Chapter 4

Event Reconstruction

Events from the ZEUS detector which are triggered at the FLT, SLT and TLT are written to tape for storage. This analysis begins with the EVB (see section 3.3.5.2), which combined all detector signals from the event into a single ADAMO table entry. Entries were then processed by the ZEPHYR program, which used calibration constant tables for each component and each run to interpret detector signals as energies, times, and positions. Offline analysis in this work was performed using the ORANGE software framework, which is based on the EAZE framework [?]. ORANGE and EAZE provide routines which retrieve data records, perform specialized and modifiable reconstruction of event variables, integrate the CERN libraries, and perform basic event selections.

4.1 Track and Vertex Reconstruction

While the nominal interaction point for ep interactions is at the origin of the ZEUS coordinate system, finite dimensions of the beam bunches caused real interactions to occur with a spread in all three dimensions. Information from the tracking components of the ZEUS detector were used to reconstruct the tracks of individual

spontaneously. The behavior of each PMT was monitored during operation and calibrations to ensure proper functioning, PMTs listed in a table of unreliable PMTs for each run were ignored during reconstruction. To supress uranium radiation in reconstruction, EMC cells with $E < 80$ MeV and HAC cells with $E < 140$ MeV were neglected. Cells were also ignored if the imalance between left and right PMTs, $I_{\text{cell}} = \left|(E_{\text{L}} - E_{\text{R}}) / (E_{\text{L}} + E_{\text{R}})\right| > 0.7$, and that cell had an energy greater than 700MeV. Cell selection was performed in this analysis by the Noise04s routine [?].

Simulations of the energy response of the CAL differ from that of the true response [23]. While this could be corrected by adjusting the energy values of the simulation, historically this is performed by altering the final CAL cell energies in the ZEUS data. For each FEMC cell, the energy was muliplied by a factor of 1.024. The same procedure was performed using factors of 0.941 for FHAC, 1.053 for BEMC, 1.096 for BHAC, and 1.022 for all RCAL cells, using the Escale03 routine. [?].

4.3 Electron Reconstruction

A major source of background to CC DIS is NC DIS, which has a pronciple signal as an isolated electron deposit located in the calorimeter. In order to reject events arising from NC interactions, the electron finder SINISTRA95 [24] was applied to all events. SINISTRA is a neural network which has been trained on simulated low-Q^2 NC DIS data to produce the best separation betweeu electromagnetic and hadrouic energy deposits by using the properties of shower profiles to establish the characteristics of the initiating particle.

Scattered particles snch as single electrons often deposit their energy into more than one calorimeter cell, and often in ways which involve cells in both the BCAL and

RCAL. Before input into SINISTRA, groups of calorimeter cells called *islands* [24] were formed, where each island is a potential deposit from an electron. The 17 inputs to the neural network were the total energy of the cells in the island, and 16 showering profile parameters called Zernike moments and Legendre functions. The output of SINISTRA is a real number between 0 and 1, which is interpreted as a liklihood that the island in question resulted from an isolated electron. Clusters with SINISTRA output near one are considered to be electron-like. In this analysis, only the candidate with greatest SINISTRA liklihood was considered, and events were only considered to have a candidate if the liklihood was greater than 0.9. The reconstructed energy of an electron canididate was the sum of the cell energies in its island, and the position of the candidate was determined from the energy-weighted island center. For candidates with greater than 0.9 liklihood and energy greater than 10GeV, SINISTRA was estimated to have 80% purity and nearly 100% efficiency [25].

4.4 Kinematic Reconstruction

Basic kinematic quantites were reconstructed from the calorimeter energy deposits, which were treated as massless particles.

$$\text{Total energy} \quad : \quad E_{\text{CAL}} = \sum_i E_i \tag{4.1}$$

$$\text{Longitudinal momentum in the CAL} \quad : \quad p_z = \sum_i p_{z,i} \tag{4.2}$$

$$\text{Transverse momentum in the CAL} \quad : \quad p_{T,\text{CAL}} = \sqrt{\left(\sum_i p_{x,i}\right)^2 + \left(\sum_i p_{x,i}\right)^2} \tag{4.3}$$

$$\text{Cosine of hadronic system angle} \quad : \quad cos\gamma_h = \frac{p_T^2 - (E - p_z)^2}{p_T^2 + (E - p_z)^2} \tag{4.4}$$

