Available online at www.sciencedirect.com
ScienceDirect
Nuclear Physics B Proceedings Supplement 00 (2014) 1–7
Nuclear Physics B Proceedings Supplement

Sharing ATLAS data and research with young students

Maiken Pedersen, Farid Ould-Saada, Magnar K. Bugge
on behalf of the ATLAS collaboration and IPPOG International Masterclasses
maiken.pedersen@fys.uio.no
University of Oslo

Abstract

In recent years the International Masterclasses (IMC) featured the use of real experimental data as produced by the Large Hadron Collider (LHC) and collected by the detectors. We present ATLAS-based educational material using these data allowing high-school students to learn about properties of known particles and search for new phenomena. The ambition to bring to the "classrooms" important LHC discoveries is realised using the recent discovery of the Higgs boson. Approximately 10% of the ATLAS discovery data are made available for students to search for the Higgs boson: $2\,\mathrm{fb}^{-1}$ at 8 TeV for the Z path, and $1\,\mathrm{fb}^{-1}$ at 7 TeV for the W path, in the 2014 version of IMC. The Higgs study samples constitute one third of the total sample including Z, W and other low mass resonances. The educational material is tuned and expanded to follow LHC "heartbeats".

Keywords: ATLAS, IPPOG, outreach, particle physics

1. Introduction

The International Masterclasses (IMC) [1, 2, 3] is a particle physics outreach program run by the International Particle Physics Outreach Group (IPPOG) [4]. The aim of the program is to provide an opportunity for 15- to 19-year old school students to discover particle physics through hands-on measurements with real LHC data. Several different measurements are available, of which two involve ATLAS data. These two, the so-called "W path" [5, 6] and "Z path" [7, 8] measurements, are the subjects of this article.

The goal of the Z path measurement is to rely on the invariant mass concept to identify and measure properties of known particles, such as the Z boson, inferred from the decay products, pairs of leptons. When a heavy gauge boson Z' with mass 1 TeV is mixed with the real data, the simulated signal shows up in the di-lepton mass distribution, to the surprise of students, who realize that they have mastered a discovery tool. They go on and apply the same technique to di-photons and pairs of di-leptons to search for the Higgs boson. To help the build-up and display of the invariant mass distributions, we developed OPloT, a scalable, php-based web plotting tool for submission and automatic combination of all measurements performed. This allows for prompt access of results for further discussion within institutes and during videoconferences. The W path deals with the structure of the proton by comparing the numbers of W^+ and W^-, and the search for the Higgs into a pair of W bosons by measuring the angle between the leptons stemming from the W bosons.

2. Physics background

The Standard Model (SM) describes the fundamental interactions (except gravity) between the matter particles, the leptons (such as the electron and the muon) and quarks, as mediated by certain force carrier particles. The electromagnetic force is mediated by the photon, the weak force by the $W^\pm$ and Z^0 bosons, and the strong force by 8 gluons.

At the LHC, protons are collided head on at very high energy, and the particles emerging from the collisions are measured by instruments such as the ATLAS detector. The recorded data are used to improve our understanding of nature at the smallest distance scales.

Many *pp* interactions are in fact really interactions between individual proton constituents. Such interactions are referred to as *hard scatterings*. As the proton consists of up and down quarks held together by gluons and as gluons can "split" into quark-antiquark pairs, the proton consists effectively of both the up and down quarks (called *valence* quarks), gluons, and additional quarks and antiquarks from gluon splittings (called *sea* quarks).

The Parton Density Functions (PDFs) describe how often each of the various proton constituents will go into a hard scattering with a given momentum fraction x (the fraction of the proton momentum carried by the constituent). They summarize the proton structure as seen on a given energy scale.

