Charged Current Cross Section Measurement at HERA

Grijpink, S.J.L.A.

Publication date
2004

Chapter 6

Cross Section Measurements

In the previous chapter the selection of charged current DIS events has been presented. In this chapter it will be discussed how this sample of charged current events has been used to measure the charged current *ep* cross sections. The binning of the kinematic range used in the measurement and the unfolding of the cross section will be discussed, followed by a discussion of the statistical and systematic uncertainties.

6.1. Bin Definitions

In order to measure the differential charged current cross sections the kinematic ranges are divided in bins wide enough to contain a sufficient number of events to measure the cross section in that bin. It is important to use an appropriate binning, since too narrow binning will increase the statistical error and migration effects between neighbouring bins will become too large. On the other hand, too wide binning would result in a measurement which reveals less information than it could have done otherwise. The binning chosen in this analysis ensures that the bin size is several times the resolution of the kinematic variable in which the cross section is unfolded.

The single differential cross section has been unfolded in the kinematic variables Q^2, x and y. For the measurement of the single differential cross section $\mathrm{d}\sigma/\mathrm{d}Q^2$ nine bins were defined in the Q^2 range $200-60000\,\mathrm{GeV}^2$. The Q^2 range $200-22494\,\mathrm{GeV}^2$ has been divided in eight bins with equal width in $\log Q^2$. Since the number of events drops rapidly with higher values of Q^2, the ninth bin had to be made larger and covered the Q^2 range $22494-60000\,\mathrm{GeV}^2$. For the unfolding of the single differential cross section $\mathrm{d}\sigma/\mathrm{d}x$ seven bins were defined in the x range $0.01-1.0$: three bins with equal width in $\log x$ in the x range $0.01-0.1$ and four bins with equal width in $\log x$ in the x range $0.1-1.0$. For the single differential cross section $\mathrm{d}\sigma/\mathrm{d}y$ seven bins were defined in the y range $0.0-0.9$: two bins with equal width in the y range $0.0-0.2$ and five

in the y range $0.2-0.9$. For both the e^-p and e^+p data sample the same binning was used for the single differential cross section measurements. Figures 4.6(b), 4.6(d) and 4.7(b) show the resolution in Q^2, x and y, respectively. The resolution in Q^2 is $\sim 30\%$ over the entire Q^2 range. The resolution in x improves from $\sim 30\%$ at low-x to $\sim 10\%$ at high-x. The resolution in y is $\sim 13\%$ over the entire y range.

The binning for the double differential cross section measurements in x and Q^2, $\mathrm{d}^2\sigma/\mathrm{d}x\mathrm{d}Q^2$, was based on the same binning as used in the single differential cross section measurements. The e^-p double differential cross section was measured in 26 bins, whereas in the e^+p data it was measured in 30 bins, in the x range $0.01-0.562$ and the Q^2 range $200-22494\,\mathrm{GeV}^2$. The difference in the number of bins between the e^-p and e^+p data is due to the larger beam-gas background in the e^-p data (see Sect. 5.4.1). Therefore, the cross section could not be measured in a number of low-Q^2 and high-x bins, though an additional bin was defined at high-x and high-Q^2, with Q^2 range $22494-60000\,\mathrm{GeV}$ and x range $0.316-0.562$. For the measurement of $\mathrm{d}^2\sigma/\mathrm{d}x\mathrm{d}Q^2$ in the e^+p data an additional bin was defined at low-x and low-Q^2, with Q^2 range $200-400\,\mathrm{GeV}^2$ and x range $0.006-0.01$. Figures 6.1 and 6.2 show the resolutions of Q^2, x and y respectively for the various $\mathrm{d}^2\sigma/\mathrm{d}x\mathrm{d}Q^2$ bins used in the e^+p data. The same resolutions were observed in the e^-p data.

The cross section measurements were restricted to bins with a high purity, $\mathcal{P}$, and a high acceptance, $\mathcal{A}$. In this way large corrections for detector acceptance and migration effects were avoided. The purity and acceptance of a bin are defined as:

- **purity,** $\mathcal{P}$: the number of events generated and measured in a bin divided by the number of events measured in that bin;
- **efficiency,** $\mathcal{E}$: the number of events generated and measured in a bin divided by the number of events generated in that bin;
- **acceptance,** $\mathcal{A}$: number of events measured in a bin divided by the number of events generated in that bin.

Here "measured in a bin" means that the kinematic variables of the reconstructed event were contained in that bin and that the event met the event selection criteria. Note that with this set of definitions the following relation holds

