Measurement of the high-Q^2 neutral current deep inelastic scattering cross sections with the ZEUS detector at HERA

José Ricardo Morais Silva Gonçalo

Imperial College of Science Technology and Medicine.

Thesis submitted for the degree of
Doctor of Philosophy of The University of London
and the Diploma of Imperial College.

March 2003

ABSTRACT

The present thesis presents a measurement of the electron- and positron-proton neutral current deep inelastic scattering cross sections at high four-momentum transfer. The data used in these measurements were collected by the ZEUS experiment running in the HERA collider in Hamburg. The data analysed corresponds to integrated luminosities of 15.87 pb^{-1} for e^-p and 62.53 pb^{-1} for e^+p scattering, and was collected at a centre-of-mass energy of 318 GeV. The double-differential cross sections $d^2\sigma/dxdQ^2$ for e^-p and e^+p were measured in the ranges of $0.005 \leq x \leq 0.65$ and 200 GeV$^2 \leq$ Q$^2 \leq$ 30000 GeV^2. The single-differential cross sections $d\sigma/dQ^2$, $d\sigma/dx$ and $d\sigma/dy$ were measured for $Q^2 >$ 200 GeV^2. To observe the effect of the Z-boson exchange, the single-differential cross sections $d\sigma/dx$ and $d\sigma/dy$ were measured for $Q^2 >$ 10000 GeV^2 from both e^-p and e^+p data. The measured cross sections are in good agreement with the expectations of the Standard Model.

In addition to the study of the neutral current deep inelastic scattering process, the results were used to study the structure of the proton. The structure function F_2, which depends on the total quark content of the proton, was extracted using both the e^-p and the e^+p data samples. The measurements were combined to improve the accuracy of the extracted F_2 at high Q^2. The use of both data samples also permitted the extraction of the structure function xF_3, which depends on the valence-quark content of the proton.

To my parents:
Joaquim José
and
Maria Leonor

To my brother:
João Paulo

To Isabel

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Ken Long, for much help and encouragement and for never giving up. Thank you to my friends and colleagues at DESY and IC for a great time. At IC, thanks are due to Alex Tapper for many conversations where I actually learned stuff, to Rod Walker for witty jokes and so many invaluable insights into the British and continental cultures, languages, and the differences thereof. Thanks to Elaine McLeod, Richard White, Daniela Bauer, David Colling, Barry McEvoy, Louise Aspinwall, Robert Illingworth, Eduard McKigney, Kostas Georgiou and Michele Petteni for much entertainment and being good friends. Thanks to Ashok Jamdagni for a constantly good nature and understanding. Many more people in the group than there is space to mention here have contributed for this to be a good place to work. Outside the group, lunch at Chez Alice Marques has been the rule for the last year, thank you Alice.

At DESY, thank you to Ben West, Stuart Boogert, Eileen Heaphy, Eduardo Rodrigues, Dave Bailey, Mark Sutton, Claire Gwenlan, Boris Leissner, Maurício Barbi and Chris Cormack for great times. To Susan Ketels a special thank you for looking after us. In the first year of my PhD, Joanne Cole tried to show me where they keep all the ropes. Thanks Jo!

The work in this thesis owes a lot to the people in the Structure Functions and Electroweak physics group in ZEUS. Especially those who worked with me in the neutral current analysis of 98/99 and 99/00 data. Many thanks to Amaya López-Duran Viani, Matthias Moritz, Henning Schnurbusch and Alexander Kappes.

Life would not be complete without many friends. To Henrique Araújo, thanks for many deep and meaningful conversations about physics that neither of us understood (and some about physics that we did understand) and thank you for many years of friendship. To Marta Araújo, Patricia Aguiar, Sofia Araújo, Richard Hampson, Elley Wakui, Cristina Perez, Sónia Martins, Carlos Aguiar and Graça Ribeiro and their son Rodrigo, Guida Mendes, Cristina Castilho, Luis Jordão, Zilia Tovar and their son Diogo, Carmo Tovar, Berta and Holger, João Gil and Sónia, Joao Taborda, Ana Viana, Teresa Chuva and Paulo Martins, Helena Dinamene and Pedro Bandeira, my thanks for being great friends. To Isabel Pereira, thank you for more than I can say. Now we can go to Paris...

To my parents and my brother Paulo, thank you for all the support and encouragement when things were not going well and for great times when they were.

I would like to acknowledge the financial support of the Fundação para a Ciência e a Tecnologia under the PRAXIS XXI programme, reference: BD/16263/98.

Contents

Abstract **2**

Contents **5**

List of Figures **9**

List of Tables **14**

Chapter 1. Introduction **16**

Chapter 2. Theory of Deep Inelastic Scattering **20**

2.1 Kinematics of deep inelastic scattering 21

2.2 The quark-parton model 23

2.3 Neutral current cross section and structure functions 24

2.4 The quark-parton model and QCD 28

2.4.1 The DGLAP equations 30

2.5 Sum rules 31

2.6 Current parametrisations of parton density functions 32

2.6.1 CTEQ 33

2.6.2 MRST2001 34

2.6.3 GRV 35

2.6.4 ZEUS NLO QCD fit 35

Chapter 3. HERA and the ZEUS detector **37**

3.1 The HERA collider 37

3.1.1 The bunch structure 39

3.2 The ZEUS detector 40

3.2.1 The central tracking detector 42

3.2.2 The uranium calorimeter 44

3.2.3 The presampler 47
3.2.4 The luminosity detectors 48
3.2.5 Other detector components 50
3.2.6 The ZEUS trigger system 50
3.3 Detector simulation 54
3.3.1 The longitudinal vertex distribution 55
3.4 The software environment 56

Chapter 4. Monte Carlo simulation 58
4.1 DIS Monte Carlo 58
4.1.1 The parton cascade and hadronisation 60
4.1.2 The DJANGOH sample 62
4.1.3 Diffractive Monte Carlo 64
4.1.4 Reweighting to the cross section including F_L 66
4.2 Photoproduction background 67
4.2.1 Normalisation of the photoproduction Monte Carlo 68

Chapter 5. Measurement of Kinematic Variables 71
5.1 The electron reconstruction method 74
5.2 The Jacquet-Blondel reconstruction method 75
5.3 The double angle reconstruction method 76
5.4 Resolution of the reconstruction methods 77

Chapter 6. Event Selection and Reconstruction 82
6.1 Event characteristics 82
6.2 Characteristics of the background 84
6.2.1 Photoproduction 84
6.2.2 Beam-gas interactions 85
6.2.3 Halo and cosmic muons 85
6.2.4 Calorimeter sparks 86
6.2.5 Elastic QED Compton events 86
6.3 Data Preselection 87
6.3.1 First level trigger 88
6.3.2 Second level trigger 90
6.3.3 Third level trigger 91
6.3.4 Data quality 92
6.4 Off-line event reconstruction 93
6.4.1 Calorimeter corrections 93

6.4.2 Track and vertex reconstruction 99
6.4.3 Electron Identification and Corrections 102
6.4.4 Reconstruction of the Hadronic Final State 109
6.5 Off-line event selection 113
6.5.1 Final data samples 120
6.5.2 Event excess at high y and Q^2 124

Chapter 7. Systematic studies: vertex distribution and QCD cascade 127
7.1 Longitudinal vertex distribution 127
7.1.1 The minimum bias methods 128
7.1.2 Measurements and tests using Monte Carlo 132
7.1.3 Measurement of the 1998/99 vertex distribution 133
7.1.4 The minimum bias vertex distribution for 1999/00 136
7.2 Simulation of the QCD cascade 138
7.2.1 The hadronic energy-flow – definitions 138
7.2.2 Data and Monte Carlo samples 140
7.2.3 Results 141
7.2.4 The hadronisation systematic uncertainty 150

Chapter 8. Extraction of the cross sections 159
8.1 Unfolding of the cross sections 159
8.1.1 Bin-by-bin unfolding 161
8.1.2 Statistical uncertainties 163
8.1.3 Bin selection 164
8.1.4 Systematic uncertainties 168
8.2 Cross section results 177
8.2.1 Single differential cross section $\mathbf{d\sigma/dQ^2}$ 177
8.2.2 Single differential cross section $\mathbf{d\sigma/dx}$ 179
8.2.3 Single differential cross section $\mathbf{d\sigma/dy}$ 182
8.2.4 Reduced cross section $\tilde{\sigma}^{\mathbf{e^{\pm}p}}$ 182
8.3 The extraction of $\mathbf{F_2}$ 186
8.3.1 Combined results on $\mathbf{F_2}$ 186
8.4 The extraction of $\mathbf{xF_3}$ 191

Chapter 9. Conclusions and outlook 195
9.1 Comparison to published results 195
9.1.1 Comparison of $\mathbf{F_2}$ with published results 195
9.1.2 Comparison of $\mathbf{xF_3}$ with published results 197
9.2 Outlook: HERA-II 199

References 202

Appendix A. Tables of single-differential cross sections 213

Appendix B. Tables of double-differential reduced cross sections 220

Appendix C. Tables of structure functions 225

Appendix D. List of acronyms 229

List of Figures

2.1 Feynman diagram of the DIS process at lowest order. 21

2.2 Schematic representation of the $e^{\pm}p$ process at increasing scales of four momentum transfer, Q^2. 22

2.3 DIS phase space covered by ZEUS, compared to the reach of older fixed-target experiments. 23

2.4 Quark and gluon density functions at $Q^2 = 100$ GeV^2. 25

2.5 First measurement of the F_2 structure function in ZEUS. 29

2.6 Feynman diagram of the NC DIS process in the quark-parton model of the proton and at leading order in QCD. 29

2.7 Sum of the valence and sea quark PDFs plotted as a function of x for different values of Q^2. 30

2.8 Parton density functions from the ZEUS NLO QCD fit at $Q^2 = 10$ GeV^2 compared to CTEQ6 and MRST2001. 32

3.1 The HERA collider and pre-accelerators. 38

3.2 Integrated luminosity delivered by HERA from 1992 to 2000. 40

3.3 Bunch structure of HERA beams showing pilot and empty buckets. 41

3.4 Schematic representation of the ZEUS detector. 41

3.5 Schematic representation of the ZEUS detector. 42

3.6 Schematic representation of the CTD sense and field-shaping wires. 43

3.7 Simplified layout of the ZEUS detector. 45

3.8 Layout of calorimeter modules. 46

3.9 Map of inactive material in the ZEUS detector. 48

3.10 Layout of the ZEUS luminosity monitors. 50

3.11 Schematic representation of the ZEUS trigger system. 51

3.12 Output rates of the ZEUS trigger system during the 2000 running period. 54

3.13 Schematic representation of the ZEUS Monte Carlo production chain. 56

4.1 Feynman diagrams for the Born level NC DIS process and for some electroweak corrections. 59
4.2 Effect of FSR and QEDC emission of a single photon on the energy distribution of the scattered electron. 60
4.3 Feynman diagrams for the QED Compton process. 61
4.4 Feynman diagrams for: **a)** the leading order process; **b)** and **c)** gluon emission; and **d)** the boson-gluon fusion processes. 61
4.5 Distribution of the maximum pseudo-rapidity found in real data compared to that observed in diffractive and non-diffractive simulated data. 65
4.6 Reduced cross sections calculated with and without considering the F_L term, shown versus x and Q^2. 66
4.7 Diagrams representing direct and resolved photoproduction. 68
4.8 Energy deposited in the Lumi-e tagger by the scattered electron in PHP events. 69

5.1 Feynman diagram representing the process of deep inelastic scattering at Born level. 72
5.2 Schematic representation of deep inelastic scattering event topology. 73
5.3 Isolines of the scattered electron angle and energy. 74
5.4 Isolines of the struck quark energy, transverse momentum, and $\delta_h = (E - p_z)_h$. 75
5.5 Isolines of the scattered electron polar angle and of the hadronic angle. 76
5.6 Resolution of the double angle, electron, and Jacquet-Blondel reconstruction methods in the measurement of Q^2. 79
5.7 Resolution of the double angle, electron, and Jacquet-Blondel reconstruction methods in the measurement of x. 80
5.8 Resolution of the double angle, electron, and Jacquet-Blondel reconstruction methods in the measurement of y. 81

6.1 High-Q^2 neutral current deep inelastic scattering event. 83
6.2 Photoproduction event. 84
6.3 Event display of a cosmic shower. 86
6.4 Event with a calorimeter spark. 87
6.5 Elastic QED Compton event. 88
6.6 Schematic representation of the Iso-e algorithm in the CFLT. 89
6.7 SLT timing cuts. 91
6.8 BCAL alignment along z. 96
6.9 RCAL alignment. 97
6.10 Island clustering in the calorimeter. 98
6.11 Electron-seed multi-vertex resolution. 102

6.12 Correction to the electron energy in the FCAL. 106
6.13 Dead material correction in the RCAL using the presampler. 108
6.14 Correction of the electron energy for calorimeter non-uniformity and dead material. 109
6.15 Effect of corrections on overall electron energy distribution. 110
6.16 Fit of function used for the correction of the backscattering of energetic particles in the calorimeter. 111
6.17 Calorimeter hadronic energy scale determination. 113
6.18 Cuts on electron probability and energy cut. 114
6.19 Cuts on electron isolation and electron track-cluster distance of closest approach. 115
6.20 Cuts on the electron track momentum and forward electron transverse momentum. 116
6.21 Scatter plot of the scattered electron position in the RCAL and the vertex distribution measured in the 1999/00 data sample. 117
6.22 Cuts on the total $E - p_z$ and on y_{el}. 118
6.23 Cuts on p_T/E_T and $p_T/\sqrt{E_T}$. 119
6.24 Cut on the projection of the hadronic angle on the FCAL and Monte Carlo simulation validity cut. 120
6.25 Layout of some off-line cuts in the kinematic plane. 121
6.26 Control plots showing a comparison between some variables in Monte Carlo and real data accepted events for the 1998/99 data samples. 122
6.27 Control plots showing a comparison between some variables in Monte Carlo and real data accepted events for the 1999/00 data samples. 123
6.28 Number of accepted events per unit of integrated luminosity as a function of the run number. 124
6.29 Ratios of currents in colliding and pilot bunches for the 1999 e^-p running period. 125
6.30 Excess of observed events at $Q^2 \gtrsim 5000$ GeV^2 and $y \gtrsim 0.65$. 126

