Summary Report for the Institute of Particle Physics

CERN Summer Student Project

Thomas G. McCarthy

August 22, 2008

1 Introduction and Motivation

The Large Hadron Collider (LHC) located at CERN near Geneva, Switzerland is now in its commissioning stage. A test beam of protons was injected into the LHC pipe in the last few weeks, where the protons were then successfully steered around a portion of the 27 km underground ring. The various detectors in the LHC collaboration will soon shift from taking cosmic data to data generated as the result of collisions between bunches of up to 10^{11} protons, each of which will interact at a rate of approximately 40 million events per second. With a total centre of mass energy of 14 TeV, the LHC promises to take experimental particle physics to an exciting new level of discovery in the upcoming years. The scientists and engineers involved in the commissioning of the ATLAS detector (Figure 1), a multi-purpose detector located approximately one hundred meters underground at one segment of the LHC beam pipe, are now running daily tests of each of the detector's subcomponents to ensure the they will work together for the first collisions between protons, set to take place in the next few months.

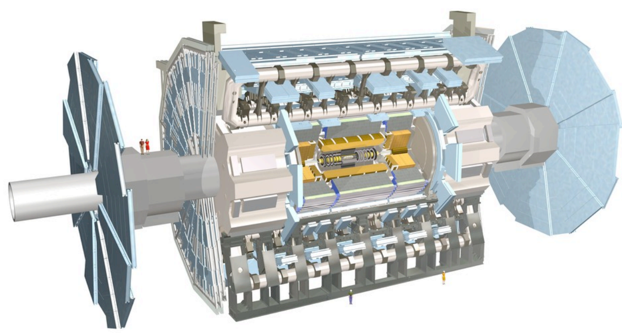

Figure 1: A computer generated view of the entire ATLAS detector

The analysis and understanding of simulated data play a crucial role in being able to interpret results obtained from the real data which will be generated once the LHC and the detector are in full operation. For this reason, there are many working groups composed of physicists who are currently running their analyses on data created by different Monte Carlo generators along with a program used to simulate interactions of particles in the different types of matter in the detector itself. Once the data have been generated through the Monte Carlo event generator and subsequently run through the detector simulation, the output data are then in the same form in which one would expect to see data coming from real interactions, though the real data will hopefully also present physicists with many unexpected results as well.

My research project involved the study of a set of simulated $Z \rightarrow e^+e^-$ events, one of the possible decay modes of the neutral, weak gauge boson. This particular channel accounts for approximately $(3.363 \pm 0.004)\%$ of Z boson decays [1]. The data set had already been generated for the general ATLAS collaboration using a Pythia event generator followed by the detector simulation. The events themselves were then reconstructed based on specific job instructions sent through ATHENA, a software package which allows those wanting to reconstruct events, whether they be from simulated or real data, to specify what exact filters they want placed on the reconstruction to keep file sizes small while still providing the exact information in which they are interested. A more detailed explanation of what exact filters were applied to the set of data used is explained more specifically in the following section. The important point from this is that all filtering of the data was done before the data were passed on to me in order to produce a summer project appropriate for the six week duration of my time at CERN.

My study involved first developing a feeling for the variables and typical quantities of the particular reaction given the selection criteria, and to look for correlations and or problems between the reconstructed values and the true values as they were produced in the event generator. In order to be able to examine my results, we first need to set the stage by introducing the relevant variables and the coordinate axes used at ATLAS.

2 Kinematic Variables of $Z \to e^+e^-$

When generating histograms of different variables for the particular interaction in which one is interested, it is first necessary to define the variables which will be used. There are many potential variables of interest associated with the electrons or positrons with which we dealt, though only a select few were made available with the idea that once these were understood, selection criteria could then be changed to request additional variables of interest when desired.

The coordinate system used for all ATLAS related analyses is as follows: the z-axis is that of the LHC beam pipe, with the positive x-axis always pointing towards the interior of the accelerating ring, and the positive y-axis pointing vertically upwards. The azimuthal angle is measured from the positive x-axis towards the positive y-axis such that $\phi = 0$ at any point on the positive x-axis. One could then define the polar angle, θ, in the typical way in which it is defined in spherical or cylindrical coordinates as an angle measured from the z-axis, though in the case of ATLAS a more useful term is the *pseudorapidity*, η, defined such that:

