# Measuring Properties of the Higgs Boson in the $H \rightarrow ZZ^{(*)} \rightarrow 4l$ channel with the ATLAS Detector at the LHC

Syed Haider Abidi*
Richard Teuscher†, Fabien Tarrade‡
*University of Toronto, *Summer Student*
†University of Toronto, *Supervisor*
‡Carleton University, *Supervisor*

## Introduction

The Higgs Mechanism is a crucial addition to the Standard Model (SM) as it is solves many internal inconsistencies present in the theory such as preserving gauge symmetry while generating mass for the Z and W bosons. However, to see if Nature really works as proposed, predictions from the Higgs theory need to be checked against experimental results. One such prediction -- that high-energy experiments have been trying to find evidence for -- is the presence of a fundamental scalar boson, namely the Higgs Boson.

As such, one of the main goals for the LHC is to solve the questions surrounding the Higgs boson. Many groups at ATLAS, a general-purpose detector, along with CMS, are looking for evidence for this elusive particle. In the summer of 2012, both experiments saw conclusive proof for the existence of a new boson [1]. Now using the full datasets collected in 2011 and 2012, experiments are trying to accurately measure the properties of the new boson to answer if it is the SM Higgs Boson or one of the other Higgs Bosons predicted by Beyond the SM theories.

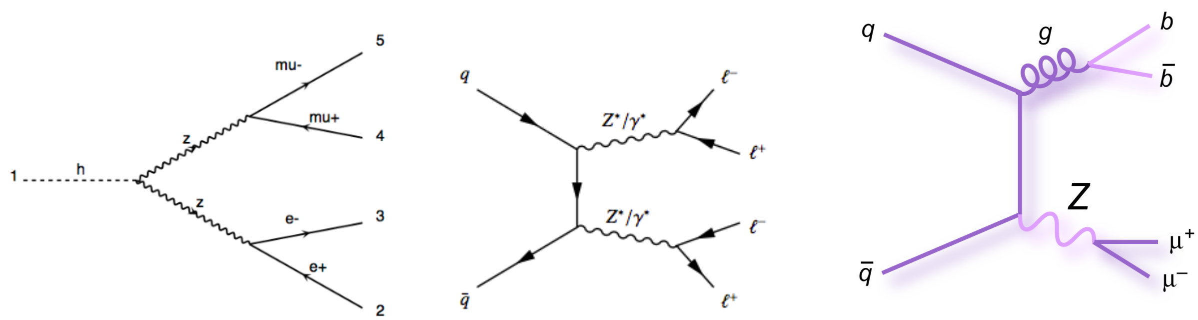
One of the most important analysis channels for measuring the properties of the Higgs Boson is $H \rightarrow ZZ^{(*)} \rightarrow 4l$[1] . While the branching ratio for this final state is approximately $1.25 \times 10^{-4}$, this channel is still referred as 'The Golden Channel' for the search and study of the Higgs Boson. This is due to the fact that the final state is very clean and leptons have a clear detector signature, making them easily identifiable. This is coupled with the fact that there are no complex physics objects such as Missing Transverse Energy (MET) or jets, which are present in other decay channels such as $H \rightarrow WW^{(*)}$ and $H \rightarrow \tau\tau$. This simple final state allows the two Z bosons and subsequently the Higgs Boson to be fully reconstructed. Furthermore, unlike $H \rightarrow \gamma\gamma$ and $H \rightarrow b\bar{b}$, there are few expected backgrounds, allowing the contributions of the Higgs to be easily seen. Hence, many different properties of the boson can be easily tested, from which conclusions about the nature of the boson can be drawn.

## Higgs to 4 leptons analysis

As the Higgs Boson has a predicted mean lifetime of $10^{-22}$s [1], only the decay products of the boson can be seen in the ATLAS detector. Hence, for this analysis, only events where 4 or more leptons are produced in the collisions are selected. However, there are many other standard model processes that have the same final state, causing them to contribute to the background seen in this analysis.

