Transverse Momentum in Nucleons, From Raw Data to TMD Extraction

David M. Riser

March 2018

Contents

Chapter 1

Introduction

This chapter is the introduction. The content of this chapter will explain some things.

Chapter 2

Basic Analysis & Corrections

2.1 Introduction

2.2 Luminosity Calculation

2.3 Determination of Good Run List

2.4 Helicity Determination

2.5 Vertex Corrections

2.6 Timing Corrections

2.7 Kinematic Corrections

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 responces to each track, the reconstruction package `recsis` 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 by the authors 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. Each negative track is considered a possible electron, a series of physically motivated cuts is applied. If a track passes all cuts, it is identified as an electron. All track indices which pass electron identification are saved, and the one with the highest momentum is used in the analysis.

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)

Each cut will now be described in more detail.

Negativity Cut

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

Drift chamber fiducial

Negative tracks which pass geometrically close to the edges of the drift chamber are, from a tracking perspective, more difficult to understand. Often these tracks originate from downstream background, or are otherwise unacceptable. 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 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 $\pm\cot(\theta)$.

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

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

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

Electromagnetic Calorimeter fiducial (UVW)

As tracks 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

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

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

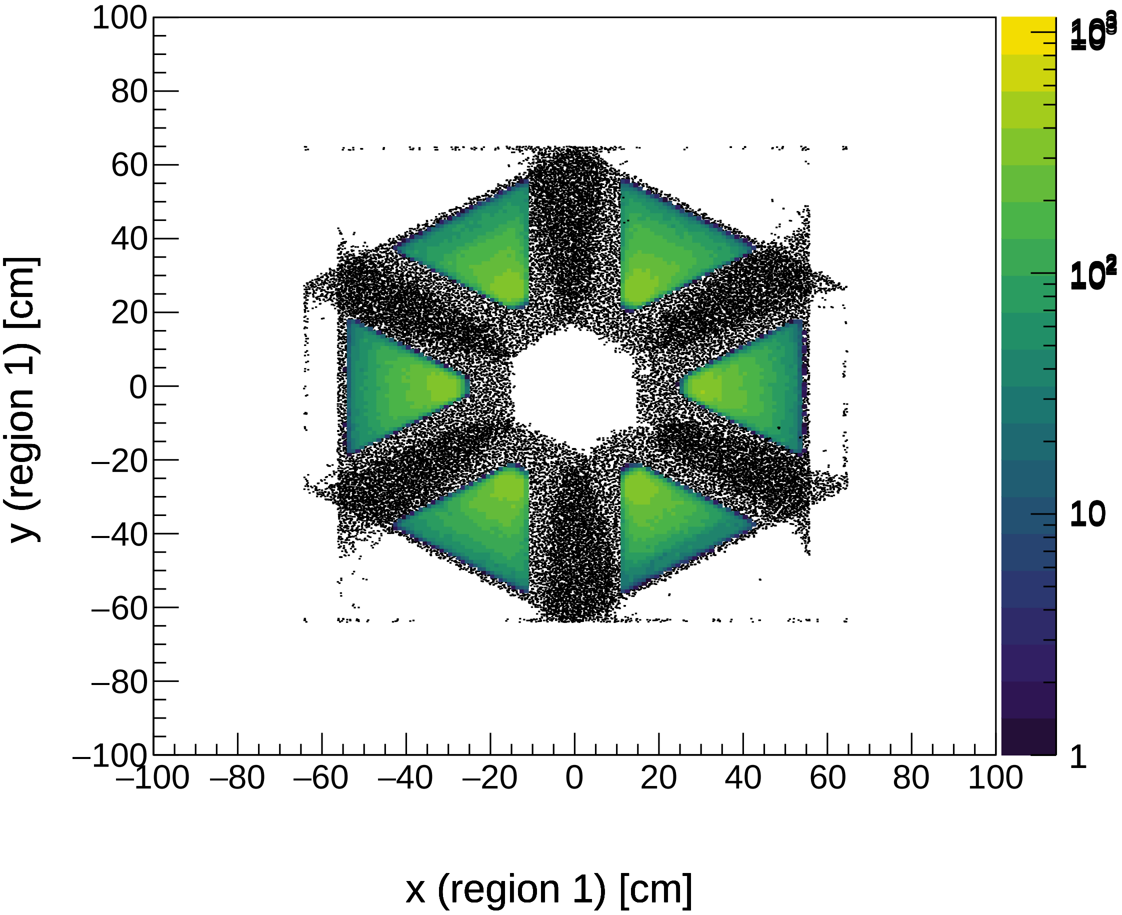

Figure 3.1: Tracks shown in color remain after the application of drift chamber region 1 fiducial cuts to all cuts, shown here as black points.