In charged current interactions at ZEUS, the exiting neutrino escapes detection, requiring kinematic quantities to be reconstructed from the hadronic system alone. To prevent bias in our estimation of these values due to particles which escape detection down the beam pipe, the Jaqnet-Blondel method [26] of reconstrnction was used. This method relies only on the transverse momenta, p_T and difference in cnergy and z-component momenta, $E - p_z$ of each particle, which are small for particles scattered near the beamline. The estimators for Q^2, x, and y are

$$\begin{aligned} y_{\mathrm{JBCAL}} &= \frac{E - p_z}{2E_e} \\ Q^2_{\mathrm{JBCAL}} &= \frac{p_{T,\mathrm{CAL}}}{1 - y_{\mathrm{JBCAL}}} \\ x_{\mathrm{JBCAL}} &= \frac{p_T}{s y_{\mathrm{JBCAL}}(1 - y_{\mathrm{JBCAL}})} \end{aligned} \tag{4.5}$$

4.5 Dead Material and Transverse Energy Corrections

In NC DIS, an ideal calorimeter wonld report zero transverse inomentum, and equal transverse energy for the electron and the hadronic system. Alternative methods for measuring y from the electron and y_{JB} would also be equal. To improve the description of the data by the simulation, a sample of NC DIS events was selected and the hadronic-electromagnetic response was compared to simulation. Linear correction functions in bins of γ_h were applied to $p_{T,\mathrm{CAL}}$ and y_{JBCAL}. The corrected values of $p_{T,\mathrm{CAL}}$ and y_{JBCAL} were used to reconstruct Q^2_{JBCAL} and x_{JBCAL}.

The Interior of ZEUS detector contained several components which are not designed to detect particles, such as the beam pipe and solenoid. As particles traversed these objects, they lost energy which is not recorded, smearing the aforementioned estimators in a way which is difficult to directly measure. To reconstruct the kinematic

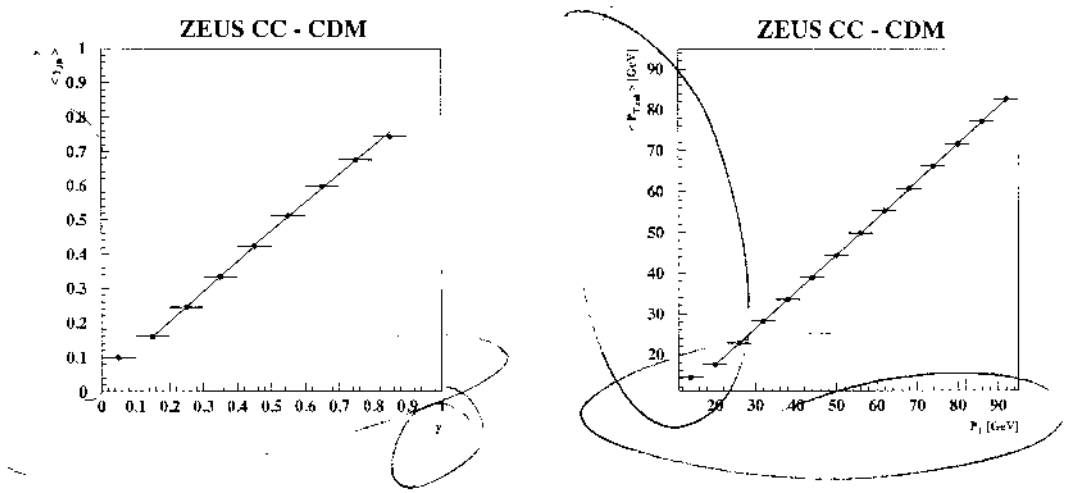

Figure 4.1: Dead Material Correction Functions for $y_{\rm JB}$ And p_T. The x-coordinate values represents values from simulated events from hadrons, and the y-coordinate values represent the calorimeter-extracted values. The method used for producing these plots is described in section 5.5

variables of hadrons for the event, simulations were used to predict linear correction functions. Functions chosen in bins of γ_h and $p_{T,\rm CAL}$ were applied to $p_{T,\rm CAL}$. Functions in bins of γ_h and $y_{\rm JB CAL}$ were applied to correct $y_{\rm JB CAL}$. These values were considered to represent the kinematics of the event after stable particles were produced, but before particles interacted with the ZEUS detector. This is referred to as *hadron level* information. The corrected values of p_T and $y_{\rm JB}$ were used to reconstruct $Q^2_{\rm JB}$ and $x_{\rm JB}$

