Lab 3: ATLAS Data Analysis

Erik Siegwarth   Physics 265

Introduction: When beams of high energy protons collide with each other an array of new particles are produced, where they then can interact and decay. Out of all the particles formed from the proton-proton interaction, this report will be focusing on the Z^o boson. We will be specifically using the ATLAS experiment to estimate the mass of the Z^o boson. The Z^o is unstable and will decay; around 10% of the time it decays into a combination of two charged leptons. These can be an electron and a positron, muon and an anti-muon, or a tau and an anti-tau pair. Each of these pairs are two particles of opposite charge, and their combined energy must be at least the mass of the original Z^o. We will then observe a large quantity of double lepton events and should then be able to estimate the mass. For reference the particle data group has published the mass as 91.2 GeV.

Part 1, The Invariant Mass Distribution and Its Fit: When observing the particles we will focus on four properties measured by the ATLAS detector in order to find the distribution and fit for the mass off the Z^o. These properties are the total energy (E), the momentum in the transverse direction (p_T), the angle the particle makes with respect to the beam line, or pseudorapidity (η), and finally the azimuthal angle (ϕ) about the beam. We can use these quantities to define the four momentum (p) of the particle as p = $(E,\ p_x,\ p_y,\ p_z)$ We can define $p_x,\ p_y,\ p_z$ respectively as: $p_x = \ p_T cos(\phi)$ $p_y = p_T\, sin(\phi)$ $p_z = p_T sin(\eta)$ Using the difference between the three momentum and the energy we can calculate the mass (M) with this equation: M = $\sqrt{E^2 - (p_x^{\ 2} + p_y^{\ 2} + p_z^{\ 2})}$ We will use data taken from the 2020 ATLAS open data set that has been filtered to only include two lepton events in their final states. For each particle in the lepton pair the data includes each of the four quantities needed. In JupyterLab we import the data and then calculate the x, y, and z momentum for each particle. We then sum each pair to get the total momentum of the system in each direction for each pair. We will then square our p_x, p_y, and p_z totals for each pair and subtract it from the squared energy. After square rooting this quantity for each pair we have the mass of the original Z^o particle. We now will have a wide range of masses for all the pairs measured. To visualize the data we will create a histogram with the x axis of bins representing varying masses and the y axis representing the frequency. Each bin will have an error (σ), which can be seen in orange on Figure 1. And it is the square root of the number of events in that respective bin, or $\sqrt{N}$.

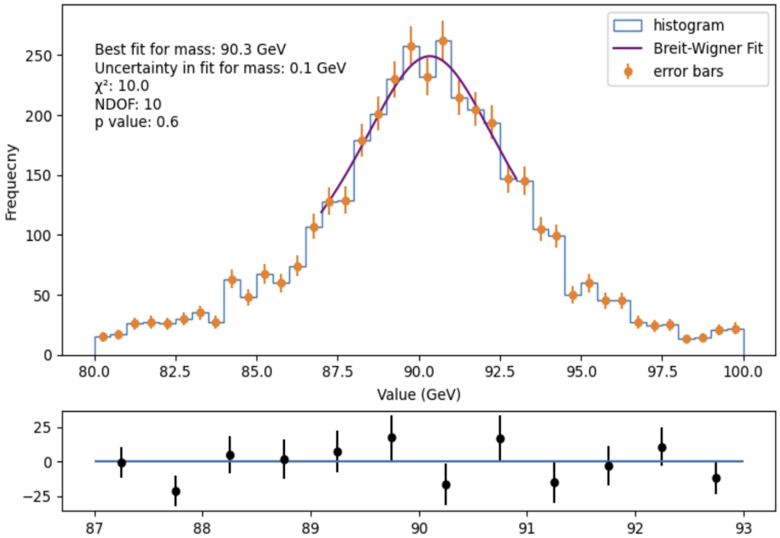


Figure 1: Mass of Original Particle Histogram with Curve Fit

We will also use the Breit-Wigner peak, which can be used to estimate, given a specific original mass (m_0) and a "width" parameter (Γ), the number of decays (D) that each reconstructed mass (m) will have. We will use python to create the function that returns the number of decays (D), based on the reconstructed mass (m), the true mass (m_0), and the width parameter (Γ):

