Big Data: ATLAS - Our research on finding the Higgs boson

The Tiffin Girls' School
Archisha
Leona
Erin
Syuan Tang

Summary

As part of the ATLAS Open Data Project, CERN has made a huge amount of data from the ATLAS detector at the Large Hadron Collider available to the public. This allows anyone with access to a device and some programming skills to carry out analysis of real data from particle collisions. We decided to take this opportunity to conduct some of our own research. One of the greatest achievements by scientists at CERN was the detection of the Higgs boson, and we felt that it would be very interesting to see if we could detect this elusive particle ourselves, even with our more limited resources.

Inevitably, we cannot be as rigorous in our investigation as actual particle physicists, since the amount of data and processing power available to us puts restrictions on what is feasible. Nevertheless, we still aimed to thoroughly analyse the data we had access to, in hopes of drawing insightful conclusions.

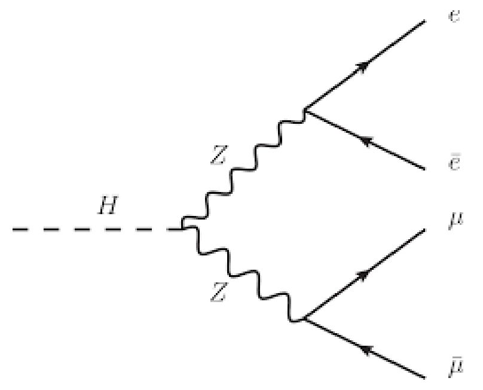
CERN, one of the largest and most respected centres for research, located in Geneva

Research aims

On the 4th July 2012, the date of the Higgs Boson's initial discovery, particle physicists at CERN discovered a particle with a mass of around 125 GeV with properties matching those of the Higgs Boson predicted by the standard model[1]. Evidence of such a particle was discovered through multiple modes of decay, including a decay into two photons and a decay into two W bosons and then a further decay into two leptons.

However, another decay route which was able to tangibly confirm the Higgs's existence was the decay into two Z bosons, resulting in a four-lepton final state[2]. This decay path had a lot of unique advantages over the other two, such as a large signal-to-background-ratio and good lepton momentum resolution, allowing crucial contributions to the accuracy of properties such as the mass, spin parity, width and fiducial cross section of the boson.

By choosing to analyse this decay route in particular, in our efforts to discover the Higgs boson, we are hoping to potentially arrive at a more intense signal around 125 GeV. Given our limited amount of data in comparison to scientists at CERN, when we were using the two-photon and W boson decay routes, the 'bump' was sometimes difficult to distinguish for the typical fluctuations from the background noise of the irrelevant decays.



