Exploring Fitting Methods for Calculating the Cross Section of the Drell-Yan Process

by
Evan Chang

Department of Physics
University of Michigan
April 2019

Advisor: Wolfgang Lorenzon

Table of Contents

Abstract

Exploring Fitting Methods for Calculating the Cross Section
of the Drell-Yan Process

by Evan Chang

Investigations into fitting methods for calculating the cross section ratio of the Drell-Yan process will be described. Our data was gathered from drift chamber tracking data from the hadron collider at Fermilab as part of the E906/SeaQuest experiment. The beam consisted of 120 GeV protons extracted from the Fermilab Main Injector and collided with a fixed target of deuterium and hydrogen. Our analysis involves the calculation of the Drell-Yan process's cross section by plotting the calculated cross section ratio against beam intensity variables. A final cross section value is determined by extrapolating and finding the y-intercept of the regression formed by the cross section ratio and the intensity. By investigating several methods for fitting and thus several methods for calculating the cross section, we can determine the best way to calculate this value moving forward in the experiment.

Acknowledgements

Firstly, I would like to thank Prof. Wolfgang Lorenzon for giving me the opportunity to work on the SeaQuest experiment for almost an entire year at this point, and offering guidance and assistance throughout. I would also like to thank Daniel Morton and Marshall Scott for their consistent support with any issues I ran into. Additionally, I want to thank my family, without whom none of this would have been possible. Lastly, I want to thank Tom Henderson for inspiring me to pursue a higher education in physics.

1. Introduction

1-1 Quantum Field Theory

Despite the great leaps we have made over the past 50 years in learning about fundamental particles, we are still far from understanding everything there is to know about the building blocks of our universe. For example, until 1964 the proton was believed to be fundamental, and it took until 1968 for Stanford's linear accelerator to provide evidence for the quark model.[1] Nevertheless, it was found that our current model of fundamental processes was largely incomplete.

Where classical mechanics ends, quantum mechanics begins. Quantum theory describes the physical processes of nature at the smallest scale of particles.[2] This includes fundamental (as far as we know) particles such as quarks. One identifying characteristic of quarks is their flavor. Quarks can be one of six flavors: up, down, top, bottom, strange, and charm. The up and down quarks have the lowest mass and the other quarks decay into up and down flavors.[3] Up and down quarks are thus the most common and most stable. While quarks have never been discovered as free particles, they make up all hadrons - most commonly protons and neutrons. Thus, they are the components of every atomic nucleus.

For every quark flavor there exists an antiquark. Each antiquark has the same mass as its respective quark but has different quantum numbers and an opposite electric charge. Virtual quark-antiquark pairs can be found in hadrons, and thus can be detected during hadron collisions. These pairs can annihilate to form a gluon, and similarly can be created by a gluon splitting. The constant annihilation and formation is referred to colloquially as "the sea."[4]

1-2 Drell-Yan Process

Despite knowing of quarks since 1968, there is still a large amount to learn about them. For example, the proton quark sea was believed to be created through quantum chromodynamics processes, primarily gluon splitting, and thus not favoring either up or down quarks.[5] Nevertheless, by examining certain interactions, it was determined that this is most likely not the case. The Drell-Yan process is one such interaction.

The Drell-Yan process occurs as a result of hadron-hadron scattering in which a quark and antiquark of two respective hadrons annihilate. This forms a virtual photon that decays into an oppositely-charged lepton pair.

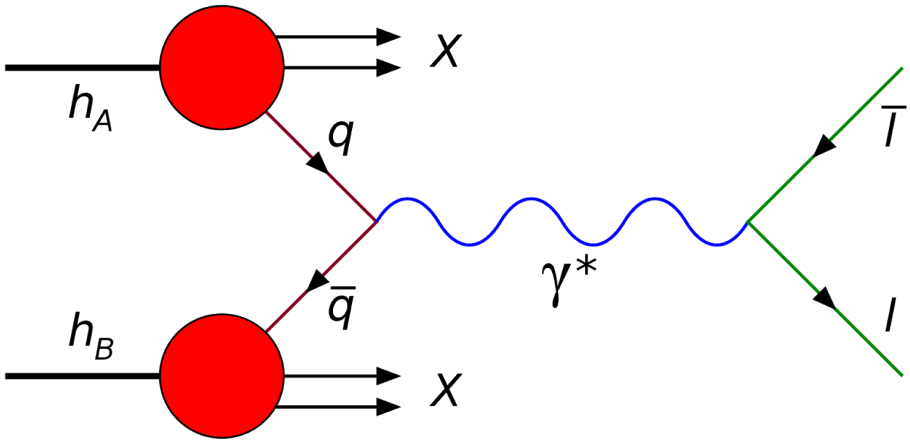

Figure 1: Feynman diagram of the Drell-Yan process. The red h_A and h_B denote the two colliding hadrons, the q's denote the quark and antiquark that form the virtual photon that decays into the lepton pair.

The cross section of this process is an important measurement, as it has implications on the flavor asymmetry in protons. For example, it was originally believed that there were an equal number of anti-up quarks as anti-down quarks in the nucleon sea. Nevertheless, deep inelastic scattering experiments at CERN demonstrated that there are actually greater numbers of anti-down quarks than anti-up quarks in protons.[6] This finding has been found to hold true at low Bjorken-*x*, while at a Bjorken-*x* of approximately 0.3 and above, the opposite is true. The results of the experiments measuring this asymmetry are displayed in Figure 2.

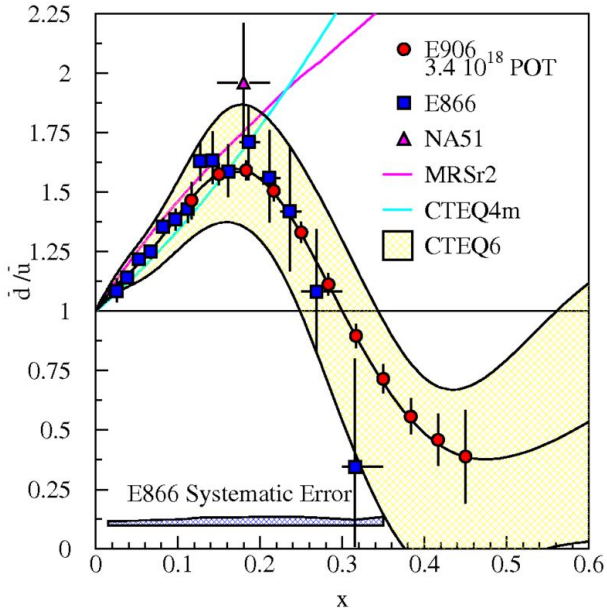

Figure 2: Ratio of anti-down to anti-up quarks vs Bjorken-*x* for several experiments. The E906, E866, and NA51 data are all displayed imposed over the CTEQ6 parton distribution function, alongside the MRSr2 and CTEQ4m functions. The E906 data extends into higher Bjorken-*x* than the other experiments.

Bjorken-*x* refers to the Bjorken scaling of elementary particles. Experimentally observed strongly interacting particles have several properties that scale based on factors such as the absolute energy of an experiment, scattering angle, and the ratio of energy to a momentum transfer.[7] In Figure 2, it also clearly affects the cross section of the Drell-Yan process.

