Studies of Quark Momentum in the Proton by Use of the SIDIS Process

David M. Riser, Ph.D.

University of Connecticut,

The three dimensional structure of protons and neutrons has been a subject of great interest since the early 1990s when advancements in theory produced the Generalized Parton Distribution (GPD) framework for understanding spatial distributions, and Transverse Momentum Dependent Parton Distribution Functions (TMD PDFs) for understanding quark momentum. In this thesis, I present two analyses of the SIDIS process which can be related to TMDs. I have used CLAS data from the E1-F run period, which recorded over 1 billion events of 5.498 GeV electrons colliding with a liquid hydrogen (LH_2) target. First, I discuss analysis of the SIDIS cross section (fully differential in 5 variables) for both charged pions over a wide kinematic range. This analysis observes a non-zero $\cos(\phi_h)$ and $\cos(2\phi_h)$ modulations, and is useful for phenomenological extractions of the Boer-Mulders and Cahn effect terms. I then present results for the beam spin asymmetry (BSA), which is only sensitive to terms containing twist-three distri-

bution functions. The analysis, performed for K^+ mesons, observes a non-zero BSA and therefore implies that twist-three effects are not negligably small in the kinematics used for the study. Unlike the large magnitude of the Sivers asymmetries observed at HERMES for K^+ mesons, the BSA reported here is of the same magnitude as the lighter π^+ meson.

Studies of Quark Momentum in the Proton by Use of the SIDIS Process

David M. Riser

B.S. Physics, Delaware State University, 2013

A Dissertation

Submitted in Partial Fullfilment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

Copyright by

David M. Riser

APPROVAL PAGE

Doctor of Philosophy Dissertation

Studies of Quark Momentum in the Proton by Use of the SIDIS Process

Presented by

David M. Riser,

Major Advisor

Kyungseon Joo

Associate Advisor

Peter Schweitzer

Associate Advisor

Harut Avakian

University of Connecticut

Dedicated to my grandfather, Abdul Qaissaunee

ACKNOWLEDGEMENTS

A network of hard working and caring people have supported and helped me through this journey. Most directly, my supervisor Dr. Kyungseon Joo has supported me from the beginning and shared with me his valuable insights and experience.

At Jefferson lab, I had the opportunity to collaborate with world class physicists. I would like to thank the CLAS collaboration and all of its members. In particular, the door of Dr. Harut Avakian was always open for discussion, and I relied heavily on his expertise in analyzing and interpreting scattering data. I have also had the pleasure of working with Dr. Youri Sharabian on the HTCC, and Dr. Maurizio Ungaro on the GEMC drift chamber geometry.

Dr. Joo has done a great job assembling a strong research group, and I thank each member for their support in this work. For more than a year I worked closely with Dr. Nick Markov, and his expertise in data analysis was the driving force behind the successful extraction of the inclusive cross section. Nick makes himself accessible for any kind of question, and I have made use of his insights many times over the last four years. For discussions of the details of c++, code design, and general computing I have found Dr. Andrey Kim an invaluable resource. I have also enjoyed collaborating with Dr. Stefan Diehl on several small but important

analysis components. From my collaborations with Dr. Nobuo Sato I learned about statistics, global analysis, and monte carlo sampling methods.

I am grateful to the educators who have contributed to my development. Drs. Gerald Dunne, Peter Schweitzer, Alexander Kovner, Richard Jones, Susanne Yelin, Elena Dormidontova, Juha Javanainen and Philip Mannheim of the University of Connecticut all delivered insightful presentations of physics and made the experience of coursework feel like following the path of discovery of the fundamental laws of physics for the first time. I will forever cherish my memories of those lectures. I was first motivated to continue to graduate school by two professors at Delaware State University. Drs. Essaid Zerrad and Gour Pati both spent countless hours with me as an undergraduate, educating me on the basic methods of computational and experimental research and giving me a real taste of research life.

Sharing this experience with others has created several friendships. Brandon Clary and Dr. Frank Cao have been with me since the beginning in 2013 when we started classes at UConn together. I have fond memories of solving problems with both of them during our first two years of classes together. I am also thankful for Kemal Tezgin sharing so many of his lunch breaks with me walking in the forest, explaining theoretical concepts and discussing philosophy. I have also enjoyed interacting with Dr. Freddy Obrecht, who is always working on an interesting project or has found an interesting challenge.

Ultimately, none of this would have been possible without the support of my family. My parents have always provided me with any type of support that I needed, and have always believed in me.

Finally, I would like to thank my wife Molly, who has selflessly followed me around the east coast for the last six years in the pursuit of my education. Not only has she supported my goals every day of my graduate school journey, she helped me in forging good study habits as an undergraduate student as well. Without her love, support, and companionship, this work wouldn't have been possible.

TABLE OF CONTENTS

1. Introduction 1

1.1 Deeply Inelastic Scattering and Structure Functions 1

1.2 The Parton Model and Parton Distribution Functions 4

1.3 HERA 6

1.4 Semi-Inclusive Deeply Inelastic Scattering 8

1.5 Transverse Momentum Dependent Functions 11

1.5.1 Factorization Theorems 13

1.5.2 Unpolarized TMD 14

1.5.3 Transversity TMD 17

1.5.4 Sivers TMD 18

1.5.5 Boer-Mulders TMD 19

1.6 This Thesis 20

2. Experiment 23

2.1 CEBAF 23

2.1.1 Electron Injection & Polarization 24

2.1.2 Acceleration of Electrons 27

2.2 CLAS in Jefferson Lab Experimental Hall B 28

2.2.1 CLAS Torus and Drift Chambers 29

2.2.2 CLAS Cherenkov Counter 31

2.2.3 CLAS Time of Flight Scintillator 32

2.2.4 CLAS Electromagnetic Calorimeter 33

3. Basic Analysis & Corrections 37

3.1 Introduction 37

3.2 Luminosity Calculation 37

3.3 Determination of Good Run List 41

3.4 Helicity Determination 42

3.5 Vertex Corrections 44

3.6 Timing Corrections 45

3.7 Kinematic Corrections 46

4. Particle Identification 52

4.1 Introduction 52

4.2 Electron Identification 52

4.2.1 Electron ID Cuts 53

4.3 Hadron Identification 66

4.3.1 Hadron ID Cuts 68

5. Inclusive Cross Section 83

5.1 Motivation 83

5.2 Inelastic Scattering Model 84

5.3 Analysis 85

5.3.1 Event Selection 85
5.3.2 Binning 87
5.3.3 Simulation 89
5.3.4 Acceptance Corrections (Theory) 90
5.3.5 Acceptance Corrections 92
5.3.6 Radiative Corrections 92
5.3.7 Bin Center Corrections 99
5.3.8 Model Comparison 100

6. SIDIS Cross Section 102
6.1 Introduction 102
6.2 Hadron Identification 104
6.3 Event Selection 106
6.4 Binning 106
6.5 MC Simulation 109
6.5.1 Acceptance Corrections 110
6.6 Radiative Corrections 112
6.7 Parameter Estimation (Fitting) 113
6.7.1 Minimal Coverage in ϕ_h for Fitting 114
6.8 Systematic Uncertainties 114
6.9 Results 120

7. Beam Spin Asymmetry Analysis 128

7.1 Introduction 128

7.2 Event Selection and Binning 129

7.2.1 Event Selection 129

7.2.2 Binning 129

7.3 Measured ϕ_h Distributions 132

7.3.1 Measured Asymmetry Values 132

7.3.2 Statistical Uncertainties 133

7.3.3 Systematic Uncertainties 135

7.3.4 Basic Formalism 135

7.3.5 Sources of Systematic Uncertainty 138

7.3.6 Electromagnetic Calorimeter Fiducial Cuts 139

7.3.7 Missing Mass 139

7.3.8 Confidence Level 141

7.3.9 Kaon Momentum 141

7.3.10 Electromagnetic Calorimeter Energy Cuts 141

7.3.11 Fiducial Cuts on DC and CC 146

7.3.12 Vertex Cuts 146

7.4 Extraction of Modulations 146

7.5 Results 151

References 153

LIST OF FIGURES

1.1 A diagrammatic representation of deeply inelastic scattering. 2

1.2 A recent PDF extraction from HERA data. 5

1.3 $F_2(Q^2)$ results from HERA 7

1.4 Diagrammatic representation of SIDIS with hadronic ϕ_h angle. 8

1.5 Twist two PDFs 11

1.6 HERMES multiplicities fit using the Gaussian ansatz [1]. 15

1.7 COMPASS multiplicities fit using the Gaussian ansatz [1]. 16

1.8 The transversity TMD (left) and Collins FF (right) extracted by global analysis of measurements at HERMES and COMPASS, described in Ref. [2]. 17

1.9 The Sivers asymmetry measured at COMPASS and HERMES for different hadron species identified in the final state, demonstrating a clear non-zero effect. Figure taken from Ref. [3]. 18

1.10 The Boer-Mulders and Cahn effects extracted from HERMES data on the $\cos(2\phi_h)$ moment of SIDIS, reported in [4]. 20

2.1 Diagrammatic representation of CEBAF. 24

2.2 This aerial photograph contains annotations that show the accelerator path and the three experimental halls. 26

2.3 In this diagrammatic representation of CLAS, several important detector sub-systems are labeled. 29
2.4 Diagrams of the CLAS torus field. This figure reproduced from [5] . . 30
2.5 Illustration of a charged particle interacting with cells in the drift chamber. This figure was reproduced from [6] 31
2.6 The CLAS Cherenkov Counter. 32
2.7 The U, V, W structure of CLAS electromagnetic calorimeter is shown above. 35
2.8 GEANT simulation of an electron showering in the EC. This figure originally appeared in [7] 36

3.1 Entries per scalar reading for run 38222. 40
3.2 Charge per scalar reading for run 38222. 40
3.3 Inclusive electrons per file normalized by the total charge accumulated for the file. This quantity is used to make a good run list. 43
3.4 The waveplate position is determined and corrected by plotting the BSA for π^+ mesons as a function of the run. The top panel shows the corrected results, the bottom shows the results before changing the helicity. 43

3.5 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. 44

3.6 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. . . . 46

3.7 This figure shows the deviation from M_p of the W spectrum peak for elastic $ep \to ep$ events (before corrections). 49

3.8 This figure shows the deviation from M_p of the W spectrum peak for elastic $ep \to ep$ events (after ϕ-dependent corrections). 50

3.9 Elastic events shown in the spectrum of W before and after momentum corrections are applied. 51

4.1 Tracks shown in color remain after the application of drift chamber region 1 fiducial cuts to all cuts, shown here as black points. . . . 56

4.2 The selection criteria shown in red is applied to drift chamber region 3. 57

4.3 All negative tracks are shown here in black. In color, the tracks which pass the EC fiducial cut are shown. 59

4.4 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). 60

4.5 The sampling fraction selection boundary is shown here for the nominal value of $N_{sigma} = 4$. 62

4.6 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. 65

4.7 The angle ϕ_{rel} is the azimuthal angle between the central line of the detector and the track. 66

4.8 Correlation between θ_{CC} and the CC segment is shown above, with our selection boundaries overlaid in red. 67

4.9 Shown above: Positive track hits on the region 1 drift chamber, events falling between the red lines are kept for analysis. 69

4.10 Shown above: The difference between the z-vertex position between detected electrons and positive tracks. 70

4.11 Shown above: The distribution of confidence level for all positive tracks after being classified by the highest probability. 72

4.12 Shown above: All positive tracks overlaid with our determination of $\mu(p) \pm \sigma(p)$ for π^+ 74

4.13 Shown above: All positive tracks overlaid with our determination of $\mu(p) \pm \sigma(p)$ for K^+ 75

4.14 Shown above: All positive tracks overlaid with our determination of $\mu(p) \pm \sigma(p)$ for p^+ 76

4.15 Positive hadrons from the Monte Carlo simulation produce a β vs. p simulation that is very similar to data. 81

4.16 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. 81

5.1 Inclusive data and simulation comparison. 85

5.2 Event distributions (θ_e vs p) for data, simulated inelastic events, and simulated elastic events with radiation are shown. The red line indicates the momentum cut applied by restricting $y < 0.7$. The band of events on the right side of the figure contains all un-radiated elastic events. The larger portion of events present on the left side are Bethe-Heitler events which we wish to avoid in our analysis. . 88

5.3 Reconstructed Monte Carlo events displayed for different sectors (by panel) and different Q^2 bins represented by different colors. 93

5.4 Generated Monte Carlo events displayed for different sectors (by panel) and different Q^2 bins represented by different colors. These events were generated with radiative effects. 94

5.5 Reconstructed events from data (black) and Monte Carlo (red) are superimposed to show that the simulation is a good approximation to the physics, a fact important for the accurate calculation of acceptance. 95

5.6 Acceptance corrections applied to the inclusive cross section are shown above in sixty panels that correspond to increasing Q^2 from left to right, and increasing sector number from top to bottom. 96

5.7 Radiative corrections applied to the inclusive cross section are shown above in sixty panels that correspond to increasing Q^2 from left to right, and increasing sector number from top to bottom. 98

5.8 The bin center correction applied to the inclusive cross section is shown for different values of Q^2 (indicated by color) as a function of W on the horizontal axis. This correction is the same for all sectors. 100

5.9 Each panel above shows one of the six sectors of CLAS, the horizontal axis shows W and superimposed on the figure are the cross sections for different bins of Q^2. For simplicity of the visualization each reported value of Q^2 is the bin center. 101

6.1 PID for π^+ used for the SIDIS analysis 105

6.2 Kinematic distributions and binning for electron variables. 107
6.3 Kinematic distributions and binning for hadron variables. 108
6.4 Acceptance corrections for SIDIS 111
6.5 Radiative corrections for π^+ 112
6.6 Corrected yield fits for π^+ 115
6.7 Systematic boundaries on region 1 117
6.8 Systematic boundaries on region 3 118
6.9 Systematic boundaries on sampling fraction 119
6.10 Systematic variations of electron z-vertex cuts. 120
6.11 Systematic variations of the CC fiducial cut for electrons. 121
6.12 Variation of EC U, V, and W cuts used to identify electrons. 122
6.13 $d^5\sigma$ for π^+ 123
6.14 The unpolarized structure functions (π^+) that don't depend on the angle ϕ_h are shown here for different values of x, z, and P_T^2. . . . 124
6.15 The unpolarized structure functions (π^-) that don't depend on the angle ϕ_h are shown here for different values of x, z, and P_T^2. . . . 125
6.16 The unpolarized structure functions (π^+) that don't depend on the angle ϕ_h are shown here for different values of x, z, and P_T^2. . . . 126
6.17 The unpolarized structure functions (π^-) that don't depend on the angle ϕ_h are shown here for different values of x, z, and P_T^2. . . . 127

7.1 Correlation between z and P_T^2 (GeV^2/c^2) for each event in our analysis sample. 130

7.2 The missing mass spectrum for the reaction $ep \to e'K^+X$ is shown after the application of all cuts used in the analyses except for the cut we apply on the missing mass. 131

7.3 Bins used for this analysis are displayed in two dimensions for the z axis. 133

7.4 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. 134

7.5 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. 134

7.6 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. 136

7.7 Boundaries (shown in dark red) and associated uncertainties (shaded red regions surrounding the boundaries) for the EC coordinate cuts. 139

7.8 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. 140

7.9 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. 142

7.10 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. 143

7.11 The relative contribution of each systematic uncertainty to the total is shown above averaged over the bins of the x axis. 144

7.12 The relative contribution of each systematic uncertainty to the total is shown above averaged over the bins of the z axis. 145

7.13 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. 147

7.14 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. 147

7.15 Counts for different helicity states are superimposed for different bins of z. 150

7.16 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. 151

7.17 In this figure the results of this study for postively 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 $A_{LU}^{\sin\phi_h}$. 152

7.18 In this figure the results of this study for postively 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 $A_{LU}^{\sin\phi_h}$. 152

LIST OF TABLES

4.1 Cut parameters used for the DC fiducial cut. 55

4.2 Cut parameters used for the EC fiducial cut. 58

4.3 μ and σ values used to construct the momentum dependent sampling fraction cut. 63

4.4 Values used to calculate the mean and resolutions for hadron probability based identification. 78

5.1 Summary of W and Q^2 binning used for the inclusive cross section. . 87

6.1 Limits for pion momentum used in fitting β. 104

7.1 Different sources of systematic effect considered in this analysis. The magnitude of the effect is shown here averaged over all bins of ϕ_h. 138

Chapter 1

Introduction

Research into the internal structure of protons and neutrons began in 1955 with the pioneering work of Hofstadter and McAllister. Working together at Stanford, the pair measured the electron-proton scattering cross section with 188 MeV electrons incident on hydrogen. Their measured cross section was inconsistent with the cross section predicted for a point-like proton, demonstrating its finite size [9], [10], [11]. Simultaneously, the Brookhaven Cosmotron produced an overwhelming number of new particles; a "particle zoo". Theoretical efforts by Gell-Mann and Zweig showed that the particle zoo could be explained by combinations of fractionally charged particles that Gell-Mann coined quarks [12].

1.1 Deeply Inelastic Scattering and Structure Functions

Learning more about quarks required scattering at higher energies. As the electron energy increases and the wavelength of the virtual photon becomes small $\lambda < r_p$ deeply inelastic scattering (DIS) occurs where the virtual photon interacts with a constituent quark inside the proton and the proton subsequently breaks up. The

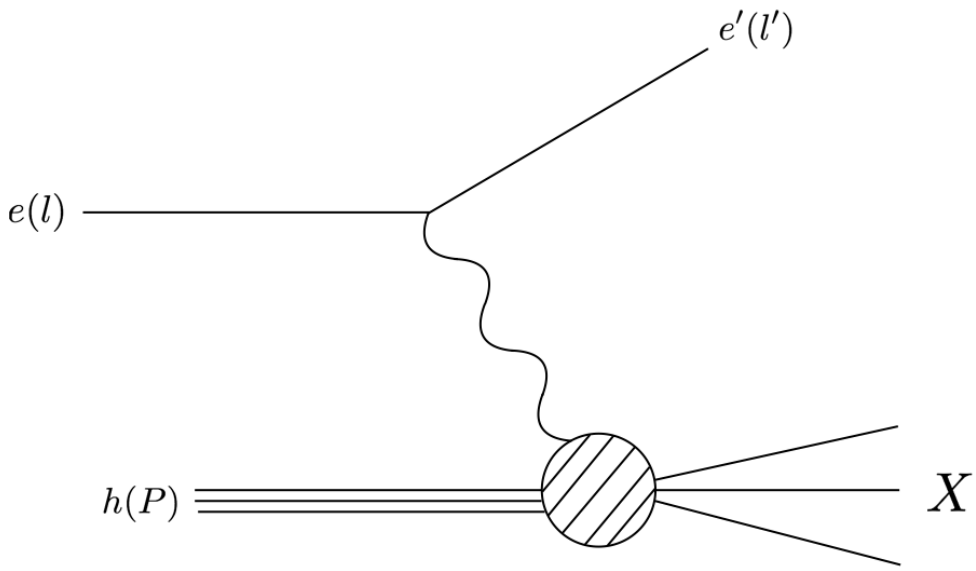

Fig. 1.1: Deeply inelastic scattering (DIS) between a lepton and a hadron, in this thesis an electron and a proton.

cross section for unpolarized DIS can be described using QED in terms of two structure functions F_1 and F_2. As we will see below, these structure functions contain information on quark momentum in the proton.

$$\frac{d^2\sigma}{dxdQ^2} = \frac{4\pi\alpha^2}{Q^4}\left\{\left(1 - y - \frac{m_p^2 y^2}{Q^2}\right)\frac{F_2(x,Q^2)}{x} + y^2 F_1(x,Q^2)\right\} \tag{1.1}$$

The electron kinematics are described by the four momentum transfer $q^\mu = l^\mu - l'^\mu$ (shown in figure 1.1), the negative momentum transfer $Q^2 = -q^\mu q_\mu$ and,

$$x = \frac{Q^2}{2P\cdot q} \qquad\qquad y = \frac{P\cdot q}{P\cdot l} \tag{1.2}$$

where x is called the momentum fraction and y is the inelasticity. The variable x (which appears throughout this thesis) describes the fraction of the total proton momentum carried by the quark struck in the scattering. The constant α is the fine structure constant of QED. For the case that $Q^2 \gg m_p^2 y^2$ the cross section can be approximated as shown below.

$$\frac{d^2\sigma}{dxdQ^2} \approx \frac{4\pi\alpha^2}{Q^4}\left\{(1-y)\frac{F_2(x,Q^2)}{x} + y^2 F_1(x,Q^2)\right\} \tag{1.3}$$

First measurements of the DIS structure functions [13] showed that $F_1(x,Q^2)$ and $F_2(x,Q^2)$ are approximately independent of Q^2 (Bjorken scaling) and that

they appear to be related to each other through the Callan-Gross relation given by:

$$F_2(x) = 2xF_1(x) \tag{1.4}$$

Both of these features are predicted by the parton model, a tool introduced by Richard Feynman in 1969 [14].

1.2 The Parton Model and Parton Distribution Functions

Feynman noted that high energy electron-proton collisions occur on a very short timescale, and can be imagined as electrons interacting electromagnetically (or electroweakly) with single quarks in the proton. Elastic electron-quark scattering is calculable, but the quark in the proton is not free. Therefore the concept of parton distribution functions $q_i(x)$ was introduced to model quark interaction with the other constituents of the proton (the index i refers to the quark flavor). The parton distribution functions (PDFs) describe the probability of finding a quark of flavor i in a proton (or other hadron) with momentum fraction x (this interpretation is no longer valid once the Q^2 evolution is calculated using DGLAP). By using the parton model the DIS cross section can be predicted in terms of the PDFs as follows:

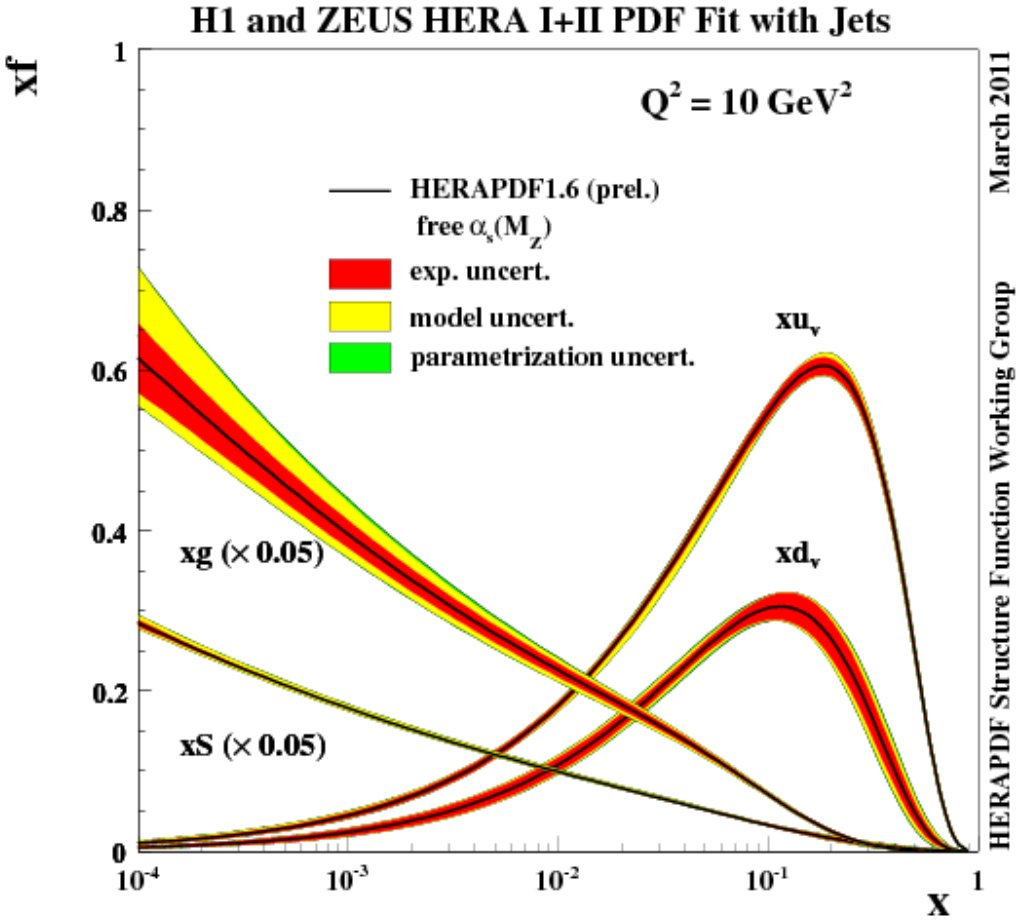

Fig. 1.2: PDFs extracted at $Q^2 = 10\ GeV^2/c^2$ from HERA data, showing the dominance of gluons and sea quarks at lower momentum fraction x.

$$\frac{d^2\sigma}{dxdQ^2} = \frac{4\pi\alpha^2}{Q^4}\left\{1 - y + \frac{y^2}{2}\right\}\sum_i e_i^2 q_i^p(x) \tag{1.5}$$

When equating (1.3) with (1.5) Bjorken scaling and the Callan-Gross relation are predicted by the parton model.

$$F_2(x, Q^2) = 2xF_1(x, Q^2) = x\sum_i e_i^2 q_i^p(x) \tag{1.6}$$

These first triumphs of the parton model led to in depth studies of the structure functions at HERA.

1.3 HERA

HERA studied DIS extensively over a large range of Q^2 by colliding 27.5 GeV electrons and protons with various energies up to 920 GeV. The H1 and ZEUS experiments mapped proton structure functions in great detail. The results 1.3 display a weak dependence on Q^2; a feature calculable using the DGLAP evolution equations. The PDFs extracted in Fig. 1.2 from these structure functions show that almost half of the proton momentum is carried by the sea. These results demonstrated that a simple picture of the proton as three valence quarks (uud) is not sufficient.

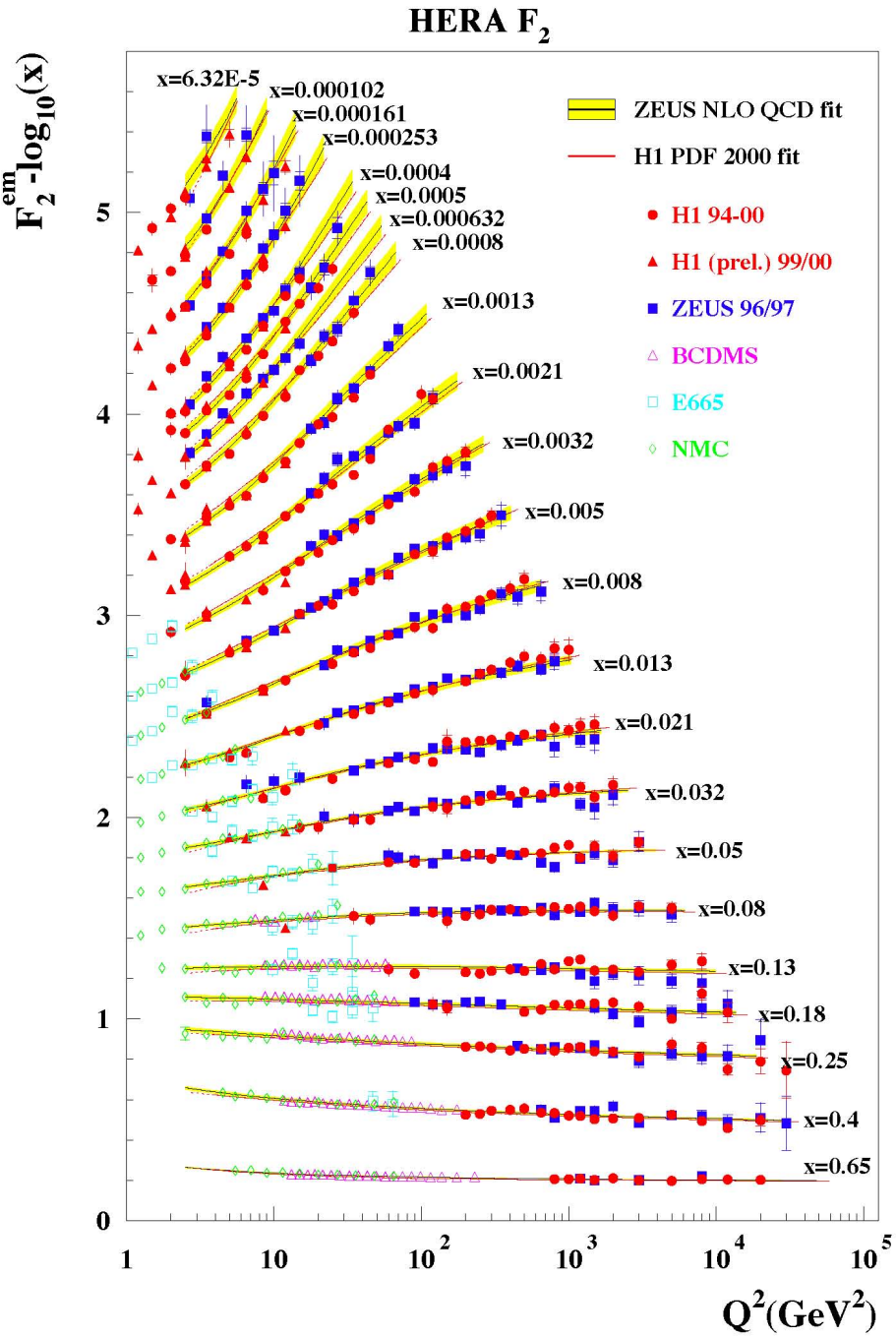

Fig. 1.3: HERA experiments ZEUS and H1 mapped out the structure function F_2, which displayed the predicted Bjorken scaling for most values of x. The mild Q^2 dependence of the PDFs can be calculated in perturbative QCD using the DGLAP evolution equations for the PDFs, providing an understanding for the observed dependence. This also provided an important test of perturbative QCD.

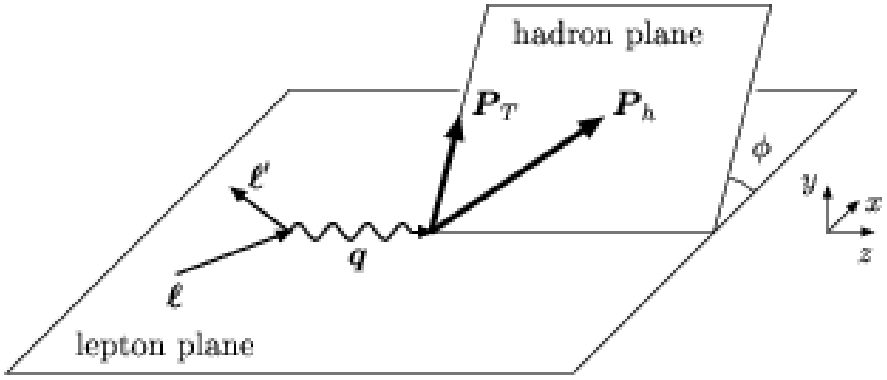

Fig. 1.4: In a SIDIS event the scattered electron is detected as well as one final state hadron, allowing for the determination of the angle ϕ_h which appears frequently in the SIDIS cross section.

1.4 Semi-Inclusive Deeply Inelastic Scattering

DIS events where one hadron is detected in the final state are called semi-inclusive DIS (SIDIS) events. The detection of a hadron in the final state provides information needed to go beyond a discussion of quark momentum along the hard scattering direction and into the plane transverse to it. A diagrammatic representation of the reaction Fig. (1.4) shows the hadron scattering out of the lepton plane with angle ϕ (also noted ϕ_h in other places).

Analogously to the manner in which the DIS cross section can be described by structure functions, the SIDIS cross section (Eq. 1.7) can be decomposed in the one-photon exchange approximation into a set of 18 structure functions (depending on the polarization direction of the electron and proton) [15], [16]. The structure functions F_{ij} are indexed by the beam polarization i and the target

polarization j, which take on values of U, L, and T for unpolarized, longitudinal, and transverse respectively.

$$
\begin{aligned}
&\frac{d\sigma}{dx\ dQ^2\ d\psi\ dz\ d\phi_h\ dP_{h\perp}^2} = \\
&\frac{\alpha^2 y}{4(k\cdot P)x^2Q^2(1-\varepsilon)}\left(1+\frac{\gamma^2}{2x}\right)\Big[F_{UU,T}+\varepsilon F_{UU,L}+\sqrt{2\varepsilon(1+\varepsilon)}F_{UU}^{\cos\phi_h}\cos\phi_h \\
&+\varepsilon F_{UU}^{\cos(2\phi_h)}\cos(2\phi_h)+\lambda_e\sqrt{2\varepsilon(1-\varepsilon)}F_{LU}^{\sin\phi_h}\sin\phi_h \\
&+s_{||}\left(\sqrt{2\varepsilon(1+\varepsilon)}F_{UL}^{\sin\phi_h}\sin\phi_h+\varepsilon F_{UL}^{\sin(2\phi_h)}\sin(2\phi_h)\right) \\
&+\lambda_e s_{||}\left(\sqrt{1-\varepsilon^2}F_{LL}+\sqrt{2\varepsilon(1-\varepsilon)}F_{LL}^{\cos\phi_h}\cos\phi_h\right) \\
&+s_\perp\Big[\sin(\phi_h-\phi_S)\left(F_{UT,T}^{\sin(\phi_h-\phi_S)}+\varepsilon F_{UT,L}^{\sin(\phi_h-\phi_S)}\right) \\
&+\varepsilon F_{UT}^{\sin(\phi_h+\phi_S)}\sin(\phi_h+\phi_S)+\varepsilon F_{UT}^{\sin(3\phi_h-\phi_S)}\sin(3\phi_h-\phi_S) \\
&+\sqrt{2\varepsilon(1+\varepsilon)}F_{UT}^{\sin\phi_S}\sin\phi_S+\sqrt{2\varepsilon(1+\varepsilon)}F_{UT}^{\sin(2\phi_h-\phi_S)}\sin(2\phi_h-\phi_S)\Big] \\
&+s_\perp\lambda_e\Big[\sqrt{1-\varepsilon^2}F_{LT}^{\cos(\phi_h-\phi_S)}\cos(\phi_h-\phi_S)+\sqrt{2\varepsilon(1-\varepsilon)}F_{LT}^{\cos\phi_S}\cos\phi_S \\
&+\sqrt{2\varepsilon(1-\varepsilon)}F_{LT}^{\cos(2\phi_h-\phi_S)}\cos(2\phi_h-\phi_S)\Big]\Big]
\end{aligned}
$$

In the SIDIS cross section the following hadronic kinematic variables are used, in addition to those defined in Eq. (1.2).

$$z=\frac{P\cdot P_h}{P\cdot q} \qquad\qquad \gamma=\frac{2Mx}{Q} \tag{1.7}$$

Additionally, the ratio ε of the longitudinal and transverse photon flux is introduced:

$$\varepsilon = \frac{1 - y - \frac{1}{4}\gamma^2 y^2}{1 - y + \frac{1}{2}y^2 + \frac{1}{4}\gamma^2 y^2} \tag{1.8}$$

.

By experimentally controlling the polarization of the electron (λ_e) and target ($s_\perp$, $s_{||}$) different terms of the cross section can be observed in isolation. For the remainder of this thesis our discussion will focus on unpolarized protons, reducing the number of structure functions to five:

$$\frac{d\sigma}{dx\ dQ^2\ dz\ d\phi_h\ dP_{h\perp}^2} = \frac{\pi\alpha^2 y}{2(k\cdot P)x^2Q^2(1-\varepsilon)}\left(1+\frac{\gamma^2}{2x}\right)\Big[F_{UU,T} + \varepsilon F_{UU,L} + \sqrt{2\varepsilon(1+\varepsilon)}F_{UU}^{\cos\phi_h}\cos\phi_h + \varepsilon F_{UU}^{\cos(2\phi_h)}\cos(2\phi_h) + \lambda_e\sqrt{2\varepsilon(1-\varepsilon)}F_{LU}^{\sin\phi_h}\sin\phi_h\Big]$$

.

The SIDIS structure functions are actively being studied for their connection to transverse momentum dependent functions.

nucleon / quark	U	L	T
U	f_1		$h_1^\perp$
L		g_{1L}	$h_{1L}^\perp$
T	$f_{1T}^\perp$	g_{1T}	h_1, $h_{1T}^\perp$

Fig. 1.5: Twist two PDFs. The leftmost column refers to the polarization of the nucleon, while the top row refers to the polarization of the quark inside of the nucleon. Although not explicit in the table, all of these functions depend on the quark flavor, on x and k_T and the scale Q^2.

1.5 Transverse Momentum Dependent Functions

Two types of transverse momentum dependent functions contribute to the SIDIS structure functions. The transverse momentum dependent parton distribution functions (TMD PDFs) depend on the momentum fraction x as well as the quark momentum in the transverse plane k_T, and are denoted $f^a(x, k_T^2)$. Fragmentation functions describe the probability of a quark flavor a fragmenting into the observed hadron h and are denoted by $D^a(z, p_T^2)$. At leading order (called twist-two) there are eight TMD PDFs (see table 1.5 for a list of twist two PDFs). At twist-three an additional sixteen TMD PDFs are introduced, suppressed by a factor of M_p/Q. Two twist-two TMD FFs are also releavant for the processes studied in this thesis. The unpolarized fragmentation function D_1 describes the probability for an unpolarized quark to fragment into an unpolarized hadron, while the Collins fragmentation function $H_1^\perp$ describes the probability of a transversely polarized quark to fragment into an unpolarized hadron.

In the TMD framework, the structure functions can be expressed as convolutions of PDFs and FFs. The following notation $\mathcal{C}[\omega f D]$ was introduced in [16]

to simplify life.

$$\mathcal{C}[\omega f D] = x \sum_a e_a^2 \int d^2\boldsymbol{p}_T d^2\boldsymbol{k}_T \delta^{(2)}\left(z\boldsymbol{k}_T + \boldsymbol{p}_T - \boldsymbol{P}_{h\perp}\right) \omega(\boldsymbol{k}_T, \boldsymbol{p}_T) f^a(x, k_T^2) D^a(z, p_T^2) \tag{1.9}$$

The convolution operation is denoted by the $\mathcal{C}$ character, while the w refers to a kinematic weighting function which depends on $\boldsymbol{k}_T$ and $\boldsymbol{p}_T$ (and can be one), f is the PDF and D is the FF. Using this notation, the five structure functions appearing in the unpolarized cross section are shown below (up to order $1/Q$):

$$F_{UU,T} = \mathcal{C}[f_1 D_1] = x \sum_a e_a^2 \int d^2\boldsymbol{p}_T d^2\boldsymbol{k}_T \delta^{(2)}\left(z\boldsymbol{k}_T + \boldsymbol{p}_T - \boldsymbol{P}_{h\perp}\right) f_1^a(x, k_T^2) D_1^a(z, p_T^2) \tag{1.10}$$

.

$$F_{UU,L} = 0 \tag{1.11}$$

$$F_{UU}^{\cos\phi_h} = \frac{2M}{Q}\mathcal{C}\Big[-\frac{\hat{\boldsymbol{h}}\cdot\boldsymbol{k}_T}{M_h}\Big(xhH_1^\perp + \frac{M_h}{M} f_1 \frac{\tilde{D}^\perp}{z}\Big) - \frac{\hat{\boldsymbol{h}}\cdot\boldsymbol{p}_T}{M}\Big(xf^\perp D_1 + \frac{M_h}{M} h_1^\perp \frac{\tilde{H}}{z}\Big)\Big] \tag{1.12}$$

$$F_{UU}^{\cos 2\phi_h} = \mathcal{C}\Big[-\frac{2(\hat{\boldsymbol{h}}\cdot\boldsymbol{k}_T)(\hat{\boldsymbol{h}}\cdot\boldsymbol{p}_T)-\boldsymbol{k}_T\cdot\boldsymbol{p}_T}{MM_h}h_1^\perp H_1^\perp\Big] \tag{1.13}$$

$$F_{LU}^{\sin\phi} = \frac{2M}{Q}\mathcal{C}\Big[-\frac{\hat{\boldsymbol{h}}\cdot\boldsymbol{k}_T}{M_h}\Big(xeH_1^\perp+\frac{M_h}{M}f_1\frac{\tilde{G}^\perp}{z}\Big)+\frac{\hat{\boldsymbol{h}}\cdot\boldsymbol{p}_T}{M}\Big(xg^\perp D_1+\frac{M_h}{M}h_1^\perp\frac{\tilde{E}}{z}\Big)\Big] \tag{1.14}$$

In the above equations, the symbol $\hat{h}$ is a unit vector that points along the direction of the outgoing hadron.

