Lab 3: ATLAS Data Analysis
Gillian La Vina
May 2025

Introduction

Using high energy proton beams, the Large Hadron Collider (LHC) at CERN breaks apart protons in order to study the fundamental particles that arise from the proton-proton (pp) collisions.

In this analysis, we focus on one particular fundamental particle, the Z^0 boson. Similar to how photons carry electromagnetic forces, the Z^0 boson is the neutral carrier for the weak nuclear force.

The Z^0 boson is difficult to measure directly as it is short-lived and quickly decays. In around 10% of all cases, the Z^0 boson decays into a pair of charged leptons, whose properties the ATLAS detector at the LHC can measure more easily. From the ATLAS detector we can measure the total energy, transverse momentum, pseudorapidity (angle of particle's momentum along the incident beamline), and azimuthal angle from the beam of each lepton.

With the lepton data, we can reconstruct some properties of the original Z^0 boson. Due to conservation of matter and energy, the energy of the leptons resulting from a Z^0 decay must at minimum sum to the mass of the Z^0. In this study, we use the measured momentum and energy of each of the lepton pairs to find the mass of the original boson. After finding the reconstructed mass, we sort the data into a histogram to show where the event peaks. Finally, we fit our data to the Breit-Wigner function to find the exact point where the data peaks, which we take to be the mass of the Z^0 boson. Additionally, we analyze the accuracy of our fitted peak by performing chi square tests along the parameter space for our fit parameters.

Invariant Mass Distribution and Fit

We use data from the ATLAS detector to reconstruct the masses of each decayed lepton pair and arrange the distributions of masses into a histogram, which should show an excess at the rest mass of the Z^0 boson.

First, we use raw ATLAS data of lepton pairs to reconstruct the mass of the original Z^0 boson. We use the data for the transverse momentum p_T, the pseudorapidity η (angle with respect to the beamline) and azimuthal angle ϕ to find the momenta in the x, y, and z directions of each lepton, using the equations:

$$p_x = p_T cos(\phi), \quad p_y = p_T sin(\phi), \quad p_z = p_T sinh(\eta)$$

where p_x, p_y, and p_z are the components of the momentum in each direction.

We then sum up the individual energies and momenta of each lepton pair to get the total energy and momentum of each pair. From here, we get the total mass of each lepton pair, and thus the reconstructed mass of the original particle using the equation:

$$M_{tot} = \sqrt{(E_1 + E_2)^2 - (\sum_{i=x,y,z} {p_{i1}}^2 + {p_{i2}}^2)}$$

where E_1 and E_2 are the measured energies, and p_{i1} and p_{i2} being the momenta of each particle in the i direction (with i = x, y, or z).

We then make a histogram showing the distribution of reconstructed masses from the double lepton events. Eventually, we expect this histogram to show an excess or peak at the mass of Z^0. First we decide how to bin the masses; we take 41 bins from 80 to 100 GeV, or masses at every 0.5 GeV starting from 80 GeV. Given our array of reconstructed masses, the Python library Matplotlib contains a function "hist" which sorts the masses into the decided bins, and creates a plot for us (Fig 1).

We also add simple error bars to our histogram. The error in counts for each bin is taken to be the square root of the number of events per bin, as per Poisson counting statistics. Notably, these error bars represent a minimum error, not taking into account errors in the measurements. These error bars can be seen in Fig 1 - Histogram of Calculated Masses.

Next, we fit the data to theory. We compare our histogram of mass values to the theoretical distribution of decays according to the Breit-Wigner formula:

$$D(m; m_0, \Gamma) = \frac{1}{\pi} \frac{\Gamma/2}{(m - m_0)^2 + (\Gamma/2)^2}$$

where m_0 represents the true rest mass, appearing as a center peak of the function, and $\mathbf{\Gamma}$ is the parameter related to the lifetime of the particle and mathematically dictates the width of the peak. Qualitatively, the Breit-Wigner distribution resembles a Gaussian with a distinct peak at the center. Physically, the Breit-Wigner distribution describes the mass profile of unstable particles, with the true rest mass of the particle as the center peak of the distribution. The Python function "curve_fit" from the library scipy.optimize is used to find the best values of m_0 and $\mathbf{\Gamma}$ which make the distribution function closest to the data. Notably, we only apply our fit to the center of data near the peak, with a range from 87 to 93 GeV. The fit function is overlaid onto the histogram of data, shown in dark blue in Fig 1.

After fitting the reconstructed mass data to the Breit-Wigner distribution, we find the rest mass of Z^0 to be **90.3 ±0.1 GeV**.

Next, we use residuals and chi-square tests to evaluate how well the data matches theory. First, we can visualize how well the data and theory agree by plotting the residuals (i.e., the difference in counts between data and theory). The bottom panel of Fig 1 shows the residuals between data and fit, with error bars, and a blue line representing zero difference between data and theory, i.e. complete agreement. Then, we can quantitatively judge the accuracy of the fit using a chi-square test. To do this, we first find the degrees of freedom. We applied our fit to the middle 12 bins of the histogram (from 87-93 GeV), giving us 12 data points, and since we performed the fit across 2 parameters (m_0 and $\mathbf{\Gamma}$), we subtract 2 from the number data points to **10 degrees of freedom**. From here, we calculate the chi-square value using the equation:

$$\chi^2 = \sum_{i=1}^{N} \frac{(y_{data} - y_{theory})}{\sigma^2}$$

where y_{data} is the counts per bin, y_{theory} is theoretical number of counts at each bin calculated using the Breit-Wigner function, and σ is the error in the number counts (according to Poisson statistics, the square root of counts). The chi squared value between our fit and the data is calculated to be **10.0**, with a reduced chi squared of 1. A Python function from scipy.stats is used to calculate the p-value to be **0.44**. The p-value is within the $0.05<p<0.95$ range for good agreement, indicating a statistically reasonable fit, i.e., our data fits the Breit Wigner distribution.