The cross section of this process is very closely tied to this asymmetry. In fact, the cross section ratio of a proton beam on a deuterium target to the cross section of a proton beam on a proton target is given by

$$\frac{\sigma^{pd}}{2\sigma^{pp}} = \frac{1}{2}\left[1 + \frac{\bar{d}(x)}{\bar{u}(x)}\right],$$

where $\bar{d}$ and $\bar{u}$ are the anti-down and anti-up quark distributions, respectively, as a function of Bjorken-*x*.[5] This equation clearly demonstrates the cross section ratio's direct relation to the antiquark asymmetry, and thus its importance to the quantum model and the E906/SeaQuest experiment.

Determining the cross section ratio of the Drell-Yan process can be done in a variety of ways. One particular method involves plotting the cross section calculation against measured variables such as chamber occupancy or beam intensity. Fitting these plots to a function and thus determining the cross section intercept as the variable goes to 0 is one method for calculating the cross section, and the method for calculating the cross section that this paper will focus on.

The reason that this method is being used for calculating the cross section is due to slight systematic errors inherent in our data. For example, the cross section of the Drell-Yan process, or any process for that matter, should not be a function of beam luminosity. Nevertheless, mechanical imperfections in detectors as well as background noise at higher intensities results in some dependence on the luminosity. Therefore, what the cross section ratio that is calculated at any given time is instead the cross section ratio with some minor added effects. This presents the need for a fitted calculation for the Drell-Yan cross section ratio. By fitting, we are able to extrapolate to a 0 point with no effects from luminosity. Obviously, it is physically impossible to gather any data at that point, yet a fit can give us an expected value for the cross section.

The difficulty in using this methodology lies in the choices that must be made when creating and fitting plots. For example, there are dozens of variables to to plot the cross section against. Additionally, it is important to consider the number of points that will be used, the fitting method, and the number of plots created. The goal of this thesis is to provide a recommendation for the fitting procedure moving forward. By investigating several different methods for fitting the cross section ratio, a more accurate calculation of the cross section ratio can be calculated - one that eliminates any erroneous beam luminosity dependence.

2. Experimental Setup and Motivation

2-1 Fermilab Spectrometer

As mentioned previously, the data used for this experiment came from the Fermilab Hadron Collider. The data comes from the spectrometer portion of the beam's path, and is shown in Figure 3.

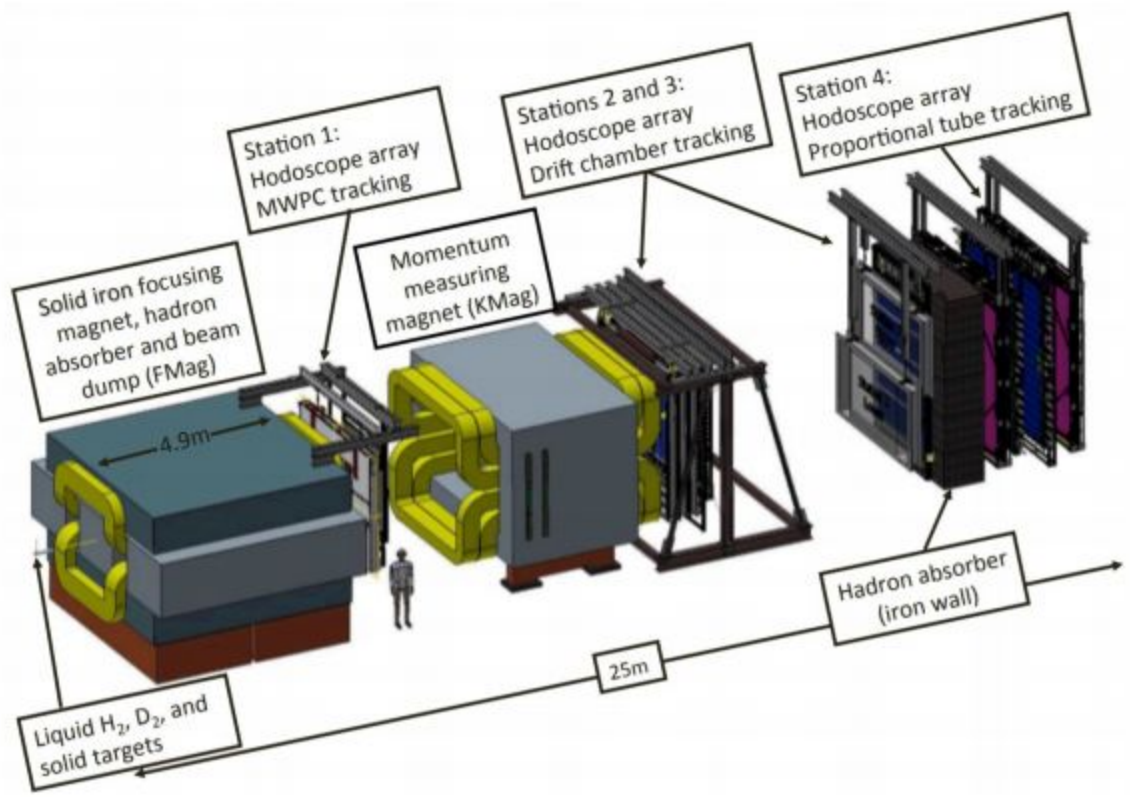

Figure 3: A schematic of the SeaQuest spectrometer with each of the individual stations labeled. The beam direction in this image would be going from left to right. A human is drawn for reference.

The beam used for the SeaQuest experiment is a 120 GeV beam of protons from Fermilab's main injector beamline. Previously, high energy collisions such as this were done using the Tevatron, although that has been closed since September of 2011. Fortunately, the main injector beamline provides greater statistics. Specifically, the Fermilab E866 experiment was another fixed target experiment that also sought to probe the asymmetry of antidown and anti-up quarks in the nucleon sea. In many ways, that experiment is a predecessor of the E906 experiment.

As the beam travels through the spectrometer, the first thing that it interacts with in the spectrometer is the target, which is fixed for this experiment. The targets are constructed of hydrogen, deuterium, as well as nuclear targets for collecting data. There is also a target simply

consisting of an empty liquid flask, and thus a "no target" position. This acts as a control detector to help eliminate background. In our calculations, we subtract the empty data from that of the hydrogen and deuterium.

Stations 1-4 all provide data on the particles that go through them. By using a combination of hodoscopes, drift chambers, and proportional tubes, we are able to determine the track of Drell-Yan dimuons through our spectrometer. The hodoscopes act as triggers, the drift chambers provide location information, and the proportional tubes allow us to determine the nature of the particles, their charge, and some more positional information. They provide rougher positional data than the drift chambers, but finer data than the hodoscopes.

It is important to note that most of the particles generated by proton-proton collisions are not the dimuons that we want to observe for this experiment.

The intensity of the beam is measured by a beam intensity monitor. It is able to determine the intensity in real-time and record it. This is extremely important for our measurements, as the intensity has a significant impact on the Drell-Yan cross section ratio.