1.5.1 Factorization Theorems

While the details of the derivation of Eq. 1.7 are left to the literature (see reference [15] and reference [16] for an overview), this introduction would not be complete without a mention of the importance of factorization theorems. The decomposition of structure functions in SIDIS (as well as Drell-Yan and Semi-Inclusive Anhillation of e^+e^-) as non-perturbative TMD functions and some calculable hard scattering part is the essence of factorization theorems. Factorization has not been demonstrated for all twist, but has been demonstrated for SIDIS at leading twist (twist-two) in recent years. The most comprehensive reference on the subject is John Collins' book (see Ref. [17]).

1.5.2 Unpolarized TMD

The unpolarized SIDIS structure functions can be studied through the observable known as multiplicity, measured by both HERMES (27.6 GeV electrons, described in Ref. [18]) and COMPASS (160 GeV muons, described in Ref. [19]). The multiplicity of HERMES is defined below (the COMPASS definition is similar):

$$M_n^h(x_B, Q^2, z, P_T) = \frac{1}{\frac{d^2\sigma_{DIS}}{dx_B\ dQ^2}} \int_0^{2\pi} \frac{d^5\sigma_h}{dx_B\ dQ^2 dz dP_T d\phi_h} d\phi_h \tag{1.15}$$

.

This can also be written in terms of the corrected number of events N^h.

$$M_n^h(x_B, Q^2, z, P_T) = \frac{1}{\frac{d^2N_{DIS}}{dx_B\ dQ^2}} \int_0^{2\pi} \frac{d^5N^h}{dx_B\ dQ^2 dz dP_T d\phi_h} d\phi_h \tag{1.16}$$

The unpolarized TMD PDF f_1 and FF D_1 have been extracted from these data with good results by using Gaussian transverse momentum dependence [1]. A representative example of the results from HERMES are shown in figure 1.6, and the COMPASS results can be seen in Fig. (1.7). At low $P_{hT} \ll Q^2$ these functions can be described as follows:

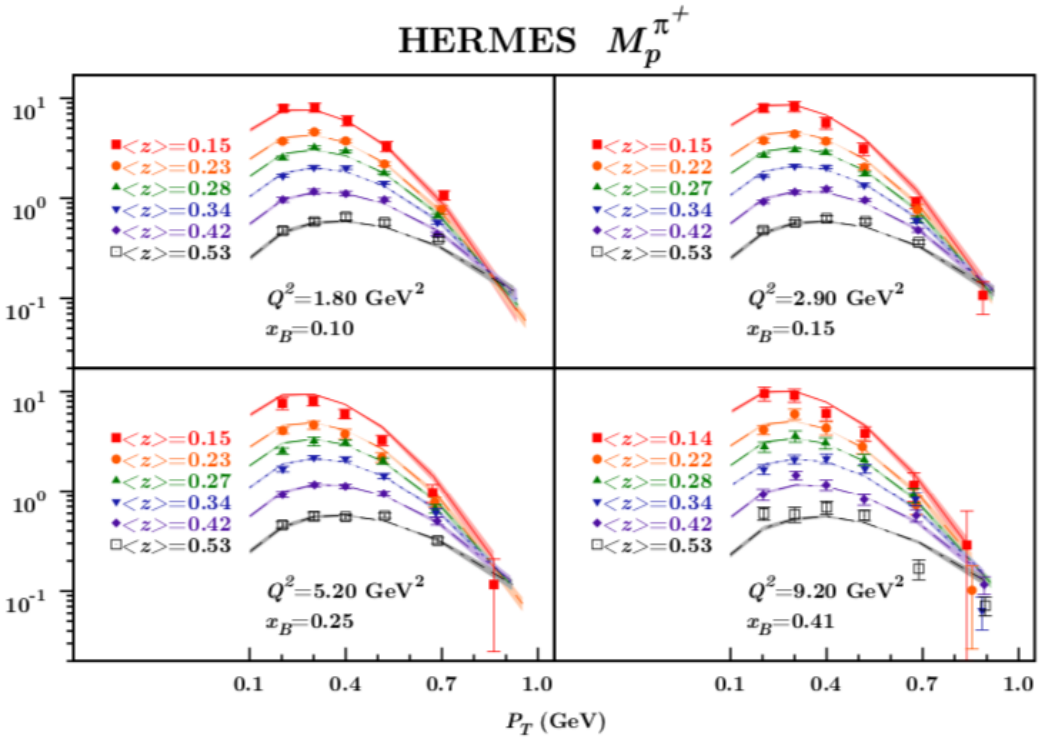

Fig. 1.6: HERMES multiplicities fit using the Gaussian ansatz [1].

$$f_{q/p}(x, k_\perp) = f_{q/p}(x) \frac{e^{-k_\perp^2/\langle k_\perp^2 \rangle}}{\pi \langle k_\perp^2 \rangle} \tag{1.17}$$

$$D_{h/q}(z, p_\perp) = D_{h/q}(z) \frac{e^{-p_\perp^2/\langle p_\perp^2 \rangle}}{\pi \langle p_\perp^2 \rangle} \tag{1.18}$$

.

Where the PDF $f_1(x)$ and FF $D_1(z)$ are (mostly) well known from DIS, Drell-Yan, and e^+e^- collisions.

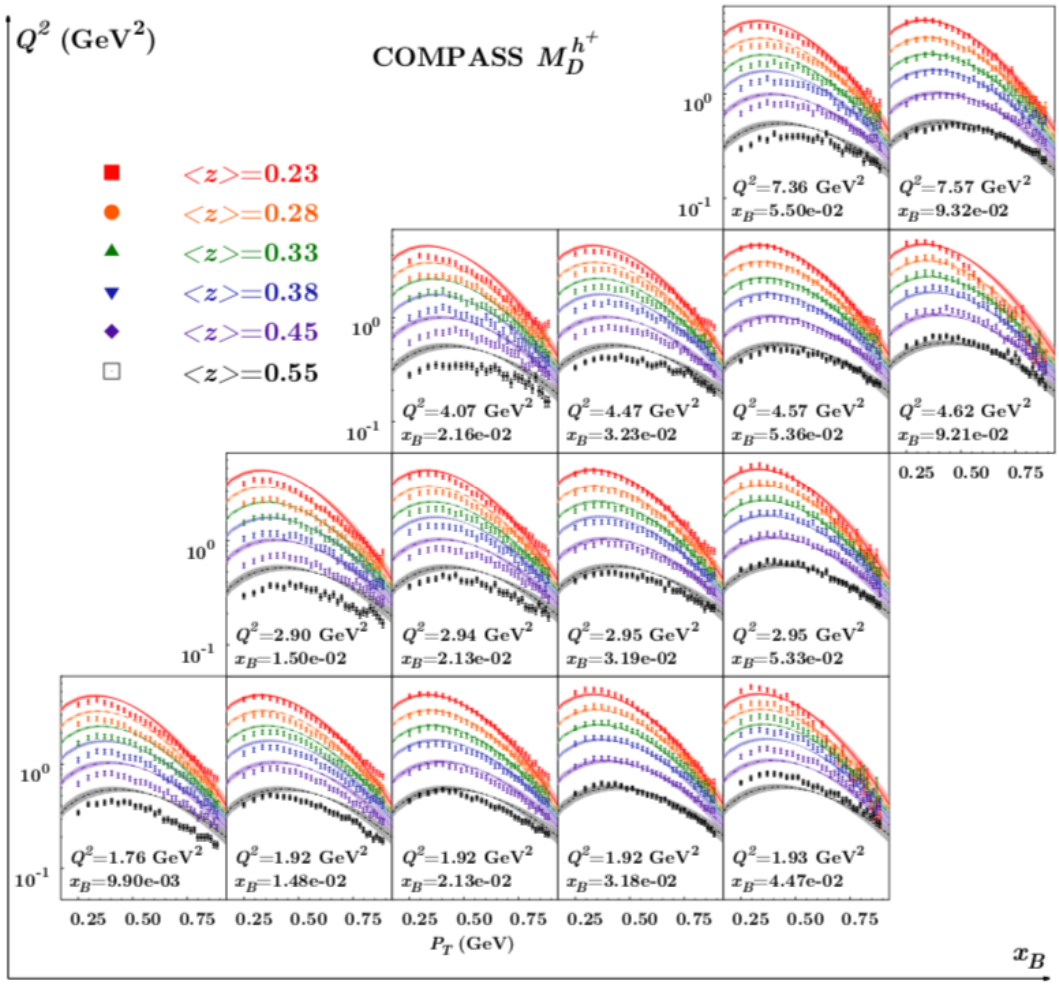

Fig. 1.7: COMPASS multiplicities fit using the Gaussian ansatz [1].

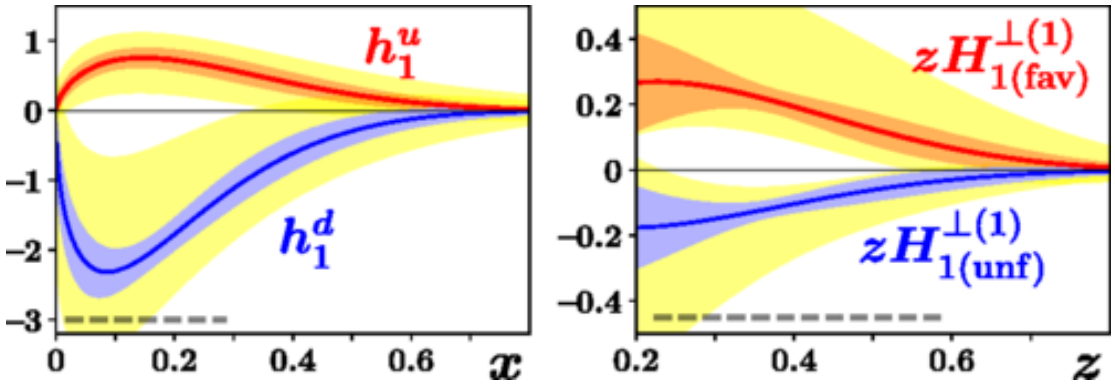

Fig. 1.8: The transversity TMD (left) and Collins FF (right) extracted by global analysis of measurements at HERMES and COMPASS, described in Ref. [2].

1.5.3 Transversity TMD

The Transversity TMD $h_1(x, k_\perp)$ is often extracted together with the Sivers TMD $f_{1T}^\perp(x, k_\perp)$ as they can both be accessed by using a transversely polarized target and measuring the target spin asymmetry defined below.

$$A_{UT}^h(\phi_h, \phi_S) = \frac{1}{s_\perp} \frac{d\sigma^h(\phi_h, \phi_S) - d\sigma^h(\phi_h, \phi_S + \pi)}{d\sigma^h(\phi_h, \phi_S) + d\sigma^h(\phi_h, \phi_S + \pi)} \tag{1.19}$$

Because of the coupling of transversity h_1 to the Collins fragmentation function (both chiral-odd), the asymmetry is often called the Collins asymmetry. The full results are described in Ref. [20], where the HERMES collaboration observed non-zero Collins asymmetries for charged pions and positively charged kaons. Transversity was recently extracted simultaneously with the Collins fragmentation function in Ref. [2].

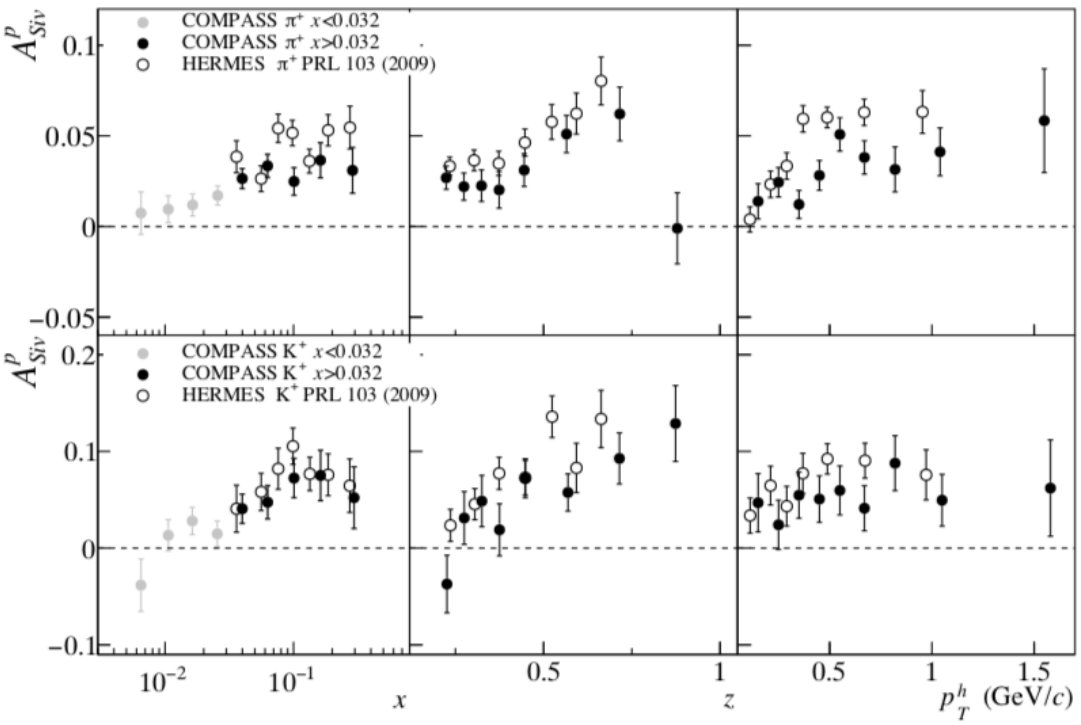

Fig. 1.9: The Sivers asymmetry measured at COMPASS and HERMES for different hadron species identified in the final state, demonstrating a clear non-zero effect. Figure taken from Ref. [3].

1.5.4 Sivers TMD

Naively, one might throw away the Sivers function altogether for its time reversal oddity, assuming that the hadronic tensor $W^{\mu\nu}$ should be time reversal invariant. However, the Sivers asymmetry has been observed by Jlab Hall A, HERMES and COMPASS and it is non-trivially large [21], [22], owing to QCD interactions in the final state. Noteworthy for this thesis, the observed asymmetries were largest for K^+.

An important property of TMDs that hasn't been stated yet is universality; the non-perturbative functions shouldn't depend on the process used to probe

them. For both the Sivers and Boer-Mulders functions there exists a *modified universality*, because the functions are expected to appear with opposite signs in Drell-Yan and SIDIS processes [23].

1.5.5 Boer-Mulders TMD

The Boer-Mulders function, like the Sivers function is time reversal odd. From measurements of unpolarized moments for $\cos\phi_h$ and $\cos(2\phi_h)$ from HERMES and COMPASS one can access the Boer-Mulders function. HERMES defined the observable as follows:

$$\langle\cos\phi_h\rangle = \frac{\int_0^{2\pi}\cos\phi_h d^5\sigma(x_B,Q^2,z,P_T,\phi_h)d\phi_h}{\int_0^{2\pi} d^5\sigma(x_B,Q^2,z,P_T,\phi_h)d\phi_h}, \tag{1.20}$$

$$\langle\cos(2\phi_h)\rangle = \frac{\int_0^{2\pi}\cos(2\phi_h) d^5\sigma(x_B,Q^2,z,P_T,\phi_h)d\phi_h}{\int_0^{2\pi} d^5\sigma(x_B,Q^2,z,P_T,\phi_h)d\phi_h} \tag{1.21}$$

.

A twist-three ($\cos(\phi_h)$) or twist-four ($\cos(2\phi_h)$) contribution to both of these observables exists, known as the Cahn effect. The Cahn term is proportional to the unpolarized functions $f_1 D_1$. Although extractions of the Boer-Mulders function are not abundant, one study has reported extracting both Boer-Mulders and Cahn from the $\cos(2\phi_h)$ moments of HERMES and COMPASS [4]. The asymmetry is non-zero, and as the authors note, the uncertainty on Boer-Mulders is still quite

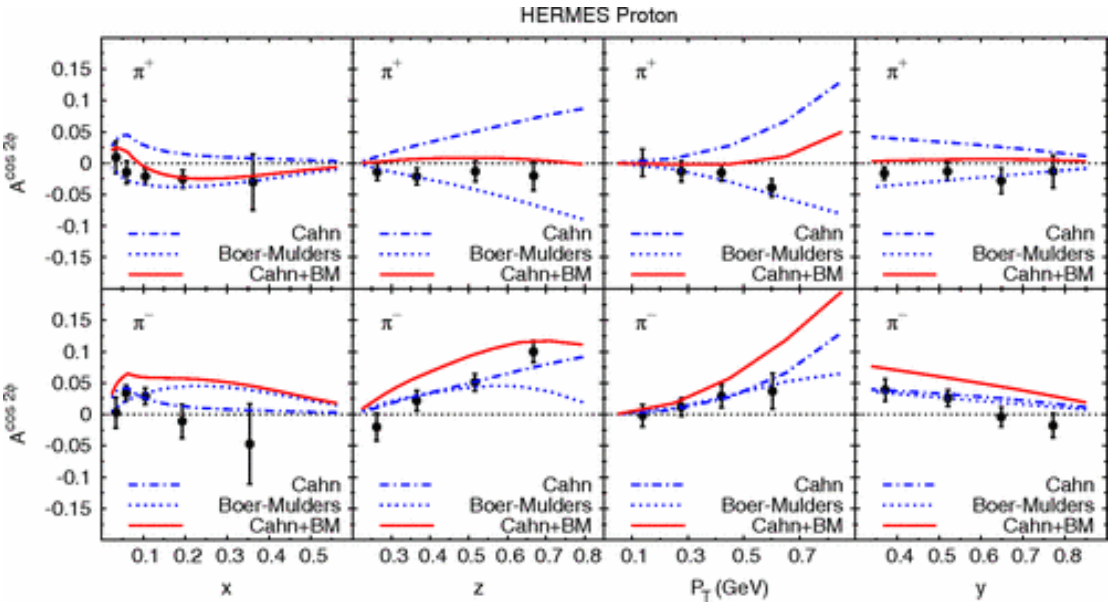

Fig. 1.10: The Boer-Mulders and Cahn effects extracted from HERMES data on the $\cos(2\phi_h)$ moment of SIDIS, reported in [4].

large, and a multidimensional analysis of SIDIS data would help in constraining the parameters from various experiments performed in complementary kinematics.

1.6 This Thesis

The unpolarized cross section for charged pions ($\pi^{\pm}$) is presented in chapter 6. This analysis extends the work of Nathan Harrison in his thesis work [24]. Our results for the cross section, useful in their own right (regardless of the framework used to explain the physics inside of structure functions), are a multidimensional measurement which can be used to constrain the Boer-Mulders and Cahn effects in the kinematics of Jefferson Lab.

In chapter 7, measurements of beam spin asymmetries (BSA) for positively

charged K-mesons (kaons or K^+) are presented. The BSA is defined as,

$$BSA = \frac{d\sigma^+ - d\sigma^-}{d\sigma^+ + d\sigma^-} = \frac{A_{LU}^{\sin\phi_h} \sin\phi_h}{1 + A_{UU}^{\cos\phi_h} \cos\phi_h + A_{UU}^{\cos(2\phi_h)} \cos(2\phi_h)} \tag{1.22}$$

where the superscript $\pm$ refers to the helicity value λ_e and the coefficient $A_{LU}^{\sin\phi}$ is shown below.

$$A_{LU}^{\sin\phi} = \sqrt{2\,\varepsilon(1-\varepsilon)} \frac{F_{LU}^{\sin\phi}}{F_{UU,T} + \varepsilon F_{UU,L}} \tag{1.23}$$

The unpolarized coefficients are defined in a similar way.

$$A_{UU}^{\cos\phi_h} = \sqrt{2\varepsilon(1+\varepsilon)} \frac{F_{UU}^{\cos\phi_h}}{F_{UU,T} + \varepsilon F_{UU,L}} \tag{1.24}$$

$$A_{UU}^{\cos(2\phi_h)} = \varepsilon \frac{F_{UU}^{\cos(2\phi_h)}}{F_{UU,T} + \varepsilon F_{UU,L}} \tag{1.25}$$

There are several interesting motivations for this measurement. First, the TMD content of $F_{LU}^{\sin\phi}$ is purely twist-three expressed in terms of quark-gluon correlators. If twist-three contributions are vanishingly small at Jefferson Lab 6 GeV kinematics, the observed asymmetry should be zero. This was not the case for pions, as observed [25] [26]. One interesting twist-three TMD PDF that appears is $e(x, k_T^2)$, which describes the transverse force acting on the struck quark

just after it has absorbed the virtual photon. Additionally, the magnitude of the Sivers effect for K^+ was surprisingly large, however in most other measurements K^+ signals have generally followed π^+ signals.

I have reviewed the experimental hardware that enables this thesis work in chapter 2. I describe basic analysis procedures that are common to both measurements in chapter 3 before discussing particle identification in chapter 4. I present results for the inclusive scattering cross section in chapter 5, followed by the SIDIS cross section in chapter 6. Finally, I present my findings for the BSA measurement in chapter 7.

Chapter 2

Experiment

Jefferson National Lab houses the continuous electron beam accelerator facility (CEBAF) which is currently capable of providing 11 GeV electron beams to three experimental end stations, and 12 GeV to a fourth. The data analyzed in this study are from the run period E1-F, which occurred (along with several other runs) before the 12 GeV CEBAF upgrade. The details in this chapter describe the accelerator and CLAS detector at the time of the E1-F run period.

2.1 CEBAF

Proposed in 1982 and constructed between 1987-1997 the CEBAF accelerator at Jefferson lab was composed of a pair of linear accelerators and 9 recirculating arcs arranged in a racetrack shape [27, 28]. Originally designed to provide 4 GeV unpolarized electrons to three experimental halls, CEBAF was fitted with twin polarized electron guns, and upgraded to 6 GeV beam energy before the E1-F run period. CEBAF was built to provide an extremely high duty factor and an

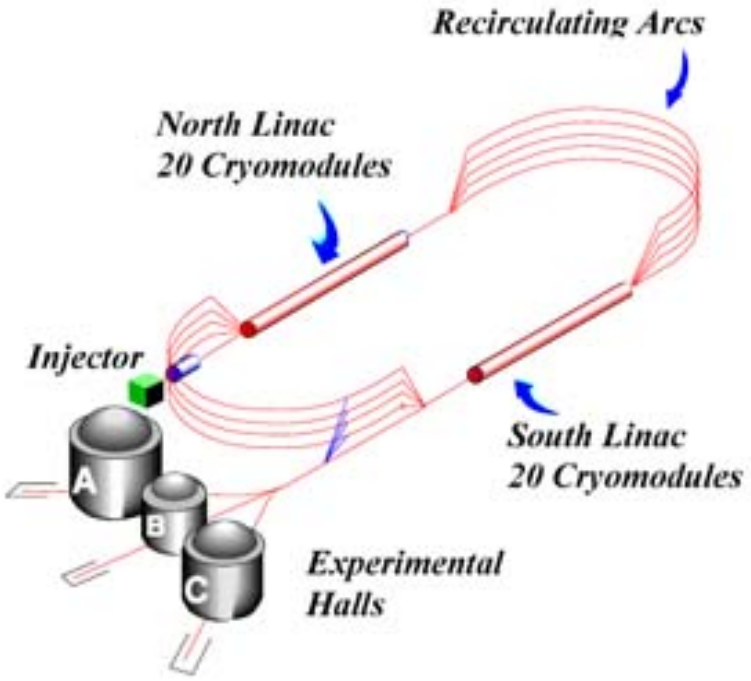

Figure 1: Layout of CEBAF.

Fig. 2.1: Diagrammatic representation of CEBAF.

average beam current of up to 200 μA.

2.1.1 Electron Injection & Polarization

CEBAF's injector provided 45 MeV electrons with 70%-80% longitudinal polarization for the main accelerator. In order to provide an apparent continuous stream of events to the detectors, electron bunches were produced at a rate of $f = 1497\ Mhz$. To accommodate different requests for energy and beam current simultaneously, the injector produced three interspersed bunch trains at a

frequency of 499 Mhz. The output injector energy $E = 45\ MeV$ was chosen so that injected electrons were *sufficiently relativistic*. In other words, when bunches of electrons at different energies simultaneously passed through the linear accelerators (LINACs), the relative phase difference (between different energy passes) accumulated over the distance remained less than 1°.

During early accelerator construction, the need for polarized electrons became apparent and the final polarized electron gun was produced in 2000. Production of polarized electrons was achieved by using twin polarized electron guns mounted at 15° with respect to the injector beam axis. Inside of each gun, electrons were liberated from gallium arsenide photo-cathodes by three independent diode lasers operated at a repetition rate of 499 Mhz. During polarized production, the diodes operated at a central wavelength of 850 nm. By manipulating the laser polarization using a Pockels cell, the electron beam spin is flipped at a rate of 60 Hz. Throughout the run period, an overall phase difference could be introduced by rotation of a wave-plate. As will be discussed in some detail later, changes in beam helicity due to wave-plate settings must be removed from the recorded data. After acceleration to 5 MeV the beam polarization was measured in the injector facility using a Mott polarimeter.

Fig. 2.2: This aerial photograph contains annotations that show the accelerator path and the three experimental halls.

2.1.2 Acceleration of Electrons

The north and south LINACs were responsible for increasing the energy of the electrons from 45 MeV up to an impressive ≈ 5.7 GeV before delivery to the experimental halls A, B, and C. In order to achieve this each bunch of electrons was accelerated through ten stages (five passes through each LINAC). The strong electric field needed to accelerate electron bunches was confined inside of superconducting 5-cell elliptical cavities. These cavities were machined from niobium, and operated at a temperature of 2.2° K. Developed at Cornell University, the cavities were operated at $1,500$ Mhz with a gradient greater than 5 MV/m and a quality factor $Q_0 \geq 3 \cdot 10^9$ (the Q factor describes the monochromaticity of the cavity and is defined $Q_0 = f_0/\Delta f$). Each cavity was sealed inside of a cryo-unit, four such units were connected together to form an 8.5 meter cryo-module. Each LINAC was composed of 20 such cryo-modules connected together to increase electron energies by more than 500 MeV per pass.

The radio frequency (RF) that powers each cavity was sourced by a water-cooled 5 kilowatt Klystron located in groups of eight above each cryo-module. Phase locking of each cavity with a master oscillator ensured that the difference in phase between all cavities was less than one degree. The important super conductivity was maintained by circulation of liquid helium at $2.2^{\circ}K$. Production of liquid helium occurred on-site at the 5 kW helium liquefaction plant.

After re-circulation of fives passes through each LINAC, full energy beams were delivered to the halls. Bunches were separated using an RF separator before entering their respective experimental halls. The beamline leading to the halls was also equipped with the ability to separate bunches before they completed five full passes delivering either three full energy beams, two full energy beams and one lower, or one full energy beam and two of lower energy. This capability, combined with the flexible beam polarization and beam current provided by the injector ensured that each hall could experiment at its desired settings simultaneously.

2.2 CLAS in Jefferson Lab Experimental Hall B

The CEBAF large acceptance spectrometer (CLAS), housed in Jefferson lab's Hall B, was used to record the E1-F dataset (as well as numerous others) that is used in this study. Capable of detecting particles over a very large angular range, the CLAS detector covered almost the full 4π solid angle around the target region. The detector was also designed to perform efficiently for particles with a large range of momenta between 0.5-6.0 GeV. The overall detector design consisted of a large superconducting magnet that produced a toroidal field (this magnet was referred to as the torus), and six ideally identical *sectors*. Each sector of CLAS contained an identical set of sub-systems. After combining information from all sub-systems and running the reconstruction algorithm, complete events

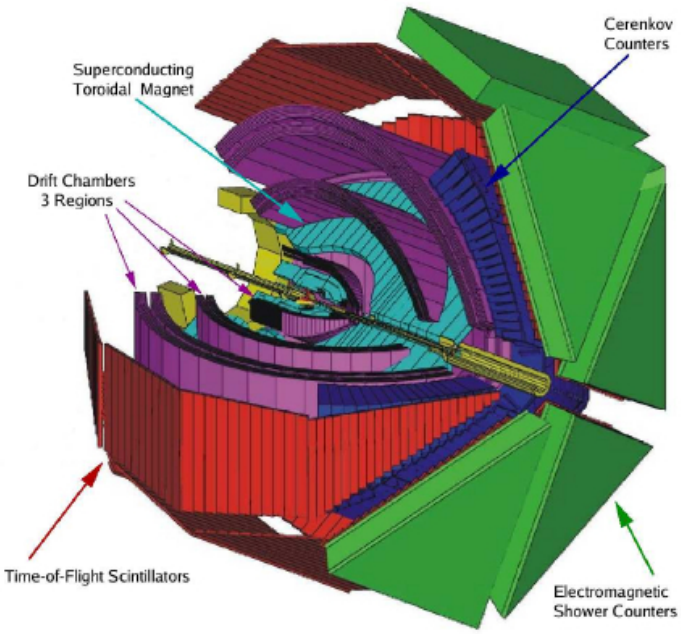

Fig. 2.3: In this diagrammatic representation of CLAS, several important detector sub-systems are labeled.

are measured. This capability made CLAS unique in comparison with the arm style spectrometers of halls A and C. The major components of CLAS are listed below.

2.2.1 CLAS Torus and Drift Chambers

Measuring the momentum of charged particles with $p > 200\ MeV/c$ was accomplished by measuring the curvature of the particle trajectory as it passed through the CLAS toroidal magnetic field. Six superconducting coils were arranged 60° apart azimuthally around the beamline to create a 2 Tesla. The field produced, which varied from two tesla at lower angles to half a tesla at angles greater than

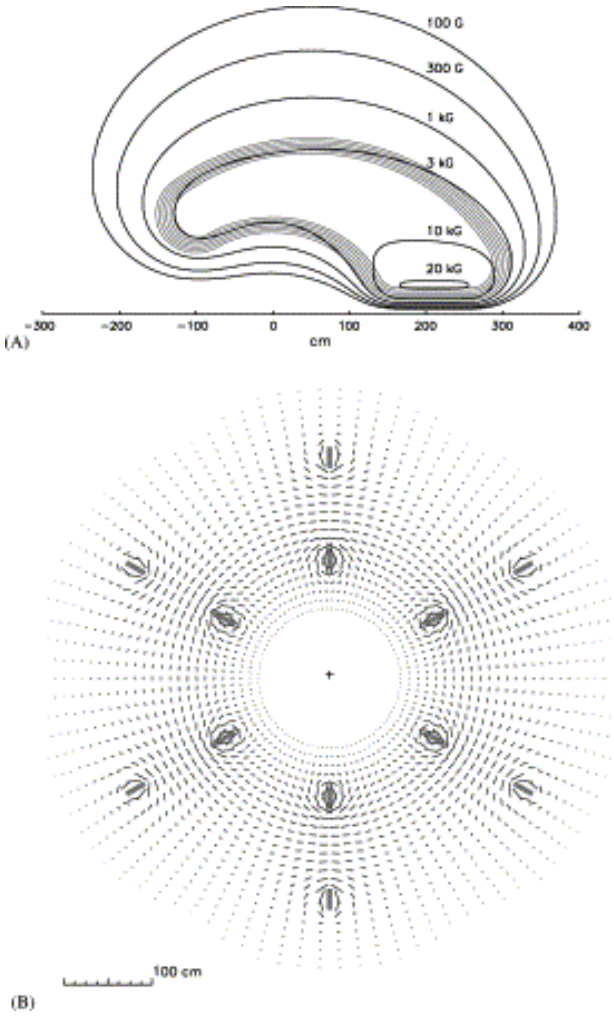

Fig. 2.4: Diagrams of the CLAS torus field. This figure reproduced from [5]

90°, curved charged particle trajectories in the θ direction without altering the azimuthal ϕ direction. The geometry of the torus magnet guided the development of the entire spectrometer.

Spatially, the 18 drift chambers were divided with 3 in each sector. In order to perform tracking before, inside, and after the torus three radially distinct drift

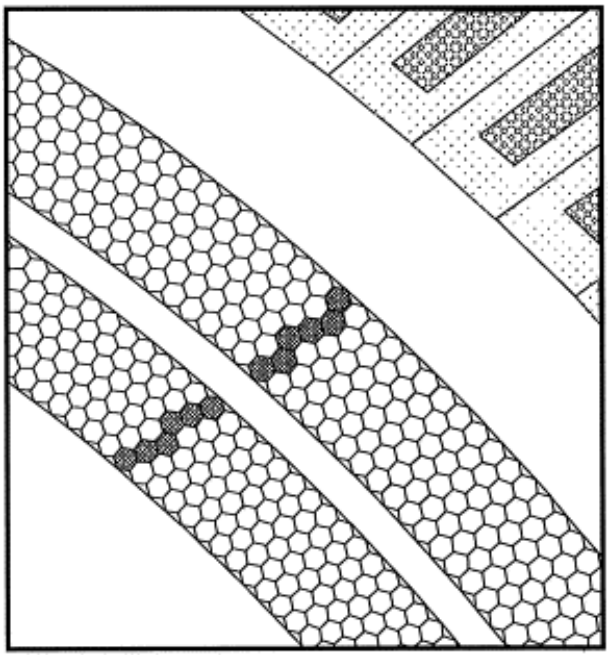

Fig. 2.5: Illustration of a charged particle interacting with cells in the drift chamber. This figure was reproduced from [6]

chambers were constructed for each sector (these were called regions 1, 2, and 3). Each drift chamber consisted of 12 superlayers of hexagonal drift chamber cells. Angular measurements in the azimuthal direction were accomplished by offsetting the first six and last six superlayers by 6°. In total 35,148 individual drift chambers cells were used for tracking [6].

2.2.2 CLAS Cherenkov Counter

The Cherenkov counters (CC), located radially outside of the region 3 drift chambers, greatly assisted in the separation of electrons and negatively charged pions for tracks with momentum less than the pion momentum threshold $p < p_\pi \approx 2.5\ GeV/c$ [5]. The CLAS CCs were divided into 6 sectors, like most other de-

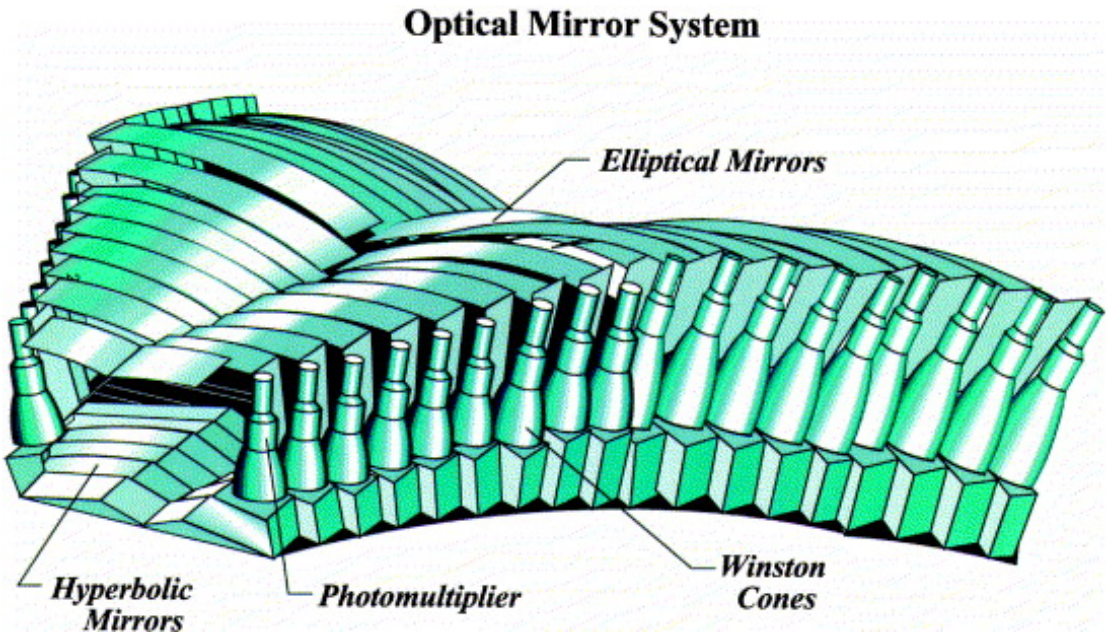

Fig. 2.6: The CLAS Cherenkov Counter.

tectors. Each sector was divided into 18 segments in the polar angle θ away from the beamline. Furthermore, these segments were divided in half azimuthally to produce 12 half-sectors. Three mirrors, a light collecting Winston cone, a magnetic shield, and a 5 inch quartz face PMT were fitted to each of the 18 segments in all 12 half-sectors. During operation each CC was filled with 6 m^3 of C_4F_{10} gas. The number of photo-electrons produced was recorded for tracks with polar angles between $8^\circ < \theta < 45^\circ$.

2.2.3 CLAS Time of Flight Scintillator

Measurements of average velocity can be made by simply knowing the distance some object has traveled in a given time period. Operating on this principle the CLAS time of flight (TOF) system allowed for the separation of π and K for

momentum $p \leq 2\ GeV/c$. Constructed of 57 scintillating bars per sector, the TOF system covered an impressive area of 206 m^2 and spanned the range of polar angles $8° \leq \theta \leq 142°$. Each of the scintillating bars measured 5.08 centimeters in thickness, 15 or 22 centimeters in width, and measured between 32 and 450 centimeters in length. Shorter bars, which covered the relatively higher rate low scattering angle, were built with an intrinsic timing resolution of 80 picoseconds. The timing resolution for longer bars was designed to be 160 picoseconds, and measurements performed after the detector contruction showed that the average timing resolution over the detector was 163 picoseconds [29].

During experiments where the electron beam was used, the start time for each event could be determined by assigning $\beta = 1$ to the fastest electron measured in the final state. However, once the tagger magnet was powered on and photon beam was delivered on target, start time information came from the start counter. Originally designed as six separate counters, the start counter was constructed with three counters spanning the full angular range of the detector. The target was surrounded by these three thin scintillators which had sufficient resolution to determine the difference between two sequentially arriving bunches ($\sigma \approx 350\ ps$).