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

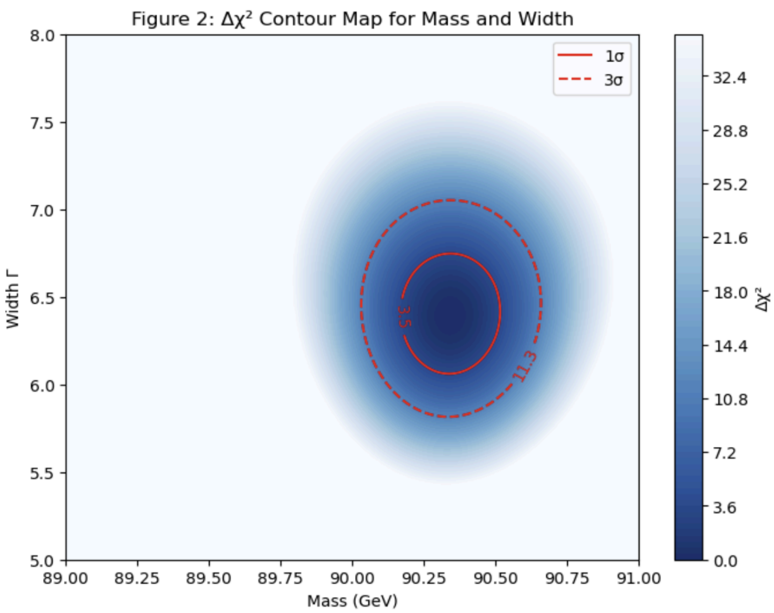
After normalizing our data set, by multiplying by half the number of data points, we can display it over top of our histogram to observe how well they agree with one another. We will fit our data for $87 \leq m \leq 91$, and can be seen as the purple line in Figure 1. Underneath our histogram is a residual plot of our data minus or fitted values, with the same error bars given by $\sqrt{N}$. We can then utilize scipy's curve fit to approximate the value of m_0 and Γ that results in a Breit-Wigner curve most suited to our data set. Curve fit also creates a covariance matrix that can be used to find the error in our estimate for each.

After executing we are given a best estimates, m_0 : 90.3 GeV $\quad \sigma_{m_o}$: .1 GeV

To check the agreement of our fit and data we can calculate the chi square total. We do this by squaring the residual of each data point and the theory and dividing by the error squared. This gives us a chi square value: χ^2= 10.0. We observed 12 data points (N) for our chi square test, and in our original equation we were testing for 2 fitting parameters (V). This means we have 10 degrees of freedom, as NDOF can be found by subtracting N-V. We can then use scipy's p value function. Which in simple terms tests for what the percent chance any difference in data and theory occurred because of random chance alone. A p value test, with a chi square of 10 and 10 degrees of freedom gives us NDOF (v) =10 $\quad P(\chi^2,v)$ = .6 $\quad$ Meaning there is a 60% chance any difference between our fit and data was because of random chance. A p value of .6 is enough to suggest agreement between the data and our fit.

Part 2, 2d Parameter Scan:

Since both the mass of the Z^o and the width parameter Γ impact the number of decays and our fit, we need to determine both parameters as dependent on one another. We can visualize this by making a 2d delta chi square scan of the mass-width parameter space. In simple terms our x axis is our different potential masses of the Z^o, the y axis is our different possible widths, and the z axis (in color) is the difference between the chi square value created by that combination of x and y and the minimum chi square value we found out of all our combinations. We can observe our graph to see how changing the mass and width parameters impact the goodness of fit of our model. To make the graph we will create a grid of the chi square values minus the minimum chi square value. We do this by looping our earlier Breit-Wigner model for each combination of width and mass and then testing those fitted values against the data we first collected. This will allow us to find the point that causes the delta chi square to be zero, which should be our best combination. We will also discard any values on the z axis greater than 35, as those are not going to be the best fit values. We will also include two lines to show the error in our chi square calculations. As can be seen the solid line is 1σ (3.5) and the dotted line is 3σ (11.3). We arrived at these values based on the fact we used 2 fitting parameters for our original fit. So 3.5 and 11.3 are standard for 1σ and 3σ.



Conclusion:

Based on our data we can estimate for the data we analyzed that the Breit Wigner fit agrees, and the estimated mass of the Z^o is 90.3 GeV with an error of .1. This is different then the widely accepted value of 91.2, as our value is 9 error bars away. Throughout our work we made a number of assumptions that simplified our calculations but could have resulted in the difference between the accepted value of Z^o and our values. We first assumed that our histogram could be treated as a Poisson distribution, or the error in each bin for our histogram was equal to $\sqrt{N}$. This is not entirely accurate and . We also assumed there was no error in our Breit wigner fit, so that the error in our residuals and for our chi square fit was the same $\sqrt{N}$. We also did not factor in any systematic uncertainties and our data was assumed to be completely accurate. We also normalized our curve's height, by multiplying by 2500, instead of letting that value change based on the data. Going forward to better improve our accuracy we would need to measure the uncertainty in the measurements made by the ATLAS machine, as well factoring in the normalizing parameter as a free parameter.