

First Observations of Beam Spin Asymmetries for K⁺

David M. Riser

December 2018

Contents

1	Introduction	5
2	Basic Analysis & Corrections	9
2.1	Introduction	9
2.2	E1-F	9
2.3	Determination of Good Run List	9
2.4	Helicity Determination	10
2.5	Vertex Corrections	10
2.6	Timing Corrections	11
2.7	Kinematic Corrections	12
3	Particle Identification	17
3.1	Introduction	17
3.2	Electron Identification	17
3.2.1	Electron ID Cuts	17
3.3	Inclusive Electron Cross Section	27
3.4	Hadron Identification	32
3.4.1	Hadron ID Cuts	32
4	Beam Spin Asymmetry Analysis	43
4.1	Introduction	43
4.2	Event Selection and Binning	43
4.3	Measured ϕ_h Distributions	47
4.4	Extraction of Modulations	56
References		65

Chapter 1

Introduction

Protons and neutrons (nucleons) are spin-half fermions. Striking results of measurements performed by the European Muon Collaboration (EMC) in 1988 [1] demonstrated that just 30% of the spin of the nucleons can be attributed to quark spin. This important experiment changed the commonly held view thatf nucleon spin structure arose from simply the spins of the constituent valence quarks, to a more complete picture in which the sea quarks and gluonic spins also contributed to the total nucleonic spin. Moreover, the orbital angular momentum possessed by the quarks is now expected to contribute, and is the subject of much interest.

Addressing the question of orbital angular momentum distributions of partons within nucleons motivates moving beyond a co-linear picture of parton interactions. During the early 1990s, theoretical tools began to emerge that are now being used to study quark dynamics in three-dimensions. Transverse momentum dependent functions (TMDs) naturally extend the co-linear parton distribution functions (PDFs) to include intrinsic quark momentum in the plane transverse to the hard probe [2, 3].

Sadly, TMDs are not directly observable. Despite this fact, single spin asymmetry (SSA) measurements of semi-inclusive deeply inelastic scattering (SIDIS) have proven useful in recent years as inputs for phenomenological extraction of TMD parton distribution functions (TMD PDFs, sometimes just called TMDs) and TMD fragmentation functions (TMD FFs or simply FFs) [4–6]. Because of the absence of a TMD PDF, semi-inclusive annihilation of $e^+e^- \rightarrow h_1h_2X$ has been successfully used as input to TMD FF extractions [7].

By assuming single photon exchange and writing the QED interaction between the virtual photon and the nucleon as a generic vertex, then applying hermiticity, parity, and naive time-reversal invariance, the cross section for SIDIS can be written in a model independent way in terms of structure functions [2, 3] (terms that arise from target polarization are omitted below).

$$\frac{d^5\sigma}{dx dQ^2 dz d\phi_h dP_T^2} = \frac{\alpha_{em}^2}{2x_B y Q^2} \frac{y^2}{1-\varepsilon} \left(1 + \frac{\gamma^2}{2x_B}\right) \left\{ F_{UU,T} + \varepsilon F_{UU,L} \right. \\ \left. + \sqrt{2\varepsilon(1+\varepsilon)} \cos\phi_h F_{UU}^{\cos\phi_h} + \varepsilon \cos(2\phi_h) F_{UU}^{\cos 2\phi_h} + \lambda_e \sqrt{2\varepsilon(1-\varepsilon)} \sin\phi_h F_{LU}^{\sin\phi_h} \right\} \quad (1.1)$$

Here, typical definitions for the SIDIS kinematic variables are used (where $q = l - l'$ and $Q^2 = -q^2$).

$$x = \frac{Q^2}{2P \cdot q} \quad y = \frac{P \cdot q}{P \cdot l} \quad z = \frac{P \cdot P_h}{P \cdot q} \quad \gamma = \frac{2Mx}{Q} \quad (1.2)$$

Additionally, the ratio ε of the longitudinal and transverse photon flux is shown below.

$$\varepsilon = \frac{1 - y - \frac{1}{4}\gamma^2 y^2}{1 - y + \frac{1}{2}y^2 + \frac{1}{4}\gamma^2 y^2} \quad (1.3)$$

The factor λ_e appearing in the cross section refers to the helicity state of the incoming lepton, and ϕ_h is the angle between the lepton and hadron scattering planes. By measuring the cross section for both electron helicity states, the beam spin asymmetry can be constructed.

$$BSA = \frac{d\sigma^+ - d\sigma^-}{d\sigma^+ + d\sigma^-} = \frac{A_{LU}^{\sin\phi_h} \sin\phi_h}{1 + A_{UU}^{\cos\phi_h} \cos\phi_h + A_{UU}^{\cos(2\phi_h)} \cos(2\phi_h)} \quad (1.4)$$

Where the coefficient $A_{LU}^{\sin\phi}$ is defined as,

$$A_{LU}^{\sin\phi_h} = \sqrt{2\varepsilon(1-\varepsilon)} \frac{F_{LU}^{\sin\phi_h}}{F_{UU,T} + \varepsilon F_{UU,L}} \quad (1.5)$$

and the unpolarized coefficients are defined in a similar way. Within the TMD framework, the structure function $F_{LU}^{\sin\phi_h}$ is a pure twist-three structure function. With the assumption of twist-three factorization (which has not been demonstrated) the structure function is composed of four terms.

$$F_{LU}^{\sin\phi_h} = \frac{2M}{Q} \mathcal{C} \left[-\frac{\hat{\mathbf{h}} \cdot \mathbf{k}_T}{M_h} \left(xeH_1^\perp + \frac{M_h}{M} f_1 \frac{\tilde{G}^\perp}{z} \right) + \frac{\hat{\mathbf{h}} \cdot \mathbf{p}_T}{M} \left(xg^\perp D_1 + \frac{M_h}{M} h_1^\perp \frac{\tilde{E}}{z} \right) \right] \quad (1.6)$$

The notation \mathcal{C} is shorthand presented in [3] as a way to write structure functions in terms of the convolutions of PDF and FF objects.

$$\mathcal{C}[\omega fD] = x \sum_a e_a^2 \int d^2\mathbf{p}_T d^2\mathbf{k}_T \delta^{(2)}(z\mathbf{k}_T + \mathbf{p}_T - \mathbf{P}_{h\perp}) \omega(\mathbf{k}_T, \mathbf{p}_T) f^a(x, k_T^2) D^a(z, p_T^2) \quad (1.7)$$

Here the summation over quark flavors a is explicitly shown.

At twist-three four TMD PDFs appear in the structure function, one of which is known as the Boer-Mulders function h_1^\perp . The Boer Mulders TMD is a twist-two time-reversal odd function. Additionally,

g^\perp is a twist-three time reversal odd TMD, that has been compared to a higher twist analog of the Sivers function. The remaining TMDs are e , a chiral odd twist-three TMD and f_1 the unpolarized TMD. Since the TMD PDF is accompanied by a TMD fragmentation function (FF), four fragmentation functions also appear in this structure function, at leading order the Collins H_1^\perp and unpolarized D_1 , and at twist-three \tilde{G}^\perp , \tilde{E} .

Despite measurements of $A_{LU}^{\sin \phi_h}$ for π^+ , π^- , and the neutral π^0 mesons [8], little is known about the contribution of each individual PDF/FF term to the asymmetry. Still fewer are the SSA measurements which have tagged kaons in the final state. Previous analyses have reported results for the Sivers asymmetry $A_{UT}^{\sin(\phi_h - \phi_S)}$, which was recently reported by Hall A [9] and was observed with a small magnitude. However, measurements by the HERMES collaboration indicate that the Sivers asymmetry has quite a large magnitude [4].

This measurement of K^+ beam spin asymmetries is non-zero, and provides further evidence that at the presently studied kinematics, intrinsically twist-three distributions functions are not vanishingly small.

Chapter 2

Basic Analysis & Corrections

2.1 Introduction

This chapter discusses analysis procedures that are used in this analysis. These procedures can be divided into two groups. The first type of basic analysis described is the aggregation or calculation of scalar values over the run-period (examples include luminosity and helicity). The second type of analysis procedure described is a correction to measured values. Vertex corrections, timing corrections, and kinematic corrections will be discussed.

2.2 E1-F

This study uses the dataset collected between April and July of 2003 known as E1-F. During this run period the beam energy was 5.498 GeV and the target was a 5cm liquid hydrogen cell. The torus current was set to 2250 Amperes, to maximize pion acceptance.

2.3 Determination of Good Run List

The total dataset contains 831 runs. Due to the complexities of the CLAS experimental setup, it is not uncommon for run conditions to change between runs such that a portion of the data collected are not of analysis quality. For this reason, a good run list is constructed.

Good runs are selected for the list by counting good electrons in each file and normalizing by the accumulated charge for the associated file. For each file, the difference between subsequent Faraday cup readings is summed to calculate the charge for the file. Creation of the total charge for the run includes an additional contribution from endpoints in adjacent files. This extra charge builds up after the last

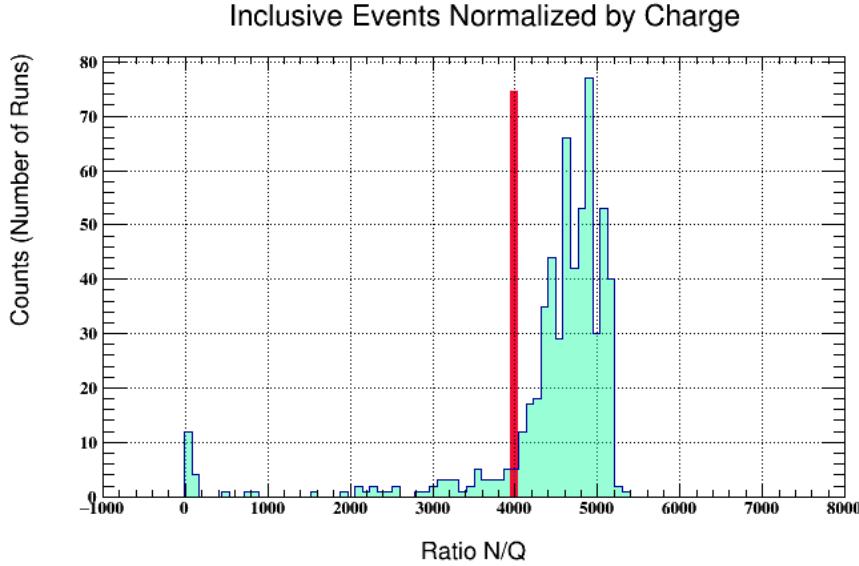


Figure 2.1: Inclusive electrons per file normalized by the total charge accumulated for the file. This quantity is used to make a good run list.

scalar reading of one file and before the first scalar bank in the new adjacent file. While the number of events collected varies from run to run the ratio defined above is a stable quantity – provided that the run conditions do not vary greatly. Good runs were chosen to have $N/Q > 4000$ based on inspection of the figure 2.3. The good run list used for this analysis contains 522 runs.

2.4 Helicity Determination

During the course of the E1-F run period the beam helicity convention was changed by the insertion of a half-wave plate at the injector. The definition of \pm helicity must change in accordance with these wave-plate insertions. To monitor these changes, the value of $A_{LU}^{\sin \phi}$ for π^+ is recorded for every run. Whenever the asymmetry (which has a magnitude of around 3%) changes sign, the sign convention has changed (see figure 2.4). These changes are taken into account in the data analysis.

2.5 Vertex Corrections

The track vertex position (v_x, v_y, v_z) is calculated based on the intersection of each track with the mid-plane (the plane which contains the beamline and bisects the sector at $\phi_{rel} = 0$). If the beam is not centered at $(x, y) = (0, 0)$, the vertex position calculation needs to be corrected by shifting the mid-planes in accordance with the target offset. The offset (x, y) is identified by plotting events from the control foil placed near the target, which has a z position of 20 cm. For the E1-F run period, the beam

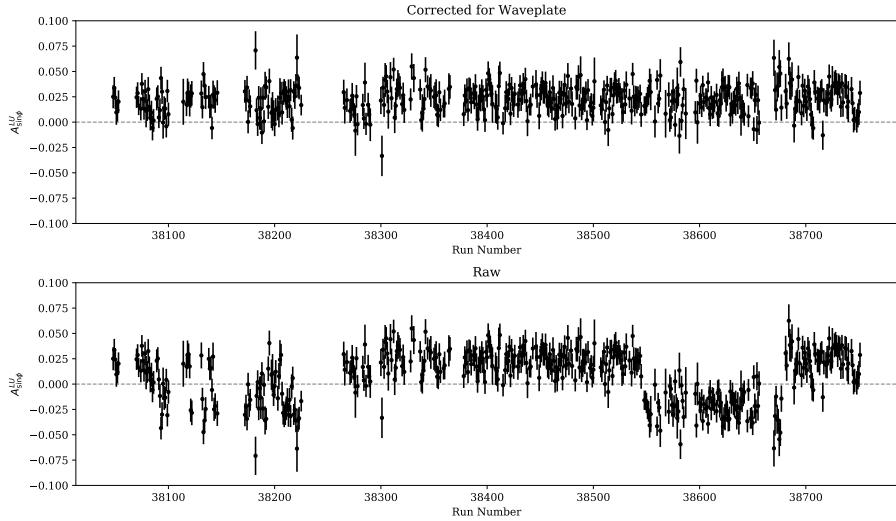


Figure 2.2: The waveplate position is determined and corrected by plotting the BSA for π^+ mesons as a function of the run number. The top panel shows the corrected results, the bottom shows the results before changing the helicity.

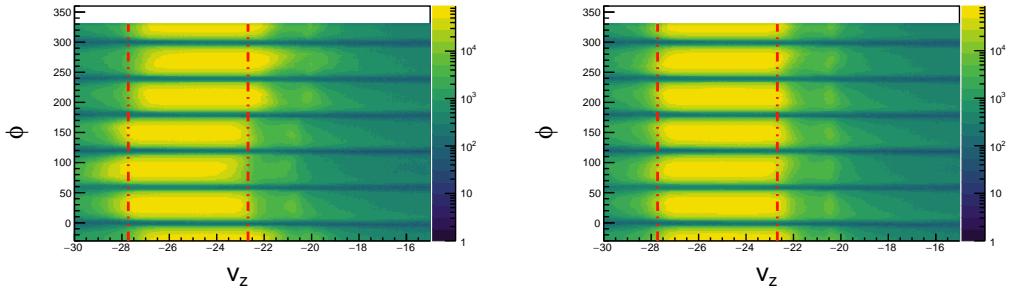


Figure 2.3: The z-vertex v_z position shown for different values of ϕ the azimuthal angle in the hall. The left figure shows the distribution before corrections are applied, the right after. The vertical red lines bound the region which we define as acceptable for electrons in our analysis.

position was $(0.15, -0.25)$ cm. This correction is applied to the entire data set based on this position, and its successful effect is shown in figure 2.5. In practice the beam position may vary more frequently than in our case, and the correction would be applied run-by-run, here it is not necessary.

2.6 Timing Corrections

Timing information comes from the time-of-flight detector system. After calibration, small offsets in timing between time of flights paddles still exist for the E1-F dataset. These biases can be removed on a run-by-run and paddle-by-paddle basis by adding a small shift t_{corr} . In order to determine this shift t_{corr} for each paddle, charged pions are used.

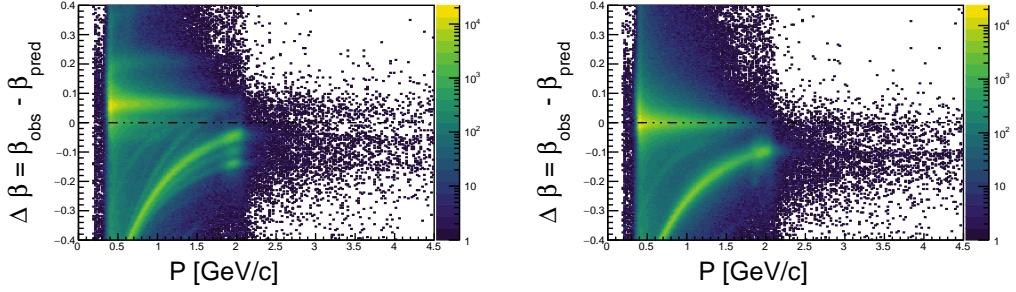


Figure 2.4: Timing corrections are shown for paddle 24 of sector 1. The left image shows the $\Delta\beta$ distribution before corrections. On the right the same is shown after correction of the timing for this paddle. We assume the mass of the track to be the pion, these show up as the green band. Heavier protons are visible below the pion band.

Using momentum information from the drift chambers the value of β can be predicted and the difference $\Delta\beta$ can be determined for each pion.

$$\Delta\beta = \beta_{obs} - \beta_{pred} = \frac{d}{t_{obs}} - \sqrt{1 + (m/p)^2} \quad (2.1)$$

Here m (the mass of the particle detected) is assumed to be m_π , t_{obs} refers to the measured time of flight for the particle, and d is the distance traveled in time t_{obs} which comes from the fit track. The offset $\Delta\beta$ from 0 is used to define the value of t_{corr} for each paddle. If this value is exceedingly small, no correction is applied. For some paddles with low statistics a reasonable value for t_{corr} cannot be obtained and these paddles are excluded from the analysis.

In the method described above, the calibrated paddle is the one which is struck by the pion. The electron paddle which was struck could also require calibration. In practice the magnitude of the correction term t_{corr} is small, and the paddle offset is (likely) randomly distributed about 0 when considering all paddles. By including events from many different (electron) paddles, miscalibration effects from the electron side cease to be important. This is demonstrated by the success of the technique in centering the $\Delta\beta$ distributions. This work was first described in [10].

2.7 Kinematic Corrections

The magnetic field map used in reconstruction to swim particle tracks cannot perfectly match the real magnetic field of the hall. As a result of this the reconstructed momentum of particles is often slightly off (of order 1%). Small misalignment in detector positions also contribute to this effect. In order to correct for these small differences, the momentum (p_x, p_y, p_z) and hence θ of charged tracks is corrected.

Various procedures exist for the correction of kinematic variables of measured particles, and they all rely on energy and momentum conservation applied to standard processes (such as elastic scattering). The procedure used to derive corrections for the E1-F dataset was developed and described by Marco Mirazita in [11].

First, elastic ($ep \rightarrow ep$) events are selected by identifying events that contain at least one electron and one proton, then requiring that the missing mass M_X of the ($ep \rightarrow epX$) system is close to 0. The kinematics of the event are then calculated.

$$k^\mu = (k, 0, 0, k) \quad (2.2)$$

$$p^\mu = (M_p, 0, 0, 0) \quad (2.3)$$

$$k'^\mu = (k', k' \sin \theta, 0, k' \cos \theta) \quad (2.4)$$

$$p'^\mu = (E_p, -p' \sin \alpha, 0, p' \cos \alpha) \quad (2.5)$$

Above, the four momenta correspond to the incoming electron k , the outgoing electron k' , the stationary target proton p and the scattered proton p' . Applying energy and momentum conservation to the equations above yields 3 equations.

$$k + M_p = k' + \sqrt{M_p^2 + p'^2} \quad (2.6)$$

$$k' \sin \theta = p' \sin \alpha \quad (2.7)$$

$$k = k' \cos \theta + p' \cos \alpha \quad (2.8)$$

Using these equations, the electron angle θ and the proton angle α can be predicted by using the momenta (k', p') . These values are compared with measured values and iteratively corrected by tuning the parameters of a phi-dependent 2nd order polynomial.

$$\cos \theta = 1 - M_p \frac{k - k'}{kk'} \quad (2.9)$$

$$\tan \alpha = \frac{1}{p'} \frac{k' \sin \theta}{k - k' \cos \theta} \quad (2.10)$$

After θ corrections are applied, the momenta of the electrons are corrected by using an analogous procedure for k' instead of θ and α . The momentum corrections are calculated as functions of ϕ for

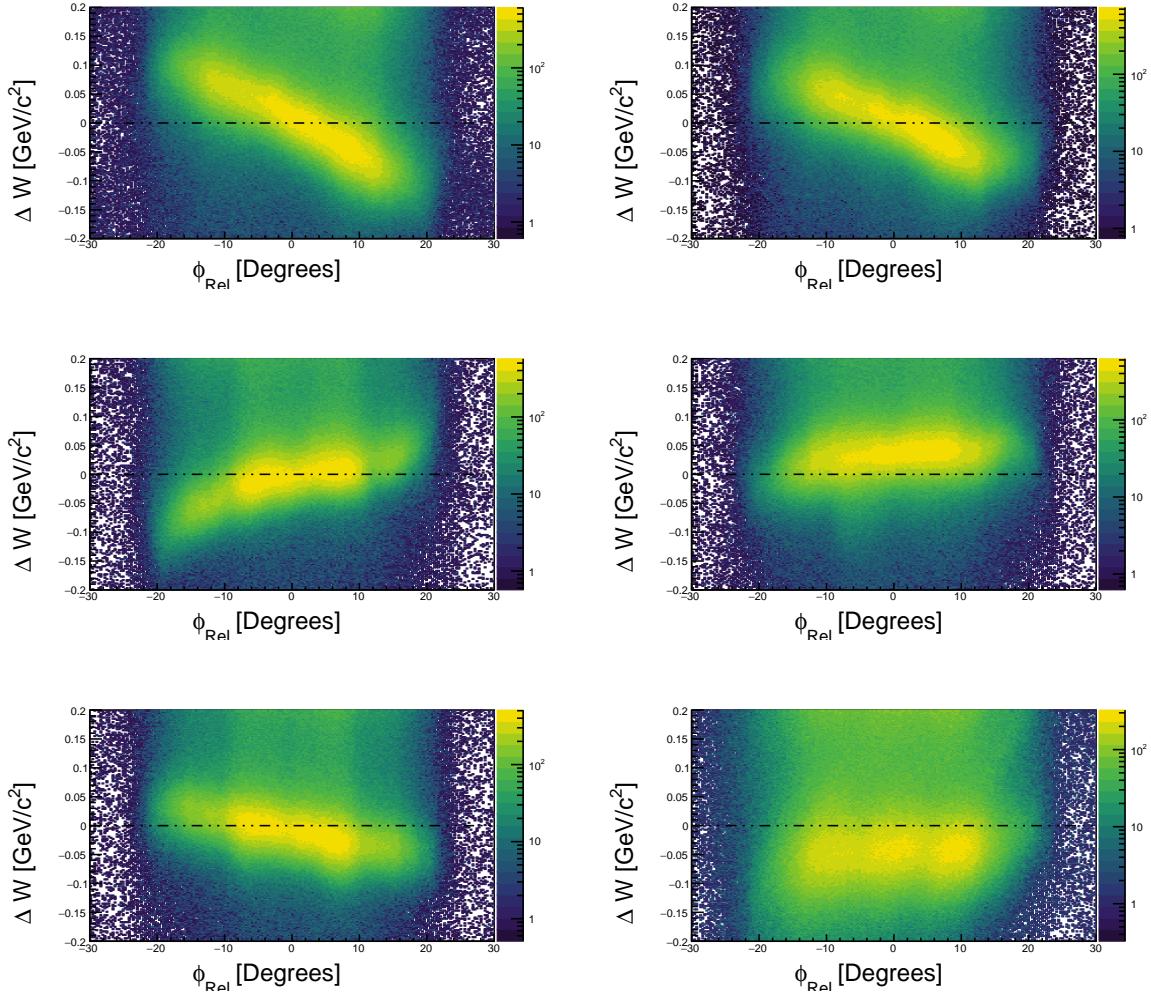


Figure 2.5: This figure shows the deviation from M_p of the W spectrum peak for elastic $ep \rightarrow ep$ events (before corrections). The six sectors of CLAS are shown here starting from 1 in the top left and ending with 6 at the bottom right.

each sector in one degree bins of θ . Finally, the positively charged particles momenta are corrected by selecting the exclusive event ($ep \rightarrow e\pi^+ N$). In this reaction the scattered electron and pion are detected and the neutron is selected using a missing mass cut. Assuming the electron momentum, electron angle, and pion angle to be correct, the pion momentum correction is then calculated by iteratively improving the central position of the neutron mass peak to coincide with M_N . Marco Mirazita shows in his note that these corrections can be satisfactorily applied to all negative and positive particles.