$$\mathcal{A} = \mathcal{E}/\mathcal{P} \tag{6.1}$$

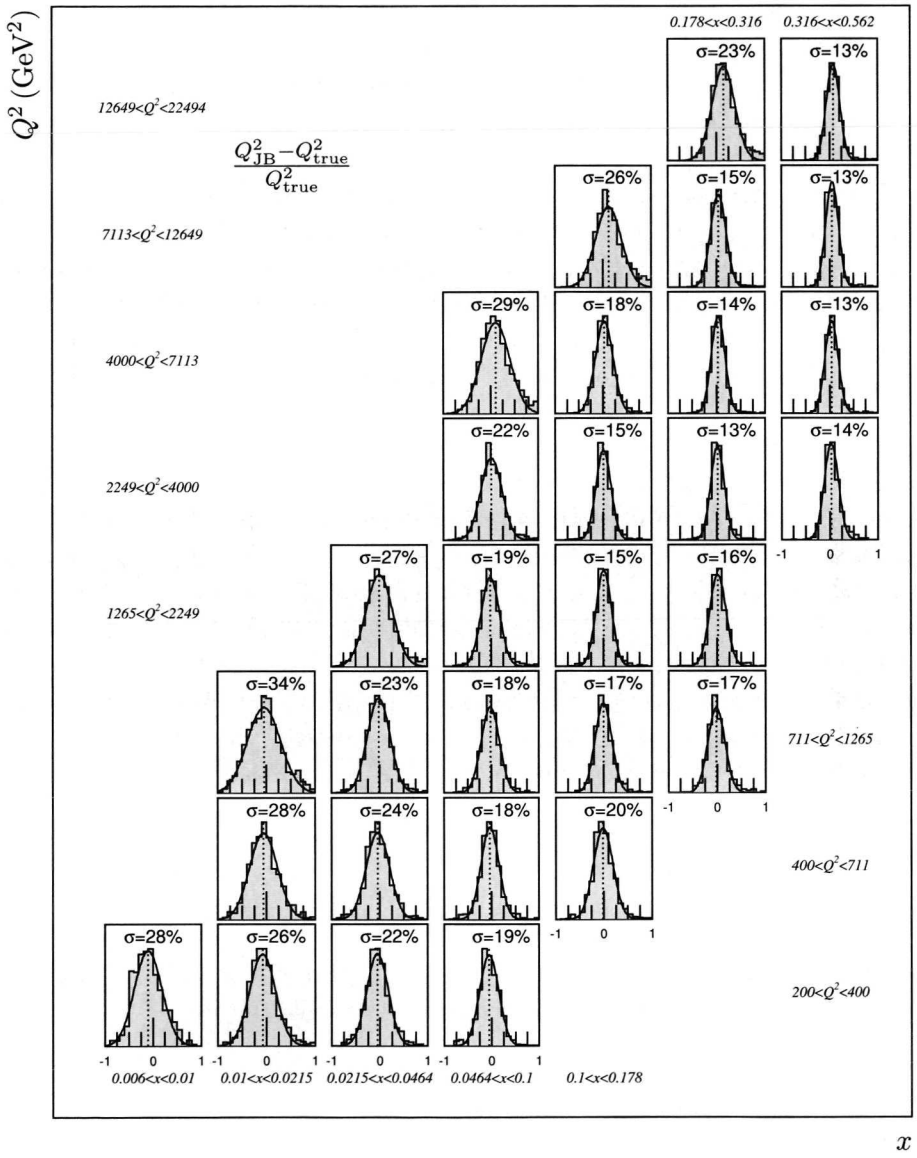

Figure 6.1. Resolution of Q^2 determined from the $(Q^2_{JB}-Q^2_{true})/Q^2_{true}$ distribution, shown for the x, Q^2 bins used in the unfolding of the e^+p double differential cross section. The best resolution are at high-x and high-Q^2.

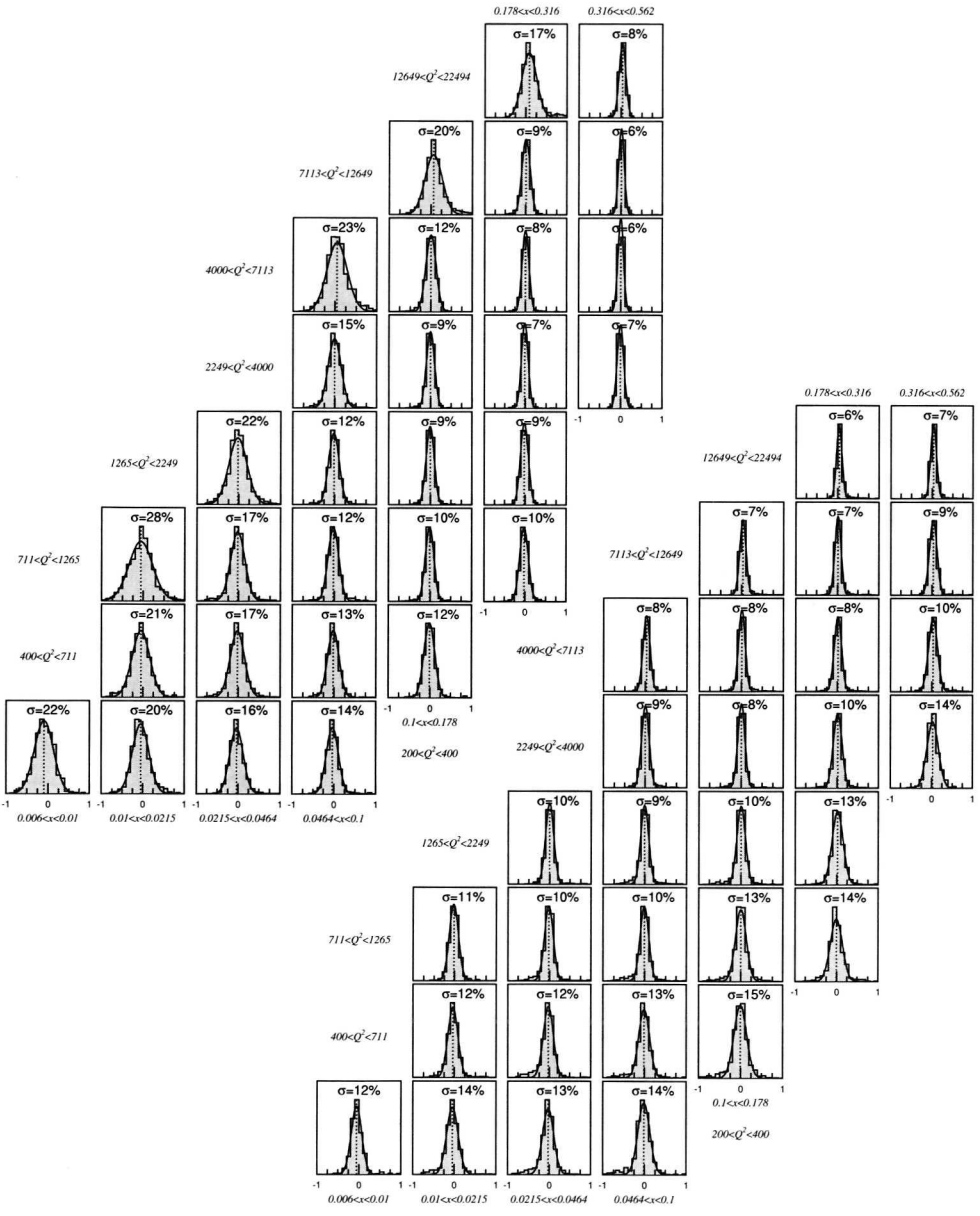

Figure 6.2. Resolutions of x (left), determined from the $(x_{JB} - x_{true})/x_{true}$ distribution, and y (right), determined from the $(y_{JB} - y_{true})/y_{true}$ distribution, shown for the x, Q^2 bins used in the unfolding of the e^+p double differential cross section.

