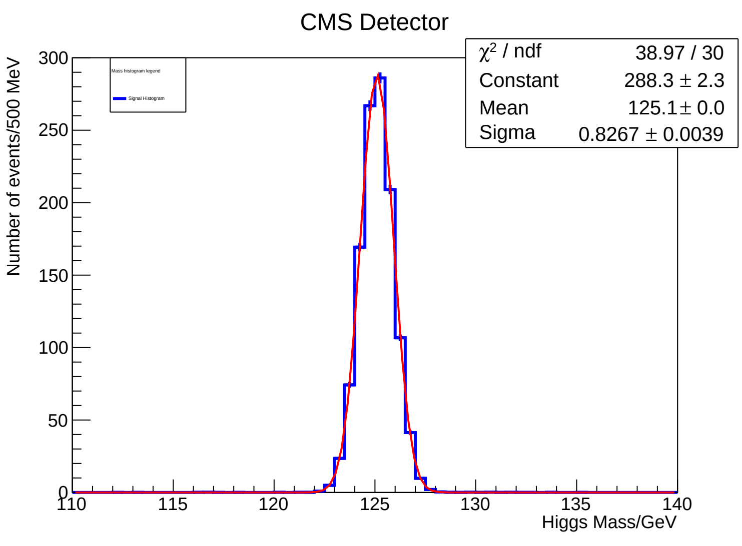


Figure 5: Higgs mass reconstruction with the CMS detector.

Gaussian is a key aspect with regard to the achievable precision on the measurement of the cross-section times branching ratio. Generally, the lower the width, the greater the

Table 2: Detector parameters and their effect on the width of the Gaussian.

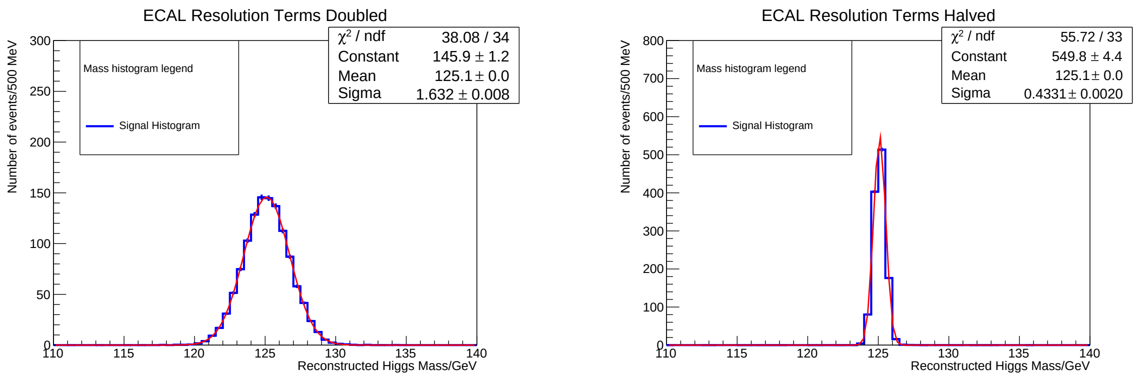
Width/GeV	Doubled	Halved
Cluster Size	0.8327±0.0040	0.8249±0.0039
All ECAL Resolution terms	1.632±0.008	0.4331±0.0020
All HCAL Resolution terms	0.8288±0.0040	0.8277±0.0039

precision of the measurement. The width of the Gaussian from this fit was 0.8267±0.0039 GeV.

After this baseline had been established with the CMS detector, some of the detector components were varied. Specifically, the cluster size, ECAL resolution terms and HCAL resolution terms were all halved and doubled individually to see what effect they had on the Gaussian width. The results can be seen in table 2. The resolution is given by:

$$Resolution = \sqrt{(\frac{a}{E})^2 + (\frac{b}{\sqrt{E}})^2 + c^2}, \quad (3)$$

where a, b and c are the noise, stochastic and constant coefficients. All three of these coefficients are varied simultaneously for the study. It can be seen that varying the granularity has little effect on the width measurement, although it would be expected to have a relatively significant effect. It is likely that halving/doubling the cluster size is not a significant enough change to produce a strong effect. It is probably more reasonable that the cluster size could be varied by a factor of five. The ECAL resolution clearly has a very large effect on the Gaussian width. The effect seems to be almost linear as doubling the values doubled the width and similarly for halving. The histograms from the ECAL resolution changes can be seen in figure 6. The histograms illustrate clearly



(a) Mass histogram from doubling ECAL resolution using the signal only.

(b) Mass histogram from halving the ECAL resolution using the signal only.

Figure 6: ECAL resolution change mass histograms using the signal only.

how large the effect of changing the ECAL resolution is.

The background was used to produce a histogram and try and fit a first, second and third order polynomial to it. The first order polynomial fit had a χ^2/d.o.f. of 1.249; the second order polynomial fit had a χ^2/d.o.f. of 1.244 and the third order polynomial fit had a χ^2/d.o.f. of 1.210. This meant that the third order fit was the best and therefore was used for the main signal plus background histogram. The signal and background histograms were first normalised to their individual cross-sections and branching ratios using 5 ab^{-1} of data. The two histograms were then added together and plotted with a third order polynomial plus Gaussian fit and the background was overlayed to get an idea of what the real measurement would look like. This histogram can be seen in figure 7. The width of the Gaussian from this plot was found to be 1.4606 ± 0.0050 GeV.