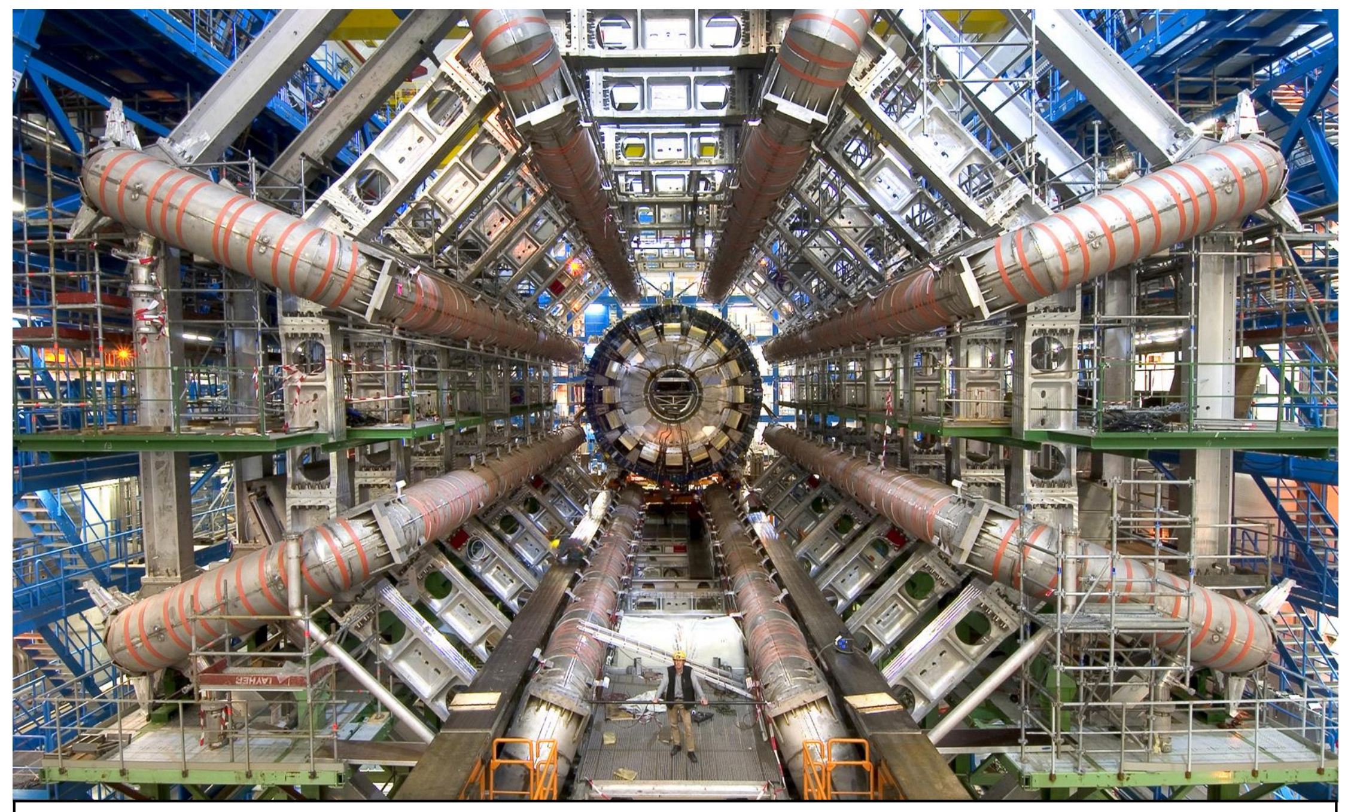
A Feynman diagram of the Higgs boson decaying into 4 leptons

Background information

Particle physics is the study of fundamental particles and forces which make up the Standard Model as fermions and bosons. While fermions comprise what we would traditionally class as 'matter', bosons tend to act as force carriers. When two particles exert forces on each other, they do so by exchanging bosons between each other, as this leads to an exchange of momentum that can be perceived as a mysterious 'push' or 'pull' between the interacting particles . For each of these force carrying particles, there is a corresponding field associated with the force that they mediate. For example, the electromagnetic force of the electromagnetic field is mediated by exchange of virtual photons between charged particles. Furthermore, the weak nuclear force is mediated by the exchange of W or Z bosons.

According to theoretical predictions, all bosons would have to be massless in order to not violate the laws of quantum field theory. The electromagnetic force has an infinite range, suggesting that photons are indeed massless. However, the weak force has a very limited range, suggesting that the mass of its associated boson restricts how far it can travel. This incongruence between theory and reality meant that there was clearly something missing from the accepted model of particle physics at the time. Enter: the Higgs boson.

Rather than mass simply being an intrinsic property of certain fundamental particles, Peter Higgs proposed in 1964 that the property of mass was due to another field: the Higgs field. To convince the wider scientific community, Higgs proposed the existence of the Higgs field's associated boson: the Higgs boson and waited for it to be discovered. However, due to the immense machinery required, it was not until 2012 that the world would see Higgs's theory confirmed through the ATLAS and CMS experiments at CERN. The ATLAS experiment is the largest detector at the LHC, consisting of six different subsystems, such as the tracker, calorimeter and muon spectrometer, to identify and measure properties of the individual particles produced at each collision by recording their trajectory, momentum, and energy.