7.1 Topology of events used to determine the vertex distribution in the Q&R method. 129
7.2 Topology of events used to determine the vertex distribution in the W method. 132
7.3 Selection efficiency of the methods employed in the vertex distribution determination for the 1998/99 data sample. 133
7.4 Effect of the photoproduction background in the minimum bias vertex distribution determination. 134
7.5 Comparison of the vertex distributions determined for the 1998/99 data sample by two minimum bias methods. 135

7.6 Fit of the measured minimum bias vertex distribution for the 1998/99 data sample. 136
7.7 Systematic uncertainties in the z_{VXT} distribution for the 1998/99 data sample. 137
7.8 Minimum bias vertex distribution for the 1999/00 data sample. 137
7.9 Verification of the calorimeter energy scale corrections used in the study of the QCD cascade. 141
7.10 Hadronic energy-flow versus polar angle in bins of the hadronic angle. 142
7.11 Measurement of the hadronic energy-flow in the BCAL for several bins of the hadronic angle. 143
7.12 Measurement of the hadronic energy-flow in fired bins versus the radius on the face of FCAL. 144
7.13 Schematic representation of the forward plug calorimeter. 146
7.14 Measurement of the hadronic energy-flow measured in calorimeter cells in the FCAL. 147
7.15 The effect of the island splitting algorithm on the reconstructed hadronic angle. 148
7.16 Ratio of average deposited energy, in the FCAL and RCAL, in real data and Monte Carlo events. 150
7.17 Ratio of average deposited energy, in the BCAL, in real data and Monte Carlo events. 151
7.18 Hadronic energy-flow versus the ratio $\theta\gamma_h$. 152
7.19 Comparison of the hadronic energy-flow measured in real data to the MC models shown as an error band defined by $\overline{E}_{Ariadne} \pm \mid \overline{E}_{Ariadne} - \overline{E}_{MEPS} \mid$. 153
7.20 Determination of the size of the hadronisation systematic uncertainty. 156
7.21 Hadronic energy-flow measured in 1999/00 data compared to the hadronisation systematic uncertainty. 157

8.1 Efficiency and purity in $d\sigma/dQ^2$ bins. 165
8.2 Efficiency and purity in $d\sigma/dx$ and $d\sigma/dy$ bins. 166
8.3 Efficiency and purity in double differential bins. 167
8.4 Effect of the event migration in double differential bins. 168
8.5 Bin-to-bin uncorrelated systematic errors δ_1^u to δ_{10}^u in $d\sigma/dQ^2$. 172
8.6 Single differential cross section $d\sigma/dQ^2$: bin-to-bin correlated systematic uncertainty sources δ_1^c to δ_6^c. 174
8.7 Total systematic uncertainty in $d\sigma/dQ^2$. 175
8.8 Note on PDF errors. 176
8.9 Single differential cross section $d\sigma/dQ^2$. 178

8.10 Ratios of the single differential cross sections $d\sigma^{e-p}/dQ^2$ and $d\sigma^{e+p}/dQ^2$ to the Standard Model predictions. 179
8.11 Single differential neutral current cross section $d\sigma/dQ^2$ compared to the charged current cross section. 180
8.12 Single differential cross section $d\sigma^{e\pm p}/dx$ for $Q^2 > 200$ GeV^2 and $Q^2 > 10000$ GeV^2. 181
8.13 Ratios of single differential cross section $d\sigma/dx$ to the Standard Model prediction for $Q^2 > 200$ GeV^2 and $Q^2 > 10000$ GeV^2. 183
8.14 Single differential cross section $d\sigma/dy$ for $Q^2 > 200$ GeV^2, and $Q^2 > 10000$ GeV^2 184
8.15 Single differential cross section $d\sigma/dy$ for $Q^2 > 200$ GeV^2, and $Q^2 > 10000$ GeV^2 185
8.16 Double differential reduced cross section $\tilde{\sigma}^{e^\pm p}$ 187
8.17 The F_2^{em} structure function measurements from the e^+p and e^-p data samples. 188
8.18 Ratios of relative systematic uncertainties in the F_2 measurements from the e^-p and the e^+p data samples. 189
8.19 Relative normalisation of the 1998/99 and the 1999/00 data sets. 190
8.20 F_2^{em} measured from the combined e^+p and e^-p data samples. 192
8.21 Double differential reduced cross sections $\tilde{\sigma}^{e^-p}$ and $\tilde{\sigma}^{e^+p}$ measured in the bins used for the extraction of xF_3. 193
8.22 Structure function xF_3 measuremet. 194

9.1 Structure function F_2 measurement compared to most recent measurements from ZEUS and H1 and to fixed-target results from E665, NMC and BCDMS. 196
9.2 Structure function F_2 measurement at high Q^2 compared to most recent measurements from ZEUS and H1. 198
9.3 Structure function xF_3 measurement compared to most recent measurements from ZEUS and H1. 199
9.4 Double differential reduced cross section and the effect of lepton beam polarisation. 201

List of Tables

1.1 fundamental particles: quarks and leptons. 17

2.1 The charge and third component of weak isospin for the interacting leptons and quarks. 26

3.1 Parameters of the HERA collider. 39
3.2 Summary of parameters of the CTD detector. 44

4.1 Summary of non-diffractive Monte Carlo data used in the analysis of e^-p and e^+p events. 63
4.2 Summary of diffractive Monte Carlo data used in the analysis of e^-p and e^+p events. 64
4.3 Summary of the photoproduction Monte Carlo samples. 68

6.1 Alignment of the BCAL and RCAL. 96
6.2 Correction factors for the hadronic energy scale of the different calorimeter sections. 113

7.1 Angular cuts applied to events in the three CTD regions and effective Q^2 limits. 131
7.2 Parameters of the analytic function fitted to the vertex distribution for 1998/99 data. 135

A.1 Table of measured $d\sigma^{e-p}/dQ^2$ values. 214
A.2 Table of measured $d\sigma^{e-p}/dy$ values for $Q^2 > 200$ GeV^2. 215
A.3 Table of measured $d\sigma^{e-p}/dy$ values for $Q^2 > 10000$ GeV^2. 215
A.4 Table of measured $d\sigma^{e-p}/dx$ values for $Q^2 > 200$ GeV^2. 216
A.5 Table of measured $d\sigma^{e-p}/dx$ values for $Q^2 > 10000$ GeV^2. 216
A.6 Table of measured $d\sigma^{e+p}/dQ^2$ values. 217
A.7 Table of measured $d\sigma^{e+p}/dy$ values for $Q^2 > 200$ GeV^2. 218
A.8 Table of measured $d\sigma^{e+p}/dy$ values for $Q^2 > 10000$ GeV^2. 218

A.9 Table of measured $d\sigma^{e+p}/dx$ values for $Q^2 > 10000$ GeV^2. 219
A.10 Table of measured $d\sigma^{e+p}/dx$ values for $Q^2 > 10000$ GeV^2. 219

B.1 Measured values of the double differential reduced cross section $\tilde{\sigma}^{e-p}(x, Q^2)$. 221
B.2 Measured values of the double differential reduced cross section $\tilde{\sigma}^{e-p}(x, Q^2)$. 222
B.3 Measured values of the double differential reduced cross section $\tilde{\sigma}^{e+p}(x, Q^2)$. 223
B.4 Measured values of the double differential reduced cross section $\tilde{\sigma}^{e+p}(x, Q^2)$. 224

C.1 Table with results of the F_2^{em} measurement. 226
C.2 Table with results of the F_2^{em} measurement. 227
C.3 Table of measured xF_3 values. 228

Chapter 1

Introduction

There are interesting times ahead for the study of the fundamental constituents of matter and their interactions. The Standard Model [1] of particle physics has stood undefeated for the last three decades as the most successful and accurate verifiable set of theories that attempt to explain the ultimate structure of the world we live in. Only one fundamental particle predicted by the Standard Model, the Higgs boson, remains to be observed. The combined results of the four experiments at the Large Electron Positron (LEP) collider at CERN have established a lower Higgs-mass limit of 114.4 GeV at the 95% confidence level [2]. In the last year of LEP running, tantalising hints of the direct observation of a Higgs boson with a mass of 116 GeV were found but could not be confirmed. Apart from the yet undetected Higgs boson, decades of experimentation in one of the most organised scientific efforts of all times has found no major inconsistency between theory and observations. Yet, some indications exist that the Standard Model (SM) cannot be the whole truth but must be, at best, an effective model valid up to the limits which could so far be explored. Putting aside conceptual questions such as why there are three generations of quarks and leptons which seem to be identical copies of each other, the breakdown of the SM at the Planck scale, where gravitational effects become important at a microscopic level, indicates that the SM must be incomplete.

Much closer to our present experimental horizon, recent indications that the neutrinos have a small but finite mass [3, 4] raise several questions and indicate that the SM is incomplete. In fact, a finite neutrino mass implies that right-handed neutrinos must exist, which may indicate the need for important changes to the chiral structure of the SM. Other experimental indications of phenomena that don't seem to fit in the theory come from astrophysics. The dynamics of spiral galaxies, for example, indicate that most of the matter in the universe is of an unknown nature, the so-called *dark matter*. Some authors argue that the inferred dark matter is composed of as yet unobserved particles that are predicted by supersymmetric theories [5] of particle physics that go beyond the SM.

Leptons	ν_e	ν_μ	ν_τ
	e^-	μ^-	τ^-
Quarks	u (up)	c (charm)	t (top)
	d (down)	s(strange)	b (beauty)

Table 1.1: fundamental particles: quarks and leptons.

In a few years, the Large Hadron Collider (LHC), will start colliding protons at a centre-of-mass energy of 14 TeV. This machine will open a new domain for exploration. What will be observed in that domain is unknown. The possibility exists that new phenomena will be observed and that, as a result, the SM theories will need to be extensively reviewed or recognised as being an approximation to a more general theory.

Central to the SM is the assumption that all matter is composed of a small set of fundamental pointlike particles of spin $\frac{1}{2}$, shown in table 1.1, which interact through the exchange of a set of gauge bosons (of integral spin). Electrically charged particles interact electromagnetically through the exchange of photons. All particles are sensitive to the weak interaction, described by the exchange of the Z^0 and $W^\pm$ gauge bosons.

Protons and neutrons have been found to be composite objects, being composed of fundamental pointlike particles called quarks. Protons and neutrons are, themselves, the constituents of the nuclei of atoms that compose the bulk of the matter which is visible to us. Quarks inside the proton interact with each other through the exchange of gauge bosons called gluons. This interaction is responsible for the quarks being kept together inside the proton. Quarks possess a new kind of charge which is characterized by a *colour* quantum number. This is the charge of the strong interaction, and is a property of quarks and gluons. There are eight kinds of gluon. As gluons are coloured, they interact both with quarks and with other gluons.

The work described in this thesis focuses on the study of the electroweak interactions and on the determination of the structure of the proton. This structure is studied by probing the proton with a beam of electrons or positrons in a process called deep inelastic scattering (DIS). The work described here was performed within the ZEUS collaboration. The ZEUS experiment is situated in one of the interaction regions of the Hadron-Elektron Ring Anlage (HERA) accelerator in Hamburg. Electrons or positrons and protons are accelerated in the HERA collider and made to interact at the centre of the two general purpose experiments, ZEUS and H1, at a centre-of-mass energy of 318 GeV. These experiments are in an unique position to probe the proton at very small distance scales through the DIS process. The information obtained by probing the structure of the proton is not only important in itself but is also useful to present and future particle physics experiments using beams of protons.

Deep inelastic scattering was observed for the first time in 1969 at the Stanford Linear Accelerator Centre (SLAC) [6, 7]. In the 1960's, the composite nature of the proton was not yet established. In fact, Murray Gell-Mann had only proposed his then controversial quark model in 1964 [8], to explain the observed proliferation of hadrons. In the same way as Rutherford's experiment in which alpha particles were scattered off gold atoms revealing the hard nucleus at its centre, the SLAC experiments showed that hard pointlike objects existed inside nucleons. At around the same time, Feynman [9] introduced a model in which hadrons were composed of an infinite number of fundamental particles which he called *partons*. The two descriptions of the nucleon were harmonised in the quark-parton model, in which Feynman's partons are identified with the quarks. The proton is assumed to be composed of three *valence* quarks (uud) and a *sea* of quark-antiquark pairs that are continually created and annihilated. It was found that the quarks carried only around half of the proton's momentum, and inferred that the other half must be carried by particles with no electric charge. In the 1970's the theory of quantum chromodynamics was proposed, in which the interactions between quarks are transmitted by electrically neutral gluons.

Electrons and positrons interact only with the quarks in the proton and not with the gluons. This interaction can be described by the exchange of photons or Z^0s, which are electrically neutral, or of $W^{\pm}$ which are charged. The processes involving the exchange of a photon or a Z^0 between an electron and a quark are therefore known as neutral current (NC) processes. The exchange of a W^+ or a W^- fall under the category of charged current. Both kinds of process are thus sensitive to the quark content of the proton. In the present work, neutral current interactions between the electron or positron probe and the quarks in the proton were analysed. As this kind of interaction can originate from the exchange of two different types of gauge boson, the photon and the Z^0, the study of this processes also probes the structure of the interaction itself. Furthermore, the interactions of electrons and positrons with the proton are different at high momentum transfer due to the chiral structure of the weak interaction. Using both electron and positron beams, the theoretical predictions on the contribution of photon and the Z^0 to the process can be tested.

The work presented in this thesis has benefited from the experience accumulated in several years of HERA running. The level of precision of the measurements presented here are a reflection of the size of the data samples collected by the first phase of the exploitation of this unique collider and also of the accumulated knowledge of the machine and detectors and of the experimental methods used in the measurements. HERA-II running has now begun, with improvements to the accelerator and detectors which are expected to yield a

five-fold increase to the luminosity and much improved experimental capabilities. Many important results are expected from this new phase in the life of HERA experiments.