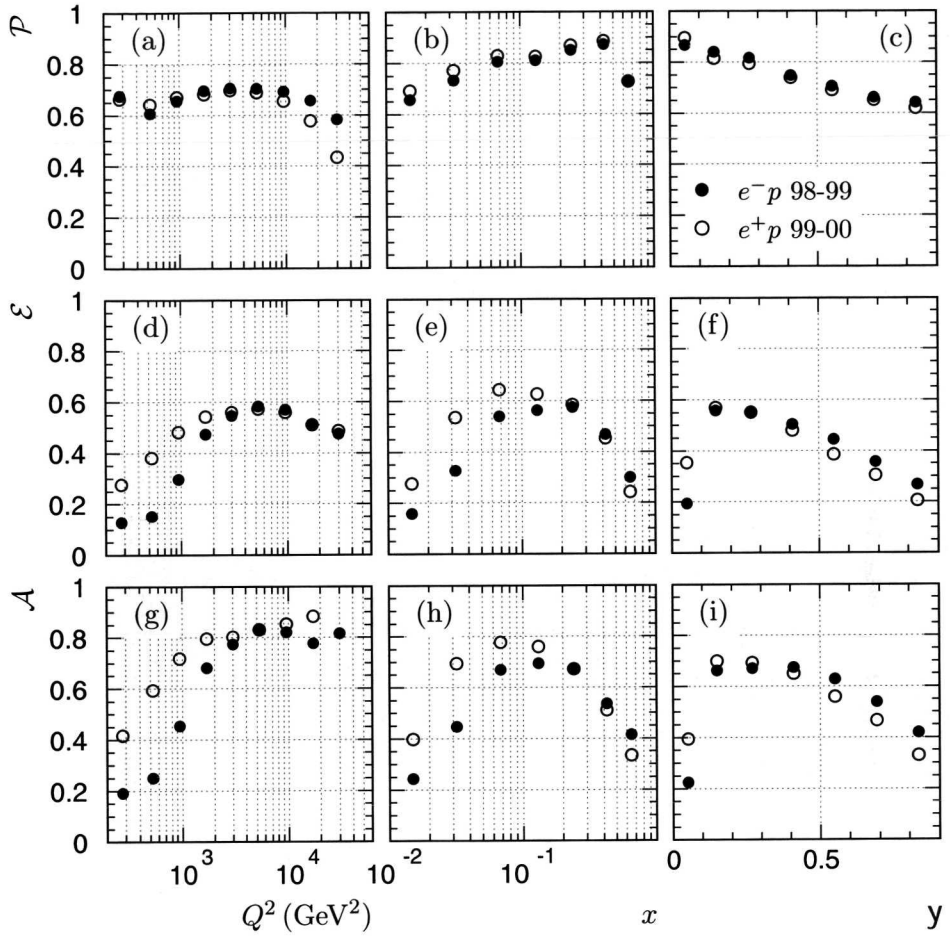

Figure 6.3. Various bin quality variables for the single differential bins in the kinematic variables Q^2, x and y. (a),(b) and (c) the purity $\mathcal{P}$; (d), (e) and (f) the efficiency $\mathcal{E}$; and (g), (h) and (i) the acceptance, $\mathcal{A}$. The solid (open) dots represent the e^-p (e^+p) data.

Figure 6.3 shows the various bin quantities for the different single differential bins in Q^2, x and y, respectively. The acceptance is above 30% for all bins, except for the lowest bins in Q^2,x and y. The purity is well above 50% for all bins, except for the highest bin in Q^2 which has a purity just below 50%. The

various bin quantities for all bins used in the analysis are listed in Tab. A.1 to A.8.

6.2. Cross Section Unfolding

The kinematic variables used in the measurement of the cross section are subject to various distortions like smearing effects, detector geometry effects and electroweak radiative effects. Hence, the measured values differ from the true values. The procedure to correct the measurement for these distortions is called unfolding. The cross section is extracted in bins of the various kinematic variables. The integrated cross section including radiative correction in a bin of Q^2 can be written as

$$\sigma_{\text{rad}}(\Delta Q^2) = \frac{N_{\text{data}} - N_{\text{bg}}}{\mathcal{A}\mathcal{L}_{\text{data}}}, \tag{6.2}$$

where $\mathcal{L}_{\text{data}}$ is the total integrated luminosity. N_{data} is the number of observed data events in the bin that passed the charged current event selection and N_{bg} is the number of background events in the bin, as estimated from MC simulation. The acceptance, $\mathcal{A}$, of the bin which is defined as $\mathcal{A} = N^{\text{MC}}_{\text{meas}}/N^{\text{MC}}_{\text{gen}}$, was used to correct for the effects from smearing and detector geometry. Where $N^{\text{MC}}_{\text{meas}}$ is the observed number of charged current MC events in the bin that passed the CC event selection and $N^{\text{MC}}_{\text{gen}}$ is the number of CC MC events generated in that bin. Re-weighting $N^{\text{MC}}_{\text{meas}}$ and $N^{\text{MC}}_{\text{gen}}$ to the measured luminosity Eq. (6.2) can be rewritten as

$$\sigma_{\text{rad}}(\Delta Q^2) = \frac{N_{\text{meas}}}{N^{\text{MC}}_{\text{meas}}} \frac{N^{\text{MC}}_{\text{gen}}}{\mathcal{L}_{\text{data}}} \tag{6.3}$$

$$= \frac{N_{\text{meas}}}{N^{\text{MC}}_{\text{meas}}} \sigma^{\text{MC}}_{\text{rad}}(\Delta Q^2), \tag{6.4}$$

where $N_{\text{meas}} = N_{\text{data}} - N_{\text{bg}}$ and $\sigma^{\text{MC}}_{\text{rad}}(\Delta Q^2)$ is the integrated radiative cross section in bin ΔQ^2 evaluated by the CC MC events. To determine the electroweak Born level cross section a correction factor was introduced

$$C_{\text{rad}} = \frac{\sigma^{\text{SM}}_{\text{Born}}(\Delta Q^2)}{\sigma^{\text{SM}}_{\text{rad}}(\Delta Q^2)}, \tag{6.5}$$

where $\sigma^{\text{SM}}_{\text{Born}}(\Delta Q^2)$ is the integrated Standard Model, SM, Born level cross section in bin ΔQ^2 and $\sigma^{\text{SM}}_{\text{rad}}(\Delta Q^2)$ is the integrated SM radiative cross section

in bin ΔQ^2. Applying this correction factor, the integrated Born level cross section in bin ΔQ^2 can be obtained from

$$\sigma_{\text{Born}}(\Delta Q^2) = C_{\text{rad}}\sigma_{\text{rad}}(\Delta Q^2), \tag{6.6}$$

$$= \frac{\sigma_{\text{rad}}(\Delta Q^2)}{\sigma^{\text{SM}}_{\text{rad}}(\Delta Q^2)}\sigma^{\text{SM}}_{\text{Born}}(\Delta Q^2), \tag{6.7}$$

where $\sigma^{\text{SM}}_{\text{rad}}(\Delta Q^2)$ was obtained using the same Monte Carlo simulation which had been used to calculate the acceptance, i.e. $\sigma^{\text{SM}}_{\text{rad}}(\Delta Q^2) = \sigma^{\text{MC}}_{\text{rad}}(\Delta Q^2)$. Therefore, combining Eq. (6.4) and Eq. (6.7) the Born level cross section can be written as

$$\sigma_{\text{Born}}(\Delta Q^2) = \frac{N_{\text{meas}}}{N^{\text{MC}}_{\text{meas}}}\sigma^{\text{SM}}_{\text{Born}}(\Delta Q^2). \tag{6.8}$$