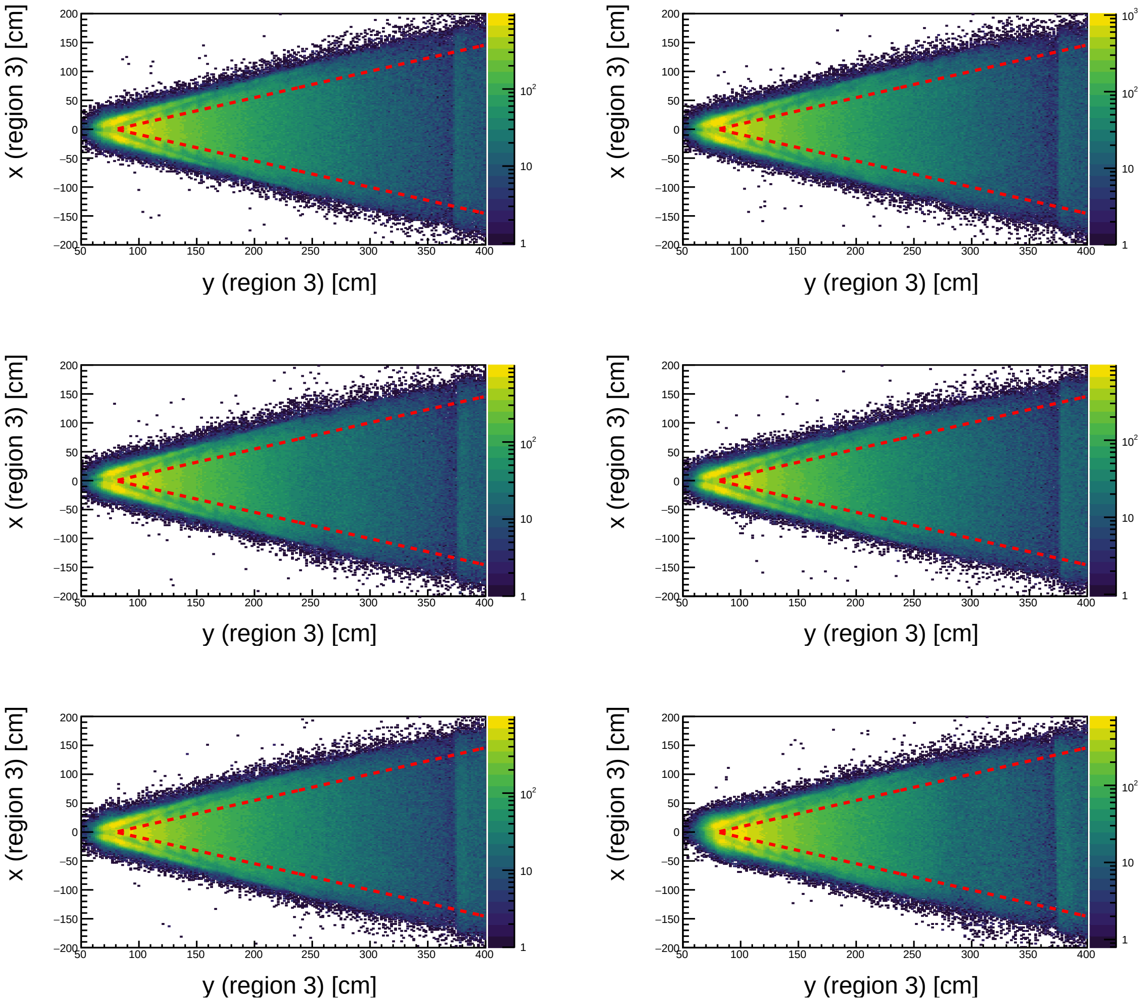

Figure 3.2: The selection criteria shown in red is applied to drift chamber region 3.

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

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

within the calorimeter volume (it spills out the edges). For this reason, it has become standard to remove the hits which fall within the outer 10 centimeters of each layer of the EC (10 centimieters is the width of a scintillator bar). This cut is applied in the U, V, W coordinate system.

EC minimum energy deposition

The negative tracks that start out as electron candidates are primarily composed of electrons and negative π mesons. One way to differentiate between these two species is to exploit the difference in energy deposition between the two in the electromagnetic calorimeter. Electron typically develop a much larger 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.

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 distribution, 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 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, the authors choose a functional form for the mean and standard deviation of the distributions to be a third order polynomial in momentum.