2-2 Data

The spectrometer provides a lot of data on the dimuon tracking, the beam intensity, the drift chamber intensity, etc.. They provide position and hit information, along with timing. This timing is used for consistency in tracking particles and does not measure velocity. However, for the scope of this thesis, only a few variables are relevant, as we are primarily concerned with the calculation of the cross section of the Drell-Yan process. The cross section can be calculated explicitly by performing

$$\frac{\sigma_{pd}}{2\sigma_{pp}} = \left\{ \left(\left(\frac{W_D}{P_D} - \frac{W_E}{P_E} \right) \cdot \frac{1}{G_D} \right) \Big/ \left(\left(\frac{W_H}{P_H} - \frac{W_E}{P_E} \right) \cdot \frac{1}{G_H} \right) + F_D - 1 \right\} \cdot \frac{1}{2F_D}.$$

In this calculation, W is the corrected dimuon yields, P is the number of beam protons, and G is the atom density, with their subscripts denoting either hydrogen, deuterium, or the empty target. F_D is the calculated purity of the deuterium based on the number of deuterons and protons. This can be calculated for each data point, which allows us to fit our cross section against different variables.

The corrected dimuon yields is a highly critical value to our calculations, but it is also very difficult to achieve. Even with tracking, most of the dimuons detected by the spectrometer are not from the Drell-Yan process. In fact, Drell-Yan dimuons are outnumbered by a factor of somewhere between 10^7 and 10^8. In order to isolate the data we want from the background, we used physics and machine learning to disregard most of the data points. All of the data used in this experiment has had these cuts performed on them.

As mentioned in Section 1-2, the Drell-Yan process cross section does not depend on the intensity of the beam. Nevertheless, the cross-section calculated from Fermilab's data does indeed depend on luminosity. This is why the work in this thesis is so important. Achieving 0 luminosity is impossible, as there would be no Drell-Yan events. Therefore there is inherent beam effects in all of our data. The solution is to plot the cross section against the beam luminosity and extrapolate to 0.

There are several variables that mimic the luminosity impact. Namely the trigger intensity, the chamber intensity, and the total occupancy. These three variables will all be used for fitting against the Drell-Yan cross section ratio. In fact, over the past summer, my work has also involved investigating exactly how and what these beam variables impact, beyond just the cross section ratio, but that is beyond the scope of this thesis. The main rate dependent effect is an increase in combinatoric background noise, or when two events occur simultaneously, appearing as one. This would greatly impact the efficiency of our detectors, and would result in a lot of data being lost. This thesis aims to fix that by eliminating the background effects.

3. Analysis

3-1 Fitting by Variables

The first few fitting methods tested involved fitting the cross section against different physical variables. The variables we chose for this method were trigger intensity, chamber intensity, and total drift chamber occupancy. These variables were chosen due to their effect on the cross section ratio, as mentioned in section 2-2. The E906/SeaQuest collaboration had previously been using trigger intensity as the variable for the preliminary fits and tests of the rate dependence method for calculating the cross section ratio.

For each fit, we divided our data into eight bins based on the Bjorken-*x* of the target parton momentum, or x_T. This was done due to the fact that x_T is correlated to mass, and x_T is known to have some effect on the intensity distribution. Each x_T bin is of roughly equal width, but the highest x_T bin had very low data counts and thus gave results with extremely high error bars and clearly erroneous points. Because of this, we did not count it for any of our statistical analyses. All of the statistics generated for this experiment only took into account the first seven x_T bins. The bounds of each bin are shown in Table 1.

x_T Bin Number	Lower Bound (not inclusive)	Upper Bound (inclusive)
0	0.100	0.130
1	0.130	0.160
2	0.160	0.195
3	0.195	0.240
4	0.240	0.290
5	0.290	0.350
6	0.350	0.450
7	0.450	0.580

Table 1: The bounds of the x_T bins used to separate the data points into roughly equal bins. Each plot shown in this thesis is generated using one of these bins.

The first method of fitting for each of the variables was the most simple: linear regression. Each of these regressions was calculating using ROOT's built in χ^2 minimization calculations. Trigger intensity was fit first, as it was the variable that the E906/SeaQuest

collaboration was using for the fitting before this analysis. Figure 4 displays a sample of the results of the linear fitting for trigger intensity.

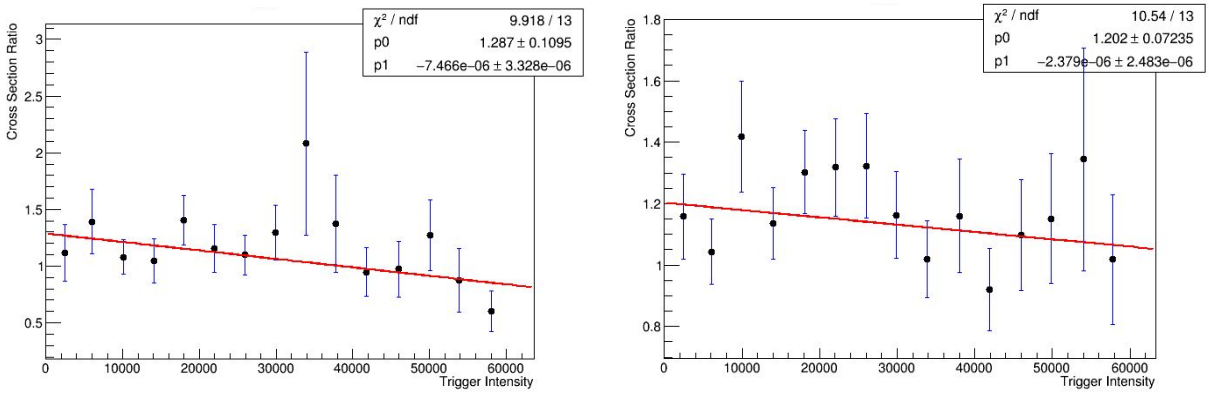

Figure 4: Linear fitting of cross section ratio vs trigger intensity. On the left is events for x_T between 0.100 and 0.130. On the right is events for x_T between 0.130 and 0.160. p0 is the calculated intercept, and thus our cross section ratio value, p1 is the calculated slope of our regression.

From these graphs alone, we began to see some of the effects of beam intensity on the cross section ratio. There is a clear downward slope on each plot. This is present in every single plot generated for the linear fitting against trigger intensity.

The next step was to perform the same fitting except against chamber intensity and total occupancy. The results of these fittings are shown in Figures 5 and 6.

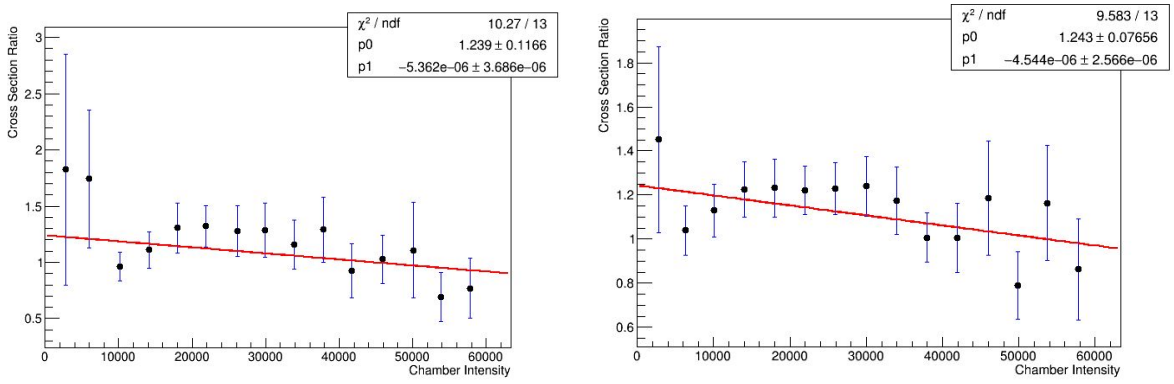

Figure 5: Linear fitting of cross section ratio vs chamber intensity. The left is events for x_T between 0.100 and 0.130. On the right is events for x_T between 0.130 and 0.160.