To obtain the differential cross section at a specific reference point in the bin, a correction factor was applied. For the differential cross section in Q^2 this bin centring correction factor was defined as

$$C_{\text{centre}} = \frac{\left.\frac{\mathrm{d}\sigma^{\text{SM}}_{\text{Born}}(Q^2)}{\mathrm{d}Q^2}\right|_{Q^2=Q^2_c}}{\sigma^{\text{SM}}_{\text{Born}}(\Delta Q^2)}, \tag{6.9}$$

where $\mathrm{d}\sigma^{\text{SM}}_{\text{Born}}(Q^2)/\mathrm{d}Q^2\big|_{Q^2=Q^2_c}$ is the SM Born level differential cross section at the reference point Q^2_c. Hence, the Born level differential cross section in Q^2 at the reference point Q^2_c can be obtained from

$$\left.\frac{\mathrm{d}\sigma_{\text{Born}}(Q^2)}{\mathrm{d}Q^2}\right|_{Q^2=Q^2_c} = C_{\text{centre}}\sigma_{\text{Born}}(\Delta Q^2). \tag{6.10}$$

Substituting Eq. (6.8) and (6.9) into Eq. (6.10) the Born level differential cross section can be written as

$$\left.\frac{\mathrm{d}\sigma_{\text{Born}}(Q^2)}{\mathrm{d}Q^2}\right|_{Q^2=Q^2_c} = \frac{N_{\text{meas}}}{N^{\text{MC}}_{\text{meas}}}\left.\frac{\mathrm{d}\sigma^{\text{SM}}_{\text{Born}}(Q^2)}{\mathrm{d}Q^2}\right|_{Q^2=Q^2_c}. \tag{6.11}$$

Finally, the unfolded Born level differential cross section at the reference point Q^2_c was obtained by

$$\left.\frac{\mathrm{d}\sigma_{\text{Born}}(Q^2)}{\mathrm{d}Q^2}\right|_{Q^2=Q^2_c} = \frac{N_{\text{data}} - N_{\text{bg}}}{N^{\text{MC}}_{\text{meas}}}\left.\frac{\mathrm{d}\sigma^{\text{SM}}_{\text{Born}}(Q^2)}{\mathrm{d}Q^2}\right|_{Q^2=Q^2_c}. \tag{6.12}$$

The SM differential cross sections were evaluated in the on-shell scheme [51] using the PDG values for the electroweak parameters and the CTEQ5D [52] parton distribution functions, PDFs. The same unfolding procedure was followed for the single differential cross sections $d\sigma/dx$ and $d\sigma/dy$ and for the double differential cross sections in bins of x and Q^2, $d^2\sigma/dxdQ^2$.

The reference points in the unfolding of $d\sigma/dQ^2$, $d\sigma/dx$ and $d^2\sigma/dxdQ^2$ were chosen to be the logarithmic centres of the bins in Q^2 and x, except for the highest Q^2 and highest x bins. The reference point for the highest Q^2 bin was set so that the logarithmic distance to the previous reference point was equal to the logarithmic distances between the other reference points. The reference point in the highest x bin was set at $x_c = 0.65$ [74]. The reference points in the unfolding of $d\sigma/dy$ were chosen to be the linear centres of the bins in y. The single differential cross sections in x and y are quoted for $Q^2 > 200\,\text{GeV}$. The calculated SM single differential cross sections in Q^2 and x include the region $y > 0.9$. Hence the acceptance loss by the y selection threshold is corrected and the obtained cross sections were extrapolated to the full y range.

6.3. Background Estimation

Various Monte Carlo samples were used to estimate the number of ep interactions other than charged current interactions passing the CC event selection. These background events were subtracted in the cross section unfolding procedure (see (6.12)). The ep backgrounds evaluated using MC samples were: NC DIS, photoproduction, charged lepton production and single W production. Section 3.2 gives an overview of the MC programs which were used to generate the background events. Tables A.1 to A.8 list the background contributions from the different ep processes in the bins used in the cross section unfolding. The smallest background contribution comes from the NC DIS interactions, whereas the photoproduction background is the largest. Over the full kinematic range the background is well below 2%, except in the lowest Q^2 bins. Here the background contamination is of the order of 5% for e^-p and 10%.for e^+p data.

6.4. Statistical Uncertainties

The quoted statistical uncertainties in the cross section measurements are determined using standard statistical data analysis techniques. The cross section

is proportional to the number of events by (see eq. (6.12))

$$\sigma \sim \frac{N_{\text{data}} - N_{\text{bg}}}{N_{\text{MC}}} \tag{6.13}$$

where N_{data} is the total number of observed data events and N_{MC} and N_{bg} are the number of measured charged current and background MC events, respectively. N_{MC} and N_{bg} were obtained by the weighted sum of all the events passing the CC event selection criteria from the various Monte Carlo samples; $N_{\text{MC}} = \sum_i w_{\text{MC},i}$ and $N_{\text{bg}} = \sum_i w_{\text{bg},i}$ where i runs over all events and the weight assigned to each of the generated events is such that the total number of events is normalised to the data luminosity. The statistical error of N_{MC} in a bin is

$$\Delta N_{\text{MC}} = \sqrt{\sum_i w_{\text{MC},i}^2} \tag{6.14}$$

and similarly for N_{bg}: $\Delta N_{\text{bg}} = \sqrt{\sum_i w_{\text{bg},i}^2}$. The weight of the observed data events is one. Therefore, the statistical error of the number of data events in a bin is

$$\Delta N_{\text{data}} = \sqrt{N_{\text{data}}} \tag{6.15}$$

The statistical error of the cross section measurements can now be obtained from

$$\delta_{stat}^i = \sqrt{\frac{(\Delta N_{\text{data}})_i^2 + (\Delta N_{\text{bg}})_i^2}{(N_{\text{data}} + N_{\text{bg}})_i^2} + \left(\frac{\Delta N_{\text{MC}}}{N_{\text{MC}}}\right)_i^2} \tag{6.16}$$

where i denotes the bin number. For bins with less than 12 events a 67% confidence interval was calculated using Poisson statistics; the boundaries of this confidence interval were taken as the statistical uncertainty.

6.5. Systematic Uncertainties

Systematic effects in the measurement can give a bias in the unfolding of the cross section. Various sources of systematic uncertainties have been studied. The most important ones were found to be the energy scale of the calorimeter, QCD cascade models and the effects of the selection thresholds. Other sources of systematic uncertainties which have been studied were: effects of the parton density functions, effects of the NLO QCD corrections, energy leakage, CTD

vertex finding efficiency and the MC vertex distribution. The systematic uncertainties have been studied in the same bins as used in the unfolding. The final systematic error will be obtained by the quadratic sum of all the systematic uncertainties.