The $W^{\pm}$ and Z^0 bosons may be produced in *pp* collisions via hard scatterings between quarks and antiquarks. For example, an up quark and a down antiquark may interact to form a W^+ boson, or a down quark and an up antiquark may interact to form a W^- boson. An up or down quark may interact with its respective antiquark to form a Z^0 boson. The $W^{\pm}$ and Z^0 bosons are short lived (with lifetimes of the order 10^{-25} s), and decay in practice immediately after their creation. They can therefore only be observed via their decay products, and decays into leptons are ideal for such detection.

The decays of the $W^{\pm}$ and Z^0 bosons into leptons provide characteristic experimental signatures. The decay of the Z^0 boson into a pair of oppositely charged leptons gives such a pair among the particles emerging from the collision (among the *final state* particles), where each lepton will typically have a large transverse momentum (perpendicular to the *beam line*). The decay of a W^+ or W^- boson into a charged lepton and a neutrino gives one charged lepton with a large transverse momentum. In addition, there will be a momentum imbalance in the transverse plane due to the neutrino, which is not measured directly by the detector. This momentum imbalance motivates the calculation of the *missing transverse momentum*, which can be thought of as an indirect measurement of the neutrino's transverse momentum.

With a much smaller probability, the Higgs boson may also be produced in a *pp* collision, mainly through gluon-gluon fusion via intermediate states involving the heavy top quark. The Higgs boson can decay to final states with two photons or four leptons via intermediate states of heavy particles, W^+W^- and $t\bar{t}$, or Z^0Z^0 respectively. The Higgs boson decays to final states with two photons or four leptons are golden channels for the detection of the Higgs boson at the LHC. Another important channel is the Higgs decay to a final state with two oppositely charged leptons and two neutrinos, which happens via a W^+W^- intermediate state.

The *branching fractions* of a particle quantify how often the particle will decay into each of the different possible sets of decay products. The Z^0 boson decays 6% of the time to an electron-positron pair (e^+e^-) or a muon pair ($\mu^+\mu^-$). The Higgs boson decays to two photons or four leptons (e^+e^- and/or $\mu^+\mu^-$) 0.2% and 0.01% of the time respectively. There are many processes occurring frequently in *pp* collisions that involve neither Z^0 nor Higgs bosons, and these constitute *backgrounds* to Z^0 and Higgs searches. Looking for the Z^0 boson and in particular the Higgs boson is therefore a bit like looking for the proverbial needle in the haystack. A good decay channel for a Z^0 or Higgs boson search must therefore not only have a sizeable branching fraction, but must also provide an experimental signature which is as easily as possible distinguished from those of the most frequent background processes. The demand for such a "clean" experimental signature favours decays into leptons and photons.

3. The *W* path and the structure of the proton

The electric charge of a W boson produced in a *pp* collision depends on the species of quark and antiquark that entered into the hard scattering. An up quark and a down antiquark will produce a W^+ boson, while a down quark and an up antiquark will produce a W^- boson. This follows simply from the conservation of electric charge. Furthermore, the decay of a W boson into a charged lepton and a neutrino leads to a final state charged lepton of the same charge as the W boson. Hence, by studying the charges of final state charged leptons from W boson decays in *pp* collisions, one can learn about the content of quarks and antiquarks inside the proton, i.e. about the *structure* of the proton.

In the W path measurement, the students learn how to identify collision events where a W boson may have been produced (a "W event") with a subsequent decay into a charged lepton (electron or muon) and a neutrino. As mentioned before, such events will in general contain one charged lepton with a large transverse momentum and large missing transverse momentum caused by the neutrino. This is exactly what the students need to look for to identify W events when they are analyzing the events one by one in the event display program MINERVA [9]. The students then count the number of