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

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

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

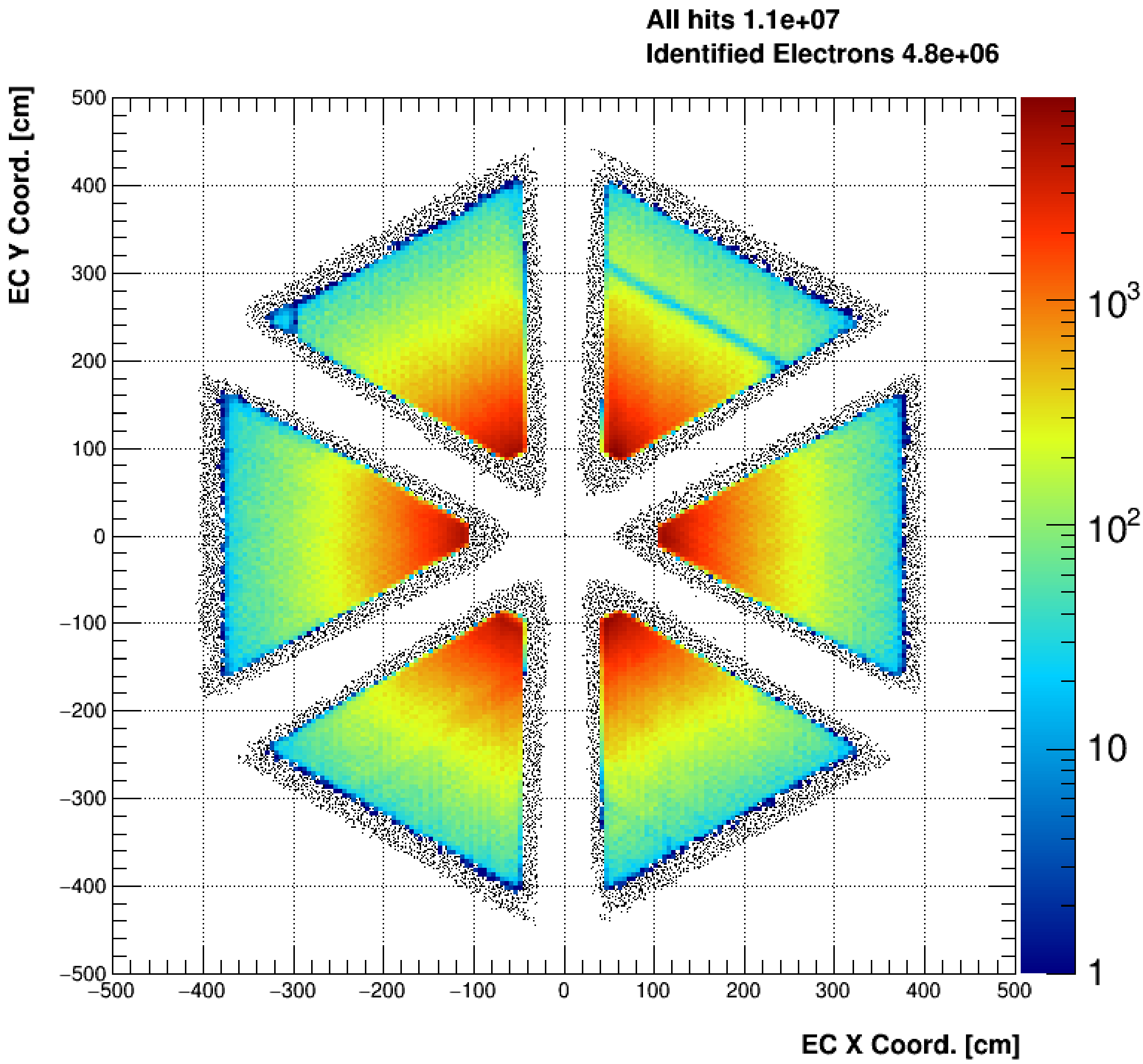

Figure 3.3: All negative tracks are shown here in black. In color, the tracks which pass the EC fiducial cut are shown.

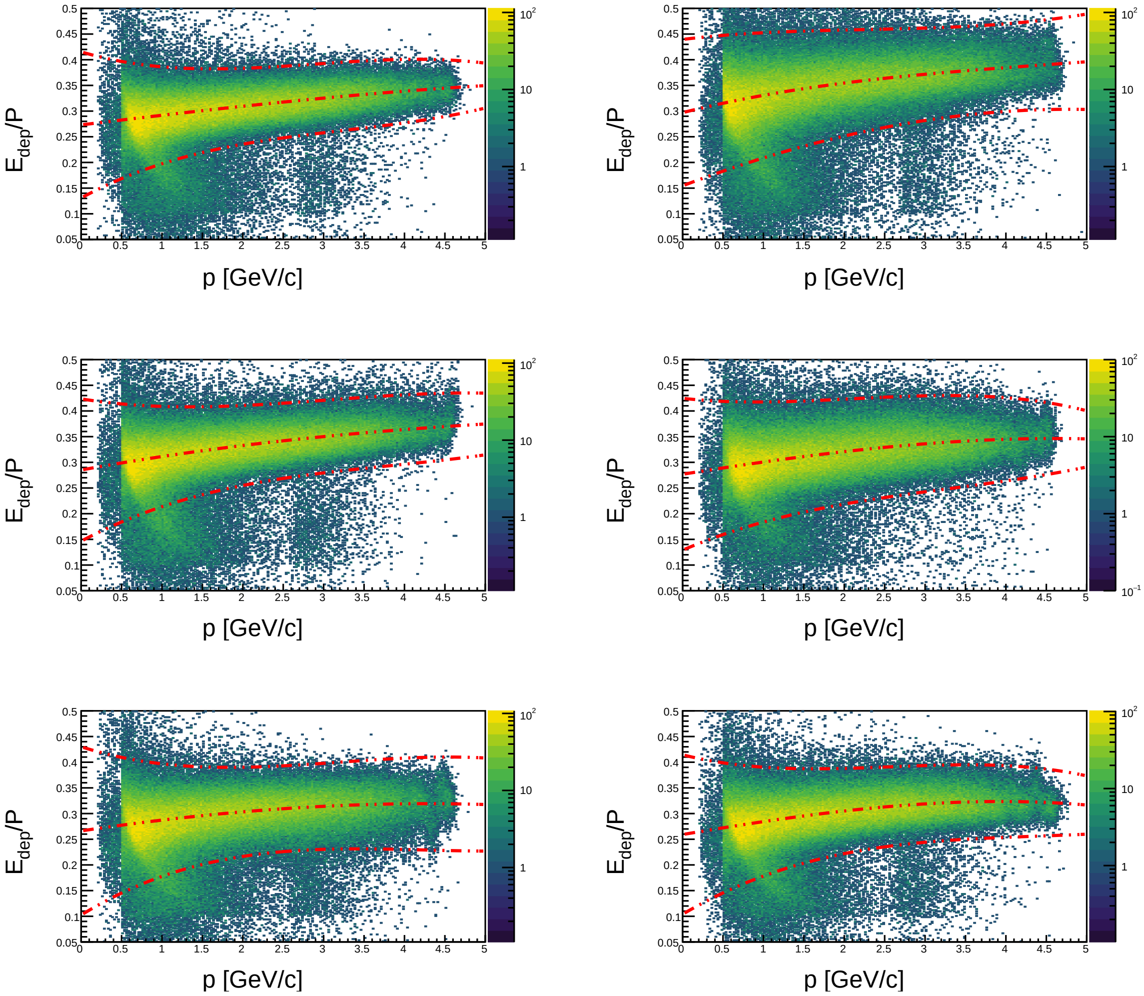

Figure 3.4: The sampling fraction selection boundary is shown here for the nominal value of $N_{sigma} = 4$.

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

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

Due to slight differences between the 6 sectors of the CLAS detector, the authors choose to calibrate and apply this cut for each sector individually. The results are shown in table 3.3.

z-vertex position

Electrons can be produced as part of e^+e^- pairs. For this analysis, these are not of interest. The authors choose to select only electrons which originate from the target and are believed to be the scattered incoming electron. For this reason the authors accept only electron candidates which have a z-vertex $v_z \in [-27.7302, -22.6864]$. This cut is applied after the vertex position has been corrected (this correc-

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.

tion will be discussed in a subsuquent chapter).

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

The placement of photo-multiplier tubes (PMT) in the Cherenkov counter allows for additional consistency conditions to be applied. The placement of 18 PMTs increasing in polar angle away from the beamline means that the PMT segment number is correlated to the angle which the electron has with the beamline at the Cherenkov counter θcc. 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 code 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 $\phi_{rel} < 0$ are allowed to pass. Similarly if the right PMT fired (code = 1), only events with $\phi_{rel} > 0$ are allowed to pass. Technical note: the integers in question can be obtained from the ntuple22 format tree by doing the following.

```
for (int index = 0; index < event.gpart; index++){
  int pmt = event.cc_segm[index]/1000 - 1;
  int segment = event.cc_segm[index]%1000/10;
}
```

3.3 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 [**bes-physics**].