In this work momentum corrections are applied to the scattered electron. We find the corrections to be very effective for all sectors, while sector 6 has a proton resonance width approximately 10 MeV wider than the other sectors, this is on the order of the resolution of the detector and is likely caused by dead wires. The central position of the proton resonance peak is corrected successfully.

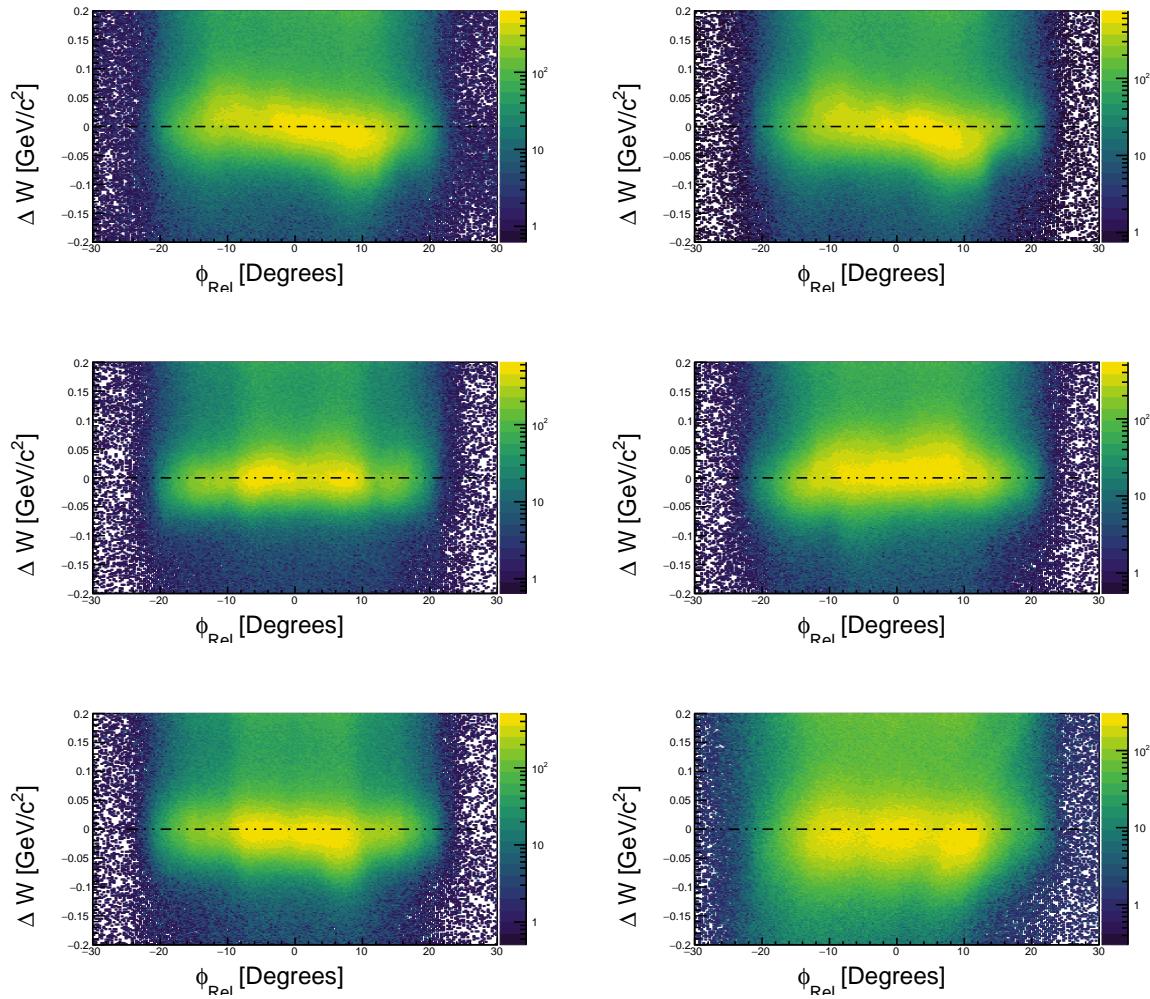


Figure 2.6: This figure shows the deviation from M_p of the W spectrum peak for elastic $ep \rightarrow ep$ events (after ϕ -dependent corrections). The momentum corrections are applied as a function of ϕ and this plot demonstrates that the correction significantly improves the position and width of the elastic resonance in W as a function of the azimuthal angle ϕ . The six sectors of CLAS are shown here starting from 1 in the top left and ending with 6 at the bottom right.

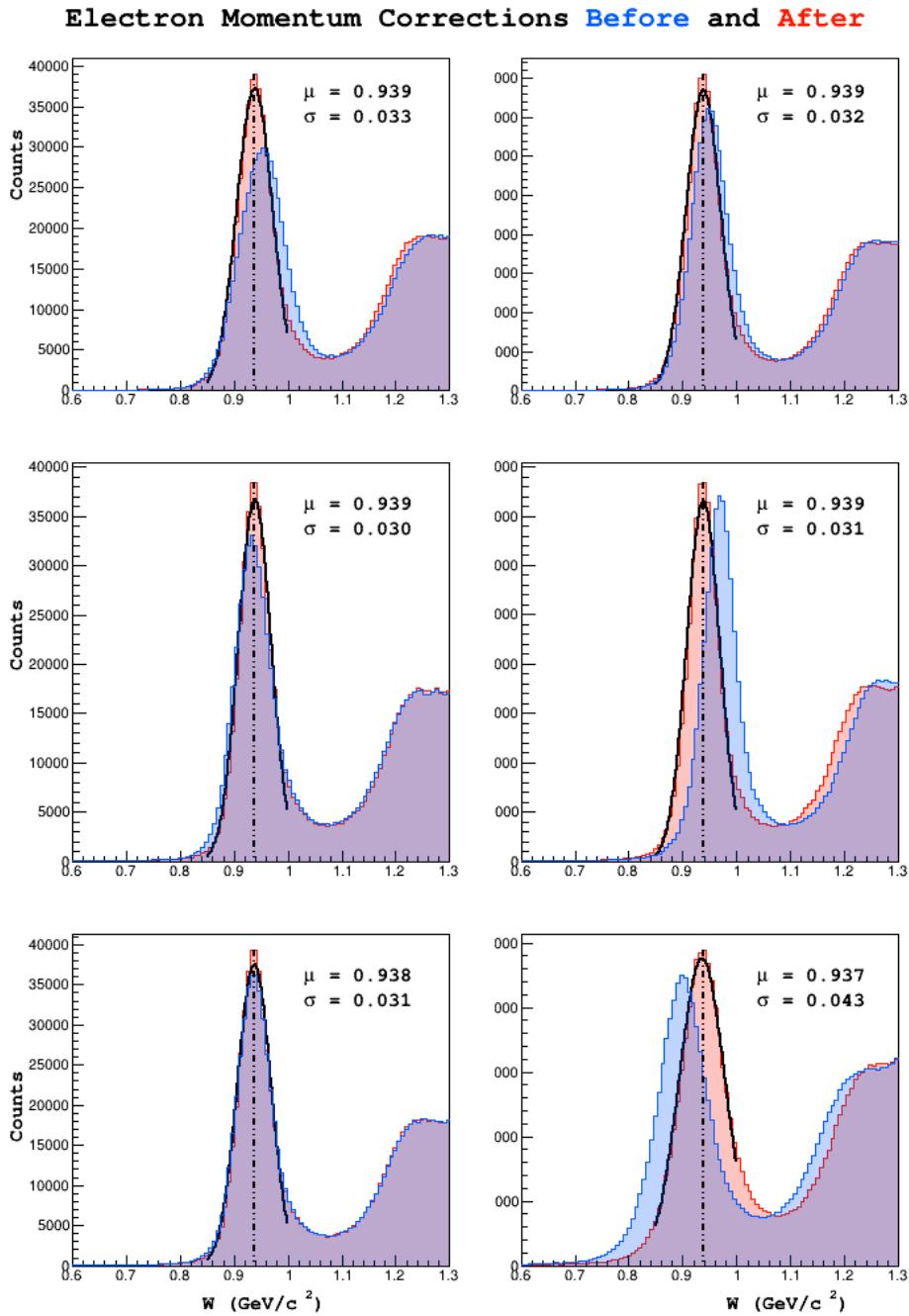


Figure 2.7: Elastic events shown in the spectrum of W before and after momentum corrections are applied. Each corrected sector's histogram has been fit with a Gaussian, whose mean μ and standard deviation σ are shown on each figure. This figure demonstrates that they are very effective in correcting the electron kinematics. The apparent disparity between the number of events in sectors 1-5 and sector 6 is largely an illusion due to the larger width in sector 6, in reality there is an approximate 2 percent difference in total counts between sector 6 and the other sectors average number of counts.

Chapter 3

Particle Identification

3.1 Introduction

Particle identification (PID) is the process of classifying tracks as known particles. After reconstruction and matching of detector responses to each track, the reconstruction package `reccsis` assigns a preliminary particle identification based on loose selection criteria. In this analysis, tracks are classified based on a more stringent criteria. This chapter discusses the methodology used to classify particles.

3.2 Electron Identification

Electrons in CLAS are abundant, and the detection of an electron is a basic necessity for every event that will be analyzed. The most naive approach to performing electron identification would be to call all negatively charged tracks electrons. Doing this would provide an extremely efficient identification of electrons (none of them are missed), however the purity of the sample (the fraction of tracks identified as electrons that are actually electrons) would be low due to the vast quantity of negatively charged pions that are produced in during the experiment. Additionally, doing this would completely eliminate the possibility of identifying negatively charged pions or kaons, as all negative tracks would be called electrons. In practice then, the identification of electrons is concerned with removal of negative pions and kaons from the sample of negative tracks. This is accomplished by applying a series of cuts on measured variables that distinguish between electrons and pions (pions are the dominant background).

3.2.1 Electron ID Cuts

The cuts used to select electrons are enumerated below.

- Negative charge

- Drift chamber region 1 fiducial
- Drift chamber region 3 fiducial
- Electromagnetic Calorimeter fiducial (UVW)
- EC minimum energy deposition
- Sampling Fraction (momentum dependent)
- z-vertex position
- Cherenkov counter θ_{cc} matching to PMT number
- Cherenkov counter ϕ_{rel} matching to PMT (left/right)
- Cherenkov counter Fiducial Cuts

Each cut is now be described in more detail.

Negativity Cut

Each track is assigned a charge based on the curvature of its trajectory through the magnetic field of the torus. This is done during the track reconstruction phase. Tracks are eliminated as electron candidates if they are not negatively charged.

Drift chamber fiducial

The fiducial region or volume is a term used to refer to the region of a sensitive detector which is unimpeded in its acceptance of physics events. In practice, shadows from other detectors, poorly understood edge effects, or geometric obstacles may impede the flight of particles from the target, and render regions of sensitive detectors unreliable (to use the vocabulary presented above, these events fall outside of the fiducial region of the detector).

Negative tracks which pass geometrically close to the edges of the drift chamber are, from a tracking perspective, more difficult to understand. Additionally, tracks which fall outside of the fiducial region of the drift chambers are likely to fall outside of the fiducial region of the downstream detectors as well. For these reasons, it is common to remove tracks which are geometrically close to the boundaries of the drift chambers in region 1 as well as region 3 coordinate systems.

To implement this cut the (x, y) coordinates of the drift chambers are rotated into one sector. Then boundaries y_{left}, y_{right} are defined as linear functions of x . The boundary lines are parametrized by an

Region	Height h (cm)	Angle θ (degrees)
1	22	60
3	80	49

Table 3.1: Cut parameters used for the DC fiducial cut.

EC Coordinate	Min (cm)	Max (cm)
U	70	400
V	-	362
W	-	395

Table 3.2: Cut parameters used for the EC fiducial cut.

offset h and an angle of the boundary line with respect to the center of the sector at $x = 0$. The slope of these lines is $\pm \cot(\theta)$.

$$y_{right} = h + \cot(\theta)x \quad (3.1)$$

$$y_{left} = h - \cot(\theta)x \quad (3.2)$$

Tracks passing this criterion are those which have $y > y_{left}(x)$ and $y > y_{right}(x)$.

Electromagnetic Calorimeter fiducial (UVW)

As particles traverse the electromagnetic calorimeter they develop electromagnetic showers. If the track passes close to the edges of the detector, there is a chance that the shower will not be fully contained within the calorimeter volume (it spills out the edges). For this reason, it is standard to remove the hits which fall within the outer 10 centimeters of each layer of the EC (10 centimeters is the width of a scintillator bar). This cut is applied in the U, V, W coordinate system.

EC minimum energy deposition

One way to differentiate between these electrons and pions is to exploit the difference in energy deposition between the two in the electromagnetic calorimeter. Electron typically develop a much larger and more energetic shower than π mesons, which minimally ionize the calorimeter material. The result is that the total energy deposition is typically larger for electrons than π mesons. In this analysis we require that at least 60 MeV was deposited in the inner calorimeter for electron candidates 3.2.1.

Sampling Fraction (momentum dependent)

The electromagnetic calorimeter is designed such that electrons will deposit $E_{dep}/p \approx 0.3$ approximately one-third of their energy, regardless of their momentum. In contrast to this, the ratio E_{dep}/p for π mesons decreases rapidly with momentum. To develop a momentum dependent cut for this distri-

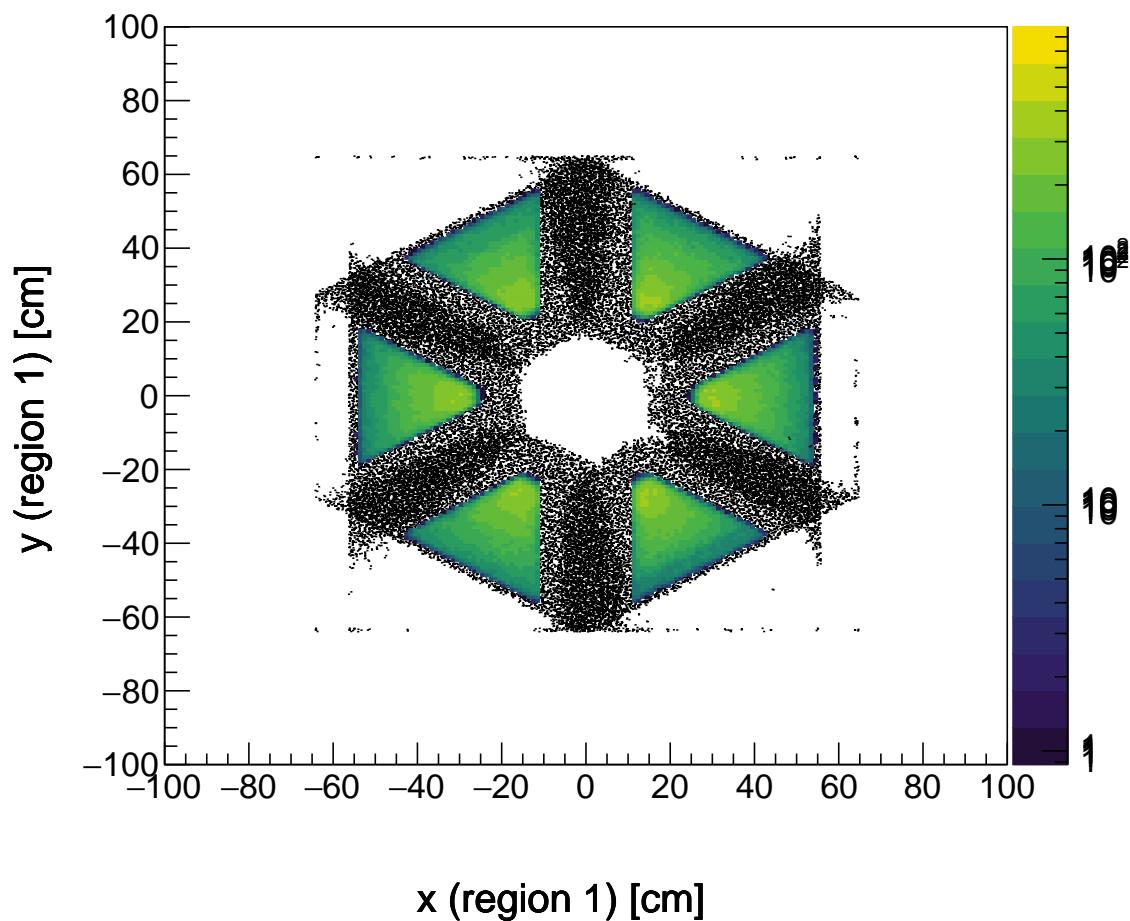


Figure 3.1: Tracks shown in color remain after the application of drift chamber region 1 fiducial cuts to all cuts, shown here as black points.

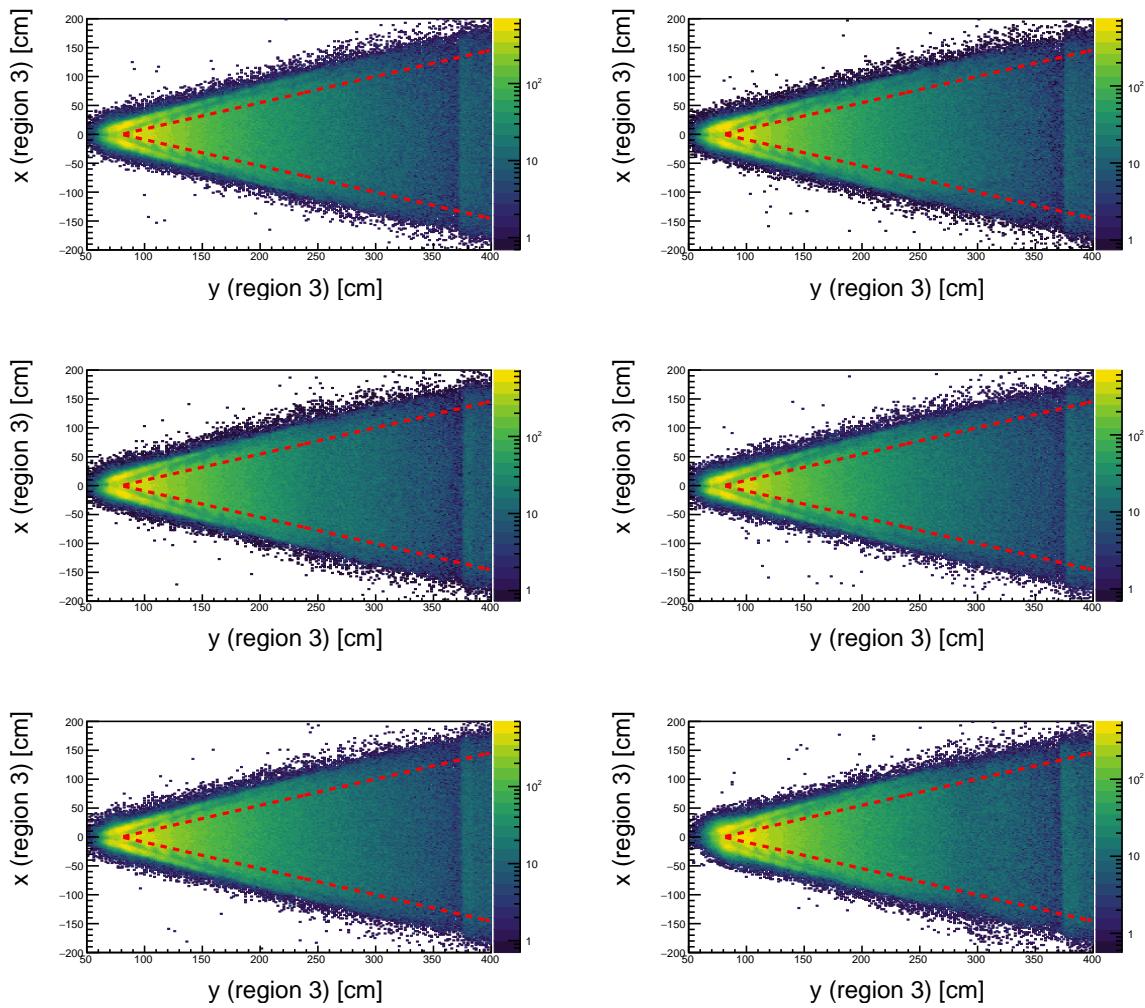


Figure 3.2: The selection criteria shown in red is applied to drift chamber region 3.