6.5.1. Calorimeter Energy Scale

A very important systematic uncertainty is the uncertainty of the energy scale of the calorimeter. This energy scale has a direct effect on the reconstruction of the kinematic variables and therefore on the measurement of the cross sections. Especially at high-Q^2 the effect can be relatively large due to the steeply falling of the cross section. The energy scale and the associated uncertainty of the energy scale were determined, using NC DIS events, from the ratios of the total hadronic transverse momentum, $P_{T,\mathrm{h}}$, to $P_{T,\mathrm{DA}}$ and $P_{T,e}$, where $P_{T,\mathrm{DA}} = \sqrt{Q^2_{\mathrm{DA}}(1-y_{\mathrm{DA}})}$ is the transverse momentum obtained from the double-angle method (see (4.16) and (4.17)) and $P_{T,e}$ is the measured transverse momentum of the scattered electron. In order to restrict the hadronic activity to particular polar regions, a sample of NC DIS events with a single jet was selected. By applying suitable cuts on the location of the current jet and evaluating $P_{T,\mathrm{h}}/P_{T,\mathrm{DA}}$ and $P_{T,\mathrm{h}}/P_{T,e}$ event by event, the hadronic energy scales of the FCAL and BCAL were determined. The responses of the HAC and EMC sections of the individual calorimeters were determined by plotting $P_{T,\mathrm{h}}/P_{T,\mathrm{DA}}$ and $P_{T,\mathrm{h}}/P_{T,e}$ as a function of the fraction of the hadronic energy measured in the EMC section of the calorimeter. In each case, the uncertainty was found by comparing the determinations from data and MC. In order to study the hadronic energy scale in the RCAL, a sample of diffractive DIS events was selected. Such events are characterised by a large gap in the hadronic energy flow between the proton remnant and the current jet. $P_{T,\mathrm{h}}/P_{T,\mathrm{DA}}$ was evaluated event-by-event for events with hadronic activity exclusively in the RCAL and the energy scale and associated uncertainty determined.

The relative uncertainty of the energy scale was determined to be 2% for the RCAL and 1% for the FCAL and BCAL [80]. Varying the energy scale of the calorimeter sections by these amounts in the detector simulation induces small shifts of the kinematic variables. The variations of the energy scale of each of the calorimeters simultaneously up or down by these amounts gave the systematic uncertainty on the total measured energy in the calorimeter. By increasing (decreasing) the FCAL and RCAL energy scales together while

the BCAL energy scale was decreased (increased) the uncertainty in the cross sections from the effect of the energy scale on the measurement of γ_h was obtained. The uncertainty stemming from the method used to determine the relative uncertainty was determined by simultaneously increasing the energy measured in the EMC section of the calorimeter by 2% and decreasing the energy measurement in the HAC section by 2% and *vice-versa*. This was done separately for each of the calorimeters.

The effect of the uncertainty of the energy scale is maximal in high-Q^2 and high-x bins. These are also the bins with the lowest number of events. Using both data and MC to estimate the systematic uncertainty on the cross section measurement yields an overestimate of the error due to statistical fluctuations in the number of events in these bins. To circumvent this effect only the MC simulation was used to determine the systematic error on the cross section, in the following way:

$$\delta_E^i = \frac{N_{\text{nom}} - N_i}{N_i}, \tag{6.17}$$

where i denotes a particular energy scale variation. N_{nom} is the number of events in the nominal, i.e. not scaled, MC data and N_i is the number of events in the scaled MC data. The systematic error on the cross section, due to the uncertainty of the calorimeter energy scale was obtained by quadratic summation of the three estimates. The uncertainties from this check reach $\sim 15\%$ in the highest Q^2 bins and $\sim 20\%$ in the highest x bins.

6.5.2. QCD Cascade Model

The QCD cascade model used in the Monte Carlo event simulation in this analysis was provided by the colour dipole model, CDM as implemented in the ARIADNE [53] program. As an alternative to the CDM from ARIADNE the matrix element parton shower, MEPS, model as implemented in the LEPTO [50] program can be used for the simulation of the QCD cascade. Both models are successful in describing data from high-Q^2 DIS events [81]. The sensitivity of the cross section measurement to the higher order QCD effects in the hadronic final state was estimated by using the MEPS model from LEPTO instead of the CDM from ARIADNE. The systematic error on the cross section was obtained by the difference in acceptance between the two models

$$\pm\, \delta_{\text{MEPS}} = \pm \left| \frac{\mathcal{A}_{\text{CDM}} - \mathcal{A}_{\text{MEPS}}}{\mathcal{A}_{\text{CDM}}} \right| \tag{6.18}$$

where $+\delta_{\text{MEPS}}$ ($-\delta_{\text{MEPS}}$) is the error in the positive (negative) direction, and $\mathcal{A}_{\text{CDM}}$ and $\mathcal{A}_{\text{MEPS}}$ are the acceptances calculated using the CDM model and MEPS model respectively. The largest uncertainty is found in the e^+p data in the highest Q^2 bin where it reaches $\sim 20\%$ and $\sim 12\%$ in the e^-p. In the highest x bins the uncertainty is $\sim 7\%$.

6.5.3. Selection Thresholds

Many selection thresholds were varied in order to verify the stability of the cross section measurement in terms of efficiency and purity. Generally the selection thresholds for a selection variable were varied by an amount comparable with the resolution of the variable. Furthermore, the thresholds were varied by such an amount that the selection efficiency was still good, and the number of background events, i.e. beam-gas, cosmic muons, etc., did not become too large. Most of the varied selection thresholds did not change the measured cross section, and were therefore not included in the uncertainty [82]. The uncertainty on the cross section due to the selection threshold variation was obtained from the difference between the nominal cross section and the cross section calculated with the threshold variation

$$\delta_{\text{T}}^{i} = \frac{\sigma_i - \sigma_{\text{nom}}}{\sigma_{\text{nom}}} = \frac{N_{\text{data}} - N_{\text{bg}}}{N_{\text{meas}}^{\text{MC}}} \cdot \frac{N_{\text{meas}}^{\text{MC},i}}{N_{\text{data}}^{i} - N_{\text{bg}}^{i}} - 1, \tag{6.19}$$

where i denotes the threshold variation and σ_{nom} the cross section unfolded with the nominal event selection. The selection thresholds which, when shifted, significantly changed the cross section, and for which it was not possible to estimate the uncertainty in an other way, were included in the systematic error. Statistical fluctuations, due to limited statistics in some bins, were suppressed by demanding that changes in $N_{\text{data}} - N_{\text{bg}}$ did not exceed 5%. If so, the uncertainty in the bin for that particular threshold variation was set to zero. In order not to overestimate the uncertainties, the threshold variations were separated in two sets, transverse momentum, T1, and tracking quantities, T2. The largest uncertainty in a set was selected as the uncertainty of the threshold variation for that set.