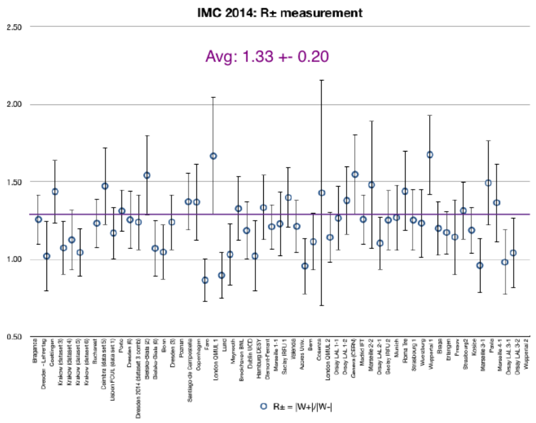


Figure 1: Example results from the 2014 *W* path Masterclasses. The charge ratio $R_{\pm} = N_{W^+}/N_{W^-}$ is shown with one point for each participating institute. The average, taking into account all submitted data, is 1.33 ± 0.20, which can be compared to the official ATLAS number 1.50 ± 0.03.

identified *W* events where the final state charged lepton is positive (i.e. the number of W^+ events), N_{W^+}, and the corresponding number of events where the charged lepton is negative, N_{W^-}. The final result of the analysis is the charge ratio N_{W^+}/N_{W^-} (fig. 1), a number which is sensitive to the structure of the proton.

Since the proton has two valence up quarks and only one valence down quark, a first naive approximation to the charge ratio is $N_{W^+}/N_{W^-} = 2$. The presence of sea quarks complicates the picture, and brings the charge ratio down to a value between 1 and 2.

As previously mentioned, the structure of the proton is summarized by the PDFs, so a charge ratio measurement can be used to test our knowledge of or constrain the PDFs of the proton. Such a measurement is an important physics result. The students doing the *W* path measurement are thus performing a measurement which is close to an actual important physics measurement performed by the ATLAS collaboration.

In the *W* path measurement, the students also look for events where two oppositely charged *W* bosons may have been produced with subsequent decays into charged leptons and neutrinos. Such events are characterized by two oppositely charged final state leptons with large transverse momenta and large missing transverse momentum caused by the two neutrinos. The production and decay of the Higgs boson may produce such a signature, and the students look at the distribution of the angle between the charged leptons in the transverse plane to search for the Higgs boson.

The *W* path measurement is performed using 7 TeV *pp* collision data consisting of about 1000 *W* candidates, 700 W^+W^- candidates, and 3300 background events (jets and *Z* candidates). Combination of results from individual student groups is performed using an online spreadsheet accessible via the *W* path web site [5]. Figs. 1 and 2 show some results of the 2014 *W* path Masterclasses. A typical *W* path Masterclass day proceeds much as described for the *Z* path in section 5.

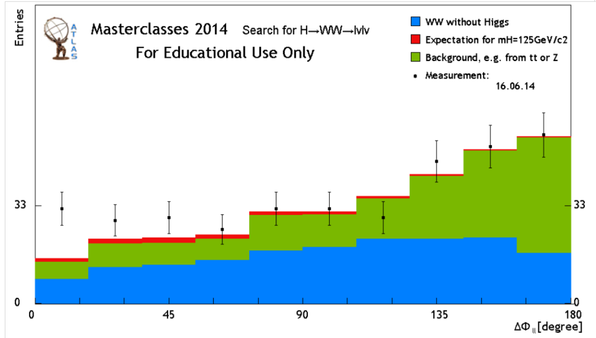


Figure 2: Example results from the 2014 *W* path Masterclasses. The distribution of the angle between the charged leptons in the transverse plane in W^+W^- candidate events is shown. An excess is seen at low values, but the excess is much larger than that expected from the Higgs signal.

4. The *Z* path and the invariant mass technique

The *Z* path measurement deals with the *invariant mass* technique for particle identification and discovery, which will now be presented. As mentioned, the *W* and *Z* bosons are short lived. The existence and properties of these and other short lived particles must be inferred from measurements of their decay products, and in this context, the invariant mass is a very useful concept.