$$\eta \equiv -\ln\tan\left(\frac{\theta}{2}\right) \tag{1}$$

For a given particle, in this case electrons or positrons, energy E and momentum p are both useful variables to examine, and we may also be interested in the three main components of the particle's momentum: p_x, p_y, and p_z. It is often also useful to know a particle's total momentum in the x-y plane which we define as the *transverse momentum*, p_T, such that:

$$p_T = \sqrt{{p_x}^2 + {p_y}^2} \tag{2}$$

We further define *Electron0* and *Electron1* as the name given to either of the leptons produced in the $Z \to e^+e^-$ reaction, where *Electron0* can represent either the electron or positron, and similarly for *Electron1*. In this sense, it does not make sense to speak of the charge of *Electron0* or *Electron1*; it is understood that each carries a positive or negative unit electron charge. Whether the lepton in question is defined as *Electron0* or *Electron1* is set in the event generator stage of the data simulation, such that the lepton is defined as *Electron1* when $0 \leq \phi \leq \pi$ and as *Electron0* when $-\pi \leq \phi \leq 0$. The asymmetric result of this nomenclature in the y direction can be seen in Figure 2 below — an example of one of the early histograms initially plotted in order to acquaint myself with the relevant variables.

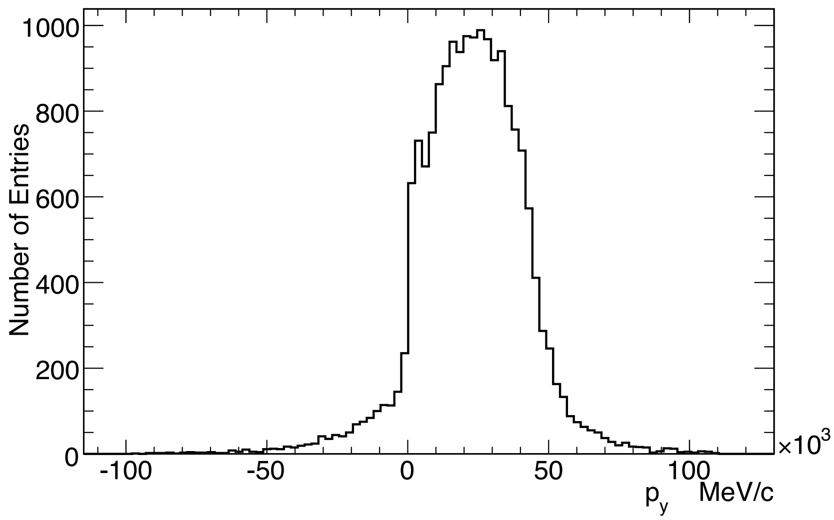

Figure 2: Distribution of the component of Electron1's momentum in the y-direction

When events were generated by Pythia, a set of pre-selection cuts were made to the data. The first of such cuts disregards any electron or positron (i.e. *Electron0* or *Electron1*) whose total p_{T} is less than 10 GeV/c, even if physical processes would allow such values. A further pre-selection cut disregards all electrons or positrons whose trajectory is such that $|\eta| > 2.7$.

In addition to the pre-selection cuts which were made during the event generation, further selection criteria were then placed on the data. One such pre-selection cut is the requirement to have an integer number of electron-positron pairs. And finally, due to gaps between the different components of the ATLAS detector, it was desired to ignore all electrons or positrons whose trajectory lies in the region where $1.4 \leq |\eta| \leq 1.5$.

Finally, one of the ultimate goals in using this simulated data is to reconstruct the mass of the initial particle — in this case the Z boson. We would then like to see how this reconstructed mass value compares with the true mass value, where the two values are indicated by m_Z^{Reco} and m_Z^{Truth}, respectively. These final two variables will be examined in the following section of the report.

We are now at a point where we may begin to plot histograms of some of the variables of interest. A selection of histograms which I generated are presented here in order to provide the reader with a sense of the typical distributions of relevant variables in this particular data set, in much the same way that I began my own familiarization with the process. Note that, prior to further cuts on the data set aside from those mentioned above, the total number of entries in the histogram is 18251.

Figure 3 shows the distribution of *pseudorapidity* values corresponding to the trajectories of *Electron1*. The gaps correspond to the disregarded η values in the selection criteria mentioned above, and one can see that the distribution is approximately symmetric about $\eta = 0$, as expected. Figure 4 shows the distribution of p_{T} values, where one can notice the results of the pre-selection cut in the region where p_T <10 GeV/c.