T1, transverse momentum

The first set of threshold variations, T1, is concerned with the transverse momentum selection cuts:

- $P_{T,\text{miss}} > 12 \pm 1.2\,\text{GeV}$, for high-$\gamma_0$ events;
- $P_{T,\text{miss}} > 14 \pm 1.4\,\text{GeV}$, for low-$\gamma_0$ e^+p events;
- $P_{T,\text{miss}} > 25 \pm 2.5\,\text{GeV}$, for low-$\gamma_0$ e^-p events;
- $P'_{T,\text{miss}} > 10 \pm 1.0\,\text{GeV}$, for high-$\gamma_0$ events;
- $P'_{T,\text{miss}} > 12 \pm 1.2\,\text{GeV}$, for low-$\gamma_0$ e^+p events;
- $P'_{T,\text{miss}} > 25 \pm 2.5\,\text{GeV}$, for low-$\gamma_0$ e^-p events;

where the $P_{T,\text{miss}}$ and $P'_{T,\text{miss}}$ cuts are described in Sect. 5.3 and Sect. 5.4, respectively. The selection thresholds are varied by the resolution of P_T, which is of the order of 10%. The uncertainty arising from these variations are up to $\sim 3\%$ in the lowest-x and highest y bins and up to $\sim 8\%$ in the lowest-Q^2 lowest-x bin of the double differential cross section in the e^+p data.

T2, track quantities

The second set of threshold variations, T2, is concerned with the selection thresholds on tracking variables:

- $\theta^{\text{vtx}}_{\text{trk}} > 15° + 18.5°$;
- $P^{\text{vtx}}_{T,\text{trk}} > 0.2 + 0.02\,\text{GeV}$;
- $N^{\text{good}}_{\text{trk}} > 0.25 N_{\text{trk}} - 5 \pm 1$;
- $N^{\text{good}}_{\text{trk}} > N_{\text{trk}} - 5 \pm 1$, for e^-p events;
- $N^{\text{good}}_{\text{trk}} > 10 \pm 1$, for e^-p events;

The first two thresholds concern the definition of a "good" track and are described in Sect. 5.4. The $\theta^{\text{vtx}}_{\text{trk}}$ threshold is tightened to select only tracks passing six super-layers of the CTD instead of five, and the $P^{\text{vtx}}_{T,\text{trk}}$ thresholds was varied with a somewhat arbitrary 10%. Both the N_{trk} and $N^{\text{good}}_{\text{trk}}$ thresholds are also described in section Sect. 5.4 The additional threshold selection for the e^-p data is described in Sect. 5.4.1. The uncertainty arising from these variations is $\sim 4\%$ in the lowest-x bins. In the e^-p data the uncertainties are $\sim 12\%$ in the lowest-Q^2 bin and up to 17% in the lowest-Q^2 lowest-x bin.

6.5.4. Background Subtraction

The backgrounds discussed in Sect. 6.3 were subtracted in the cross section unfolding procedure. Hence, uncertainties in the normalisation or shapes of these backgrounds can bias the cross section measurement. The largest background contribution came from the direct and resolved photoproduction events. The contribution to the systematic error on the cross section due to the uncertainty of the normalisation is presented in this section.

Figures 6.4(a) and 6.4(c) show the P_T/E_T distribution for high-γ_0 events with $P_T < 20\,\text{GeV}$ for e^-p and e^+p, respectively. The arrows in the figures indicate the selection thresholds as applied in the CC event selection (see Sect. 5.7). Hence, only the background events with $P_T/E_T > 0.55$ were subtracted in the cross section unfolding. Below the P_T/E_T threshold, a large number of photoproduction events is observed in both e^-p and e^+p. The uncertainty in the normalisation of the direct and resolved photoproduction events was obtained by a χ^2 fit, using MINUIT [83], to the total P_T/E_T distribution, with the following function:

$$N_{\text{MC}} = \alpha(\beta f_{\text{dir}} + (1-\beta) f_{\text{res}}) + N_{\text{CC}} + N_{\text{other}} \tag{6.20}$$

where α and β are the fit parameters. Parameter α is the sum of all photoproduction events, i.e. the total photoproduction normalisation, N_{php}; Parameter β is the fraction of direct photoproduction events of the total number of photoproduction events, F_{dir}; N_{CC} is the total number of CC MC events and N_{other} is the sum of all other background MC events (NC DIS, charged lepton production and single W production); f_{dir} and f_{res} are defined as

$$f_{\text{dir},i} \equiv N_{\text{dir},i} / \sum_{i=\text{bin}} N_{\text{dir},i} \qquad f_{\text{res},i} \equiv N_{\text{res},i} / \sum_{i=\text{bin}} N_{\text{res},i}$$

where i denotes the histogram bin number. $N_{\text{dir},i}$ and $N_{\text{res},i}$ are the number of direct and resolved photoproduction events in histogram bin i, respectively. The sum runs over all histogram bins included in the fit. From the above the following χ^2-square definition is obtained

$$\chi^2 = \sum_{i=\text{bin}} \frac{(N_{\text{data},i} - N_{\text{MC},i})^2}{(\delta N_{\text{data},i})^2 + (\delta N_{\text{MC},i})^2} \tag{6.21}$$

where $N_{\text{data},i}$ is the number of data events in histogram bin i, and $N_{\text{MC},i}$ is the sum of the number of events from all MC simulations in histogram bin i,

determined from (6.20). $\delta N_{\text{data},i}$ and $\delta N_{\text{MC},i}$ denote the statistical errors on $N_{\text{data},i}$ and $N_{\text{MC},i}$, respectively. (6.20) was chosen as the fit function, since it separates the relative normalisation between the direct and resolved photoproduction MC from the overall photoproduction MC normalisation. Therefore, it was possible to fit the normalisation, N_{php}, and the fraction of direct and resolved photoproduction, F_{dir}, separately.

First a fit was performed to determine N_{php}, with F_{dir} fixed at the values provided by the MC generator; this was followed by a fit of F_{dir} with N_{php} fixed at the fitted value. These fits were performed once in the P_T/E_T range $0.1-1.0$ and once in the range $0.25-0.8$. The results from the fits are listed in Table 6.1, and Figs. 6.4(b) and 6.4(d) show the χ^2/ndf distributions. From these distributions it is clear that no sensitivity for F_{dir} is observed in the P_T/E_T distributions. Hence, no contribution to the systematic error on the cross section measurement was obtained for the fraction of direct and resolved photoproduction.

The fit of N_{php} in both P_T/E_T regions for e^+p, resulted in an uncertainty of the normalisation of $\sim 10\%$. For e^-p, the fit of both N_{php} and F_{dir} failed in the larger P_T/E_T range $0.1-1.0$, due to a lack of statistics. This lack of statistics also influenced the fit for e^-p in the tighter P_T/E_T range $0.25-0.8$, resulting in a large uncertainty of the normalisation of $\sim 25\%$. Since no difference between the photoproduction background in e^-p and e^+p is expected, and the fits of the e^-p data were very much influenced by lack of statistics, the same uncertainty of the normalisation found for e^+p was applied for e^-p. To determine the contribution of the uncertainty on the cross section measurement due to the normalisation of photoproduction, the photoproduction background was varied up and down by 20%, corresponding to twice the value of the uncertainty given by the fit, in both e^-p and e^+p. The systematic error was than obtained by

$$\delta^{\pm}_{\text{php}} = \frac{N_{\text{nom}} - N_{\pm}}{N_{\text{nom}}} \tag{6.22}$$

were N_{nom} is the number of MC events in a sample with the subtracted photoproduction background normalised to the generator cross section. $N_{\pm}$ is the number of MC events with the photoproduction background varied up and down, as described above. The systematic errors were typically less than 1%. Only in one of the lowest-Q^2 bins of the double differential cross section the systematic error was $\sim 4\%$.