Consider a massive short lived particle decaying into several lighter particles. The energy *E* of the short lived particle is related to its momentum $\mathbf{p}$ and mass *m* by

$$E^2 = \mathbf{p}^2c^2 + m^2c^4, \tag{1}$$

where *c* is the speed of light in vacuum. Assume now that the energies and momenta of the decay products are measured. Conservation of energy and momentum implies that

$$E = \sum_i E_i \quad \text{and} \quad \mathbf{p} = \sum_i \mathbf{p}_i, \tag{2}$$

where the sums run over the decay products and E_i ($\mathbf{p}_i$) is the energy (momentum) of decay product number *i*. Equations (1) and (2) lead to

$$m = \sqrt{\frac{1}{c^4}\left(\sum_i E_i\right)^2 - \frac{1}{c^2}\left(\sum_i \mathbf{p}_i\right)^2}. \tag{3}$$

The expression on the right hand side is known as the invariant mass, and can be calculated for any set of measured final state particles. In the case that the final state particles are the decay products of a short lived particle, the invariant mass is equal to the mass of the short lived particle.

One can search for short lived particles by plotting distributions of invariant masses of final state particles in *pp* collisions at the LHC. Short lived particles will give rise to peaks in such distributions. The *width* of a peak in an invariant mass distribution depends on the *natural width* of the corresponding short lived particle and the experimental resolution. In the limit of perfect experimental resolution, the width of each peak in an invariant mass distribution would be the corresponding particle's natural width, which is inversely proportional to its lifetime: the larger the width, the higher the decay probability, and the shorter the lifetime.

The students doing the Z path measurement learn how to identify electrons and positrons, muons, and photons in the ATLAS detector. They look for events containing (i) two oppositely charged leptons, (ii) two photons, or (iii) two pairs of oppositely charged leptons, i.e. four charged leptons in total. The goal of the measurement is to produce invariant mass distributions and look for peaks corresponding to short lived particles. The analysis performed by the students follows closely the general procedure used in many important physics analyses performed by the ATLAS collaboration.

In the distribution of the invariant mass of pairs of oppositely charged leptons, the students may discover the Z boson as well as the J/ψ and Υ *mesons*, each consisting of a quark and an antiquark bound together by the strong force. The peaks corresponding to these particles will be around $90\,\mathrm{GeV}/c^2$, $3\,\mathrm{GeV}/c^2$, and $10\,\mathrm{GeV}/c^2$ respectively. In addition to these well known particles, the students may discover a new, heavier, version of the Z boson, called the Z'. The latter is expected in theories involving hypothetical new weak interactions. Simulated events with the production and decay into leptons of this particle have been mixed in with the real data given to the students. This gives the students the possibility of really discovering something new and unexpected, and allows them to see how a new particle could be discovered at the LHC.

The distributions of the invariant mass of two photons and four charged leptons are sensitive to the production and decay of the Higgs boson because of its decays to these final states.

5. The Z path Masterclass

When attending a standard Z path Masterclass event, the students spend one full day at their local university. The program begins in the morning with lectures on both theoretical and experimental aspects of particle physics. In the theoretical lectures, the particles and forces of the SM are introduced. The experimental lecture introduces the invariant mass technique and explains how one can learn about short lived particles by studying their decay products. Furthermore, it deals with the experimental detection of particles using a particle detector. The structure of the ATLAS detector is introduced, and the students learn how different particles are "seen" in the detector.

After lunch, the students proceed with the actual practical measurement. Before they begin, there is a short demonstration where some key elements of the morning lectures are repeated. In particular, the procedures for identifying electrons and positrons, muons, and photons in ATLAS events are reviewed with some examples. Pairs of students share a computer, and proceed to analyze their own set of real LHC collision events recorded by ATLAS. Tutors are available for questions and guidance.