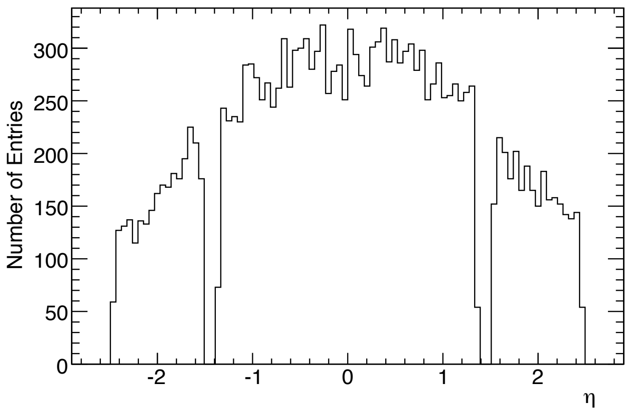

Figure 3: η distribution for Electron1

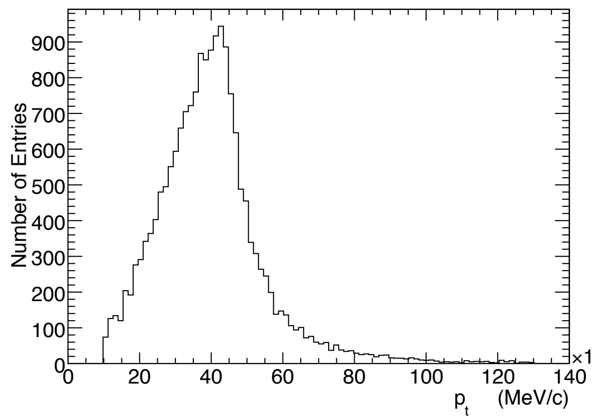

Figure 4: p_T distribution for Electron0

Next it is interesting to look at the momentum of the electrons or positrons in more detail, specifically to get a sense of how their momentum is divided up into the different components. Figure 5 shows the correlation, particularly at higher total momentum, between p_z and p_{total} for *Electron0*. It can be seen that the greater the lepton's total momentum, the more likely it is that most of its momentum will be in the z-direction, and again there is the symmetry one would expect for particles travelling either in the positive or negative z-direction. The relationship between *Electron0*'s transverse momentum and z-component of momentum, and specifically how the momentum is divided unevently between the compoments, can be seen in Figure 6. The two gaps visible in this figure, though not quite as intuitive as in previous plots, again are the result of the gap in η which can be seen by plotting one of the variables versus η itself. Alternatively it can be seen by convincing oneself this is the case by choosing a point in one of the gap regions and using it to construct a right-angle triangle to determine the θ value corresponding to that particle's trajectory, and subsequently applying equation (1) to determine the pseudorapidity.

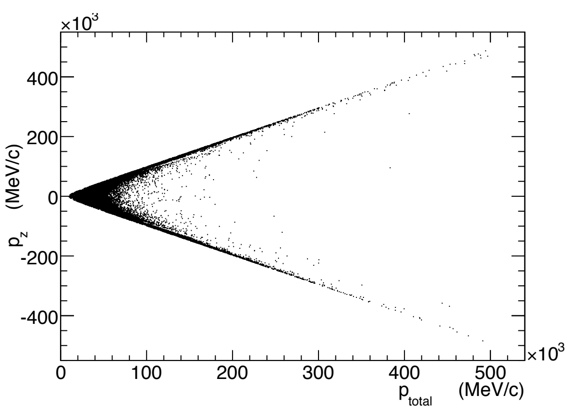

Figure 5: Relationship between Electron0's p_z and total momentum

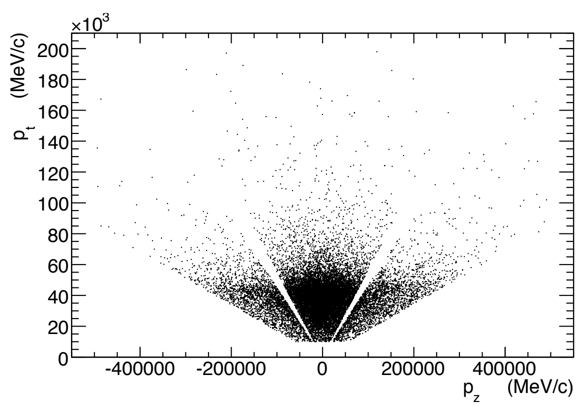

Figure 6: Relationship between p_{T} and p_z for Electron0

Many other interesting histograms could of course be examined, but we would like to now move on to examining the plots of both the reconstructed mass and the true mass of the *Z* boson.