2.2.4 CLAS Electromagnetic Calorimeter

The outermost later of the CLAS detector was the electromagnetic calorimeter (EC). This sampling calorimeter was a main component of the CLAS trigger

and had several important roles. Foremostly, the EC detected and triggered on electrons with $E > 0.5\ GeV$. Detecting neutral particles such as photons and neutrons was a secondary role. By detecting photons with energies higher than 200 MeV, and having sufficient granular resolution, strongly decaying (π^0 or η) $\rightarrow \gamma\gamma$ particles were measured. Separation of neutrons and photons was achieved by combining information from the EC with timing information from the CLAS time of flight system [7].

Structurally, the EC was composed in total of 1296 PMTs and 8424 scintillating strips in the six EC modules (one per sector). Alternating layers of lead and scintillator material were used to create a sampling fraction E_{dep}/p of approximately 0.3 for electrons. Measuring just 10 mm thick, and with a width of 10 cm (to balance cost of PMTs and granularity), the length of the scintillating strips depended on the angular location. Each EC module contained 3 sets of 13 layers offset by 120° to provide spatial information, these layers were referred to as U, V, and W.

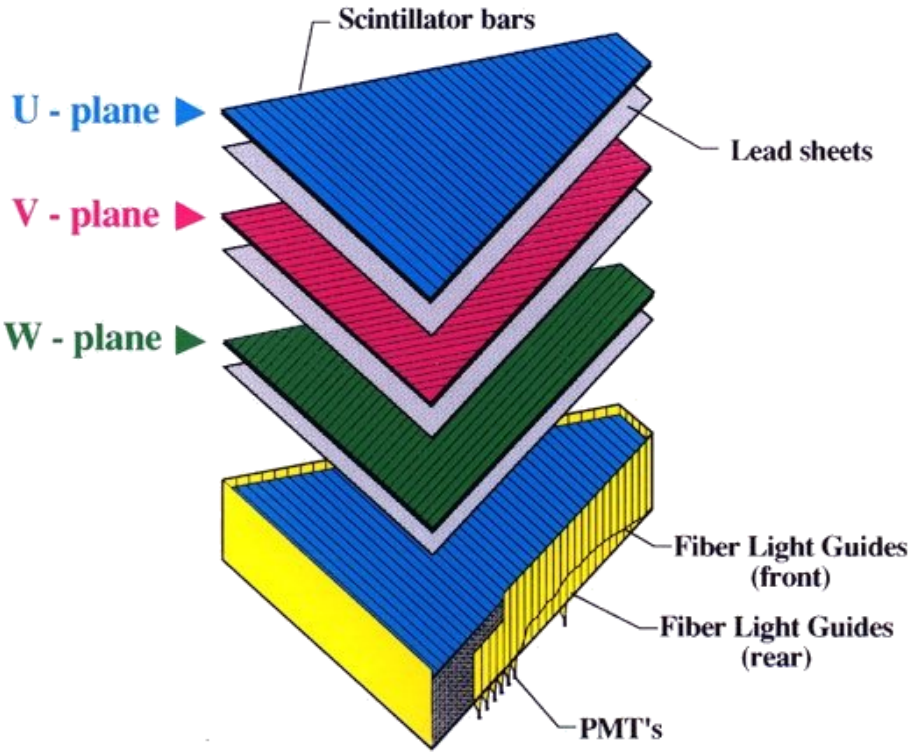

Fig. 2.7: The U, V, W structure of CLAS electromagnetic calorimeter is shown above.

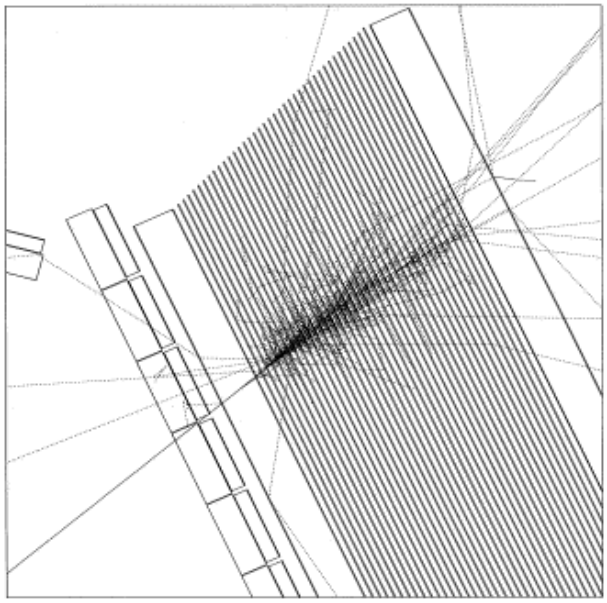

Fig. 2.8: GEANT simulation of an electron showering in the EC. This figure originally appeared in [7]

Chapter 3

Basic Analysis & Corrections

3.1 Introduction

Several analysis components are common to both main analyses presented in this thesis. During this chapter, I discuss the calculation of luminosity, vertex corrections, timing corrections, and kinematic corrections.

3.2 Luminosity Calculation

A useful concept in accelerator/collider physics is the luminosity $\mathcal{L}$. Luminosity is defined as the number of collisions per unit area per unit time that could lead to some process of interest. Consider as an example elastic scattering of electrons from protons, the luminosity is the number of electron-proton collisions per unit time per unit area. The rate $\frac{dN}{dt}$ of the occurance of events for some process X can be written in terms of this luminosity and the cross section for the process.

$$\frac{dN_X}{dt} = \mathcal{L}\sigma_X \tag{3.1}$$

For the fixed target case, the luminosity has a simple expression.

$$\mathcal{L} = \frac{j_e \rho_p l_T}{e} \tag{3.2}$$

Here l_t is the target length, ρ_p is the proton number density in the target, and j_e is the beam current. To find the total number of events which accumulate in some time t_{exp} the event rate is integrated with respect to time.

$$N_X = \int_0^{t_{exp}} \frac{j_e \rho_p l_T}{e} \sigma_X dt \tag{3.3}$$

$$= \frac{\rho_p l_T}{e} \sigma_X \int_0^{t_{exp}} j_e dt \tag{3.4}$$

$$= \frac{\rho_p l_T}{e} \sigma_X \Delta Q \tag{3.5}$$

Thus the experimentally observed cross section for some process X is,

$$\sigma_X = \frac{N_X}{\mathcal{L}_{int}} \tag{3.6}$$

where the number of events N_X is corrected for all effects and $\mathcal{L}_{int}$ is the integrated luminosity as shown above.

Experimentally, the factor ΔQ can be calculated from charge deposition measurements by the Faraday cup. The Faraday cup charge is a scalar value writ-

ten periodically into the output event stream, not with every recorded event. This information is stored in the output BOS files in a bank called `TRGS`, the variable is named `FCUP_G2`.

For this data analysis, the final ntuple (root files) used did not contain the Faraday cup charge information. For this reason, the authors used the BOS files directly and recorded the value of `FCUP_G2` for every scalar reading, as well as the event number directly after each scalar entry (from the `HEAD` bank). This event number correlates directly to the event number stored in the root files used for analysis.

The total accumulated charge over a run is simply the sum over consecutive differences in the Faraday cup charge.

$$\Delta Q = \sum_{i=1}^{n-1} q[i+1] - q[i] \tag{3.7}$$

Here n denotes the number of scalar entries for a given file. Due to the periodic nature of the scalar bank writing events are also recorded after the last reading of the file, and before the first scalar reading of the next file in the run. To account for this the difference between consecutive files last and first readings is added to the total.

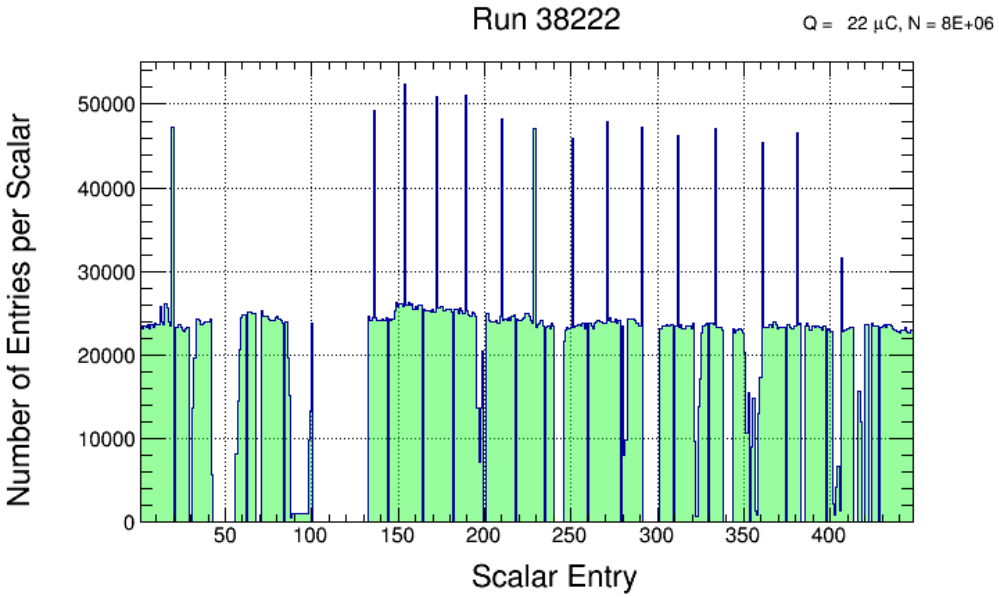

Fig. 3.1: Entries per scalar reading for run 38222.

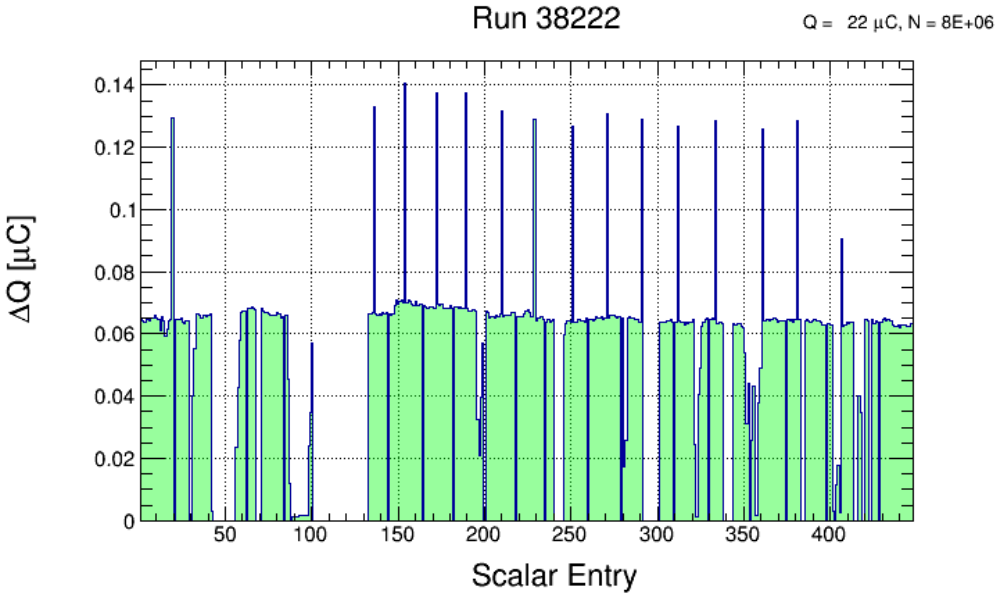

Fig. 3.2: Charge per scalar reading for run 38222.

For the E1-F dataset, a run typically contains around 20 files, each representing a raw file size of 2 gigabytes. These files are named by run number, and given an index from 0 to $n_{files} - 1$. It is not uncommon that a run will contain missing files in the middle of the range. If this occurs, the charge difference between last/first reading is not added to the total.

Any charge which accumulates in a period of time where the number of events did not change is not added to the total. Similarly, any events which occur within regions where no charge is recorded need to be discarded. This is accomplished by recording the bad event ranges for every file and removing these events from our analysis.

The result of this procedure is a numerical value of charge for each run. For practitioners, it is important to note that this value needs to be scaled by the DAQ scaling factor before it represents a value of charge. In our analysis, the numerical value of charge for a typical file is a few tens of micro-Coulombs.

3.3 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 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 3.3.

3.4 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. These changes are taken into account in the data analysis.

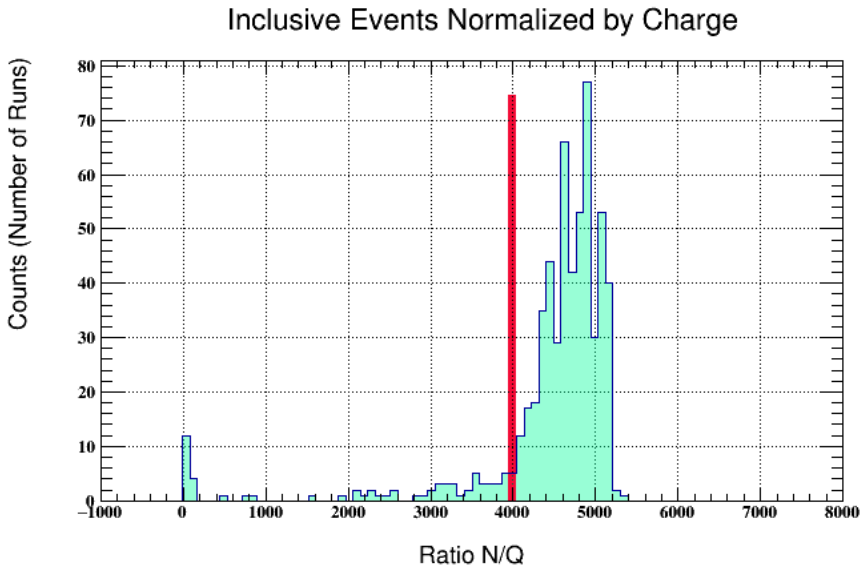

Fig. 3.3: Inclusive electrons per file normalized by the total charge accumulated for the file. This quantity is used to make a good run list.

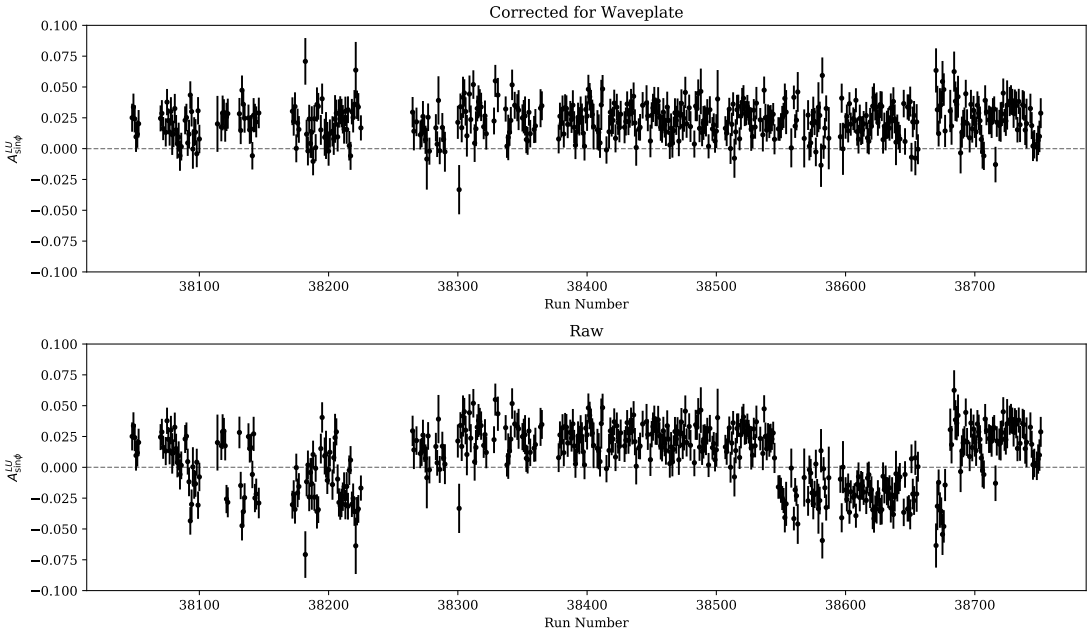

Fig. 3.4: The waveplate position is determined and corrected by plotting the BSA for π^+ mesons as a function of the run. The top panel shows the corrected results, the bottom shows the results before changing the helicity.

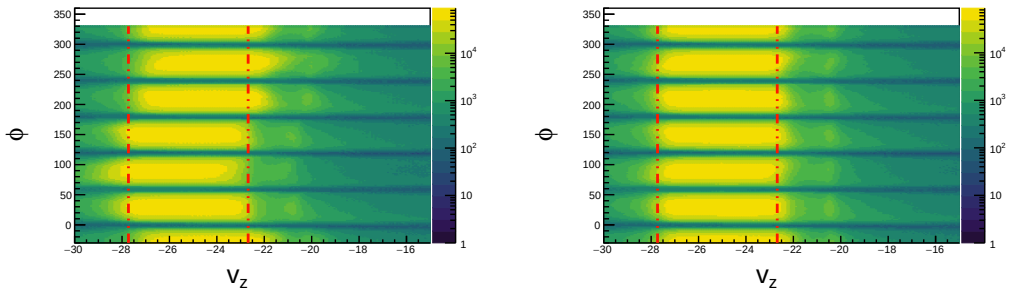

Fig. 3.5: The z-vertex v_z position shown for different values of ϕ the azimuthal angle in the hall. The left figure shows the distribution before corrections are applied, the right after. The vertical red lines bound the region which we define as acceptable for electrons in our analysis.

3.5 Vertex Corrections

The track vertex position (v_x, v_y, v_z) is calculated based on the intersection of each track with the midplane (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 midplanes 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 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 3.5. 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.

3.6 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.

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

$$\Delta\beta = \beta_{obs} - \beta_{pred} = \frac{d}{ct_{obs}} - \sqrt{1 + (m/p)^2} \tag{3.8}$$

Here m is assumed to be m_π. 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 pad-

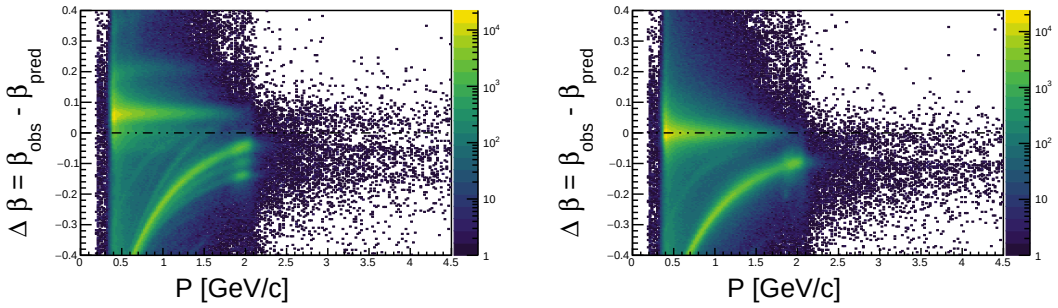

Fig. 3.6: Timing corrections are shown for paddle 24 of sector 1. The left image shows the $\Delta\beta$ distribution before corrections. On the right the same is shown after correction of the timing for this paddle. We assume the mass of the track to be the pion, these show up as the green band. Heavier protons are visible below the pion band.

dles. 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 [24].

3.7 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 proceedures 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 [30].

As mentioned previously, the need for correction to θ (the polar angle measured from the beamline) arises from misalignments in the drift chambers. This implies that the correction will be the same for positives and negatives, and this assumption is used in the correction algorithm. First, elastic ($ep \to ep$) 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 ($ep \to epX$) system is close to 0. The kinematics of the event are then calculated.

$$k^\mu = (k, 0, 0, k) \tag{3.9}$$

$$p^\mu = (M_p, 0, 0, 0) \tag{3.10}$$

$$k'^\mu = (k', k' \sin\theta, 0, k' \cos\theta) \tag{3.11}$$

$$p'^\mu = (E_p, -p' \sin\alpha, 0, p' \cos\alpha) \tag{3.12}$$

Applying energy and momentum conservation to the equations above yields

3 equations.

$$k + M_p = k' + \sqrt{M_p^2 + p'^2} \tag{3.13}$$

$$k' \sin\theta = p' \sin\alpha \tag{3.14}$$

$$k = k' \cos\theta + p' \cos\alpha \tag{3.15}$$

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.

$$\cos\theta = 1 - M_p \frac{k - k'}{kk'} \tag{3.16}$$

$$\tan\alpha = \frac{1}{p'} \frac{k' \sin\theta}{k - k' \cos\theta} \tag{3.17}$$

After θ corrections are applied, the momentum 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 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

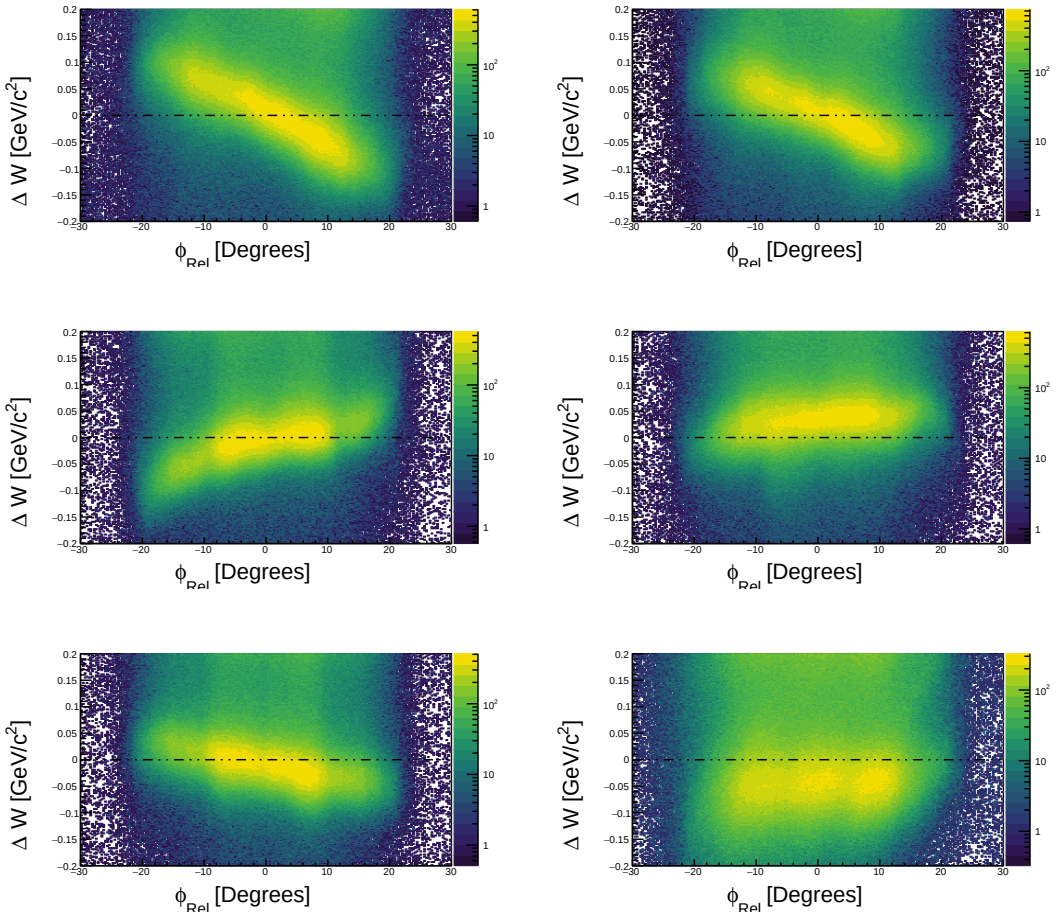

Fig. 3.7: This figure shows the deviation from M_p of the W spectrum peak for elastic $ep \rightarrow ep$ events (before corrections).

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.

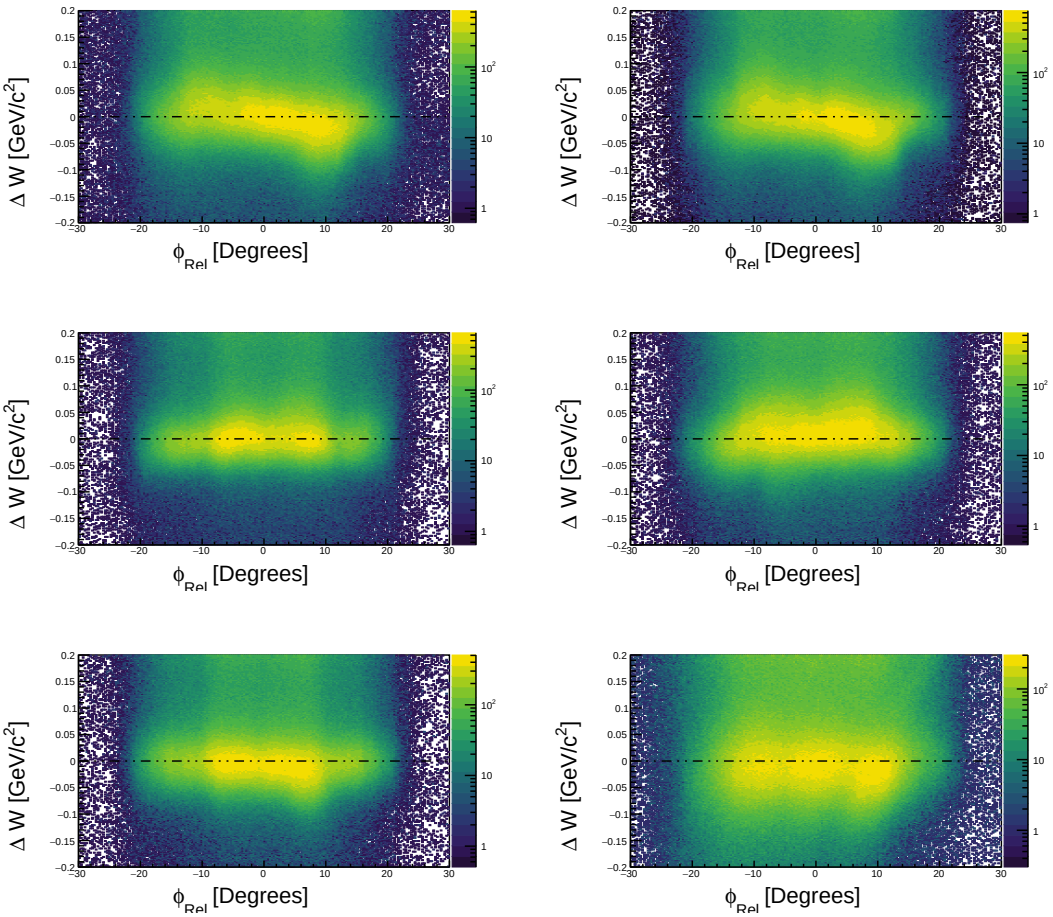

Fig. 3.8: This figure shows the deviation from M_p of the W spectrum peak for elastic $ep \rightarrow ep$ events (after ϕ-dependent corrections).

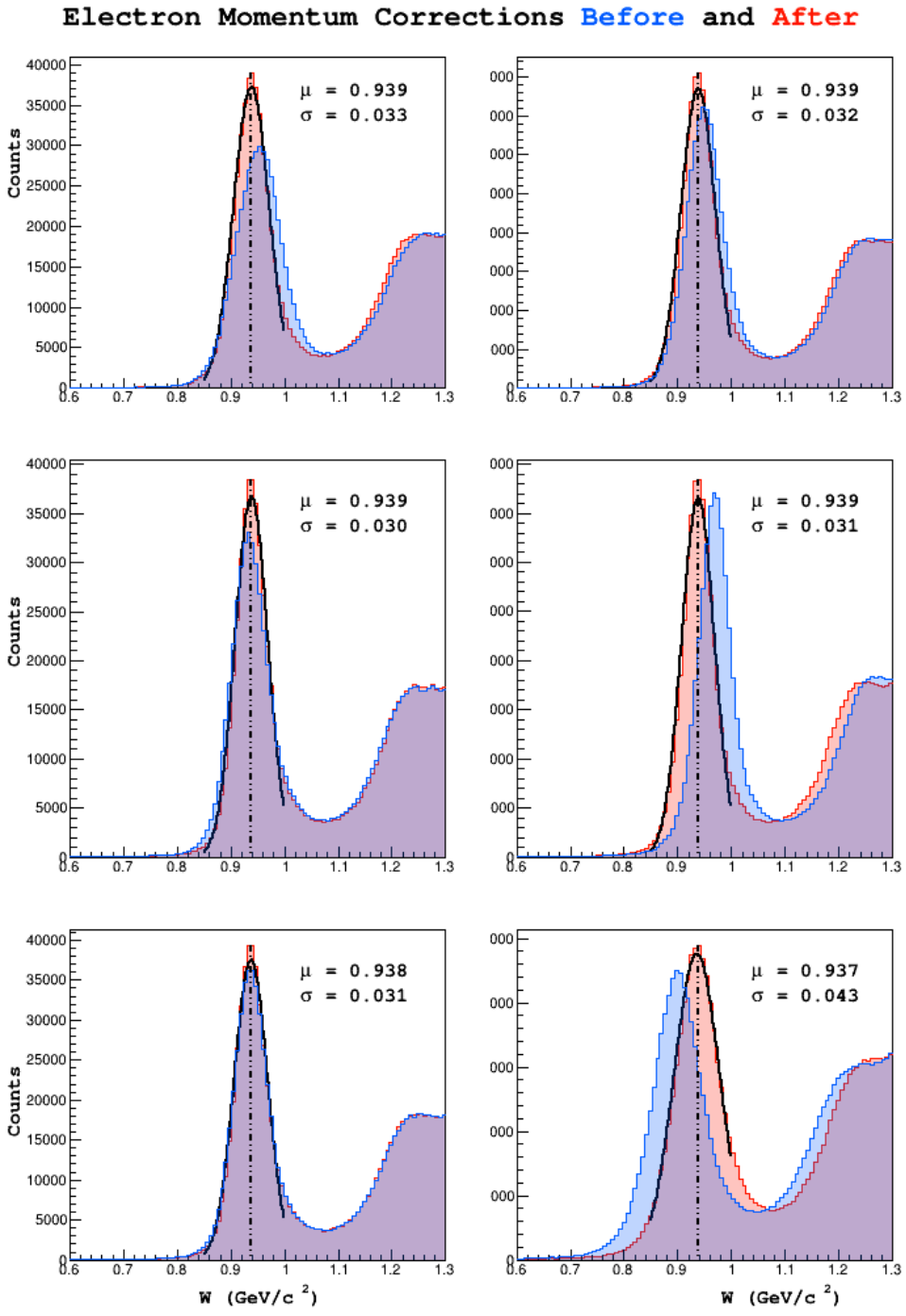

Fig. 3.9: Elastic events shown in the spectrum of W before and after momentum corrections are applied.

Chapter 4

Particle Identification

4.1 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 `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 to classify particles.

4.2 Electron Identification

Electrons in CLAS are abundant, and the detection of an electron is a basic necessity for every event that is 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 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, 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).

4.2.1 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 is 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. 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 it's 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

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

Table 4.1: Cut parameters used for the DC fiducial cut.

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 $\pm \cot(\theta)$.

$$y_{right} = h + \cot(\theta)x \tag{4.1}$$

$$y_{left} = h - \cot(\theta)x \tag{4.2}$$

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

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

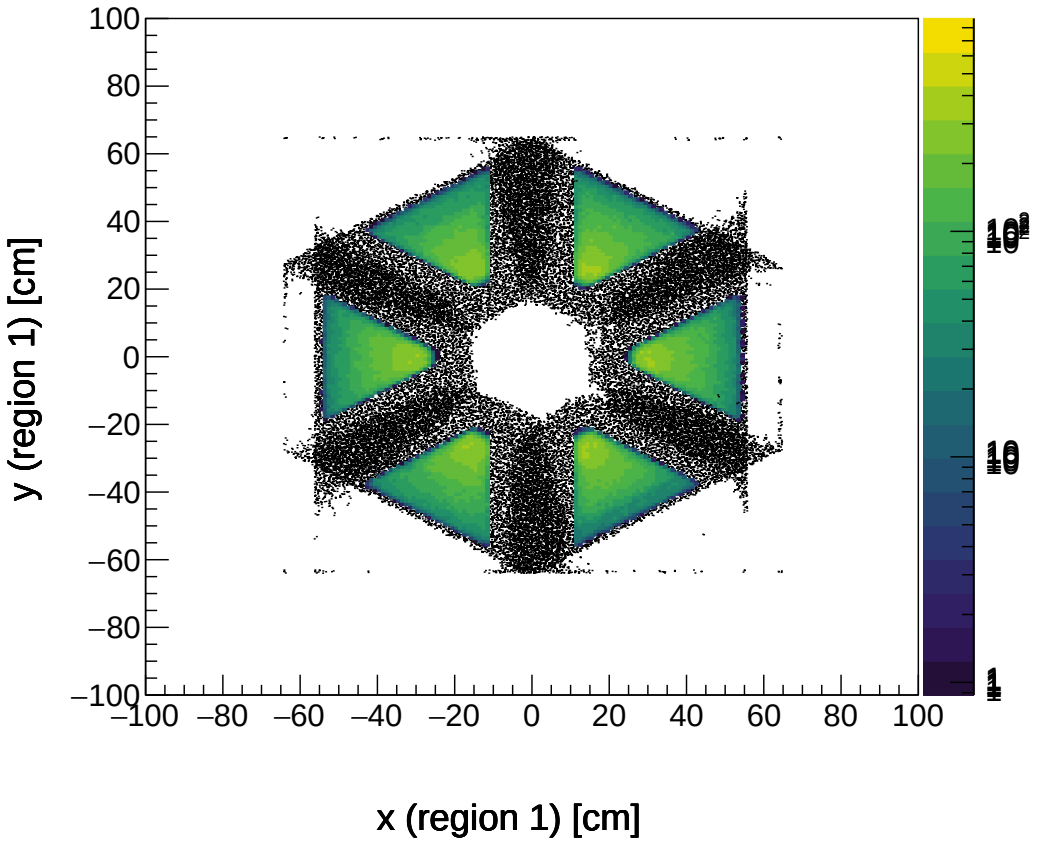

Fig. 4.1: Tracks shown in color remain after the application of drift chamber region 1 fiducial cuts to all cuts, shown here as black points.

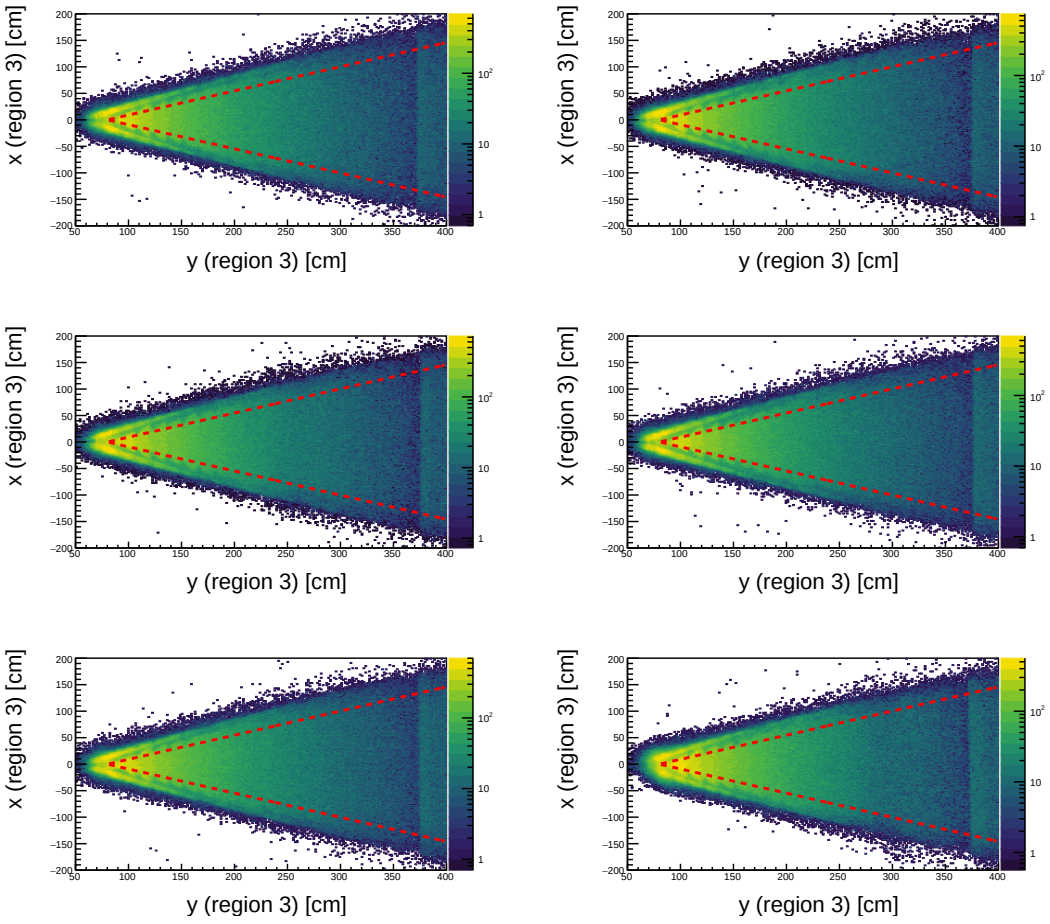

Fig. 4.2: The selection criteria shown in red is applied to drift chamber region 3.

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

Table 4.2: Cut parameters used for the EC fiducial cut.