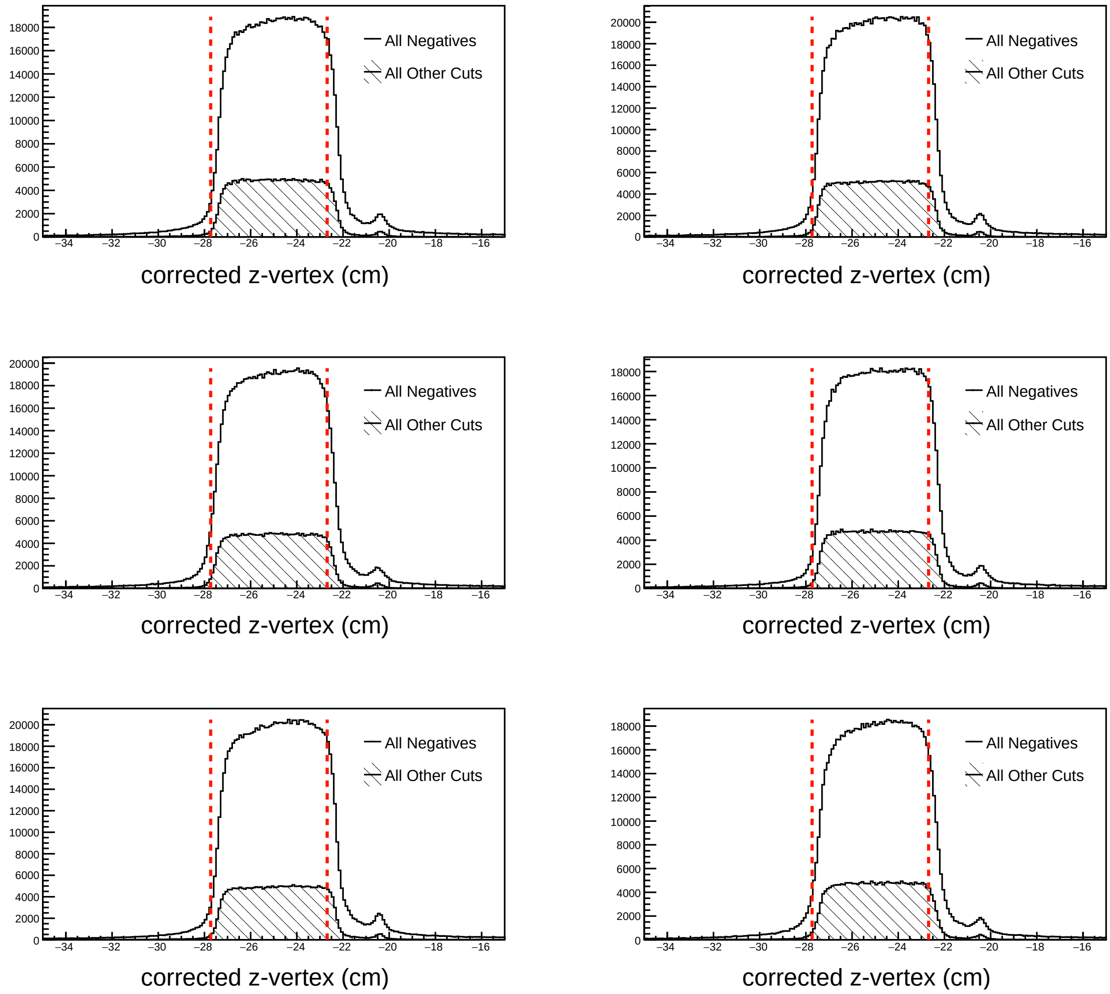

Figure 3.5: The track vertex cut is shown above. All negative tracks are shown in white, while the tracks passing all other criteria are shown in black hatch. The cut bounday is displayed as red lines. For E1-F the target center was located at -25 cm.

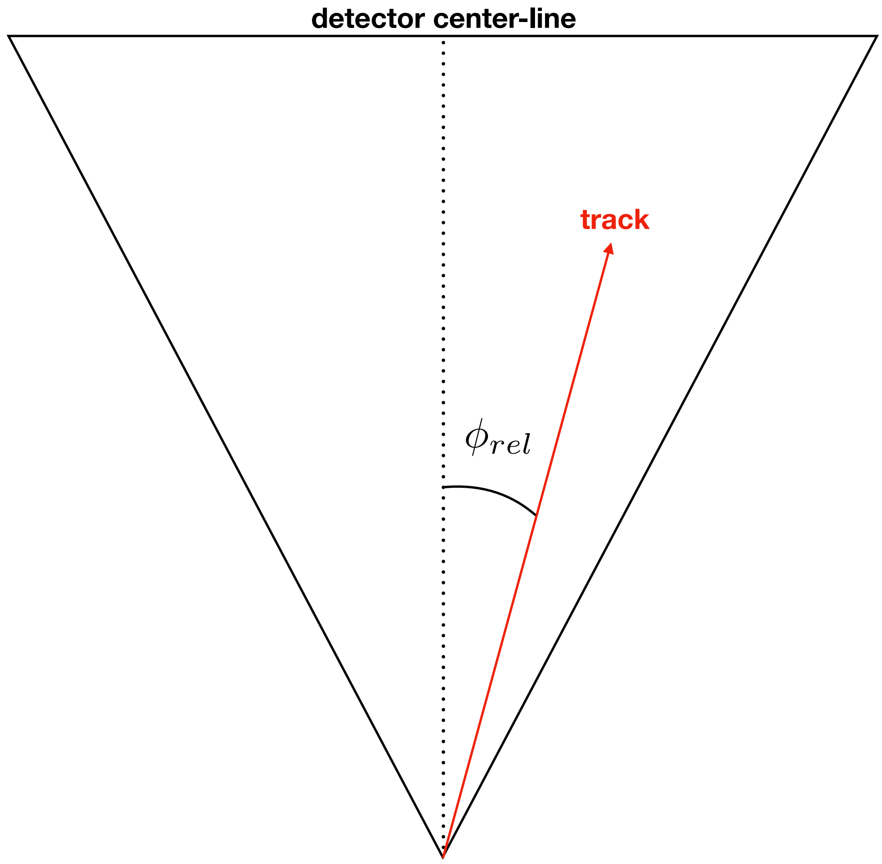

Figure 3.6: The angle ϕ_{rel} is the azimuthal angle between the central line of the detector and the track.