4.6 Jet Reconstruction and Corrections

Calorimeter cells were combined using the k_T cluster algorithm [27] to form jets of calorimeter cells. The algorithm was also used to cluster jets of hadrons and partons

in the simulations, so in this section we will refer only to "objects" being clustered for generality. All clustering objects were considered massless. The algorithm defines the distance of each particle from the beam line as $d_i = E_{T,i}^2$, and the distance between two objects as $d_{ij} = \min\left(E_{T,i}^2, E_{T,j}^2\right) \times \left[(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2\right]$. The algorithm begins by identifying $\min(d_i, d_{ij})$. If $d_{kl} = \min(d_i, d_{ij})$ for some k, l, then objects k, and l are merged into a cluster (here indexed as m) with properties

$$\begin{aligned} E_{T,m} &= E_{T,j} + E_{T,k} \\ \eta_m &= \frac{E_{T,j}\eta_j + E_{T,k}\eta_k}{E_{T,j} + E_{T,k}} \\ \phi_m &= \frac{E_{T,j}\phi_j + E_{T,k}\phi_k}{E_{T,j} + E_{T,k}}. \end{aligned} \tag{4.6}$$

If $d_n = \min(d_i, d_{ij})$ for some n, then object n is considered a jet, and no longer participates in the merging procedure. In this analysis, the k_T algorithm was applied in the lab frame in inclusive mode, meaning that all objects are merged into jets. Phase space requirements were applied to each jet, neglecting objects outside acceptance of with small E_T, which is described in section 6.5.

Similarly to correction of detector effects for the kinematic variables, detector jet variables were also corrected. Resolution studies (shown in section 5.5) showed that η^{jet} and ϕ^{jet} shoed no systematic smearing from detector effects, and no corrections were applied. Linear correction functions for dead material were prepared from simulations. Functions in bins of calorimeter *etjet* and *rapjet* were applied to jets of calorimeter cells to produce jet E_T from jets of hadrons.

4.7 Jet Energy Scale ~~Uncertainties~~ Corrections

In order to improve the description of hadronic jets, additional correction factors were applied to the jets in a similar fashion as the transverse energy corrections described in section 4.5. A sample of NC DIS was extracted where each event was required to have single a jet. Because the jet would ideally have equal E_T to the scattered electron in this sample, the E_T of the electron was used to calibrate the CAL response to each jet. This was extracted in terms of linear functions of E_T in bins of η^{jet}, which were then applied to the hadronic jet E_Ts in the CC sample. Reference

Chapter 5

Event Simulation

pQCD is not presently calculable to all orders in α_s. Interpretation of measurements at HEP experiments therefore requires the simulation of higher order effects. which is often achieved by the Monte Carlo (MC) method. Here the term MC is exclusively used to refer to the generation and complete simulation of discrete ep events in the detector. This begins with probability distributions for the desired hard-scale processes. which are chosen by a pseudorandom[1] number generator. The outgoing state ~~with have~~ free partons which become parton showers by radiating gluons and spliting into $q\bar{q}$ pairs, which become hadrous via a phenomenological model. This results in a list of observable, on-shell particles with definite momenta, who's detector response is simulated. The final output is a table containing the initial process, partous, on-shell particles before detector interaction, and a list of detector component simulated responses, which should closely resemble real world data.

[1] A pseudorandom number generator is a deterministic algorithm for the production of numerical sequences which satisfy some statistical tests of randomness, such as uniform frequency and distribution of the numbers produced.

5.1 PDF and Hard Scatter

Initially, the incoming state is an electron and proton, which will interact via the protons partons. Which flavor parton will interact, and what the x and Q^2 value of the interaction is determined by the proton PDFs and a pseudorandom number generator. The PDFs in this analysis come from the Coordinated Theoretical-Experimental Project on QCD (CTEQ5D) [28], and Martin Roberts Stirling Thorne (MRST)[29], and The ZEUS Collaboration. The relative probability of a subclass of physical processes is computed to fixed order, which are stored in tables for efficient recall. This component is referred to as the *hard scatter* to differentiate it from QCD effects at lower scales. numbers are generated, and used to select from these processes.

5.2 Parton Cascade

Two methods of simulating a parton shower will be considered here: Matrix-Element plus Parton Shower (MEPS) [30], and the Color Dipole Model (CDM) [31]. The MEPS approach generates parton showers from the DGLAP splitting functions. Partons become less virtual as they radiate and split until an arbitrary virtuality is reached, typically on the order of 1 GeV. Because the parton shower uses the DGLAP evolution equations, radiation is ordered strongly in decreasing k_T and increasing proton fractional momenta. Samples of LEPTO MC presented in this thesis employ MEPS as implimented via JETSET 7.4 [30, 32]. CDM describes the proton remnant and its dissociated parton as a color dipole. As the parton and remnant separates, the dipole energy increases until sufficient energy is reached to radiate a gluon, which bifurcates the dipole. The two halves continue to stretch and split as more gluons

are emitted. Unlike MEPS, CDM does not order its radiation in k_T. Samples of ARIADNE 4.08 [33] MC presented in this thesis employ CDM.