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 I 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

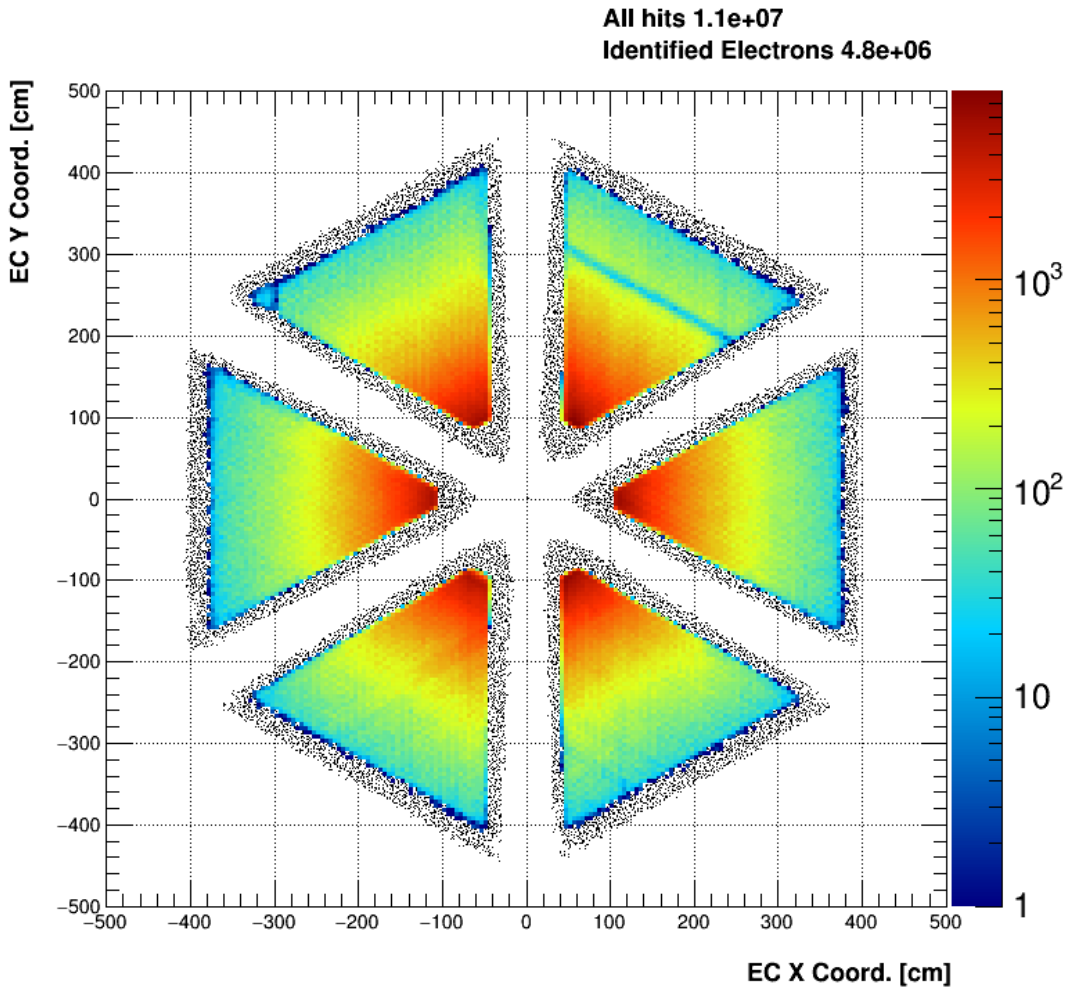

Fig. 4.3: All negative tracks are shown here in black. In color, the tracks which pass the EC fiducial cut are shown.

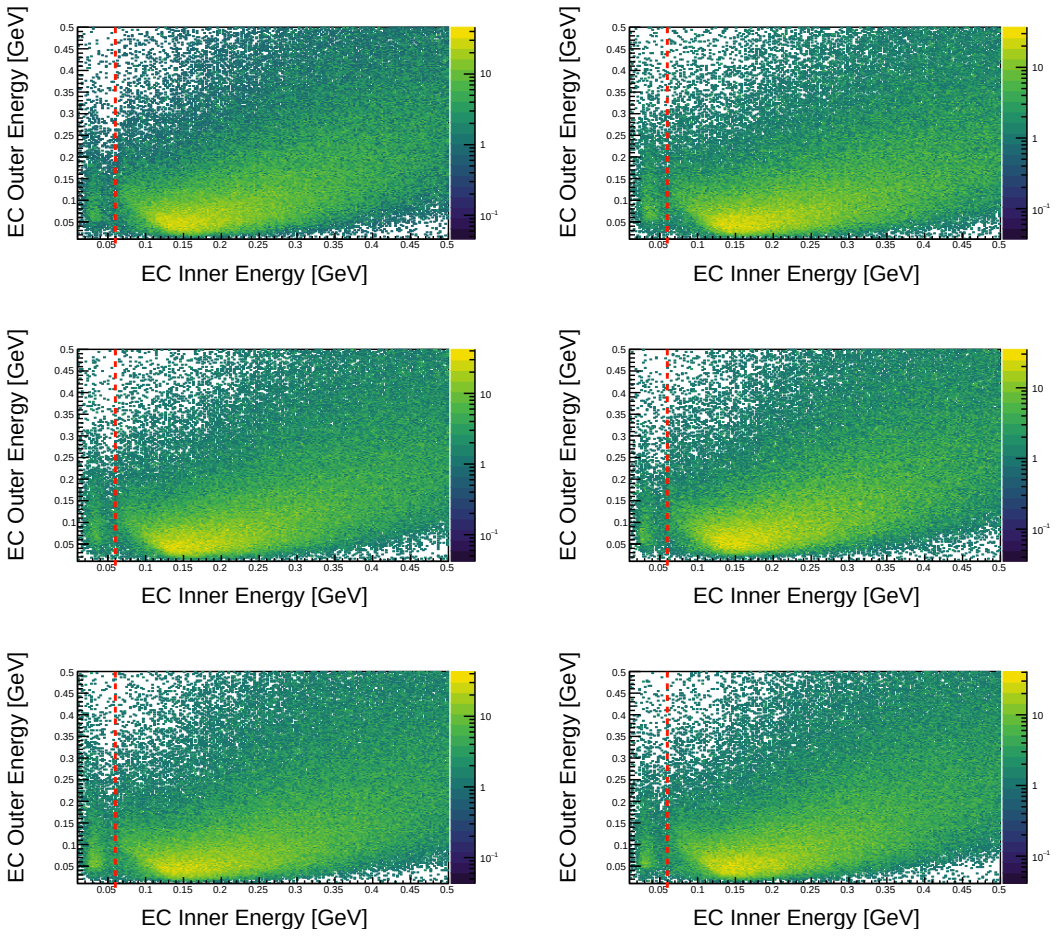

Fig. 4.4: Each panel shown above contains events from one sector, increasing from 1-6 from top left to bottom right. The value selected of 60 MeV is applied to all sectors and separates the negatively charged pions (left) from the electrons (right).

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.

$$\mu(p) = \mu_0 + \mu_1 p + \mu_2 p^2 + \mu_3 p^3 \tag{4.3}$$

$$\sigma(p) = \sigma_0 + \sigma_1 p + \sigma_2 p^2 + \sigma_3 p^3 \tag{4.4}$$

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

$$f_{max}(p) = \mu(p) + n_\sigma \sigma(p) = (\mu_0 + n_\sigma \sigma_0) + (\mu_1 + n_\sigma \sigma_1)p + (\mu_2 + n_\sigma \sigma_2)p^2 + (\mu_3 + n_\sigma \sigma_3)p^3 \tag{4.5}$$

$$f_{min}(p) = \mu(p) - n_\sigma \sigma(p) = (\mu_0 - n_\sigma \sigma_0) + (\mu_1 - n_\sigma \sigma_1)p + (\mu_2 - n_\sigma \sigma_2)p^2 + (\mu_3 - n_\sigma \sigma_3)p^3 \tag{4.6}$$

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 4.3.

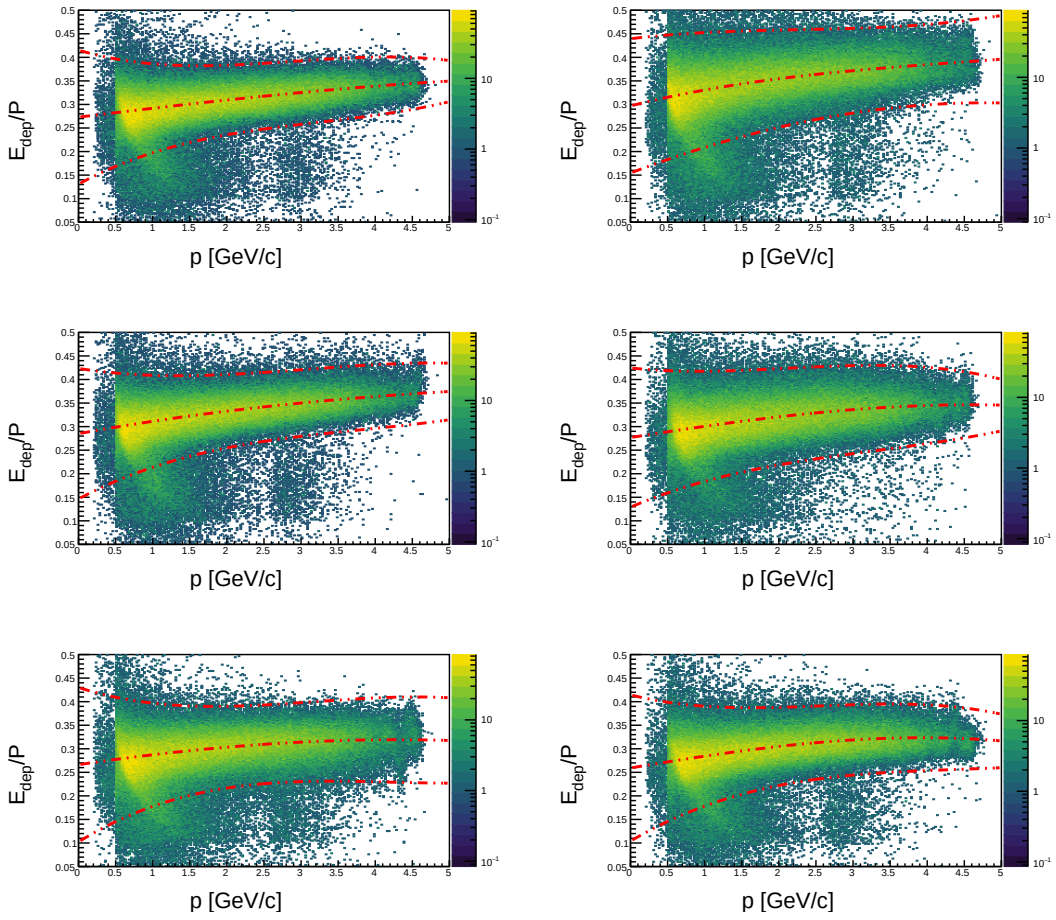

Fig. 4.5: The sampling fraction selection boundary is shown here for the nominal value of $N_{sigma} = 4$.

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 4.3: 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).

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 θcc 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

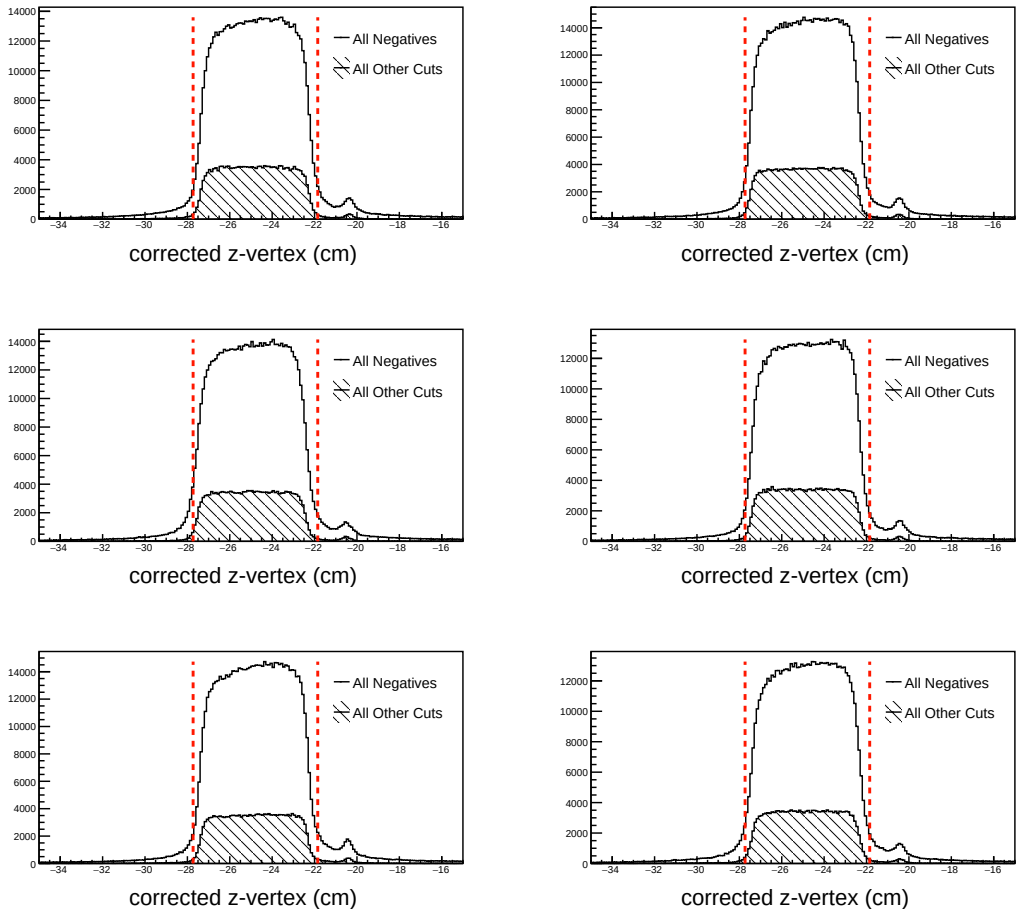

Fig. 4.6: The track vertex cut is shown above. All negative tracks are shown in white, while the tracks passing all other criteria are shown in black hatch. The cut boundary is displayed as red lines. For E1-F the target center was located at -25 cm.

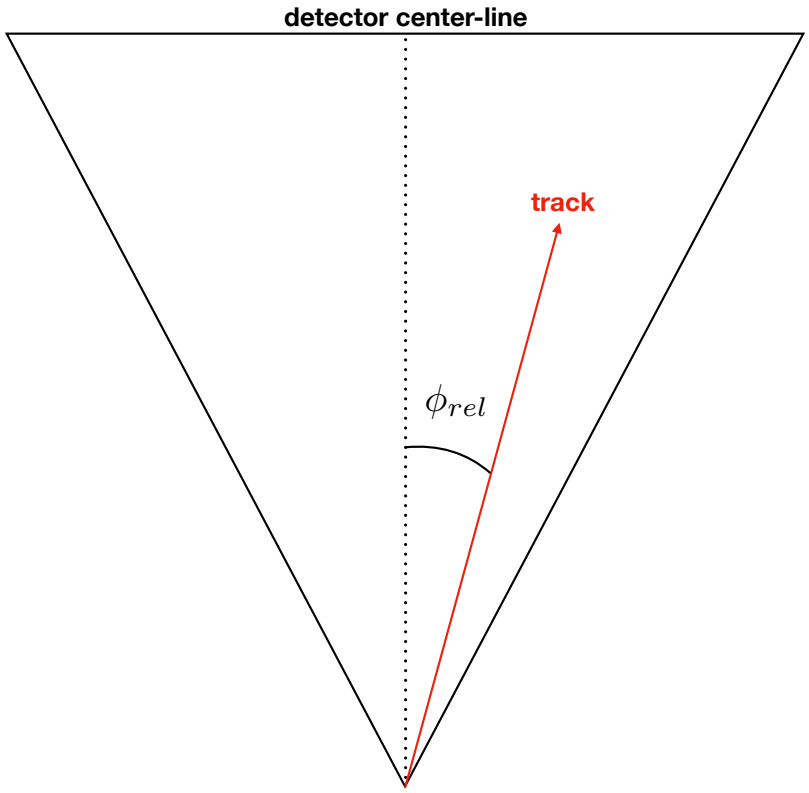

Fig. 4.7: The angle ϕ_{rel} is the azimuthal angle between the central line of the detector and the track.

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.

4.3 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 probabalistic

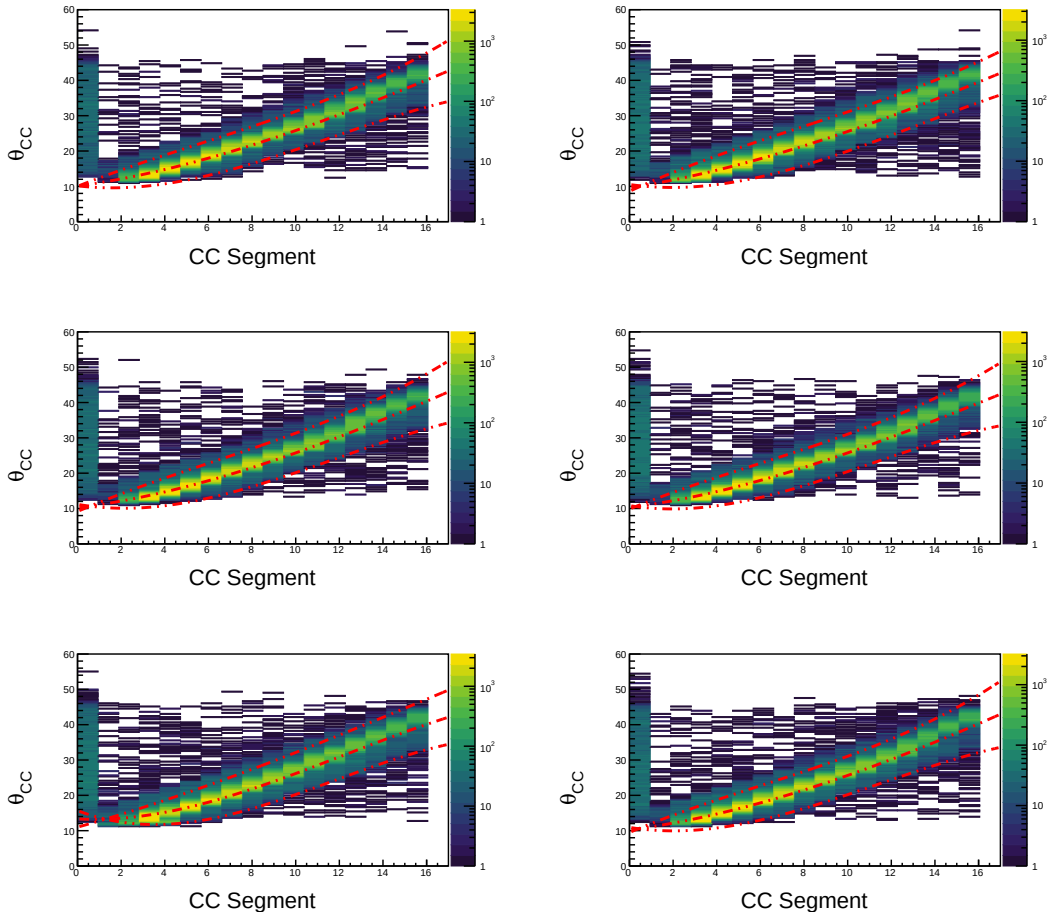

Fig. 4.8: Correlation between θ_{CC} and the CC segment is shown above, with our selection boundaries overlaid in red.

methodology described in this section is based on the discussion provided by the BES collaboration in [31].

4.3.1 Hadron ID Cuts

The cuts used for hadron classification are enumerated below.

- Drift chamber fiducial
- Hadron-electron vertex difference
- Probability 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 $h = 10$, $\theta = 60$.

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 4.3.1. This cut would not be applicable to studies where a significantly detached vertex is expected.

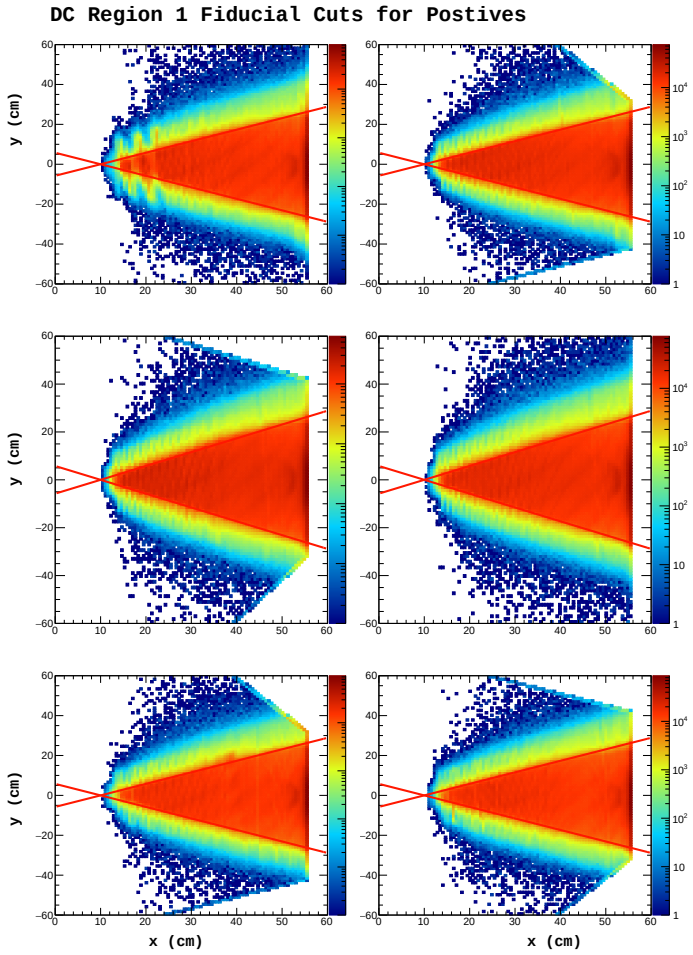

Fig. 4.9: Shown above: Positive track hits on the region 1 drift chamber, events falling between the red lines are kept for analysis.

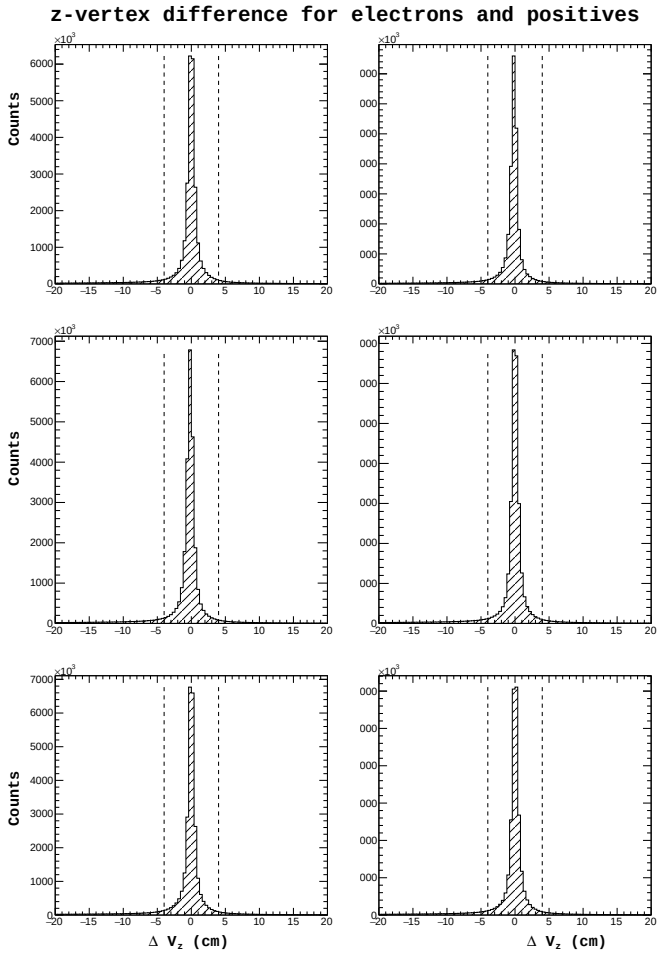

Fig. 4.10: Shown above: The difference between the z-vertex position between detected electrons and positive tracks.

Probability of $\beta(p,h)$

The observed deviation between β and the expected value for each different species of hadron h with a given momentum p is Gaussian distributed about zero.

$$P(\beta;p,h) = \frac{1}{\sqrt{2\pi}\sigma_\beta(p,h)} exp\left\{-\frac{1}{2}\left(\frac{\beta-\mu_\beta(p,h)}{\sigma_\beta(p,h)}\right)^2\right\} \tag{4.7}$$

If the expected value $\mu_\beta(p,h)$ and resolution $\sigma_\beta(p,h)$ of beta is known for each hadron type the identity of each candidate can be assigned by choosing the hadron h which has the largest probability.

$$h* = argmax P(\beta;p,h) \tag{4.8}$$

Using this method, every positive (negative) track is assigned a particle identification. However, at times the probability value is quite small. 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.

$$\alpha = 1 - \int_{\mu-\beta_{obs}}^{\mu+\beta_{obs}} P(\beta;p,h)d\beta \tag{4.9}$$

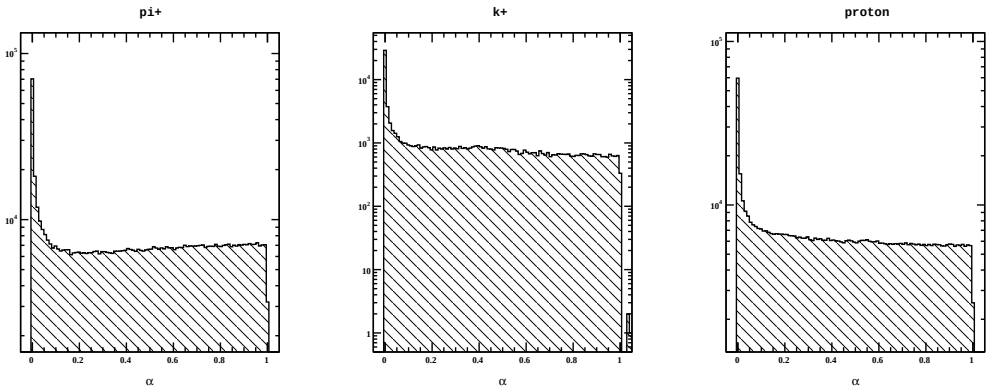

Fig. 4.11: Shown above: The distribution of confidence level for all positive tracks after being classified by the highest probability.

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.

Although here only one variable is used to define the hadron probability this method into the likelihood by including information from other detectors, not present in CLAS.

Determination of probability density functions for probabilistic method

The most important and most difficult part of constructing the probabilistic identification is the determination of the mean and standard deviation of the probabil-

ity 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 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). Negative pions and kaons are fit directly (as is normally done).

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

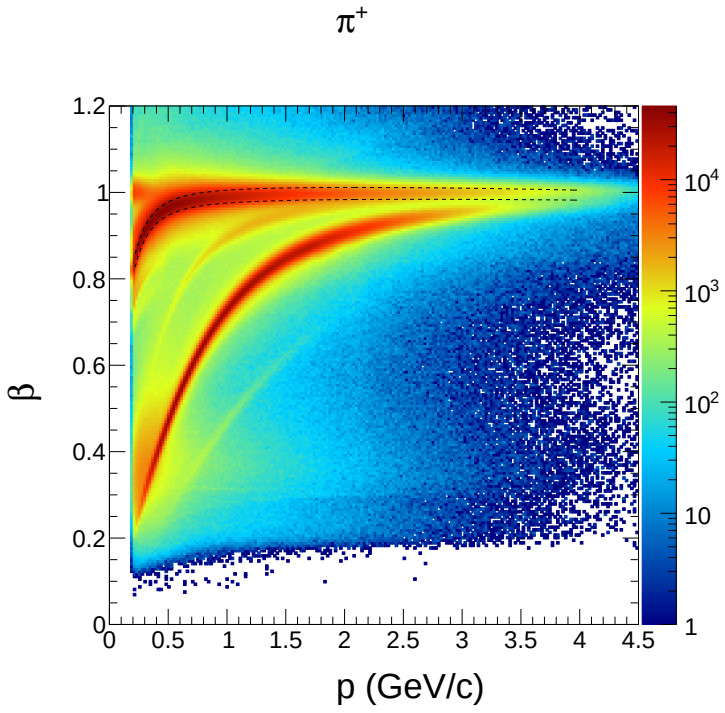

Fig. 4.12: Shown above: All positive tracks overlaid with our determination of $\mu(p) \pm \sigma(p)$ for π^+

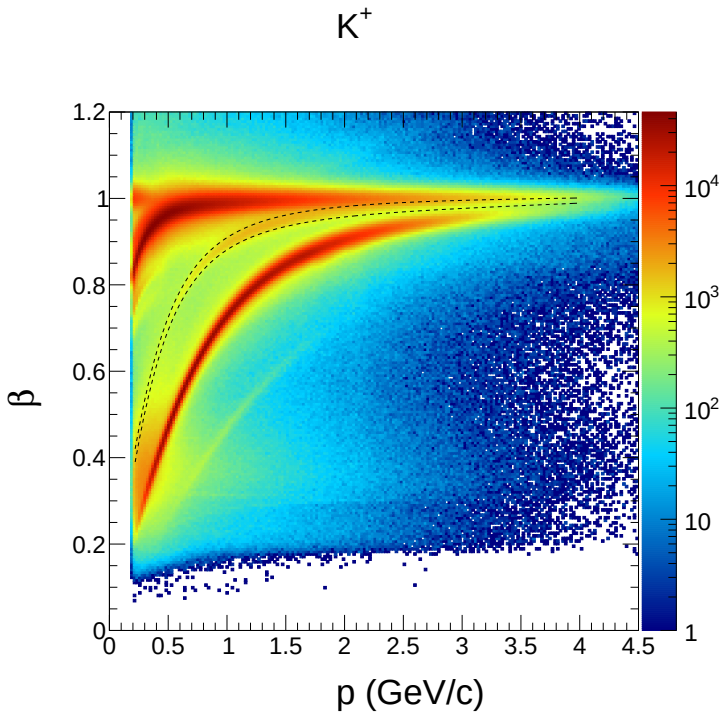

Fig. 4.13: Shown above: All positive tracks overlaid with our determination of $\mu(p) \pm \sigma(p)$ for K^+

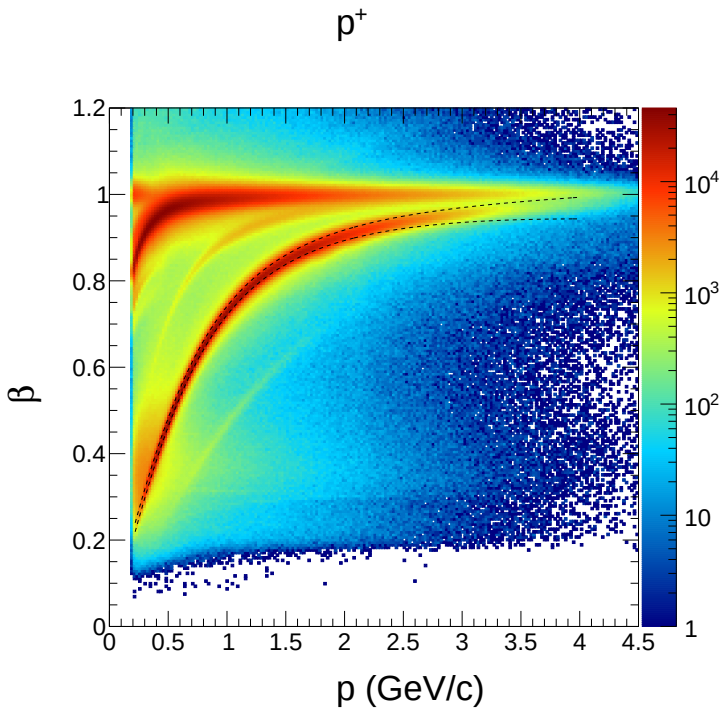

Fig. 4.14: Shown above: All positive tracks overlaid with our determination of $\mu(p) \pm \sigma(p)$ for p^+

$$\mu(p,h) = \mu_{theory} + \Delta\mu \tag{4.10}$$

$$\mu_{theory} = \frac{1}{\sqrt{1+(m_h/p)^2}} \tag{4.11}$$

$$\Delta\mu = \mu_0 + \mu_1 p + \mu_2 p^2 \tag{4.12}$$

$$\sigma(p,h) = \sigma_0 + \sigma_1 p + \sigma_2 p^2 \tag{4.13}$$

The values are displayed in the table below.

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 4.4: Values used to calculate the mean and resolutions for hadron probability 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 ($p_{max} = 2.0\ GeV/c$) 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 $E_{beam} = 5.498\ GeV$. 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 I correlated reconstructed particles to their generated counter-part by requiring that magnitude of the three momentum difference was small ($\delta P < 0.05$).

$$\delta P = \frac{|\vec{P_{gen}} - \vec{P_{rec}}|}{|\vec{P_{gen}}|} \tag{4.14}$$

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, I 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,

$$P = \frac{tp}{tp + fp} = \frac{N_{K^+}}{N_{K^+} + N_{\pi^+} + N_{P^+}} \tag{4.15}$$

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.

$$\epsilon = \frac{tp}{tp + fn} = \frac{N_{K^+}^{identified}}{N_{K^+}^{total}} \tag{4.16}$$

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 4.3.1.

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 0.02, if the sample is comprised of 20% pions (which is the worse case in our measurement) then the

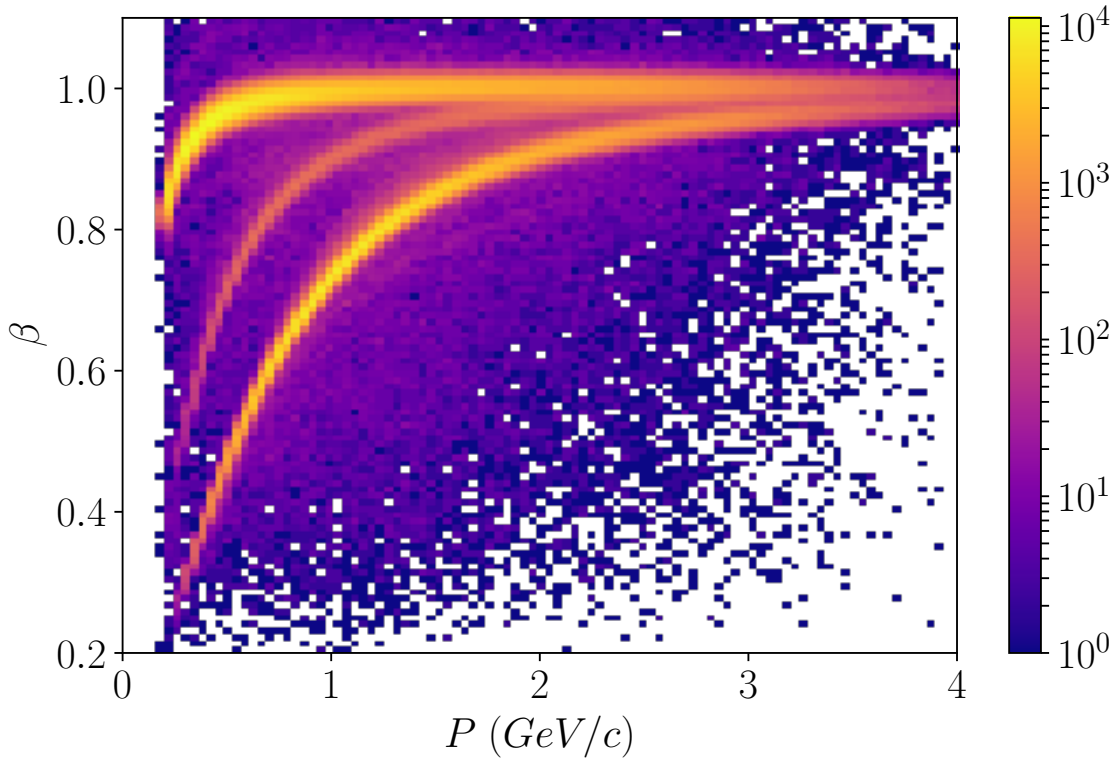

Fig. 4.15: Positive hadrons from the Monte Carlo simulation produce a β vs. p simulation that is very similar to data.

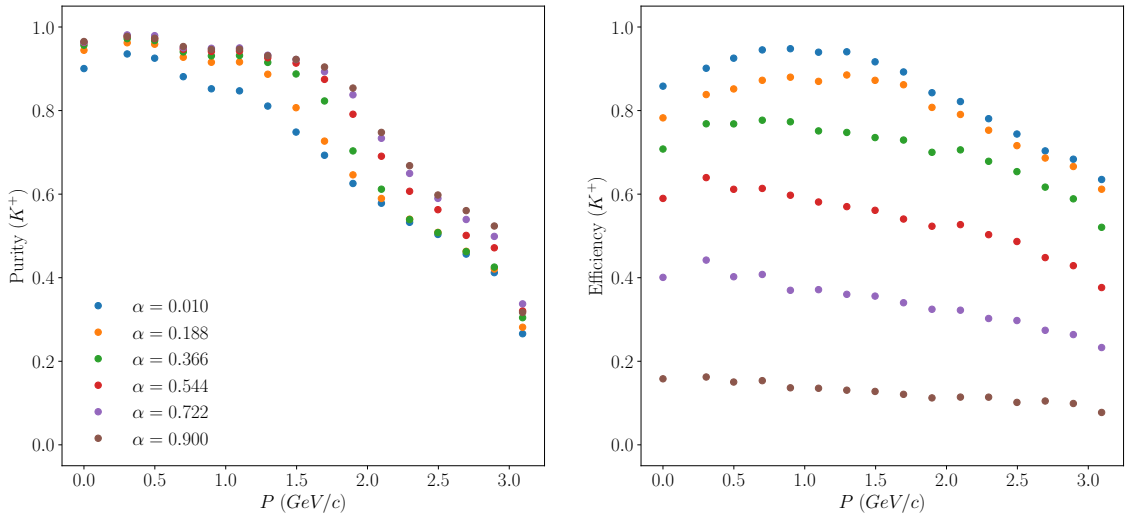

Fig. 4.16: The efficiency and purity of our kaon sample are studied by using a Monte Carlo simulation. Here, the results are studied as a function of the confidence level, and of the track momentum.

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 5

Inclusive Cross Section

One of the primary goals of this thesis work is the accurate determination of cross sections for charged π mesons in SIDIS. The measurement of a cross section for a well known process such as elastic scattering or inclusive electron scattering is typically calculated (using the same data) before new cross sections are provided. Not only does this procedure provide validation of the integrated luminosity factor, it also builds confidence that the overall quality of the dataset is high, the electron identification is accurate, and the list of files used in the analysis is reasonable. In this chapter, the inclusive electron scattering cross section is discussed.

5.1 Motivation

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 verification is achieved here by studying the resonance region. This is due principally to the excess of Bethe-Heitler events which collect in the $2 < W < 3$ region for $E_{beam} = 5.498$ (such events are difficult to remove when detecting only the final state electron).

Measurement of the inclusive cross section is performed for each sector independently by counting reconstructed events N_i in each bin i of the phase space of W and Q^2. Experimentally the cross section for bin i is,

$$\frac{d\sigma_i}{dW\ dQ^2} = \frac{N_i}{\mathcal{L} A_i B_i R_i} \frac{1}{\Delta W\ \Delta Q^2} \tag{5.1}$$

where the factor $\mathcal{L}$ is the integrated luminosity for the time period over which the events N were collected. The correction factors A_i, B_i, and R_i refer to the acceptance, bin center, and radiative corrections for bin i respectively.