The contribution to the cross section measurement from the other backgrounds (NC, charged lepton production and single W production) was very

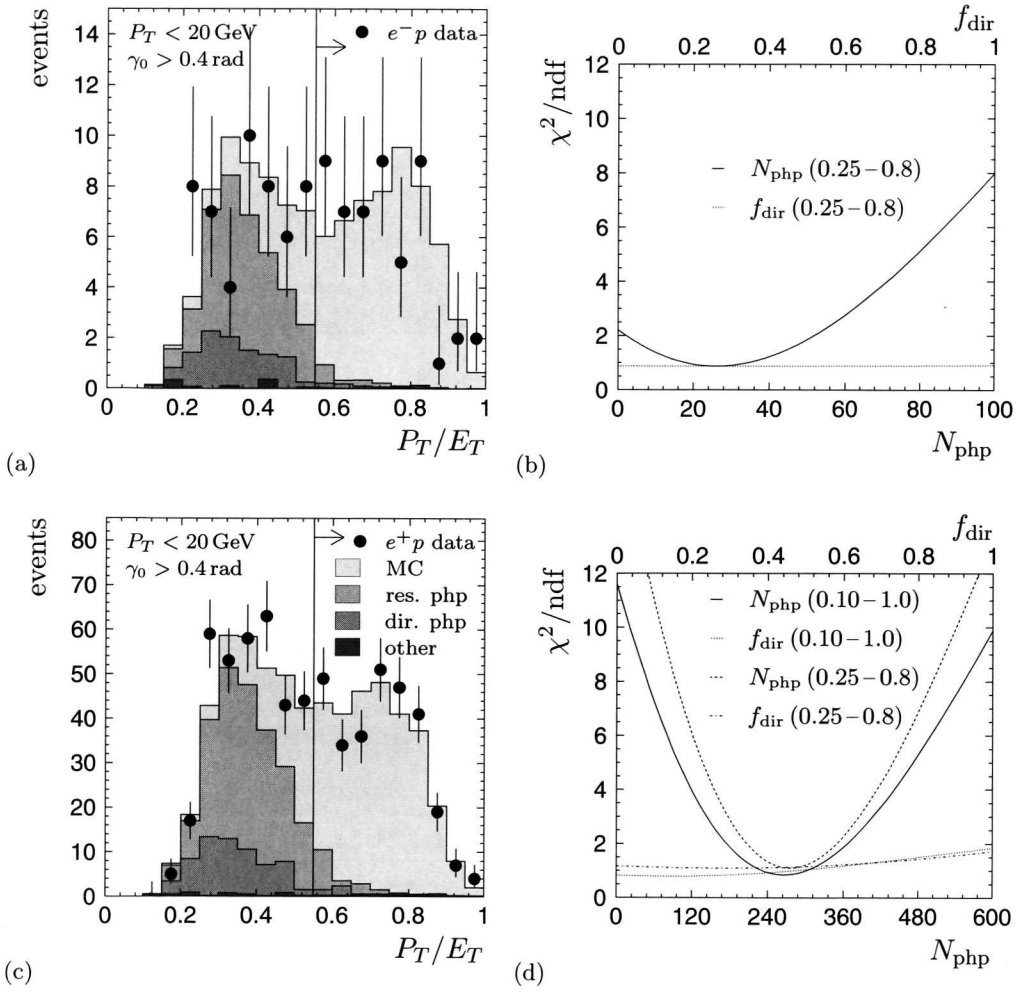

Figure 6.4. (a) The P_T/E_T distribution for events with high-γ_0 and $P_T < 20\,GeV$ for e^-p and, (c) for e^+p. (b) The χ^2/ndf distributions of the four fits performed to the P_T/E_T distribution as function of the fraction of direct photoproduction of the fit (upper axis), and as function of the total number of photoprodcution events (lower axis) for e^-p and, (d) for e^+p.

small, and variations of the normalisation of these background by 100% resulted in variations in the cross section well below 0.5% in the full kinematic

Table 6.1. Results for the fit to the P_T/E_T distribution. The numbers for the nominal situation are not fitted but derived from the cross sections given by the MC generator. The fits to N_{php} and f_{dir} are performed separately, e.g. N_{php} is fitted while $f_{dir}$is fixed and vice versa.

fit condition	fit range	N_{php}	F_{dir}	χ^2/ndf
Nominal (e^-p)		38.9 ± 1.4	0.27 ± 0.07	
N_{php} fit	0.25 – 0.8	26.0 ± 6.8	0.27	9.6/11
F_{dir} fit	0.25 – 0.8	26.0	0.16 ± 0.48	9.0/11
Nominal (e^+p)		280.9 ± 7.4	0.31 ± 0.05	
N_{php} fit	0.10 – 1.0	265.8 ± 21.1	0.31	15.2/18
F_{dir} fit	0.10 – 1.0	265.8	0.14 ± 0.17	14.2/18
N_{php} fit	0.25 – 0.8	275.5 ± 22.7	0.31	12.1/11
F_{dir} fit	0.25 – 0.8	275.5	0.28 ± 0.28	12.0/11

range. Therefore the contribution to the total systematic uncertainty from the subtraction of these backgrounds was neglected.

6.5.5. Parton Distribution Functions

The Monte Carlo events used in unfolding the cross section were generated with the CTEQ5D [52] PDFs. The same PDFs were used in the calculation of the bin centring corrections. In this way a consistent unfolding of the cross section was achieved. The influence on the cross section from variations of the PDFs were investigated using the ZEUS-S NLO QCD fit [84] via the difference in acceptance. The Monte Carlo events were re-weighted to the total experimental uncertainty of the prediction of the cross sections evaluated from the ZEUS-S fit. Note that no HERA CC data is included in the fit. The cross sections were unfolded using the re-weighted MC, and compared with the nominal cross sections. The differences in the measured cross sections for the e^-p data were below 0.5% in the full kinematic region, and therefore the contribution to the total systematic error was neglected. For the e^+p data the differences were below 1% except for the highest Q^2 bin where it was -5% and the highest x bin where it was $+4\%$. Hence, the effect of the uncertainty in the PDFs, δ_{PDF}, was included in the total systematic error for the e^+p data.