The thesis is organised as follows. Chapter 2 introduces the kinematic variables used in the study of DIS and gives an outline of the theory that is necessary for the following chapters. Chapter 3 describes the experimental setup, focussing on the relevant aspects of the HERA collider and on the ZEUS detector. Chapter 4 outlines the Monte Carlo simulation programs that were used extensively in the present analysis to estimate experimental acceptance and resolution, to estimate the background contamination and to extract the cross sections. Chapters 5 to 7 describe the bulk of the issues related to the selection of events. The methods used in the measurement of kinematic variables described in chapter 5. The event reconstruction, the trigger and offline event selection and the corrections applied to quantities measured in the detector are described in chapter 6. Chapter 7 describes two studies of systematic effects which were performed in the context of this thesis. The methods used in the extraction of the cross sections and structure functions are described in chapter 8, where the measurement results are also shown. Finally, in chapter 9, the structure function results are compared to published data and theoretical predictions, and an outlook is given.

The analysis of NC DIS events in ZEUS using data collected from 1998 to 2000 was the work of several people. While the results and conclusions reported here are of my sole responsability, the analysis methods have greatly benefited from the collaboration with the ZEUS neutral current analysis team. Conversely, my major contributions to the work of the analysis team are also reported here, namely the studies described in chapter 7. My analysis also followed, in some cases, different routes from the "official" ZEUS analysis, such as in the choice of the used electron finder algorithm (see chapter 6), and the combination of data sets used to obtain the structure functions F_2 and xF_3, which is described in chapter 8.

Chapter 2

Theory of Deep Inelastic Scattering

Deep inelastic scattering (DIS), denotes the scattering of leptons on nucleons with high momentum transfer. In this process, the lepton exchanges a gauge boson with a quark inside the proton. The exchanged gauge boson has sufficiently short wavelength to resolve a small area inside the nucleon, hence the adjective "deep". The process is inelastic if the proton breaks up as a result of the interaction. This often happens, as the interacting quark is knocked off the proton, and the remaining partons tend to fragment into several hadrons known as the proton remnant. In charged lepton-proton DIS, at Born level, the reaction is described through the exchange of either a charged, virtual gauge-boson ($W^{\pm}$), or a neutral virtual-photon or Z^0. These different processes are called charged current (CC) and neutral current (NC) deep inelastic scattering, respectively.

The work reported here focuses on the study of neutral current $e^{\pm}p$ DIS events. The usual final state topology of DIS events thus consists of a scattered lepton ($e^{\pm}$ in the case of NC DIS or a neutrino in the case of CC events), a proton remnant travelling in the same direction as the proton beam, and one or more jets of particles travelling approximately in the direction of the struck quark. The cross sections for *inclusive* NC $e^{\pm}p$ DIS have been measured and are presented in chapter 8. All possible hadronic final states contribute to the inclusive cross section. Structure functions, which characterise the momentum distribution of the proton constituents, have also been extracted from the measured cross sections.

Natural units, which correspond to the relation $\hbar = c = 1$ between the reduced Planck constant, $\hbar$, and the speed of light, c, will be used throughout the present thesis. For simplicity, unless stated otherwise, the incoming and scattered lepton will be referred to as electrons even though both $e^{-}p$ and $e^{+}p$ events were analysed.

2.1 Kinematics of deep inelastic scattering

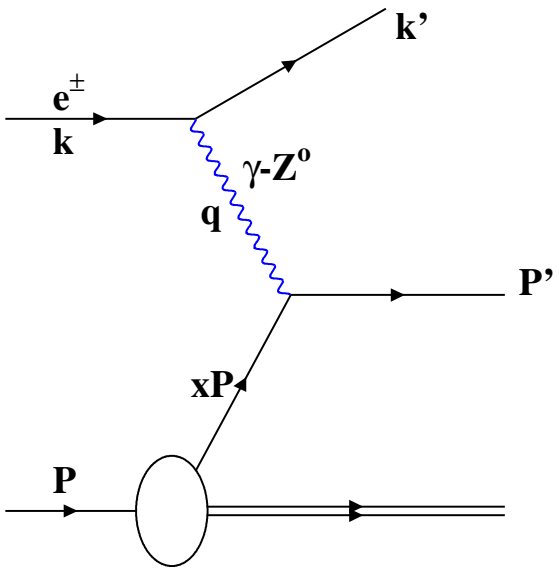

Figure 2.1: Feynman diagram of the DIS process at lowest electroweak order.

Figure 2.1 schematically represents the process of neutral current DIS at lowest perturbative order, also known as the Born level. The exchanged gauge boson (a virtual photon or Z^0) has four-momentum $q = k - k'$, where k and k' are the four-momenta of the incoming and scattered lepton. P is the four-momentum of the incoming proton. The following Lorentz-invariant variables are usually employed in the description of DIS:

$$Q^2 = -q^2 = (k - k')^2 \tag{2.1}$$

$$s = (k + P)^2 \tag{2.2}$$

$$x = \frac{Q^2}{2P.q} \tag{2.3}$$

$$y = \frac{P.q}{P.k} \tag{2.4}$$

$$W^2 = (P + q)^2 \tag{2.5}$$

Q^2 is the virtuality of the exchanged boson, i.e. the (negative of the) invariant-mass squared of the exchanged virtual-boson. It provides a measure of the length scale, Δx, at which the proton is being probed by a photon carrying momentum Δp. The Heisenberg uncertainty relation provides a relation between Q^2 and the wavelength, λ, of the virtual photon exchanged in the interaction, according to:

$$\Delta p \cdot \Delta x \approx \hbar \Leftrightarrow \sqrt{Q^2} \approx \frac{\hbar}{\lambda} \tag{2.6}$$

At high Q^2, the proton is probed at very small distance scales by a short wavelength photon. This is shown schematically in figure 2.2. The shortest scale within the reach of HERA experiments is around 10^{-18} m [10, 11], or three orders of magnitude smaller than the size of the proton itself, corresponding to the upper limit of the four momentum transfer. This limit is set by $Q^2 = s$, where s is the square of the lepton-proton centre-of-mass energy. With the beam energies of $E_e^{beam} = 27.5$ GeV and $E_p^{beam} = 920$ GeV for electrons and protons stored in HERA, respectively, the centre of mass energy is $\sqrt{s} = 318$ GeV (see equation 2.2).

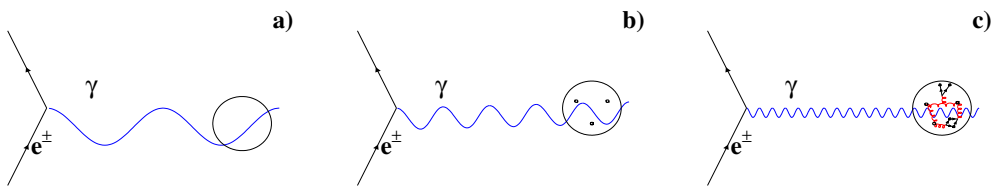

Figure 2.2: Schematic representation of $e^{\pm}p$ scattering at increasing scales of four-momentum transfer (Q^2) or, equivalently, at shorter distance scales. In figure **a)**, the virtual photon interacts with the whole proton. As the wavelength of the photon probing the proton decreases, the proton starts to behave as a collection of structureless particles (figure **b)**). For even shorter wavelengths, more particles become visible inside the proton, which are created by quantum fluctuations.

The Bjorken *scaling* variable, x, defined by equation 2.3, can be interpreted in the quark-parton model discussed in section 2.2 below, as the fraction of the momentum of the proton that is carried by the interacting quark. It can thus take values between 0 and 1. Given the electron- and proton-beam momenta in HERA, the centre-of-mass frame of the electron-quark interaction coincides with the laboratory reference frame for $x = 27.5/920 \approx 0.03$. For higher values of x, the centre of mass of the interaction is boosted in the direction of the moving proton, and vice versa for lower x.

In the proton rest frame the *inelasticity*, y, corresponds to the fraction of the energy of the incoming electron that is transferred to the proton. y takes values between 0 and 1. In the electron quark centre-of-mass frame, y is related to the electron scattering angle, θ^*, by:

$$1 - y = \frac{1 + \cos\theta^*}{2} \tag{2.7}$$

W is the invariant mass of the hadronic final state. W can vary between the mass of the proton, m_p, and the centre of mass energy, $\sqrt{s}$.

Not all of these variables are independent. s is fixed by the beam energies, $s = 101200$ GeV^2. The variables, Q^2, x, and y are related through equation 2.8, where m_e and m_p are the electron and proton masses, respectively.

$$Q^2 = xy(s - m_p^2 - m_e^2) \approx sxy \tag{2.8}$$

The approximation is very good due to the large value of s. Another relation exists between W^2, x and Q^2:

$$W^2 = Q^2 \frac{1-x}{x} + m_p^2 \tag{2.9}$$

Only two independent variables remain. These can be used to describe DIS events completely. A common choice for NC DIS is to use the x and Q^2 variables to classify events in phase space. The kinematic region within the reach of ZEUS is shown in figure 2.3 and compared to that of fixed target experiments. The present work is concerned with the measurement of the NC DIS cross section for $Q^2 > 200$ GeV^2.

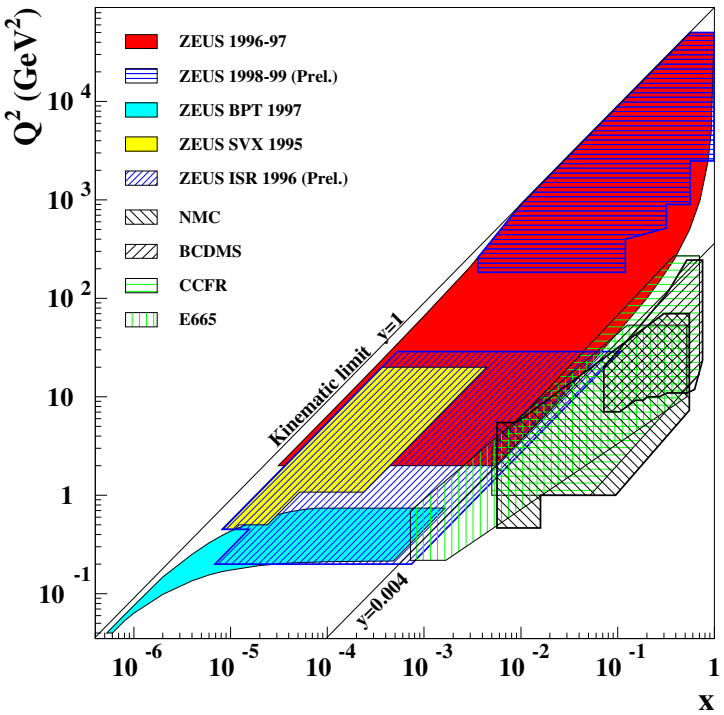

Figure 2.3: DIS phase space covered by ZEUS, compared to the reach of older fixed-target experiments.

2.2 The quark-parton model

In the naive quark-parton model (QPM), the proton is considered to be composed of free, pointlike partons. The approximation of free partons can be understood in the scattering of leptons off protons by considering that in the reference frame where the proton has infinite momentum, the proton is relativistically contracted into a thin disk and at the same time the rate of interaction between its constituents is slowed down by the relativistic time dilation [12]. The lepton-proton interaction takes place on a much shorter time-scale than that of the interactions between the quarks in the proton which are then approximately free. In this picture, the lepton-proton interaction corresponds to the elastic scattering of a lepton by a single parton carrying longitudinal momentum xP,

where P is the momentum of the proton, and negligible transverse momentum and mass. The naive-QPM partons correspond to the quarks postulated by Gell-Mann [8] and Zweig. Three *valence* quarks (uud) give the proton its flavour and charge properties. In addition, a *sea* of quark-antiquark pairs, with no net charge or flavour, exist in the proton. The antiquark content of the proton thus belongs only to the sea.

As the partons do not interact with each other, in the naive QPM approximation, the distributions of quarks in the proton, known as *parton density functions* (PDFs), can depend only on x and not on the distance scale at which the proton is probed, given by Q^2. This was predicted by Bjorken [13, 12] and observed in early DIS experiments [6, 7], confirming the assumption that the proton is made of almost free pointlike constituents. This independence of Q^2 is known as *"scaling"*. This behaviour is approximately verified for the structure function F_2 (see section 2.3), for example, at values of x close to $x \approx 0.2$, as is shown in figure 9.1 of chapter 9.

Once interactions between quarks are considered, which are described in quantum chromodynamics (QCD) by the exchange of gluons, the radiation of gluons by quarks and the splitting of gluons into quark-antiquark pairs must be taken into account. In this more realistic picture, the proton becomes a dynamic system with the sea quarks being generated by the splitting of gluons, and the quarks and antiquarks radiating other gluons. The partons, the bare charges of the theory, now include both quarks and gluons. The total number of partons now depends on the scale at which the proton is probed, as in the case of quantum electrodynamics, in which the bare electron charge is surrounded by a cloud of virtual charges. As Q^2 increases, the density of valence quarks falls due to the radiation of gluons, which must carry some of the quark momentum. At high x, where the cross section is sensitive mainly to the valence quarks, this results in a decrease in the cross section compared to the QPM prediction, and thus in a violation of the scaling behaviour. At low x, on the other hand, the cross section is dominated by the quark-antiquark pairs in the sea, produced by the splitting of gluons. At small distance scales, or high Q^2, the number of $q\overline{q}$ pairs increases, leading to an increase in the cross section and again to scaling violation. The behaviour of the parton density functions as a function of Q^2 is given a more formal basis in section 2.4.1.

2.3 Neutral current cross section and structure functions

The neutral current cross section for deep inelastic scattering can be obtained from an incoherent sum over electron-quark scattering cross sections weighted by the quark charge and momentum. For unpolarised charged-lepton-proton scattering due to the exchange of a virtual photon or Z^0 the cross section can be written in a general form as [14]:

$$\frac{d^2\sigma_{Born}(e^{\pm}p)}{dxdQ^2} = \frac{2\pi\alpha^2}{xQ^4}\left[Y_+F_2(x,Q^2) - y^2F_L(x,Q^2) \mp Y_-xF_3(x,Q^2)\right] \quad (2.10)$$

where x, y and Q^2 have already been defined, $Y_+ = (1 + (1-y)^2)$, $Y_- = (1 - (1-y)^2)$, $\alpha = e^2/4\pi\hbar c$ is the fine structure constant, and F_2, F_3 and F_L are *structure functions* of the proton.