5.2 Inelastic Scattering Model

A particularly important aspect of this validation study is the reference model. Not only is the model used to predict the cross section for comparison, it is also sampled in the Monte Carlo event generators used in study detector acceptances

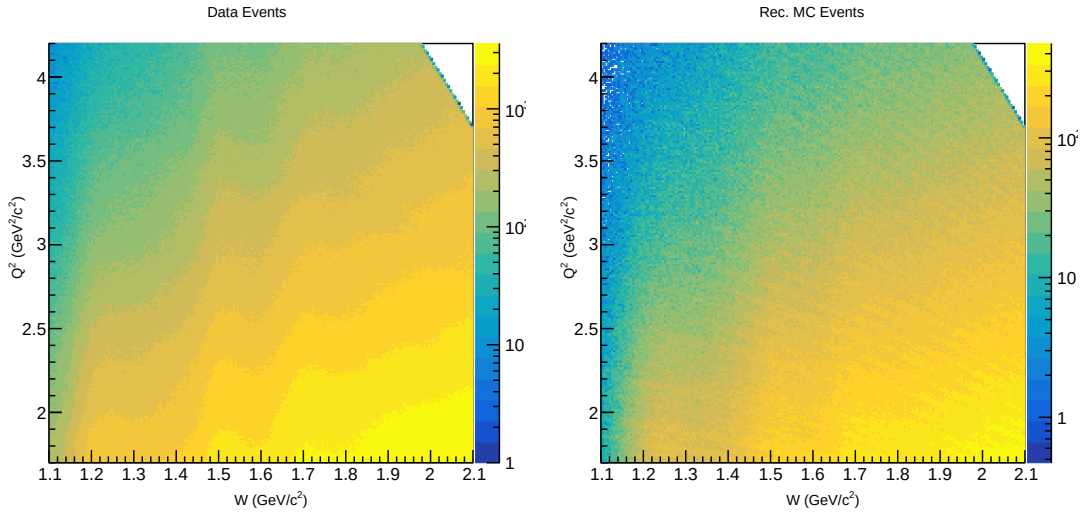

Fig. 5.1: Data and Simulation (reconstructed) events are compared.

and radiative corrections. The model which is used in this study was first developed by Cynthia Keppel [32]. Data taken by SLAC experiments `NE11` and `E133` was used to fit a 24 parameter model, 3 background terms and 3 resonance terms in W as well as polynomial in Q^2. While the model is constrained at higher W by data from the DIS region, it is designed to operate in the resonance region (starting at the pion production threshold).

5.3 Analysis

5.3.1 Event Selection

The problems solved by event selection are two-fold. Typically it is required that the negative momentum transfer squared $Q^2 > 1\ GeV^2/c^2$ is sufficiently large so that the measurement can be considered part of the DIS region. Analyses in

CLAS have typically used 1 GeV^2/c^2 as the working assumption for what constitutes the lowest acceptable momentum transfer to be considered DIS, and this value is used in the SIDIS presented in this work. Because we study the resonance region with $W \in [1.1, 2.1]$ and the beam energy is $E_{beam} = 5.498\ GeV$ the kinematics are sparsely populated for $Q^2 < 1.6\ GeV^2/c^2$ and this limit is applied in our study. The second important function of event selection in this study is to avoid the largest source of background, Bethe-Heitler events.

As mentioned in the introduction, Bethe-Heitler events (where a real photon is radiated by the electron in the initial or final state) are the dominant source of background for this measurement. If when studying elastic scattering the final state electron and the scattered target proton are detected, it's possible to apply energy-momentum constraints to the system and eliminate radiative Bethe-Heitler events. However, in this measurement only the final state electron is detected and used to workout the event kinematics. There is no way to directly apply energy-momentum conservation to eliminate Bethe-Heitler events. Faced with this problem, we choose to restrict the values of *inelasticity* $y = 1 - E'/E < 0.7$ that are included in our event sample to limit this contribution. This restriction is applied because events with large-y have a significantly higher probability to be Bethe-Heitler events (this is verified by simulation and shown below in figure 5.3.1). This restriction is equivalent to enforcing a minimum energy for the scattered electron.

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

Table 5.1: Summary of W and Q^2 binning used for the inclusive cross section.

$$E_{min} = E_{beam}(1 - y_{max}) \approx 1.6\ GeV \tag{5.2}$$

Figure 5.3.1 shows the kinematic distribution of events from simulation and data, and illustrates the need for this y cut. The number of Bethe-Heitler events in our sample is further reduced by limiting the measurement to the resonance region. Simulations for $E_{beam} = 5.498\ GeV$ show that the majority of Bethe-Heitler events occur between $W \in [2, 3]$.

5.3.2 Binning

Detected inclusive events are divided into 10 bins of equal size in Q^2 and 40 bins of equal size in W (for a total of 400 bins for each sector). Bins are chosen small enough in W to detect resonance features, but larger than the resolution of CLAS to avoid having a large numbers of events reconstructed in a bins that they do not belong in. Equal sized bins are chosen for their simplicity and because for cross sections (unlike asymmetry measurements) we can tolerate bins with comparatively low statistics. The binning details are displayed in table 5.1.

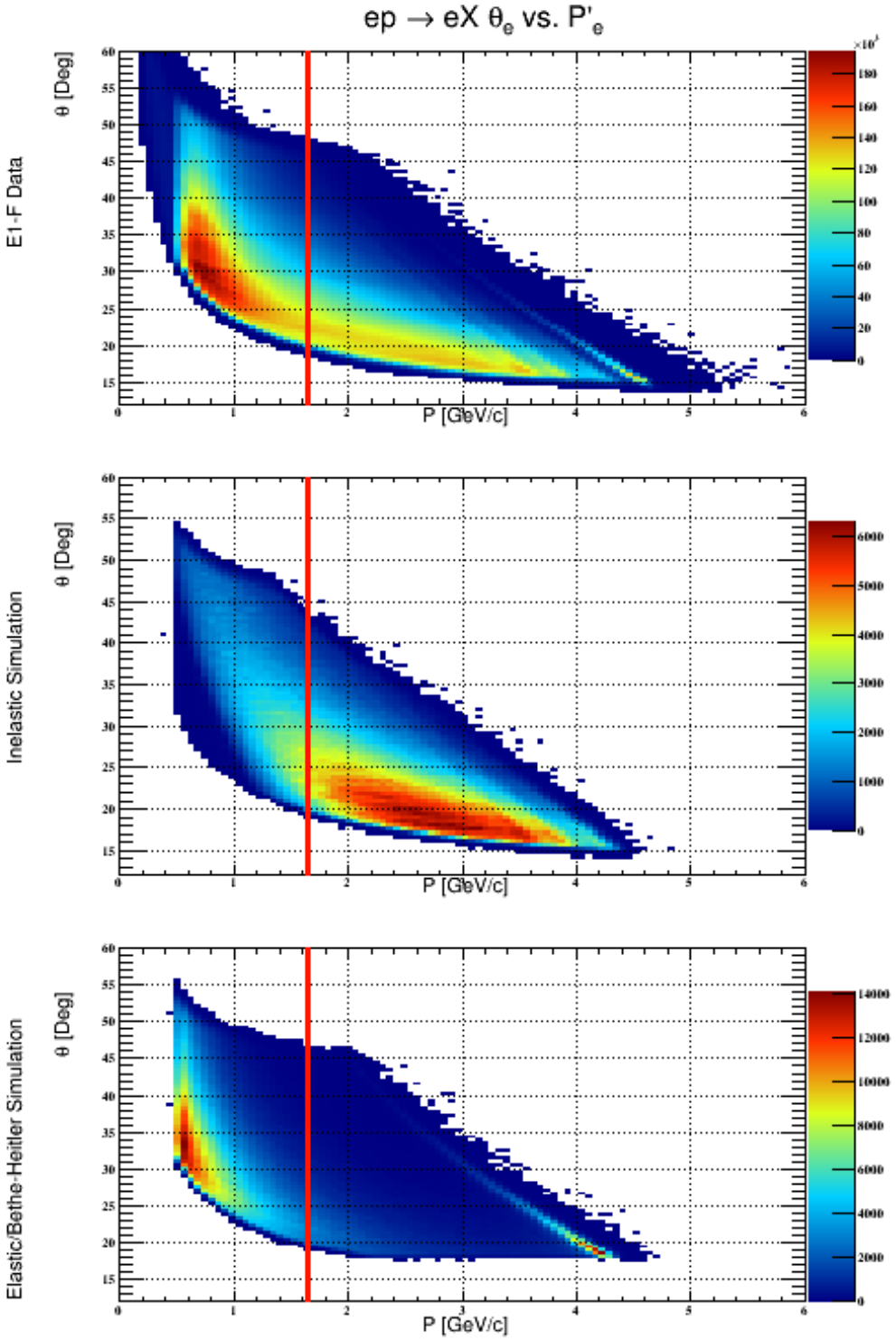

Fig. 5.2: Event distributions (θ_e vs p) for data, simulated inelastic events, and simulated elastic events with radiation are shown. The red line indicates the momentum cut applied by restricting $y < 0.7$. The band of events on the right side of the figure contains all un-radiated elastic events. The larger portion of events present on the left side are Bethe-Heitler events which we wish to avoid in our analysis.

5.3.3 Simulation

All processes that CLAS measures are observed through the combination of signals from several sub-detectors. During analysis all sub-systems are calibrated accurately, but such a complicated system still often produces distributions that do not look like the true physical distribution. This discrepancy arises from the combination of several effects.

1. Holes, barriers, obstructions, shadows of other detectors, and any other physical effects that prevent events from being measured in some range of θ, ϕ are known as geometrical acceptance effects. An important geometrical acceptance effect is the presence of the torus coils in between every sector. These represent a complete loss of information for a small range of ϕ between each sector.

2. Inefficiencies due to the probabilistic nature of particle interaction in the detector subsystems.

3. Detectors have finite spatio-temporal resolutions.

In order to understand and limit the impact of these effects on the physics extracted from the experiment, a mock experiment is simulated. In the simulated experiment everything is modeled as realistically as possible. The simulation used for CLAS is called GSIM and is based on the CERN package GEANT3 (GEometry ANd Tracking).

In this controlled environment, samples of events can be generated and fed into the simulation. The output of GSIM is a bos file that is similar to the raw data from the data acquisition system. This is then reconstructed using the same reconstruction algorithm that is applied to data (`userana`).

5.3.4 Acceptance Corrections (Theory)

By retaining the truth information for all particles that are generated, the effect of the detector can be studied completely. These concepts can be stated more formally by considering the true $t(x')$ and measured $m(x)$ distributions of some observable. In the absence of background processes, the relationship between these distributions is expressed as a Fredholm integral equation of the first kind.

$$m(x) = \int_{\Omega} K(x, x')t(x')dx' \tag{5.3}$$

Here $K(x, x')$ is a kernel which encodes information about detector acceptance due to the effects described above. The goal of the Monte Carlo simulation is to *unfold* the measured distribution $m(x)$ by providing an estimate of $K(x, x')$ and finally corrected the data to get $t(x)$.

Observed events are usually aggregated into bins and the problem is naturally discretized and written in vector-matrix form.

$$\mathbf{Ax} = \mathbf{y} \tag{5.4}$$

In this notation $\mathbf{A}$ represents the response matrix, a discretized version of the kernel function K, the vector $\mathbf{y}$ represents the measured distribution in the bins, and the vector $\mathbf{x}$ is the true distribution over the bins. The matrix elements A_{ij} can be estimated by using generating events, passing them through a Monte-Carlo detector simulation, and then counting the number of events that are reconstructed in bin i when generated in bin j. This quantity is then normalized by the total number of events generated in the j^{th} bin. In the absence of bin migration and with perfect acceptance this matrix is the identity matrix I^n where n is the total number of bins.

$$A_{ij} = \frac{n_{rec=i,gen=j}}{n_{gen=j}} \tag{5.5}$$

The binned true distribution can be recovered by inverting the response matrix and correcting the observed distribution.

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{y} \tag{5.6}$$

In the absence of bin migration, the matrix becomes diagonal with efficiency elements ϵ_i that represent the fraction of events reconstructed in the bin i.

$$\mathbf{A} = \begin{pmatrix} \epsilon_0 & 0 & 0 \\ \vdots & \ddots & \\ 0 & & \epsilon_n \end{pmatrix} \tag{5.7}$$

The inverse is,

$$\mathbf{A}^{-1} = \begin{pmatrix} 1/\epsilon_0 & 0 & 0 \\ \vdots & \ddots & \\ 0 & & 1/\epsilon_n \end{pmatrix} \tag{5.8}$$

and the corrected observation for the i^{th} bin is simply given by the observation over the efficiency.

$$t_i = \frac{m_i}{\epsilon_i} = m_i \frac{n_{gen=i}}{n_{rec=i}} \tag{5.9}$$

This is the simple *bin-by-bin* acceptance correction method, which is widely used and produces accurate results provided that bin migration is not significant. In this analysis the simple bin-by-bin acceptance correction is used.

5.3.5 Acceptance Corrections

5.3.6 Radiative Corrections

Inclusive events detected in CLAS are really *radiated* inclusive events. By using the term *radiated*, one implies that the electron could have radiated a real photon in the initial or final state, or that internal radiative correction diagrams could

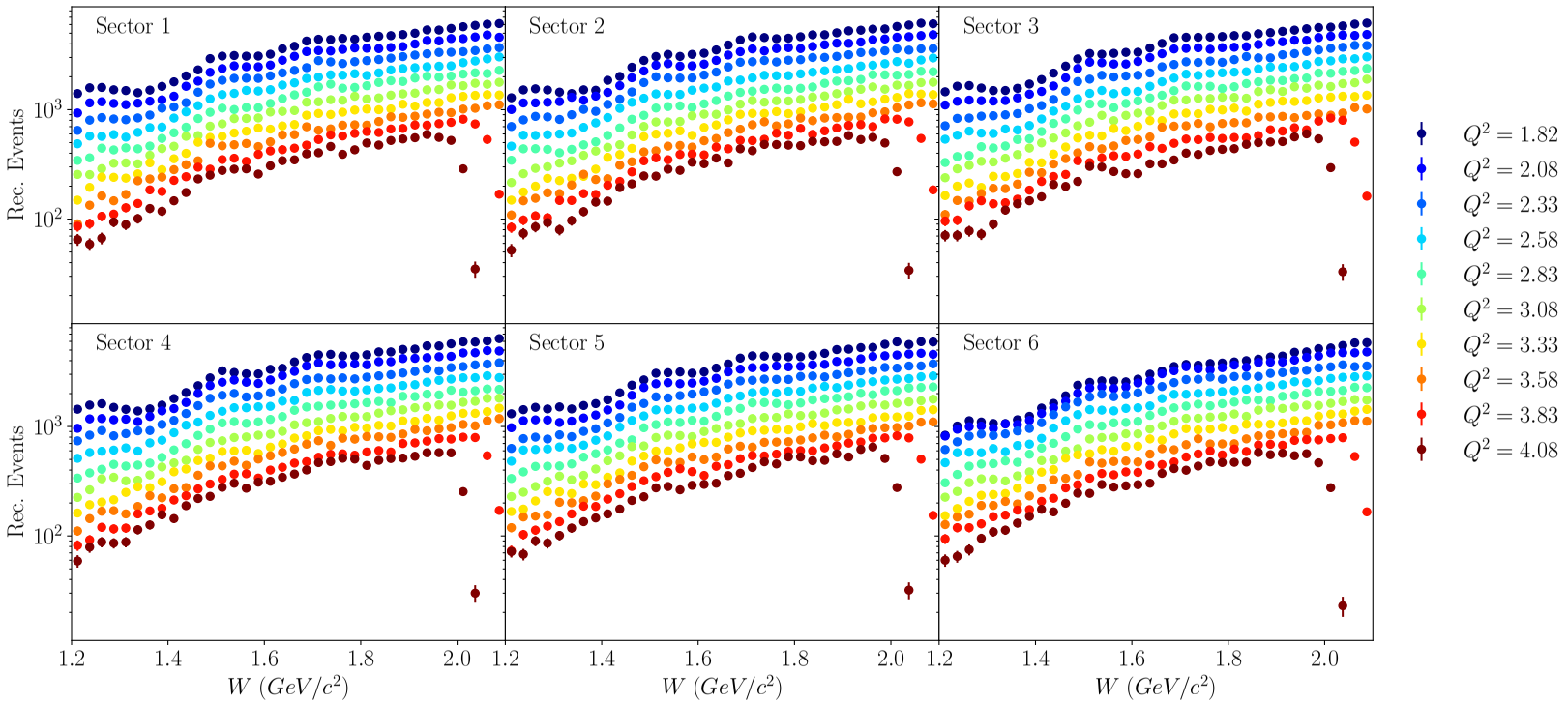

Fig. 5.3: Reconstructed Monte Carlo events displayed for different sectors (by panel) and different Q^2 bins represented by different colors.

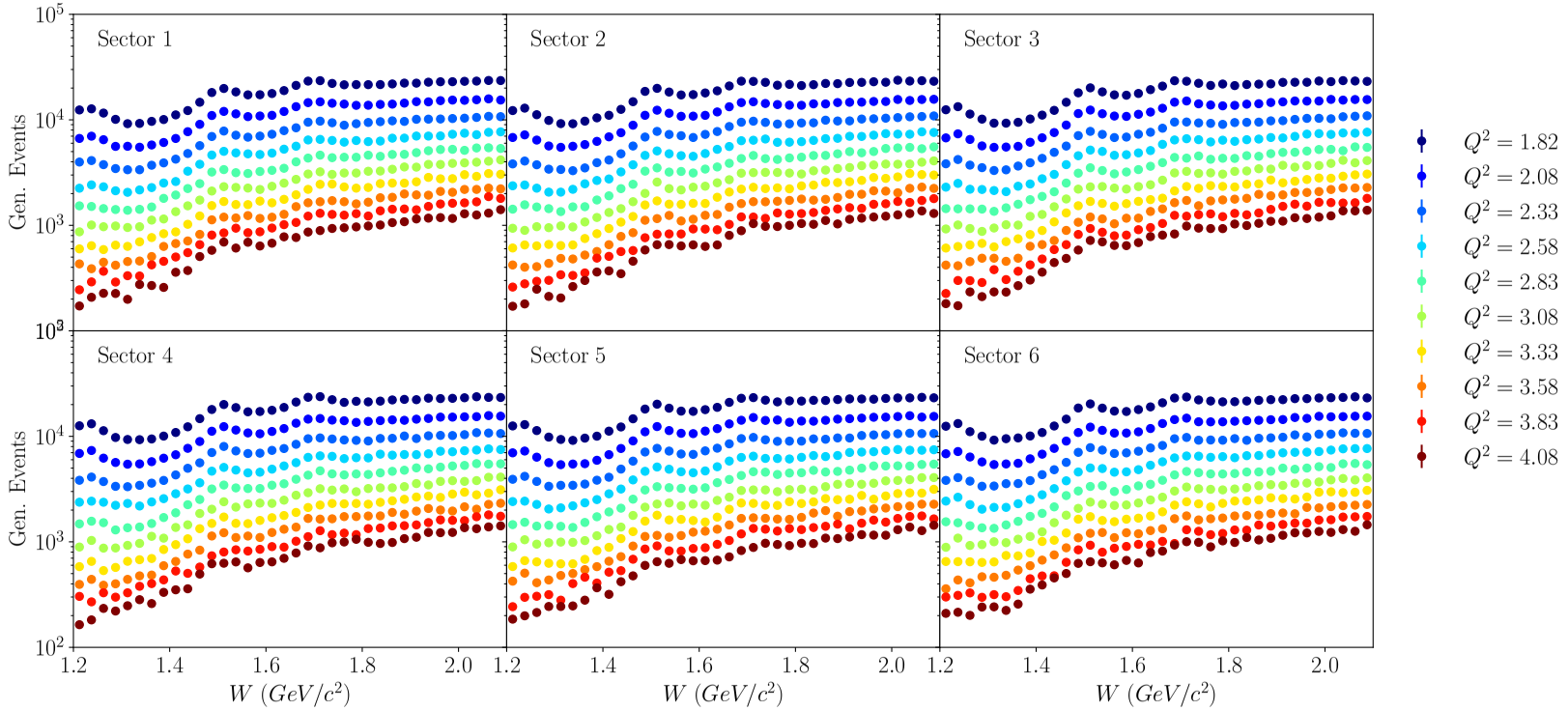

Fig. 5.4: Generated Monte Carlo events displayed for different sectors (by panel) and different Q^2 bins represented by different colors. These events were generated with radiative effects.

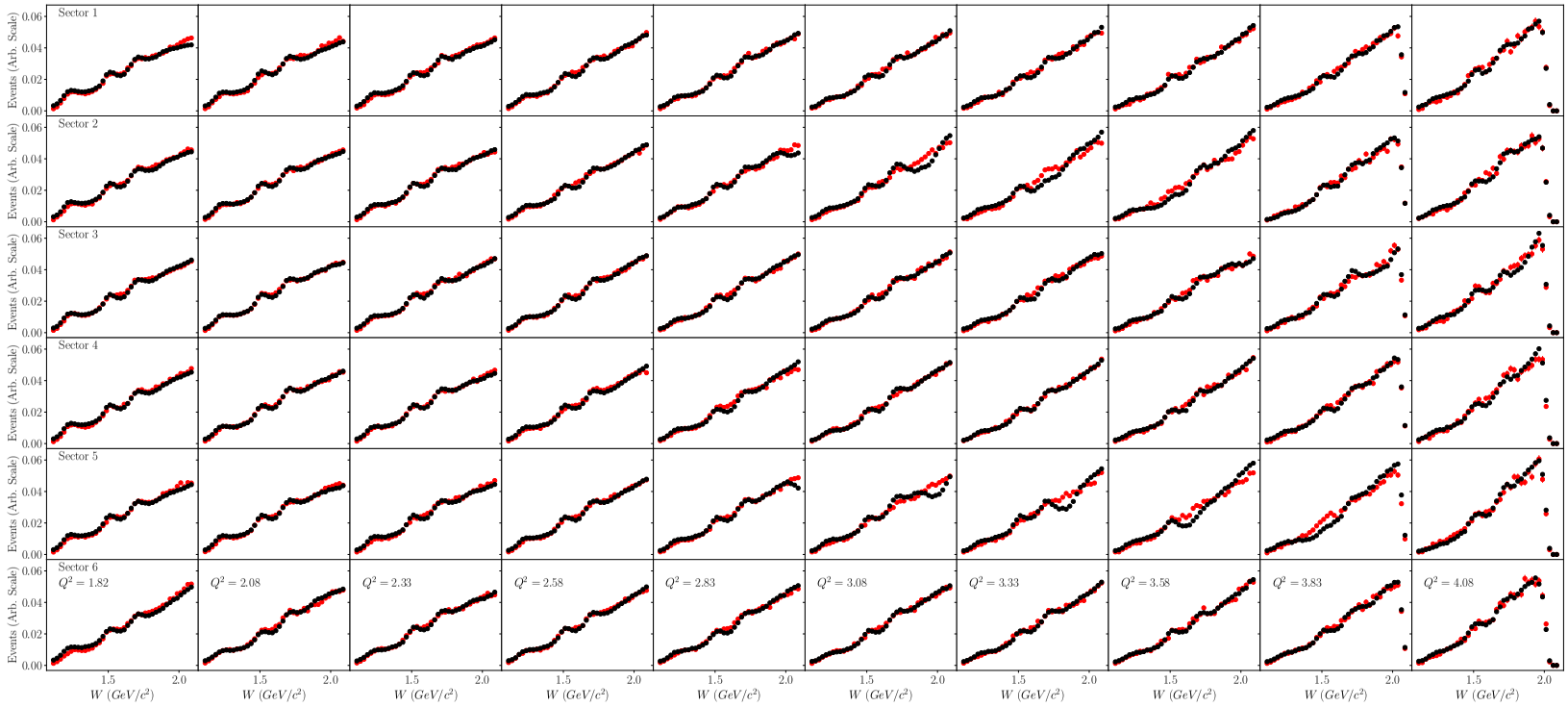

Fig. 5.5: Reconstructed events from data (black) and Monte Carlo (red) are superimposed to show that the simulation is a good approximation to the physics, a fact important for the accurate calculation of acceptance.

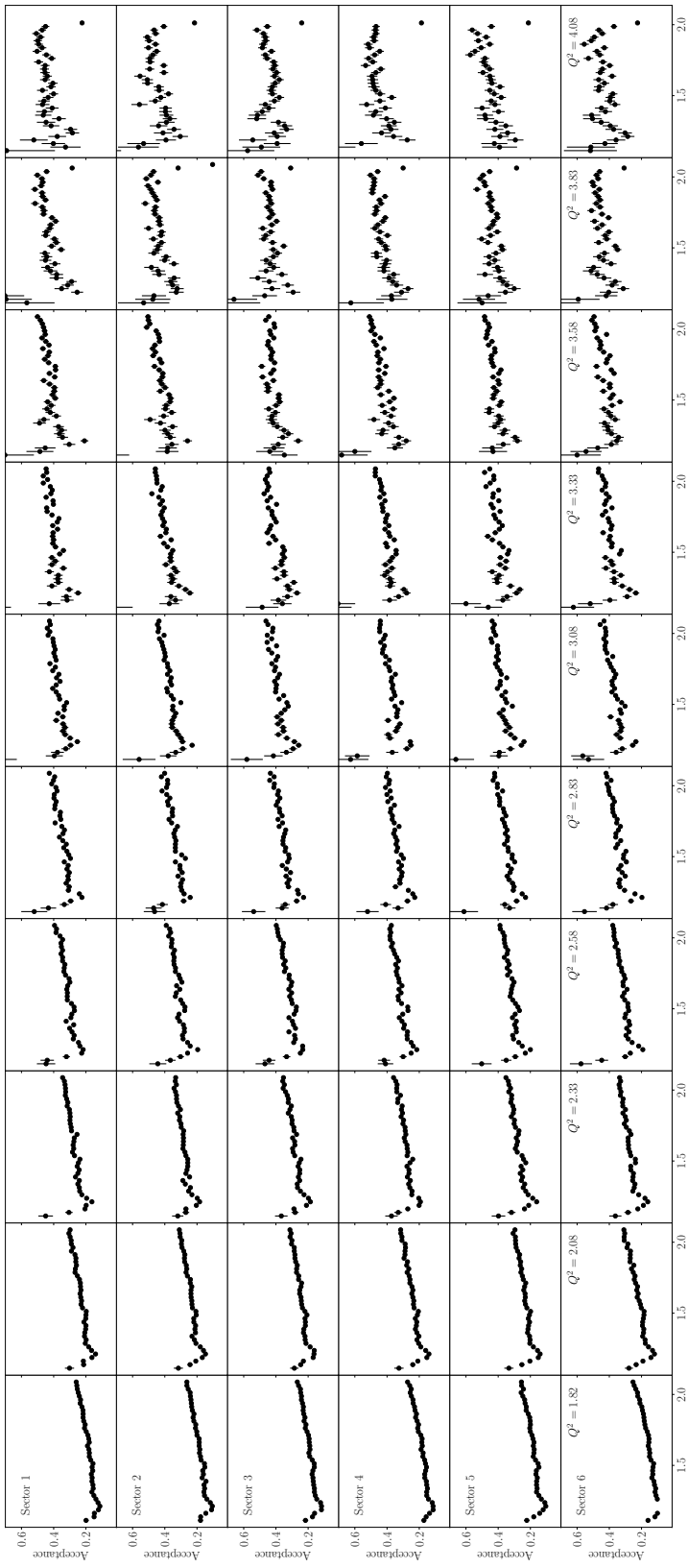

Fig. 5.6: Acceptance corrections applied to the inclusive cross section are shown above in sixty panels that correspond to increasing Q^2 from left to right, and increasing sector number from top to bottom.

have influenced the event kinematics. Unfortunately, there is no way of differentiating these events. For this reason, a radiative correction procedure that unfolds these effects from the cross section is applied to our measured distributions.

Two Monte Carlo event generators are used to calculate radiative corrections for each bin, both sample from the cross section model discussed in the beginning of the chapter. The first program generated events with no radiative corrections by sampling directly from the cross section model. The second program uses the same cross section model but generates radiated events following the procedure of Mo and Tsai [33] which includes corrections due to internal and external Bremsstrahlung. The correction ratio for the i^{th} bin $R^{(i)}$ is defined for the i^{th} bin as shown below.

$$R^{(i)} = \frac{n^{(i)}_{unrad}}{n^{(i)}_{rad}} \tag{5.10}$$

This factor can be estimated without passing events through the simulation (this is true if the acceptance is the same for both models, an assumption which is only invalid if the models are appreciably different) and the results are used directly from the output of the event generator to correct the cross section.

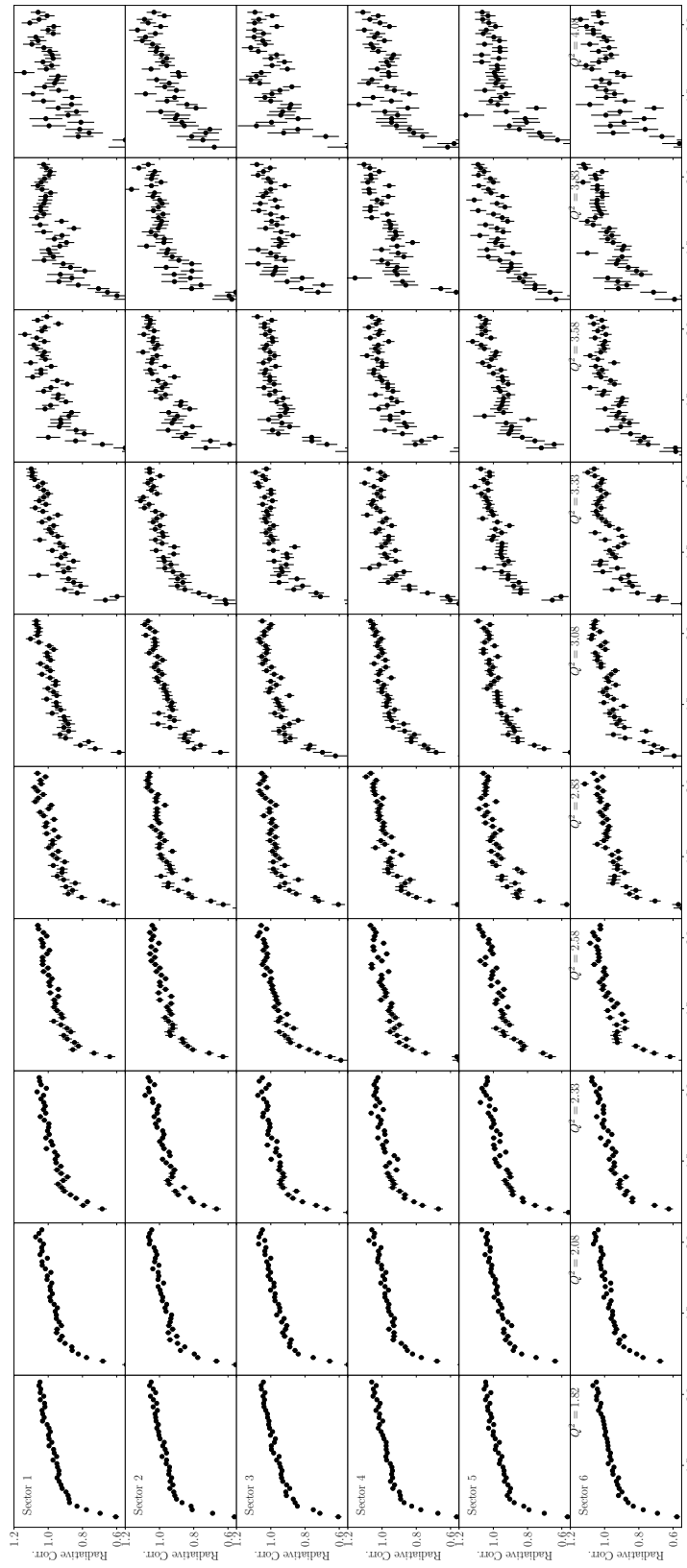

Fig. 5.7: Radiative corrections applied to the inclusive cross section are shown above in sixty panels that correspond to increasing Q^2 from left to right, and increasing sector number from top to bottom.

5.3.7 Bin Center Corrections

Measurement of the cross section is done in bins, meaning that what is actually measured is the average cross section over some finite range of W and Q^2 for each bin. When reporting and plotting the results the W and Q^2 value at the center of the bin is typically quoted. If a model is available for comparison, the model is usually queried at the bin center, and the resulting prediction is often incorrect. In order to obtain the correct prediction, one should average the model prediction over the measured bin. In this study, an accurate model prediction is available (described in detail below) and a correction factor is applied to the measured cross section value to remove the effect of plotting and reporting the cross section at the central bin value. The factor,

$$B_i = \frac{\sigma_{center}}{\sigma_{avg}} \tag{5.11}$$

B_i is calculated from the model for each bin i and applied to the measured cross section. This factor is calculated once for each bin, and does not depend on the sector. We observe that the correction factor depends more strongly on Q^2 than W which is sensible because the cross section varies more rapidly over Q^2. Additionally, the correction factor is larger near the resonances in W.

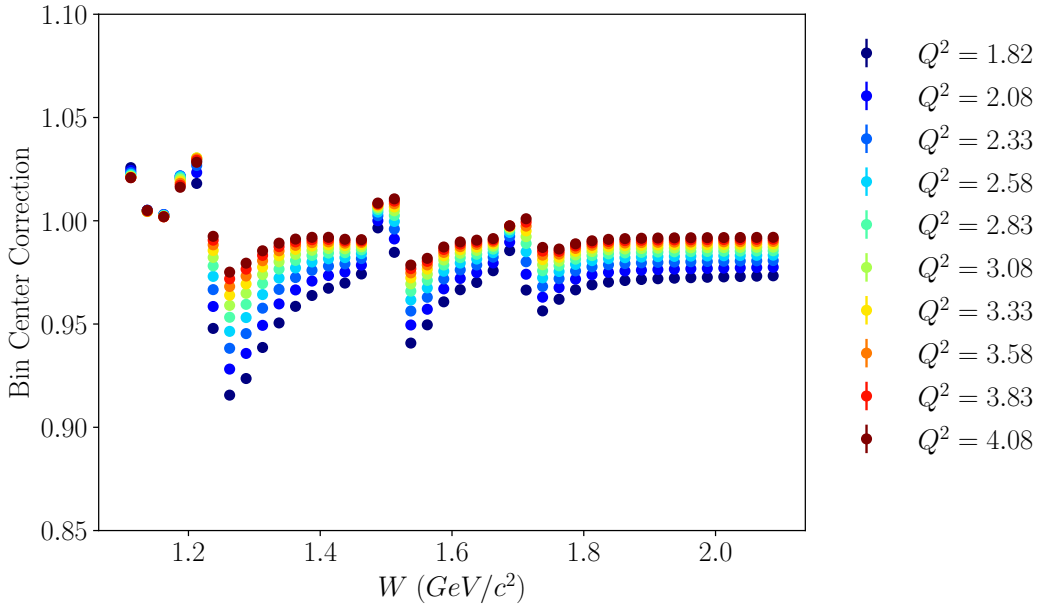

Fig. 5.8: The bin center correction applied to the inclusive cross section is shown for different values of Q^2 (indicated by color) as a function of W on the horizontal axis. This correction is the same for all sectors.

5.3.8 Model Comparison

The cross section extracted in this study is in agreement with the predictions from the model, shown in figure 5.3.8.

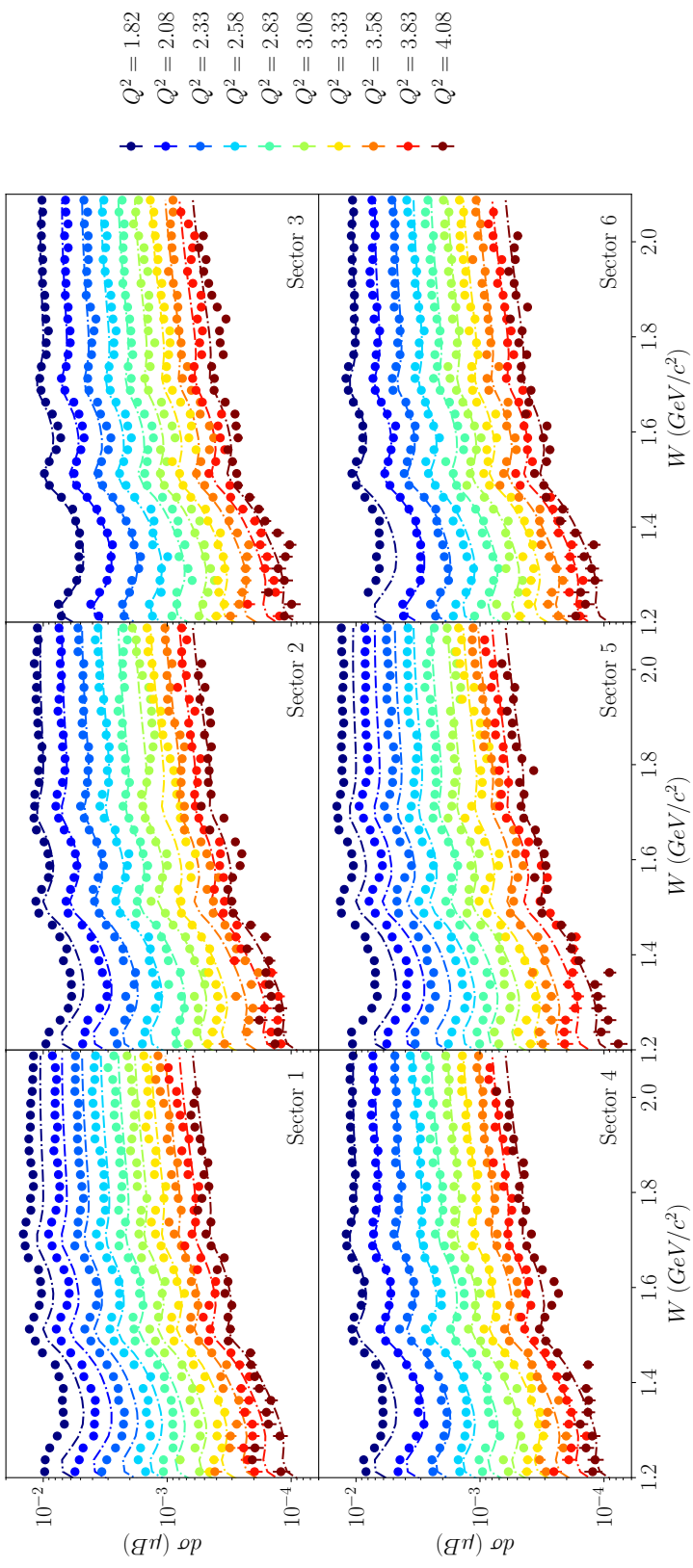

Fig. 5.9: Each panel above shows one of the six sectors of CLAS, the horizontal axis shows W and superimposed on the figure are the cross sections for different bins of Q^2. For simplicity of the visualization each reported value of Q^2 is the bin center.