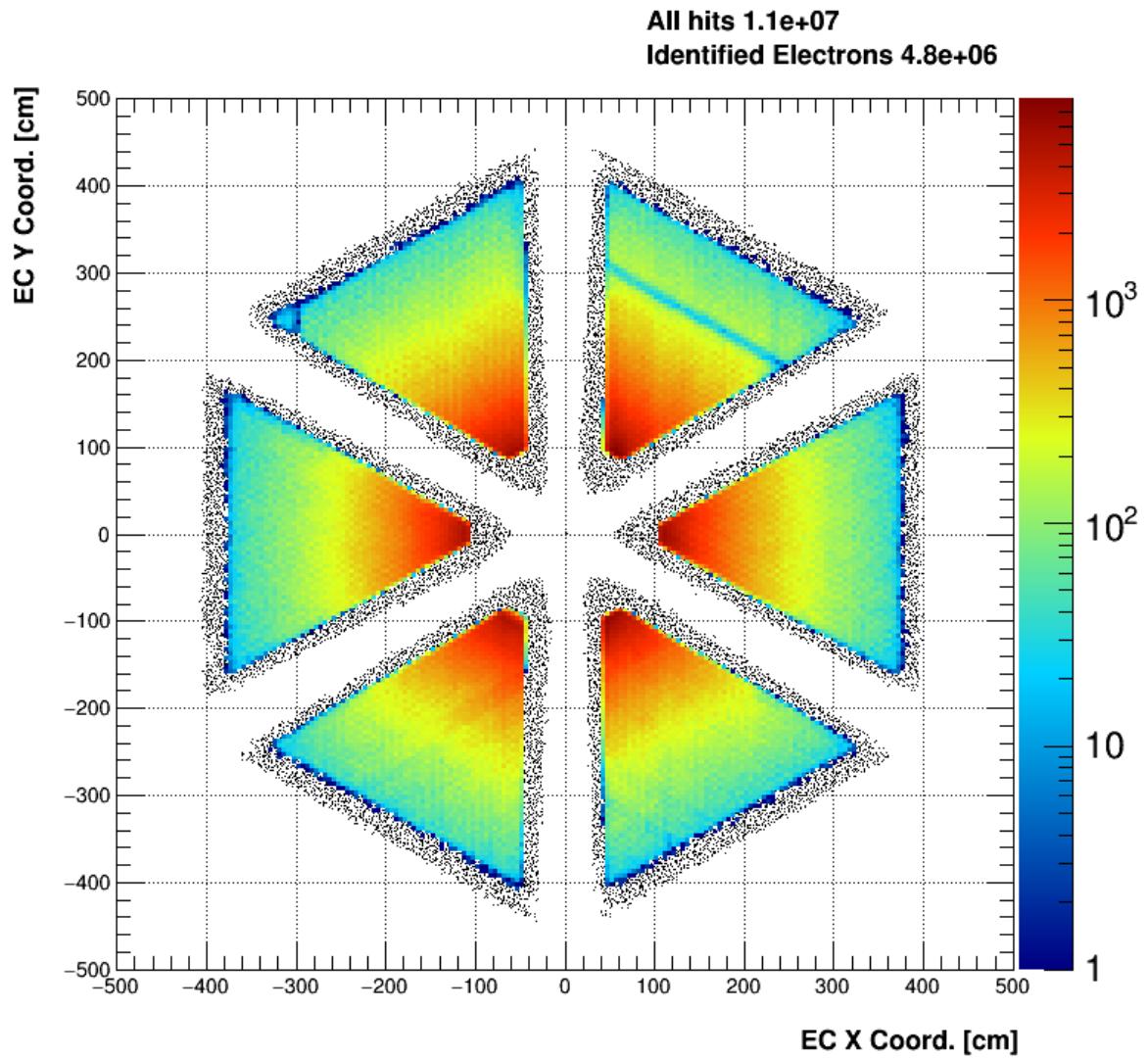


Figure 3.3: All negative tracks are shown here in black. In color, the tracks which pass the EC fiducial cut are shown.

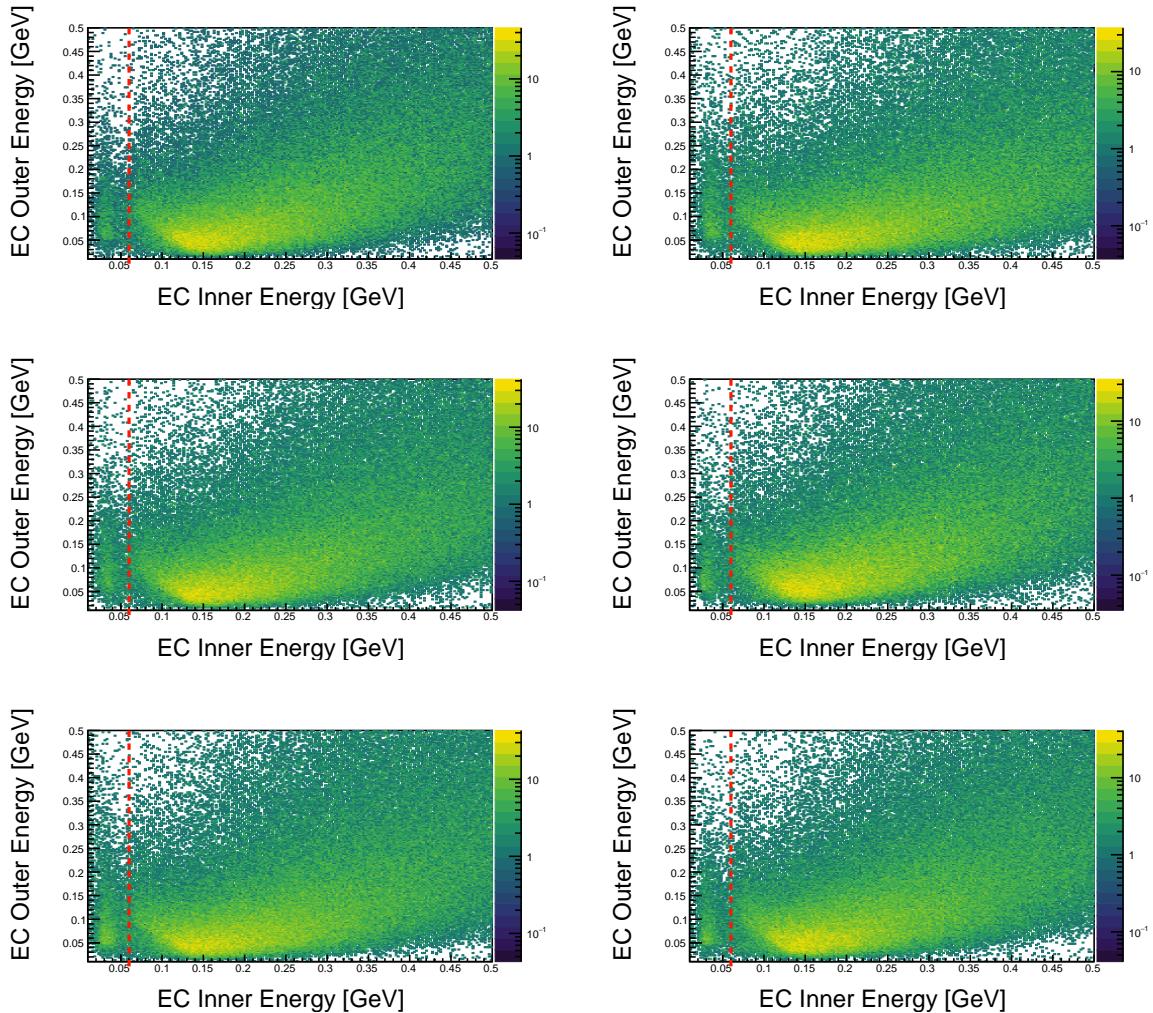


Figure 3.4: Each panel shown above contains events from one sector, increasing from 1-6 from top left to bottom right. The value selected of 60 MeV is applied to all sectors and separates the negatively charged pions (left) from the electrons (right).

bution, all negative candidates are first filled into a two-dimensional histogram of E_{dep}/p vs. p . The histogram is then binned more coarsely in momentum, and projected into a series of 40 momentum slices. Each of these slices is fit with a Gaussian to extract the position μ_i and width σ_i of the electron peak. Finally, a functional form for the mean and standard deviation of the distributions is chosen to be a third order polynomial in momentum.

$$\mu(p) = \mu_0 + \mu_1 p + \mu_2 p^2 + \mu_3 p^3 \quad (3.3)$$

$$\sigma(p) = \sigma_0 + \sigma_1 p + \sigma_2 p^2 + \sigma_3 p^3 \quad (3.4)$$

Boundaries are constructed from this information by adding (subtracting) n_σ from the mean. In the nominal case, we use $n_\sigma = 2.5$.

$$f_{max}(p) = \mu(p) + n_\sigma \sigma(p) = (\mu_0 + n_\sigma \sigma_0) + (\mu_1 + n_\sigma \sigma_1)p + (\mu_2 + n_\sigma \sigma_2)p^2 + (\mu_3 + n_\sigma \sigma_3)p^3 \quad (3.5)$$

$$f_{min}(p) = \mu(p) - n_\sigma \sigma(p) = (\mu_0 - n_\sigma \sigma_0) + (\mu_1 - n_\sigma \sigma_1)p + (\mu_2 - n_\sigma \sigma_2)p^2 + (\mu_3 - n_\sigma \sigma_3)p^3 \quad (3.6)$$

Due to slight differences between the 6 sectors of the CLAS detector, this cut is calibrated and applied for each sector individually. Results are shown in table 3.3.

Parameter	Sector 1	Sector 2	Sector 3	Sector 4	Sector 5	Sector 6
μ_3	-8.68739e-05	0.000459313	9.94077e-05	-0.000244192	-7.65218e-05	-0.000392285
μ_2	-0.000338957	-0.00621419	-0.00267522	-0.00103803	-0.00222768	-0.00105459
μ_1	0.0191726	0.0393975	0.02881	0.0250629	0.0233171	0.0265662
μ_0	0.2731	0.296993	0.285039	0.276795	0.266246	0.25919
σ_3	-0.000737136	0.000189105	-0.000472738	-0.000553545	-0.000646591	-0.000633567
σ_2	0.00676769	-0.000244009	0.00493599	0.00434321	0.00717978	0.00626044
σ_1	-0.0219814	-0.00681518	-0.0180929	-0.0140827	-0.0246181	-0.022029
σ_0	0.0474188	0.0475098	0.0461743	0.0492728	0.0546257	0.0517508

Table 3.3: μ and σ values used to construct the momentum dependent sampling fraction cut.

z-vertex position

Electrons can be produced as part of e^+e^- pairs, or by other processes. For this analysis, these are not of interest. For the purposes of this analysis it is then natural to accept only electron candidates which have a z-vertex $v_z \in [-27.7302, -22.6864]$ within the expected target region. This cut is applied after the vertex position has been corrected (which is discussed in the basic analysis section).

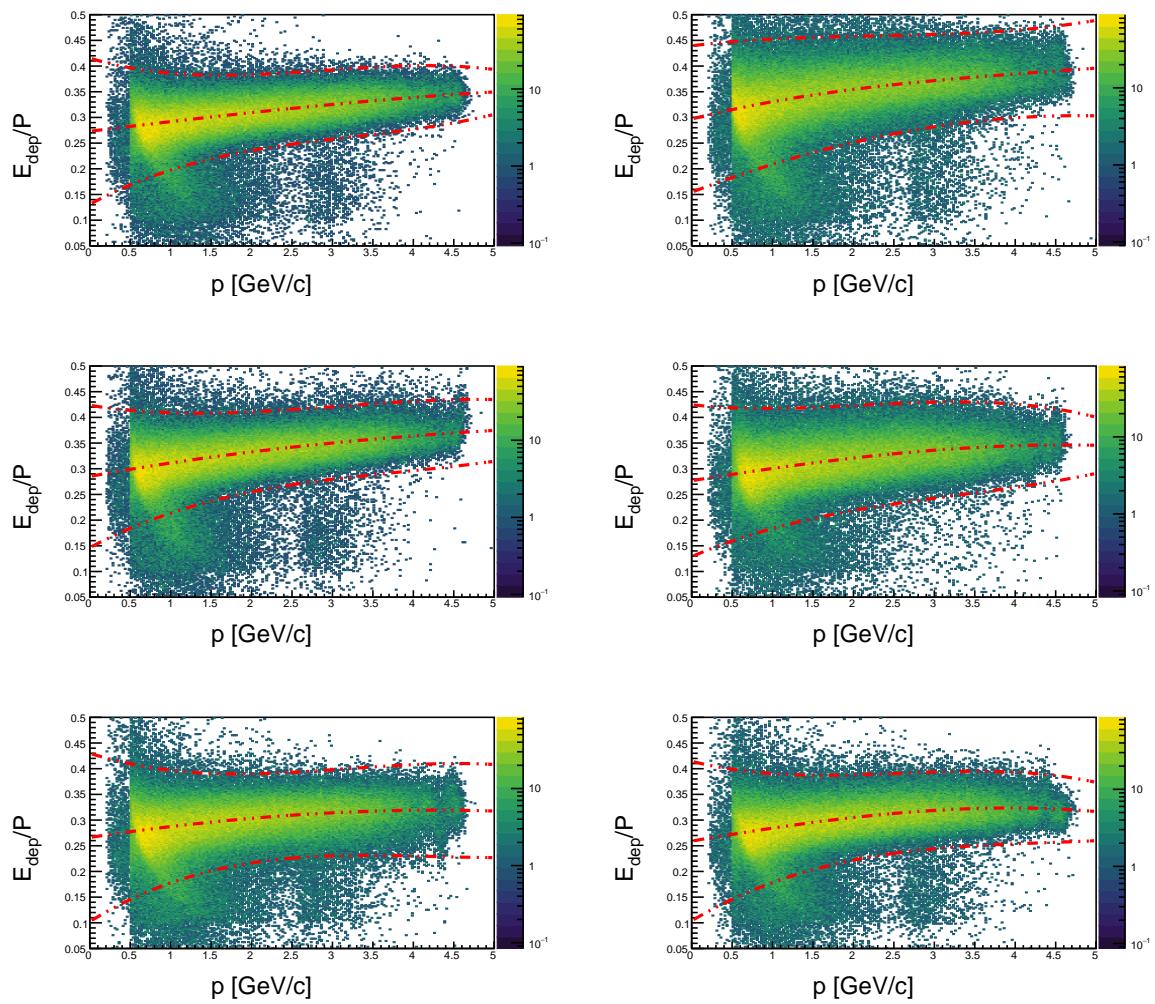


Figure 3.5: The sampling fraction selection boundary is shown here for the nominal value of $N_{\text{sigma}} = 4$.

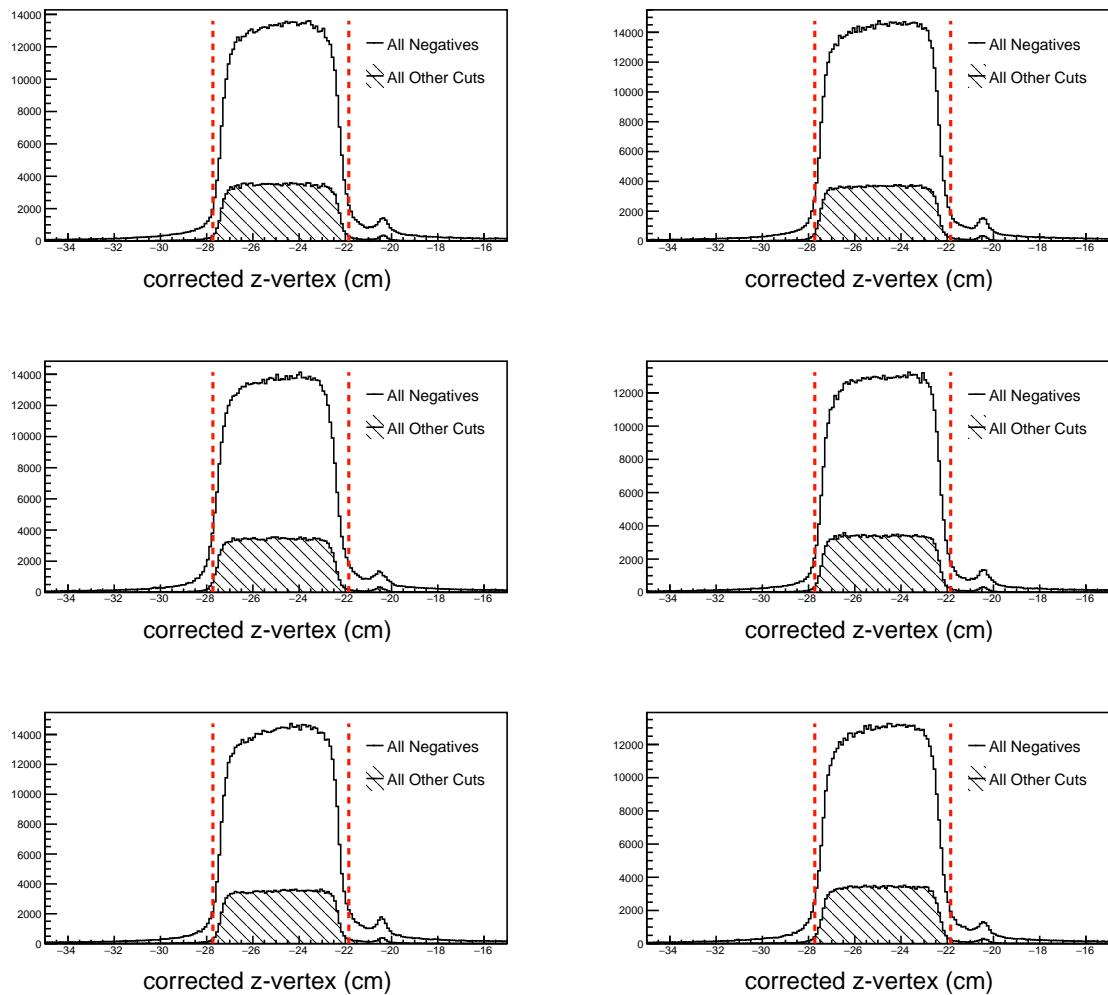


Figure 3.6: The track vertex cut is shown above. All negative tracks are shown in white, while the tracks passing all other criteria are shown in black hatch. The cut boundary is displayed as red lines. For E1-F the target center was located at -25 cm.

Cherenkov counter fiducial cuts

The value of the polar angle θ_{cc} and the azimuthal angle ϕ_{rel} with respect to the center of the sector is used for each event to define a fiducial cut at the Cherenkov counter. In this study, the functional form below is used for the boundary function.

$$\theta_{cc} > 46 - 35\sqrt{1 - \phi_{rel}^2/350} \quad (3.7)$$

The parameters were chosen empirically based on a study of the distributions with and without other electron identification cuts applied.

Cherenkov counter θ_{cc} and ϕ_{rel} matching to PMT

The angular arrangement of photo-multiplier tubes (PMTs) in the Cherenkov counter allows for additional consistency conditions to be applied. Each half-sector of the CC contains 18 PMTs increasing in polar angle away from the beamline, these divisions are known as segments. The polar angle measured at the Cherenkov counter θ_{cc} is then correlated to the segment in which the track was detected. Additionally, PMTs that are placed on the left and right of the detector can be used to check consistency with the azimuthal angle the track forms with the central line of the detector (ie $\phi_{rel} > 0$ means the track was in the right half of the sector, $\phi_{rel} < 0$ means the track was in the left half of the sector). An integer is used to describe the PMT associated with the track. The left PMT is assigned value -1, the right 1, and a signal in both PMTs is assigned 0. If both PMTs have a signal, the track is allowed to pass. If the left PMT was the one that had a signal, only events with $\phi_{rel} < 0$ passes. Similarly if the right PMT fired (code = 1), only events with $\phi_{rel} > 0$ are allowed to pass.