The backgrounds are classified into two general categories, namely irreducible and reducible. The irreducible background consists of events in which two Z bosons or two virtual photons are produced. They then subsequently decay leptonically, mimicking the same final decay state. As seen from the first-order Feynman

[1] '*l*' refers to lepton, which for the purpose of this report only includes electron and/or muon




(a) Higgs signal (b) Irreducible Background (c) Reducible Background: Zbb

Figure 1: Leading order Feynman diagrams of process that contribute to the $H \to ZZ^{(*)} \to 4l$ analysis

diagram in Figure 1, the only difference between the signal, Figure 1a and the irreducible background, Figure 1b is the intermediate Higgs Boson. This leads to the irreducible background having very similar event topologies and kinematic distributions as the signal.

On the other hand, the irreducible background consists mainly of events where the Z boson is produced in association with quarks, which mainly consist of $c\bar{c}$ and $b\bar{b}$, or $t\bar{t}$ pair decaying semileptonically. In both cases, two leptons are created in the decay in association with at least two jets, which if mis-reconstructed as leptons, leads to the same final state. Figure 1c shows the leading order Feynman diagram of one of the contributing processes. Here, the Z boson decays into either electrons or muons, and the jets from the b quarks are misidentified as leptons, causing this event to fake the same final state as the signal. As the jets behave differently when compared to leptons, the event topology and kinematic distributions of the irreducible background are different from that of the signal.

As described above, due to background processes that have the same final state as the signal, simply requiring events with 4 leptons will make it very difficult to distinguish the contribution of the Higgs. Hence, further selection is needed to reduce the background and improve the signal significance.

### *Event Selection*

The data used for the analysis are chosen by using the single-lepton or di-lepton triggers. There are five overall triggers: the single and di-muon triggers select events that contain a muon above a given $p_T$. Similarly, the single and di-electron triggers select events where the $E_T$ of the electron exceeds a certain threshold. Lastly, there is an electron-muon trigger, which has a lower threshold than the dedicated triggers; however, it requires leptons of each flavour to exceed the respective cut value [2]. Furthermore, for each event, it is also required that the ATLAS detector satisfy data quality cuts; thereby any event collected in a period where a part of the detector is not working as designed is vetoed.

Once the event has been selected, initial detector level cuts are imposed on the leptons to ensure that the reconstructed objects are found in a region of the detector where the lepton can be detected and to reduce the number of fake leptons. Each electron (muon) must satisfy $E_T > 7$ GeV ($p_T > 6$ GeV), $|\eta| < 2.47$ ($|\eta| < 2.7$)[2] and related identification cuts [2]. Also, to ensure that leptons come from the main interaction vertex in the events, the impact parameter for each lepton is required to be within 10 mm of the primary

[2] The ATLAS experiment uses a right-handed coordinate system with its orgin at the nominal interaction point. The z-axis is along the beam pipe, the x-axis points to the centre of the LHC ring and the y-axis is defined as pointing upwards. Polar coordinates (r, $\phi$) are used in the transverse plane $\phi$ being the azimuthal angle around the beam pipe. The pseudo-rapidity $\eta$ is defined as $\eta = ln[tan(\theta/2)]$ where $\theta$ is the polar angle.

vertex. Lastly, to reduce cosmic rays muons, each muon must have a transverse impact parameter of less than 1mm with respect to the beam axis.

As there are many reconstruction algorithms, each optimized for certain regions of the detectors, overlaps between the muons (electrons) from different reconstruction algorithms are removed by choosing only one lepton if many particles share the same track or are close together in $\eta - \phi$ space. For example, for electrons, if multiple other electrons are close together, the one with the highest cluster $E_T$ is kept. Moreover, one of the reconstruction algorithms uses Electromagnetic Calorimeter data to reconstruct muons; thus, electrons can fake the signature of muons. Therefore, calorimeter-reconstructed muons that are close to electrons in $\eta - \phi$ space are also vetoed. Lastly, as stated in the previous section, jets may be reconstructed as electrons; therefore, electrons close to jets in $\eta - \phi$ space are vetoed [2].