Chapter 6

SIDIS Cross Section

6.1 Introduction

The SIDIS cross section for scattering of unpolarized electrons on an unpolarized proton target can be written in terms of three coefficients A_0, $A_{UU}^{cos(\phi_h)}$, and $A_{UU}^{cos(2\phi_h)}$ as shown below. Each of these coefficients depends on the electron and hadron kinematic variables $A \rightarrow A(x, Q^2, z, P_T^2)$ which have been omitted for readability.

$$\frac{d^5\sigma}{dx\ dQ^2\ dz\ dP_T^2\ d\phi_h} = A_0\Big[1 + A_{UU}^{\cos\phi_h}\cos\phi_h + A_{UU}^{\cos(2\phi_h)}\cos(2\phi_h)\Big] \tag{6.1}$$

The A coefficients are defined in terms of structure functions F, and were measured from the E1-F dataset by N. Harrison as reported in [24].

$$A_0 = \frac{\pi\alpha^2 y(1+\gamma^2/2x)}{2EM_p x^2 Q^2(1-\varepsilon)}\Big(F_{UU,T} + \varepsilon F_{UU,L}\Big) \tag{6.2}$$

$$A_{UU}^{\cos\phi_h} = \sqrt{2\varepsilon(1+\varepsilon)}\frac{F_{UU}^{\cos\phi_h}}{F_{UU,T} + \varepsilon F_{UU,L}} \tag{6.3}$$

$$A_{UU}^{\cos(2\phi_h)} = \varepsilon\frac{F_{UU}^{\cos(2\phi_h)}}{F_{UU,T} + \varepsilon F_{UU,L}} \tag{6.4}$$

To measure the structure functions, the five dimensional differential cross section is measured 6.5. The ϕ_h measurements in each kinematic bin (x, Q^2, z, P_T^2) are used to perform parameter estimation of the terms A_0, $A_{UU}^{cos(\phi_h)}$, and $A_{UU}^{cos(2\phi_h)}$. The structure functions can then be measured by rearranging equation 6.4 as shown in 6.6.

$$\frac{d^5\sigma}{dx\ dQ^2\ dz\ dP_T^2\ d\phi_h} = \frac{1}{\Delta x\ \Delta Q^2\ \Delta z\ \Delta P_T^2\ \Delta\phi_h}\frac{N_{obs}^{(i)}}{\mathcal{L}A^{(i)}R^{(i)}} \tag{6.5}$$

$$F_{UU,T} + \varepsilon F_{UU,L} = A_0\frac{2(1-\varepsilon)E_B M_p x^2 Q^2}{\pi\alpha^2 y(1+\gamma^2/2x)} \tag{6.6}$$

$$F_{UU}^{\cos\phi_h} = A_{UU}^{\cos\phi_h}\frac{F_{UU,T} + \varepsilon F_{UU,L}}{\sqrt{2\varepsilon(1+\varepsilon)}} \tag{6.7}$$

$$F_{UU}^{\cos(2\phi_h)} = A_{UU}^{\cos(2\phi_h)}\frac{F_{UU,T} + \varepsilon F_{UU,L}}{\varepsilon} \tag{6.8}$$

Results for the extracted structure functions as well as the five dimensional

Hadron	p_{min} (GeV)	p_{max} (GeV)
π^+	0.2	3.75
π^-	0.2	3.25

Table 6.1: The limits used for pion momentum used in the fitting of β for both charged pi-mesons.

differential cross section are presented at the conclusion of this chapter.

6.2 Hadron Identification

In chapter 4, the hadronic identification methods described used a probabilistic interpretation of the difference between the observed values of β and the expected values (determined by fitting small deviations from the theoretical case). This study was performed before that identification routine was developed, and therefore uses a slightly different boundary and interpretation (described first by N. Harrison in his thesis work [24]). Fiducial cuts used for positives and negatives are those described in chapter 4, and will not be discussed here.

In order to determine the cut boundaries, the events are fit with an unnormalized Gaussian distribution in each of 70 momentum bins and the mean μ and standard deviation σ are recorded. Pions which fall between $\mu - 3\sigma \leq \beta \leq \mu + 3\sigma$ are kept for analysis. At higher momentum positive pi-mesons are difficult to separate from protons than at lower momentum values. To accommodate this, the σ value is reduced after 2 GeV.

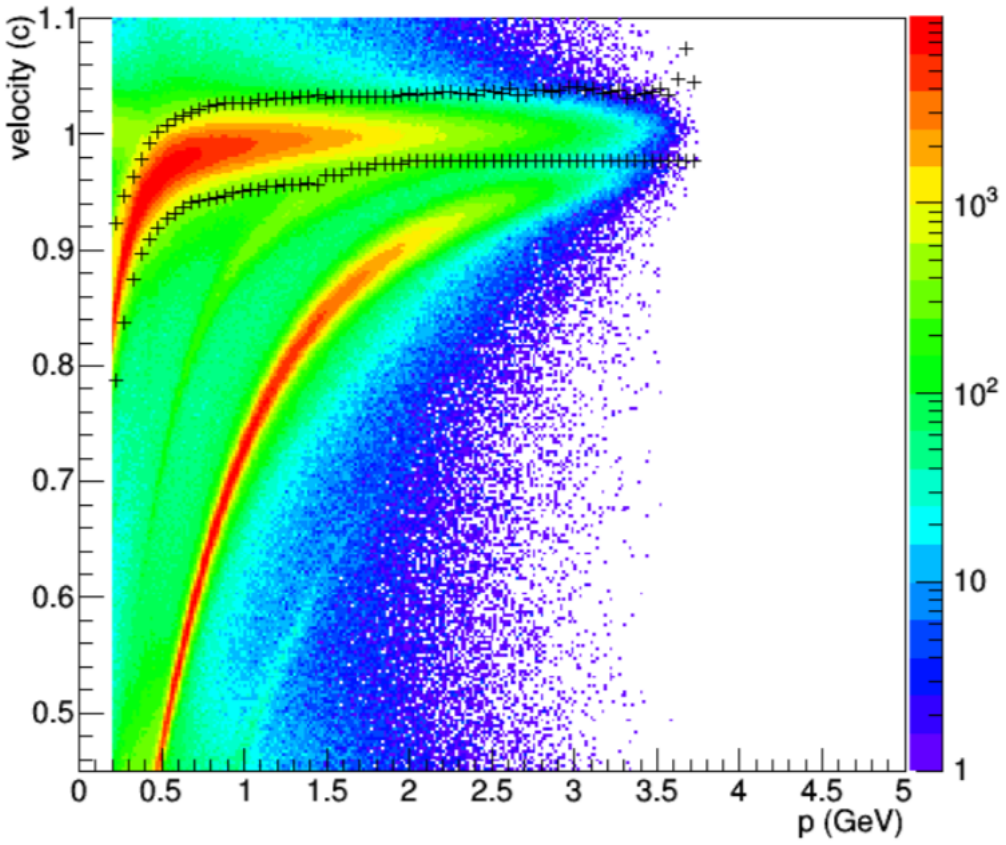

Fig. 6.1: Positively charged π mesons are identified by applying cuts to the momentum dependent β.

6.3 Event Selection

The identification of an event which contains a good electron and a good pion ($\pi^{\pm}$) is the first step in selecting SIDIS events from the dataset. Next, events that meet the working assumption for what constitutes the DIS region ($Q^2 > 1.0\ GeV^2$ and $W > 2.0\ GeV$) are selected. A missing mass cut is used in this analysis to remove low lying exclusive resonances which are not produced by the SIDIS mechanism. Accordingly it is required that the missing mass $M_X(ep \rightarrow ehX)$ be greater than 1.35 GeV (in the positive channel this excludes the delta resonance).

One additional kinematic restriction is applied at the event selection stage, a cut on the maximum allowable inelasticity y_{max}. Full radiative corrections will be discussed in detail, but this kinematic restriction excludes events which have extremely high values of Q^2 and are much more likely to have radiated a photon in the initial or final state. In this analysis $y_{max} = 0.85$ is used.

6.4 Binning

Both π^+ and π^- are binned using the same scheme. The momentum fraction x is divided into five equally sized bins from 0.1 - 0.6. With the exception of the highest x bin, each is split into two Q^2 bins. This binning scheme for $x - Q^2$ is displayed in figure 6.2.

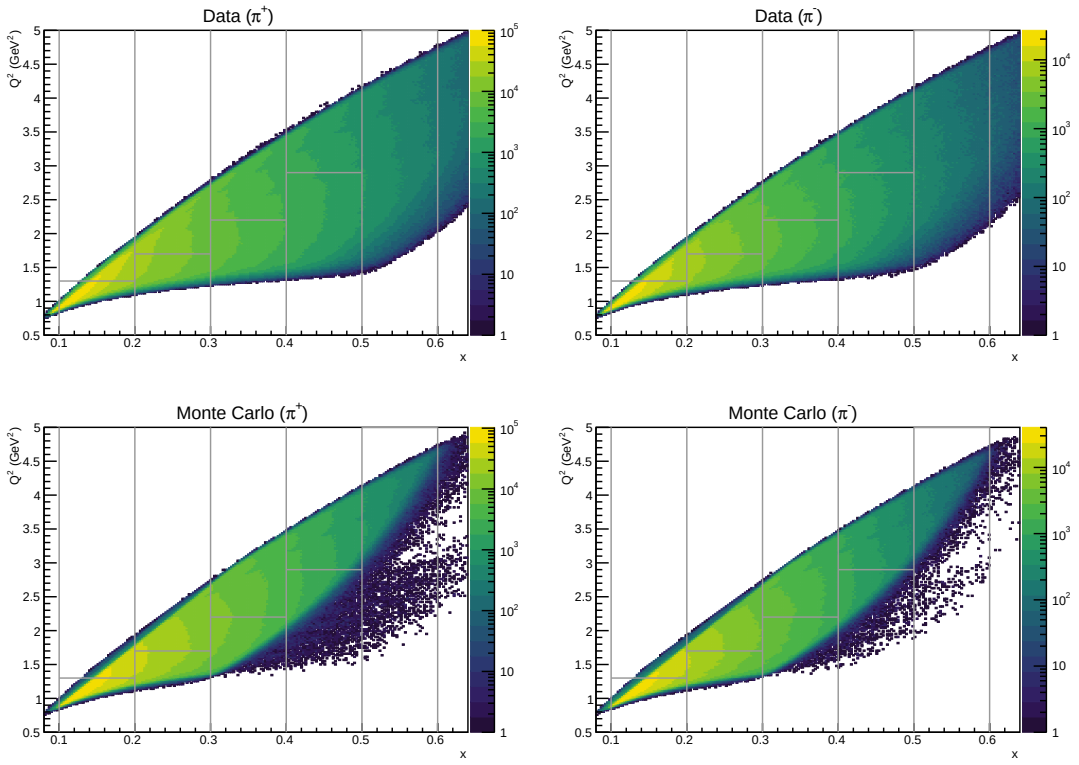

Fig. 6.2: The kinematic distributions for $\pi^{\pm}$ electron variables shown with binning overlaid.

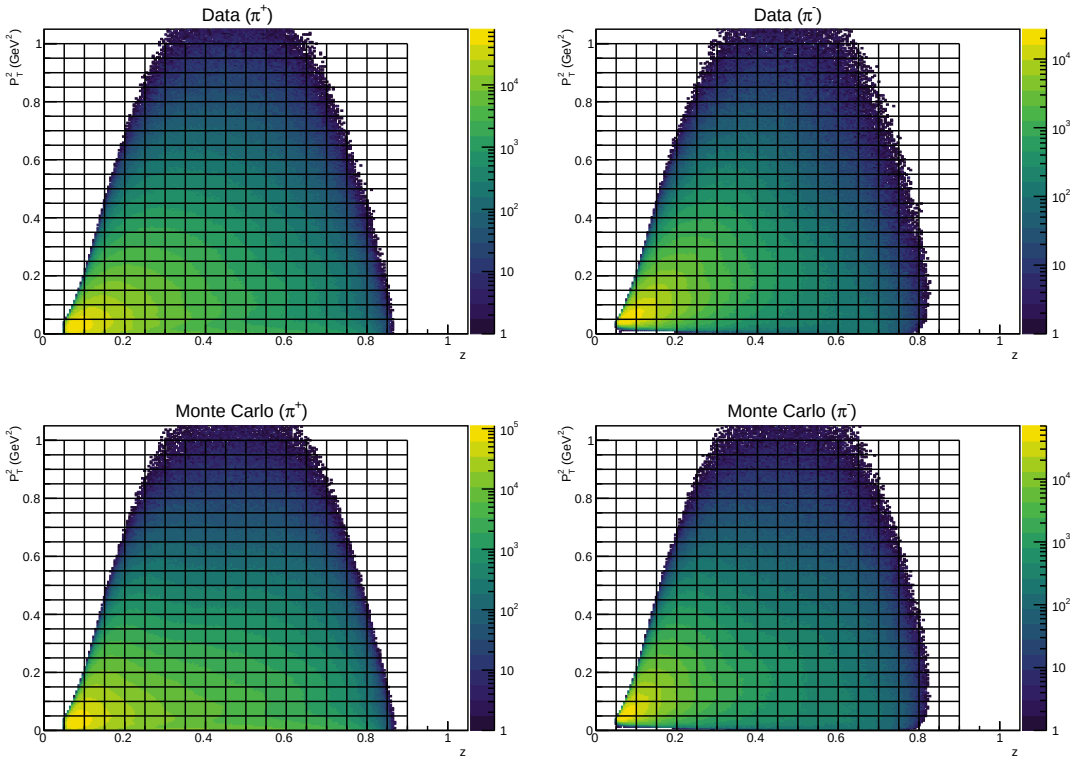

Fig. 6.3: The kinematic distributions for $\pi^{\pm}$ hadron variables shown with binning overlaid.

The hadronic variables are binned finer, the z axis is divided into 18 equal sized bins between 0-0.9, and the transverse momentum squared P_T^2 is binned in 20 equally sized bins from 0-1 GeV^2. This simple grid is shown on figure 6.3. The average value of the kinematic variables in each bin is calculated from the event sample and displayed in the results table at the end of this thesis.

6.5 MC Simulation

Acceptance corrections based on Monte Carlo simulation of the CLAS detector are vital in the accurate calculation of cross sections, and are performed in this work by using a modified version of `PYTHIA` called `clasDIS`, the detector simulation `GSIM`, and the resolution smearing program called `GPP`. Two main challenges existed in calculating acceptance corrections for this cross section measurement. The first challenge, which is common to all acceptance calculations, is that the acceptance depends weakly on the model input used. The second challenge facing this calculation is that at present the values of the coefficients A which shape the ϕ_h distributions are not known, and therefore the events are simulated flat.

In order to resolve the first challenge, the acceptance model used is included as a source of systematic uncertainty. This quantifies the amount to which the answer changes based on different input physics models. In answering the second challenge posed, the model is made more realistic by an iterative procedure of

using the extracted $A^{(0)}$ coefficients to weight the generated events and the and produce the coefficients $A^{(1)}$, where the superscript denotes the iteration. After two iterations, the extracted values of A no longer vary outside of parameter estimation errors. The last (2^{nd}) iteration is used for the simulation of events in this analysis, and the subsequent iteration (1^{st}) is used to study systematic uncertainties associated with the difference between both models.

6.5.1 Acceptance Corrections

The generator `clasDIS` is used with the modified ϕ_h dependence and more than 800 million π^+ events are simulated, as well as more than 600 million π^- events. The acceptance is calculated over the phase space and used to correct the cross section. A sample of the acceptance is shown in figure 6.4. Two main features are evident. First, the segmented nature of the CLAS detector in the azimuthal direction correlates to the center of mass hadron-electron angle ϕ_h and causes a periodic modulation of acceptance. Additionally, the central region of ϕ_h has the lowest acceptance of any value. For this reason, bins in the central region are excluded if they have large statistical errors from acceptance or have extremely low acceptance values. The resulting lack of points in the central region impacts the parameter estimation, and is discussed later in the section on parameter estimation.

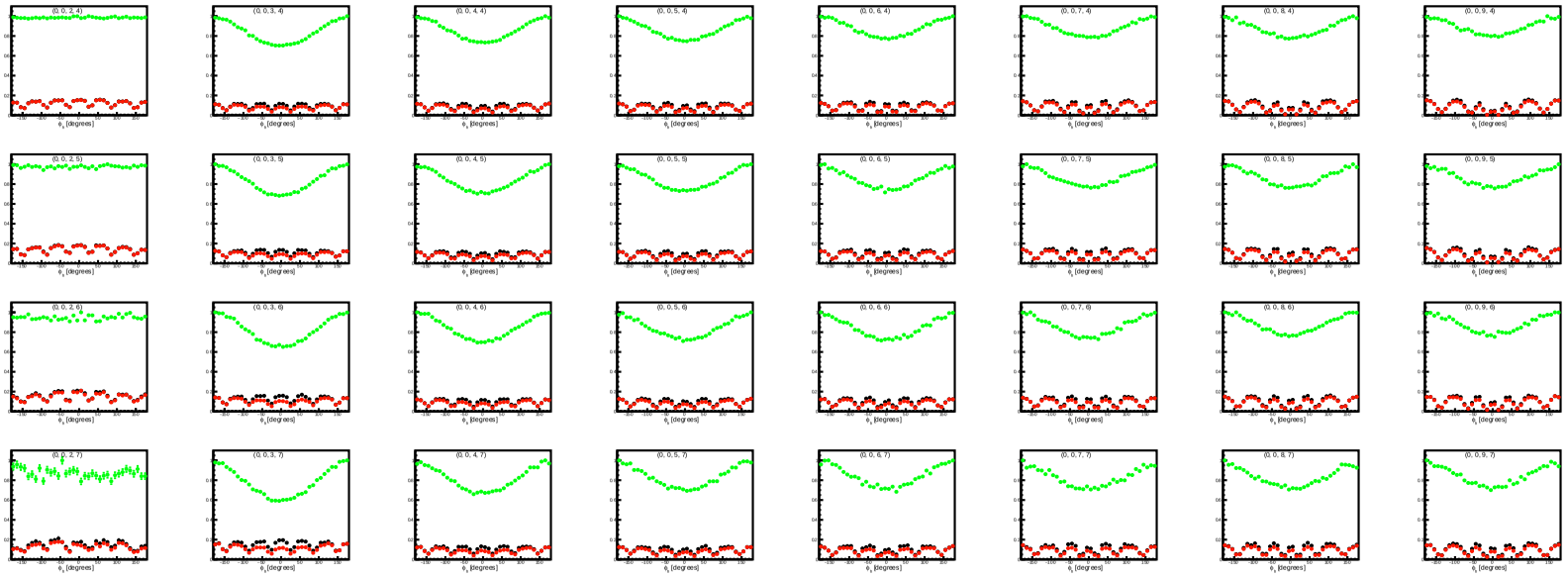

Fig. 6.4: Acceptance corrections shown for different z bins (increasing over the horizontal axis) and P_T^2 bins (increasing down the vertical axis) are displayed in red. Generated events are displayed in green, and in red the reconstructed events are shown normalized by the maximum number of generated events in any bin. On each figure the complete bin index is given in the format (x, Q^2, z, P_T^2).

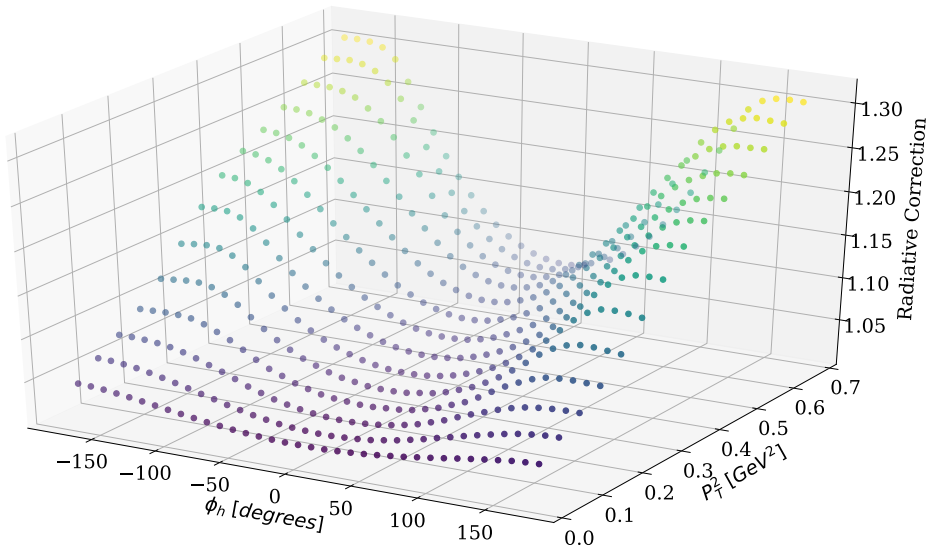

Fig. 6.5: Radiative corrections are displayed as a 3-dimensional scatter plot for the bin with indices $(x = 0, Q^2 = 1, z = 4)$ for π^+. This correction includes exclusive contributions σ_{tail}.

6.6 Radiative Corrections

The Born cross section, in which no radiation takes place during the scattering, is the cross section that we intend to measure. However, the incoming and scattered electron can emit photons, and therefore the cross section measured in the lab is not the Born cross section but the *radiated* cross section. In addition to initial and final state radiation, radiated exclusive events can have long tails and enter into the SIDIS kinematics. These effects are calculable by using the software package called `HAPRAD`, which takes as an input a Born cross section model σ_{Born} and produces the radiated version of the cross section. This calculation is performed at each kinematic point that is measured.

$$R = \frac{\sigma_{rad}(x, Q^2, z, P_T^2, \phi_h)}{\sigma_{Born}(x, Q^2, z, P_T^2, \phi_h)} \tag{6.9}$$

This software package is used to correct our measured cross section by calculating the ratio R_i for each bin i (as shown in equation 6.9). Because the correction depends on the model used, two different models are used and the difference in the extracted structure functions is assigned as a systematic uncertainty in our final analysis.

6.7 Parameter Estimation (Fitting)

Once measurements of the 5-differential corrected yield are made, χ^2 minimization is used to estimate the value of the A coefficients in every bin of (x, Q^2, z, P_T^2). The difference between the data points and the model prediction is called the *residual*, and the square sum of these residuals is known as the mean squared error. The χ^2 is a simply extends this by the addition of an error term in the denominator.

$$\chi^2 = \sum_{i=1}^{n} \frac{(d_i - t_i)^2}{\sigma_i^2} \tag{6.10}$$

Here n refers to the degrees of freedom, or the number of data points used in fitting the model parameters. The `Minuit` minimization package which is included in CERN's ROOT is used to minimize the χ^2 function for each bin and provides

the parameters and their associated errors.

6.7.1 Minimal Coverage in ϕ_h for Fitting

As points are removed from the central ϕ_h region, fitting becomes less stable. This is particularly problematic as the periodicity of the function increases. Because the central region of ϕ_h has been removed in some cases due to low acceptance, this problem needs to be addressed before the fitting algorithm is applied to all the kinematic bins.

In order to alleviate this concern, the bins of (x, Q^2, z, P_T^2) that have a gap from $\Delta\phi_h > 60^\circ$ are not fit (-30 to 30 or larger). This value was set by generating psuedo-data with a known ϕ_h dependence given in terms of the three A coefficients. The distribution was fit, and then points were removed and the distribution was fit again. This was repeated for several different values of A, particularly the cosinusoidal terms. This study is described in detail in [24].

6.8 Systematic Uncertainties

Thirteen possible sources of systematic effect have been identified 6.8, and the values for each are varied slightly from the nominal values in order to observe the affect that each has on the final results. Each of the values is increased and decreased slightly in accordance with the amount of uncertainty that exists around

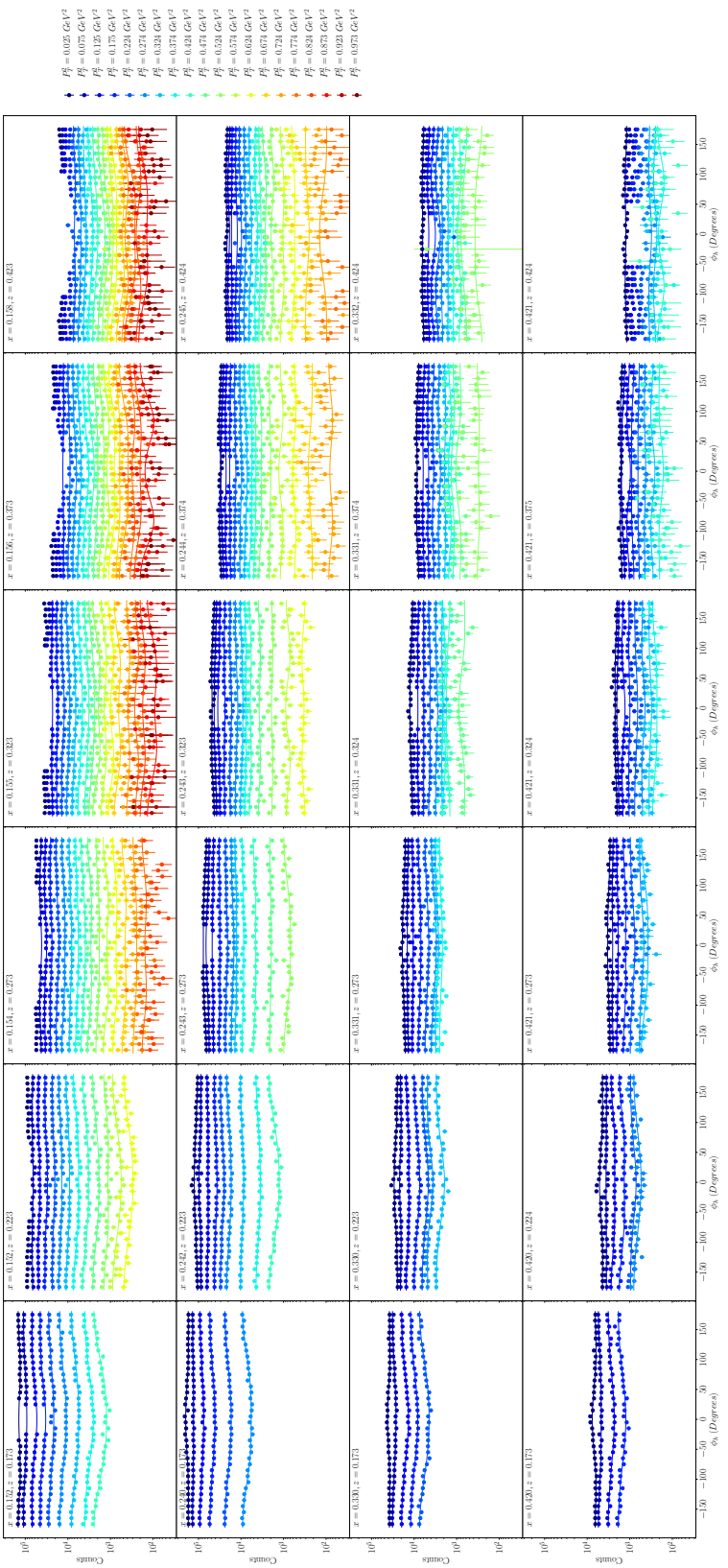

Fig. 6.6: Corrected π^+ yields are fit and the coefficients A_0, $A_{UU}^{cos(\phi_h)}$, and $A_{UU}^{cos(2\phi_h)}$ are estimated for each kinematic bin. In this figure bins of P_T^2 are superimposed. Rows in the figure indicate increasing x, while z increases with the column.

Label	Source	Description
0	electron z-vertex cut	Varied by ± 0.2 cm on each side
1	electron sampling fraction cut	shown in figure 6.9
2	electron EC E_{dep} cut	± 5 MeV
3	electron EC U, V, W cut	shown in figure 6.12
4	electron θ_{CC} matching cut	Varied by ± 0.2 cm on each side
5	electron region 1 fid. cut	shown in figure 6.7
6	electron region 3 fid. cut	shown in figure 6.8
7	electron CC fid. cut	shown in figure 6.11
8	pion β cut	± 0.25 σ
9	pion region 1 fid. cut	-
10	ϕ_h fid. cut	±10° on each side
11	acceptance model	second to last iteration used
12	radiative correction model	second to last iteration used

its ideal value. The uncertainty due to source i is calculated from these variations by calculating the RMS of the deviations, as shown in equation 6.11.

$$\Delta^{(i)} = \sqrt{\frac{1}{N_{vars}} \sum_{j=1}^{N_{vars}} (r^{(0)} - r^{(j)})^2} \tag{6.11}$$

Here N_{vars} is the number of variations performed for source i, and $r^{(0)}$ is the measured result with the nominal set of parameters and j is a an index for the sum over variations of this parameter. For the acceptance and radiative corrections model dependence instead of increasing and decreasing a value the model is changed and these are treated as variations.

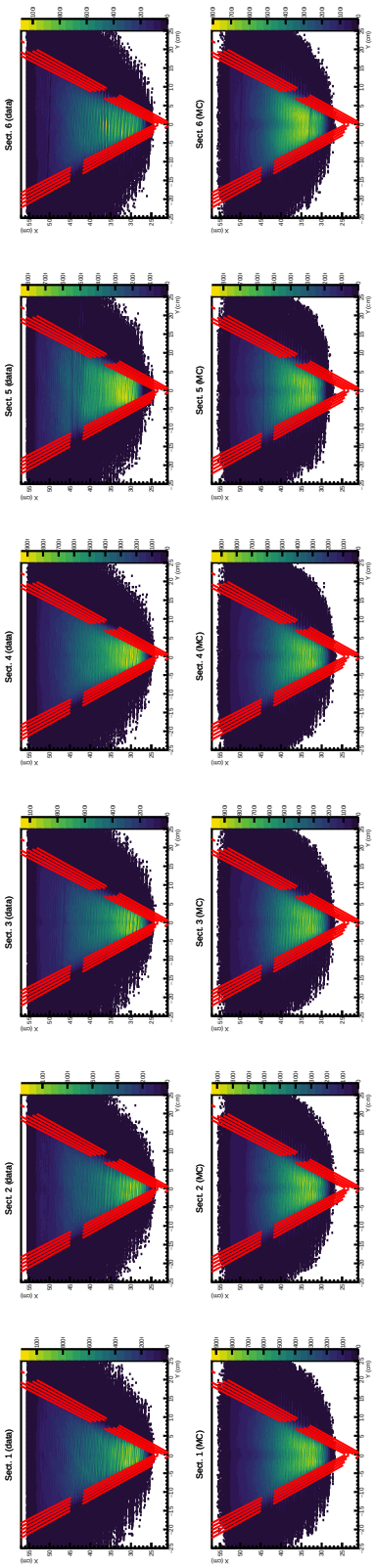

Fig. 6.7: Boundaries for electron identification cuts placed on the region 1 drift chambers are shown for data (top row) and Monte Carlo (bottom row).

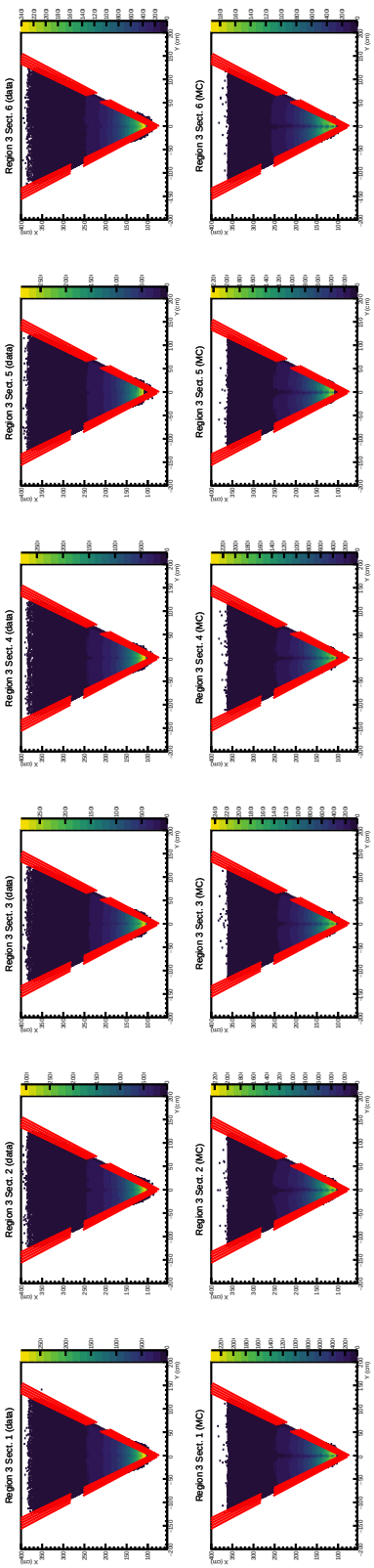

Fig. 6.8: Boundaries for electron identification cuts placed on the region 3 drift chambers are shown for data (top row) and Monte Carlo (bottom row).

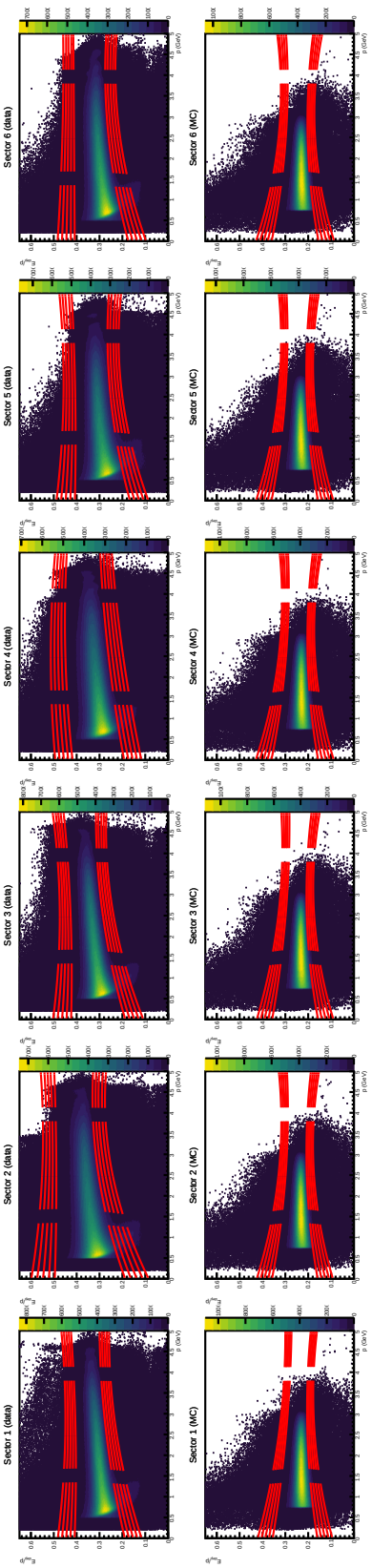

Fig. 6.9: The sampling fraction boundaries used to identify electrons in data (top) and Monte Carlo (bottom) are shown over the distributions.

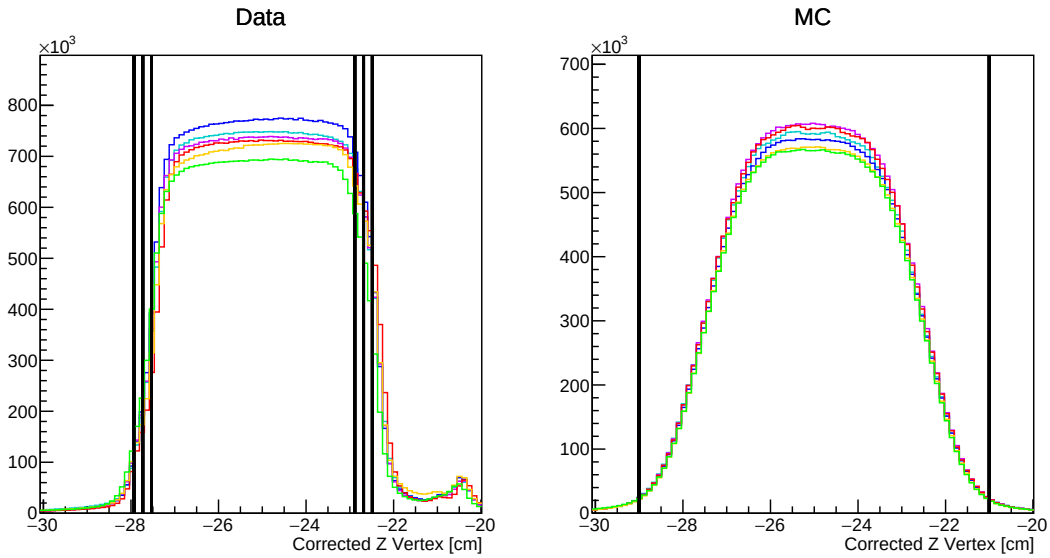

Fig. 6.10: The cut boundaries for z-vertex used to identify electrons in data (left) and Monte Carlo (right). This histograms shown for data have been corrected before being filled.

6.9 Results

The analysis of the unpolarized cross section presented in this chapter can be used immediately to constrain the kinematic dependence of the unpolarized TMD $f_1(x, k_\perp)$ for both favored and un-favored distributions. Interestingly, the results can also be compared with the results from HERMES and COMPASS unpolarized multiplicity studies. We additionally provide results for the azimuthally dependent structure functions $F_{UU}^{\cos\phi_h}$ and $F_{UU}^{\cos(2\phi_h)}$ which are sensitive to the Boer-Mulders TMD. Although the errors associated with these extractions is comparatively larger the values should be useful in constraining the parameters of Boer-Mulders in the higher-x region, and adding to the discussion of the contribution from the twist-four Cahn effect term.

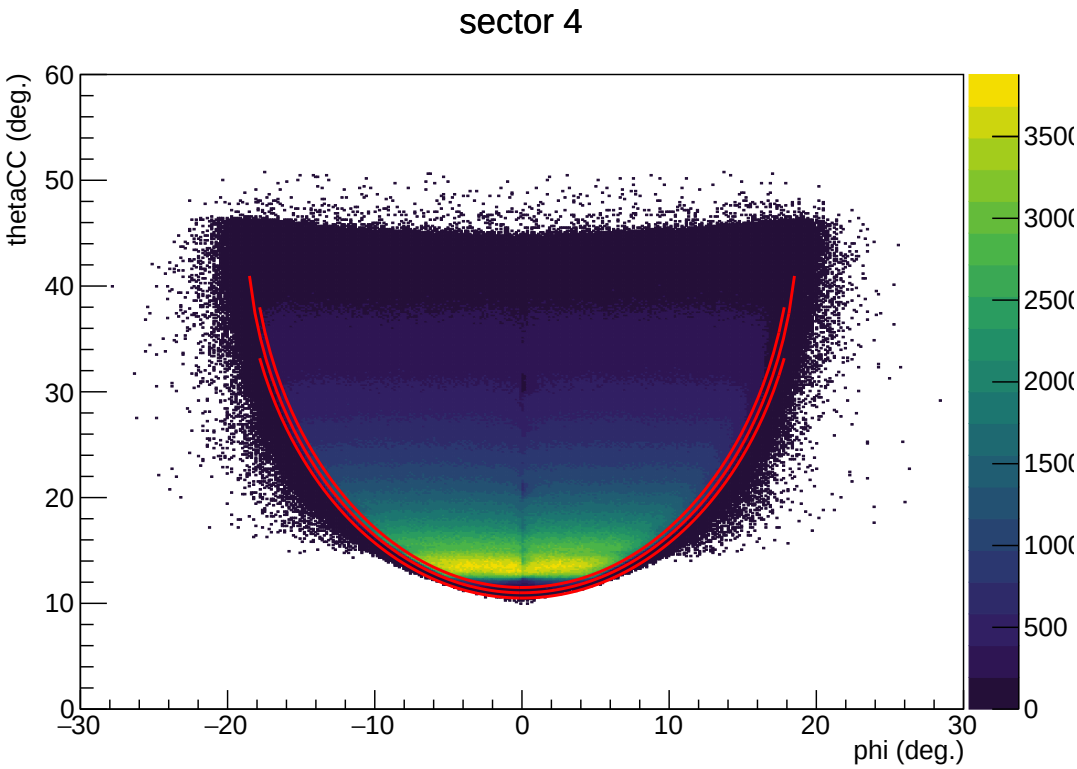

Fig. 6.11: The loose, nominal, and tight boundaries applied to data for the fiducial cuts on the Cherenkov Counter.