3 Investigation of the Z Mass

Current measurements [1] of the mass of the Z boson set its value at (91.1876 ± 0.0021) GeV/c^2 with a full width of (2.4952 ± 0.0023) GeV/c^2 and the distributions of m_Z^{Reco} and m_Z^{Truth} below (Figure 7 and Figure 8, respectively) seem to agree reasonably well with that value. As one would expect, the truth peak in Figure 8 is narrower than that of Figure 7 and has a full width of approximately 3 GeV/c^2 — close to what is expected in this case.

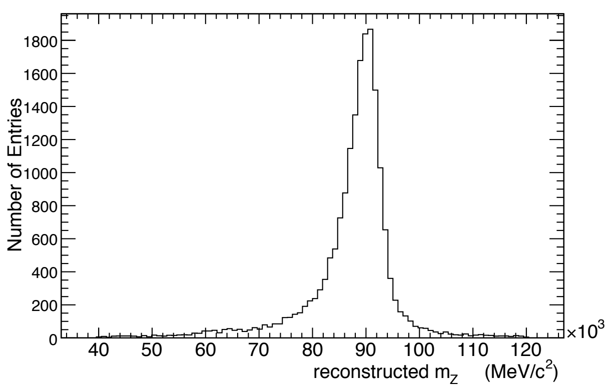

Figure 7: Z reconstructed mass distribution

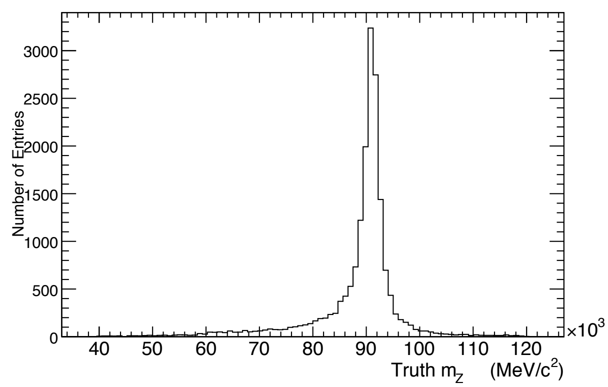

Figure 8: Z true mass distribution

Although the m_Z^{Truth} distribution is a sharper peak, events in either the truth or reconstructed distribution are matched using a certain algorithm, meaning that when an event is reconstructed to obtain a value for m_Z^{Reco} we would expect that event to have a similar value for m_Z^{Truth}. An interesting result is obtained, however, when a plot is produced showing m_Z^{Reco} versus m_Z^{Truth} which can be seen in Figure 9 below. For a perfect correlation, one would expect to see a smooth diagonal line in this plot. The horizontal and vertical lines visible in Figure 9 corresponding to near-constant m_Z^{Reco} and m_Z^{Truth} values, respectively, were an unexpected result which required further investigation. From this plot, one may note that, for example, a reconstructed Z mass was found to be approximately 90 GeV/c^2 whereas for the same event, the true value of the Z mass was only approximately 20 GeV/c^2.

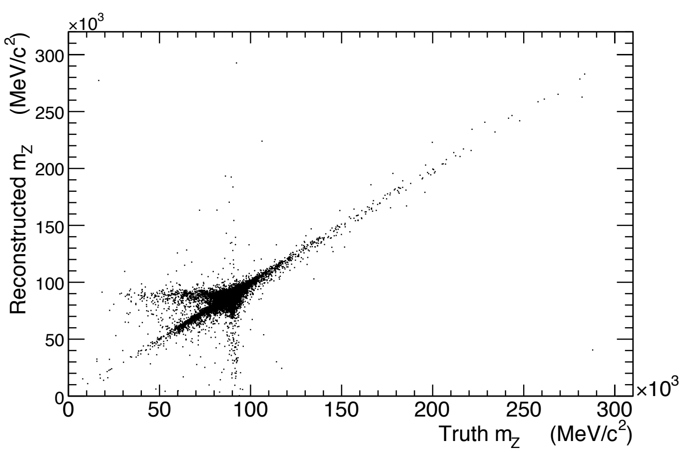

Figure 9: Plot of m_Z^{Reco} versus m_Z^{Truth}

As a next step, we could now eliminate all points corresponding to events in which both reconstructed and truth values produce the same Z mass to within an arbitrary value — the points one would expect to see with a very strong correlation. The events which remain would then potentially provide some clues as to the pattern we see in Figure 9. If we select, for example, only events for which the reconstructed mass differs from the true mass by more than 19 GeV/c^2, that is we keep only events such that:

$$|m_Z^{Reco} - m_Z^{Truth}| > 19 \;\; GeV/c^2 \tag{3}$$

we can then recreate a histogram of m_Z^{Truth} with the condition given by (3). Figure 10 shows this distribution, and one can note the second peak with centre located somewhere between 60 GeV/c^2 and 70 GeV/c^2. It should be noted that whereas we started with 18251 events, after having made the selection given by (3), we are down to 518 events; this does not mean the second peak shown in Figure 10 is negligible, but it is important to get a sense of how much of an effect this has on the overall mass spectrum.

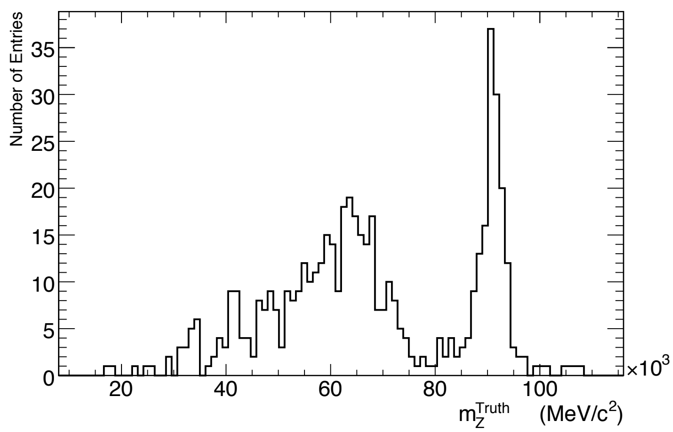

Figure 10: m_Z^{Truth} distribution after having made the cuts given by equation (3)

Now keeping the same set of data, we plot $\frac{p_T^{Reco}}{p_T^{Truth}}$ for *Electron0* versus m_Z^{Truth} (Figure 11). We already know that all of the events which produced a point on this plot already disagree in their value of m_Z by at least 19 GeV/c^2. In addition to this, it can be seen that while several of the events agree in their value of transverse momentum, there are many points that disagree quite substantially. Note, for example, that there are even a few events in which p_T^{Reco} and p_T^{Truth} disagree with each other by nearly a factor of ten.

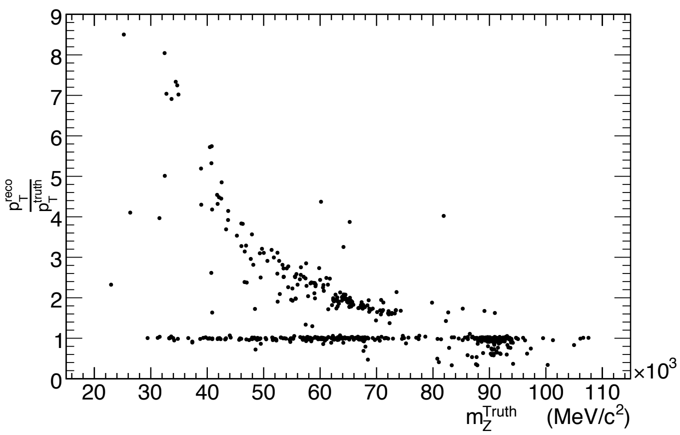

Figure 11: Plot of reconstructed to true transverse momentum of Electron0 versus m_Z^{Truth}

As a final step, we remove the condition given by equation (3) and instead specify that we will only look at events corresponding to suspicious points in Figure 11, that is to say we keep only those events such that:

$$\frac{p_T^{Reco}}{p_T^{Truth}} > 1.2 \tag{4}$$

Using these events, we can now plot a distribution of m_Z^{Truth} a third time (Figure 12), to see that the true Z mass for such events is indeed lower than it should be. Perhaps more insightful is the plot showing m_Z^{Reco} versus m_Z^{Truth}, as we see in Figure 13, which shows that while the reconstructed Z mass values are where one might expect them to be, the true Z mass values are spread from 93 GeV/c^2 down to values as low as 10 GeV/c^2. It should be noted as well that simply plotting p_T^{Reco} or p_T^{Truth} on their own produces a range in values similar to the range which would be observed by plotting either transverse momentum distribution without applying either (3) or (4), so this definitely dismisses the possibility that they simply happen to be erroneous points.