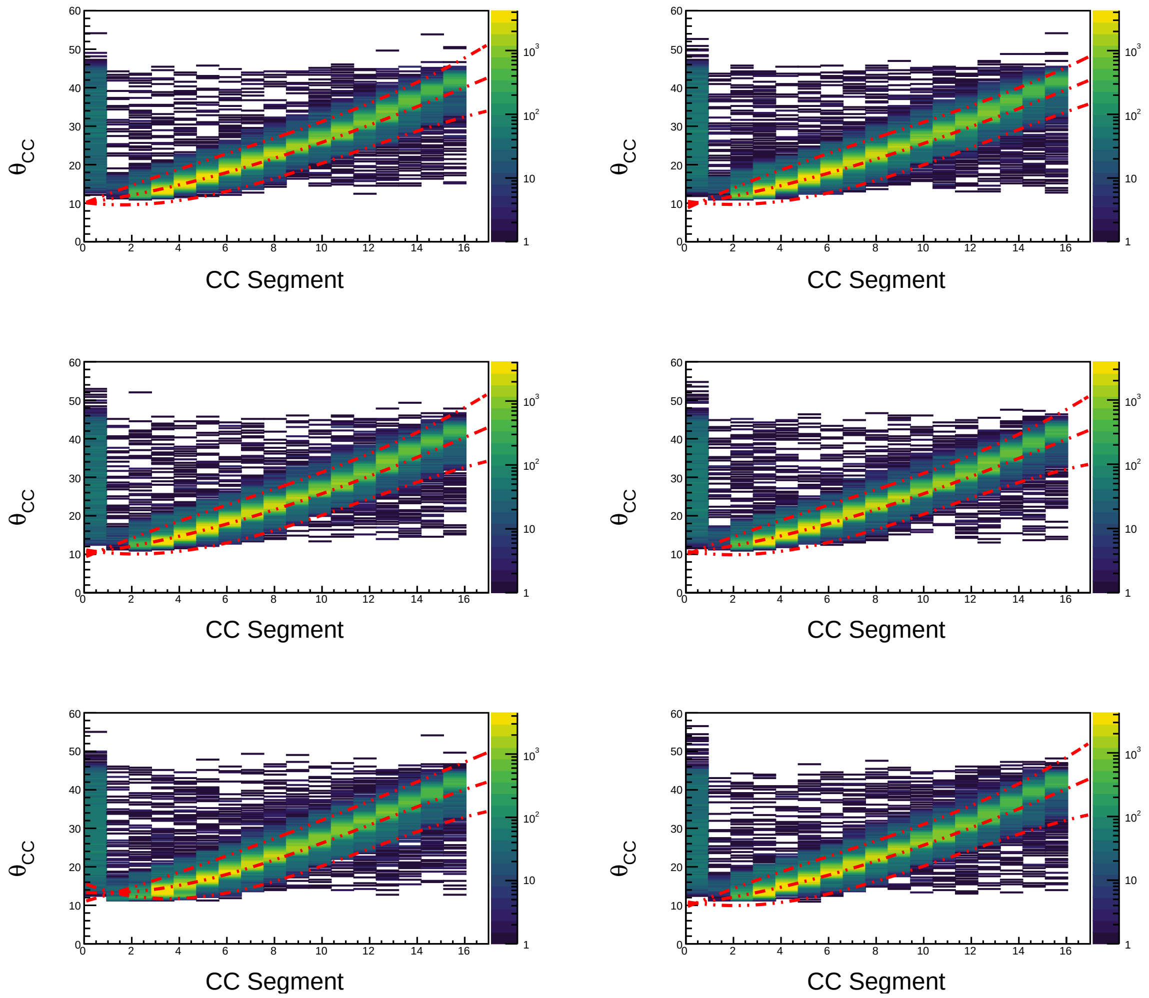

Figure 3.7: Correlation between θ_{CC} and the CC segment is shown above, with our selection boundaries overlaid in red.

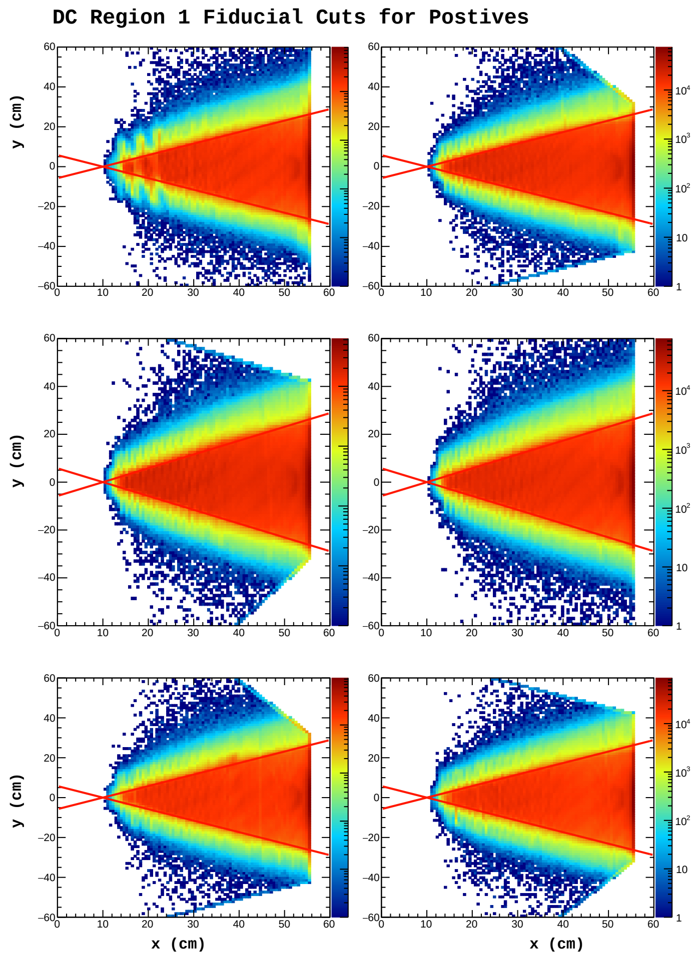

Figure 3.8: Shown above: Positive track hits on the region 1 drift chamber, events falling between the red lines are kept for analysis.

