Lab 3: ATLAS Data Analysis

Isaac Sherwood | May 6 2025

1 Introduction

The A Toroidal LHC Apparatus (ATLAS) is an experiment at CERN in Geneva. ATLAS operates by colliding high energy protons together, and the resulting byproducts can be studied. One notable byproduct of the proton-proton (pp) interactions is the Z^0 boson, the carrier of the weak force. The Z^0 boson is unstable and decays. Approximately 10% of the time, a Z^0 boson decays into a pair of oppositely charged leptons in the process $Z^2 \to \ell\bar{\ell}$. This pair of leptons consists of an oppositely charged particle/antiparticle pair, namely an electron and an anti-electron, a muon and an anti-muon, or a tau and an anti-tau.

Because energy and mass are conserved, the total energy of the lepton pair must sum to at least the mass of the Z^0 boson from which they decayed. Using this fact, the mass of the Z^0 boson can be measured. A series of measurements of the Z^0 boson have been taken and are analyzed here. The masses are calculated from energies and momenta, and a Breit-Wigner fit is applied to the corresponding distribution. The Z^0 proton mass is extracted from this fit, and the data is statistically analyzed via a chi-squared test and a two dimensional parameter scan.

2 Invariant Mass Distribution and Fit

The ATLAS experiment is capable of measuring four fundamental properties of the the particles resulting from the decay of the Z^0 boson. The first of these are the total energy of the particles, E, and the momentum in the transverse direction of each particle, p_{T1} and p_{T2}. The pseudorapidities, η_1 and η_2, which describe the angle the particles make relative to the beamline, are also measured. When $\eta \to \infty$ the particle continues straight along the beamline, and $\eta = 0$ for a $90°$ angle. The final property measured is the the azimuthal angle of each particle, ϕ_1 and ϕ_2. These data were recorded for 5,000 instances of Z^0 boson decay in the LHC. The four-momentum of each particle, (E, p_x, p_y, p_z) can be calculated using this data according to:

$$p_y = p_T \cos\phi$$

$$p_y = p_T \sin\phi$$

$$p_z = p_T \sinh\eta$$

For each of the two particles in each measurement, the four-momentum was calculated. The total momentum p_{tot} can thus be calculated by taking the sum of the four-momenta of the two particles:

$$p_{tot} = p_1 + p_2$$

From this total momentum, the mass of the original Z^0 boson can be calculated from:

$$M = \sqrt{E^2 - (p_x^2 + p_y^2 + p_z^2)}$$

For each of the 5,000 lepton pairs, the Z^0 mass was calculated. These calculations were undertaken using the numpy Python package, which uses vectorized operations written in a mixture of Python and C to efficiently perform calculations on large arrays of numbers. The average mass was found to be 84.4 GeV/c^2, with a standard deviation of 29.2 GeV/c^2. 73.9% of the calculated mass values were found to be between 80 GeV/c^2 and 100 GeV/c^2. Fig. 1 shows a histogram of the masses in this range, split into 41 equal bins. The uncertainty in each data point in the histogram was calculated under the simplifying assumption that the data follows a Poisson distribution, giving uncertainty in the number of items in each bin as:

$$\sigma = \sqrt{N}$$

Here, N is the number of occurrences in each bin. These uncertainties are shown by the error bars on the plot.

In scattering theory it is known that the the distribution of decays at a reconstructed mass m follows the Breit-Wigner distribution. The distribution $\mathcal{D}$ is given by:

$$\mathcal{D}(m) = \frac{1}{2\pi}\frac{\Gamma}{(m - m_0)^2 + \left(\frac{\Gamma}{2}\right)^2}$$

Here m_0 is the true rest mass of of the Z^0 boson, and Γ is a width parameter with units of mass.

The distribution of calculated masses was fit with the Breit-Wigner distribution. The overall normalization was fixed to half the number of data points in the set. That is, the data was fit to $2500 \times \mathcal{D}$. The fit was applied only to the bins on the previous histogram with centers between 87 GeV/c^2 and 93 GeV/c^2. This is the region in which 47.6% of the masses occur.