The ATLAS experiment taken place in The Large Hadron Collider at CERN

Other bosons like photons, W/Z bosons and gluons have a spin of 1 and so are classed as 'vector bosons'. This makes sense as their effects all manifest as forces, which are vector quantities. The Higgs boson, on the other hand, has a spin of 0, making it a 'scalar boson'[3]. This means that its effect has no associated direction. Again, this makes sense because mass itself is purely a scalar quantity.

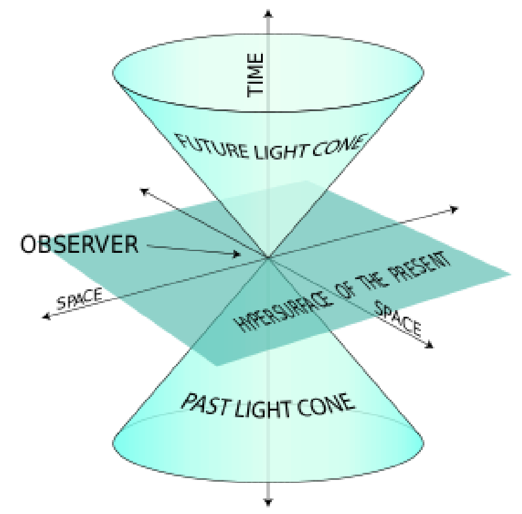
Experimental Method

The Higgs boson itself has far too short a lifespan to be detected directly, as it decays before coming anywhere near the detectors. Therefore, we rely on the fundamental principles of momentum and energy conservation to reconstruct the original particle from its decay products. In classical mechanics, momentum is defined as a vector quantity, with a number of spatial components depending on the number of dimensions. In our three-dimensional reality, this results in a 3-momentum vector formed from 3 components of momentum along 3 orthogonal axes

Experimental Method (continued)

On a macroscopic scale with velocities far less than that of light, this 3-momentum vector is conserved in all interactions. However, when velocities begin to approach light-speed (as in particle accelerators), the principles of classical mechanics begin to fall apart, and the 3-momentum vector changes depending on the velocity of the reference frame from which it is observed. To amend this discrepancy, a fourth component is introduced into the momentum vector: a temporal one.

While the individual components of the four-momentum vector will change under different reference frames, the result of a special operation (often referred to as the dot-product of the four-momentum vector) remains invariant across all reference frames. In other words, the four-momentum vector is covariant under a Lorentz transformation. Therefore, even when factoring in relativistic effects, the sum of all four-momentum vectors before and after a decay will remain constant. It is this key principle that allows the original particles to be reconstructed from the less massive particles they break down into.



A diagrammed representation of space-time, used in four-momentum

To find the mass of The Z_0 boson, we were able to simply add the masses of the decay products. However, because the combined mass of two Z_0 bosons is greater than that of the Higgs, we must instead carry out a non-resonant search. This involves using the transverse mass of the decay products rather than regular mass. Also, we must subtract the background readings to arrive at a sharper Higgs signal. We used a Monte Carlo simulation of the background, which is an integration tool to grant a survey of possible outcomes and is used when there are multiple factors involved so that it is too complicated to find the analytical probability distribution of the outcome.