In the late afternoon, there is a results session. First, the results obtained by the students at the given university are discussed in a plenary session. Finally, the students take part in a video conference with all the other universities that participated in the IMC on that given day. The conference is led by moderators based at CERN. It includes discussion of the results obtained by all the universities, a quiz, and a question session where the students can ask the moderators about anything, for example what it is like to be a scientist and to be working at CERN.

5.1. The measurement

For the actual measurement, each group of two students is assigned a unique (except for intentional duplication of four lepton events) dataset containing 50 LHC collision events recorded by ATLAS. Some details on the event mixture are given in table 1.

The students analyze the events one by one by inspecting them visually in the event display program HYPATIA [10, 11]. For each event, the students should decide whether it could fall into one of the categories mentioned in section 4. If so, the students select the particles they believe to be electrons/positrons, muons, or photons, and HYPATIA calculates the invariant mass.

After the students have analyzed all their 50 events, a plain text file containing the calculated invariant masses

Event type	Number of events	Fraction in mixture
$Z \to l^+l^-$	18 500	50%
$J/\psi \to l^+l^-$	1 850	5%
$\Upsilon \to l^+l^-$	1 850	5%
$Z' \to l^+l^-$	1 850	5%
Four lepton	40	5%
Two photon	11 100	30%

Table 1: Details on the event mixture for the 2014 Z path Masterclasses. All the real data (everything except for the Z' events) are selected from runs 204769-206971 (period B12-C6, 2 fb^{-1}) recorded by ATLAS in June and July 2012. The numbers of events used from each category are shown, and correspond in general to the relevant fractions except in the case of four lepton candidate events, which are replicated many times in the students' datasets to allow all students the possibility of discovering such events. For all categories of l^+l^- events, there is a democratic division between electron and muon pairs.

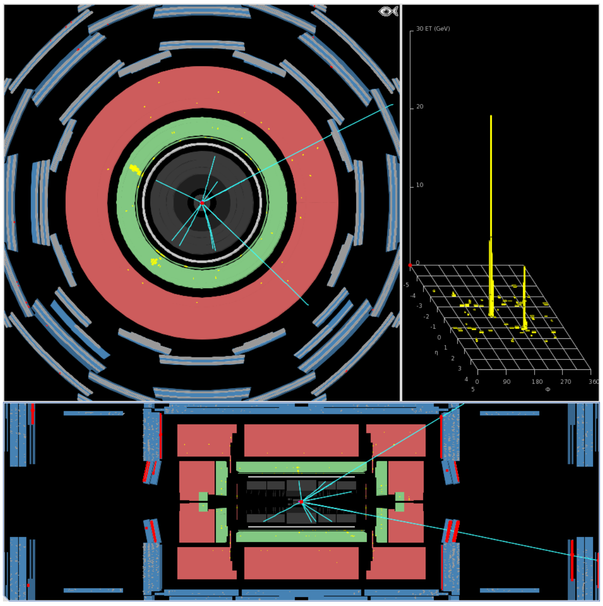


Figure 3: A HYPATIA event display showing two projections of the ATLAS detector, as well as the "lego plot" (top right) showing the distribution of measured energy in an "unfolded" calorimeter geometry. This event contains both an e^+e^- and a $\mu^+\mu^-$ pair.

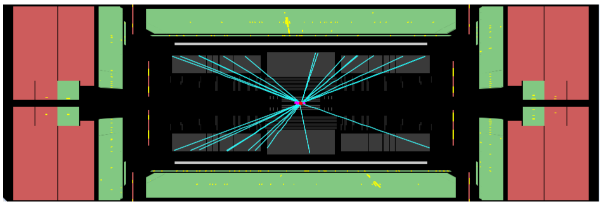


Figure 4: A HYPATIA event display of a typical event containing two photons. Only the longitudinal projection is shown.

is exported. The students upload this file to the online plotting tool OPloT [12], where they can look at the invariant mass distributions they have obtained. The results uploaded to OPloT are used for combination plots shown in the afternoon results session at the university and in the video conference.