5.3 Hadronization

Free partons are never observed due to color confinement, so the partons simulated during the cascade phase discussed in the last section must be converted into color singlet hadrons in order to describe physical data. This process is referred to as hadronization, and because it descibes confinement, at scales where α_s is large, it necessarily cannot be described by perturbative QCD. Several hadronization models are currently implimented in MC models, but only the Lund String Model[34] is used in the MC results presented in this thesis.

In the Lund String Model, the color field between a $q\bar{q}$ pair is represented as a linear potential, described as a flux-tube or string of gluons. Spatial stretching of this string results in a potential energy of $\approx$ 1GeV/fm. If the initial $q\bar{q}$ pair has sufficient energy, they will separate and the color field will contain enough potential energy to produce a new $q\bar{q}$ pair. This is described in the model as occurring between the initial $q\bar{q}$ pair, as if the string had been cut. This mechanism produces quarks with an approximately gaussian p_T spectrum. This process continues until some preset cutoff, when the $q\bar{q}$ pairs have small enough relative kinetic energy to be considered on-shell hadrons. Groups of 2 and 3 quarks are then chosen to form mesons and hadrons.

5.4 Description of Initial- and Final-State QED Radiation

While MC models can generate interactions at fixed electron and proton energies, interactions in real world data are not produced with this kind of kinematic

spectrum, however, because of radiation of photons immediately before and after the hard interaction. This type of process is referred to as initial- and final-state radiation (ISR/FSR). For this analysis, the MC implimentations used for event simulation were interfaced with the HERACLES 4.6.1 [35] program via DJANGO v1.3[36]. The HERACLES program includes QED effects up to $\mathcal{O}(\alpha_{\rm EM}^2)$. Physical parameters used to compute QED effects were

$$\begin{aligned} \alpha_{\rm EM} &: 1/137.035999 \\ G_{\rm F} &: 1.664 \times 10^{-5}\ {\rm GeV}^{-2} \\ M_Z &: 91.1876\ {\rm GeV} \end{aligned} \tag{5.1}$$

5.5 Detector Simulation and Reconstruction

Once an event has been simulated to produce real particles with definite momenta, these particles are used as input into a full detector simulation. This simulation includes potential decays of short lived particles, the interactions with the active and passive materials in the detector, and the response of the electronics at all levels in the data-taking and reconstruction chain. Information for the modeling of the detector response comes from test-beam data, and the simulation is referred to as the *Monte Carlo for ZEUS Analysis, Reconstruction and Trigger* (MOZART vXXX)[?], which is based on GEANT v3.21[37].

The trigger simulation, called Complete ZGANA[2] Analysis Routine (CZAR

[2]The ZEUS GEANT Analysis (ZGANA) and TLT_ZGANA are simulations of the ZEUS 1st+2nd and 3rd trigger systems, respectively. CZAR is the combination of the two, encompassing the com-

vXXX[?]) provides complete description of the ZEUS trigger system response to simulated events. The output of MOZART and CZAR is an ADAMO table of detector signals which is organized similarly to real world data from the ZEUS detector. The principle difference from real data is the additional information about the generator, parton, and hardonic partos of the simulation. These ADAMO tables were then submitted to the ZEPHYR program discussed in the beginning of chapter 4, and were reconstructed identically as the real world data was. This source of information is referred to as at the *detector level*.

5.6 Reconstruction and Selection of Simulated Events

To produce parton and hadron level distributions, the only selection criteria apart from requiring jets were of $Q^2_{\rm app} > 200$ GeV^2 and $y_{\rm app} < 0.9$ were required, where $Q^2_{\rm app}$ and $y_{\rm app}$ differ from generated values $Q^2_{\rm gen}$ and $y_{\rm gen}$ due to ISR and FSR. Jets were found in the lab frame by applying the k_T cluster algorithm to partons and hadrons, as described in chapter 4.6. Jets were selected with the same phase space as the corrected detector level jets, described in section 6.5

To produce detector level and corrected detector level distributions, all reconstruction procedures and selection criteria were performed identically as the data as in chapters 4 and 6, with the following exceptions: No adjustment to calorimeter cell values by Escale03 as described in section 4.2 was performed to the MC samples. Selection cuts using DST bits, EVTAKE and POLTAKE flags were not applied to the MC samples.

It is useful to note that because the matric elements used in the MCs here are

plete trigger chain.