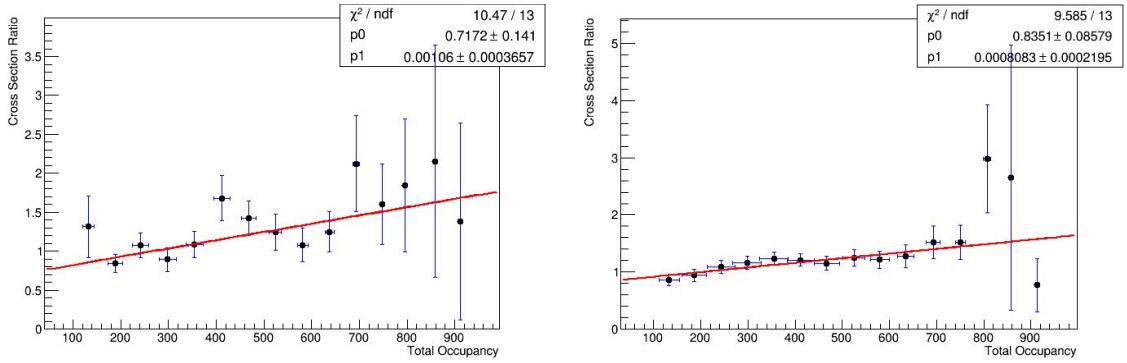

Figure 6: Linear fitting of cross section ratio vs total occupancy. The left is events for x_T between 0.100 and 0.130. On the right is events for x_T between 0.130 and 0.160.

In Figures 5 and 6, there is a strong linear dependence on the beam intensity variable - although in Figure 6, we see how the total occupancy and cross section ratio are directly related, as opposed to the inverse relation in chamber and trigger intensity. Nonetheless, it is still a very strong relation.

In addition to performing a linear fit for each variable, we also performed a quadratic fit. Samples of the results are in Figures 7 through 9.

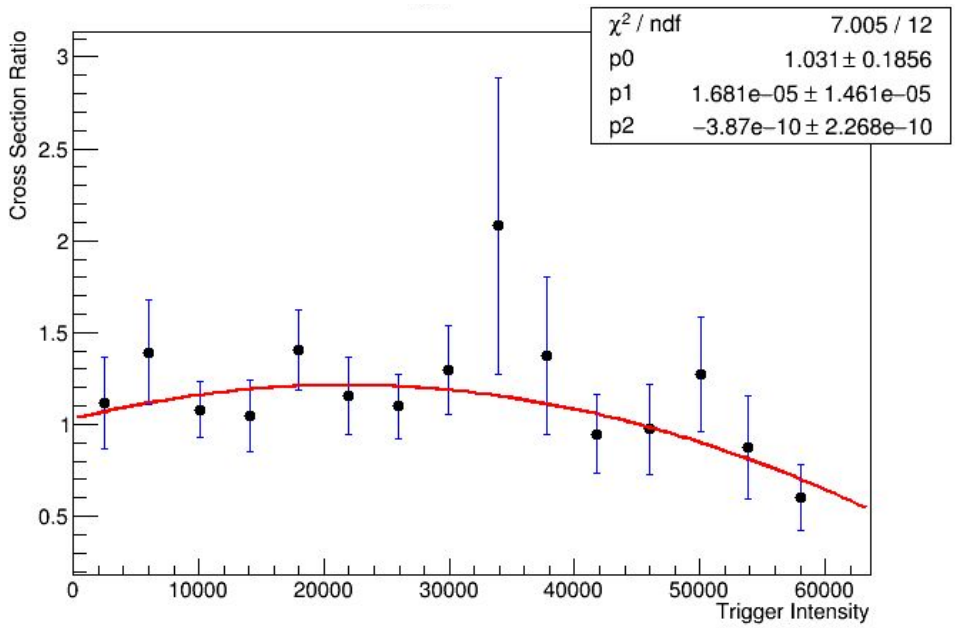

Figure 7: Quadratic fitting of cross section ratio vs trigger intensity. p2 is the coefficient of the quadratic term of the regression, and p1 is still the linear term. The events all have a value of x_T between 0.100 and 0.130.

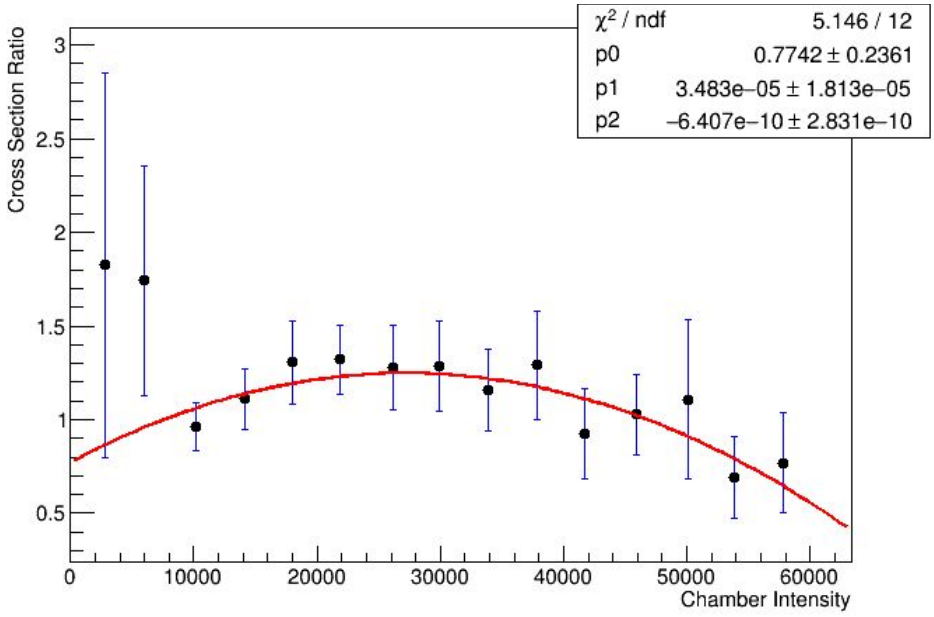

Figure 8: Quadratic fitting of cross section ratio vs trigger intensity. p2 is the coefficient of the quadratic term of the regression, and p1 is still the linear term. The events all have a value of x_T between 0.100 and 0.130.