3.3.1 Hadron ID Cuts

The cuts used by the author 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 are for negative hadrons are those which are used for the electron. The parameters used for positive tracks are $h = 10$, $\theta = 60$.

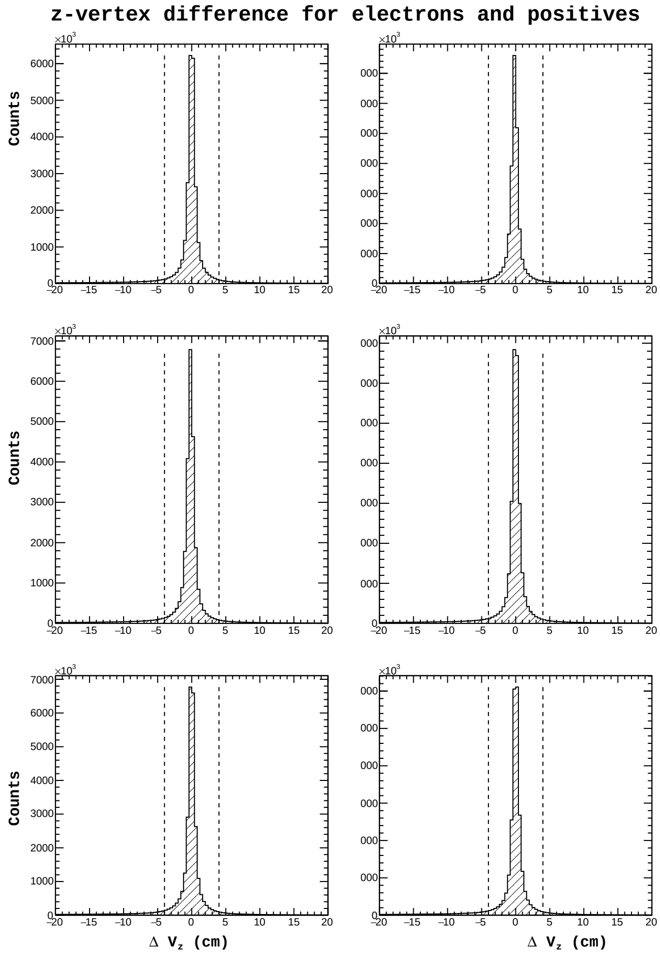

Figure 3.9: Shown above: The difference between the z-vertex position between detected electrons and positive tracks.

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 3.3.1. If the analyst desires to look at events where the hadron is produced as the result of a decaying hadron, this cut should be removed.

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

In this section, positive hadrons are used as an example. The same method is applied to the negative hadrons. 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 catagorize 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 hadron being hypothesized (eg: in our case the possible values for positive hadrons are pion, kaon, proton). In general if one uses a set of N variables $x = (x_1, x_2, ..., x_N)$, the likelihood for a hypothesis h is defined below.

$$\mathcal{L}_h = \prod_{i=1}^{N} P_i(x_i; p, h) \tag{3.7}$$

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 (β).

$$P(\beta; p, h) = \frac{1}{\sqrt{2\pi}\sigma_\beta(p, h)} exp \left\{ -\frac{1}{2} \left(\frac{\beta - \mu_\beta(p, h)}{\sigma_\beta(p, h)} \right)^2 \right\} \tag{3.8}$$

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

$$\frac{\mathcal{L}_h}{\mathcal{L}_\pi + \mathcal{L}_K + \mathcal{L}_p} \tag{3.9}$$

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.

$$\alpha = 1 - \int_{\mu - \beta_{obs}}^{\mu + \beta_{obs}} P(\beta; p, h) d\beta \tag{3.10}$$

This quantity represents the probability to observe a value of β as far from the mean as β_{obs}. Confidence levels of 0 then 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 sigma 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 ascertation of the mean and standard deviation of the probability density function (which depends on momentum) for the different particle hypothesis. 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, the authors chose to create an enhanced sample of candidates for each of the three positive

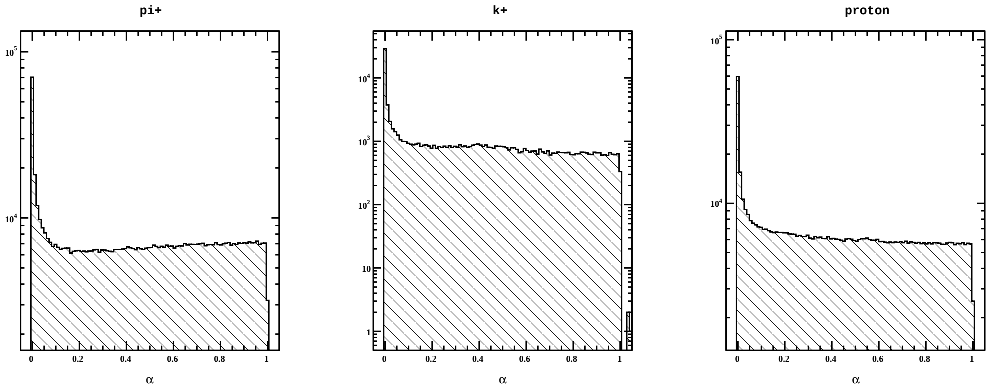

Figure 3.10: Shown above: The distribution of confidence level for all positive tracks after being classified by the likelihood ratio.

particles in question before doing the fitting. In this way, we hope to get a better quality fit of the true mean, and resolutions for the different species. For fitting of pion and proton resolutions, positive tracks are assumed to be pions and the missing mass of the event is calculated. Then, a cut is placed around the neutron mass. In doing so, we are selecting mainly two types of exclusive events. 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). The negative tracks π^-, K^- are fit directly as is normally done.

