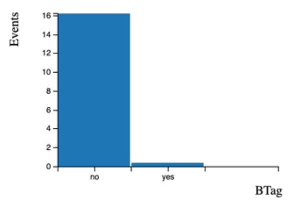
Project by **Haberdashers' Aske's Boys' School**:

Optimising the search for the Higgs Boson

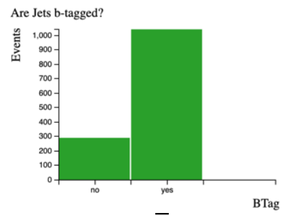
B-Tagging

The decay products of proton-proton collisions may include bottom quarks and antiquarks. Since quarks cannot exist alone, some of their kinetic energy is converted into the mass of more quarks and consequently b-hadrons are formed. These hadrons have relatively long lifetimes, so travel some distance before decaying within the detector.

The majority of $t\bar{t}$ events are b-tagged as top quarks often decay into bottom quarks, whilst most Higgs events have no b-tagging. Restricting the data to no b-tagging eliminates a large proportion of $t\bar{t}$ events.



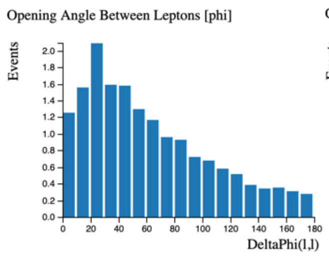
Higgs



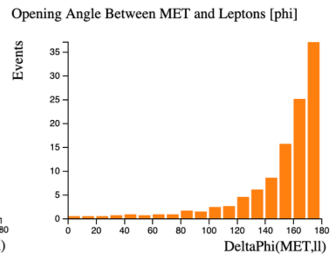
$t\bar{t}$

Opening Angle

Higgs events are likely to have a small opening angle between the leptons, whilst the $t\bar{t}$, Z and WW events are more likely to have obtuse opening angles between leptons.



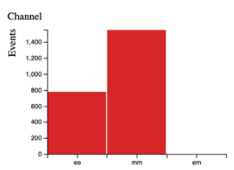
Higgs



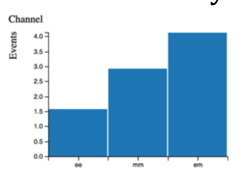
WW

Channel

All Z events occurred in the di-electron and di-muon channels, whilst 48% of Higgs events occurred in the electron-muon channel. Hence we restrict events to the em channel only.



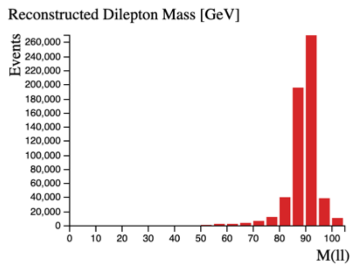
Z



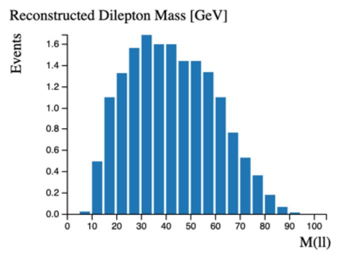
Higgs

Reconstructed Dilepton Mass

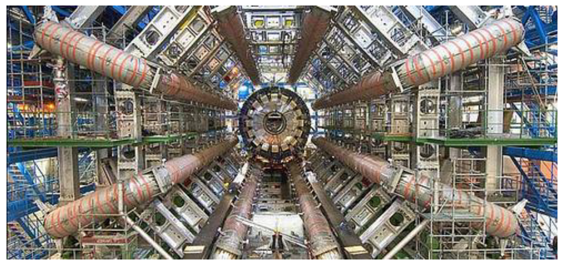
The reconstructed dilepton mass for Z events has a peak at 90 – 95 GeV, which we would expect since the Z boson itself has a mass of 91 GeV before decay. The peak reconstructed dilepton mass for a Higgs event is at around 30 – 35 GeV. Restricting this mass to values less than 80 GeV eliminates a significant proportion of the Z background events.