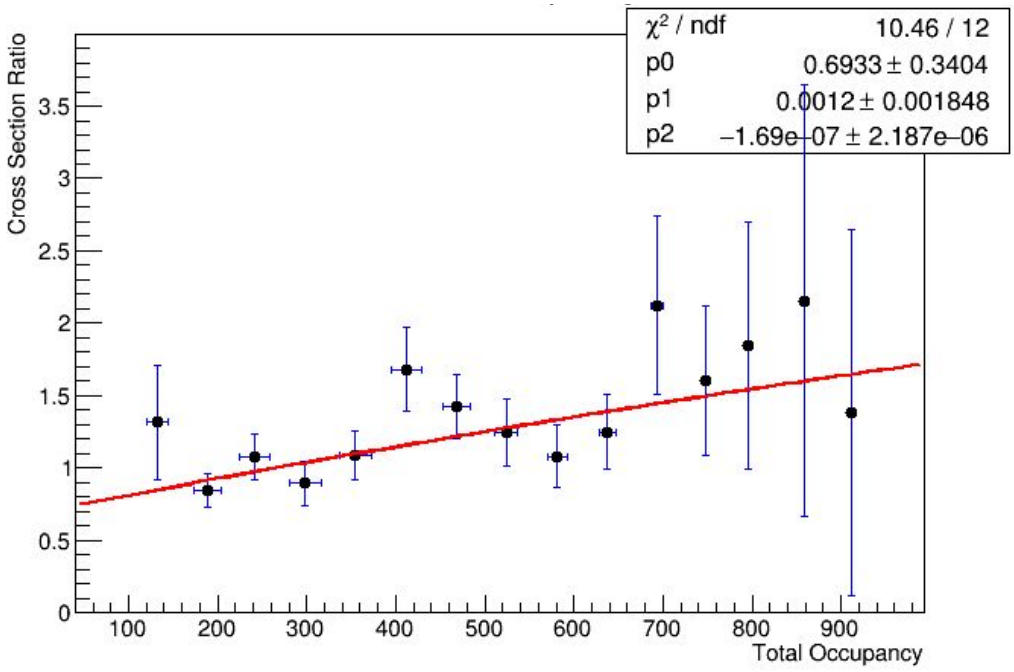

Figure 9: Quadratic fitting of cross section ratio vs trigger intensity. p2 is the coefficient of the quadratic term of the regression, and p1 is still the linear term. The events all have a value of x_T between 0.100 and 0.130.

While we can easily determine that the fitting for each of these variables displays the rate dependence that we had predicted, it was important to determine which variables demonstrate the

most clear relation to the cross section ratio. The results of averaging several statistics across all x_T bins are displayed in Table 2.

Variable	χ^2 of Linear Fitting	AIC of Linear Fitting	BIC of Linear Fitting	Linear -2ln(L)	Quadratic χ^2	Quadratic AIC	Quadratic BIC	Quadratic -2ln(L)
Chamber Intensity	12.139	4.96	6.376	0.960	9.954	6.063	9.187	1.063
Occupancy	7.874	4.008	6.757	1.341	7.024	6.677	9.801	1.677
Trigger Intensity	9.333	4.594	6.010	0.594	7.977	6.509	8.633	0.509

Table 2: Average statistics over all x_T bins for each of the simple fittings for each variable, where -2ln(L) is the log likelihood function and AIC and BIC are -2ln(L) + 2k and ln(n)*k-2ln(L), respectively, where k is the model parameters and n is the number of points.

The results in Table 2 are averaged over all x_T bins, except for the high x_T bin, as mentioned previously. For each of the results in the tables throughout this thesis, the values are going to be averaged over x_T. This is valid because we do not expect x_T to impact the luminosity effects, or to have a significant impact on the Drell-Yan cross section or any of the goodness-of-fit parameters when performing a regression on the plots. The most important columns for deciding upon a variable were the χ^2 and the log likelihood columns, as these columns measured how well the regression fit the data. We found that fitting against trigger intensity produced the lowest χ^2 and log likelihood for both the linear and quadratic fits, which indicated that it was going to be the best variable for analyzing the cross section ratio moving forward.

The AIC (Akaike information criterion) and BIC (Bayesian information criterion) columns for the linear and quadratic fits were to help decide between the linear and quadratic fits. AIC and BIC are both penalized-likelihood criteria, meaning that they measure a goodness-of-fit but "penalize" the score based on a few factors. Using these two numbers is standard for model selection. AIC and BIC are calculated by -2ln(L) + 2k and ln(n)*k-2ln(L), respectively, where k is the model parameters, n is the number of points in the plot, and -2ln(L) is the log likelihood function. Since AIC and BIC are related to the number of parameters, these statistics helped give us insight as to whether or not it was truly worth it to use higher order functions when fitting. Obviously the χ^2 and log likelihood would be lower for the quadratic fittings, but since AIC and BIC were still lower for the linear fittings, we determined that the linear fittings were better.

At this point, it was clear that trigger intensity was the best variable for our calculations. For the rest of our fitting techniques, we made sure to fit the cross section ratio versus chamber

intensity and total occupancy as well to be thorough, but they are not relevant for the sake of this thesis.

3-2 Using Log Likelihood Minimization

For the fittings in section 3-1 and 3-2, we had used a χ^2 minimization to fit our plots. However, it was also important to use a log likelihood minimization to generate a linear regression. The χ^2 value was what we believed would be most important to measure, as it measures a deviation of the observed data from the fit. However, we decided that it would also be important to generate fittings using a log likelihood minimization. This type of fit would allow us to generate a regression that would instead fit the most probable behavior of the data. It would almost certainly increase the χ^2 of our plots but would serve to lower the log likelihood and also our AIC and BIC statistics.

While this idea seems fine in theory, in practice it is slightly more difficult. ROOT, the interpretive C++ framework used for this experiment, makes using χ^2 minimizations very simple, while log likelihood is very difficult to use. Since ROOT is fairly standard for high energy physics research, this is an important consideration for others, not just our work or E906/SeaQuest's work.

Fitting the trigger intensity plots from section 3-1 using log likelihood produces the statistics shown in Table 3. The trigger intensity statistics are the same as in Table 1 but have been placed in this table as well for an easier comparison.

Fitting Method	χ^2	AIC	BIC	-2ln(L)
χ^2	9.333	4.594	6.010	0.594
Log Likelihood	11.410	4.518	5.934	0.518

Table 3: A comparison of χ^2 versus log likelihood regression statistics for a linear fit of the cross section ratio vs trigger intensity. The results are averaged over every x_T bin.

From Table 3, we can see that the statistics changed exactly as we expected, with χ^2 increasing while the rest of the statistics decreased. Since neither χ^2 nor log likelihood changed significantly enough to determine whether either fitting technique is superior, we had to ask ourselves whether using a different and more technically difficult method would even create a significant difference. Additionally, the AIC and BIC for the two methods are quite similar.

We calculated the difference in calculated cross section ratio for each x_T bin and found that they only differed by an average of 4.89%, which was actually consistently within the error bounds of the calculated cross section ratio for each fitting. We decided that this was too low of a difference to justify using the log likelihood fitting method, especially since its benefit over χ^2 minimization was debatable to begin with.

3-3 Fitting by Bin Number

For the first set of fits, we used 15 points for each of our plots. This was not by any conscious and deliberate choice, however; rather the collaboration had been using 15 bins in each of their histograms for each of their fits when using this method for calculating the cross section ratio and we just followed suit. Since the number of points can have a drastic impact on the fitting for each plot, we decided to vary the number of bins used in our plots. The resulting statistics are shown in Table 4.