6.5.6. Effect of NLO QCD Corrections

The computer program DJANGOH [48] does not take into account contributions to the cross section from the longitudinal structure function, F_L, and NLO QCD corrections to xF_3 when generating Monte Carlo events. However, at high-y the contribution of F_L to the cross section is of the order of 10% [18]. In the calculation of the bin centring corrections the contribution of NLO QCD corrections were also neglected, yielding a consistent unfolding of the cross sections, and effects from neglecting the NLO QCD corrections can only originate from differences in the acceptance. The uncertainty is obtained by re-weighting the MC events to the ratio between the cross section calculated with and without NLO QCD corrections. The systematic errors, δ_{QCD}, were typically less than 1% for both e^-p and e^+p. The largest effect was observed in the e^+p data in the highest Q^2 bin where it was $\sim 6\%$ and in the highest x bin where is was $\sim 4\%$.

6.5.7. Energy Leakage

For an accurate measurement of the kinematic variables, it is important that the hadronic system is fully contained within the CAL. Energy leakage of the hadronic system out of the CAL can have an effect on the cross section measurement. The CAL is surrounded by the backing calorimeter, BAC (see Sect. 2.3.2), which was used to measure the effect of energy leakage of the CAL. It was found that 4% of the accepted events had a measurable energy leakage from the CAL into the BAC. The average energy fraction in the BAC w.r.t. to the total energy was 5%. Both the fraction of events with leakage and the average amount of leakage were well modelled by the MC simulation and the effect on the cross section measurement is negligible.

6.5.8. Vertex Finding Efficiency

A difference in the CTD vertex finding efficiency, $\mathcal{E}_{\mathrm{CTD}}$, between data and Monte Carlo can bias the measurement of the cross section. To obtain the $\mathcal{E}_{\mathrm{CTD}}$ in the γ_0 range of $0.0-0.6$ rad the CC event selection was redone with the γ_0 threshold set to 0.6 rad (see Sect. 5.2). $\mathcal{E}_{\mathrm{CTD}}$ was determined as the ratio of events with a CTD vertex and all events passing the CC event selection (events in the forward direction always have a timing vertex). Figures 6.5(a) and 6.5(b) show the $\mathcal{E}_{\mathrm{CTD}}$ for the e^-p and e^+p data and MC as a function of γ_0 . The turn

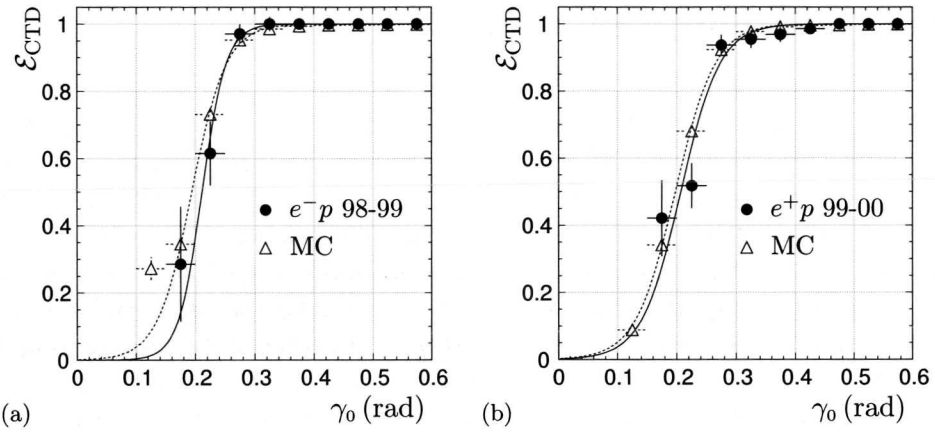

Figure 6.5. The CTD vertex finding efficiency as function of γ_0 for the (a) e^-p and (b) e^+p analysis. The solid dots represent the data events and the open triangles represent the MC events. Also shown are the turn on curves for data (solid line) and MC (dashed line) obtained from a fit.

on curves shown in Fig. 6.5 were obtained by a χ^2 fit to the function

$$\mathcal{E}_{\mathrm{CTD}} = \left(\frac{1}{2}\tanh\left(\frac{\gamma_0 - \alpha}{\beta}\right) + \frac{1}{2}\right)\epsilon, \tag{6.23}$$

with α, β and ϵ as free parameters. Parameter α is the turn on point, β is the slope and ϵ is the saturation value. It can been seen from the figure that good agreement is observed as γ_0 increases towards the 0.4 rad threshold where a CTD vertex is required in this analysis. Also it can be observed from the figure that the efficiency approaches 100% at the threshold of 0.4 rad for both e^-p and e^+p. Hence, the contribution from the CTD vertex finding efficiency to the systematic error is insignificant.

6.5.9. Vertex Distribution in Monte Carlo

The distribution of the Z position of the reconstructed vertex depends on the run period, due to changes of the beam conditions over time. The vertex distributions used in the Monte Carlo samples were corrected for these effects using the method described in Sect. 4.4. Changes in the measured cross section were found to be less than 0.5% and the contribution to the overall systematic uncertainty is insignificant.

6.5.10. Summary of the Systematic Uncertainties

To obtain the total systematic uncertainties the systematic uncertainties from each of the sources described in this section were added in quadrature for the positive and negative deviations from the nominal cross section values separately.

Figures B.1–B.6 show the various systematic checks described in the above sections for the single differential bins. The various systematic errors in all bins used in the analysis are listed in Tab. B.1 to B.8. Table 6.2 shows the systematic errors in the total cross section measurement for e^-p and e^+p charged current DIS in the kinematic region $Q^2 > 200\,\mathrm{GeV}^2$. The largest systematic uncertainty in e^-p came from the selection thresholds based on tracking and in e^+p from the QCD cascade modelling. Note that the largest error on the cross section measurements still came from the limited statistics.

Table 6.2. Uncertainties on the total cross section measurement for e^-p and e^+p charged current deep inelastic scattering in the kinematic region $Q^2 > 200\ GeV^2$.

source	error (%, e^-p)	error (%, e^+p)
calorimeter energy scale	$^{+0.34}_{-0.43}$	$^{+0.48}_{-0.26}$
QCD cascade model	± 0.57	± 1.08
selection thresholds, T1	± 0.65	± 0.25
selection thresholds, T2	± 0.95	± 0.40
php subtraction	$^{+0.18}_{-0.40}$	$^{+0.39}_{-0.68}$
PDF uncertainty	$^{+0.06}_{-0.10}$	± 0.50
NLO QCD corrections	-0.57	-0.85
total systematic error	$^{+1.3}_{-1.5}$	$^{+1.4}_{-1.7}$
statistical error	± 4.0	± 2.6

The uncertainties on the measured total luminosity were 1.8% and 2.25% for the e^-p and e^+p data, respectively, and were not included in the total systematic uncertainty.

6.6. Summary

The binning of the kinematic range used in the measurement of the cross section and the unfolding strategy together with an overview of the various systematic uncertainties were presented in this chapter.

In the next chapter the final results for the charged current cross section for e^-p and e^+p data will be discussed.