At low Q^2, the weak contribution to the cross section can be neglected, the interaction occurring only through virtual photon exchange. In this case, the structure function xF_3 is zero and F_2 is given by:

$$F_2(x, Q^2) = \sum_q e_q^2 \left(xq(x, Q^2) + x\overline{q}(x, Q^2)\right) \tag{2.11}$$

where the sum runs over all quark flavours q in the proton, e_q^2 is the quark charge in units of the positron charge, and $q(x, Q^2)$ and $\overline{q}(x, Q^2)$ correspond to the number density of quarks q and antiquarks $\overline{q}$ in the proton. Thus, $xq(x, Q^2)$ and $x\overline{q}(x, Q^2)$ give the momentum density functions of quarks and antiquarks. Figure 2.4 shows some parton density functions as a function of x. The F_2 structure function is proportional to the total quark content of the proton, including valence and sea quarks.

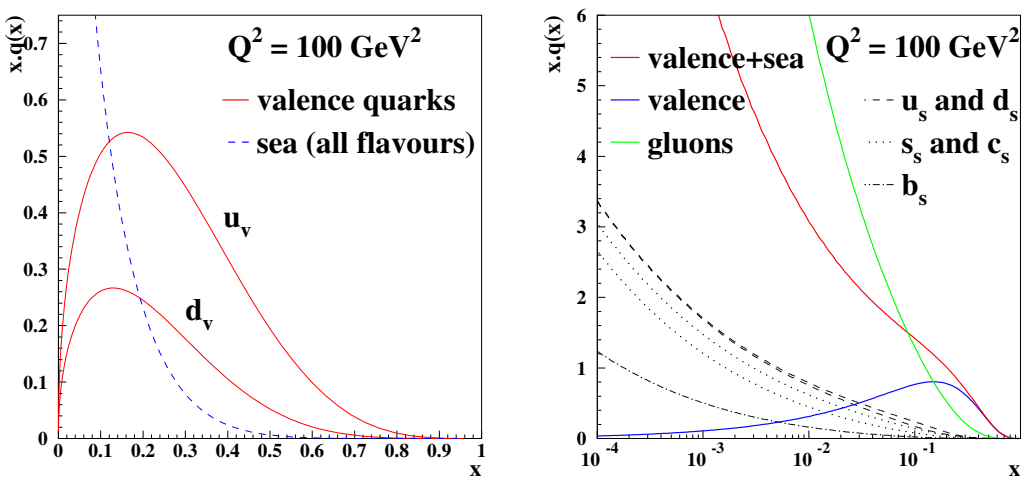

Figure 2.4: Quark and gluon density functions at $Q^2 = 100$ GeV^2.

The F_2 term parametrises the transverse part of the cross section, in which a transversely polarised photon is exchanged. The longitudinal cross section depends on the F_L structure function. In the QPM, the interacting quark has a negligible mass, and the longitudinal cross section must be zero due to helicity conservation. Extending the quark-parton model to include QCD changes this picture by allowing the radiation of gluons by the interacting quark. This quark becomes virtual and may acquire a non-negligible mass, and is no longer an eigenstate of helicity. It may then absorb a longitudinally polarized photon, giving rise to a non-zero longitudinal cross section. The F_L term parametrises

this part of the cross section and becomes non-zero. The F_L term is still suppressed by a factor y^2, and constitutes a negligible contribution to the cross section for $y \lesssim 0.5$.

As Q^2 approaches the square of the mass of the Z^0 boson, the weak contribution to the cross section becomes significant. The coupling of the fermions to the currents now depend on the polarisation of the lepton beam (L or R, for left or right polarisation, respectively). This dependence is encoded in functions $A_q^{L,R}(Q^2)$ and $B_q^{L,R}(Q^2)$ of Q^2, which contain the vector and axial vector couplings as parameters. F_2 and xF_3 including the weak interaction can be written as [14]:

$$F_2(x,Q^2) = \sum_q A_q^{L,R}(Q^2)\,(xq(x,Q^2) + x\overline{q}(x,Q^2)) \tag{2.12}$$

$$xF_3(x,Q^2) = \sum_q B_q^{L,R}(Q^2)\,(xq(x,Q^2) - x\overline{q}(x,Q^2)) \tag{2.13}$$

For (left- or right-handed) electrons interacting with a quark flavour q, the $A_q^{L,R}(Q^2)$ and $B_q^{L,R}(Q^2)$ functions are given by:

$$A_q^{L,R}(Q^2) = e_q^2 - 2e_e e_q (v_e \pm a_e) v_q \chi_Z(Q^2) + (v_e \pm a_e)^2 (v_q^2 + a_q^2)\chi_Z^2(Q^2) \tag{2.14}$$

$$B_q^{L,R}(Q^2) = \mp 2e_e e_q (v_e \pm a_e) a_q \chi_Z(Q^2) \pm 2(v_e \pm a_e)^2 v_q a_q \chi_Z^2(Q^2) \tag{2.15}$$

where e_e and e_q are the electron and quark charges in units of the positron charge, and v and a are the electroweak vector and axial-vector couplings of quarks and lepton. The upper signs correspond to left-handed leptons and the lower signs to right-handed leptons. The term χ_Z is proportional to the ratio of the Z^0 and photon propagators:

$$\chi_Z(Q^2) = \frac{1}{4\sin^2\theta_W \cos^2\theta_W} \frac{Q^2}{Q^2 + M_Z^2} \tag{2.16}$$

where M_Z is the mass of the Z^0 and θ_W is the weak mixing angle.

The electroweak couplings can be written for any fermion f in terms of the electric charge, e_f, the third component of weak isospin, T_3^f (see table 2.1), and the weak mixing angle:

$$v_f = T_3^f - 2e_f \sin^2\theta_W \quad \text{and} \quad a_f = T_3^f \tag{2.17}$$

leptons	e_e	T_3^e	quarks	e_q	T_3^q
e_L^-	-1	$-\frac{1}{2}$	u_L, c_L, t_L	$\frac{2}{3}$	$\frac{1}{2}$
e_R^-	-1	0	u_R, c_R, t_R	$\frac{2}{3}$	0
e_L^+	1	$\frac{1}{2}$	d_L, s_L, b_L	$-\frac{1}{3}$	$-\frac{1}{2}$
e_R^+	1	0	d_R, s_R, b_R	$-\frac{1}{3}$	0

Table 2.1: The charge and third component of weak isospin for the interacting leptons and quarks.

The structure function xF_3 is proportional to the difference of quark and antiquark density functions in the proton, i.e. it is proportional to the *valence* quark content of the proton, as the sea quark distributions must come from quark-antiquark pair production.

For positrons interacting with quark flavour q, functions $A_q(Q^2)$ and $B_q(Q^2)$ are obtained from expressions 2.14 and 2.15 by interchanging the L and R labels.

Interactions of leptons with left- and right-handed quarks are taken into account in expressions 2.14 and 2.15, considering that the proton beam is not polarised and therefore contains as many left-handed as right-handed quark states. The remaining dependence on polarisation is only on that of the lepton beam. When the parity operator is applied to the cross section, $B_q^L(Q^2)$ transforms into $B_q^R(Q^2) = -B_q^L(Q^2)$. This happens because the parity violating terms in the cross section were grouped into the xF_3 term.

The NC DIS process was studied in the present work with unpolarised e^+ and e^- beams and with unpolarised proton beams. To understand the origin of different sign of the xF_3 term in the expressions for the e^-p and the e^+p cross sections (equation 2.10), one may consider as an example the interaction between a left-handed electron and a left-handed quark, $e_L^- q_L$. Applying the parity operator, P, to the interaction gives $P(e_L^- q_L) = e_R^- q_R$. Recall that applying P to the $e_L^- p$ cross sections, i.e., to the sum of all possible combinations $e_L^- q_L$, $e_L^- q_R$, $e_L^- \overline{q}_L$ etc., results in a relative minus sign in the xF_3 term. Now, applying the combined charge conjugation, C, and parity operators to the electron-quark interaction gives: $CP(e_L^- q_L) = C(e_R^- q_R) = e_R^+ \overline{q}_R$. As the neutral current interaction is invariant under the CP operation, the cross sections for $e_L^- q_L$ and $e_R^+ \overline{q}_R$ must be the same. In particular, the xF_3 term does not change under the transformation. Now, when the CP transformation is applied to the initial e^-p state it also changes the proton into an antiproton. It can be assumed that the antiquark distributions in antiprotons are the same as the quark distributions in protons. So, the right-handed antiquark PDF in an antiproton, which can be denoted $\overline{q}_R^{\overline{p}}$, is the same as the right-handed quark distribution in the proton: $\overline{q}_R^{\overline{p}} = q_R^p$, where q_R^p now stands for the right-handed quark PDF in the proton. No difference should be found between the density of right-handed and left-handed quarks in unpolarised protons. Applying P a second time, however, changes the right-handed positron and quark into left-handed particles: $P(CP(e_L^- q_L)) = P(e_R^+ \overline{q}_R^{\overline{p}}) = e_L^+ q_L^p$. The NC DIS cross section for e^+p can now be computed by summing all contributions from $e_L^+ q_L$, $e_L^+ q_R$, $e_L^+ \overline{q}_L$ etc. It can also be obtained by applying the PCP operation to the e^-p cross section and exchanging the quark PDFs for antiquark PDFs and vice versa. From expression 2.13, it becomes clear that this results in a negative sign for the xF_3 term of the NC cross section [15]. This reasoning is represented schematically above.

e^-p	$\to CP \to$	$e^+\overline{p}$	$\to P \to$	e^+p
$e^-_L q$	$\to$	$e^+_R \overline{q}_{\overline{p}} = e^+_R q_p$	$\to$	$e^+_L q_p$
$e^-_R q$	$\to$	$e^+_L \overline{q}_{\overline{p}} = e^+_L q_p$	$\to$	$e^+_R q_p$
$e^-_L \overline{q}$	$\to$	$e^+_R q_{\overline{p}} = e^+_R \overline{q}_p$	$\to$	$e^+_L \overline{q}_p$
$e^-_R \overline{q}$	$\to$	$e^+_L q_{\overline{p}} = e^+_L \overline{q}_p$	$\to$	$e^+_R \overline{q}_p$
xF_3	$\to$	xF_3	$\to$	$-xF_3$

To obtain the unpolarised cross sections measured in ZEUS, one must average over the left- and right-handed functions $A_q^{L,R}(Q^2)$ and $B_q^{L,R}(Q^2)$. This yields the functions:

$$\mathcal{A}_q(Q^2) = e_q^2 - 2e_e e_q v_e v_q \chi_Z + (v_e^2 + a_e^2)(v_q^2 + a_q^2)\chi_Z^2 \quad (2.18)$$

$$\mathcal{B}_q(Q^2) = -2e_e e_q a_e a_q \chi_Z + 4 v_q a_q v_e a_e \chi_Z^2 \quad (2.19)$$

And the structure functions can be written in the more usual form:

$$F_2(x,Q^2) = \sum_q \mathcal{A}_q(Q^2)\left(xq(x,Q^2) + x\overline{q}(x,Q^2)\right) \quad (2.20)$$

$$xF_3(x,Q^2) = \sum_q \mathcal{B}_q(Q^2)\left(xq(x,Q^2) - x\overline{q}(x,Q^2)\right) \quad (2.21)$$

The structure functions can now be written as a sum of terms proportional to different powers of χ_Z. This decomposes the structure functions into the contributions from pure photon exchange (denoted as em below), pure Z^0 exchange, and the interference between those contributions (γ/Z^0):

$$F_2^{NC}(x,Q^2) = F_2^{em} + \frac{Q^2}{Q^2+M_Z^2}F_2^{\gamma/Z} + \left(\frac{Q^2}{Q^2+M_Z^2}\right)^2 F_2^Z \quad (2.22)$$

$$xF_3^{NC}(x,Q^2) = \frac{Q^2}{Q^2+M_Z^2}xF_3^{\gamma/Z} + \left(\frac{Q^2}{Q^2+M_Z^2}\right)^2 xF_3^Z \quad (2.23)$$

F_2^{em} is clearly the largest contribution to F_2^{NC}. The contribution from γ^*/Z^0 interference is suppressed both by the ratio $Q^2/(Q^2+M_Z^2)$, which is small for $Q^2 \ll M_Z^2$, and by the v_e coupling ($v_e = 0.03760$ for $\sin^2\theta_W(M_Z) = 0.2312$ [16] in the $\overline{\text{MS}}$ renormalisation scheme). The contribution from pure Z^0 exchange is suppressed by $Q^4/(Q^2+M_Z^2)^2$, and is also small up to very high Q^2. For xF_3, the largest contribution comes from the interference term between the photon and the Z^0 exchange diagrams, since the latter is suppressed by $Q^4/(Q^2+M_Z^2)^2$. Figure 2.5 shows the first measurement of F_2^{em} that was published by the ZEUS collaboration [17].