3.3 Inclusive Electron Cross Section

A brief discussion is provided here regarding our extraction of the inclusive electron scattering cross section in the resonance region. The inclusion of this study here is intended to add credibility to our electron identification and good run list. Inclusive electron scattering is the process $ep \rightarrow eX$, where only the final state electron is detected and the rest of the event is not (anything apart from the electron that is detected is not analyzed). As a function of W (the invariant mass of the final state ($\gamma^* + p$) system) the region below 2 GeV contains resonances and is often referred to as the resonance region. Resonance structures are difficult to detect higher than about 2 GeV, and this region is typically called the *deeply inelastic* region. While the deeply inelastic region is used extensively for measurements in nuclear/- particle physics, the goal of luminosity and electron identification verification is more easily achieved in the resonance region. This fact is due principally to the excess of Bethe-Heitler events which collect

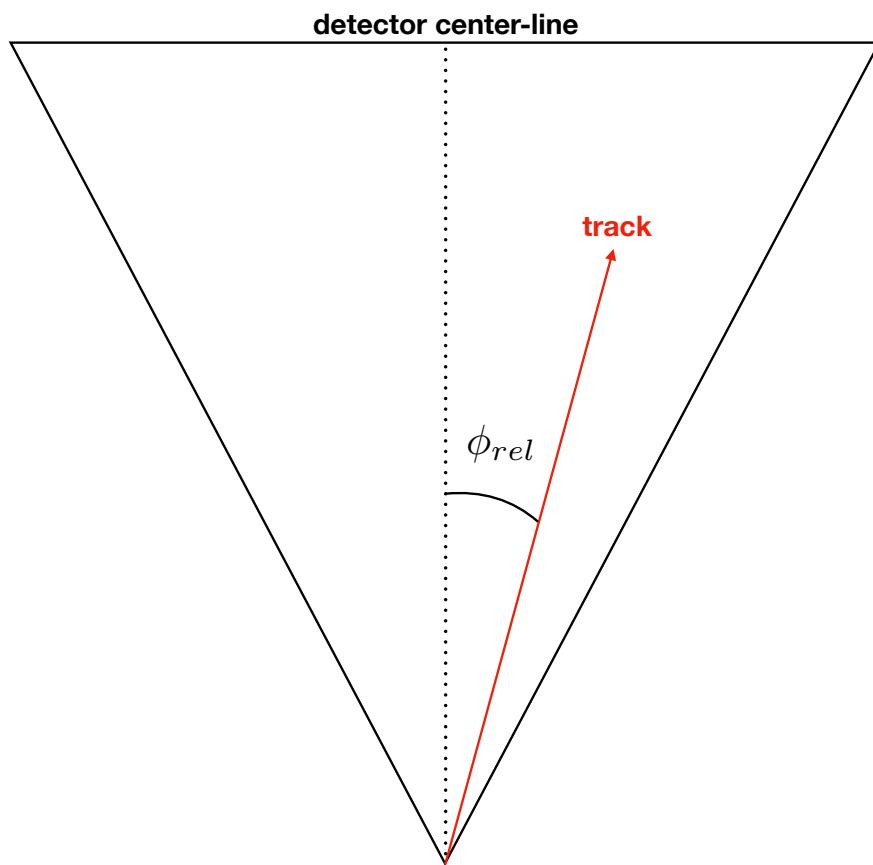


Figure 3.7: The angle ϕ_{rel} is the azimuthal angle between the central line of the detector and the track.

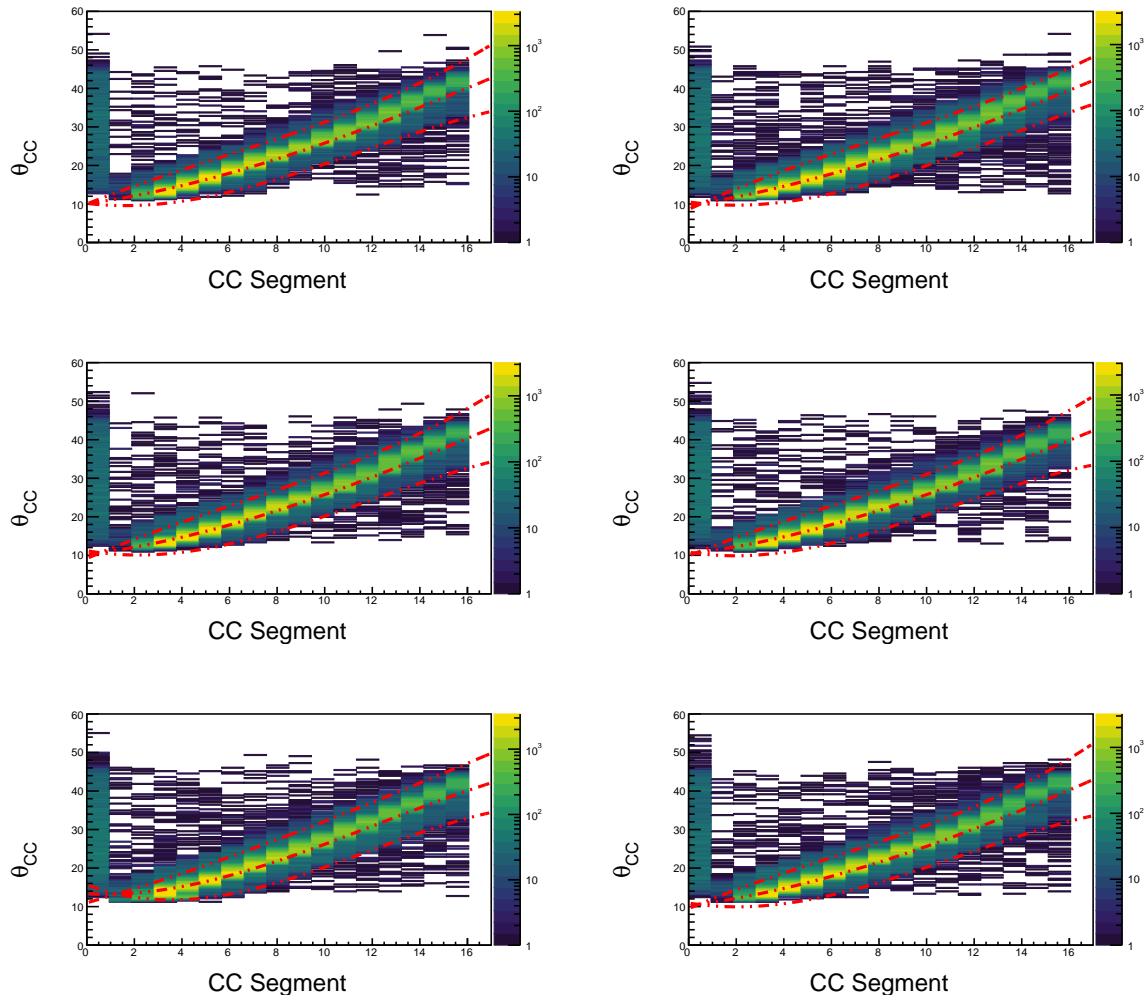


Figure 3.8: Correlation between θ_{CC} and the CC segment is shown above, with our selection boundaries overlaid in red.

Variable	N	Min.	Max	Width
W	40	1.1	2.1	0.25
Q^2	10	1.7	4.2	0.25

Table 3.4: Summary of W and Q^2 binning used for the inclusive cross section.

in the $2 < W < 3$ region for $E_{beam} = 5.498$ (such events are difficult to remove when detecting only the final state electron). The cross section is measured by counting events in each kinematic bin, and normalizing by the integrated luminosity over files analyzed. Additionally, acceptance and radiative corrections are applied bin-by-bin as shown in 3.8.

$$\frac{d^2\sigma}{dW dQ^2} = \frac{1}{\Delta W \Delta Q^2 \mathcal{L}} \frac{N_i}{A_i R_i} \quad (3.8)$$

Here A_i and R_i refer to the acceptance and radiative corrections for the bin i , and \mathcal{L} is the integrated luminosity.

Event Selection and Binning

A simple choice of 10 bins in Q^2 and 40 bins in W is used. This choice is mainly driven by the desire to keep bin migration effects small. Events are generated and reconstructed in some bins $R^{(j)}$ and $G^{(i)}$ respectively. Due to finite detector resolution, it is not always the case that $i = j$. This effect is known as bin migration, and negatively impacts the acceptance calculation.

The only kinematic restriction that is imposed is applied to the *inelasticity* $y = 1 - E'/E < 0.7$. This restriction is applied because events with large- y have a significantly higher probability to be Bethe-Heitler events. This cut is equivalent to enforcing a minimum energy for the scattered electron.

$$E_{min} = E_{beam}(1 - y_{max}) \approx 1.6 \text{ GeV} \quad (3.9)$$

Results

All sectors individually were analyzed and our results are consistent with the model parametrization by Cynthia Keppel 3.3. In most of the kinematics that we studied, our extracted cross section is within 5 % of the model prediction. The most notable difficulty that appeared was the residual elastic tail at lower W values that caused some kinematic bins to be slightly off. There also exist bin migration effects on the resonance peaks, but these are small.

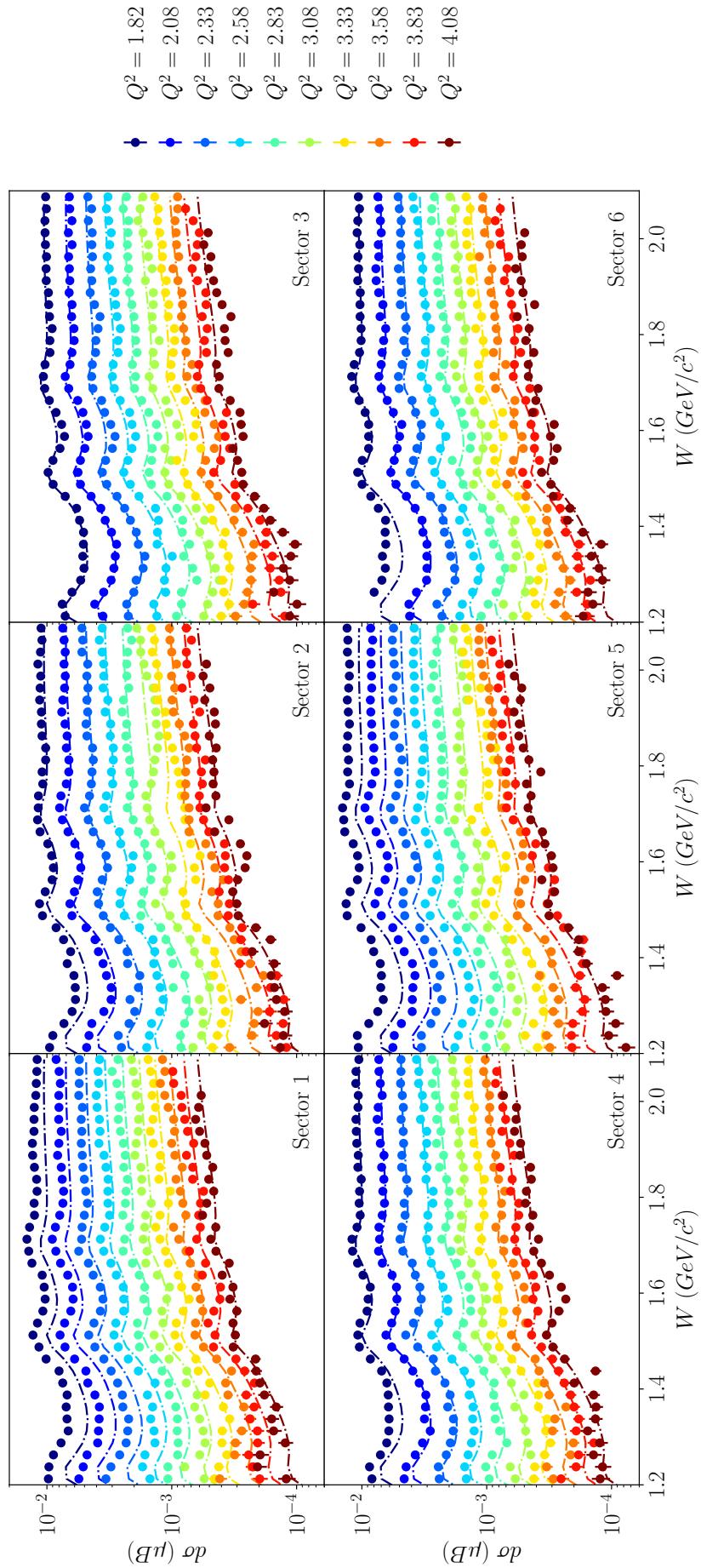


Figure 3.9: Inclusive cross section results compared with model predictions. Each panel shows the result for all 10 Q^2 bins and one sector.

3.4 Hadron Identification

Hadron identification in CLAS is done by correlating particle momentum from the drift chambers with timing information supplied by the time of flight detector. In this analysis some quality assurance cuts are applied preliminarily, but they do not discriminate between different species of particle. The likelihood methodology described in this section is based on the discussion provided by the BES collaboration in [12].

3.4.1 Hadron ID Cuts

The cuts used for hadron classification are enumerated below.

- Drift chamber fiducial
- Hadron-electron vertex difference
- Likelihood maximization of $\beta(p, h)$

Drift chamber fiducial

Drift chamber fiducial cuts are applied (only region 1) using the same procedure as described for electrons. The parameters used for positive tracks are $h = 10, \theta = 60$.

Hadron-electron vertex difference

The distance between the electron vertex and the hadron candidate track vertex is computed ($\delta v_z = v_z^e - v_z^+$). This distance is constrained to be within the length of the target (5 cm) see figure 3.4.1. For events where the electron-kaon vertex difference is larger than the target size, we cannot assert that the kaon came from the electron-proton collision. Although the number of events excluded by this cut is not large, those events are considered to be outliers.

Likelihood maximization of $\beta(p, h)$

In this section, positive hadrons are used as an example. For each particle species considered, a normalized probability density function $P(x; p, h)$ is constructed for each input into the likelihood analysis. Here, x corresponds to the feature being used to categorize different particles (in our case, x is the β value measured by CLAS time-of-flight), p is the particle momentum, and h is the label for the hadron being hypothesized (eg: the possible values for positive hadrons are pion, kaon, proton). In general if one uses a set of N variables $x = (x_1, x_2, \dots, x_N)$, the likelihood for a hypothesis h is defined below.

$$\mathcal{L}_h = \prod_{i=1}^N P_i(x_i; p, h) \quad (3.10)$$

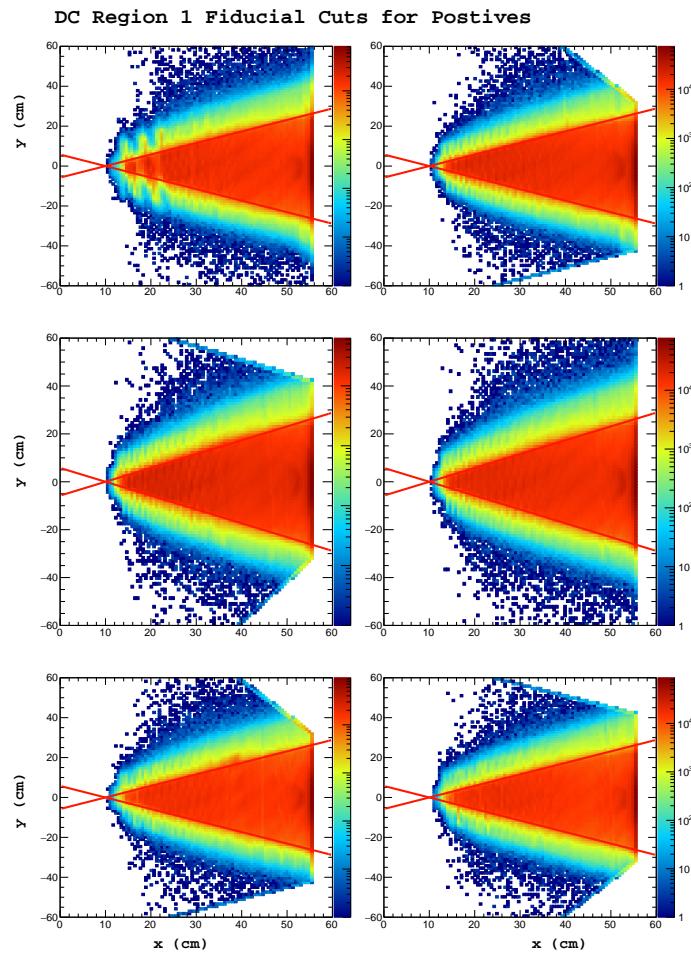


Figure 3.10: Positive track hits on the region 1 drift chamber, events falling between the red lines are kept for analysis.

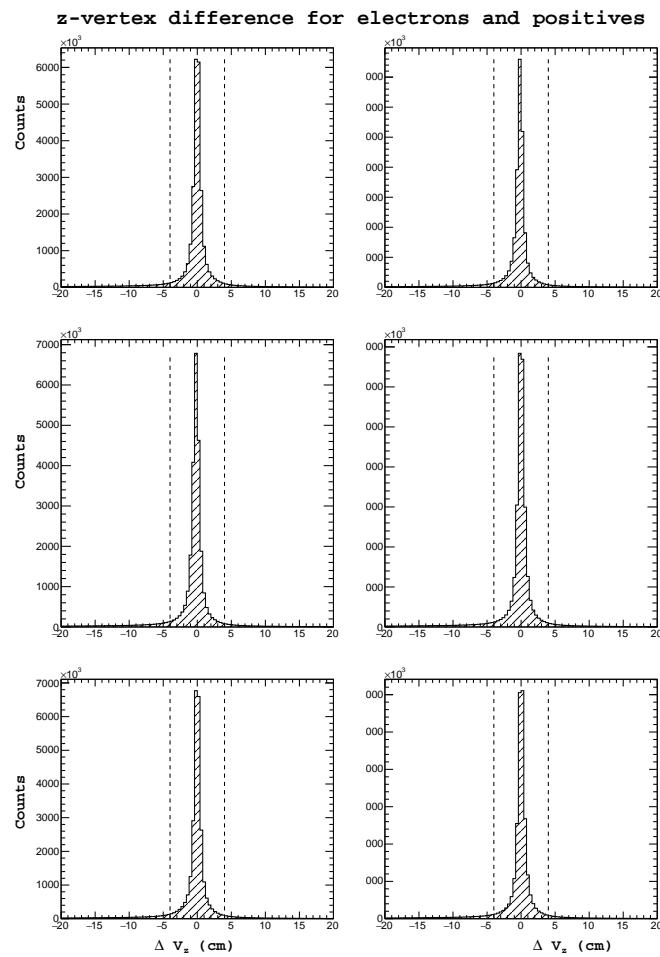


Figure 3.11: The difference between the z-vertex position between detected electrons and positive tracks.

In our case, the only random variable we consider is β , and the likelihood is just the PDF. Here, and in many cases where the choice is statistically appropriate, it is possible to use a Gaussian PDF for the variable x_i (here β). The variable $\sigma_\beta(p, h)$ is the standard deviation of the PDF at a given value of momentum and for a given hadronic species, similarly μ_β is the mean of the Gaussian PDF for the same value of momentum and same hadron.

$$P(\beta; p, h) = \frac{1}{\sqrt{2\pi}\sigma_\beta(p, h)} \exp \left\{ -\frac{1}{2} \left(\frac{\beta - \mu_\beta(p, h)}{\sigma_\beta(p, h)} \right)^2 \right\} \quad (3.11)$$

The identity is assigned by choosing the particle hypothesis h which maximizes the likelihood ratio.

$$\frac{\mathcal{L}_h}{\mathcal{L}_\pi + \mathcal{L}_K + \mathcal{L}_p} \quad (3.12)$$

Using this method, every positive track is assigned a particle identification. However, at times the likelihood value is quite small when compared with the maximum likelihood for that species. This is the case for positrons which are classified by this method as positive pions, because they are the closest particle for which a hypothesis has been provided. To avoid these situations, the confidence level α of each track is calculated and a cut is applied on the minimum confidence. This cut can be easily varied to see how it changes the analysis result.

$$\alpha = 1 - \int_{\mu - \beta_{obs}}^{\mu + \beta_{obs}} P(\beta; p, h) d\beta \quad (3.13)$$

This quantity represents the probability to observe a value of β as far or farther from the mean as β_{obs} . Confidence levels close to zero correspond to tracks which are poorly identified as the class h . In the case that the PDF is Gaussian, the standard 1, 2, and 3 σ cuts on β vs. p can be understood simply as confidence levels of approximately $0.32 = 1 - 0.68$, $0.05 = 1 - 0.95$, and $0.01 = 1 - 0.99$.

Determination of probability density functions for likelihood method

The most important and most difficult part of constructing the likelihood ratio identification is the determination of the mean and standard deviation of the probability density function (which depends on momentum) for the different hypotheses. In the case where exceptionally accurate monte carlo (MC) simulations of the detector are available, one can use the truth information and track matching to construct the β vs. p 2-dimensional histograms, and fit the $\mu(p)$ and $\sigma(p)$. In the absence of high quality MC, analysts typically fit directly the spectrum of β vs. p and extract the mean and variance. In this work, an enhanced sample of candidates for each of the three positive particles in question is created before doing the fitting. In this way, we hope that our fit better represents the true μ and σ for each particle. For fitting of pion and proton resolutions, positive tracks are assumed to be pions and

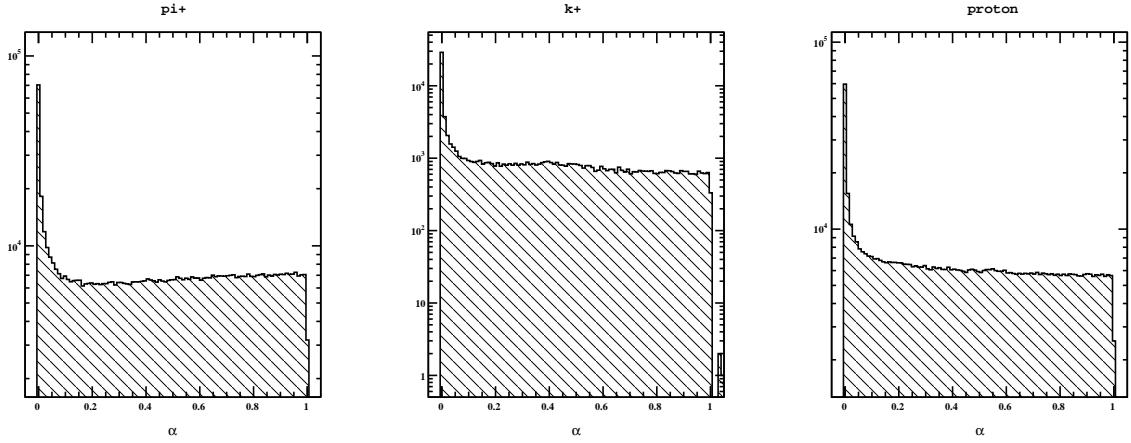


Figure 3.12: Shown above: The distribution of confidence level for all positive tracks after being classified by the likelihood ratio.

the missing mass of the event is calculated. Then, a cut is placed around the neutron mass. In doing so, two main exclusive reactions are selected. The first is $ep \rightarrow e\pi^+ N$, and the second is $ep \rightarrow ep\pi^0$. In this way most positrons, and positive kaons are removed from the sample prior to fitting. The mean and variance are fit using a third order polynomial in p (MINUIT χ^2 minimization is used). For accessing the kaon distribution, the opposite condition is applied. This is not nearly as effective as the pion case, but we believe is still an improvement over fitting the spectrum directly.