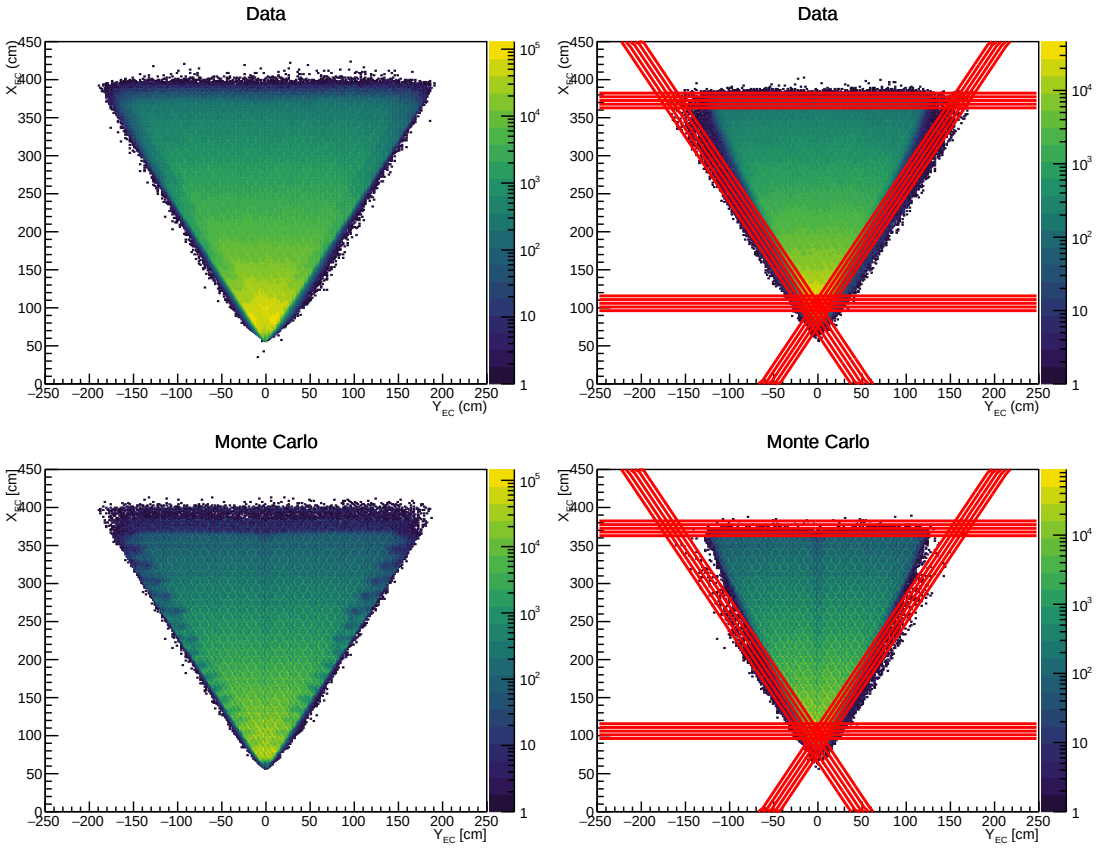

Fig. 6.12: Electron identification cuts on U, V, and W coordinates are shown in x-y space.

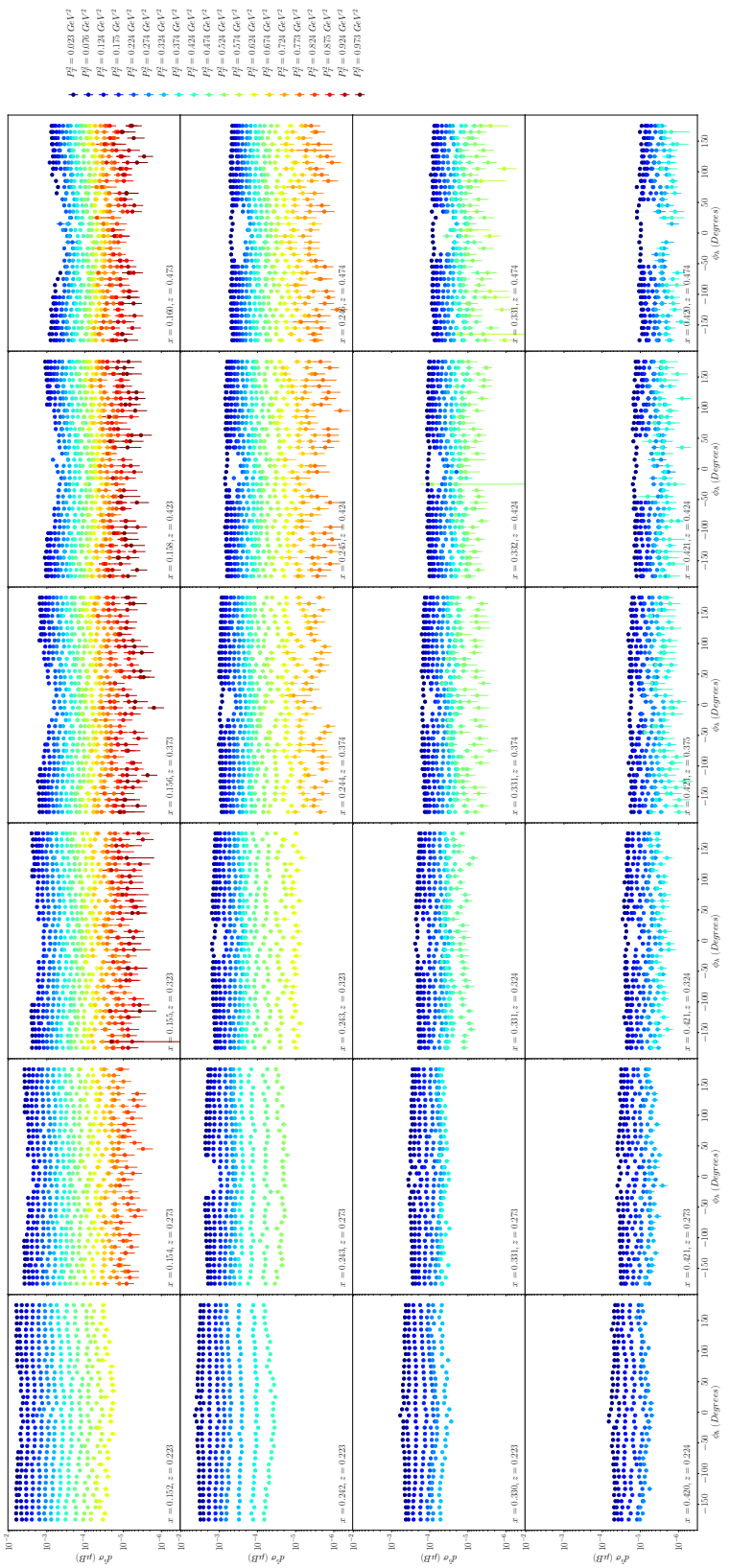

Fig. 6.13: The 5-differential cross section for π^+ shown for different values of x and z in different panels of the figure. In each panel, the P_T^2 dependence of the cross section is shown by superimposing different bins in different colors. The legend at the right of the figure contains information about the colors and their corresponding P_T^2 value.

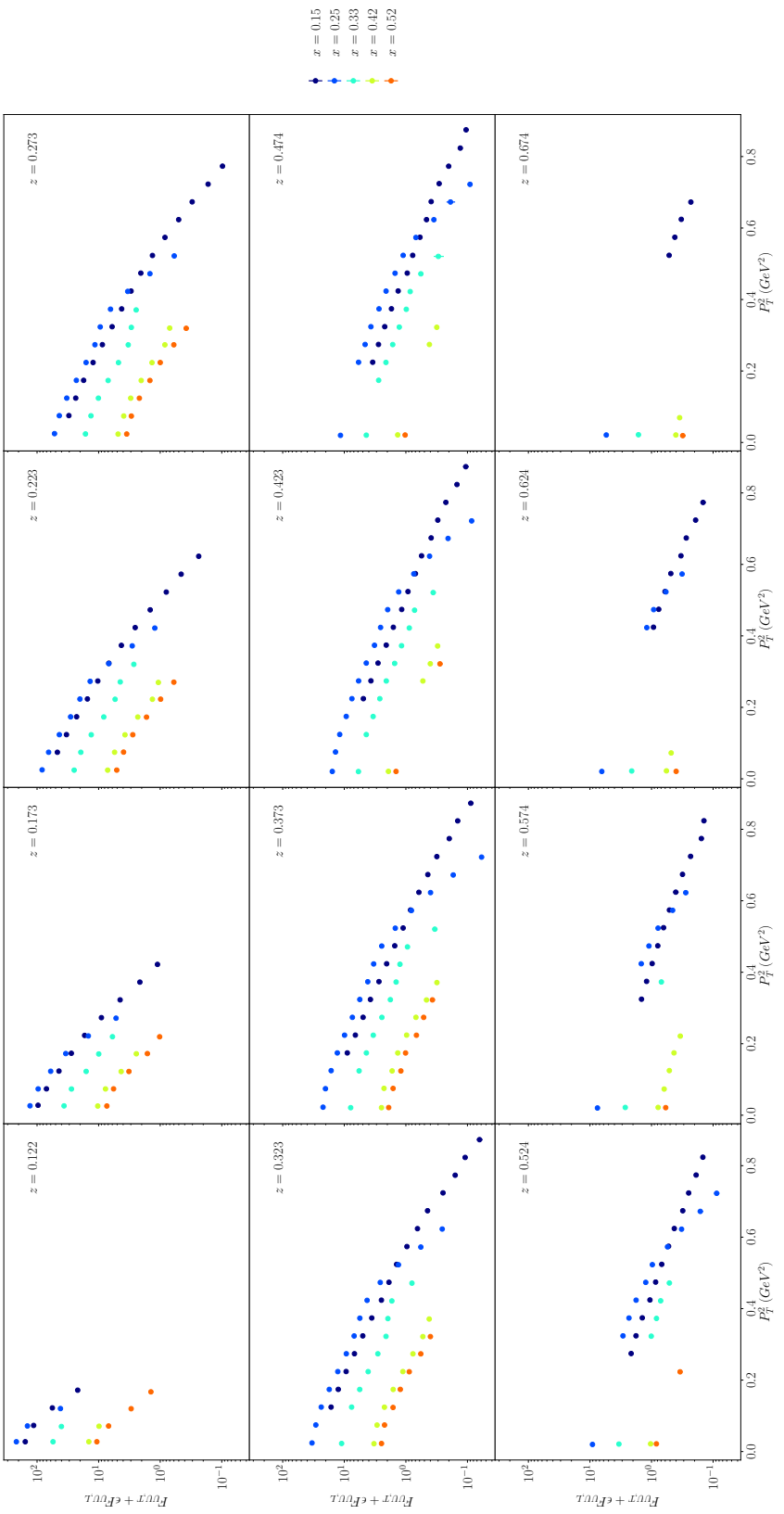

Fig. 6.14: The unpolarized structure functions (π^+) that don't depend on the angle ϕ_h are shown here for different values of x, z, and P_T^2.

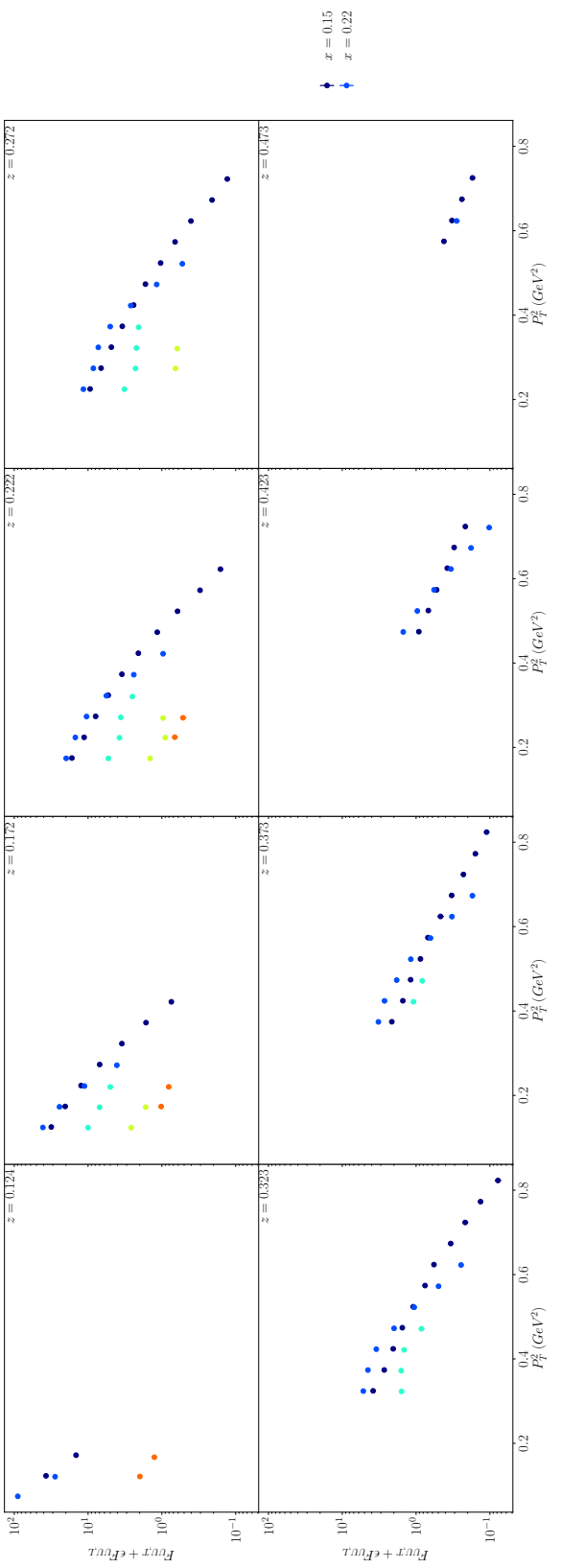

Fig. 6.15: The unpolarized structure functions (π^-) that don't depend on the angle ϕ_h are shown here for different values of x, z, and P_T^2.

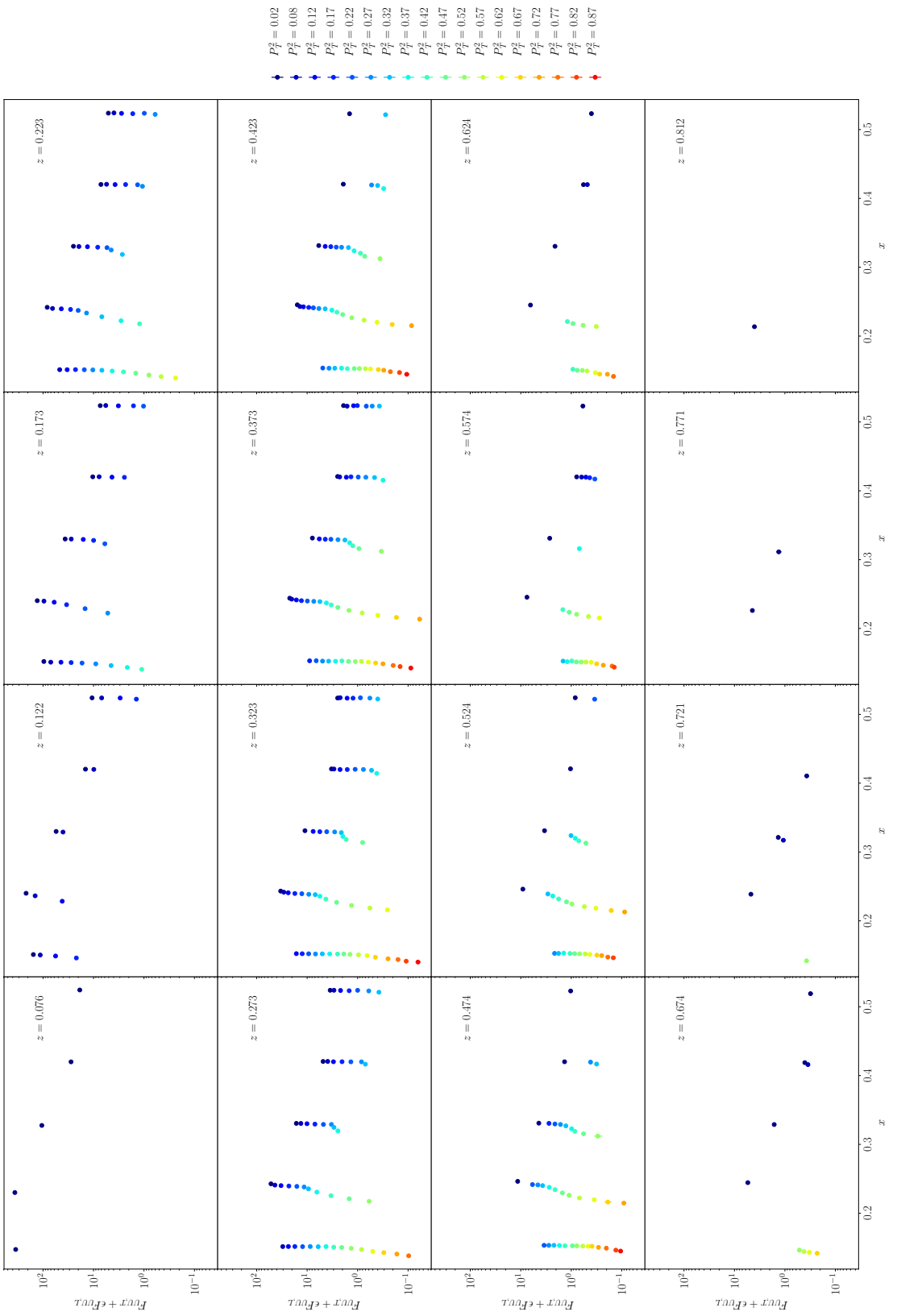

Fig. 6.16: The unpolarized structure functions (π^+) that don't depend on the angle ϕ_h are shown here for different values of x, z, and P_T^2.

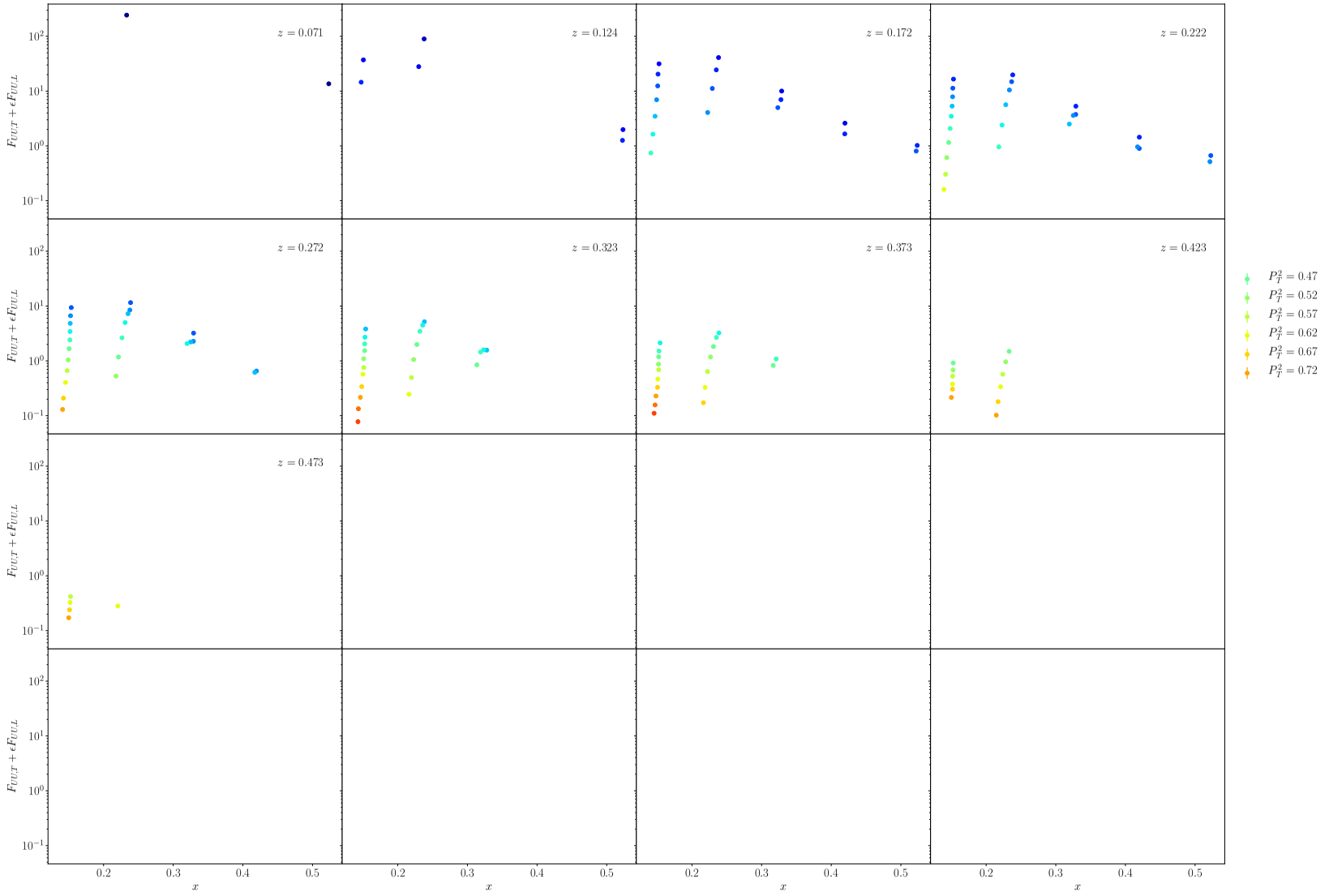

Fig. 6.17: The unpolarized structure functions (π^-) that don't depend on the angle ϕ_h are shown here for different values of x, z, and P_T^2.

Chapter 7

Beam Spin Asymmetry Analysis

7.1 Introduction

Analysis 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.

7.2 Event Selection and Binning

7.2.1 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 $W > 2\ GeV/c^2$ and $Q^2 > 1\ GeV^2/c^2$, because they are not considered part of the deeply inelastic region. Additionally, to avoid exclusive resonances in the $ep \to eK^+X$ spectrum, a minimum value is imposed on the missing mass of the final state M_X $(ep \to eK^+X)$. Here, we use $M_X(ep \to eK^+X) > 1.25\ GeV/c^2$. While it is not yet clear exactly how to constrain kinematics to the current fragmentation region, we remove events with $z < 0.25$ values in an attempt to do so (TMD factorization has been shown at leading order for the current factorization region). It is additionally required that $z < 0.75$ where exclusive events are dominant. 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.

7.2.2 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.

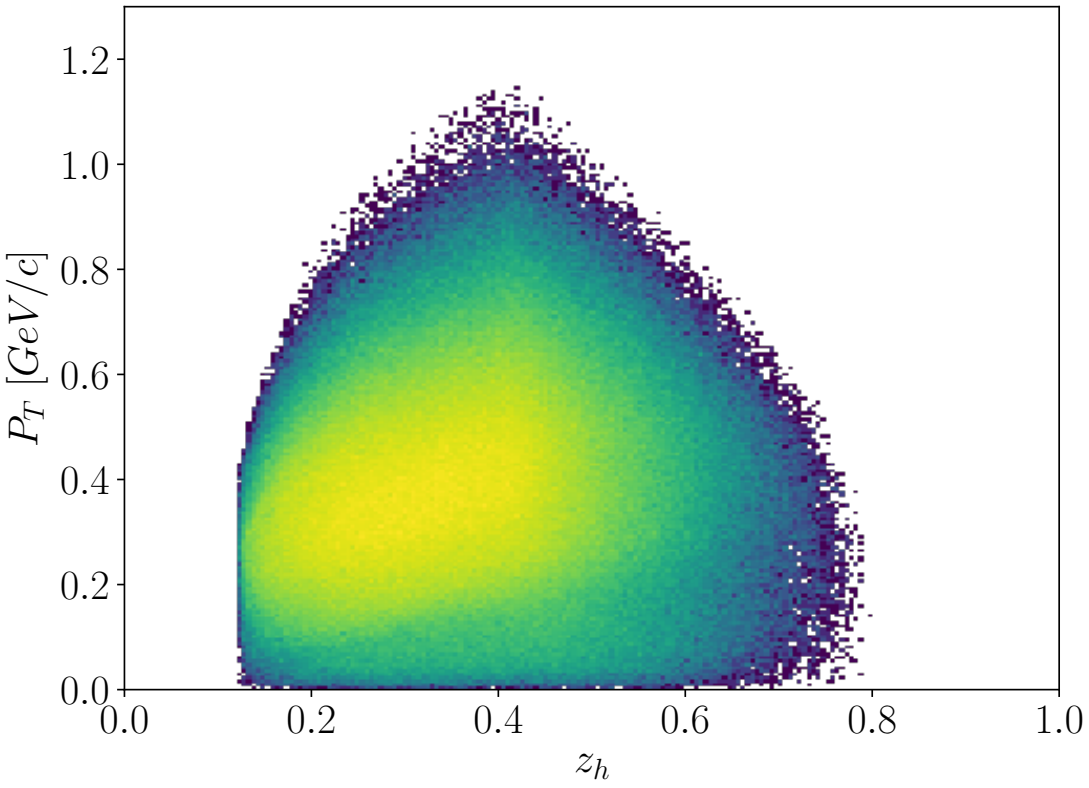

Fig. 7.1: Correlation between z and P_T^2 (GeV^2/c^2) for each event in our analysis sample.

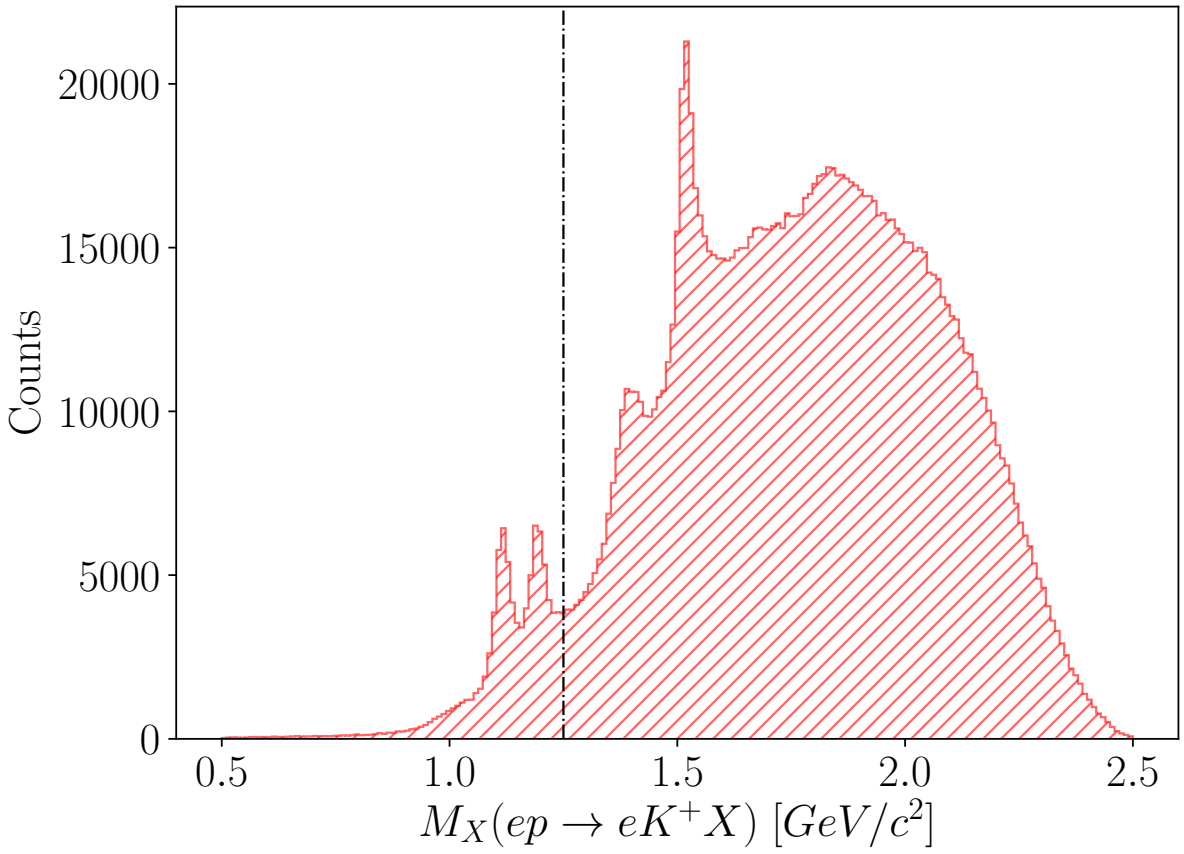

Fig. 7.2: The missing mass spectrum for the reaction $ep \to e'K^+X$ is shown after the application of all cuts used in the analyses except for the cut we apply on the missing mass.

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.

$$\vec{b} = (x_1, x_{N/M}, x_{2N/M}, ..., x_N) \tag{7.1}$$

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.

7.3 Measured ϕ_h Distributions

7.3.1 Measured Asymmetry Values

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

$$A_i = \frac{1}{P_e} \frac{n_+^i - n_-^i}{n_+^i + n_-^i} \tag{7.2}$$

where P_e is the average beam polarization over the dataset (74.9%). The

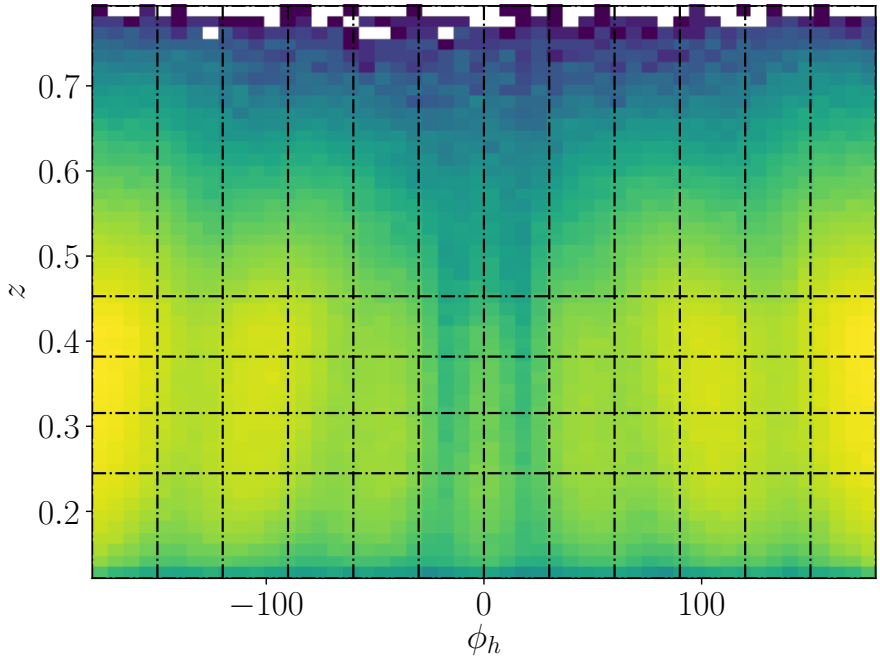

Fig. 7.3: Bins used for this analysis are displayed in two dimensions for the z axis.

symbols $n^i_\pm$ refer to the number of events counted in bin i with helicity $\pm$.

7.3.2 Statistical Uncertainties

The uncertainty on the measured value of A_i can be attributed to statistical uncertainty on the counts $n^\pm_i$, 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).

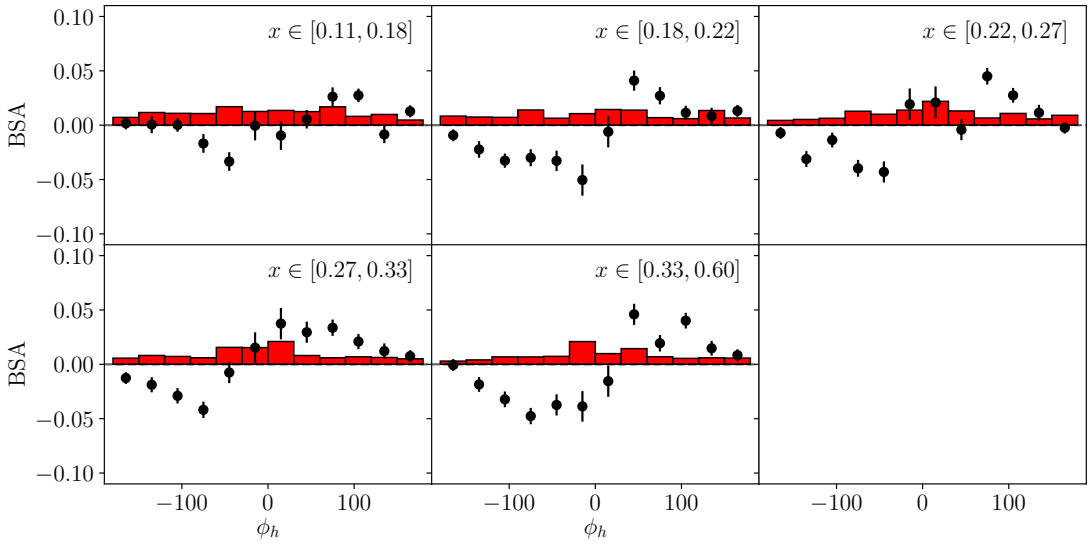

Fig. 7.4: The ϕ_h dependence is shown for each bin of x, increasing in value from the top left to the bottom right. The statistical uncertainty is shown as black error bars on each point. The total systematic uncertainty is shown as a red bar centered at zero.

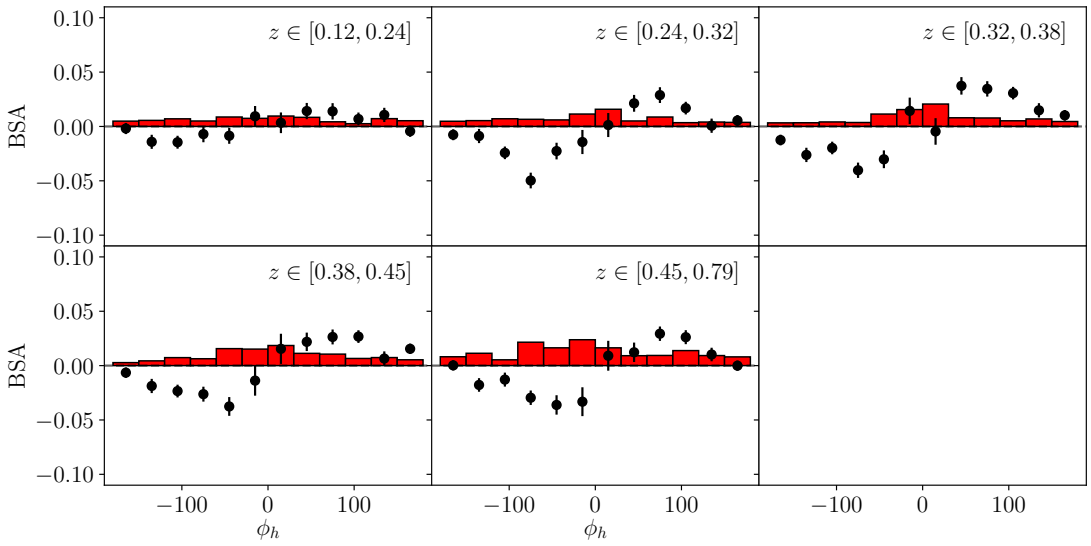

Fig. 7.5: The ϕ_h dependence is shown for each bin of z, increasing in value from the top left to the bottom right. The statistical uncertainty is shown as black error bars on each point. The total systematic uncertainty is shown as a red bar centered at zero.

$$\sigma_{\mathcal{O}}^2 = \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\partial\mathcal{O}}{\partial\theta_i}\frac{\partial\mathcal{O}}{\partial\theta_j}\rho_{ij}\sigma_i\sigma_j \tag{7.3}$$

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.

$$\sigma_A^2 = \frac{A^2}{P_e^2}\sigma_{P_e}^2 + \frac{4(n_-^2\sigma_+^2 + n_+^2\sigma_-^2)}{P_e^2(n_+ + n_-)^4} \tag{7.4}$$

The first term is the contribution from the variance in the measurements of beam polarization and is 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 distributed, 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:

$$\sigma_A^2 = \frac{4n_+n_-}{P_e^2(n_+ + n_-)^3} \tag{7.5}$$

7.3.3 Systematic Uncertainties

7.3.4 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

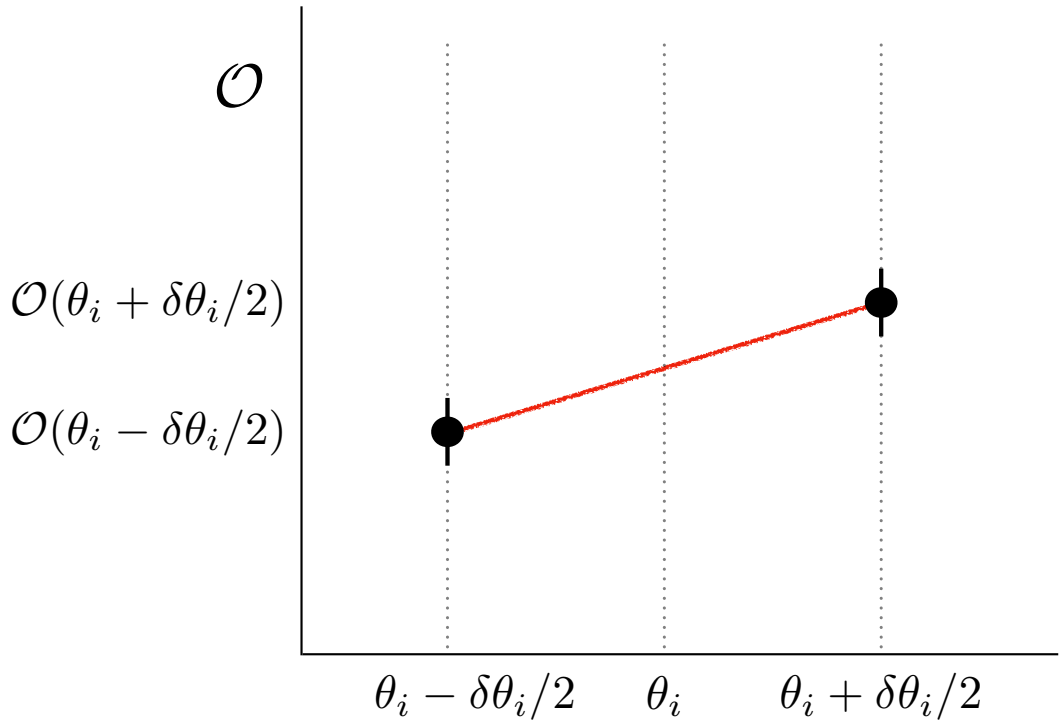

Fig. 7.6: The analysis is run for variations in the input parameters θ_i to calculate the dependence of the result $\mathcal{O}$ on each parameter, as described in this section.

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 [34].