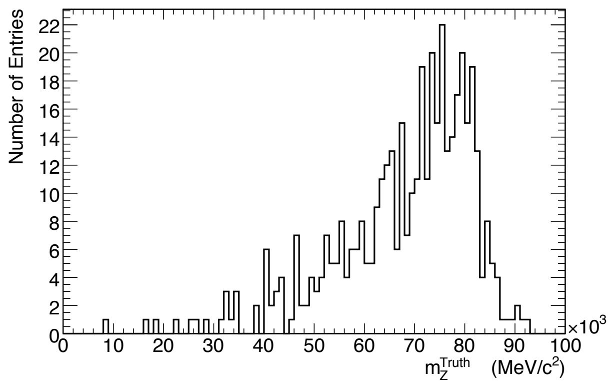

Figure 12: m_Z^{Truth} distribution after having made the cuts from (4)

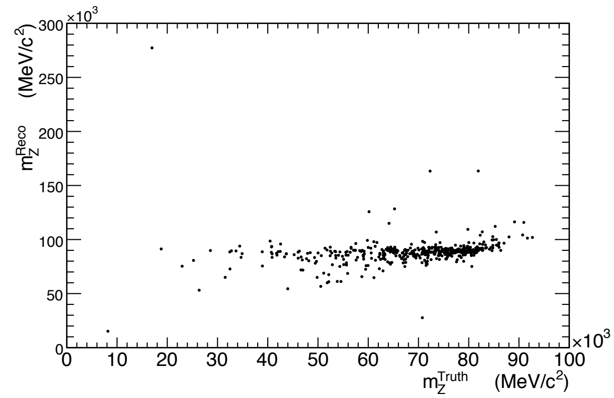

Figure 13: Plot of m_Z^{Reco} vs m_Z^{Truth} subject to equation (4)

At this point, it seems clear that when looking at either the truth of reconstructed values separately, the numbers are more or less what one expects they might be; it is only when one examines the reconstructed values versus the truth values for particular events that we get unexpected results, which seems to indicate that there are issues with the matching algorithm between reconstructed and truth values.

4 Summary and Suggested Next Steps

From my investigations of this set of data, it seems that there is a problem in being able to match some truth events to reconstructed events, which is made evident by some of the later plots in the preceding section. To proceed further, one needs to understand the matching algorithm used.

At a simplified level, this algorithm which matches truth events to reconstructed events does so by examining the neighboring area around an electron's or a positron's reconstructed trajectory, where the area ΔR in question, in this case, is defined such that:

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} \tag{5}$$

In order to further investigate the suspected problems with this matching algorithm, my next step would be to look at the source file in an effort to understand exactly how the algorithm is structured, and in fact it is felt that without going on to this next step, it would be unable to determine exactly where the problem lies.

In terms of this current project, the past weeks at CERN has provided the opportunity to become familiar with how to generate histograms and plots to look at several variables of interest in typical particle physics reaction, and these skills will no doubt be of use in similar projects in the future.

5 Acknowledgments

I wish to thank the Institute of Particle Physics for providing the funding in order to allow me to participate in the CERN Summer Student Program.

6 Additional Experiences While at CERN

In addition to the summer research project on which this report is based, participating in the CERN Summer Student Program has provided me with several other chances to develop the skills I will need as I continue a career in experimental particle physics. The following is a list of the different professional experiences I have had over the course of the last six weeks:

- Involvement in summer research project involving $Z \rightarrow e^+e^-$ investigation (Supervisor: Dr. Manuella Vincter)
- Development of computing skills with the following programs or systems: C/C++, familiarity with Linux systems, and the ROOT software package
- Participation in ATLAS Software Tutorial with an introduction to the GRID. This provided me with a basic overview and understanding of the following computing tools or software packages used for performing physics analysis: Athena, Ganga, Python & Atlantis
- Participation in the Liquid Argon Group meetings, a subgroup of the entire ATLAS collaboration, as well as the chance to train for a full weekend as a shifter in the ATLAS control room, helping to perform calibration runs for the Liquid Argon components of the detector as well as monitoring our section during cosmic data collection
- Attendance at the morning summer student lectures held in the main auditorium at CERN on a variety of topics relevant to experimental and theoretical physics today
- Opportunities to tour different parts of CERN including the ATLAS detector and control room, one of the linear accelerators, CERN's computing centre, and most recently a tour of the LHC control room

References

[1] C. Amsler *et al.* (Particle Data Group), Phys. Lett. **B667**, 1 (2008).