Z



Higgs



LEFT:

The ATLAS detector at CERN

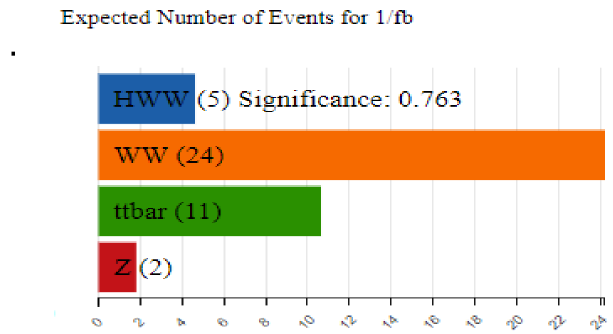
Total Lepton Transverse Momentum

Higgs events have a total lepton transverse momentum peak at around 50 – 60 GeV whilst Z events have a total lepton transverse momentum peak at 0 – 10 GeV.

Jets

Over 80% of Z events and over 70% of WW events have 0 jets, whilst over 50% of Higgs events have greater than or equal to 1 jet. Hence restricting the number of jets to one or more is an important step in increasing the significance of Higgs events.

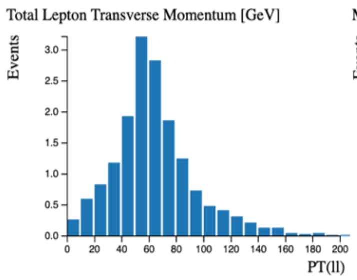
Our Optimised Significance:

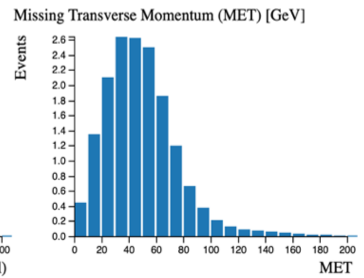


Missing Transverse Momentum

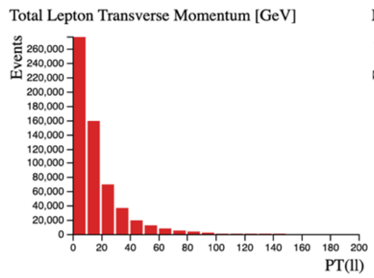
Before the collision, the momenta of the protons in the lateral direction relative to the axis of collision is 0. If after the collision there is a discrepancy between the momenta of the products in the direction perpendicular to the axis of collision, this indicates the production of undetected particles, such as neutrinos.

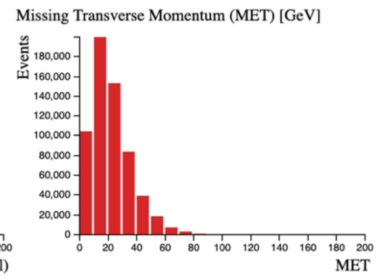
The peak missing transverse momentum for Higgs events occurs at 30 – 50 GeV whilst the peak missing transverse momentum for Z events occur at 10 GeV.





Higgs





Z

Estimation of amount of data necessary for statistically significant detection of the Higgs Boson

Having optimised the cuts for the histograms to attain **5** Higgs events and **37** background events, we assume that the relative proportions of Higgs events and background events will remain constant.

$$\sigma = \frac{Signal\ Events}{\sqrt{Background\ events}}$$

Let k be some constant which scales both the number of signal and background events proportionally. What value of k will yield a significance greater than 2σ?

$$\frac{5k}{\sqrt{37k}} > 2$$

$$k(25k - 148) > 0$$

$$k > 5.92$$

The amount of data analysed with these histograms was 1/fb. Assuming the distribution of all the events is uniform in larger data sets, we can estimate that roughly 6/fb of data will be necessary to find statistically significant evidence of the Higgs boson.

Monte Carlo Simulations vs Actual Data

Monte Carlo simulations refer to statistical techniques which model stochastic systems and produce probability distributions of possible outcome values.

By generating a certain amount of simulated signal and background data, MC simulations can show how many background events there are per filtering step and therefore improve the filters and estimate and reduce the systematic uncertainty in the analysis. However, where the simulated data and actual data could differ is normally because of experimental conditions during the collision that were unanticipated by the MC simulations, such as collision energies.

It is also because not all background events are included in the simulation; whilst the simulated signal data is a fixed, known amount, the background data does not need to be a certain value to estimate the relative uncertainty.