The parametrization used for the mean $\mu(p, h)$ and resolutions $\sigma(p, h)$ are shown below.

$$\mu(p, h) = \mu_{theory} + \Delta\mu \quad (3.14)$$

$$\mu_{theory} = \frac{1}{\sqrt{1 + (m_h/p)^2}} \quad (3.15)$$

$$\Delta\mu = \mu_0 + \mu_1 p + \mu_2 p^2 \quad (3.16)$$

$$\sigma(p, h) = \sigma_0 + \sigma_1 p + \sigma_2 p^2 \quad (3.17)$$

The values are displayed in the table below.

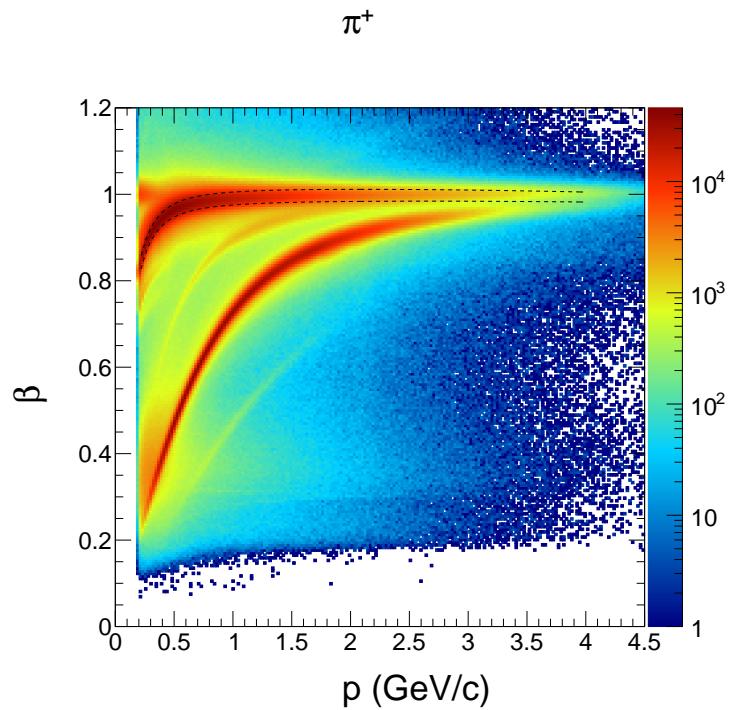


Figure 3.13: All positive tracks overlaid with our determination of $\mu(p) \pm \sigma(p)$ for π^+ (black dashed lines).

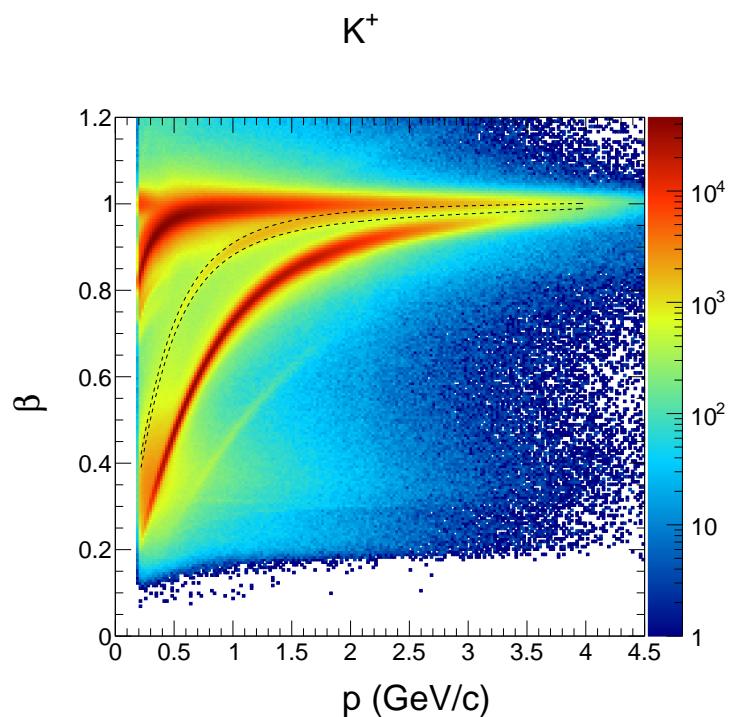


Figure 3.14: All positive tracks overlaid with our determination of $\mu(p) \pm \sigma(p)$ for K^+ (black dashed lines).

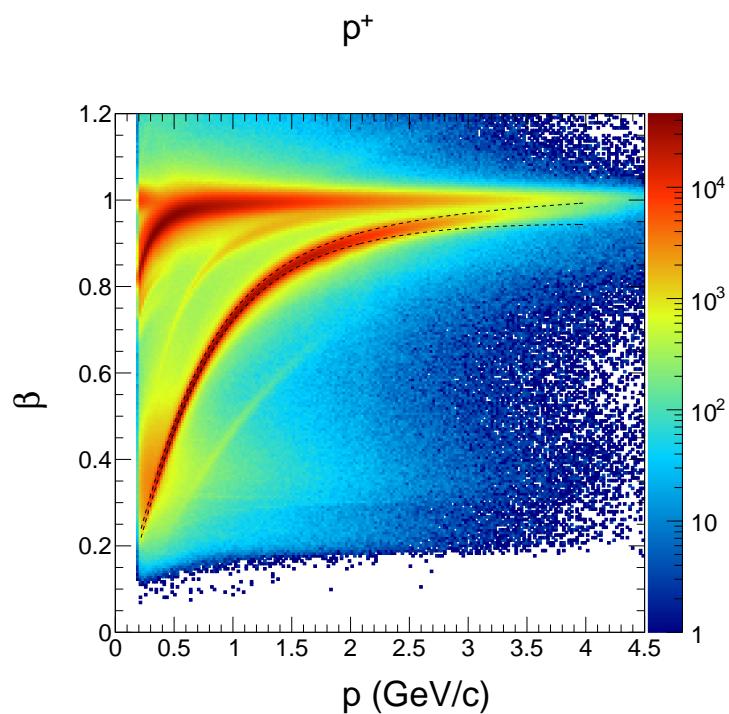


Figure 3.15: All positive tracks overlaid with our determination of $\mu(p) \pm \sigma(p)$ for p^+ (black dashed lines).

Hadron	Parameter	Sector 1	Sector 2	Sector 3	Sector 4	Sector 5	Sector 6
K^+	μ_2	0.00111554	-8.97687e-05	4.78796e-05	0.000376425	-0.00204856	0.000652209
K^+	μ_1	-0.00468038	6.19414e-05	-0.00081741	-0.00107931	0.00629181	-0.00264143
K^+	μ_0	0.00361012	0.00134921	0.00299674	0.00220194	0.000117821	0.00162582
K^+	σ_2	-0.000331838	-0.00105807	-0.000712404	-0.000573934	-0.000259289	0.000508389
K^+	σ_1	-0.00105857	0.00236686	0.000509169	0.000163467	-0.00233617	-0.00461598
K^+	σ_0	0.0154964	0.0117702	0.0140748	0.0143761	0.0184055	0.0180945
π^+	μ_2	-0.000962041	-0.000300602	-0.000306326	-3.2245e-05	-0.00226511	-0.000330818
π^+	μ_1	0.00296349	0.0016512	0.0021962	0.00176045	0.00750862	0.00126443
π^+	μ_0	-0.00225794	-0.00047045	0.000370406	0.000435526	-0.000449409	-0.00131045
π^+	σ_2	-0.000127659	0.000691895	-0.000289961	0.000315041	-0.000936521	-0.000131269
π^+	σ_1	-0.000489092	-0.0033948	0.00196853	-0.00197841	0.00212778	-0.000339411
π^+	σ_0	0.0155195	0.0167998	0.0124066	0.0157476	0.0145571	0.0141728
p^+	μ_2	-0.00039358	-0.000701003	-0.000347651	0.0004854	-0.00121666	0.000563786
p^+	μ_1	-0.000295423	0.00170899	0.000794901	-0.000744446	0.00376887	-0.00353545
p^+	μ_0	0.00227353	0.00231676	0.00364672	0.00276859	0.00128827	0.00439605
p^+	σ_2	0.001429	0.00144256	0.00124456	0.00190709	0.00141039	0.0011516
p^+	σ_1	-0.0021472	-0.00262226	-0.00196308	-0.00385218	-0.00186708	-0.00186749
p^+	σ_0	0.0107541	0.0109091	0.0104381	0.0115449	0.0109969	0.0107759

Table 3.5: Values used to calculate the mean and resolution for hadron likelihood based identification.

Validation of Kaon Identification

A Monte Carlo simulation was used to study particle identification of positive hadrons as a function of the hadronic momentum. As a result of this study the minimum confidence level ($\alpha = 0.55$) and maximum momentum ($p_{max} = 2.0 \text{ GeV}/c$) for K^+ were determined.

To study this SIDIS events were generated using `clasdis`. Our simulation includes π^+ , K^+ and protons (here denoted P^+) over a range of kinematics consistent with the E1-F beam energy of $E_{beam} = 5.498 \text{ GeV}$. After passing these events through the CLAS detector simulation `GSIM`, reconstruction was used to fit tracks. The truth information for generated kinematics was stored in the output files, and we correlated reconstructed particles to their generated counter-part by requiring that magnitude of the three momentum difference was small ($\delta P < 0.05$).

$$\delta P = \frac{|\vec{P}_{gen} - \vec{P}_{rec}|}{|\vec{P}_{gen}|} \quad (3.18)$$

In this simple equation P_{rec} and P_{gen} are the reconstructed and generated three momentum of the track being matched. After this matching procedure has been applied, we calculate two simple metrics, the *purity* and the *efficiency*. The purity refers to the fraction of tracks that are classified as kaons that are truly kaons. More formally it is written,

$$P = \frac{tp}{tp + fp} = \frac{N_{K^+}}{N_{K^+} + N_{\pi^+} + N_{P^+}} \quad (3.19)$$

where tp and fp are true positives and false positives respectively, and the variables N refer to the true number of tracks with that label in the sample. The efficiency is simply the fraction of all true kaons which are identified as kaons, and can be written as shown below.

$$\epsilon = \frac{tp}{tp + fn} = \frac{N_{K^+ \text{ identified}}}{N_{K^+}^{total}} \quad (3.20)$$

As a simple illustrative example, consider the case when all hadrons are called kaons, in this case the efficiency is 1, but the purity will be at its minimum (related to the fraction of total particles that are kaons). As the purity of the sample is increased, the efficiency drops. For this study the efficiency drives our statistical errors, but the purity is the more important metric. As the cut boundaries are restricted by raising the minimum confidence level of hadrons identified as kaons, the purity goes up and the efficiency goes down 3.4.1.

Based on this study, a maximum momentum of 2.0 GeV and a minimum confidence α_{min} of 0.55 is required for all kaons in our analysis, which provides a purity of 80% or more (depending on the kinematics). The magnitude of the π^+ asymmetry is known in these kinematics to be on the order of

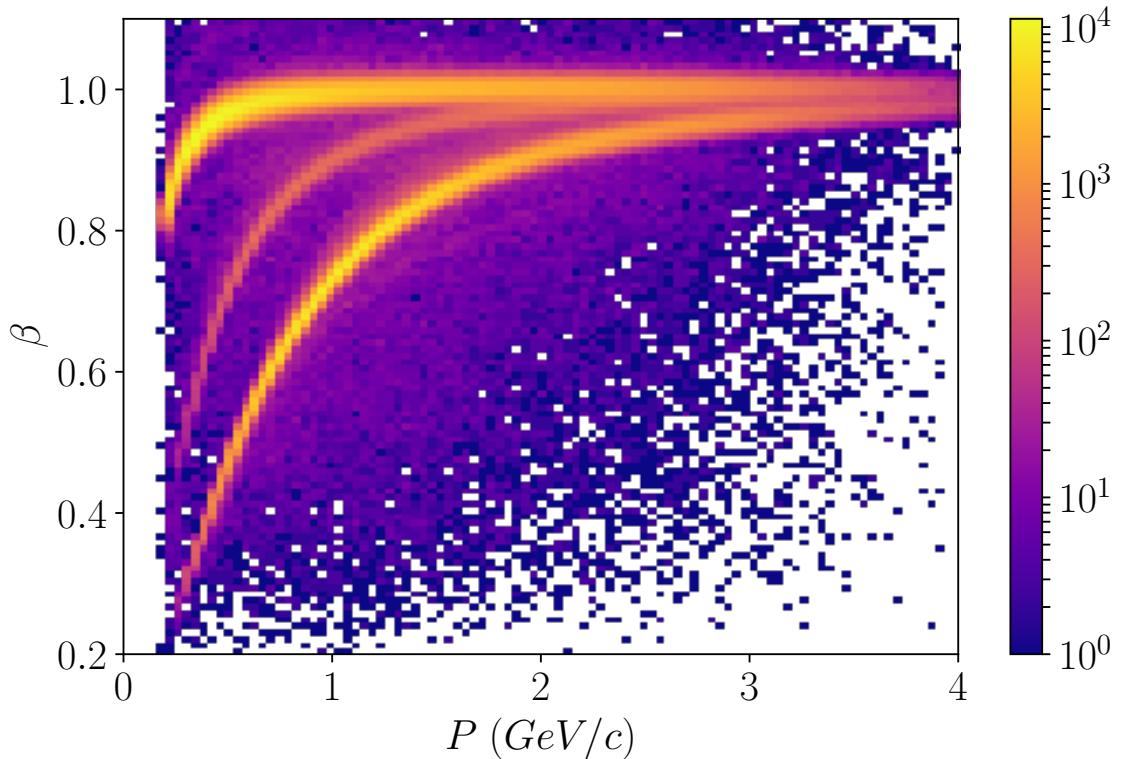


Figure 3.16: Positive hadrons from the Monte Carlo simulation produce a β vs. p simulation that is very similar to data.

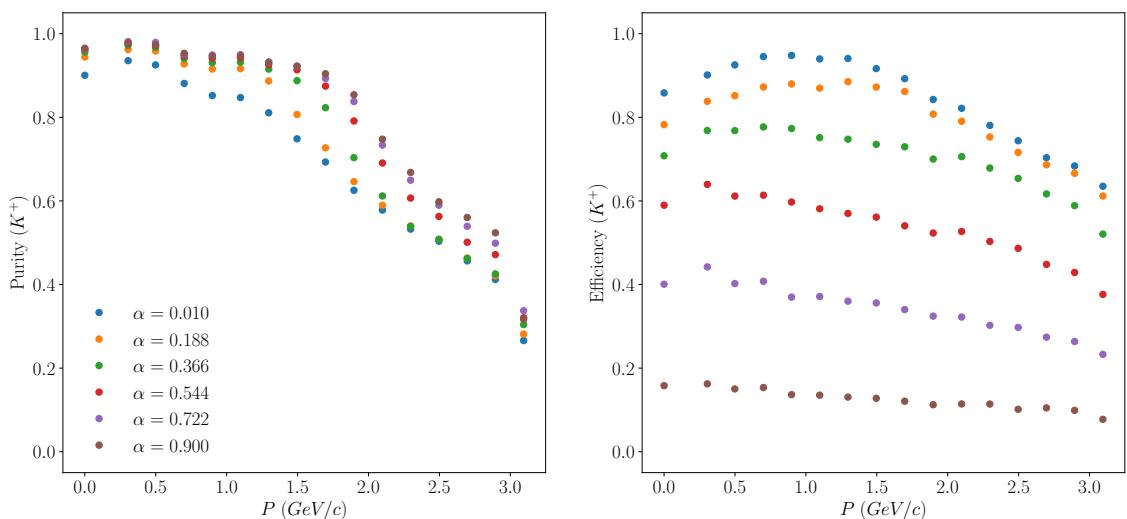


Figure 3.17: The efficiency and purity of our kaon sample are studied by using a Monte Carlo simulation. Here, the results are studied as a function of the confidence level, and of the track momentum.

0.02, if the sample is comprised of 20% pions (which is the worse case in our measurement) then the contribution to the total asymmetry is equivalent to a K^+ asymmetry of 0.005, which is much smaller than our errors. This level of contamination is therefore very tolerable, and should have no significant impact on our analysis.

Chapter 4

Beam Spin Asymmetry Analysis

4.1 Introduction

Measurement of the beam spin asymmetry is carried out for the positively charged K-meson. As discussed in the introduction, the beam spin asymmetry theoretically depends on $F_{UU,L}$, $F_{UU,T}$, $F_{UU}^{\cos\phi}$, $F_{UU}^{\cos 2\phi}$, and $F_{LU}^{\sin\phi}$. By dividing the electron-kaon events into several bins, beam spin asymmetry measurements are taken at different average values of the kinematic variables x , Q^2 , z_h , and P_T . Finally, the structure function ratios $A_{LU}^{\sin\phi}$, $A_{UU}^{\cos\phi}$, and $A_{UU}^{\cos 2\phi}$ are extracted from each bin. In this chapter a discussion is provided of SIDIS event selection, the binning used in this analysis, measurement values with associated systematic uncertainties, and the extraction of structure function ratios using the ϕ_h dependence in each kinematic bin.

4.2 Event Selection and Binning

Event Selection

After particle identification, events which have a trigger electron and a positive kaon are kept for analysis. Events are discarded that do not have $W > 2 \text{ GeV}/c^2$ and $Q^2 > 1 \text{ GeV}^2/c^2$, because they are not considered part of the deeply inelastic region. Additionally, to avoid exclusive resonances in the $ep \rightarrow eK^+X$ spectrum, a minimum value is imposed on the missing mass of the final state M_X ($ep \rightarrow eK^+X$). Here, we use $M_X(ep \rightarrow eK^+X) > 1.25 \text{ GeV}/c^2$. Finally, a cut is applied to exclude low values of z to constrain our kinematics to the current factorization region where TMD factorization has been demonstrated at leading order. It is additionally required that $z < 0.75$ to avoid exclusive events. This restriction on z is applied when z is integrated over (for the axes x , Q^2 , and P_T) but not to the z axis itself, where we measure the asymmetry across the entire experimentally observed kinematic range.

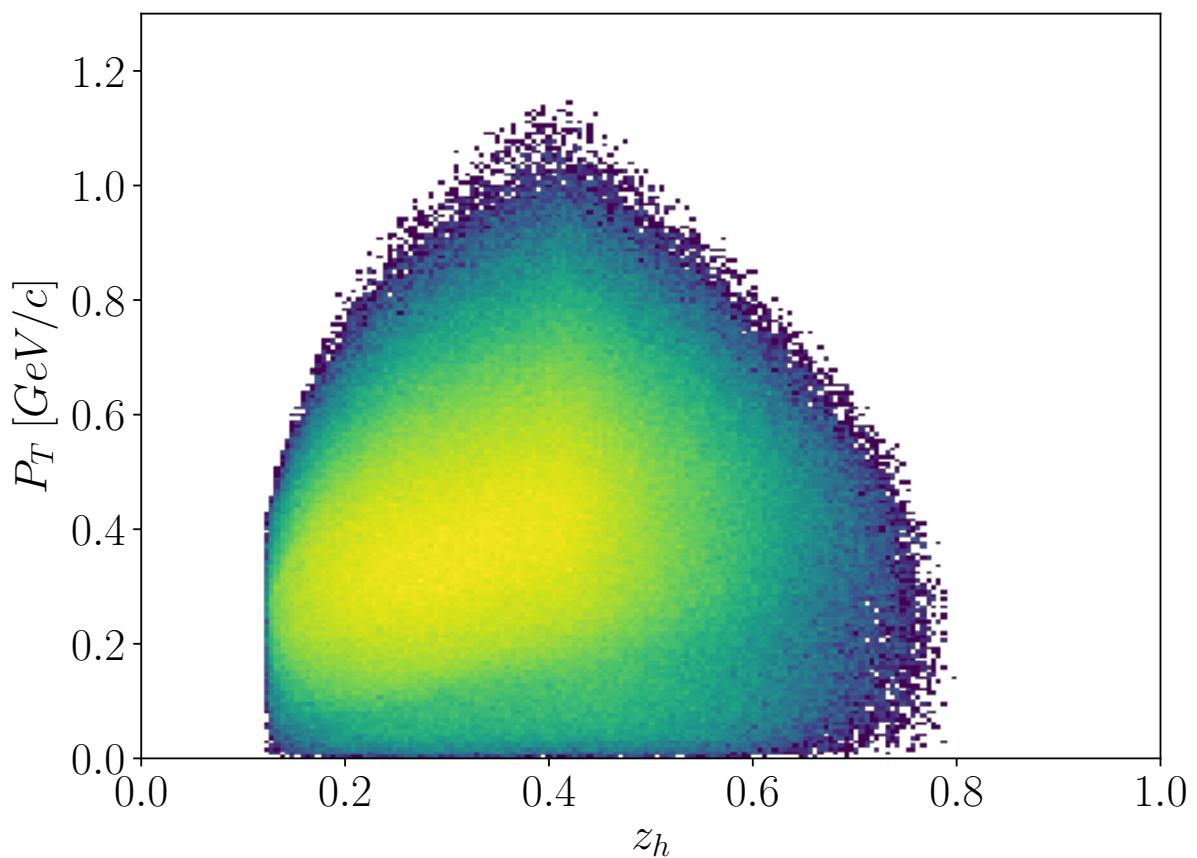


Figure 4.1: Correlation between z and P_T^2 (GeV^2/c^2) for each event in our analysis sample.