Systematic uncertainties are included using the standard equation for error propagation. For the beam polarization, a few other cases it is possible to analytically calculate the effect of an analysis parameter θ_i on the observable $\mathcal{O}$,

however in many cases it is not. 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.

$$\frac{\partial \mathcal{O}}{\partial \theta_i} \approx \frac{\mathcal{O}(\theta_i + \sigma_{\theta_i}/2) - \mathcal{O}(\theta_i - \sigma_{\theta_i}/2)}{\sigma_{\theta_i}} \tag{7.6}$$

After inserting the above into equation 7.3 one finds,

$$\sigma^2_{\mathcal{O}} = \sum_{i=1}^{n}\sum_{j=1}^{n} \rho_{ij}(\mathcal{O}(\theta_i+\sigma_{\theta_i}/2)-\mathcal{O}(\theta_i-\sigma_{\theta_i}/2))(\mathcal{O}(\theta_j+\sigma_{\theta_j}/2)-\mathcal{O}(\theta_j-\sigma_{\theta_j}/2)) \tag{7.7}$$

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.

$$\sigma^2_{\mathcal{O}} = \sum_{i=1}^{n}\Big[\mathcal{O}(\theta_i + \sigma_{\theta_i}/2) - \mathcal{O}(\theta_i - \sigma_{\theta_i}/2)\Big]^2 \tag{7.8}$$

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

7.3.5 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 are not used in the fitting of structure function ratios A. Table 7.1 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 correc-

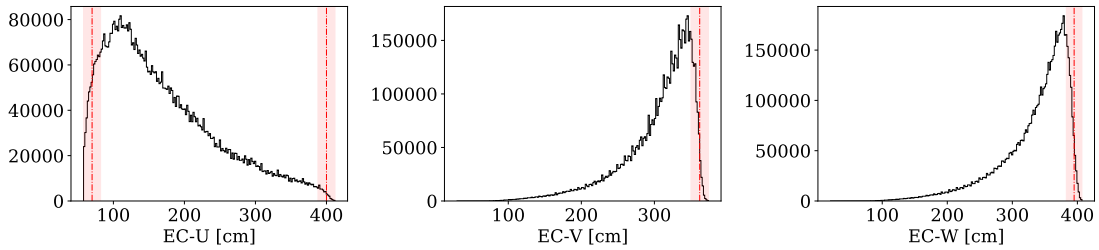

Fig. 7.7: Boundaries (shown in dark red) and associated uncertainties (shaded red regions surrounding the boundaries) for the *EC* coordinate cuts.

tions (0.8%).

7.3.6 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 7.3.6. 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).

7.3.7 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

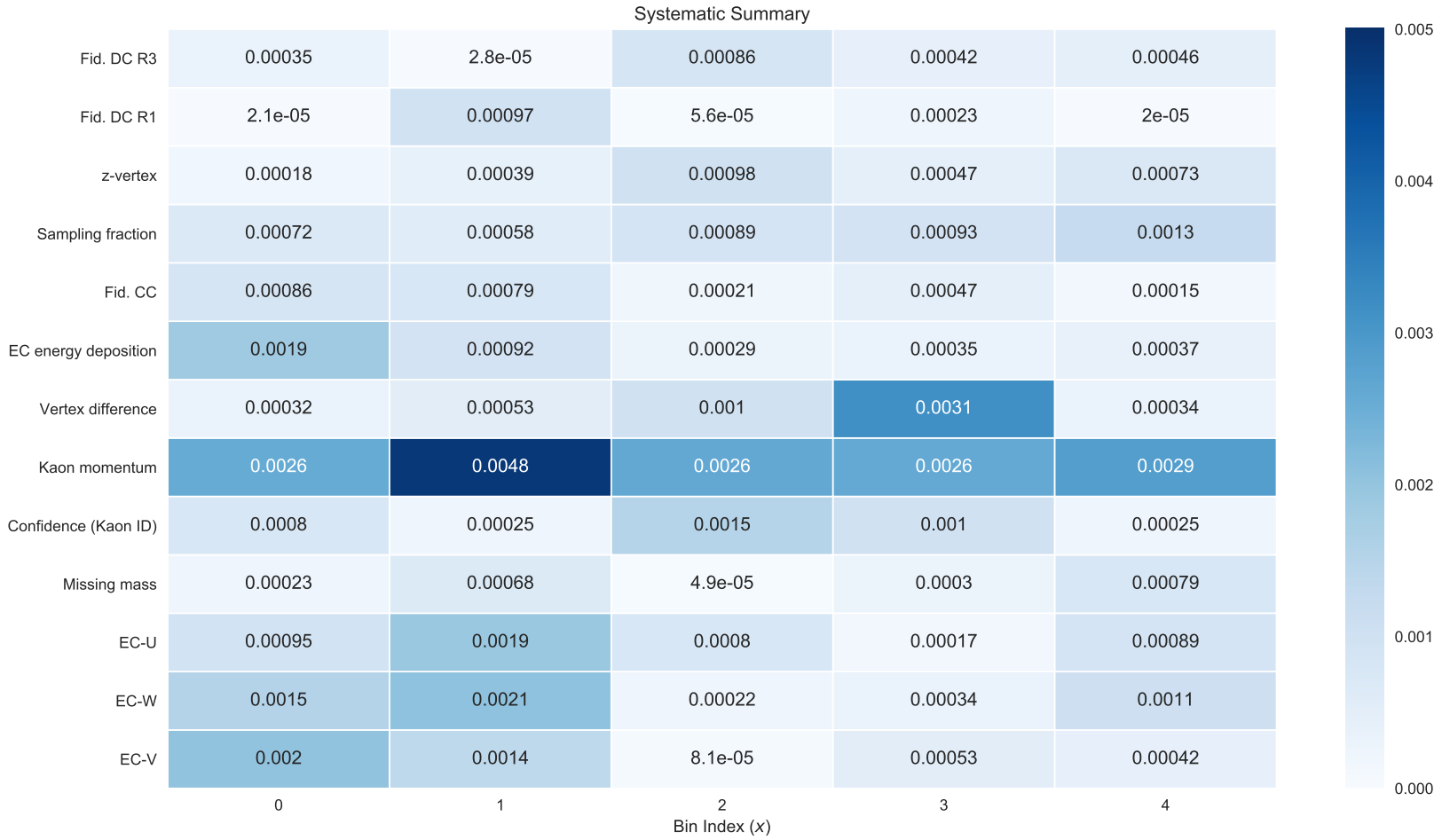

Fig. 7.8: The magnitude of each systematic error source considered is shown above in a heat-map for the x axis. The vertical scale maximum corresponds to half of a percent.

in the figures 7.3.6 7.3.8 7.3.8.

7.3.8 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.

7.3.9 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.

7.3.10 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.

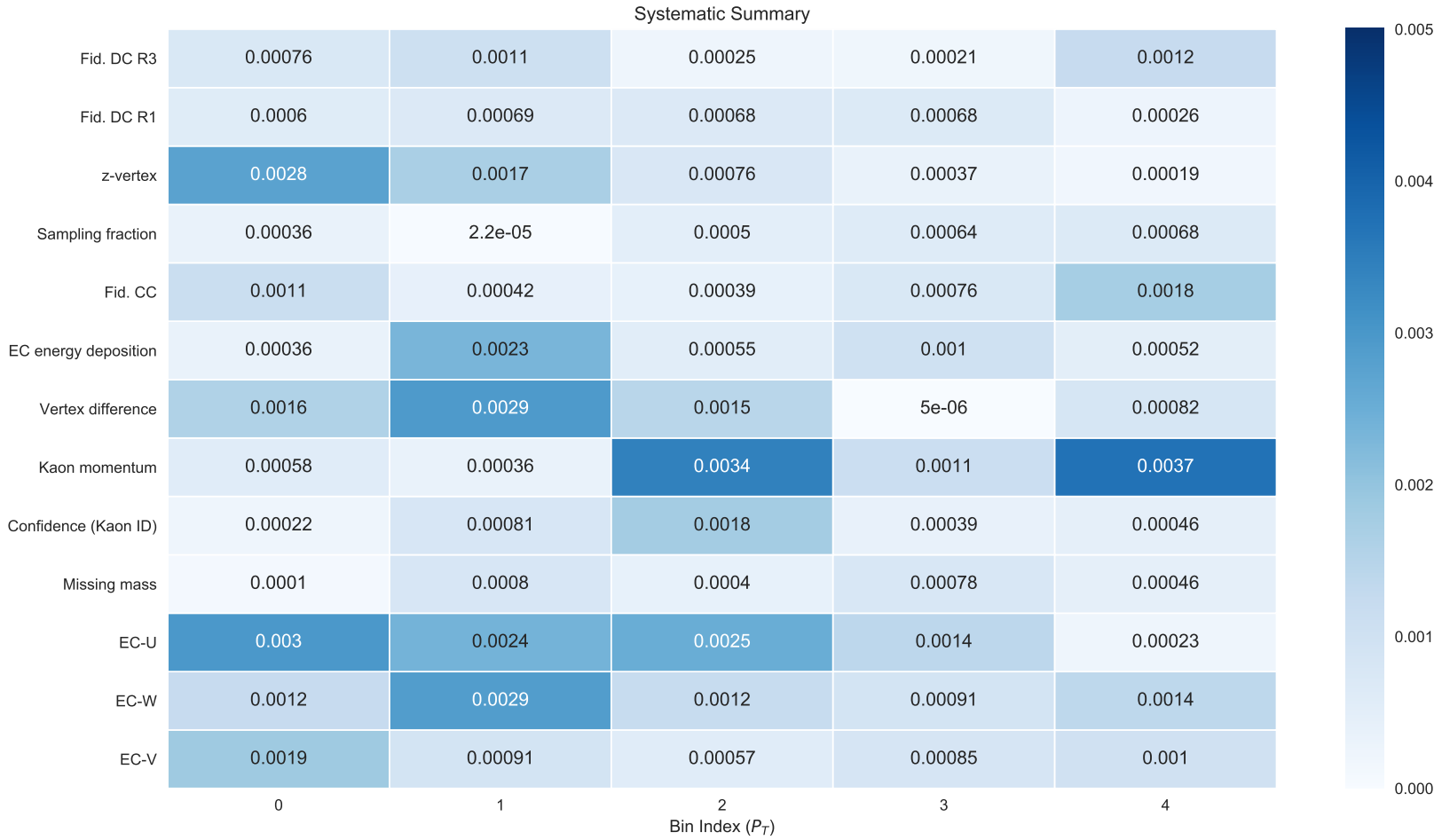

Fig. 7.9: The magnitude of each systematic error source considered is shown above in a heat-map for the P_T axis. The vertical scale maximum corresponds to half of a percent.

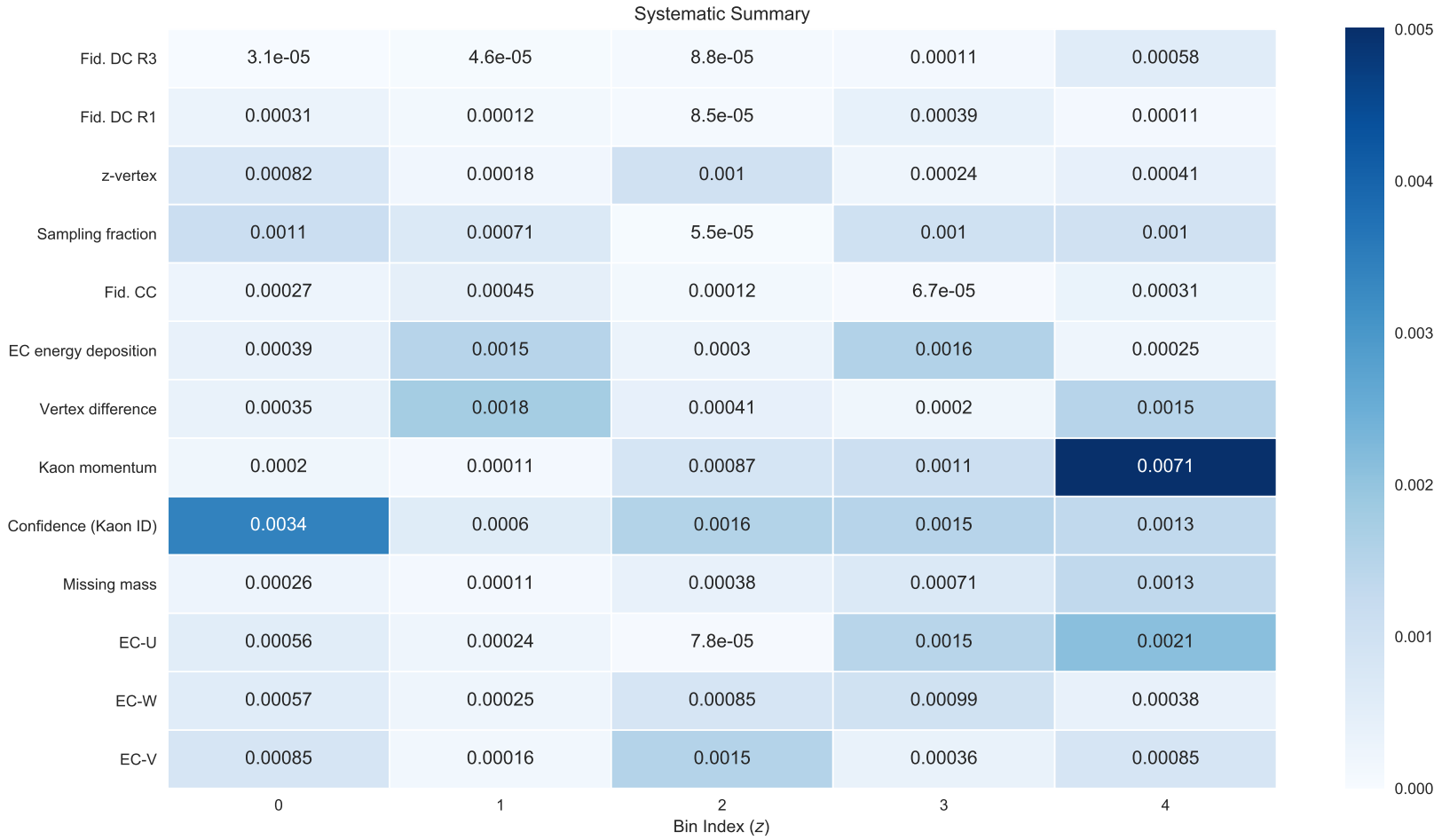

Fig. 7.10: The magnitude of each systematic error source considered is shown above in a heat-map for the z axis. The vertical scale maximum corresponds to half of a percent.

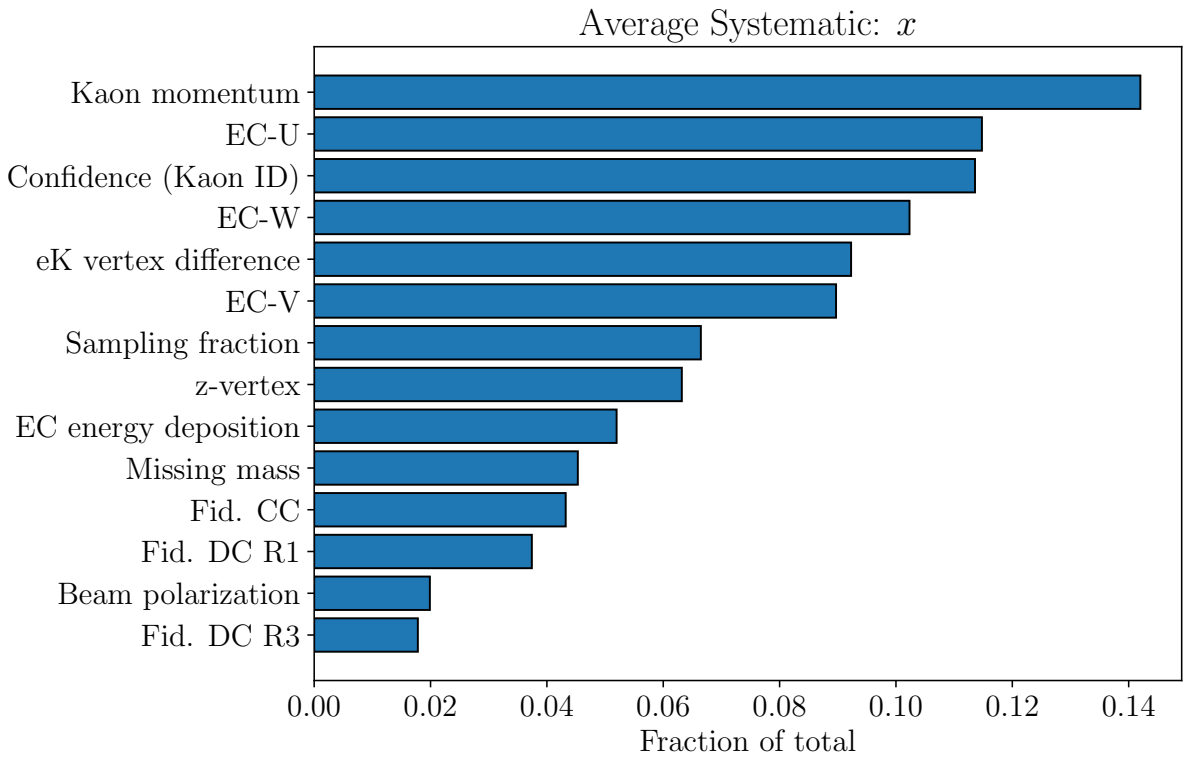

Fig. 7.11: The relative contribution of each systematic uncertainty to the total is shown above averaged over the bins of the x axis.

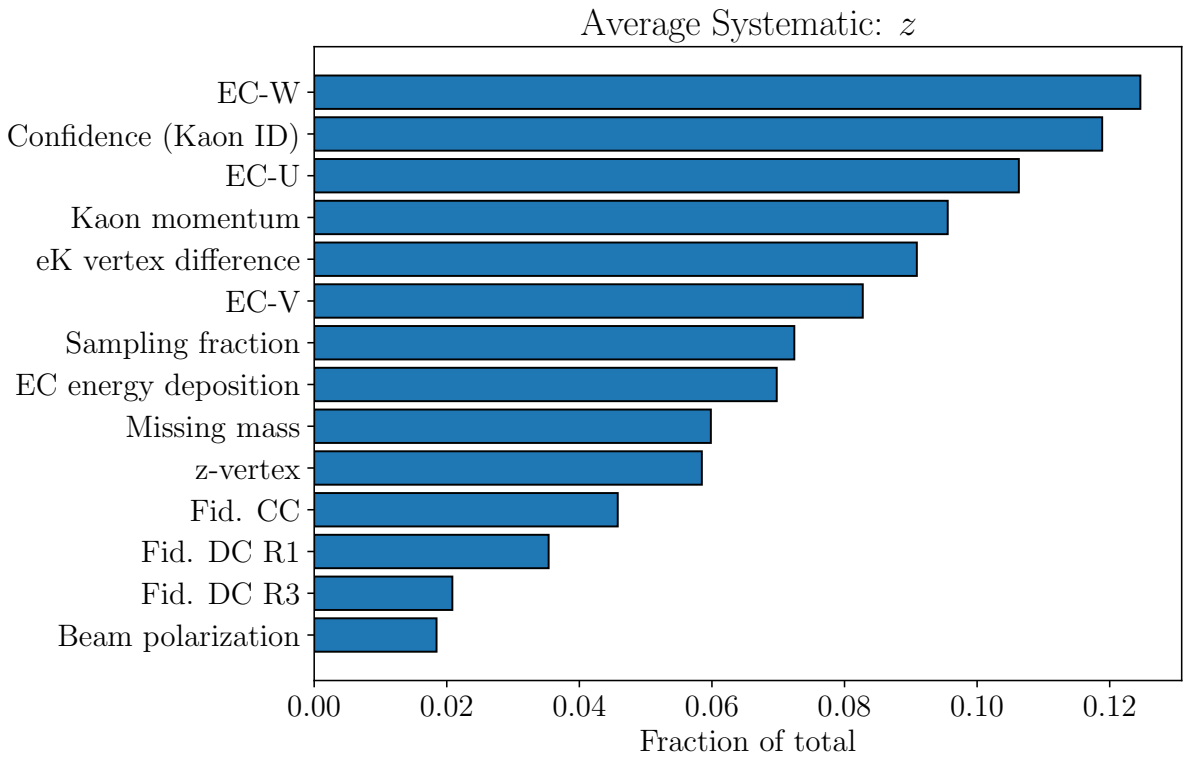

Fig. 7.12: The relative contribution of each systematic uncertainty to the total is shown above averaged over the bins of the z axis.

7.3.11 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.

7.3.12 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.

7.4 Extraction of Modulations

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.

$$f(\phi_h, \vec{a}) = \frac{a_0 \sin \phi_h}{1 + a_1 \cos \phi_h + a_2 \cos(2\phi_h)} \tag{7.9}$$

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

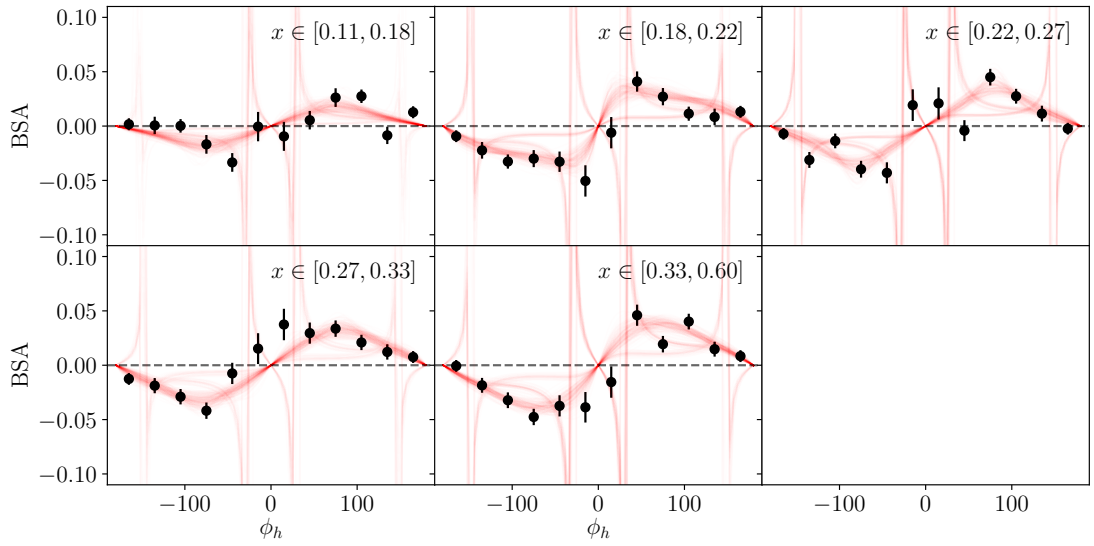

Fig. 7.13: The ϕ_h dependence is shown for each bin of x, increasing in value from the top left to the bottom right. The statistical uncertainty is shown as black error bars on each point. Fits to 256 replicas have been superimposed on the figure.

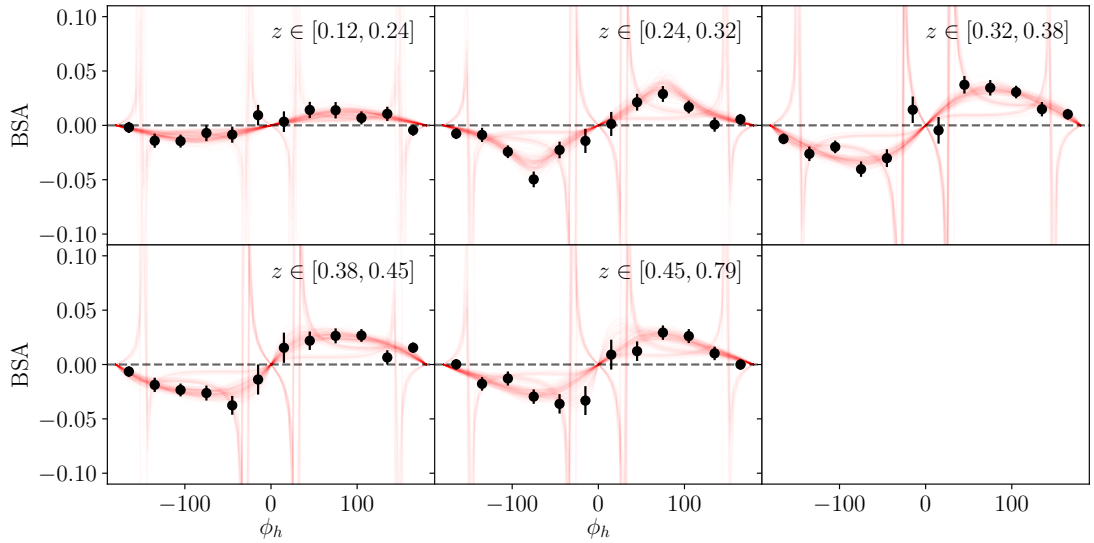

Fig. 7.14: The ϕ_h dependence is shown for each bin of z, increasing in value from the top left to the bottom right. The statistical uncertainty is shown as black error bars on each point. Fits to 256 replicas have been superimposed on the figure.

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-Newtons 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 (additional research into exactly what the problem is would be interesting). 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, and 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 distrubition located at the observed value, and with a standard deviation equal to the statistical errors on the associated data point.

$$\vec{A}_{rep} = \mathcal{N}(\vec{A}, \vec{\sigma_A}) \tag{7.10}$$

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

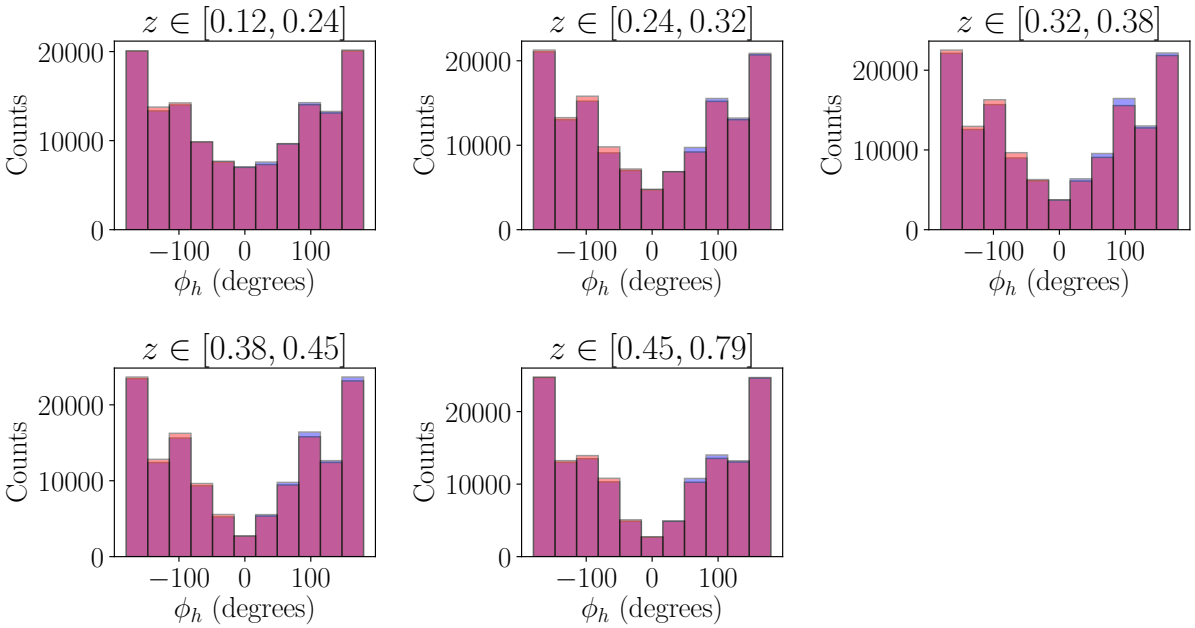

Fig. 7.15: Counts for different helicity states are superimposed for different bins of z.

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 [35].

$$\langle [a_j] \rangle = \frac{1}{N_{rep}} \sum_{i=1}^{N_{rep}} a_j^{(i)} \tag{7.11}$$

$$\sigma_{a_j}^2 = \frac{1}{N_{rep} - 1} \sum_{i=1}^{N_{rep}} (a_j^{(i)} - \langle [a_j] \rangle)^2 \tag{7.12}$$

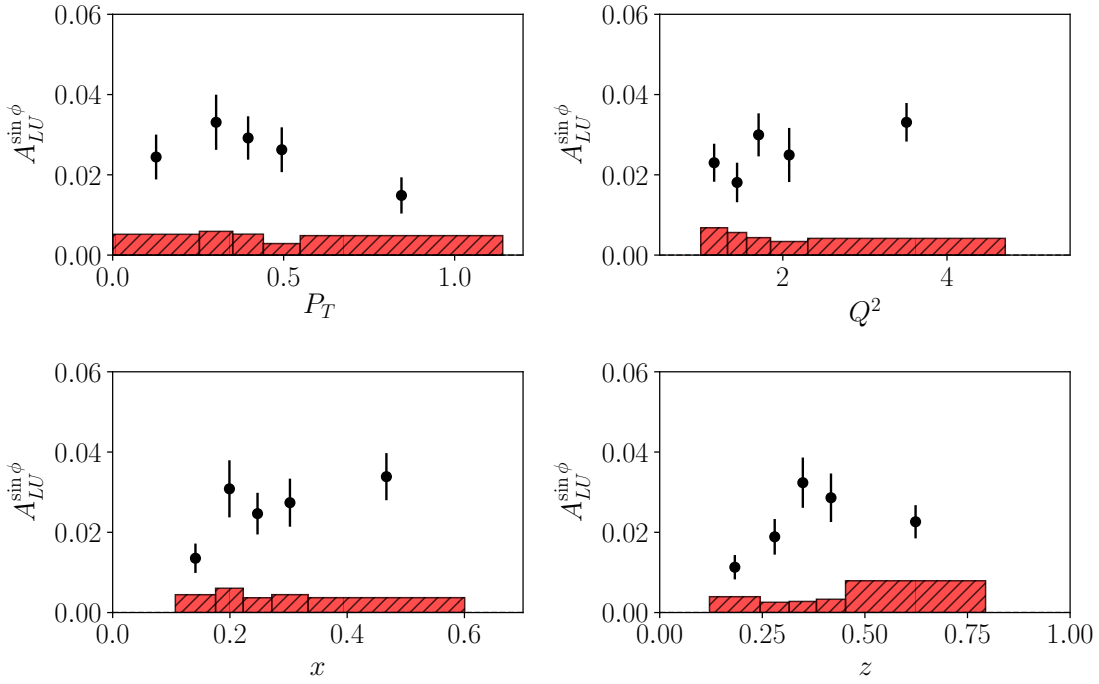

Fig. 7.16: Our extraction of $A_{LU}^{\sin\phi}$ for the kinematic bins described above. The black error bars represent uncertainty in the extraction of the parameter value. Red error bars are systematic uncertainties.

7.5 Results

Our most salient observation is that the magnitude of the BSA is non-zero, implying that twist-three terms are important in our kinematics. As in the case for positive pions, the observed structure function ratio $A_{LU}^{\sin\phi}$ is positive for all kinematic points that were measured. 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 ratio of error to asymmetry value is around 1.5% for most measured points.

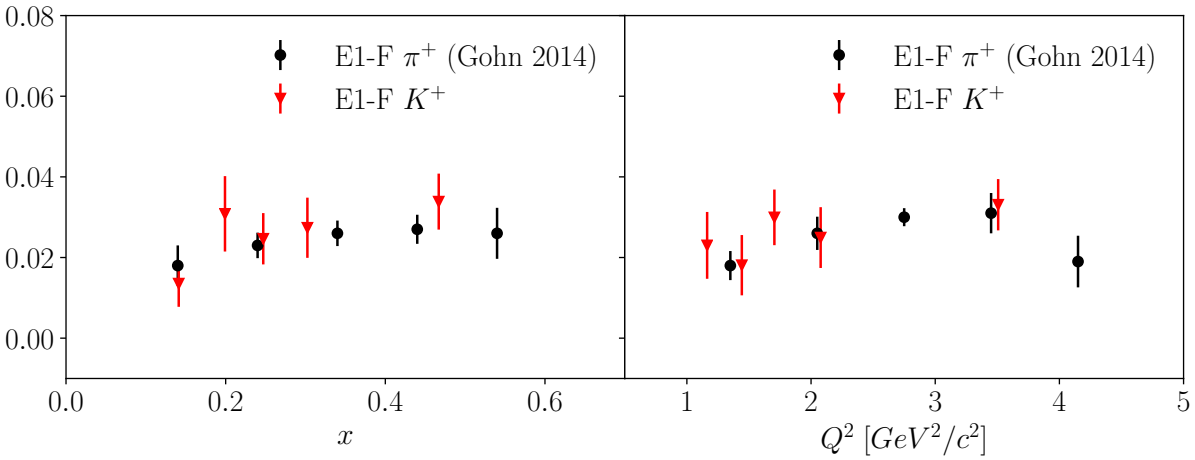

Fig. 7.17: In this figure the results of this study for postively charged kaons are compared with previous results from the same dataset produced by [8] for positively charged pions. This figure shows the x and Q^2 dependence of $A_{LU}^{\sin\phi_h}$.

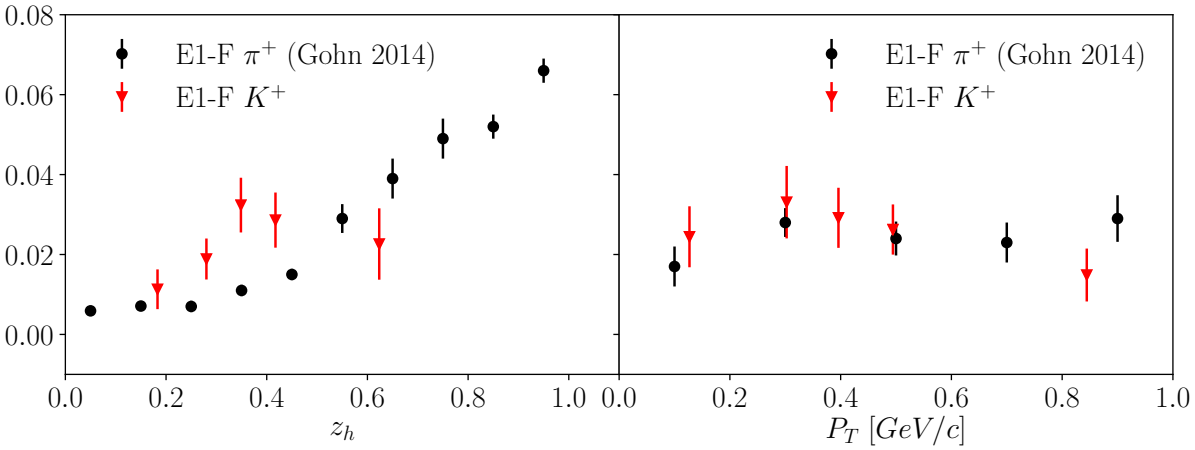

Fig. 7.18: In this figure the results of this study for postively charged kaons are compared with previous results from the same dataset produced by [8] for positively charged pions. This figure shows the z and P_T dependence of $A_{LU}^{\sin\phi_h}$.

References

[1] M. Anselmino, M. Boglione, J. O. Gonzalez Hernandez, S. Melis, and A. Prokudin, JHEP **04**, 005 (2014), arXiv:1312.6261 [hep-ph] .

[2] H.-W. Lin, W. Melnitchouk, A. Prokudin, N. Sato, and H. Shows, Phys. Rev. Lett. **120**, 152502 (2018), arXiv:1710.09858 [hep-ph] .

[3] A. Airapetian *et al.* (HERMES), Phys. Rev. Lett. **103**, 152002 (2009), arXiv:0906.3918 [hep-ex] .

[4] V. Barone, S. Melis, and A. Prokudin, Phys. Rev. **D81**, 114026 (2010), arXiv:0912.5194 [hep-ph] .

[5] G. Adams *et al.*, Nucl. Instrum. Meth. **A465**, 414 (2001).

[6] M. D. Mestayer *et al.*, Nucl. Instrum. Meth. **A449**, 81 (2000).

[7] M. Amarian *et al.*, Nucl. Instrum. Meth. **A460**, 239 (2001).

[8] W. Gohn *et al.* (CLAS), Phys. Rev. **D89**, 072011 (2014), arXiv:1402.4097 [hep-ex] .

[9] R. Hofstadter and R. W. McAllister, Phys. Rev. **98**, 217 (1955).

[10] R. Hofstadter, Rev. Mod. Phys. **28**, 214 (1956).

[11] E. E. Chambers and R. Hofstadter, Phys. Rev. **103**, 1454 (1956).

[12] G. Zweig, (1964).

[13] J. D. Bjorken and E. A. Paschos, Phys. Rev. **185**, 1975 (1969).

[14] R. P. Feynman, Phys. Rev. Lett. **23**, 1415 (1969), [,494(1969)].

[15] P. J. Mulders and R. D. Tangerman, Nucl. Phys. **B461**, 197 (1996), [Erratum: Nucl. Phys.B484,538(1997)], arXiv:hep-ph/9510301 [hep-ph] .

[16] A. Bacchetta, M. Diehl, K. Goeke, A. Metz, P. J. Mulders, and M. Schlegel, JHEP **02**, 093 (2007), arXiv:hep-ph/0611265 [hep-ph] .

[17] J. Collins, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **32**, 1 (2011).

[18] A. Airapetian *et al.* (HERMES), Phys. Rev. **D87**, 074029 (2013), arXiv:1212.5407 [hep-ex] .

[19] M. Aghasyan *et al.* (COMPASS), Phys. Rev. **D97**, 032006 (2018), arXiv:1709.07374 [hep-ex] .

[20] A. Airapetian *et al.* (HERMES), Phys. Lett. **B693**, 11 (2010), arXiv:1006.4221 [hep-ex] .

[21] C. Adolph *et al.* (COMPASS), Phys. Lett. **B744**, 250 (2015), arXiv:1408.4405 [hep-ex] .

[22] X. Qian *et al.* (Jefferson Lab Hall A), Phys. Rev. Lett. **107**, 072003 (2011), arXiv:1106.0363 [nucl-ex] .

[23] J. C. Collins, Phys. Lett. **B536**, 43 (2002), arXiv:hep-ph/0204004 [hep-ph] .

[24] N. A. Harrison, *Exploring the Structure of the Proton via Semi-Inclusive Pion Electroproduction*, Ph.D. thesis, Connecticut U. (2010).

[25] H. Avakian *et al.* (CLAS), Phys. Rev. **D69**, 112004 (2004), arXiv:hep-ex/0301005 [hep-ex] .

[26] W. Gohn, *Probing the Proton's Quark Dynamics in Semi-inclusive Pion Electroproduction*, Ph.D. thesis, Connecticut U. (2012).

[27] C. W. Leemann *et al.*, Ann. Rev. Nucl. Part. Sci. **51**, 413 (2001).

[28] Y. C. Chao *et al.*, J. Phys. Conf. Ser. **299**, 012015 (2011).

[29] E. S. Smith *et al.*, Nucl. Instrum. Meth. **A432**, 265 (1999).

[30] M. Mirazita, "Kinematic corrections for e1-f," .

[31] R.-G. Ping *et al.*, Int. J. Mod. Phys. **A24S1**, 23 (2009).

[32] C. Keppel, *Inclusive nucleon resonance electroproduction at large momentum transfer*, Ph.D. thesis, American U. (1994).

[33] L. W. Mo and Y.-S. Tsai, Rev. Mod. Phys. **41**, 205 (1969).

[34] 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] .

[35] G. Watt and R. S. Thorne, JHEP **08**, 052 (2012), arXiv:1205.4024 [hep-ph] .