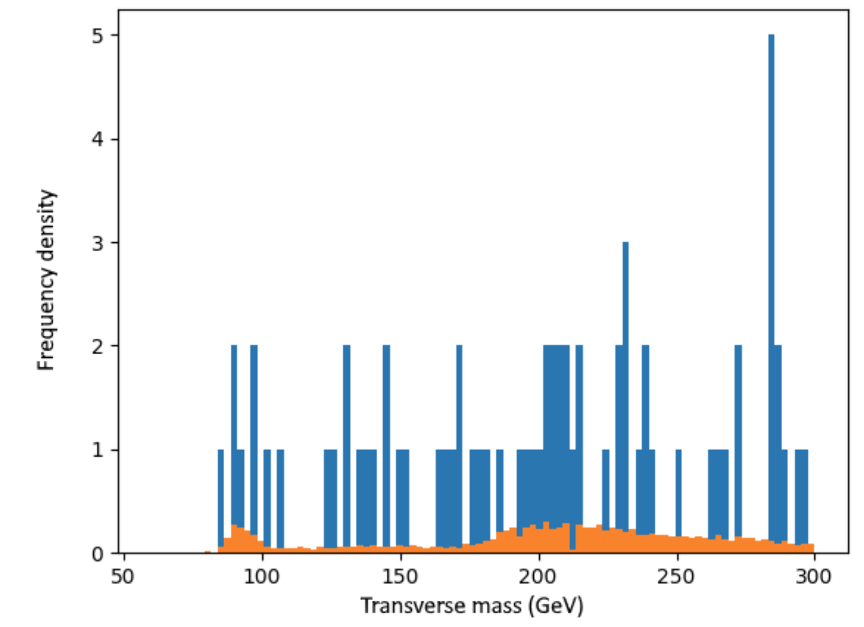
In the decay path we have chosen to analyse, the Higgs boson decays into two Z bosons, which then further decay to result in a four-lepton final state. To ensure we were only analysing events which resulted in 4-lepton final states, we made sure to only draw from the '4lep' folder of the ATLAS Open Data Repository. This would help us to save processing time, as we didn't have to check through thousands of irrelevant events to find those with the correct number of leptons. Z_0 bosons have lepton number 0 since the Z_0 boson is its own antiparticle. One of the fundamental laws of particle physics is that at each 'junction' of a Feynman diagram, a number of particular quantities must be conserved. One of these is the lepton number. Therefore, the lepton numbers of the four resulting leptons must sum to 0 so as to not violate any conservation laws. This can be achieved as long as for each lepton, a corresponding antilepton is produced. These can be neutrinos, or the antiparticle of the lepton itself. In this case, we expect an electron-positron and muon-antimuon pair. The ATLAS detector registers the type of particle in the lep_type event by its corresponding number (e.g. an electron has the number 11). By verifying that the four leptons produced are 2 electron-flavoured leptons and 2 muon-flavoured leptons, we narrow down the events to ones that are more likely to be decays of Higgs bosons.

Results

The transverse mass of the Higgs is around 173.2Gev. In the final result, there is a noticeable peak around this value. The source of this peak very well could come from the Higgs boson. However, the peak itself is not very well-defined, and it is possible that without the prior expectation of a Higgs signal that it may even go unnoticed.

Results (continued)

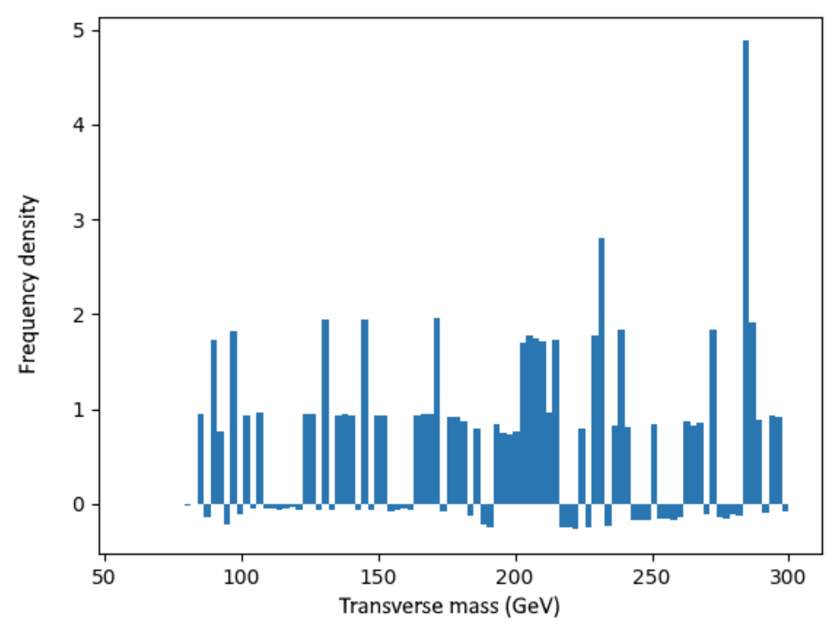
The main reason for this is likely the small amount of ATLAS sample data for events resulting in 4 leptons (in the range of megabytes), as these decays are relatively rare compared to 2lep decays. According to the standard model, there are other pathways through which the Higgs boson can decay, such as into two W bosons, resulting in a 2 lepton final state, or a decay into 2 photons. The ATLAS detector provides data for both leptons and photons, so either one of these decay paths can be feasibly analysed with the ATLAS Open Data, and since 2 lepton and 2 photon final states are far more common, they have far more available ATLAS data (multiple gigabytes). Even for scientists at CERN, the Higgs signal identified in 2012 was small enough that it could be almost disregarded as an insignificant blip by the untrained observer, so a large amount of data as well as extensive analysis is certainly necessary to draw meaningful conclusions.