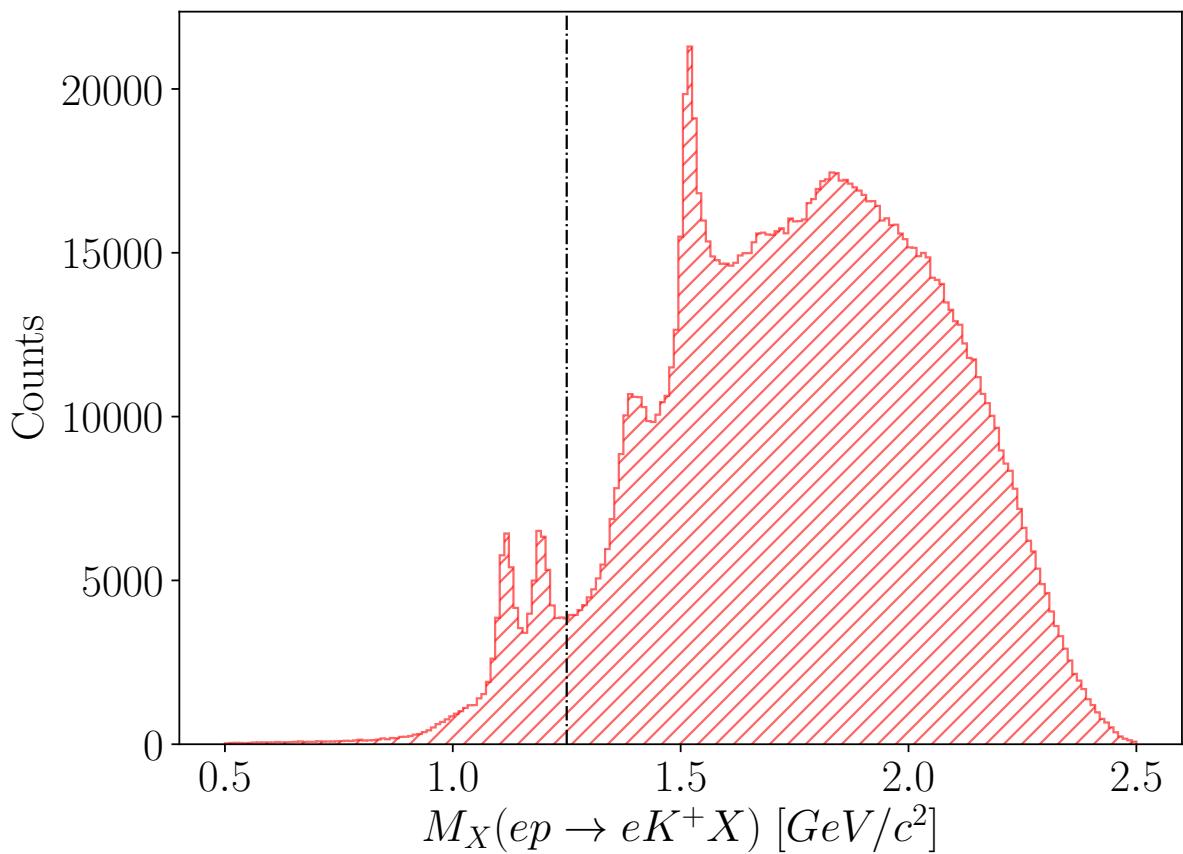


Figure 4.2: The missing mass spectrum for the reaction $ep \rightarrow e'K^+X$ is shown after the application of all cuts used in the analyses except for the cut we apply on the missing mass.

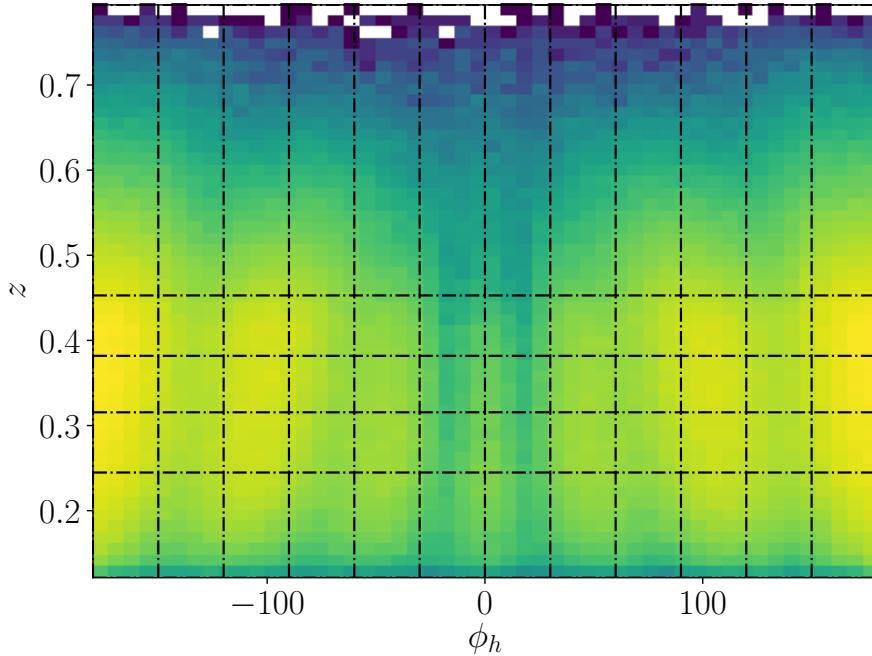


Figure 4.3: Bins used for this analysis are displayed in two dimensions for the z axis.

Binning

The beam spin asymmetry measurement is performed for the kinematic variables x , Q^2 , z , and P_T . For each variable 5 bins are chosen, as well as 12 bins in ϕ for a total of 60 analysis bins.

Bins were chosen using a simple method to ensure equal statistics in each kinematic variable bin (the phi bins do not have equal statistics). The procedure is described using the axis x as an example. First, all events are sorted by their x value from smallest to largest. Then, the smallest and largest values are recorded, which are x_1 and x_N if there are N events in the sample. Next, the target number of bins M is chosen (this choice depends on each analysis). Finally, the limits of each bin can be chosen by calculating the number of events per bin N/M and then using the value of x which corresponds to multiples of N/M in the sample.

$$\vec{b} = (x_1, x_{N/M}, x_{2N/M}, \dots, x_N) \quad (4.1)$$

Here, the symbol \vec{b} denotes a vector of $(M+1)$ x values which represent bin limits. The binning in ϕ is chosen to be regularly spaced between -180 and 180 degrees.

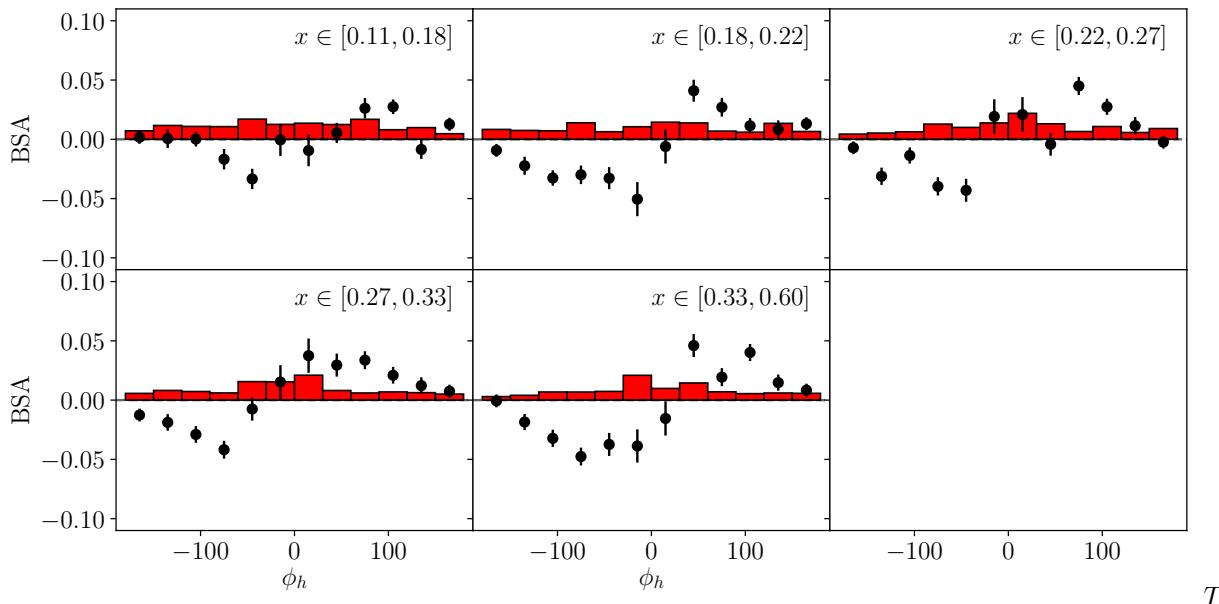
4.3 Measured ϕ_h Distributions

Measured Asymmetry Values

In each bin i the beam spin asymmetry (here A_i) is calculated according to,

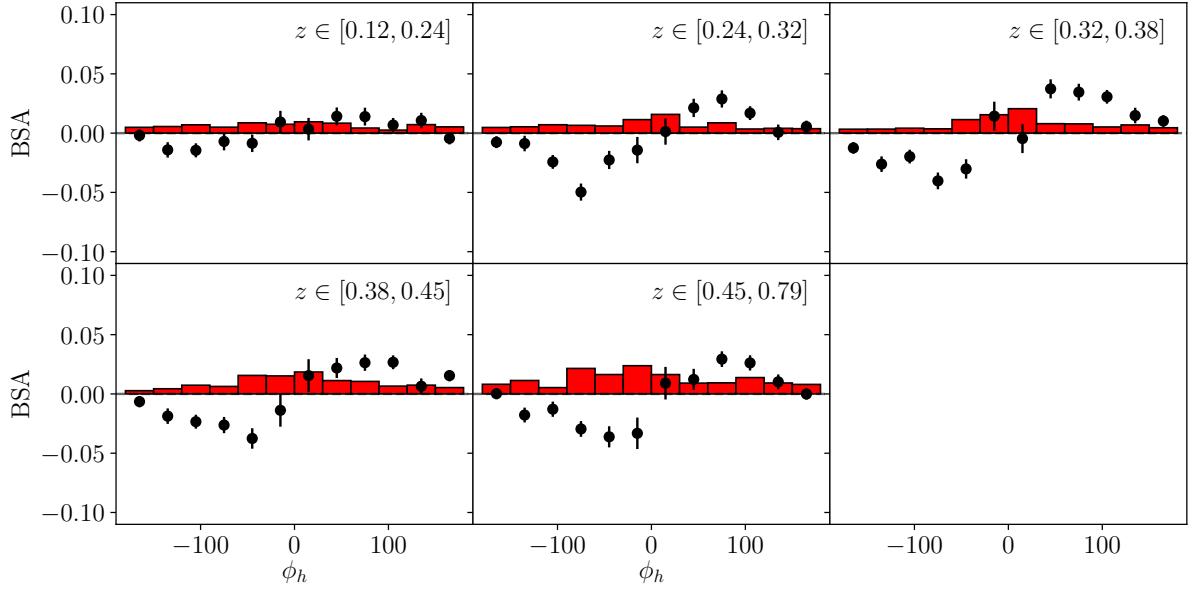
$$A_i = \frac{1}{P_e} \frac{n_+^i - n_-^i}{n_+^i + n_-^i} \quad (4.2)$$

where P_e is the average beam polarization over the dataset (74.9%). The symbols n_\pm^i refer to the number of events counted in bin i with helicity \pm .



The

ϕ_h dependence is shown for each bin of x , increasing in value from the top left to the bottom right. The statistical uncertainty is shown as black error bars on each point. The total systematic uncertainty is shown as a red bar centered at zero. (4.3)



The

ϕ_h dependence is shown for each bin of z , increasing in value from the top left to the bottom right. The statistical uncertainty is shown as black error bars on each point. The total systematic uncertainty is shown as a red bar centered at zero. (4.4)

Statistical Uncertainties

The uncertainty on the measured value of A_i can be attributed to statistical uncertainty on the counts n_i^\pm , and the uncertainty associated with the measurement of P_e . The statistical uncertainty reported on the measurement includes the contribution from counts, but not from the uncertainty in P_e which is included in the systematic errors. In general, the uncertainty in a measured observable \mathcal{O} depends on the uncertainty of the parameters (here denoted by $\vec{\theta}$) used in the analysis in the following way (see appendix for derivation).

$$\sigma_{\mathcal{O}}^2 = \sum_{i=1}^N \sum_{j=1}^N \frac{\partial \mathcal{O}}{\partial \theta_i} \frac{\partial \mathcal{O}}{\partial \theta_j} \rho_{ij} \sigma_i \sigma_j \quad (4.5)$$

For the beam spin asymmetry in the i^{th} bin A_i one finds that without correlations ($\rho_{ij} = \delta_{ij}$) the error propagation proceeds as shown below.

$$\sigma_A^2 = \frac{A^2}{P_e^2} \sigma_{P_e}^2 + \frac{4(n_-^2 \sigma_+^2 + n_+^2 \sigma_-^2)}{P_e^2 (n_+ + n_-)^4} \quad (4.6)$$

The first term which is the contribution from the variance in the measurements of beam polarization

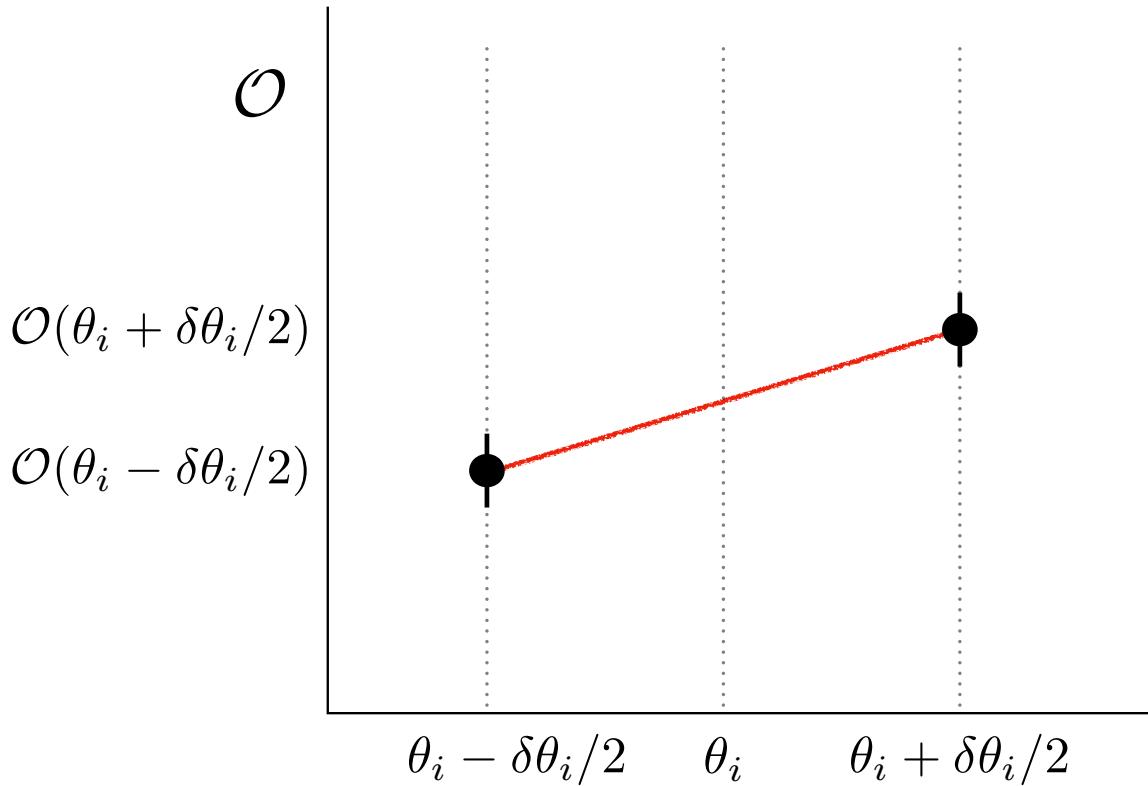


Figure 4.4: The analysis is run for variations in the input parameters θ_i to calculate the dependence of the result \mathcal{O} on each parameter, as described in this section.

will be included as a systematic error. The second term is used as the statistical error bars shown through the analysis. The counts n_{\pm}^i for the i^{th} bin are assumed to be Poisson in nature, and therefore have a variance equal to the expected number of counts $\sigma_{\pm}^2 = n_{\pm}^i$. With this expression for the statistical uncertainty on the counts, and dropping the beam polarization term for now, the expression becomes:

$$\sigma_A^2 = \frac{4n_+n_-}{P_e^2(n_+ + n_-)^3} \quad (4.7)$$

Systematic Uncertainties

Basic Formalism

Systematic effects are shifts or biases in the measured result of some observable as a result of the procedure used in the measurement. Systematic effects can typically be identified and corrected for, or removed all together from the measurement. In the cases where an effect cannot be completely removed, the degree to which the correction for the effect is uncertain is included in the result of the measurement as a systematic uncertainty [13].

Systematic uncertainties are included using the standard equation for error propagation. In some cases it is possible to analytically find the derivatives needed to calculate the dependence of the observable on a source of systematic uncertainty. This is the case for effect of the variance of the beam polarization on the beam spin asymmetry observable. However in many cases, it is not possible to analytically calculate the effect of an analysis parameter θ_i on the observable \mathcal{O} . Since the observable is usually calculated using some computational chain which starts with the input parameters $\vec{\theta}$, it is possible to find the dependence of the observable \mathcal{O} on the inputs numerically.

$$\frac{\partial \mathcal{O}}{\partial \theta_i} \approx \frac{\mathcal{O}(\theta_i + \sigma_{\theta_i}/2) - \mathcal{O}(\theta_i - \sigma_{\theta_i}/2)}{\sigma_{\theta_i}} \quad (4.8)$$

After inserting the above into equation 4.5 one finds,

$$\sigma_{\mathcal{O}}^2 = \sum_{i=1}^n \sum_{j=1}^n \rho_{ij} (\mathcal{O}(\theta_i + \sigma_{\theta_i}/2) - \mathcal{O}(\theta_i - \sigma_{\theta_i}/2)) (\mathcal{O}(\theta_j + \sigma_{\theta_j}/2) - \mathcal{O}(\theta_j - \sigma_{\theta_j}/2)) \quad (4.9)$$

where ρ_{ij} is the correlation $V_{ij}/\sigma_i \sigma_j$. In most cases, these correlations are assumed to be zero. In some cases, when the parameters θ_i, θ_j come from a fit one may have a correlation provided by the covariance matrix and it should be used. In the case where correlations are assumed to be zero, the total systematic uncertainty is simply the quadratic sum of the shifts in the observable within the uncertainty window on each parameter.

$$\sigma_{\mathcal{O}}^2 = \sum_{i=1}^n \left[\mathcal{O}(\theta_i + \sigma_{\theta_i}/2) - \mathcal{O}(\theta_i - \sigma_{\theta_i}/2) \right]^2 \quad (4.10)$$

Sources of Systematic Uncertainty

Systematic uncertainties are calculated using the techniques described above for both the ϕ_h dependent asymmetry measurement as well as the results of the parameter estimation for each kinematic bin. The systematic errors on the phi dependent asymmetry bins is not used in the parameter estimation for the structure function ratios A . Table 4.1 below summarizes the sources of systematic uncertainty considered in this analysis.

Except for the beam polarization and the momentum of the kaon track, all parameters listed in the table are treated using the formalism outlined above. The beam polarization uncertainty quoted at 2.4% contains contributions from the standard deviation of the Moller polarimetry measurements (0.2%), residual target polarization effects (1.4%), and atomic motion/finite acceptance corrections (0.8%).

Source	Variation
Beam polarization	0.024
DC Region 1 Fid.	1 (cm)
DC Region 3 Fid.	3 (cm)
EC-W	12 (cm)
EC-V	12 (cm)
EC-U	12 (cm)
Kaon Confidence (α)	0.5-0.6
θ_{cc} Matching	σ
EC Energy Deposition	0.01 (GeV)
p_{K^+}	1.9-2.1
EC Sampling Fraction	0.5σ
Z-Vertex	0.5 (cm)
Vertex diff.	1 (cm)

Table 4.1: Different sources of systematic effect considered in this analysis. The magnitude of the effect is shown here averaged over all bins of ϕ_h .

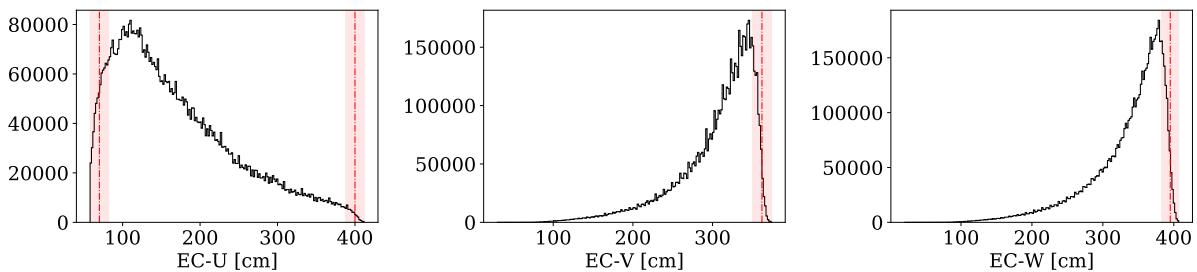


Figure 4.5: Boundaries and associated uncertainties for the *EC* coordinate cuts.

Electromagnetic Calorimeter Fiducial Cuts

The electron identification cuts used on ECAL are varied in order to estimate the dependence of the asymmetry on these parameters. These boundaries (which are sometimes excluded from systematic uncertainties) produce some of the largest changes in our analysis, associated with the large deviation of the distributions around the cut 4.3. The shift in the measured beam spin asymmetry is large particularly for low x , low Q^2 , low P_T and high z . These shifts likely arise from the reduction of statistics in the low angle region (x is correlated with *ECU* with coefficient 0.49, Q^2 is correlated with *ECU* with coefficient 0.8).

Missing Mass

To investigate the effect of varying our missing mass cut on the analysis we shift it left and right by 50 MeV, which produces little to no effect. This is demonstrated in the figures 4.3 4.3 4.3.