5.2. *Event identification in HYPATIA*

Fig. 3 shows a HYPATIA event display. We see two projections of the ATLAS detector. In the *transverse* projection (top left), the beam line is perpendicular to the plane of the paper, while in the *longitudinal* projection (bottom), the beam line is horizontal and in the plane of the paper. In the tracking detectors (grey), we see reconstructed tracks corresponding to the trajectories of charged particles. In the electromagnetic (green) and hadronic (red) calorimeters, yellow dots correspond to measured energy deposits. In this event, there are two pronounced clusters of energy deposits in the electromagnetic calorimeter. Since there are also tracks in the tracking detector pointing in the direction of these clusters, we identify this event as containing an e^+e^- pair. Only muons pass through to and are detected in the muon spectrometer, the outermost part of the detector, and we see also a $\mu^+\mu^-$ pair in the event. An example of a typical two photon event is shown in fig. 4. Photons deposit energy in the electromagnetic calorimeter, but do not leave tracks in the tracking detector.

A photon may *convert* into an e^+e^- pair when interacting with the material of the tracking detector. In this case, there will be two tracks close together in the tracking detector pointing towards an energy cluster in the electromagnetic calorimeter. The students can calculate the invariant mass of the two tracks, which should be compatible with zero if they are indeed the result of a converted photon.

5.3. *The Oslo Plotting Tool (OPloT)*

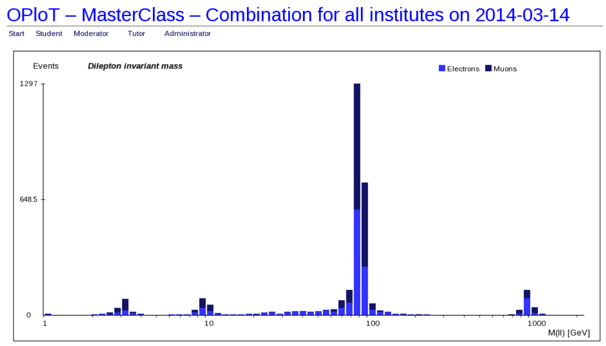
The Oslo online Plotting Tool (OPloT) is developed specifically for the analysis, combination, and presentation of results for the Z path Masterclass. Immediately after uploading their file, the students can study their own results in the form of invariant mass histograms. They can interactively change the invariant mass axis range, choose between linear and logarithmic binning, and set the number of bins. The invariant mass of two charged leptons, two photons, and four charged leptons can be viewed individually and together.

OPloT makes invariant mass distributions where the data from many students are combined. In particular, one can choose to combine all the student data from a given university on a given day, or all the student data from all universities taking part in the IMC on a given day. The former combination is used for the plenary discussion of results locally at each university, while the latter combination is used when results are discussed in the video conference at the end of the day.

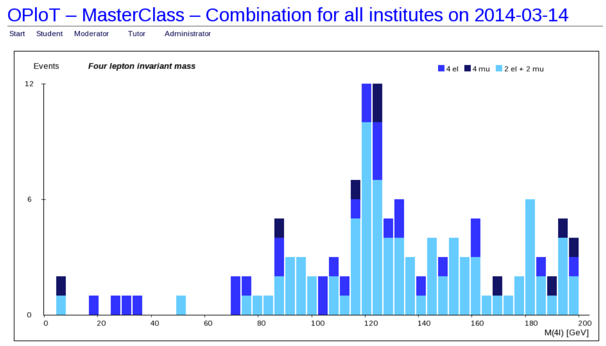
Fig. 5 shows the invariant mass distributions resulting from the combination of all submitted results from all universities taking part in the IMC on the 14th of March 2014. The two lepton invariant mass distribution in fig. 5(a) shows clear evidence of the J/ψ and Υ mesons as well as the Z boson. It is also clear that the students have discovered a new particle with a mass of $1\,\mathrm{TeV}/c^2$. Although this is because of the simulated events mixed in with the real data as mentioned earlier, it allows the students to see what the discovery of a new particle may look like. Obviously, it is explained to the students during the results session that the peak at $1\,\mathrm{TeV}/c^2$ is due to the simulated events. We also observe a smooth "continuum" distribution between the peaks. This must be coming primarily from misidentification by the students. It is interesting to note that e^+e^- events dominate completely the regions between the peaks, as expected from the fact that an electron-like experimental signature is more easily mimicked by hadrons, of which there are always plenty in pp collision events.