Bin Number	χ^2	AIC	BIC	-2ln(2)
5	2.001	4.032	3.251	0.032
6	3.120	4.103	3.687	0.103
8	6.149	4.189	4.348	0.189
10	5.478	4.210	4.815	0.210
12	8.797	4.487	5.457	0.487
14	8.079	4.495	5.773	0.495

Table 4: Table comparing average statistics for different bin numbers. Each fitting was done linearly with trigger intensity as the beam intensity variable. These results are averaged over each x_T bin.

From the data in Table 4, it was somewhat difficult to say which bin number was really the best. The χ^2 and log likelihood would obviously favor the plots with lower data points. However, using lower numbers of data points obscures the actual effects of the beam intensity on the cross section ratio, and it would definitely not be worth essentially losing data just to lower some statistics on our plots.

By the time we had conducted our survey on variable bin numbers, the collaboration had already decided to use eight bins instead of the previously used 15. Since we saw no obvious reason to do otherwise, we decided to use eight bins from this point on as well. It seemed as though there was somewhat of a jump in χ^2 and log likelihood from 10 to 12 bins, so as long as the number of bins was below 12, we felt confident in using eight bins moving forward. Additionally, this kept our results standardized across the experiment.

3-4 Equal Bin Content

Similar to section 3-3, the next fitting we performed involved manipulating the bins we used to generate our plots. For this fitting, our goal was to have an equal number of data points in each bin. As we can see from the histogram in Figure 10, trigger intensity follows a curve, and thus we have significantly less high and low values data points.

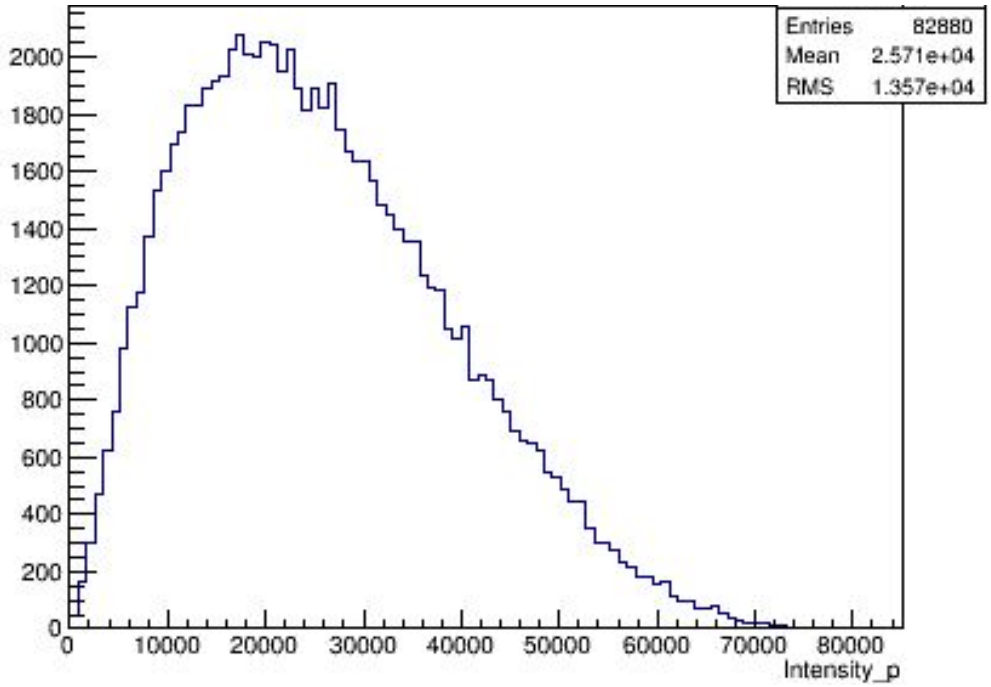

Figure 10: Histogram of all trigger intensity data points from the LH2 data.

Since our bins for the past fittings have all been equal width, the bins with low and high trigger intensity have much fewer data points. Since χ^2 treats each of the points in our plots equally (save for error), this leads to points with lower statistics having disproportionately high impacts on our regressions, and thus our calculations for the cross section ratio.

To remedy this, we decided to divide the trigger intensity data into eight equally populated bins. This meant that the bins that previously had fewer points and thus less reliable data would now be just as reliable as the bins with more points. We sorted the intensity and then counted off by the number of data points divided by eight. The trigger intensity values that we stopped on would be the bounds of each of our bins. We completed this while still binning for x_T, since each x_T bin would have slightly different sets of trigger intensity points. An example of one of the plots fitted in this way is shown in Figure 11.

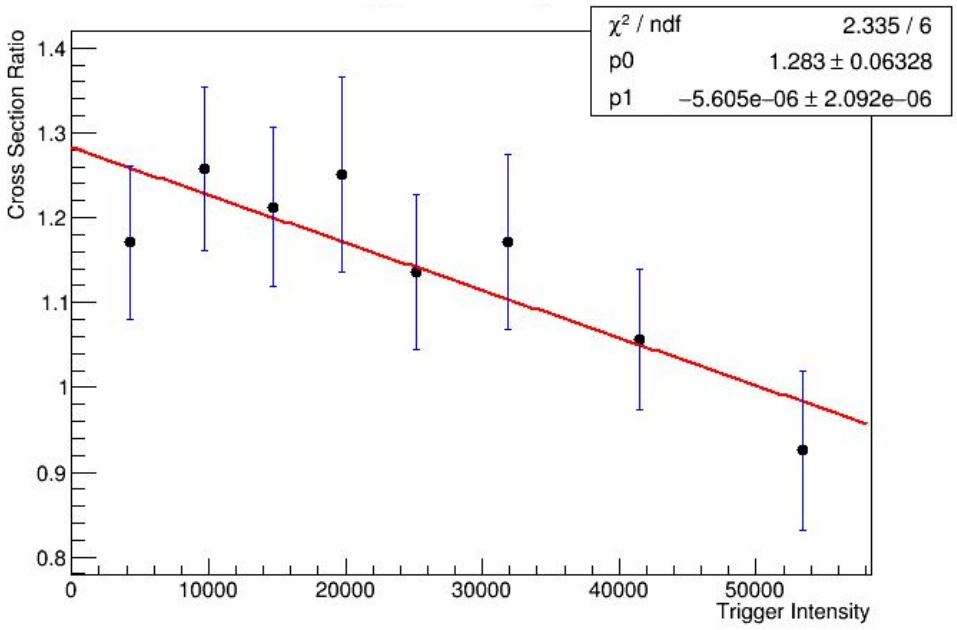

Figure 11: Plot with equally sized bins.

Figure 11 makes it very easy to see the results of our differently spaced binning. The bins on the right are clearly much more spaced apart than those to the left, as they need to encompass a wider range of intensity to account for the lack of data at those values. The leftmost bin is also very clearly wider than the few immediately to the right of it. The results of this fitting in comparison the the equally spaced bin linear fitting are displayed in Table 5.