Confidence Level

The minimum acceptable confidence level is varied between 0.5 at the loosest and 0.6 at the tightest. For the x , Q^2 , and P_T axes the observed shift is roughly constant regardless of the kinematic bin.

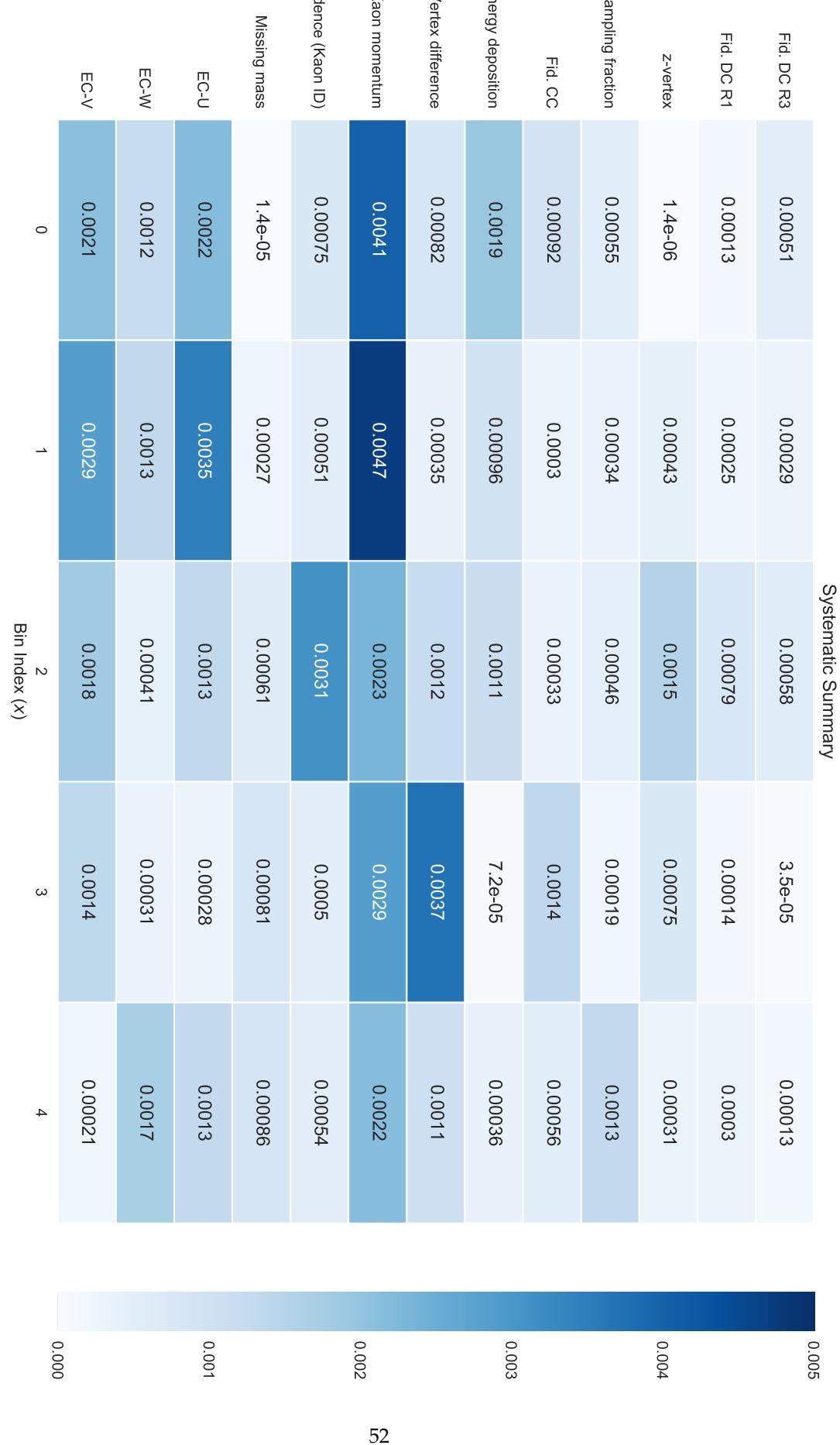


Figure 4.6: The magnitude of each systematic error source considered is shown above in a heat-map for the x axis. The vertical scale maximum corresponds to half of a percent.

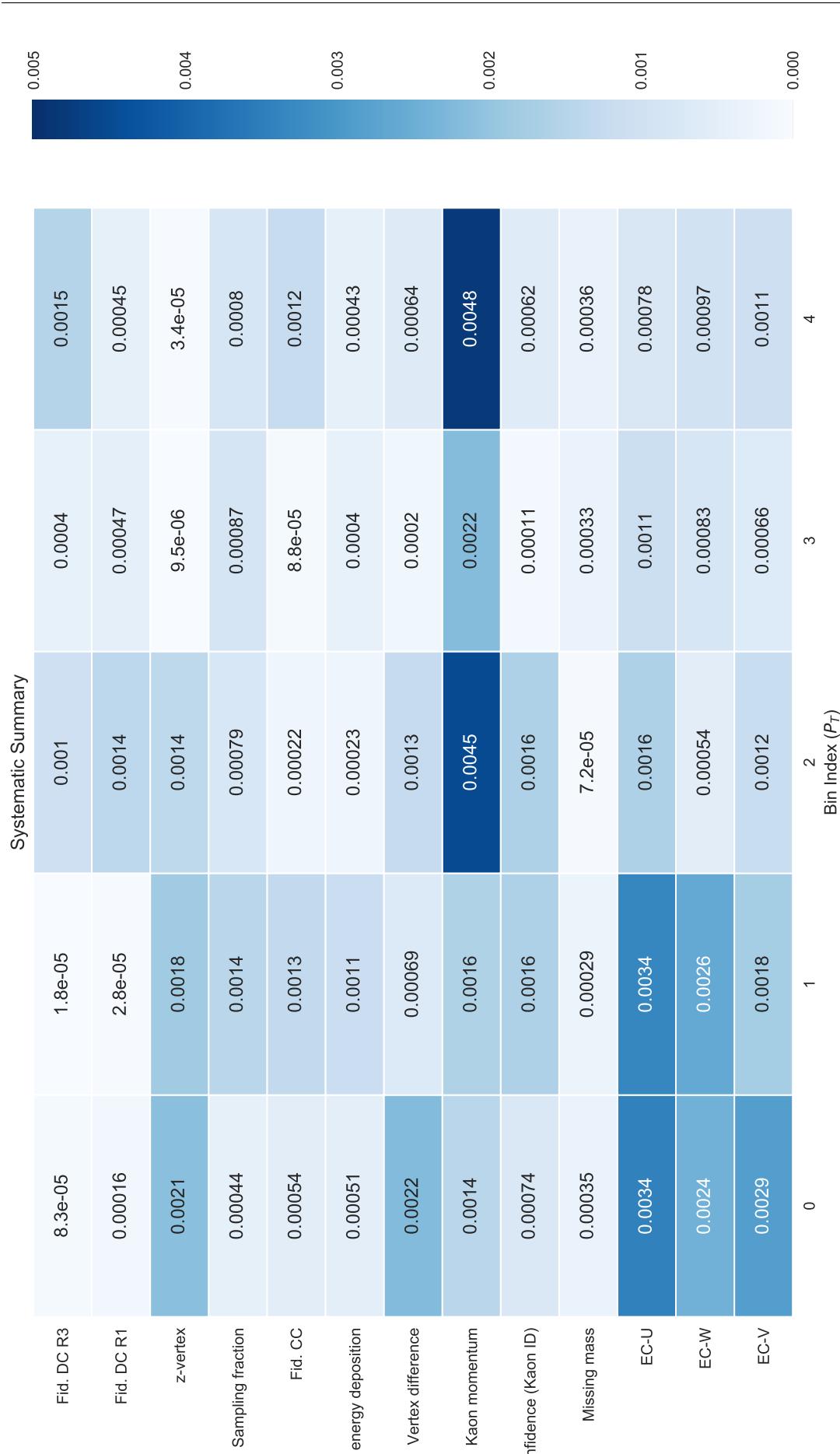


Figure 4.7: The magnitude of each systematic error source considered is shown above in a heat-map for the P_T axis. The vertical scale maximum corresponds to half of a percent.

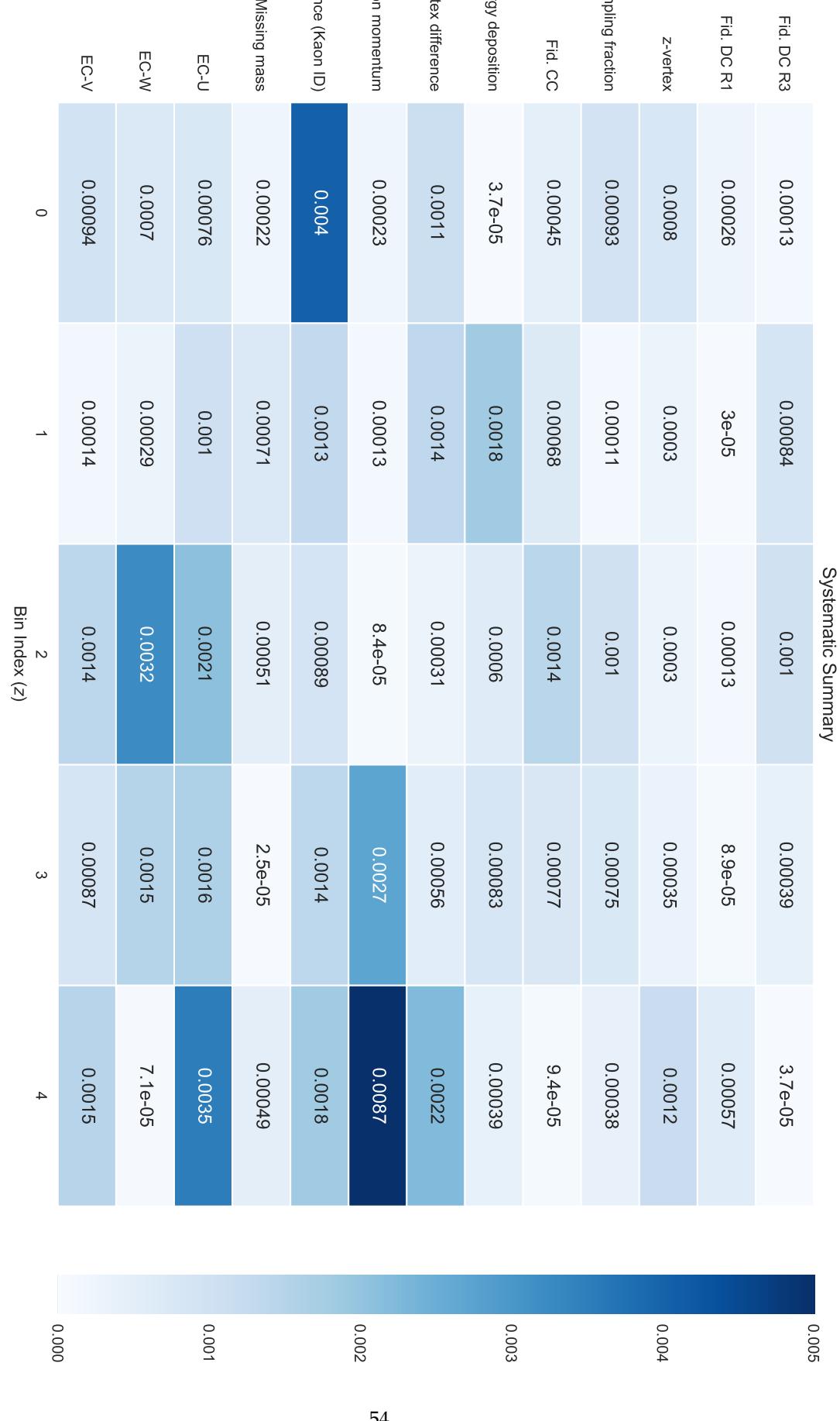


Figure 4.8: The magnitude of each systematic error source considered is shown above in a heat-map for the z axis. The vertical scale maximum corresponds to half of a percent.

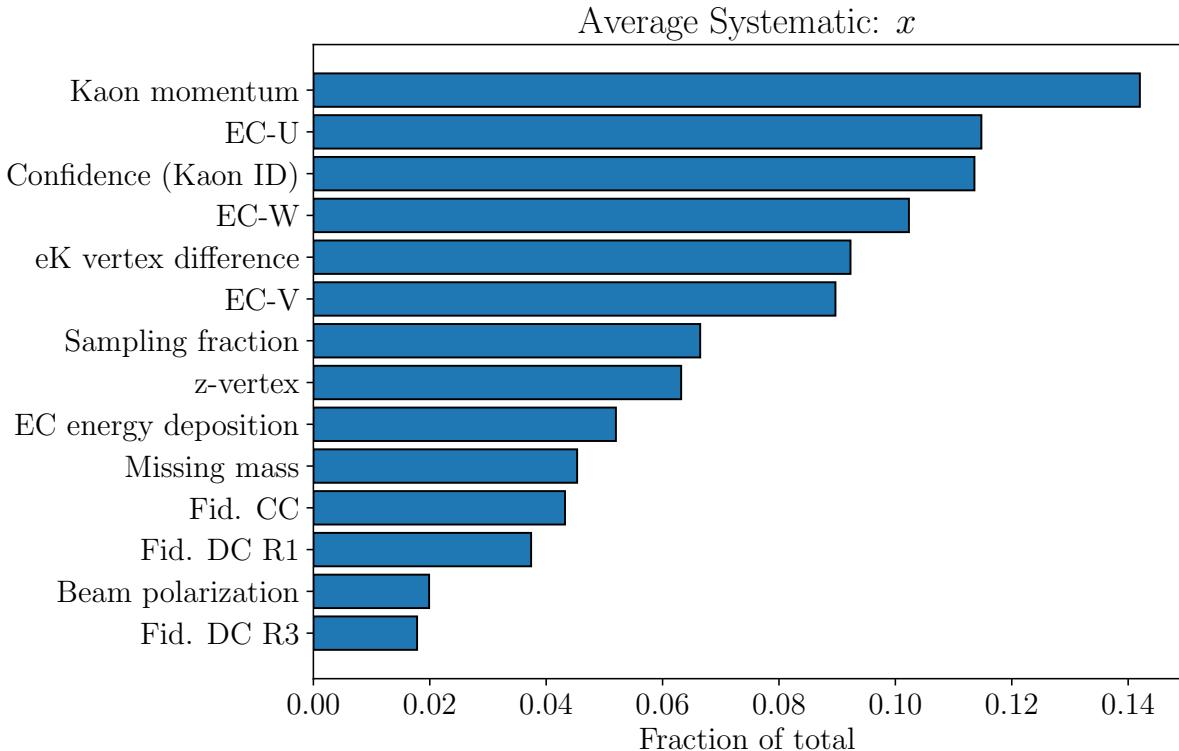


Figure 4.9: The relative contribution of each systematic uncertainty to the total is shown above averaged over the bins of the x axis.

Kaon Momentum

Monte Carlo analysis of kaon identification purity and efficiency was used to establish a maximum acceptable momentum for kaons included in our analysis. In order to study the impact of that value (2.0), the value is varied by 100 MeV and the result is included as a systematic error.

Electromagnetic Calorimeter Energy Cuts

The momentum dependent sampling fraction cut, as well as the energy deposition cut placed on the inner electromagnetic calorimeter do not contribute much to the total systematic uncertainty. In this study, the variation of the energy deposited cut by 10 MeV did not have a strong impact on the result. Additionally, the observed shift was mostly constant over the kinematic variables. The same is true for the sampling fraction cut.

Fiducial Cuts on DC and CC

The variation of our fiducial cuts on the drift chambers regions 1 and 3, as well as the Cherenkov counter produced no major shift in our measured asymmetries. This reflects the redundancy of using several fiducial cuts.

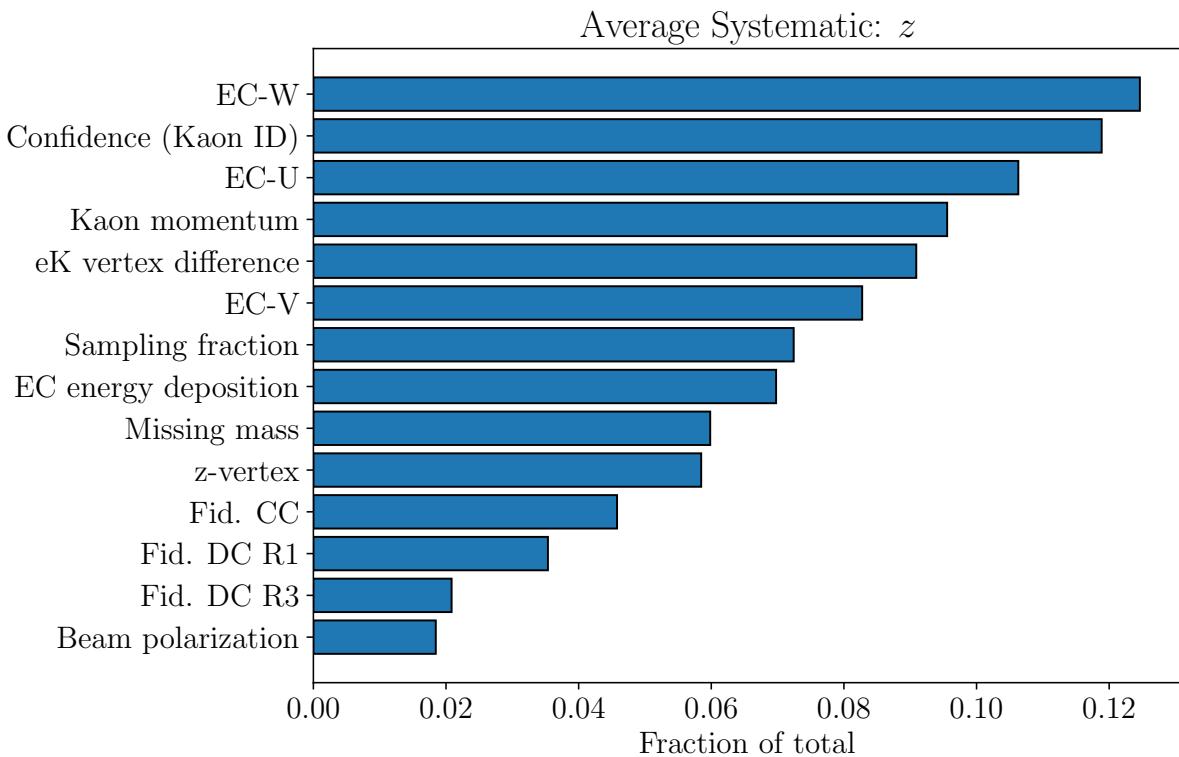
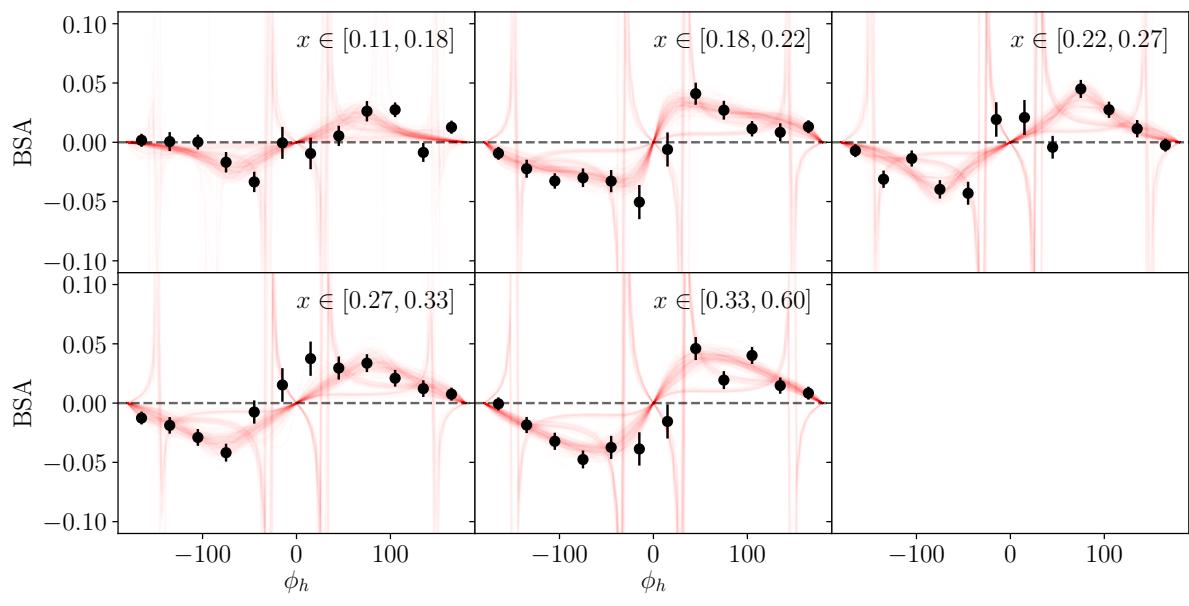


Figure 4.10: The relative contribution of each systematic uncertainty to the total is shown above averaged over the bins of the z axis.