The four lepton ($e^+e^-e^+e^-$, $e^+e^-\mu^+\mu^-$, or $\mu^+\mu^-\mu^+\mu^-$) invariant mass distribution in fig. 5(b) shows that the students are very eager to look for such events, and possibly that they should be more critical in their particle identification. In fact, only 40 four lepton events were selected and mixed into the event samples for the 2014 Masterclasses, while the students have identified more than 200. We note the dominance of events of the type $e^+e^-\mu^+\mu^-$, which could be due to the misidentification of hadrons as electrons and positrons in events with a real muon pair.

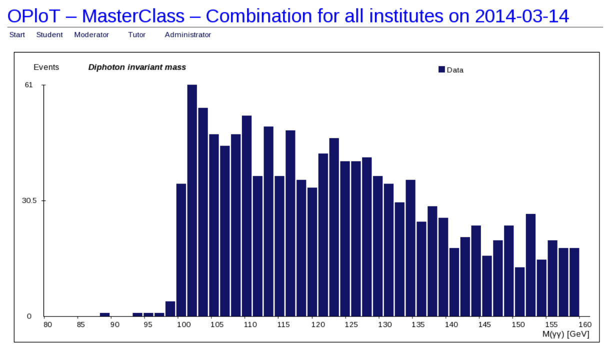
The two photon invariant mass distribution in fig. 5(c) looks pretty much as expected, but the statistical fluctuations are clearly too large for a small peak due to the Higgs boson to be discovered. It is important that the students understand that this is a limitation of the size of the data sample, and that the two photon invariant mass distribution was in fact a key ingredient in the Higgs discovery at the LHC. This is discussed in the plenary results session and the video conference, and to aid the discussion, one can in OPloT choose to display simulated data corresponding to different data sample sizes



(a) The combined two lepton invariant mass distribution.



(b) The combined four lepton invariant mass distribution.



(c) The combined two photon invariant mass distribution.

Figure 5: The invariant mass distributions resulting from the combination of all submitted results from all universities taking part in the IMC on the 14th of March 2014. The two lepton (a), four lepton (b), and two photon (c) invariant mass distributions are shown. The interested reader can view these and other results interactively via the "Moderator" tab in OPloT [12]. The students find too many four lepton events. While the four lepton events are replicated many times in the students' data, this fact is not causing the four lepton excess as OPloT automatically removes the overlap.

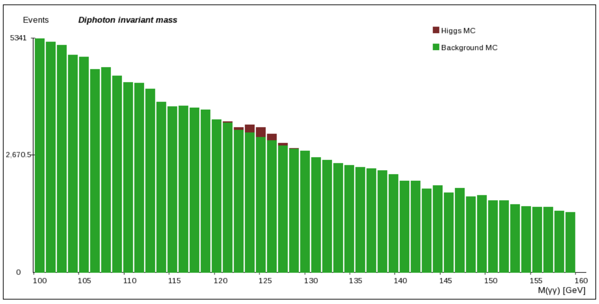


Figure 6: The two photon invariant mass distribution presented by OPloT when set to display simulated background and Higgs signal corresponding to an integrated luminosity of 25 fb^{-1}.

in order to show how a Higgs peak becomes more apparent as the amount of data increases. The simulated data corresponding to a large data sample (25 fb^{-1} of integrated luminosity, which can be compared to the 2 fb^{-1} available to the students[1]) is shown in fig. 6. Here, it should be possible to convince oneself that the peak due to the Higgs boson would be visible even if it were the same colour as the background.