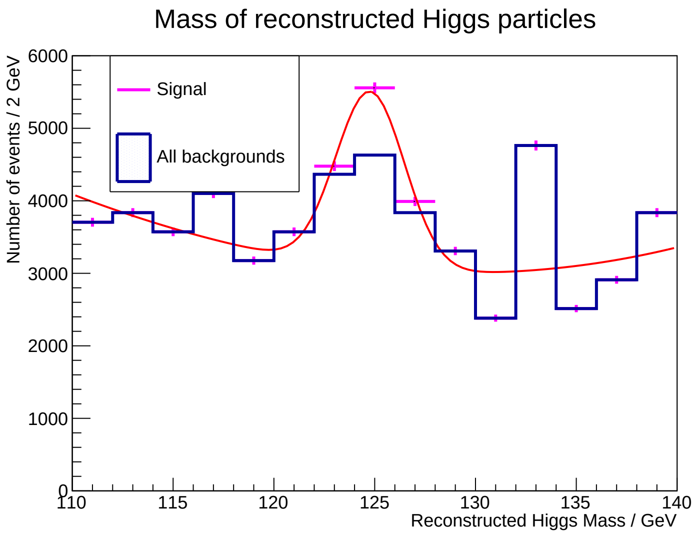


Figure 7: Distribution of the reconstructed Higgs mass in the H→ $\gamma\gamma$ channel using background and signal plus background.

5 Conclusion

Using the expected 5 ab^{-1} of data at $\sqrt{s}$ = 240 GeV from the FCC, the plot in figure 7 could be reproduced and if the CMS detector were used then the width measurable would be at least 1.4606 ± 0.0050 GeV.
The most important detector component for measuring the process $H \to \gamma\gamma$ was shown to be the ECAL resolution, halving the resolution terms halves the width of the Gaussian.

6 Improvements

The first main improvement which could be made to the project is related to the main signal plus background histogram. As can be seen in figure 7, the background has a lot of statistical fluctuations. It would be much better to have an increased number of background events, probably 10M, to try and reduce these statistical fluctuations. This would allow for a much better polynomial fit to the background and therefore a

much better fit overall for the combined histogram and therefore a better estimate of the Gaussian width. Another interesting addition for the main histogram is to add the background for hadronic and leptonic Z decays separately, to see how much each contributes to the background.
Another improvement for the project would be to investigate the effects of the tracking detector and the solenoid magnet. These should have an effect on the reconstruction of the Z mass, which in turn affects the reconstruction of the Higgs mass through the minimisation of the quantity in equation 2.
The final main improvement to the project would be to use a different particle simulation software to produce the background events through the process $e^+e^- \to Z\gamma\gamma$ at higher order in perturbation theory.

7 Acknowledgements

I would like to give a big thank you to my supervisor, Katharina Behr, for being so helpful throughout the project and ensuring that I was able to complete the work that I planned to do.
I would also like to thank Olaf Behnke and the other members of the Summer student organisation team for making this Summer so brilliant and giving us all the opportunity to take part in this programme.

8 Appendix

8.1 How to use the analysis code

The code can be acquired at: https://github.com/jones117a/FCC/tree/master/FCC. First of all the heppy framework also needs to be acquired as it can't be accessed from this repository. Then the folder zh_hgamgam needs to be moved into the folder heppy/analyzers/examples. Within this folder there are a few important analyzers used in the analysis code. The first is the initial selection cuts, this can be altered to make different cuts on the particles in the event before reconstructing the Higgs. The second important file is the Higgs reconstruction, this can likely be left alone as it just reconstructs the Higgs and Z using an optimisation method. The final important file is the second selection file, which makes cuts on the photons used for the Higgs reconstruction or the Higgs itself.
The analysis code itself is in the WorkDir folder and is called
analysis_ee_ZH_Z_Hgamgam_cfg.py. This needs to be accessed to change the files that are being analysed, this can be done in the definition of comp. The other analysis files won't function properly and need to be changed to include the Selection_2 analyser. However, the only other difference is the files that they are set up to analyse so they aren't much use. To run the analysis code type the command "heppy (some path to the Out Directory) analysis_ee_ZH_Z_Hgamgam_cfg.py".

The main histogram plotting script is Mass_Histograms_TEST.cxx and can reproduce the plot in figure 7 with the correct ROOT tree. The path to the initial files in the code need to be changed to use your own ROOT trees from running the analysis. The histogram script can be run using the command ".x Mass_Histograms_TEST.cxx" in ROOT.

The folder BatchDir contains files for creating various events on the batch and also running the analysis on the batch. First a text file needs to be written for a PYTHIA submission script and then in generate.py the process needs to be changed to match the name of the text file. Then the file run_batch_gen.sh needs to be changed to move the output files into your own workspace. Then just type "python generate.py" and the files should be produced. The file run_batch_analysis.sh needs to be altered similarly to move the OutDir to your own workspace and then the analysis can be run using "python run_analysis.py".

References

[1] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett.*, B716:1–29, 2012.

[2] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012.

[3] Azzi et al. Prospective studies for lep3 with the cms detector, 2012. Submitted to the European Strategy Preparatory Group.

[4] Torbjorn Sjöstrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun*, 178:852–867, 2008.

[5] FCC Software Group. Software for future circular colliders (hh, ee, he), 2017. https://github.com/HEP-FCC [Online; accessed 26-07-2017].