The real data overlaid with the simulated data

Analysis & conclusions

After subtracting the background MC data from the real data and plotting the difference in a histogram, we see a small bump around the 125 GeV mark. This could indicate that a significant portion of the reconstructed 4 lepton final states could be correlated to a Higgs decay, since a standard object-selection criteria and event-selection criteria was used to identify this event. However, a different decay channel may have a larger sample of decays so that a clearer peak to signal the Higgs boson decaying can be seen. Perhaps, if we had access to a more powerful processor, we could carry out analysis of multiple sources of data from multiple different decay paths, and combine the results to further increase the accuracy of our findings, as this was the method by which the Higgs signal was originally uncovered. With regards to official scientific discoveries, a '5 sigma significance' is required to be able to deem the signal statistically meaningful; the probability of the signal being due to chance must be less than 1 in 3 million. Given our smaller sample size, it is almost impossible to achieve such a significance, but perhaps conducting further analysis of the statistical significance of our obtained signal could help testify to the legitimacy of our discovery.



The result after subtracting the simulated data

References:

[1] Aad, G., Abajyan, T., Abbott, B., Abdallah, J., Abdel Khalek, S., Abdelalim, A.A., Abdinov, O., Aben, R., Abi, B., Abolins, M., AbouZeid, O.S., Abramowicz, H., Abreu, H., Acharya, B.S., Adamczyk, L., Adams, D.L., Addy, T.N., Adelman, J., Adomeit, S. and Adragna, P. (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*, [online] 716(1), pp.1–29. doi:https://doi.org/10.1016/j.physletb.2012.08.020.

[2] Chatrchyan, S., Khachatryan, V., Sirunyan, A. M., Tumasyan, A., Adam, W., Bergauer, T., Dragicevic, M., Erö, J., Fabjan, C., Friedl, M., Frühwirth, R., Ghete, V. M., Hartl, C., Hörmann, N., Hrubec, J., Jeitler, M., Kiesenhofer, W., Knünz, V., Krammer, M. and Krätschmer, I. (2014). Measurement of the properties of a Higgs boson in the four-lepton final state. *Physical Review D*, 89(9). doi:https://doi.org/10.1103/physrevd.89.092007.

[3] Sirunyan, A.M., Tumasyan, A., Adam, W., Ambrogi, F., Asilar, E., Bergauer, T., Brandstetter, J., Brondolin, E., Dragicevic, M., Erö, J., Flechl, M., Friedl, M., Frühwirth, R., Ghete, V.M., Grossmann, J., Hrubec, J., Jeitler, M., König, A., Krammer, N. and Krätschmer, I. (2017). Constraints on anomalous Higgs boson couplings using production and decay information in the four-lepton final state. *Physics Letters B*, 775, pp.1–24. doi:https://doi.org/10.1016/j.physletb.2017.10.021.

L004