Chapter 7

Fixed-Order pQCD Calculations

7.1 The MEPJET Program

The calculations in this thesis were produced with the MEPJET 2.1 [38, 39] program, which was written for DIS by E. Mirkes. D. Zeppenfeld and S. Willfahrt. It is capable of computing partonic cross sections in DIS via exchange of photons, Z, and W bosons at $\mathcal{O}(\alpha_s)$ or $\mathcal{O}(\alpha_s{}^2)$. For this analysis, only W exchange was calculated, the mass of the W bosons were set to 80.40 GeV. The number of active quark flavors was set to 5, and $\Lambda^{(5)}_{\overline{\rm MS}}$ was set to 226 MeV, which corresponded to a strong coupling constant $\alpha_s(M_Z) = 0.118$. $\alpha_{\rm EM}$, G_F, and M_Z were set to the same values as chosen for the MC described in section 5.4. For the inclusive-jet calculations, $\mathcal{O}(\alpha_s)$ predictions were produced, which are NLO. For dijet and three-jet cross sections cross sections, $\mathcal{O}(\alpha_s{}^2)$ predictions were produced. which were NLO and LO, respectively.

7.2 Phase Space Slicing Method

The MEPJETprogram uses the phase space slicing method [40] to cancel infrared and collinear divergences. A parameter $s_{\rm min}$ was introduced, which acts as a mini-

mum resolving power to separate partons, and effectively separates finite and infinite portions of the phase space. Integrals in regions where particles i and j satisfied $s_{ij} \cong 2p_i \cdot p_j < s_{min}$ were calculated analytically. Some divergences were matched and exactly cancelled with virtual contribution divergences, and others were subrtracted as described in the next section. It has been shown that the final results are independent of the unphysical parameter $s_{\rm min}$ when $s_{\rm min}$ is less than or on the order of 0.1 GeV [39]. The remaining phase space, where the integrals were free of these divergences, were left to be caculated numerically by MC integration techniques. Leaving the finite portion of the phase space integral for the user provides a flexible system for custom modification of phase space cuts.

7.3 $\overline{\rm MS}$ Renormalization Scheme

In MEPJET, the one-loop divergences in the diagrams are handled in the $\overline{\rm MS}$ renormalization scheme. The scheme begins by introducing a dimensional regulator for loop momentum, by transforming the 4-dimensional integral over virtual particle momenta into $4 - 2\epsilon$ dimensions. This parameterizes the divergences in ϵ, but artificially introduces additional constant terms. The name stands for modified minimal subtraction, because the earlier method of subtracting only divergent parts [41] was later modified by additionally subtracting the constant terms that were introduced [42].

7.4 Scale Dependence and Uncertainty

In MEPJET, ultraviolet divergences were renormalized and absorbed into the bare coupling constant, introducing a dependance of the renormalized coupling $\alpha_{\rm s}(\mu_{\rm R})$. For

this analysis. μ_R was set to Q for the central values extracted from the calculations. The dependance on μ_R would cancel for calculations performed at all orders in α_s, but for our fixed order calculations, some residual scale dependance persists. To quantify the uncertainty introduced by our choice, the calculations were also performed for $\mu_R \equiv Q/2$ and $\mu_R \equiv 2Q$.

In MEPJET, initial state collinear divergences were factorized into the bare parton densities, introducing a dependence of the final cross sections on the factorization scale μ_F. This scale was also set to Q, and modified by a factor of 2 up and down to quantify uncertianty based on this choice.

7.5 PDFs and PDF Uncertainty

To produce *ep* predictions, hard scattering functions must be convolved with PDFs. Calculations in this thesis were performed using the ZEUS-S[43] parameterization of the proton PDFs. PDFs must be experimentally determined, and contain both theoretical and experimental uncertainties. To understand the experimental uncertainty of the PDFs effect on results, calculations were performed using multiple different sets of PDFs, each of which contains different experimental input. To understand the theoretical uncertainty involved in the choice of parameterization, alternative calculations were performed using the CTEQ6[44, 45] and MRST [46].

7.6 Comparisons with Other Theoretical Implimentations

It was observed in [47] that while MEPJETand the DISENT [48] agree for inclusive NC DIS jet cross sections on the order of 1%, the relative disagreement for NC DIS dijet cross sections is on the order of 5%. MEPJETis currently the only available

implimentation of $\mathcal{O}(\alpha_s{}^2)$ CC DIS, and therefore no systematic study in the choice of implimentation can be performed, but these values can be taken as rough estimates of theoretical uncertaintiy in our choice of theorertical impliunetnation.