2.4 The quark-parton model and QCD

The naive quark-parton model described in section 2.2 does not take into account the interactions between quarks described by quantum chromodynamics. When QCD is considered, the QPM is modified to accomodate the emission of gluons by quarks and the

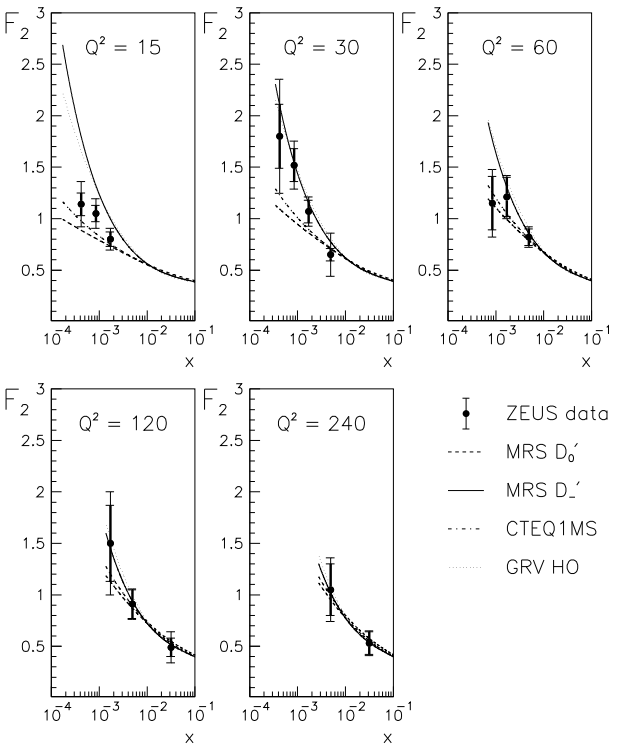

Figure 2.5: First measurement of the F_2 structure function in ZEUS [17].

splitting of gluons into $q\overline{q}$ or gluon pairs. Figure 2.6 **a)** corresponds to the QPM description of DIS. Considering QCD, new diagrams must be added to the QPM cross section. Two of these, corresponding to diagrams of order α_s, or leading order (LO), are represented in figures 2.6 **b)** and **c)**.

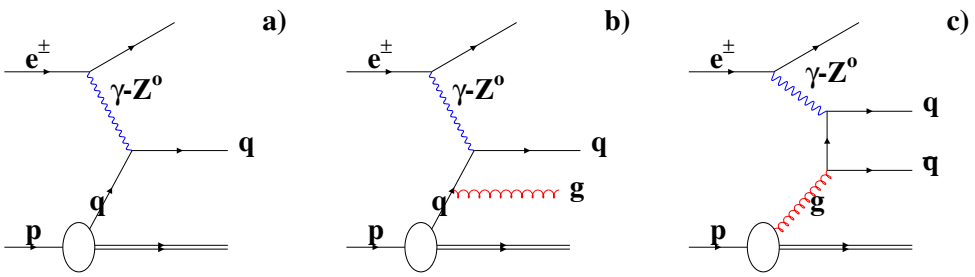

Figure 2.6: Feynman diagram of the NC DIS process in the quark-parton model of the proton (**a**)), and contributions from gluon radiation (**b**)) and $q\overline{q}$ pair production from gluon splitting (**c**)) at leading order in QCD.

Considering the diagram in figure 2.6 **b)**, for example, the struck quark, carrying a fraction x of the proton momentum, originates from a quark carrying momentum fraction η such that $x < \eta < 1$. All quark distributions $q(\eta, Q^2)$ with $\eta > x$ contribute to the process. In the same way, the gluon momentum distribution $g(\eta, Q^2)$ such that $x < \eta < 1$

contributes to the process represented by the diagram in figure 2.6 **c)**. As Q^2 increases, the distance scale at which the proton is probed decreases and more such quantum fluctuations may give rise to the struck quark. The number of $q\overline{q}$ pairs and of gluons in the parton sea thus increases with Q^2. While at low Q^2 and high x the valence quarks carry most of the momentum of the proton, at high Q^2 that momentum is shared with sea partons and the fraction of the momentum carried by the valence quarks becomes smaller. The PDFs are said to *evolve* with Q^2. This is shown in figure 2.7, where the sum of the valence quark PDFs is represented as a function of x for different values of Q^2. As Q^2 grows, the F_2 structure function which is sensitive to both sea and valence quark distributions, increases at low x and decreases at high x. Thus, the scaling behaviour predicted by the naive quark-parton model is violated when QCD is taken into account.

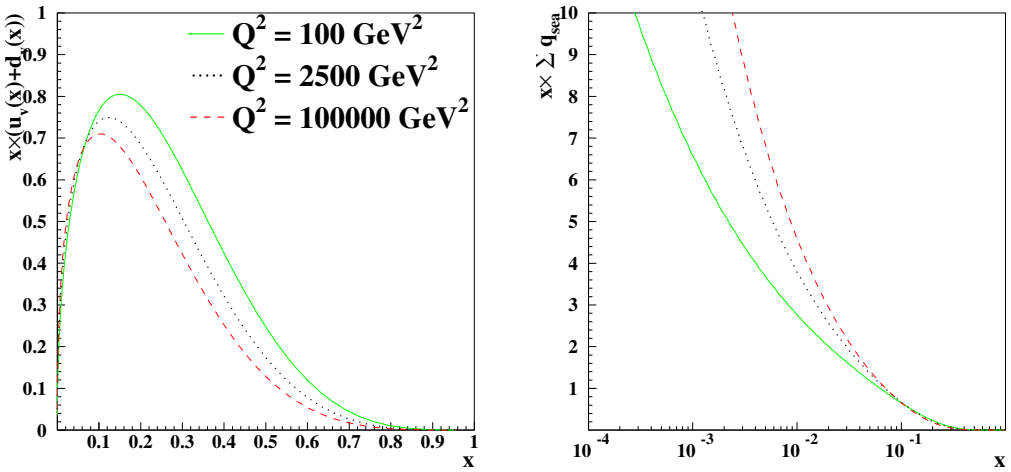

Figure 2.7: Sum of the valence (left) and sea (right) quark parton density functions plotted as a function of x for $Q^2 = 100$ GeV^2, $Q^2 = 2500$ GeV^2, and at $Q^2 = 100000$ GeV^2.

2.4.1 The DGLAP equations

The evolution of the quark and gluon PDFs, $q_i(x, Q^2)$ and $g_i(x, Q^2)$, respectively, is expressed by the Dokshitser-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [18, 19, 20, 21]. Considering n_f quark flavours in the proton, the DGLAP equations are a set of $2n_f+1$ coupled integro-differential equations given by:

$$\frac{d\,q_i(x,Q^2)}{d\,\ln Q^2} = \frac{\alpha_s(Q^2)}{2\pi}\int_x^1 \frac{d\eta}{\eta}\left[\sum_j q_j(\eta,Q^2)\,P_{q_iq_j}(\frac{x}{\eta}) + g_j(\eta,Q^2)\,P_{q_ig}(\frac{x}{\eta})\right] \quad (2.24)$$

$$\frac{d\,g(x,Q^2)}{d\,\ln Q^2} = \frac{\alpha_s(Q^2)}{2\pi}\int_x^1 \frac{d\eta}{\eta}\left[\sum_j q_j(\eta,Q^2)\,P_{gq_j}(\frac{x}{\eta}) + g_j(\eta,Q^2)\,P_{gg}(\frac{x}{\eta})\right] \quad (2.25)$$

where the functions $P_{ij}(z)$, known as *splitting functions*, represent the probability that a parton j (quark or gluon) carrying a fraction η of the proton's momentum emits a parton i, with a fraction x of the proton's momentum, so that i carries a fraction $z = x/\eta$ of the momentum of j.

The splitting functions can be calculated at order α_s, α_s^2, etc. At leading QCD order (LO), including the diagrams shown in figure 2.6 **b)** and **c)**, the splitting functions are given by:

$$\begin{aligned} P_{qq}(z) &= \frac{4}{3}\frac{1+z^2}{1-z} \qquad & P_{qg}(z) &= \frac{1}{2}(z^2+(1-z)^2) \\ P_{gq}(z) &= \frac{4}{3}\frac{1+(1-z)^2}{z} \qquad & P_{gg}(z) &= 6\left(\frac{z}{1-z}+\frac{1-z}{z}+z(1-z)\right) \end{aligned} \tag{2.26}$$

At next-to-leading order (NLO), the struck quark may originate from one extra splitting of partons. The PDFs then become dependent on the order in α_s at which they are defined. Also, different renormalisation schemes may be used. Two such schemes are common in DIS, the *modified minimal subtraction* ($\overline{\text{MS}}$) scheme and the *deep inelastic scattering* (DIS) scheme. The DIS scheme is defined such that F_2 is given by equation 2.20 to all orders. This scheme makes the physical interpretation of the F_2 more intuitive. At LO both the $\overline{\text{MS}}$ and the DIS renormalisation schemes give identical results, but they become different at NLO. Physical observables such as the cross sections and structure functions are independent of the renormalisation scheme. However, the PDFs and the definitions of structure functions do depend on the scheme used.

2.5 Sum rules

Summing over all the quark density functions must yield the quantum numbers of the proton. This was one of the early experimental verifications of the quark-parton model. In particular, two net u-quarks and one d-quark should be obtained:

$$\int_0^1 u_v dx = 2 \quad \text{and} \quad \int_0^1 d_v dx = 1 \tag{2.27}$$

where u_v and d_v are valence quarks number densities. This leads, for example, to the Gross-Llewellyn-Smith sum rule [22], which was verified in neutrino-nucleon DIS:

$$\int_0^1 F_3^{\nu N} dx = \int_0^1 (u_v + d_v) dx = 3 \tag{2.28}$$

For charged-lepton nucleon DIS, the Gottfried sum rule was approximately verified in early data. It is given by:

$$\int_0^1 (F_2^{lp} - F_2^{ln})\frac{dx}{x} \approx \frac{1}{3}\int_0^1 (u_v - d_v) dx = \frac{1}{3} \tag{2.29}$$

where F_2^{ln} is F_2 obtained from charged lepton-neutron DIS. This sum rule was found to be violated because it assumed equal densities of different flavours in the partonic sea. In fact, the $\bar{d}$-antiquark contribution has been found to be larger than the $\bar{u}$ contribution for $x \lesssim 0.2$ [23].

Finally, the momentum sum rule can written as:

$$\int_0^1 xg + x\sum_q (q + \overline{q})dx = 1 \tag{2.30}$$

where the sum runs over all valence and sea quark-flavours and xg is the gluon PDF. It was found that only about 50% of the proton momentum was carried by quarks or antiquarks, the rest being carried by gluons.

2.6 Current parametrisations of parton density functions

In this section, the parametrisations of the PDFs which were used in the present work, either in the simulation of events or in comparisons of the experimental results to the expectations from the Standard Model, are briefly described. Figure 2.8 shows a comparison between some of the PDFs that are described in the following sections.

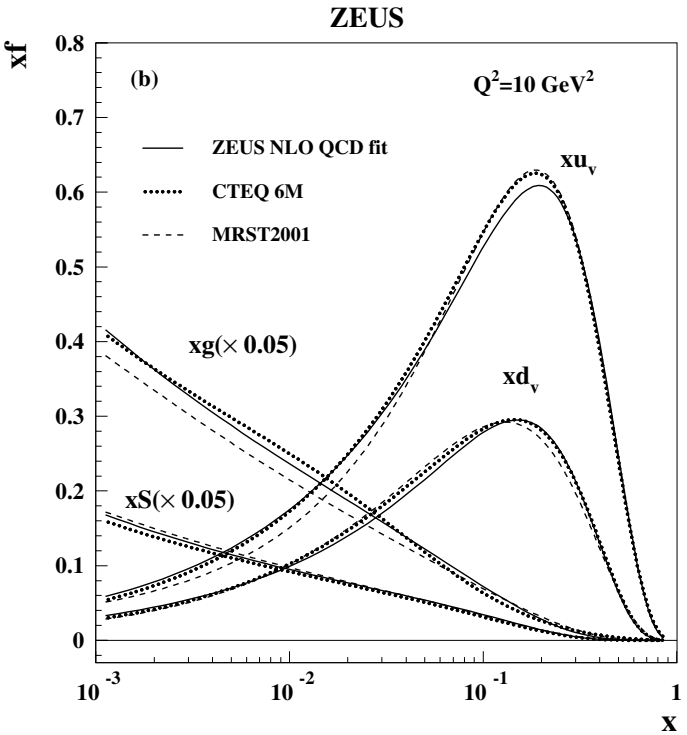

Figure 2.8: Parton density functions from the ZEUS-S set at $Q^2 = 10$ GeV^2 compared to CTEQ6 and MRST2001 PDFs. Reproduced from [24].

Parton density functions are usually determined from global fits to structure function or cross section measurements from various processes and experiments, assuming a functional form for the PDFs and employing the evolution equations. DIS experimental results naturally play an important role in constraining the PDFs, but other types of measurements are also used. The global analyses that yield the PDF sets also constitute stringent tests of QCD. Furthermore, the PDF sets are subsequently employed in the simulation and analysis of various types of interaction and often constitute one of the important sources of experimental uncertainty.

The global fits to the experimental data start from defining the PDFs as an analytic function of x that is assumed to be valid at some starting scale $Q^2 = Q_0^2$, where α_s is small enough to allow the use of perturbative QCD. The DGLAP equations are then used to evolve the PDFs to the Q^2 of the measurements and used to calculate the measured quantities. The calculations are then fitted to the experimental measurements [14]. Sum rules are used to constrain the distributions.

The u-quark PDF is the most accurately known. NC DIS data from several experiments have contributed to this, as the largest contribution to the cross section is from photon exchange, which couples to the square of the quark charge. The d-quark density is less well constrained by the data, especially at high x, where the existing data does not strongly constrain the ratio of the u and d PDFs. The gluon PDF is the least well known.

2.6.1 CTEQ

The Coordinated Theoretical-Experimental Project on QCD (CTEQ) has produced several determinations of the parton density functions. In the present analysis, the CTEQ5D set of PDFs, obtained in the DIS renormalisation scheme, was used for the simulation of NC DIS data and the CTEQ5-LO, obtained at leading-order in QCD, was used for the simulation of the photoproduction background (see chapter 4). Comparisons are made in later chapters between the SM predictions calculated using both the CTEQ5 and CTEQ6 sets and the measured cross sections. The parametrisation of the PDFs at the starting scale Q_0^2 used in CTEQ fits is given by:

$$xf(x) = p_1 x^{p_2} (1-x)^{p_3} (1 + p_4 x)^{p_5} \tag{2.31}$$

where xf is a quark or gluon PDF and p_1 to p_5 are parameters to be determined from the fit.