Vertex Cuts

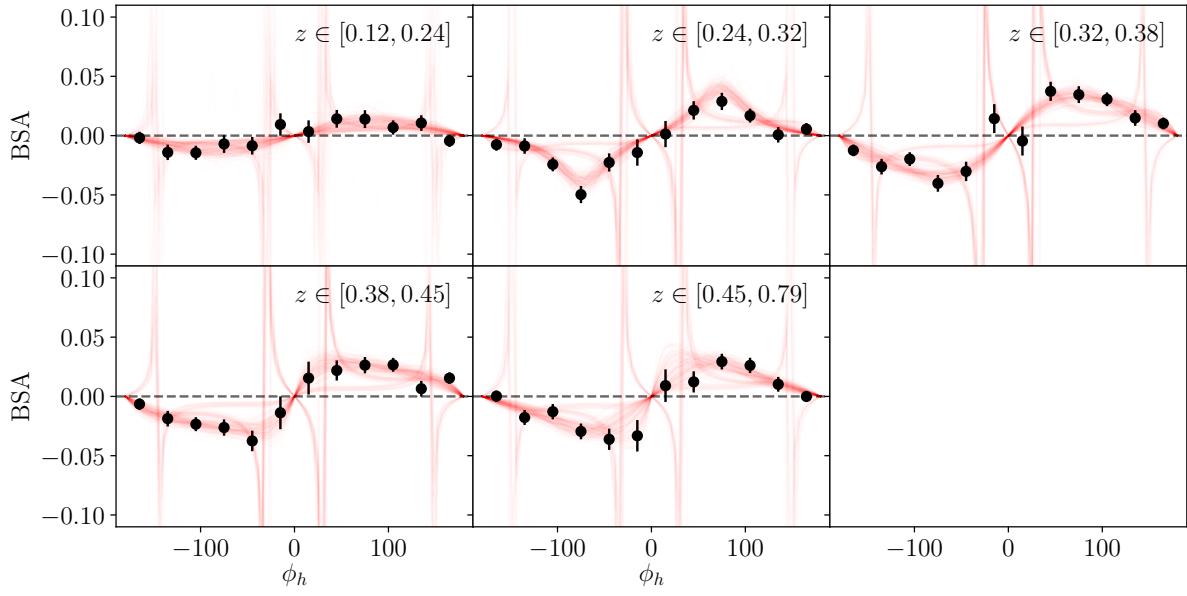
The vertex cut position was varied by ± 0.5 (cm) and small changes were observed in the extracted beam spin asymmetries. Additionally, no kinematic dependence was observed in the shifts.

4.4 Extraction of Modulations



The

ϕ_h dependence is shown for each bin of x , increasing in value from the top left to the bottom right. The statistical uncertainty is shown as black error bars on each point. Fits to 256 replicas have been superimposed on the figure. (4.11)



The

ϕ_h dependence is shown for each bin of z , increasing in value from the top left to the bottom right. The statistical uncertainty is shown as black error bars on each point. Fits to 256 replicas have been superimposed on the figure. (4.12)

The motivation to measure the beam spin asymmetry in several kinematic bins as well as bins of ϕ_h is to perform an estimate of the value of structure functions at the kinematic points (more precisely the average value of the structure functions over the range of values included in a point). To do this, the authors perform parameter estimation on the ϕ_h distributions taking as a model the theoretical dependence of the beam spin asymmetry on ϕ_h .

$$f(\phi_h, \vec{a}) = \frac{a_0 \sin \phi_h}{1 + a_1 \cos \phi_h + a_2 \cos(2\phi_h)} \quad (4.13)$$

The parameters \vec{a} are the structure function ratios to be extracted. The simplest way to extract these parameters is to use χ^2 minimization implemented in a standard fitting package. In these approaches, χ^2 is defined as the square difference between the observed data values and those predicted by the model, normalized by the error. If the fluctuation between the data and theory predictions is on the

order of the error, the χ^2 is simply on the order of the number of data points. The parameters \vec{a} which best describe the data are those which make the $\chi^2(\vec{a})$ assume its minimum value. This minimization is done in practice with gradient descent or quasi-Newton's method based algorithms like those provided in `Minuit` or `scipy.optimize.minimize`, the details of such algorithms will not be discussed here. It is sufficient to say that these minimization methods produce the parameters \vec{a} , and an estimate of the covariance matrix V . The parameters and their errors become the extracted value and uncertainty of the structure function ratio in each bin.

Unfortunately, applying the standard single-fit procedure described above does not always produce stable results. In some cases, the resulting parameter sets are reasonable, in other cases however the parameters in the denominator become nonphysically large and oppose each other. This effect has motivated previous analysts to search for other means of extracting the dominant $\sin \phi_h$ behavior from the distributions. One common technique is to assume that the coefficients a_1 and a_2 of above are small compared to 1. The analyst can then fit the ϕ_h distribution with just one linear parameter a_0 . This produces a stable result, but has the disadvantage that one needs to introduce a systematic uncertainty associated with the difference observed between using the full model (with a restricted range for the parameters in the denominator) and the results obtained using the single parameter model. Additionally, the structure function decomposition of the SIDIS cross section relies on theoretically solid ground, therefore it should be used in its full form. If the data contain little information regarding the structure function ratios in the denominator, the authors believe it more valuable to demonstrate this by extracting those parameters with (large) errors, rather than ignore their contribution. In order to accomplish this, the method of replicas (or parametric bootstrapping) is used to perform the parameter estimation. The replica method consists of generating N_{rep} pseudo-data ϕ_h distributions which have a normal distribution located at the observed value, and with a variance equal to the statistical errors on the associated data point.

$$\vec{A}_{rep} = \mathcal{N}(\vec{A}, \sigma_A) \quad (4.14)$$

Here \vec{A} is a vector of length n_{phi} bins, representing the measured beam spin asymmetry for each value of ϕ_h in a given kinematic bin. Each of these distributions is fit with the full model, and the resulting parameter values are saved. The final reported value for each fit parameter, as well as its uncertainty can be reported as the mean, and standard deviation of the fit results. This procedure which is similar to bootstrapping, can be seen as an attempt to fit the underlying distribution that generated the data while avoiding the statistical noise. This technique has been discussed in [14].

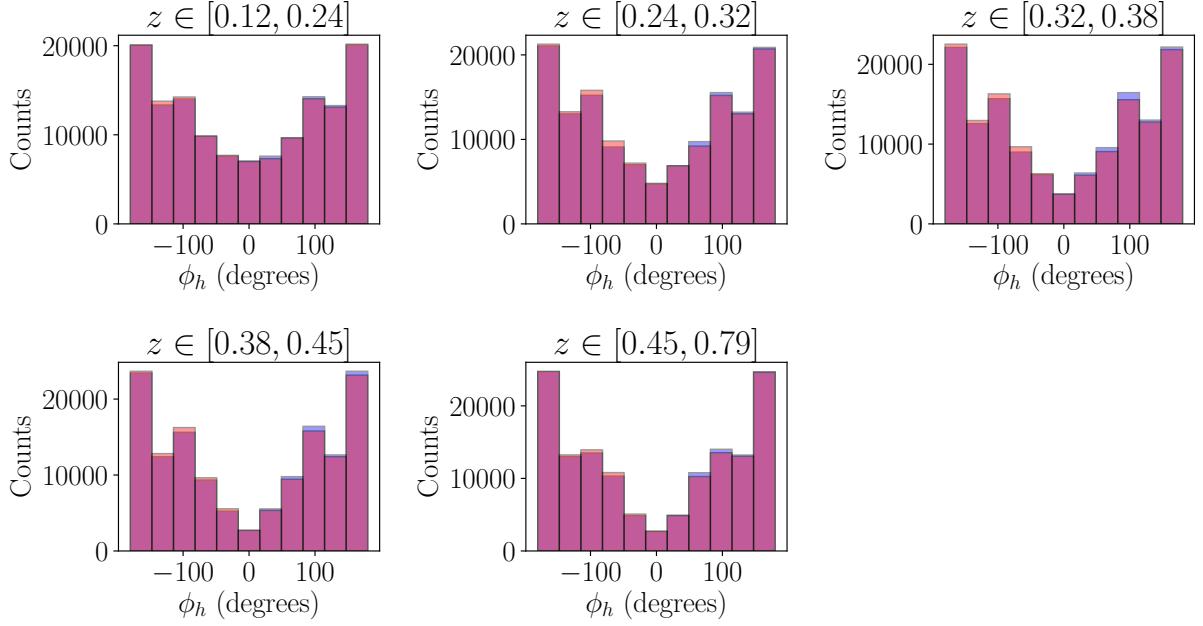


Figure 4.11: Counts for different helicity states are superimposed for different bins of z .

$$\langle a_j \rangle = \sum_{i=1}^{N_{rep}} a_j^{(i)} \quad (4.15)$$

$$\sigma_{a_j}^2 = \frac{1}{N_{rep}-1} \sum_{i=1}^{N_{rep}} (a_j^{(i)} - \langle a_j \rangle)^2 \quad (4.16)$$

Results

As is the case for positive pions, the observed structure function ratio $A_{LU}^{\sin \phi}$ is positive for all kinematic points that were measured. In general, this extraction reveals that the $\sin \phi_h$ moment has a magnitude around 3% for most kinematic points, and depends weakly on the kinematic variables used in this analysis. The relative asymmetry value to total error ratio is around 1.5% for most measured points.

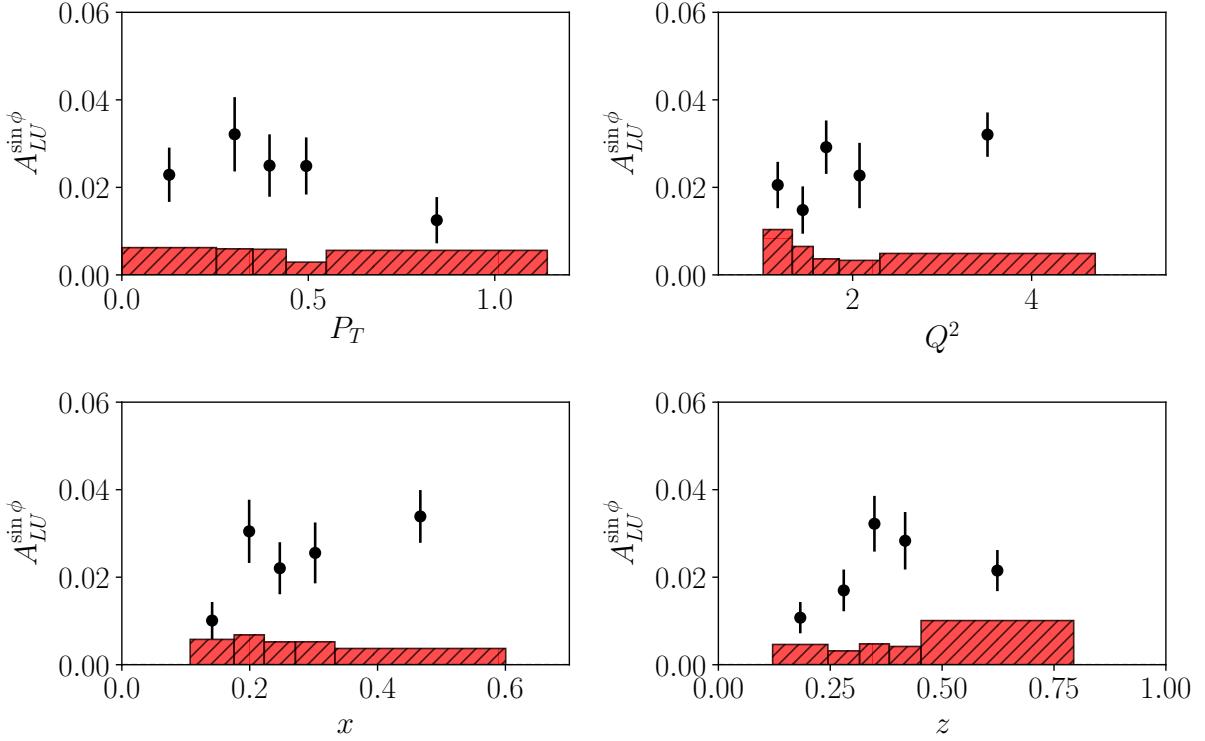


Figure 4.12: Our extraction of $A_{LU}^{\sin \phi}$ for the kinematic bins described above. The black error bars represent uncertainty in the extraction of the parameter value. Red error bars are systematic uncertainties.

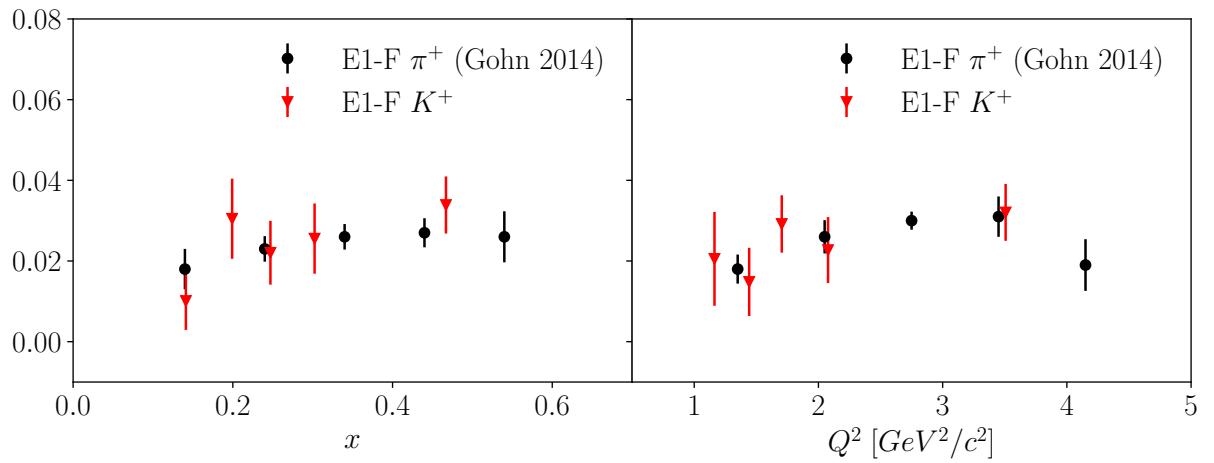


Figure 4.13: In this figure the results of this study for positively charged kaons are compared with previous results from the same dataset produced by [8] for positively charged pions. This figure shows the x and Q^2 dependence of $A_{LU}^{\sin \phi_h}$.

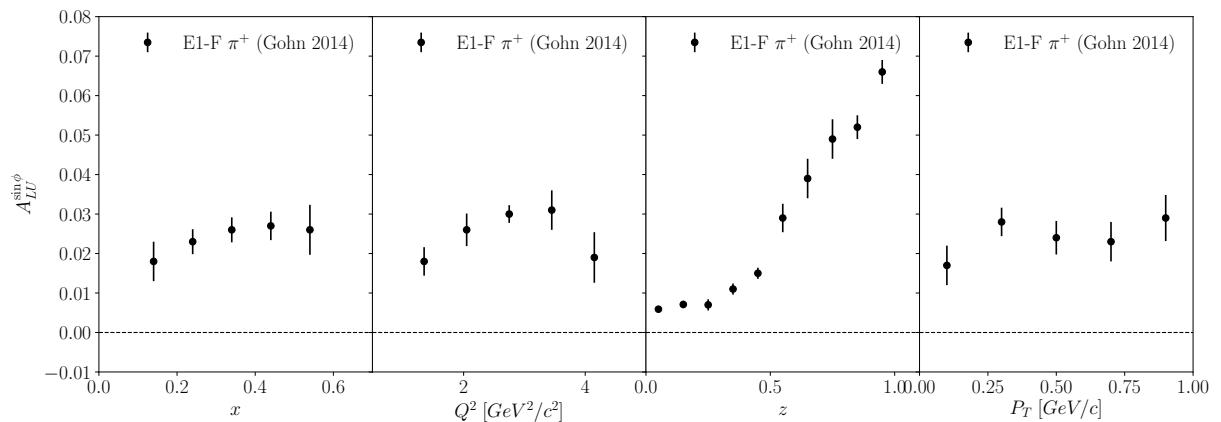


Figure 4.14: In this figure the results of this study for positively charged kaons are compared with previous results from the same dataset produced by [8] for positively charged pions. This figure shows the z and P_T dependence of $A_{LU}^{\sin \phi_h}$.

Appendix A: Derivation of formulas related to errors

Propagation of errors

Let \vec{x} be a set of n random variables $\vec{x} = (x_1, x_2, \dots, x_n)$ and known mean $\mu_i = \langle x_i \rangle$ and covariance $V_{ij} = \langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle$. Suppose that we measure a function $f(\vec{x})$ that depends on the variables \vec{x} and we want to understand how the covariances V_{ij} on \vec{x} will show up manifest themselves as errors on our measurement of $f(\vec{x})$. We can start by expanding our function around the expected value of x_i .

$$f(\vec{x}) \approx f(\vec{\mu}) + \sum_{i=1}^n \frac{\partial f}{\partial x_i} \Big|_{x_i=\mu_i} (x_i - \mu_i) \quad (17)$$

We can then take the expectation value of our function.

$$\langle f(\vec{x}) \rangle = \langle f(\vec{\mu}) \rangle + \sum_{i=1}^n \frac{\partial f}{\partial x_i} \Big|_{x_i=\mu_i} \langle x_i - \mu_i \rangle \quad (18)$$

Where here the term $\langle x_i - \mu_i \rangle$ is zero.

$$\langle x_i - \mu_i \rangle = \langle x_i \rangle - \mu_i = \mu_i - \mu_i = 0 \quad (19)$$

It is apparent then that the expectation value of our function f evaluated close to the expected values of our variables \vec{x} is just the function evaluated at the expectation value of the random variables \vec{x} .

$$\langle f(\vec{x}) \rangle = \langle f(\mu) \rangle = f(\mu) \quad (20)$$

We can also consider the second moment $\langle f^2(\vec{x}) \rangle$,

$$\langle f^2(\vec{x}) \rangle \approx \left\langle \left(f(\vec{\mu}) + \sum_{i=1}^n \frac{\partial f}{\partial x_i} \Big|_{x_i=\mu_i} (x_i - \mu_i) \right)^2 \right\rangle \quad (21)$$

which is,

$$= \langle f^2(\mu) \rangle + \sum_{i=1}^n \sum_{j=1}^n \frac{\partial f}{\partial x_i} \Big|_{x_i=\mu_i} \frac{\partial f}{\partial x_j} \Big|_{x_j=\mu_j} \langle (x_i - \mu_i)(x_j - \mu_j) \rangle + 2 \left\langle f(\mu) \sum_{i=1}^n \frac{\partial f}{\partial x_i} \Big|_{x_i=\mu_i} (x_i - \mu_i) \right\rangle \quad (22)$$

and the last expectation value vanishes due to the same logic used when calculating the first moment. We recognize the term $\langle (x_i - \mu_i)(x_j - \mu_j) \rangle$ as the element of the covariance matrix V_{ij} . Our second moment is then complete as follows.

$$\langle f^2(\vec{x}) \rangle = f^2(\mu) + \sum_{i=1}^n \sum_{j=1}^n \frac{\partial f}{\partial x_i} \Big|_{x_i=\mu_i} \frac{\partial f}{\partial x_j} \Big|_{x_j=\mu_i} V_{ij} \quad (23)$$

We can then calculate the variance of the function.

$$\sigma_f^2 = \langle f^2(\vec{x}) \rangle - \langle f(\vec{x}) \rangle^2 \quad (24)$$

$$= (f^2(\vec{\mu}) - f^2(\vec{\mu})) + \sum_{i=1}^n \sum_{j=1}^n \frac{\partial f}{\partial x_i} \Big|_{x_i=\mu_i} \frac{\partial f}{\partial x_j} \Big|_{x_j=\mu_i} V_{ij} \quad (25)$$

$$= \sum_{i=1}^n \sum_{j=1}^n \frac{\partial f}{\partial x_i} \Big|_{x_i=\mu_i} \frac{\partial f}{\partial x_j} \Big|_{x_j=\mu_i} V_{ij} \quad (26)$$

This is the standard error propagation formula which is widely used. These correlations σ_{ij} can arise from several sources.

- Common measurement uncertainties.
- Correlations in $x_i x_j$ leading to correlations in $\sigma_i \sigma_j$.

References

- [1] E. Leader and M. Anselmino, *Proceedings, 8th International Symposium on High-energy Spin Physics: Minneapolis, Minnesota, USA, September 12-17, 1988. Vol. 1, 2*, Z. Phys. **C41**, 239 (1988), [AIP Conf. Proc.187,764(2008)].
- [2] P. J. Mulders and R. D. Tangerman, Nucl. Phys. **B461**, 197 (1996), [Erratum: Nucl. Phys.B484,538(1997)], arXiv:hep-ph/9510301 [hep-ph].
- [3] A. Bacchetta, M. Diehl, K. Goeke, A. Metz, P. J. Mulders, and M. Schlegel, JHEP **02**, 093 (2007), arXiv:hep-ph/0611265 [hep-ph].
- [4] A. Airapetian *et al.* (HERMES), Phys. Rev. Lett. **103**, 152002 (2009), arXiv:0906.3918 [hep-ex].
- [5] A. Airapetian *et al.* (HERMES), Phys. Rev. **D87**, 074029 (2013), arXiv:1212.5407 [hep-ex].
- [6] M. Aghasyan *et al.* (COMPASS), Phys. Rev. **D97**, 032006 (2018), arXiv:1709.07374 [hep-ex].
- [7] M. Anselmino, M. Boglione, U. D'Alesio, J. O. Gonzalez Hernandez, S. Melis, F. Murgia, and A. Prokudin, Phys. Rev. **D92**, 114023 (2015), arXiv:1510.05389 [hep-ph].
- [8] W. Gohn *et al.* (CLAS), Phys. Rev. **D89**, 072011 (2014), arXiv:1402.4097 [hep-ex].
- [9] Y. X. Zhao *et al.* (Jefferson Lab Hall A), Phys. Rev. **C90**, 055201 (2014), arXiv:1404.7204 [nucl-ex].
- [10] N. A. Harrison, *Exploring the Structure of the Proton via Semi-Inclusive Pion Electroproduction*, Ph.D. thesis, Connecticut U. (2010).
- [11] M. Mirazita, “Kinematic corrections for e1-f,” .
- [12] R.-G. Ping *et al.*, Int. J. Mod. Phys. **A24S1**, 23 (2009).
- [13] R. Barlow, in *Advanced Statistical Techniques in Particle Physics. Proceedings, Conference, Durham, UK, March 18-22, 2002* (2002) pp. 134–144, arXiv:hep-ex/0207026 [hep-ex].
- [14] G. Watt and R. S. Thorne, JHEP **08**, 052 (2012), arXiv:1205.4024 [hep-ph].