Even though examples are not shown here, there are further interesting possibilities in OPloT for the two photon invariant mass distribution. One can choose to compare the students' distribution to simulated background and signal distributions corresponding to the size of the data sample analyzed by the students. Doing so for the results of the 14th of March 2014, we find that the students have only identified about half as many two photon events as expected. A significant part of the mismatch between expected and observed event counts is assumed to be due to student groups which do not manage to analyze all their 50 events within the available time. The student data can also be replaced by the "correct" distribution, resulting from the selection of events using ATLAS software analysis procedures.

Also in the four lepton invariant mass distribution, one can choose to display simulated Higgs signal in OPloT. One can see the expected size of a Higgs signal in this distribution, and discuss in the plenary session and video conference how it could be used to discover the Higgs boson. The students should understand that also the four lepton invariant mass distribution was a key ingredient in the Higgs discovery at the LHC, but that it looks very different from the two photon distribution, as the expected numbers of background and signal events are both much smaller in the four lepton case.

6. Summary and outlook

The ATLAS *W* and *Z* path Masterclass measurements have been presented. While the *W* path deals with both the structure of the proton and the search for the Higgs boson, the *Z* path is devoted completely to searches for short lived particles and mass and width measurements.

The interpretation of the *Z* path results in terms of short lived particles has been reviewed, and while evidence of the Higgs boson could not be observed in the students' distributions, the two photon and four lepton invariant mass distributions were both discussed in terms of the Higgs search at the LHC.

The *W* path data sample will be improved in the near future to increase the event identification success rate and decrease the complexity. The *Z* path measurement is also still evolving. In the future we hope to bring new discoveries to the public. Among new features to be implemented in the *Z* path, a signal of graviton resonances in di-lepton, di-photon mass distributions, and the exploitation of missing transverse momentum to study dilepton invariant mass endpoints of supersymmetric particles. Finally, a script is already available for more advanced university students to loop through the full *Z* path dataset.

[1] In the 2014 *Z* path data, about 60 Higgs events are expected in the two photon invariant mass distribution. In the four lepton distribution, between one and two Higgs events are expected, but there are several candidates in the range 120-130 GeV.

References

[1] U. Bilow and M. Kobel, *International Masterclasses – bringing LHC data to school children*, EPJ Web of Conferences **71** (2014)
[2] F. Ould-Saada, *International Particle Physics Masterclasses – Bringing LHC data into the Classroom*, PoS ICHEP2012 (2013) **559**
[3] International Masterclasses, `http://physicsmasterclasses.org/`
[4] International Particle Physics Outreach Group (IPPOG), `http://ippog.web.cern.ch/`
[5] K. Jende et al., ATLAS *W* path measurement, `http://atlas.physicsmasterclasses.org/en/wpath.htm`
[6] U. Bilow et al., *ATLAS W path – real data from the LHC for high school students*, EPJ Web of Conferences **71** (2014)
[7] M. Pedersen et al., ATLAS *Z* path measurement, `http://atlas.physicsmasterclasses.org/en/zpath.htm`
[8] M. K. Bugge et al., *ATLAS Masterclasses – W and Z path physics and presentation of the Z path measurement*, EPJ Web of Conferences **71** (2014)
[9] T. McLaughlan et al., MINERVA, `http://atlas-minerva.web.cern.ch/atlas-minerva/`
[10] S. Vourakis, *Bringing high energy physics to the classroom with HY.P.A.T.I.A.*, EPJ Web of Conferences **71** (2014)
[11] Ch. Kourkoumelis et al., HYPATIA, `http://hypatia.phys.uoa.gr/`
[12] V. Morisbak et al., OPloT, `http://cernmasterclass.uio.no/`