CTEQ5

The CTEQ5 [25] PDF set includes DIS data from H1 and ZEUS as well as from fixed-target experiments and $p\overline{p}$ data from the DØ and CDF experiments. The CTEQ5D PDF set used in the event simulation and analysis was performed at NLO in QCD. The HERA measurements included in the fit were made using data collected in 1994 [26, 27], and shifted vertex data from 1994 [28] and 1995 [29, 30], to constrain the region of low Q^2 and low x. Drell-Yan [31] lepton pair production data from the E866 [23] fixed-target experiment, was used to constrain the ratio $\overline{u}/\overline{d}$ of the antiquark PDFs in the range $0.03 < x < 0.35$.

CTEQ6

In addition to data used in the CTEQ5 fit, some of which were re-analysed, the CTEQ6 [32] PDF set includes the most recent published F_2 measurements from HERA, consisting of ZEUS [33] measurements from the 1996/97 running period and H1 [34] combined measurements from 1994 to 1997. The HERA data constrain the fit in the x region of $10^{-5} < x < 0.1$. Also included are inclusive jet production measurements from DØ [35]. This data set, together with precision DIS data, constrains the behaviour of the gluon PDF at high-x (for $x > 0.01$). The gluon PDF in CTEQ6 sets is significantly higher at high x than in the CTEQ5 PDF sets.

The CTEQ6 global fit is different from the previous CTEQ5 fit also in the treatment of uncertainties. Correlated uncertainties are fully integrated into the treatment of experimental errors. Also, a set of eigenvectors is supplied, which describes the behaviour of the χ^2 of the fit in a 20-dimensional space of PDF parameters. This allows a precise treatment of the uncertainties in the measurements arising from the PDF parametrisation, when this set is used to simulate and analyse the observed data. This method was not used in the present analysis (see chapter 8), as the CTEQ5D set was used in the event simulation.

2.6.2 MRST2001

The Martin-Roberts-Stirling-Thorne [36] (MRST) group has produced several generations of PDF sets. The MRST2001 parton density functions set was used in this analysis to make comparisons between measured quantities and the expectations from the Standard Model.

The experimental data set used in the MRST2001 NLO fit was similar to that used in the CTEQ6 fit. Regarding HERA measurements, high-Q^2 charged-current data from H1 [37] and ZEUS [38] from the running period of 1994 to 1997 were used, in addition to the F_2 measurements as used in the CTEQ6 fit. Measurements of the c-quark contribution

to F_2 (usually written $F_2^{c\overline{c}}$) made by H1 [39] and ZEUS [40] by selecting events containing a $D^{*\pm}$ meson were also used in the MRST2001 fit.

The initial PDF parametrisation was:

$$xf(x) = p_1 x^{p_2}(1-x)^{p_3}(1+p_4\sqrt{x}+p_5 x) \quad (2.32)$$

$$xG(x) = g_1 x^{g_2}(1-x)^{g_3}(1+g_4\sqrt{x}+g_5 x)+g_6 x^{g_7}(1-x)^{10} \quad (2.33)$$

where xf corresponds to a quark or antiquark PDF and xG is the gluon PDF. The additional freedom in the gluon parametrisation was necessary in order to fit HERA data. However, as a consequence, the gluon PDF at the starting scale becomes negative for $x < 10^{-5}$ and $Q^2 < 5$ GeV^2. This is in principle allowed by the theory, as the gluon distribution is not a physical observable. At high x, data from the Tevatron experiments [35, 41] had a large impact, as in the case of the CTEQ6 fit. The gluon PDF increased at around $x \approx 0.1$ and decreased sharply at $x \approx 0.3$ as compared to the previous MRST PDF fit.

The MRST group has recently published an update of their PDF sets, MRST2002 [42]. The new set contains a gluon PDF which is closer to the CTEQ6 gluon at high x. Also following the CTEQ6 procedure, the MRST group provides a set of eigenvectors that may be used to determine the PDF errors in physical quantities. The MRST2002 set was not used in the present work.

2.6.3 GRV

The Glück-Reya-Vogt [43] (GRV) group calculated the parton density functions in the photon (i.e. created by the fluctuation of the photon into an hadronic state). They have assumed valence-like input density functions xf^γ for both quarks (and antiquarks) and gluons given at an initial scale of 0.25GeV^2 by functions of the form:

$$xf(x) = K 4\pi\alpha x^{p_1}(1-x)^{p_2} \quad (2.34)$$

The LO set of PDFs released by the GRV group in 1994, were used in the simulation of the photoproduction background (see chapter 4)

2.6.4 ZEUS NLO QCD fit

A fit to the most recent F_2 measurement from ZEUS [33] and to fixed-target data has been performed within the ZEUS collaboration [24]. The ZEUS-S PDF set resulting from the best fit to the data, setting α_s to the world average ($\alpha_s(M_Z) = 0.118$ [16]), has been used in the present analysis to perform comparisons to measured quantities. One of the main focuses of the determination of the ZEUS-S PDF set was the evaluation of the effect of experimental uncertainties on the PDFs, including point-to-point correlated

systematic errors of the experimental data, and theoretical and model uncertainties. The PDF uncertainties were used in the present analysis to estimate the uncertainties in the present measurements arising from the PDFs used in the Monte Carlo simulation (see section 8.1.4). The parametrisation of the ZEUS-S PDFs xf at the initial scale of $Q_0^2 = 7$ GeV^2 was:

$$xf(x) = p_1 x^{p_2}(1-x)^{p_3}(1+p_4 x) \tag{2.35}$$

As in the case of the MRST2001 fit, the gluon density becomes negative at low Q^2 and low x, approximately at $Q^2 \approx 1$ GeV^2 and $x \lesssim 2 \times 10^{-3}$, but is still consistent with zero within the estimated uncertainty.

Chapter 3

HERA and the ZEUS detector

3.1 The HERA collider

The *Hadron Elektron Ring Anlage* (HERA) collider is situated in Hamburg, Germany, at the *Deutsches Elektron Synchrotron* laboratory (DESY). HERA was commissioned in 1991 [44] and has been collecting physics data since June 1992. It is currently the only high energy physics facility which collides beams of leptons and protons.

There are four experiments placed around the ring. ZEUS and H1 are general purpose experiments which take data from lepton-proton interactions. HERA-B was designed to study CP violation in the $B^0\bar{B}^0$ system by measuring J/ψ and K^0 mesons originating from the decay of B mesons. The B mesons are produced by the interaction of protons from the beam halo with a fixed tungsten wire target. The HERMES experiment measures the spin structure of the nucleon using the scattering of longitudinally polarized electrons or positrons off a jet of polarized gas consisting of hydrogen, deuterium or He^3.

The proton and electron beam energies are not the same. The proton beam is accelerated to an energy of 920 GeV while the electron beam is accelerated to 27.5 GeV. The collider is housed in a tunnel of 6.3 km circumference (figure 3.1) and consists of two separate rings. The proton ring employs superconducting magnets operating at around 4° K to produce the intense magnetic field (4.68 T) necessary to bend the very energetic proton beam. The electron ring employs conventional magnets to keep the less energetic beam in a stable orbit.

The proton beam is produced by ionising hydrogen, thus creating H^- ions which are subsequently accelerated in a 50 MeV H^- linear accelerator. The ions are stripped of electrons by making the beam pass through a thin foil. The proton beam is then injected in the proton synchrotron DESY III in up to 11 bunches separated by 96 ns, the same spacing as in HERA. After being accelerated to 7.5 GeV, the protons are transferred to

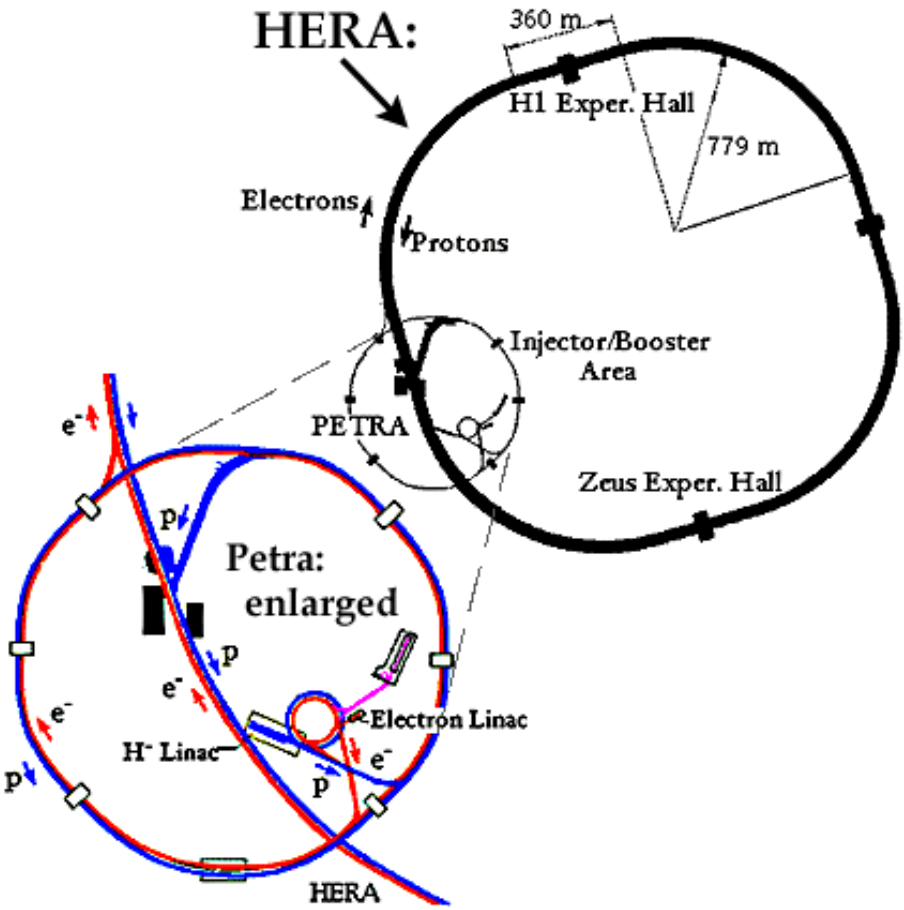

Figure 3.1: The HERA collider and pre-accelerators

the PETRA storage ring. This process is repeated until up to 70 bunches are accumulated in PETRA, at which point they are accelerated up to an energy of 40 GeV. From PETRA, the proton bunches are injected in HERA and accumulated until up to 210 bunches are filled [45]. These are then accelerated to 920 GeV (820 GeV before 1998).

Electrons are obtained from a hot filament and positrons are produced by pair production from bremsstrahlung photons, themselves produced using an electron beam [46]. After collection, the electrons or positrons are accelerated in a series of stages before being injected into HERA. First, the electron beam is accelerated by two linear accelerators LINAC I and II to an energy of 450 MeV and injected into the DESY II storage ring. Electrons are then accelerated to 7 GeV, transferred to PETRA II, where 70 bunches separated by 96 ns are accumulated, and accelerated to 14 GeV. Finally the electron bunches are transferred to HERA and accelerated to their final energy of 27.5 GeV. The electron beam has a typical lifetime of 8-10 hours.

The two beams are made to collide at the centre of the two general purpose experiments, ZEUS and H1. HERA-B and HERMES use only the proton and only the electron

circumference	6335.83 m
bunch spacing	96 ns (28.8 m)
number of buckets	220 ($\sim$ 180 colliding)
beam energies	920 GeV (p) 27.5 GeV ($e^{\pm}$)
energy spread ($\Delta E/E$)	9.47×10^{-5} (p) 1.02×10^{-3} ($e^{\pm}$)
bunch length	85 mm (p) 8.3 mm ($e^{\pm}$)
electron bunch width	280 μm (horiz.) 50 μm (vert.)
proton bunch width	265 μm (horiz.) 50 μm (vert.)
RF frequency	208.2/52.05 MHz (p) 499.7 MHz ($e^{\pm}$)
luminosity (design)	1.5×10^{31} cm^{-2}s^{-1}
beam lifetime	$\sim$ 10 hours (dominated by $e^{\pm}$ beam lifetime)
residual gas pressure	$\sim 3 \times 10^{-9}$ Torr

Table 3.1: Parameters of the HERA collider [16, 47, 45].

beam, respectively. Table 3.1 contains the key parameters that characterise the HERA collider.

An important parameter for colliding beam experiments is the luminosity. The event rate for a physical process is proportional to its cross section, σ, with the proportionality constant being the instantaneous luminosity. The luminosity is proportional to the number of particles in the colliding bunches and to the frequency of the bunch crossing, and inversely proportional to the transverse size of the beams [16]. The *integrated luminosity*, $\mathcal{L}$, expressed in inverse cross section units, is often the limiting factor on the precision to which a process can be studied, as the number of accumulated events, N, is proportional to this quantity, according to:

$$N = \mathcal{L}\sigma. \tag{3.1}$$

The HERA performance first increased above its design value in the 2000 running period, when the collider routinely delivered $\sim$ 2 pb^{-1} per week to the ZEUS and H1 experiments. Figure 3.2 shows the increase in HERA performance over the years.