Fitting was carried out using the optimize.curve_fit function in the scipy Python library. The optimize.curve_fit functions by using non-linear least squares to fit data to a function. The function returns the optimal values of m_0 and Γ, as well as the covariance matrix for the fit. Uncertainty in the mass and in Γ were extracted from the covariance matrix. The optimal value of mass was found to be $m_0 = 90.4 \pm 0.1$ GeV/c^2, and the optimal Γ found to be 6.4 ± 0.2 GeV/c^2. Fig. 1 also shows the fitted distribution, as well as the residuals. Residuals were calculated as the difference between the measured values of each bin center, and the theoretical value given by the fit. Uncertainties in the residuals were found by propagating the errors, giving identical uncertainties to those of the histogram of data.

A chi-squared test was carried out to measure how accurately the fit describes the data. The chi-squared value was determined according to:

$$\chi^2 = \sum_i \left(\frac{x_{o_i} - x_{E_i}}{\sigma_i} \right)^2$$

In the above equation x_o and x_E are the experimentally observed values and expected values from the fit respectively, and the summation is taken over all i many points in the dataset being fit. The chi-squared value of the data was found to be $\chi^2 = 10.0$. The reduced chi-squared of a fit with v degrees of freedom can be calculated from:

$$\chi^2_{red} = \frac{\chi^2}{v}$$

The number of degrees of freedom is defined as:

$$v = n - f$$

where n is the number of points in the dataset, and f is the number of fit parameters. The data has $v = 10$ degrees of freedom, corresponding to the two fit parameters, m_0 and Γ, subtracted from the twelve data points in the fit region. This gives a reduced chi-squared value of $\chi^2_{red} = 1.0$. The p value was alsoe calculated, using the stats.chi2 functionality of the scipy Python package. The probability density function of the chi-square distribution for values of x greater than zero is given as:

$$P(x) = \frac{x^{v/2-1} e^{-x/2}}{2^{v/2}\Gamma(v/2)}$$

Here v is the number of degrees of freedom, and Γ is the gamma function. The p value is found by integrating the probability density function from the calculated χ^2 value to infinity:

$$p = \int_{\chi^2}^{\infty} P(x)dx$$

The stats.chi2 functionality of scipy computes this integral numerically to determine the p value with a high level of accuracy. The p value for this fit was calculated to be $p = 0.4$. This indicates a 40% probability that a χ^2 larger of at least 10.0 would occur by random chance, assuming accurate uncertainty and the null hypothesis to be true. A p value of 0.4 represents a reasonable probability of similar results occurring, and is well within the commonly accepted range of acceptable values. It thus indicates a reasonable agreement between the data and its corresponding uncertainty, and this model.

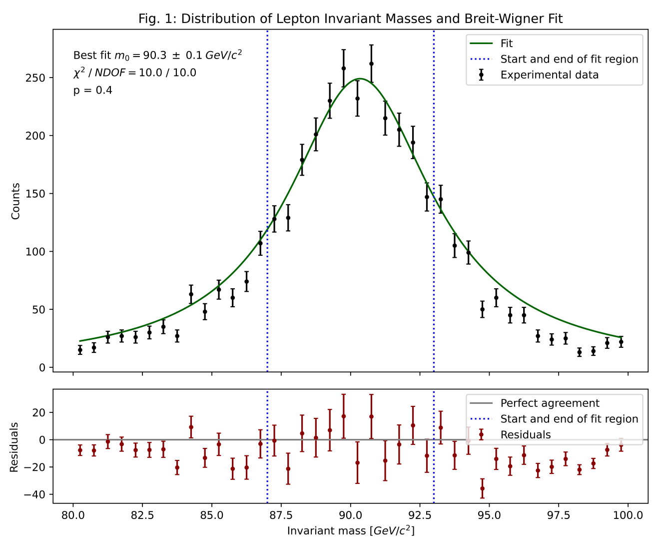

Figure 1: Distribution of lepton invariant masses, Breit-Wigner fit, and residuals

3 Two-Dimensional Parameter Scan

In order to evaluate the relationship between the m_0 and Γ parameters in the fit, and the χ^2, a two-dimensional parameter scan was conducted. For the scan, 300 m_0 values between 89 and 91 GeV/c^2, and 300 Γ values between 5 and 8 GeV/c^2 were tested. At each possible pair of m_0 and Γ values, the χ^2 of the resulting fit was calculated and recorded. Thus, a total of $300 \times 300 = 90,000$ combinations of fit parameters were evaluated. This was again accomplished using the capabilities of the numpy Python package.