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

$$\mu(p,h) = \mu_{theory} + \Delta\mu \tag{3.11}$$

$$\mu_{theory} = \frac{1}{\sqrt{1 + (m_h/p)^2}} \tag{3.12}$$

$$\Delta\mu = \mu_0 + \mu_1 p + \mu_2 p^2 \tag{3.13}$$

$$\sigma(p,h) = \sigma_0 + \sigma_1 p + \sigma_2 p^2 \tag{3.14}$$

The values are displayed in the table below.

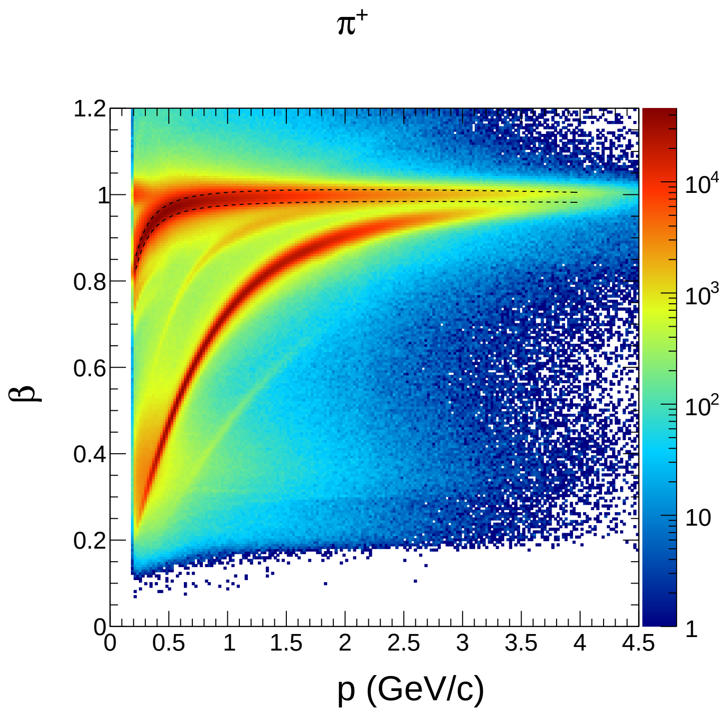

Figure 3.11: Shown above: All positive tracks overlaid with our determination of $\mu(p) \pm \sigma(p)$ for π^+

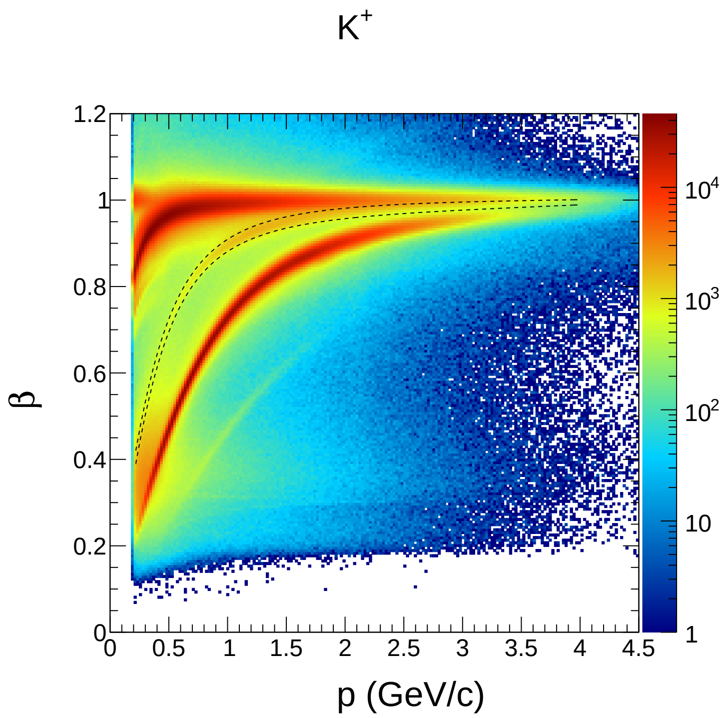

Figure 3.12: Shown above: All positive tracks overlaid with our determination of $\mu(p) \pm \sigma(p)$ for K^+

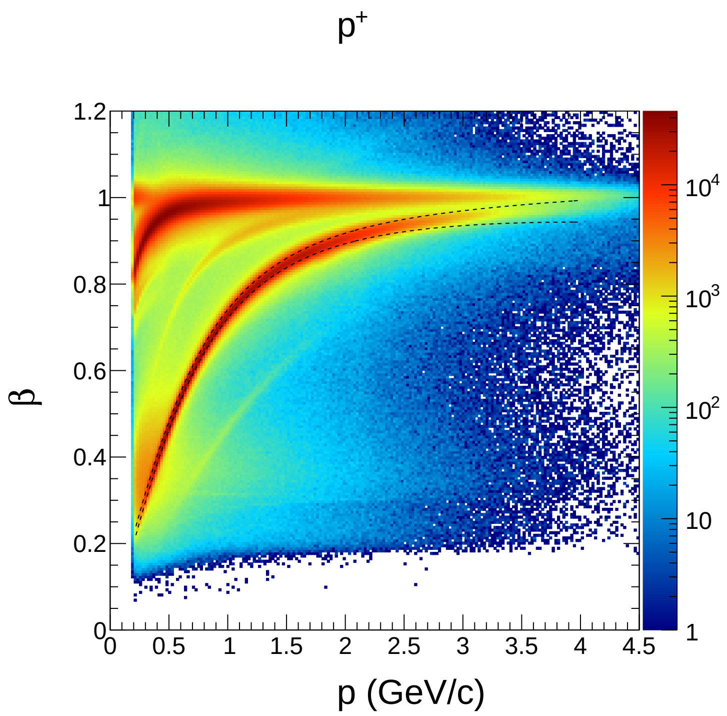

Figure 3.13: Shown above: All positive tracks overlaid with our determination of $\mu(p) \pm \sigma(p)$ for p^{+}