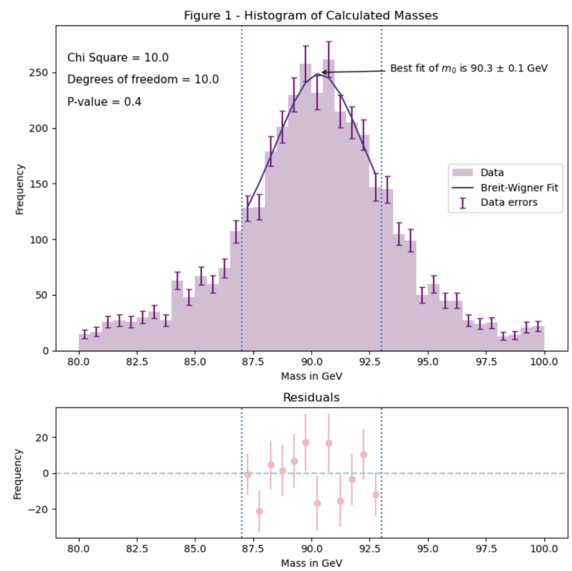

Figure 1: Calculated mass sums for each of the lepton pairs arranged into a histogram, as well as a plot of the residuals between data and (Breit-Wigner) theory

2D Parameter Scan

To further examine how the theory fits data, we perform a 2D parameter scan, calculating the chi-square across a range of values for m_0 and $\mathbf{\Gamma}$. We use the same chi-square equation, and use around 300 values of m_0 and $\mathbf{\Gamma}$ each, from 89-91 GeV and 5-8 respectively.

We can then visualize the different fits using a contour plot of $\Delta\chi^2 = \chi^2 - \chi^2_{min}$, where χ^2_{min} is the minimum possible chi-square value, or the chi-square value of the best fit set of parameters. Shown in Figure 2 is a contour plot of the different $\Delta\chi^2$ values, with the x-axis marking the different values of m_0, and the y-axis marking different values of **Γ**. The color at a given point represents the value of $\Delta\chi^2$ for the associated parameters, with the exact $\Delta\chi^2$ shown in the color bar. Qualitatively, the color at a given point symbolizes how well the Breit-Wigner distribution with the two parameters at that point fit the data, with darker regions representing better fit.

We also outline the $\Delta\chi^2$ corresponding to 1σ and 3σ significance using known properties of multi-dimensional Gaussian distributions. For two fit parameters, 1σ significance is marked by a chi-square increase of 2.30 ($\Delta\chi^2 = 2.30$), and 3σ significance is marked by $\Delta\chi^2 = 9.21$. Labelled in Fig 2 is the contour outlines of 1σ and 3σ significance. The resulting Fig 2 gives us a visualization for the goodness of fit, showing how changing different parameters brings us closer or further from the best fit.

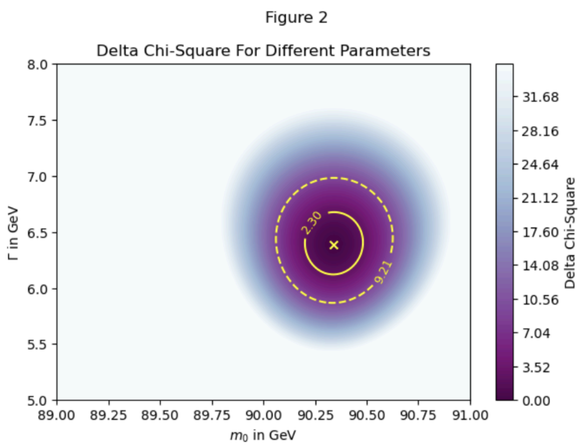

Figure 2: Delta chi-square across the parameter space. Marked in yellow are the 1σ and 3σ contours, which have a delta chi-square of 2.30 and 9.21 respectively

Discussion and Future Work

Summary of Findings

From the ATLAS lepton measurements, we find the rest mass of Z^0 to be **90.3 ±0.1 GeV**. Using 12 data points and 2 parameters to fit the raw data to the Breit-Wigner distribution, we have a total of 10 degrees of freedom, which gives us a chi-square value of 10, a reduced chi-square value of 1, and a p-value of 0.44. Our p-value of 0.44 indicates good agreement between the data and model.

Assumptions/Simplifications and Further Steps

One major simplification in this analysis comes from the histogram binning. The energy and momentum measurements from ATLAS are of very high resolution, making our reconstructed masses relatively precise. However, we only take our bins every 0.5 GeV, losing a lot of the accuracy. A simple way to increase the accuracy would be to increase the binning of the histogram to reflect a higher resolution of masses.

Secondly, our errors are highly simplified, as our error bars only account for the random Poisson errors, without any systematic uncertainties. These systematic uncertainties can come from the detector, the proton beams, and environmental conditions which may affect measurements. The fit does not account for energy resolutions of the ATLAS detector, and we assumed no error on the measurements. In particular, the width parameter of the fit function comes with experimental uncertainties that all real detectors are subject to, which we did not account for – in experimental settings, the width of the function is wider than in theory. In future analyses, we can find and account for errors in the detector measurements as well as propagate them to find accurate errors for the reconstructed masses.

Comparison to literature value

Our calculated rest mass is 90.3±0.1 GeV. The standard literature value for the mass of Z^0 is 91.1880±0.002 GeV. Unfortunately, these two values are not within error bars of each other. While our findings fit theoretical models, our values are far off from other measurements. Taking the above outlined steps to further the accuracy of our calculations could help minimize this discrepancy.