Once the leptons are selected, two pairs of the same flavour opposite sign leptons are reconstructed, forming a quadruplet. As there are multiple ways in which leptons can be paired, the quadruplet where one of the pairs has a mass closest to the PDG[3] mass for the Z boson is kept. This pair is referred to as the leading pair while the other is labelled as the sub-leading pair. If more than 4 leptons are present, multiple quadruplets are constructed; the one with the leading pair with the mass closest to the PDG Z boson mass is kept. Due to the nature of the Z boson decay, further kinematic and isolation cuts on the leptons are imposed. It is required that the 3 highest $p_T$ leptons must satisfy: $p_{T1} > 20$ GeV, $p_{T2} > 15$ GeV and $p_{T3} > 10$ GeV and for each same (opposite) flavour lepton pair, it is required that $\Delta R > 0.10(0.20)$[4]. To veto leptons coming from $J/\psi$ decays, events are rejected if any same flavour lepton pair has an invariant mass of less than 5 GeV. Lastly, it is also required that at least one lepton in the quadruplet has triggered the relevant data trigger [2].

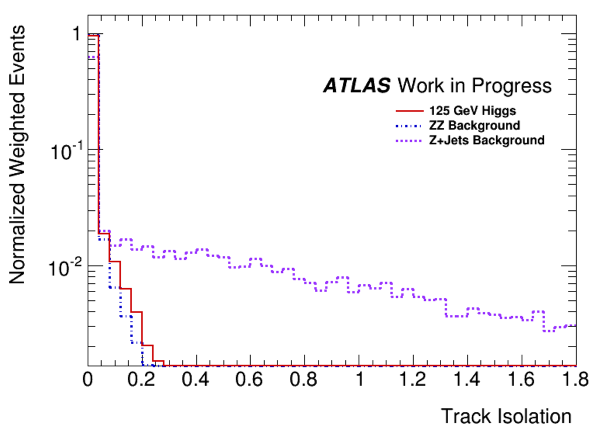



Figure 2: Distribution of the Track Isolation Discriminant

As the mass of the newly found particle is approximately 125 GeV, kinematic constraints ensure that in the Higgs decay to 2 Z's, one Z boson is mostly on-shell while the other is an off-shell particle. As such, the mass of the leading pair ($m_{12}$) is required to lie between 50 GeV and 106 GeV and the mass of the sub-leading pair ($m_{34}$) is required to be within the range $m_{min} < m_{12} < 115$ GeV, where $m_{min}$ is 12 GeV if the mass of the quadruplet ($m_{4l}$) is less than 140 GeV and it rises linearly to 50 GeV at $m_{4l} = 190$ GeV [2].

To limit the contribution of the reducible background, further cuts based on impact parameter significance, track and calorimeter isolation requirements are imposed [2]. For example, the track isolation cut is based on the sum of $p_T$ of nearby leptons divided by the $E_T$ of the particle being tested, and it is required to be less than 0.15. Figure 2 shows the distribution of this discriminant for the signal and the background processes. It can be clearly seen that the reducible background has a very different distribution shape; therefore, this cut can greatly reduce this type of background.

Over this summer, a C++/ROOT analysis program was written which implemented all the above event selection cuts, along with the smearing and reweighting for Monte Carlo (MC) samples. This allowed the production of skimmed mini-ntuples, which are being used by the group working on the Higgs to 4 lepton decay for further studies. In addition, the cutflow analysis program was used to test and validate new recommendations

[3] Particle Data Guide
[4] $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