3.1.1 The bunch structure

The beams in particle colliders normally have a *bucket* structure, reflecting the radiofrequency of the accelerating cavities. Not all possible buckets are filled. Among filled buckets, in addition to paired colliding bunches, which are filled in both beams, some buckets are filled for one of the beams only (pilot bunches). Other buckets are empty in both the electron and the proton beam. This is illustrated in figure 3.3. In the figure, empty bunches, shown as empty ellipses, are represented crossing at time $t = t_0$. Proton-pilot bunches are represented crossing at $t = t_1$ and electron-pilot bunches crossing at $t = t_2$. The remaining bunches shown are colliding bunches, with both electron and proton bunches filled. Pilot and empty bunches are used for background studies, allowing

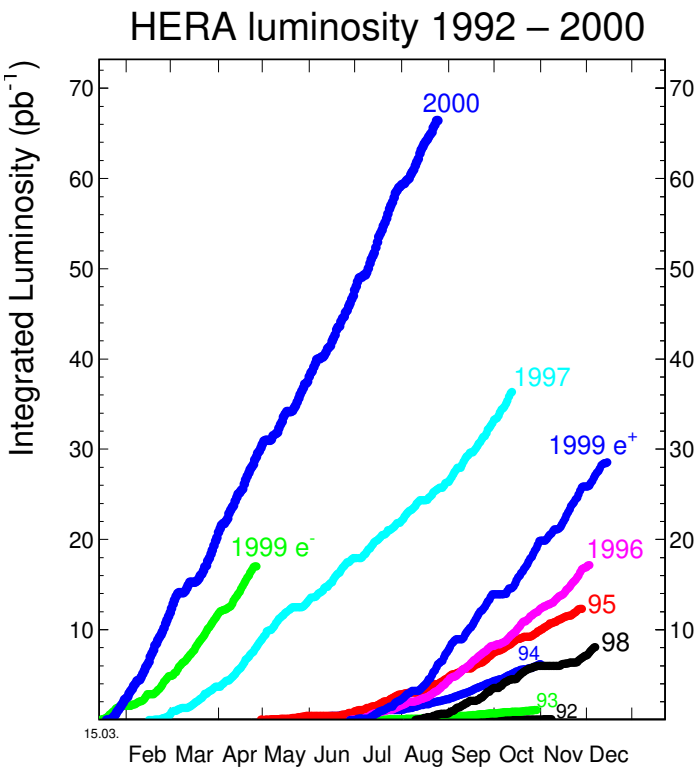

Figure 3.2: Integrated luminosity delivered by HERA to the ZEUS experiment from 1992 to 2000 [48].

the estimation of the rate of interactions between each beam and the residual gas in the beampipe (beam-gas events). This is one of the relevant sources of background to any physics analysis in ZEUS, as the interesting physics interaction rate (before the trigger selection) is much smaller than the rate of beam-gas events.

3.2 The ZEUS detector

The ZEUS detector is one of the two general purpose detectors operating in the HERA collider. ZEUS has the usual features of modern collider experiments. It consists of an inner tracking region, where a uniform magnetic field is created by a thin solenoid, surrounded by a high resolution calorimeter. This is further enclosed by an iron sampling calorimeter (backing calorimeter) which is magnetized through the use of copper coils to serve as a return path for the solenoid magnetic field. Muon chambers are placed on the inside and outside of the backing calorimeter. Other ZEUS sub-systems include detectors placed close to the beampipe for small angle tagging and detectors to measure the luminosity delivered by HERA.

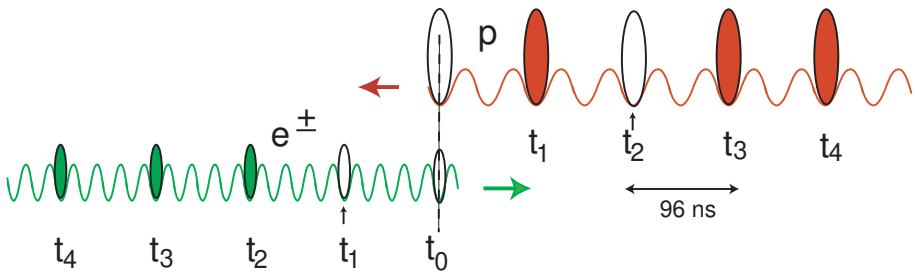

Figure 3.3: Bunch structure of HERA beams showing pilot and empty bunches. Empty buckets are represented by empty ellipses and filled bunches by filled ellipses. After two empty bunches cross at time t_0, a filled proton-pilot bunch and an empty electron bunch cross at t_1, then a filled electron-pilot bunch and an empty proton bunch cross at t_2. Subsequent crossings are between filled bunches until the cycle starts again with a new bunch-train.

The ZEUS detector is described in detail elsewhere [45]. An outline of the detector components most relevant to this analysis is given below. The detector is schematically represented in figures 3.4 and 3.5.

All positions and angles will be quoted with respect to the ZEUS reference frame. This is a right-handed reference frame with the z axis co-linear with the beams, pointing in the direction of motion of the proton beam. The y axis is perpendicular to the beams pointing upwards, and the x axis points to the centre of the HERA ring.

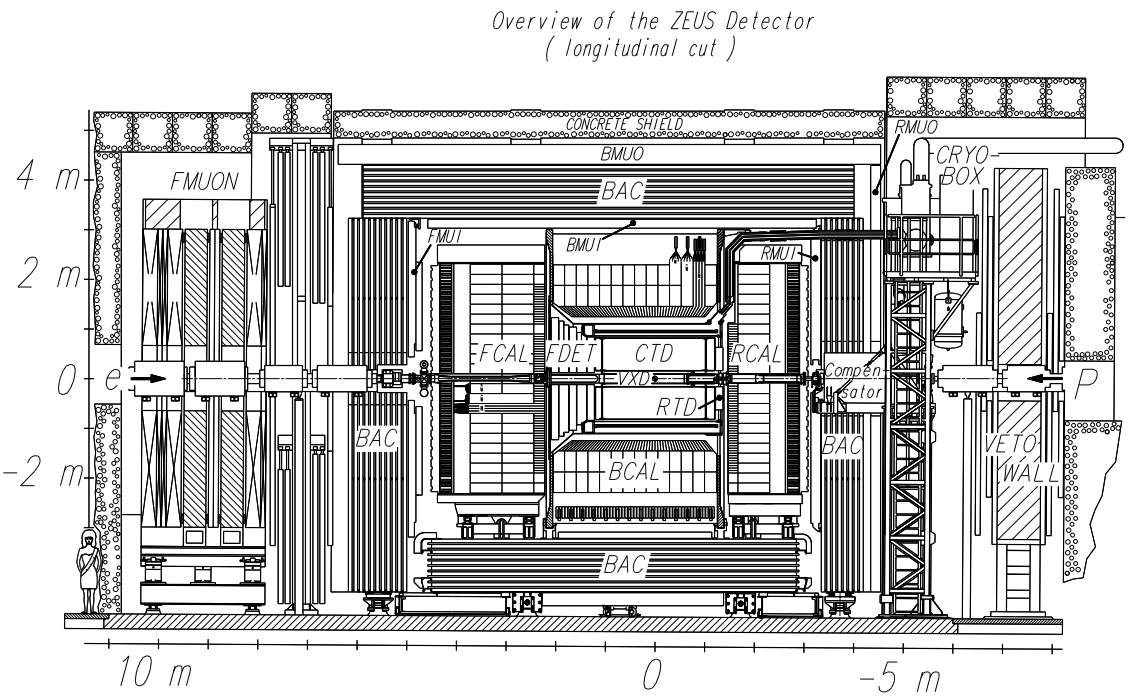

Figure 3.4: Schematic representation of the ZEUS detector [45].

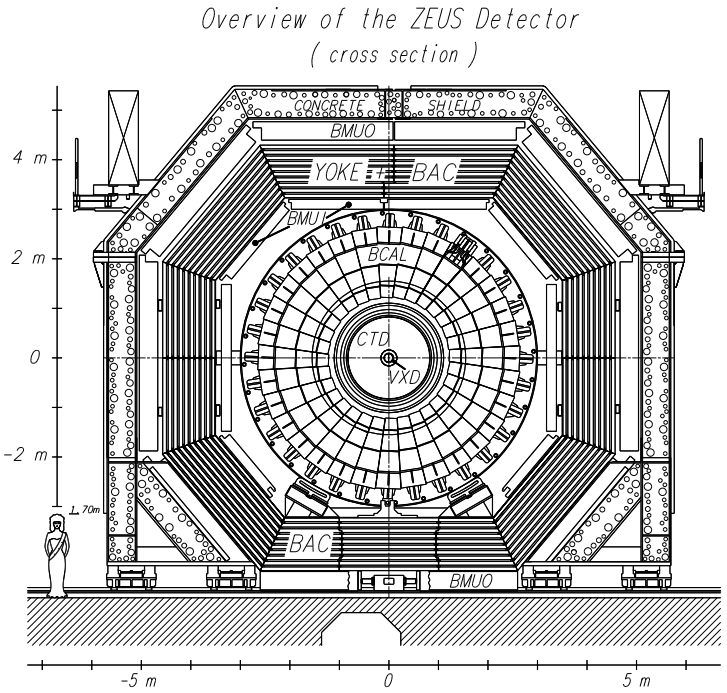

Figure 3.5: Schematic representation of the ZEUS detector [45].

3.2.1 The central tracking detector

The central tracking detector (CTD) [49] lies at the heart of the ZEUS detector. It consists of a multicell drift chamber with charge multiplication close to the anode (sense) wires. The chamber has nine "superlayers", with 8 layers of sense wires each. Each superlayer is further divided into cells consisting of 8 sense wires plus field-shaping wires. The chamber geometry ensures that a straight (high momentum) track passing through any superlayer always crosses two planes of sense wires [50], i.e. leaves hits in two adjacent cells (see figure 3.6). The CTD outer shell consists of thin inner and outer walls and electrostatic shields and of 20 mm thick endplates. A "safety volume" containing nitrogen surrounds the CTD.

The CTD is operated with a gas mixture of $Ar/CO_2/C_2H_6$, in the proportions of 83/3/14 by volume, at 3 mbar above atmospheric pressure. The mixture is bubbled through ethanol to provide protection against the aging provoked by whisker growth. The chamber has been in almost continuous operation since the beginning of the experiment, and has started to show some signs of aging during 1999. The problem was solved by adding a small amount of water to the mixture (by bubbling the gas through water), in adition to the ethanol. This mixture provides the appropriate drift velocity of the ionisation charges ($\sim$50 μm/ns) and Lorentz angle ($\sim 45°$) for a short charge collection time (maximum drift time of 500 ns) and for the operation in the intense ZEUS magnetic field.

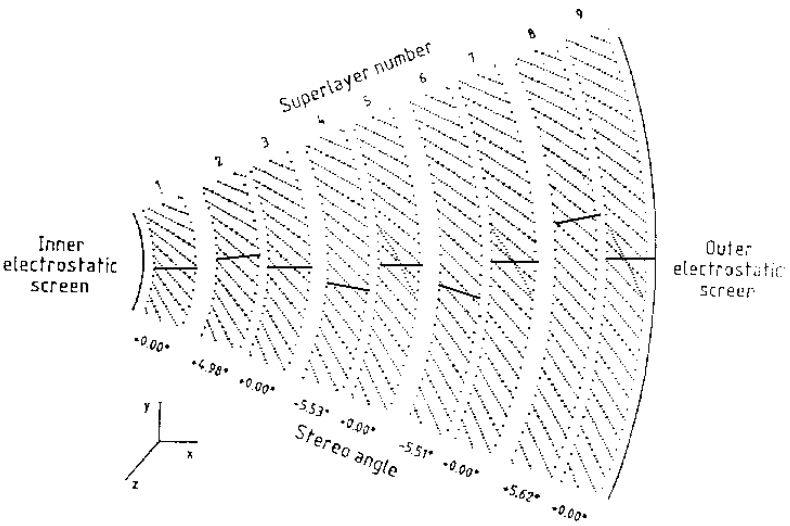

Figure 3.6: Schematic representation of one octant of the CTD. Sense wires and field-shaping wires are represented by dots. Superlayer numbers and angles of axial and stereo superlayers are noted in the figure [49]. The straight line segments indicate how a straight track at an angle to the $x - y$ plane might be viewed by the axial and stereo superlayers.

The CTD is used to measure the angles and momenta of charged tracks. Well reconstructed tracks, crossing three or more superlayers, can be extrapolated to the calorimeter face with small uncertainty. While wires in odd numbered ("axial") superlayers are parallel to the z axis, even numbered superlayers have angles of around 5° with respect to the z axis. These "stereo" superlayers provide information on the z position of ionisation clusters and thus on the track polar angle. Additional information on the z position of hits is obtained from the precise timing of the charge signals measured in both ends of sense wires in superlayers 1,3 and 5 [51]. This information can be obtained in a short enough time to be used in the first level of the ZEUS trigger system [51, 50]. This is referred to as the z-by-timing system.

The CTD provides the most accurate measurement of the position of the interaction point (vertex). The accuracy of the vertex measurement along z is around 1.5-2 mm for well measured events [52]. The z-by-time system alone has a vertex resolution which varies from 4 cm for high-multiplicity events to around 12 cm for single track events [53]. The ZEUS reference frame is determined by the CTD, with the positions of other sub-detectors being aligned with respect to the CTD.

The chamber operates in a magnetic field of 1.43 T, created by a thin superconducting solenoid ($\sim 0.9\ X_0$ [45]). Charged particle tracks originating at the nominal interaction point with a polar angle in the region $15° < \theta < 164°$ leave signals in at least two CTD

inner/outer radius	16.2 cm/85.0 cm (79.4 cm/18.2 cm active)
total length	241 cm (202.4 cm active)
polar angle acceptance	$15^\circ < \theta < 164^\circ$
limits in z	-98.7 cm $< z <$ 103.7 cm
maximum drift length/time	$\sim$ 25 mm / $\sim$ 500 ns
z-by-timing resolution (good tracks)	30 mm (nominal) / 47 mm (measured)
z stereo resolution	1.4 mm (nominal)
hit r-ϕ resolution	220 μm (measured)
polar angle resolution	$\sigma_\theta \sim 3.5$ mrad (high-Q^2 NC)

Table 3.2: Summary of parameters of the CTD from [49, 56] and from CTD data quality monitoring.

superlayers. Detailed Monte Carlo studies [54] indicated that the transverse momentum resolution for tracks crossing all the superlayers of the CTD can be parametrised as $\sigma(p_T)/p_T = 0.0058p_T \oplus 0.0065 \oplus 0.0014/p_T$ in the interval 150 MeV$< p_T <$ 5 GeV. In this expression, the transverse momentum, p_T, is expressed in GeV and the $\oplus$ sign indicates a quadratic sum. The first term in the expression comes from the CTD single-hit resolution and limits the resolution for high momentum tracks. The second term can be attributed to multiple scattering inside the CTD, which places a lower limit on the momentum resolution. The final term becomes relevant at low p_T and can be attributed to scattering in the beampipe and inner wall of the CTD.