Fitting Method	χ^2	AIC	BIC	-2ln(L)
Equally Spaced	6.149	4.189	4.348	0.189
Equal Points	4.956	4.170	4.329	0.170

Table 5: Comparison of equally spaced binning vs binning for equal points

From Table 5, we can notice that this method significantly reduced our χ^2. Therefore, this fitting method appears to model the beam effects on the cross section more accurately than the evenly spaced binning we used previously. Nevertheless, it is important to note that the AIC and BIC are practically identical.

3-5 Common Slope

The final fitting involved finding a common slope between all plots. This meant that the slope of each regression would be the same and the only variable that would change would be the intercept. The reasoning behind this was that the beam intensity dependence of the cross

section ratio should be largely independent of x_T. Thus, the slope for each plot should be the same. We saw that most of the slopes for our regressions were roughly similar, and often within the error bounds of each other, yet this fitting would allow us to force the slopes for every plot to be equal. This meant that we would be fitting the background impact on the cross section explicitly.

To perform this fitting, we needed to simultaneously fit each plot. Since we had seven plots to fit, this was a bit of a daunting task. We performed this minimization by creating a minimization script to determine the common slope and then fitting our plots with ROOT with a single parameter function that only determined the intercept. The first step was to create a χ^2 function that we could minimize. The function would have eight parameters: each of the intercepts and the common slope. The minimization program iterated through all plausible values for the potential common slope. For each of the cross section versus trigger intensity plots, it then computed the chi-square for the plot by using ROOT's built in fitting algorithm to determine the intercept. The χ^2 values of each of the plots are then summed with equal weight. By minimizing this function, we essentially simultaneously minimize the χ^2 of each plot with a common slope. A sample of the results of this fitting are in Figure 12.

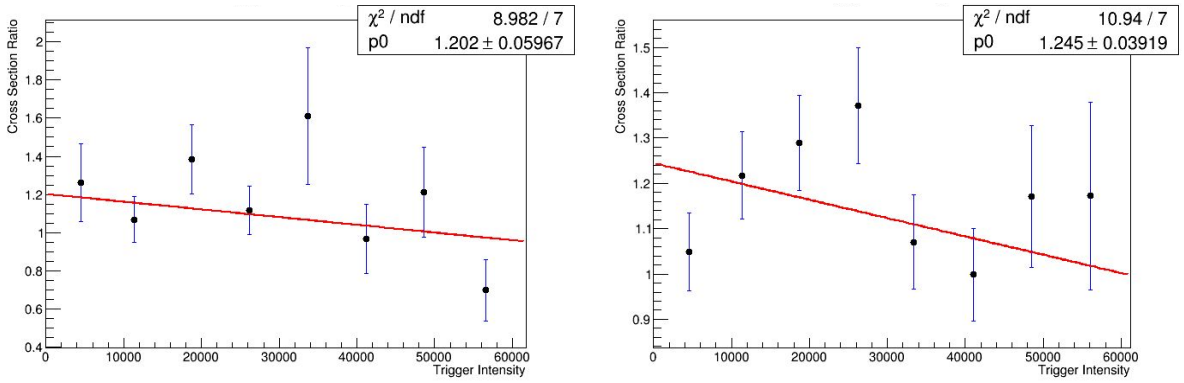

Figure 12: Fitting of two different x_T bins with the same slope. The difference in appearance of the slope in these images is due to the different y-axis scales. The left is events for x_T between 0.100 and 0.130. On the right is events for x_T between 0.130 and 0.160.

From this point, we decided to do a quadratic simultaneous fit. This fit was calculated very similar as the linear one, except another parameter to the χ^2 function was added as a common quadratic term. A sample of these results is shown in Figure 13.

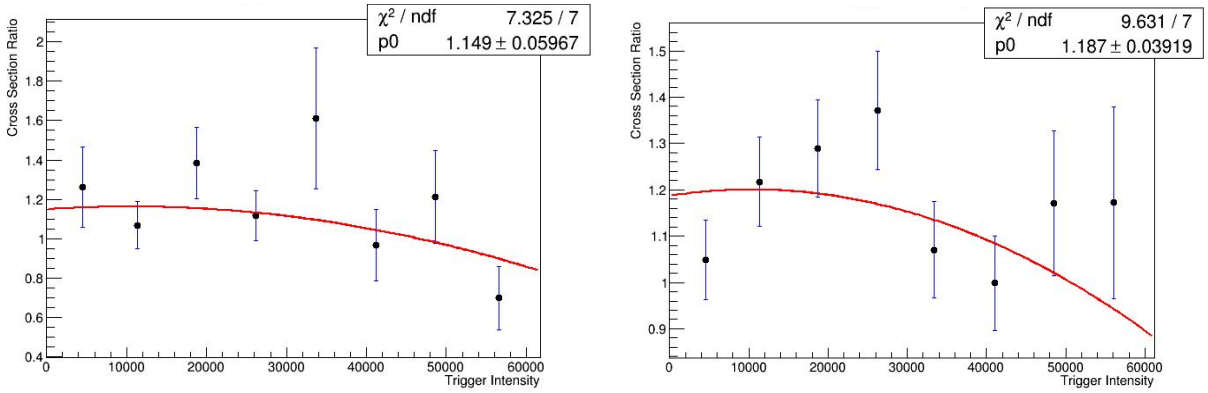

Figure 13: Simultaneous quadratic fitting of the first two x_T bins. The left is events for x_T between 0.100 and 0.130. On the right is events for x_T between 0.130 and 0.160.

These two sets of plots gave us about all of the information we had set out to gain from this method of fitting. Nevertheless, there was one more simultaneous fitting method we decided to investigate. Another member of the E906/SeaQuest group wanted to explore a fitting where the function to be fitted went as $p_0 + (p_1 + p_2 * < x_T >) * TriggerIntensity$, where p_0 would be the parameter that is changed for each plot, p_1 and p_2 are the common parameters, and $<x_T>$ is the average x_T value for each plot. A sample of results from the fitting are shown in Figure Z.

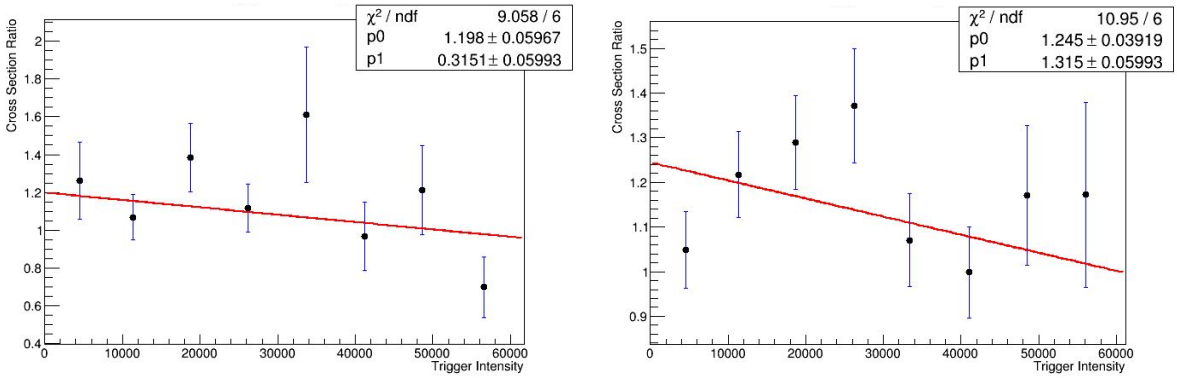

Figure 14: Simultaneous fitting of the first two x_T bins using a modified linear fit. The left is events for x_T between 0.100 and 0.130. On the right is events for x_T between 0.130 and 0.160.