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
π^-	μ_2	3.28823666e-04	-1.30673670e-05	-2.32502052e-04	-9.75619848e-04	-5.89834444e-04	5.27496718e-04
π^-	μ_1	-3.94924663e-03	-2.66028661e-03	-1.28565631e-03	9.09410075e-04	-2.01610684e-03	-4.42276918e-03
π^-	μ_0	9.48011169e-04	1.55078786e-03	1.43431985e-03	1.35056935e-03	4.59833580e-03	2.30751866e-03
π^-	σ_2	4.37635504e-04	4.38306224e-04	5.32057510e-04	3.36999845e-04	7.74135462e-04	1.36515196e-04
π^-	σ_1	-3.28011836e-03	-3.28456104e-03	-3.82847286e-03	-3.11749323e-03	-4.63110728e-03	-2.21229710e-03
π^-	σ_0	1.63296567e-02	1.62229164e-02	1.59769911e-02	1.58803427e-02	1.74670064e-02	1.51753145e-02
K^-	μ_2	-2.72020947e-03	-5.21081786e-03	-2.13868763e-02	-4.45600034e-03	-7.60703841e-03	-5.27074813e-03
K^-	μ_1	1.78610401e-02	2.30787460e-02	9.49357818e-02	1.95764575e-02	3.63245785e-02	2.92417500e-02
K^-	μ_0	-2.26190100e-02	-2.22562379e-02	-1.02704771e-01	-2.25931014e-02	-5.10484618e-02	-3.19918187e-02
K^-	σ_2	1.76905114e-02	1.62989708e-02	3.60928130e-02	1.51270521e-02	1.91308107e-02	2.38470033e-02
K^-	σ_1	-7.74901862e-02	-7.33041628e-02	-1.57454534e-01	-7.26870393e-02	-9.23654247e-02	-1.02397836e-01
K^-	σ_0	1.07082820e-01	1.00573410e-01	1.93148260e-01	1.00993689e-01	1.26963814e-01	1.30057621e-01

Table 3.4: Values used to calculate the mean and resolutions for hadron likelihood based identification.

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 of SIDIS kinematic variables, beam spin asymmetry measurements can be taken at different average values of the kinematic variables. Finally, the structure function ratios $A_{LU}^{\sin\phi}$, $A_{UU}^{\cos\phi}$, and $A_{UU}^{\cos 2\phi}$ can be extracted from each bin. In this chapter, the authors discuss the selection of SIDIS events, the binning used in this analysis, our measurement with associated systematic errors, and the extraction of structure function ratios using the ϕ dependence in each kinematic bin.

4.2 Event Selection and Binning

Event Selection

After particle identification is performed on the event, those events which have a trigger electron and a positive kaon are kept for analysis. We discard events that do not have $W > 2$ and $Q^2 > 1$, because they are not conisdered to be part of the deeply inelastic region. Additionally, to avoid exclusive resonances in the $ep \rightarrow eK^+X$ spectrum, the authors impose a cut on the missing mass of the final state X. For this analysis we use $M_X(ep \rightarrow eK^+X) > 1.25$. Finally, we attempt to perform our measurement in the current fragmentation region where factorization has been proven. This is done by excluding events with $z_h < 0.25$. We also require that $z_h < 0.75$ to avoid exclusive events. This restriction on z_h is not applied to the z_h axis, where we measure across the entire experimentally observed range. After these selection criteria have been applied, the data is sorted into kinematic bin.

Binning

For this study, the authors chose to measure the integrated beam spin asymmetry. This simply means that for a given axis (P_T for example), the events included have all observed values of the other kinematic variables (in this example x, z_h, Q^2). The axes studied are x, Q^2, z_h, and P_T. We chose to use 12 bins in ϕ and 10 bins of the other kinematic variables for a total of 120 analysis bins.

The bins were chosen using a simple method to ensure equal statistics in each bin. The procedure will be 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 just x_1 and x_N if there are N events in the sample. Next, the target number of bins M is chosen (this choice is done by the analyst based on what he/she believes to be the best choice). Finally, the limits of each bin can be chosen simply 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.

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

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 ϕ_h Distributions

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

$$A_i = \frac{1}{P_e} \frac{n_i^+ - n_i^-}{n_i^+ + n_i^-} \tag{4.2}$$

where P_e is the average beam polarization over the dataset. The symbols $n_i^{\pm}$ refer to the number of events counted in bin i with helicity $\pm$. 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 treatment of the statistical uncertainty reported on the measurement includes the contribtion from counts, but not from the uncertainty in P_e which is included in the systematic errors. The uncertainty in a measured observable $\mathcal{O}$ depends on the uncertainty of the parameters used to construct it $th\vec{e}ta$ in the following way.

$$\sigma_{\mathcal{O}}^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\partial \mathcal{O}}{\partial \theta_i} \frac{\partial \mathcal{O}}{\partial \theta_j} \rho_{ij} \sigma_i \sigma_j \tag{4.3}$$

4.4 Extraction of Modulations

Chapter 5

SIDIS Cross Section

5.1 Introduction

5.2 Inclusive Cross Section

5.2.1 Event Selection and Binning

5.2.2 Simulation

5.2.3 Radiative Corrections

5.2.4 Model Comparison

5.3 Results for SIDIS

Chapter 6

TMD Extraction

6.1 Introduction

6.2 EVA

6.2.1 Wandzura Wilzcheck Approximation

6.2.2 Parametrization of TMD Functions

6.3 Results for $cos(2\phi_h)$) Modulation

6.4 Predictions for $cos\phi_h$) Modulation