Table I: Number of normalized weighted events passing each cut for $4\mu$ channel

| **Selection Cut** | Signal ($m_H = 125$ GeV) | Irreducible Background | Reducible Background | Signal Significance |
|---|---|---|---|---|
| Total | 94.66 | 2281 | 379420 | 0.153 |
| Trigger | 52.29 | 2148 | 213289 | 0.113 |
| 4 Lepton | 7.94 | 260 | 1667 | 0.181 |
| SFOS | 7.85 | 258 | 1116 | 0.212 |
| Kinematics | 7.45 | 225 | 898 | 0.222 |
| Trigger Match | 7.45 | 225 | 897 | 0.222 |
| Z1 Mass | 7.24 | 206 | 690 | 0.241 |
| Z2 Mass | 6.60 | 142 | 376 | 0.289 |
| $\Delta R$ and $J/\psi$ veto | 6.47 | 140 | 282 | 0.314 |
| Track Isolation | 6.05 | 134 | 7.49 | 0.504 |
| Calo Isolation | 5.91 | 131 | 3.90 | 0.504 |
| Impact Parameter | 5.81 | 129 | 0.86 | 0.505 |

from different performance groups, such as e/gamma and Muon combined.

***Results***

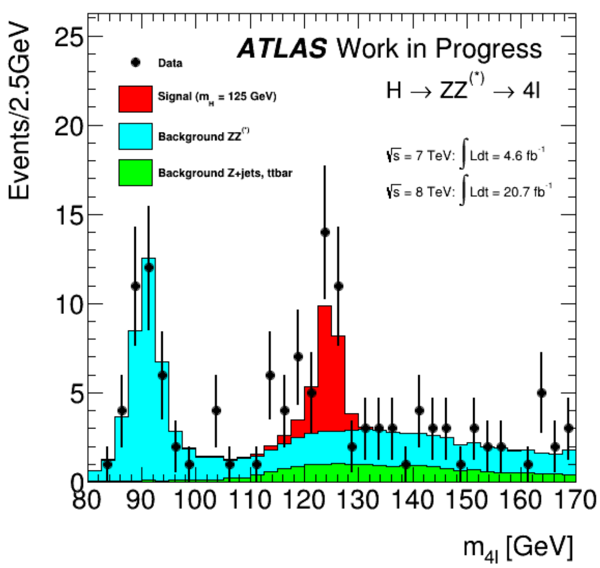
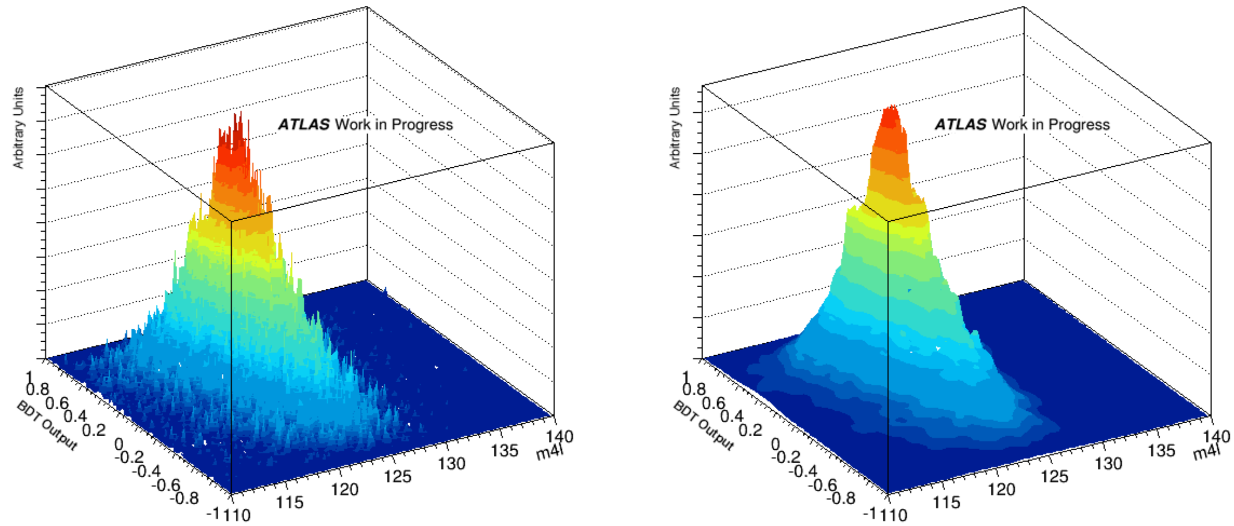