The minimum possible χ^2 value was determined to be 10.0, located at $m_0 = 90.3$ GeV/c^2 and $\Gamma = 6.4$ GeV/c^2. These values are consistent with the optimal values found by scipy, given their uncertainties. Fig. 2 shows the $\Delta\chi^2$ plotted as a function of m_0 and Γ. The $\Delta\chi^2$ was calculated by taking the difference between the χ^2 value at each location and the minimum χ^2 of the entire dataset, 10.0. The data was clipped at 35, such that all parameter combinations with a $\Delta\chi^2 \geq 35$ are plotted as $\Delta\chi^2 = 35$. This allows for greater readability of the plot by eliminating values of $\Delta\chi^2$ that are arbitrarily large.

Additionally plotted are the 1 σ and 3 σ levels. The 1 σ region is the set of points at which the $\Delta\chi^2 = 2.3$, and the 3 σ region is the set of points at which the $\Delta\chi^2 = 9.21$. The region bounded by the 1 σ level is the region in which approximately 68% of the best fit values of parameters would fall were the experiment repeated. The 3σ region is where approximately 99% of the best fit values would fall. The optimal values of m_0 and Γ are well within the 1 σ region.

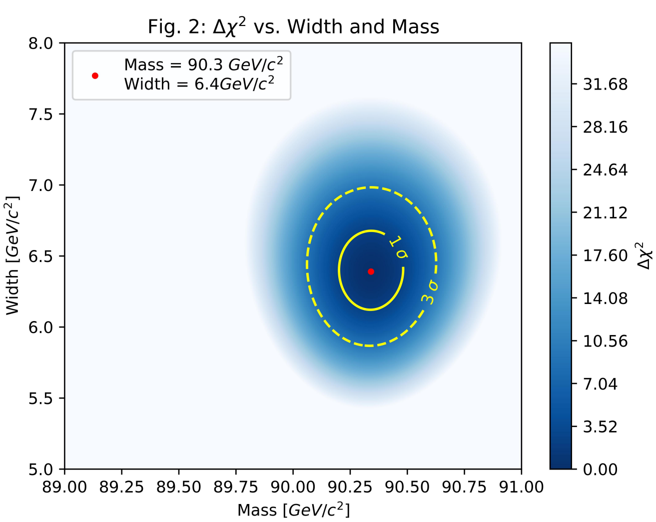

Figure 2: Two-dimensional parameter scan of possible m_0 and Γ values

4 Discussion and Future Work

The analysis carried out here fitted the experimental results of 5,000 measurements of Z^0 boson decays made by ATLAS to a Breit-Wigner distribution. The fit gave an optimal true mass of $m_0 = 90.3 \pm 0.1$ GeV/c^2, with a p value of $p = 0.4$. The currently accepted value of the mass of the Z^0 boson is 91.1880 ± 0.0020 GeV/c^2. The consistency of these two values can be examined by considering the ratio of their difference to the uncertainty in their difference, z:

$$z = \frac{|m_2 - m_1|}{\sqrt{\sigma_{m_1}^2 + \sigma_{m_2}^2}}$$

where m_1, m_2 and $\sigma_{m_1}, \sigma_{m_2}$ are the mass values and their uncertainties respectively. This ratio is calculated as 8.9. This indicates a discrepancy of > 8 σ, and this the values are in significant disagreement.

This disagreement could stem from a number of simplifications and assumptions made throughout the experiment and data analysis. Most notable, no uncertainties were recorded in the measurements taken by ATLAS. Uncertainties in the ATLAS measurements could be determined in the future, thus giving a more accurate description of the Z^0 mass. Additionally, the assumption was made that the data followed a Poisson distribution, giving a simple expression for uncertainties in the number of occurrences in each bin. This may not necessarily be true, and further research could yield more accurate descriptions of the distribution and related uncertainties.

Additionally, the fit was applied only to region of the data containing 50% of points nearest to the mean. This likely impacted the fit and the uncertainties in the parameters. Further research could apply the fit to a greater range of data, possible giving more accurate results.

The assumption was also made that the distribution of masses would follow a Breit-Wigner distribution. While this is suggested by scattering theory, further analysis of the experimental results could be conducted to determine if other distributions may more accurately describe the data.

While a value for the mass of the Z^0 boson has been determined, there is still significant disagreement between it and other accepted values. Further research should be conducted to further examine the data and expand our understanding of fundamental particle physics.