There is almost no difference between the cross section ratios calculated by this method in comparison to our original simultaneous linear fitting. Each value for p0 was within the error bounds of our original fitting, and we did not achieve any lower error bounds with this method. However, this was somewhat expected, since we expected not to see any correlation between x_T and the beam intensity effect, thus making this fitting almost the same as a normal linear fit. Additionally, this method is significantly harder to perform. The addition of another parameter, as well as the inclusion of $<x_T>$ requires some creativity when creating the regression. The results of all of the above fits are displayed in Table 6.

Simultaneous Fitting Method	Total χ^2	χ^2	AIC	BIC	-2ln(L)
Linear	49.392	7.056	4.200	4.359	0.200
Quadratic	43.701	6.243	6.255	6.493	0.255
Modified Linear	49.455	7.065	6.177	6.415	0.177

Table 6: Results of the simultaneous fittings, averaged over all the x_T bins. Total χ^2 is equal to $7*\chi^2$.

The reason that there is an extra column shown for the total χ^2 is because it was the parameter being minimized when determining the slope. It is χ^2 without the division by the number of bins, but it is important to include, as it is the parameter that was minimized for each fitting method. Using χ^2 to compare these fits is still acceptable, as the number of bins and points does not change for each of the simultaneous fitting methods, and these factors are also the same for the other methods we have investigated.

From Table 6 we can see that the simple linear fitting is statistically the best method for simultaneous fitting. The χ^2 and log likelihood statistics for the quadratic and modified linear fittings are nearly identical to those of the linear. Additionally, the quadratic and modified linear fittings have an extra parameter that gives them a much higher AIC and BIC.

4. Conclusions

4-1 Results and Recommendations

Throughout Chapter 3, we explored many methods for fitting the Drell-Yan cross section. Most of the fits we performed were very easy to rule out as possible candidates for use in future calculations. The statistics for our most viable fittings are compiled in Table 6.

Fitting Method	χ^2	AIC	BIC	-2ln(L)	p0
Linear	6.149	4.189	4.348	0.189	1.216
Equal Points	4.956	4.170	4.329	0.170	1.226
Linear Common Slope	7.056	4.200	4.359	0.200	1.216

Table 6: A comparison of each of the current best fitting methods for the Drell-Yan cross section, including a column with the average extrapolated Drell-Yan cross section ratio.

The p0, or the average calculated Drell-Yan cross section ratio, is displayed for these fittings, as now that we have our final fitting methods selected based on their ability to model the background and luminosity effects, it is important to see how much a difference the models actually make in calculating our desired value. From these fits it is clear that the equal points fitting is the best method for modelling the beam luminosity effect of the Drell-Yan process, and therefore calculating the cross section ratio. It has a significantly lower χ^2 value than the original linear fitting, which indicates that putting wider bins in the lower data regions is useful for getting a better understanding of the beam effects.

The one issue that arises from using this method is the extra work that it takes. There is no easy method for calculating the bin bounds as the plot is created, so they must be calculated separately using a script or a separate function within a fitting program. This is not too great of a hurdle, however, as once this is created it would be easy to use it for all future data.

Nevertheless, the decrease in χ^2 and log likelihood does not have a large impact on the cross section ratio. The calculated ratio for the equally spaced binning fit averaged over the xT bins was 1.226 and the ratio for the equal points fitting was 1.216. This is only a 0.819% difference, and thus may not be relevant for many studies. If a quick fitting method is needed, a simple linear fit against trigger intensity should suffice.

4-2 Further Investigations

The fitting method for calculating the Drell-Yan cross section ratio has been proven by this thesis to not only be effective, but also necessary for reducing the impact of systematic errors due to beam impacts. This is shown by the clear correlation between the cross section ratio and each of the beam intensity variables we tested. For all future calculations, our work recommends that a simple linear fit will work, although a fitting with equally populated bins is the best method for calculating the cross section ratio.

As mentioned previously, these methods are of great importance to learning more about the proton sea and its flavor asymmetry. With the techniques outlined in this thesis, future E906/SeaQuest researchers will be able to much more efficiently and accurately analyze data and reduce systematic errors.

Another avenue of investigation that may be pursued is the implementation of asymmetric error bars in the plots. Throughout this study, all error bars have been assumed to be symmetric. While this does not have any real implications on the fittings, it is important to note that in reality, the uncertainty of each of the data points will likely be skewed, with greater uncertainty on the lower end. An example of some of the research I have done beyond this thesis is displayed in Figure 15.

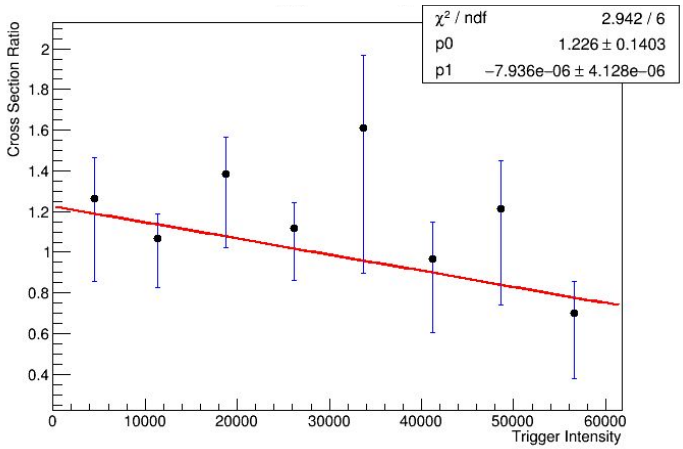

Figure 15: Linear fitting of the cross section ratio vs trigger intensity for x_T between 0.1 and 0.13. The fitting was performed with asymmetric error bars.

In every data point, we see that there is a larger uncertainty on the lower side of the cross section ratio than on the upper side. This impacts the way ROOT fits the plot, and would certainly result in a lower p0 value than a standard symmetric error bar fitting. At the time of writing this, I am still somewhat uncertain of the accuracy of the error bar calculation, however, and I would prefer to refrain from mentioning any specific findings. Nonetheless, these preliminary observations are very promising, and hopefully others will have a chance to investigate them.

References

1. Carithers, Bill, and Paul Grannis. “Discovery of the Top Quark.” *Stanford SLAC*, 1995.
2. Feynman, Richard P. (Richard Phillips), 1918-1988. The Feynman Lectures on Physics. Reading, Mass. *Addison-Wesley Pub. Co.*, 1963-1965.
3. Nave, R. "Quarks." Quarks. N.p., n.d. Web. 02 Apr. 2019.
4. Steinberger, Jack. *Learning about Particles: 50 Privileged Years*. Springer, 2005.
5. Hawker, E.A. et al "Measurement of the Light Antiquark Flavor Asymmetry in the Nucleon Sea." *ArXiv*. N.p., 12 Mar. 1998. Web.
6. M. Arneodo et al. (CERN NMC), “Reevaluation of the Gottfried sum”, Phys. Rev. D50, 1994.
7. Wu-Ki Tung, “Bjorken scaling”. Scholarpedia, 2009.