Figure 3: Invariant mass of 4 selected leptons

To understand the impact of each cut on the event selection, the number of events passing each criterion were counted for different processes using the MC generated samples. Table I lists the normalized events for the $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ decay sub-channel after the detector level cuts have been imposed on the leptons. It can be seen that the reducible background dominates the initial expected number of events. However, simply requiring at least 4 leptons in the event greatly reduces this background. The next few selection criteria reduce the overall background while keeping the expected signal events -- slowly increasing the signal significance. As mentioned in the previous section, the track isolation cut is specifically designed to decrease the reducible background and this effect is seen in the cutflow table. The calorimeter isolation and impact parameter significance cut have similar discriminating power, but after the track isolation, the reducible background is statistically limited.

Figure 3 shows the plot of the invariant mass of the leptons in the selected events for both data and MC. Comparing the data to the MC background, a clear excess is seen at around 125 GeV, which is consistent with the Higgs hypothesis. Computing the significance of the excess seen, a $6.6\sigma$ effect is seen at 124.3 GeV. This alone exceeds the $5\sigma$ requirement; using just the $H \rightarrow ZZ^{(*)} \rightarrow 4l$ channel, discovery of a new particle can be claimed.




(a) Unsmoothed Histrogram (b) Smoothed Histrogram

Figure 4: Comparsion of the smoothed and unsmoothed histogram for the 4e channel. The x-axis is the invariant mass of the 4 leptons and the y-axis is the BDT output

### *Mass Measurement*

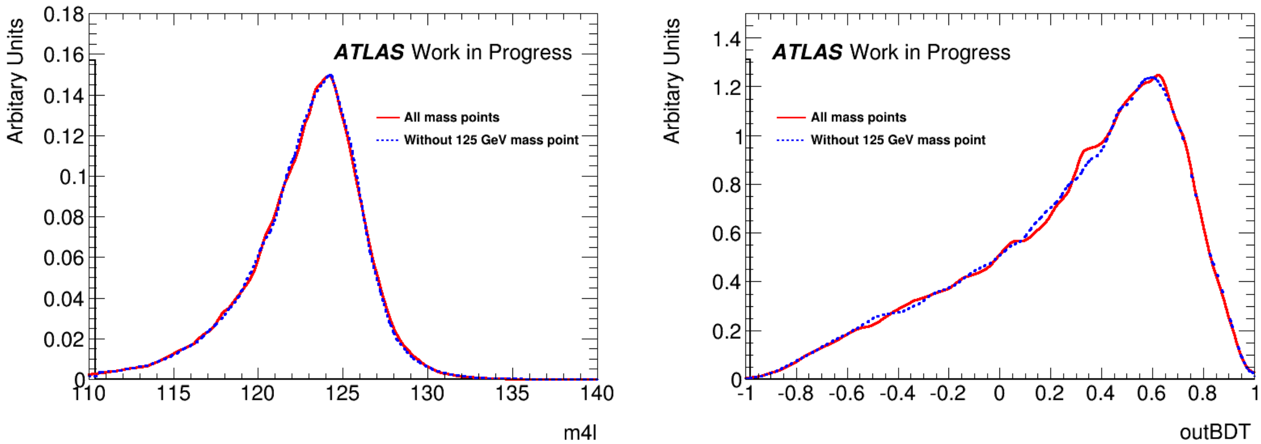
One of the most important parameters in the SM Higgs mechanism is the mass of the Higgs boson as it is the only free parameter in the theory. Once the mass is specified, all other properties such as the production cross-section and branching ratio are fixed in the SM. Since the Higgs is fully reconstructable in the 4 lepton decay channel, the mass can be measured very precisely, allowing for concrete predictions to be tested by other analyses.

The mass of the boson is measured by performing a one-dimensional unbinned likelihood fit to the invariant mass distribution. Probability density functions (PDFs), describing signal and background distributions, are reconstructed from the high statistics Monte Carlo samples. Furthermore, as the signal MC samples exist only for discrete Higgs Mass, PDFs for intermediate mass points are created by interpolating/morphing between the simulated masses. Uncertainties arising from different measurements are taken into account by distorting signal shape and changing overall normalization and these are parameterized as a function of the Higgs mass. Using this, the mass of the boson is measured to be $123.4^{+0.6}_{-0.5}$ (stat) $^{+0.5}_{-0.3}$ (sys) GeV [2].

To improve this mass measurement, studies are being carried out to see the effects of introducing a Boosted Decision Tree (BDT) output based on the kinematics and angular distribution of the reconstructed quadruplet to the fit. This discriminant is added as a second dimension to the likelihood fit and is expected to increase sensitivity to the signal and improve the mass errors.

Similar to the 1D fit, 2D PDFs are needed to describe the distribution of the signal and background. They are obtained by smoothing 2D histograms and then using them to create continuous distributions. As these are new procedures, each step in the process is validated by comparing the output to the original MC sample and ensuring that the overall smoothed shape matches the MC shape and only statistical fluctuations are reduced. Figure 4 compares a smoothed histogram with the unsmoothed original histogram for the 4e channel. It can be seen that the overall distribution shape is the same and fluctuations, especially in low statistics bins, are reduced.

For the Higgs signal, PDFs for intermediate mass points are again obtained by morphing between the simulated




(a) 2D PDF project onto mass axis

(b) 2D PDF project onto BDT axis

Figure 5: Comparsion of the morphed and smoothed PDF for 125GeV Higgs Sample

mass points. To validate these 2D interpolations, a specific Higgs mass MC sample is removed from the morphing input and the distribution for the morphed PDF at that mass is compared to the one obtained from the MC sample. For example, if the morphing is to be validated for the Higgs mass around 125 GeV, the 125 GeV MC sample is removed from input and using the other samples, the morphed PDFs are created. Afterwards, the 125 GeV morphed-PDF is compared to the 125 GeV PDF obtained from MC. Figure 5 shows the result of the above procedure; the 2D PDFs are projected on to the mass and BDT axes. For the mass axis, interpolation is extremely close to the MC shape, but for the BDT axis there are features in the MC sample that are smoothed by the morphing. However, it is expected that these will not greatly affect the measurement.

As the procedures for creating the 2D PDF have almost been validated, further studies are being performed to understand the results and performance of the fit. This is being accomplished by performing test fits to normalized MC samples and toy experiments. After each fit, the truth information is compared to the fitted results and currently, no significant deviations have been seen.

## Summary

The Higgs mechanism is an important addition to the SM; however, experimental proof is required before the theory can be accepted as the SM Higgs Boson. $H \rightarrow ZZ^{(*)} \rightarrow 4l$ is one of the important decay modes where the Higgs Boson can be seen and it is being studied in detail by the ATLAS Analysis group. As there are many background processes with the same final state, many selection cuts are imposed on the events to reduce the background while ensuring that signal is present. An analysis program was written over the summer that implemented these cuts and was used to produce the mini-ntuples for the analysis.

The mass of the Higgs is the only free parameter in the SM Higgs mechanism and previously it was measured by performing a 1D likelihood fit. Currently, studies are being undertaken to improve this fit by adding a BDT discriminant and performing a 2D fit. Over the summer, validations for the steps, such as smoothing and morphing, to create 2D PDFs were also carried out. No significant deviations in the smooth PDF shapes have been found up until now and further studies are being performed by using MC samples and toy experiments.

## References


[1] https://twiki.cern.ch/twiki/bin/view/AtlasPublic/HiggsPublicResults

[2] http://cds.cern.ch/record/1523699/files/ATLAS-CONF-2013-013.pdf