More relevant to this analysis is the resolution in the measured polar angle of charged particle tracks, in particular the polar angle of the scattered electron track. Detailed studies [55] have shown that, for high-Q^2 events, the CTD resolution is $\sim$3.5 mrad. This resolution is lower than the calorimeter resolution and thus gives a preferable method of determining the electron polar angle. In addition, such a determination isn't affected by any possible misalignment between CTD and calorimeter.

3.2.2 The uranium calorimeter

The ZEUS calorimeter (CAL) is a uranium-scintillator sampling calorimeter. Its design emphasises a good resolution for the detection of hadronic jets. To this end, an excellent degree of compensation was achieved to guarantee that the response to electrons, or the electromagnetic component of jets, is the same as the response to the hadronic component. The CAL consists of three sections: the forward calorimeter (FCAL), barrel calorimeter (BCAL), and the rear calorimeter (RCAL) [57, 58, 59]. The CAL surrounds the solenoid and tracking detectors, with a solid angle coverage of 99.7% of 4π srad (see figure 3.7).

The calorimeter is composed of 3.3 mm depleted uranium plates (98.1% ^{238}U, 1.7%Nb, 0.2% ^{235}U [57]) interleaved with 2.6 mm plastic scintillator plates (see figure 3.8). This combines the enhanced production of fission neutrons in the absorber material with an efficient detection of those neutrons in the active layers, thus enhancing the signal from

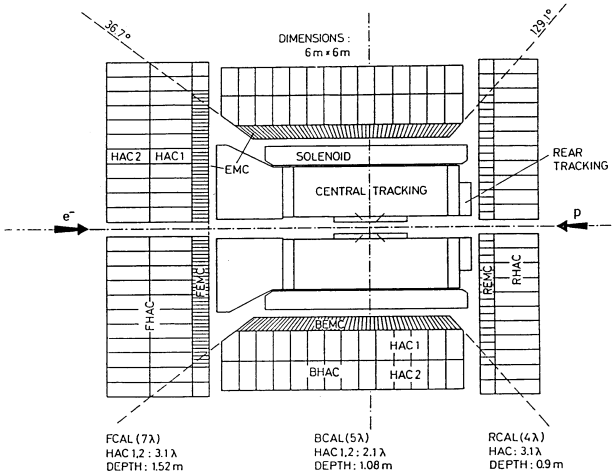

Figure 3.7: Simplified layout of the ZEUS calorimeter and tracking detectors [57].

the hadronic component of showers. The response to the electromagnetic component is reduced by the use of the high-Z absorber material (with a plate thickness of $\sim 1X_0$). Most low energy photons are converted in the absorber material [60]. The combination of the enhanced hadronic response and the reduced electromagnetic response resulted in a very good degree of compensation ($e/h = 1.00 \pm 0.03$ [59]). The choice of scintillator readout also avoids pile-up effects by providing pulses shorter than the bunch crossing time of 96 ns. The overall time resolution of the calorimeter is around 1 ns [57, 59, 61], which allows very good discrimination against background from cosmic rays and beam-gas events.

Each calorimeter section is divided longitudinally into an electromagnetic (EMC) and two (one) hadronic sections (HAC1 and HAC2) in the FCAL and BCAL (RCAL). Calorimeter sections are also segmented transversely into towers. Each tower is composed of four (two) EMC cells and two (one) HAC cells in the FCAL and BCAL (RCAL). EMC cells are 5×23 cm^2 in the FCAL, and 10×20 cm^2 in the RCAL. BCAL EMC cells have a projective geometry in polar angle with inner faces of 5×20 cm^2, wider in the azimuthal direction. All cells are read independently by two photomultiplier tubes (PMT) through wavelength shifter (WLS) sheets and light guides, which transmit the light pulses from the scintillator plates to the PMT. The difference (imbalance) in signals from the two PMT tubes is used to achieve better spatial resolution. The FCAL and RCAL are further divided into two halves, which can be moved apart to protect the instrumentation from radiation damage during beam injection in HERA.

The inner part of each CAL section (EMC section) corresponds to about 25 (21 in the BCAL) radiation lengths (X_0) for electromagnetic showers or one interaction length (λ) for hadronic interactions, at normal incidence. The total depth of the calorimeter sections

varies from 7.1 λ in the FCAL to 3.3 λ in the RCAL, reflecting the asymmetry of the beam energies. These dimensions correspond to the containment of at least 95% of the energy of 90% of hadronic jets up to an energy of 800 GeV in the forward direction. The calorimeter resolution measured under test-beam conditions is given by $\sigma_E/E = 35\%\sqrt{E(GeV)}$ for hadronic showers and by $\sigma_E/E = 18\%\sqrt{E(GeV)}$ for electromagnetic showers.

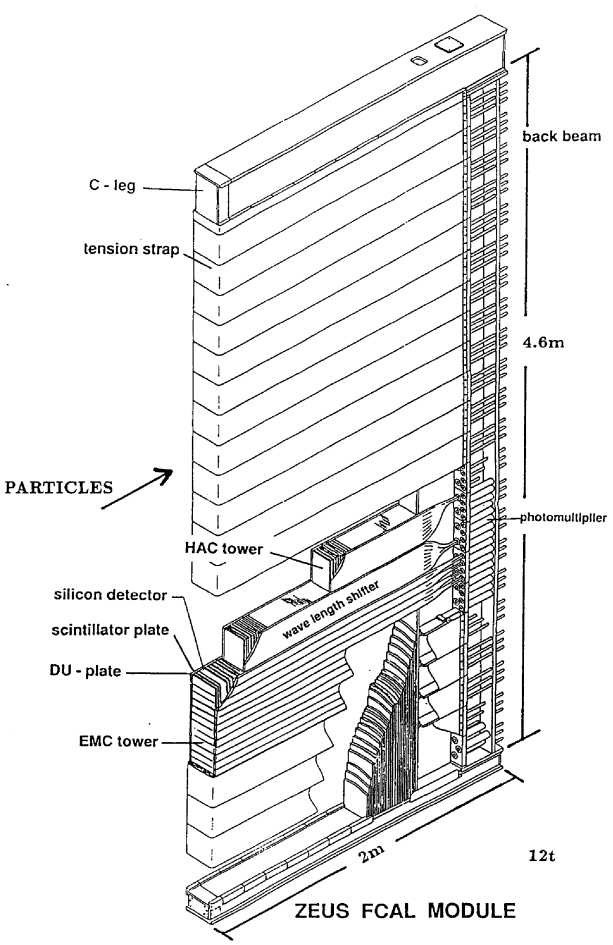

Figure 3.8: Constitution of the calorimeter modules in the FCAL [45]. A calorimeter module mechanically groups several towers. The wavelength shifter plates and light guides that are used to transmit the light signal from each cell to the photomultipliers run parallel to the module sides.

The stability, linearity and calibration of the readout system are monitored by several means. The natural radioactivity of the $^{238}_{92}$U ($\tau = 4.5 \times 10^9$ years) provides an extremely stable signal source which is integrated into a dc current measured in the PMTs. This signal originates mainly from photons produced in the decay of excited isotopes, since most α and β particles are stopped in the cladding of the uranium plates. This signal is used to monitor the gain of individual photomultipliers and to calibrate the relative gain between identical calorimeter cells. A section-to-section intercalibration accuracy close to 1% was achieved in beam tests, with a long-term calibration monitoring of better than 0.5%, from

the uranium radioactivity alone [57]. Three other calibration tools are employed: movable radioactive ^{60}Co sources, light flashers (laser and LEDs), and charge injectors. The ^{60}Co sources are used to provide local information complementary to the uranium signal, and information on aging of the scintillators and wavelength shifters. The distribution of light pulses to each PMT is used to test the response of the photomultiplier together with the subsequent pulse shaping electronics. Charge injection is done at the front end cards and is used to balance the gain of all electronic channels independent of PMT and uranium signal.

A layer of 400 μm thick silicon diodes is mounted in the EMC sections of the CAL at a depth of 3.3 X_0. This is the hadron-electron separator (HES). This detector component is used to help discriminate between electrons and hadronic jets [62], especially in situations where the electrons are not isolated. It can also be used to provide a more accurate measurement of the scattered electron position than can be obtained from the calorimeter alone [63]. The small size of the diodes (28.9×30.5 mm^2) provides a spatial resolution of around 5 mm for an electron energy greater than 5 GeV, by combining the signals from neighbouring diodes.

3.2.3 The presampler

Particles emitted from the interaction point lose energy in the inactive material between the vertex and the calorimeter. A map of the dead material in the ZEUS detector is shown in figure 3.9. This energy loss contributes to the broadening of the energy resolution of the calorimeter. A way of partially compensating for this loss is to employ a presampler detector in front of the calorimeter and to use the signals from this detector to correct the energy deposit measured in the calorimeter [60]. These corrections were used for electrons incident on the rear calorimeter.

The forward and rear presamplers consist of plastic scintillator plates of 20×20 cm^2 which match the FCAL and RCAL towers [65]. The scintillator plates are read out through wavelength-shifting fibers embedded in the scintillator. The light pulses are transported to multi-anode photomultipliers through clear fibers coupled to the WLS fibers.

A correction to the energy measured in the calorimeter is usually parametrised as a linear function of the energy deposit measured in the presampler. In test-beam conditions, it was found that this kind of parametrisation is valid for limited ranges of the amount of dead material traversed by the beam. Using such a parametrisation, a systematic precision of 2% was obtained for electron test-beam energies above 5 GeV and absorber thickness below 2 X_0 [65]. In a detector environment, the amount of dead material in front of the calorimeter is not precisely known and can vary rapidly, as shown in figure 3.9. However, in the angular region of interest to the present analysis, the dead material always

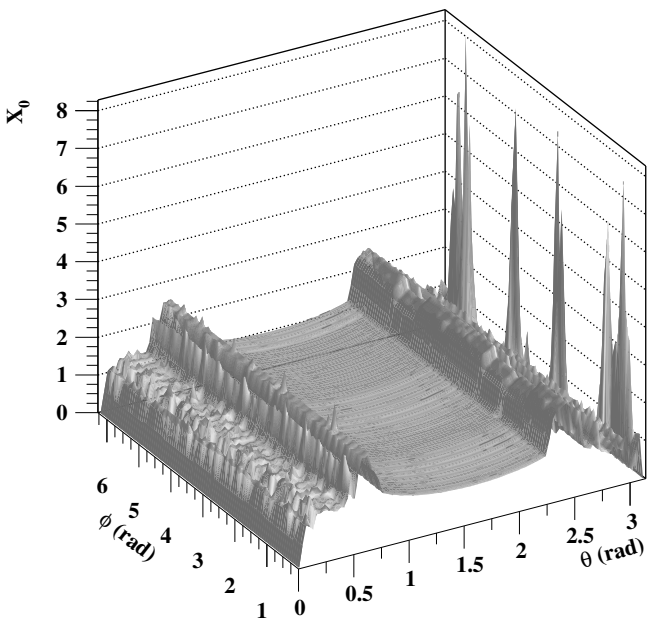

Figure 3.9: The dead material map in ZEUS obtained from the simulation of the detector. The amount of dead material is expressed in radiation lengths (X_0) as a function of polar and azimuth angle [64].

amounts to less than 3 X_0 and is generally less than 2 X_0. Linear correction parameters were determined as a function of the RCAL radius, which corresponds approximately to determining these parameters as a function of polar angle [46].

3.2.4 The luminosity detectors

The measurement of the integrated luminosity is essential for the measurement of the cross sections of physics processes. In ZEUS, the luminosity measurement relies on the detection of the photons produced in electron-proton bremsstrahlung, or "Bethe-Heitler", processes: $ep \to e\gamma p$. In this process, the incoming electron radiates a high energy photon in the electromagnetic field of the proton. The momentum transfer to the proton is negligible, so that the combined energy of final state electron and photon is equal to the electron beam energy. The final state electron and photon emerge from the interaction point at very small angles, mainly within the divergence of the electron beam (around 230 (70) μrad for the horizontal (vertical) directions) [47]). The cross section for this process is known (to about 0.3% [66])and is used to obtain a measurement of the luminosity $\mathcal{L}$ according to:

$$\mathcal{L} = R^{ep}/\sigma^{obs}_{B-H} \tag{3.2}$$

where R^{ep} is the observed rate of Bethe-Heitler events and σ^{obs}_{B-H} is the theoretical cross section corrected for detector and analysis cuts acceptance.

The luminosity monitors consist of three small electron calorimeters and a photon calorimeter (Lumi-γ). Initially, the luminosity calculation was performed by requiring the detection of an energetic photon and an electron in coincidence. However, even in 1992, the background conditions were sufficiently good to allow a precise determination of the luminosity using only the bremsstrahlung photons, with the electron detectors used for the estimation of the systematic uncertainty.

The Lumi-γ detector is placed at z = -107 m, after a 0.095 X_0 copper-beryllium window, mounted on the beampipe at z = -92.5 m, and a vacuum tube. It consists of a lead-scintillator sampling calorimeter with transverse dimensions of 18×18 cm^2. Its total depth is around 22 X_0, with x and y scintillator strips placed at a depth of 3 X_0 to give a measure of the transverse position of the shower. These give a resolution of about 3 mm, corresponding to an uncertainty of ∼30 μrad on the photon angle. The energy resolution of the calorimeter is around $\sigma_E/E = 20\%\sqrt{E(GeV)}$. The photon detector must be protected against beam-related radiation. A 3 X_0 graphite filter and the copper-beryllium window shield the detector against a large flux (∼100 W) of direct synchrotron radiation produced by the bending of the electron beam. Additionally, the detector must be moved behind a lead shield during beam injection.

Electrons resulting from a Bethe-Heitler process have an energy smaller than the beam electrons, and their trajectories are bent by the HERA magnets towards the centre of the ring. They can be detected by small calorimeters placed at z = -8 m, -35 m and -44 m on the inside of the HERA ring. The detectors placed at z = -8 m and -44 m are small tungsten-scintillator sampling calorimeters. The electrons detected at these places have average energies of 2 and 24 GeV, respectively. The detector placed at z = -35 m is of very similar construction to the Lumi-γ calorimeter. The energy of electrons detected by this calorimeter is between 5 and 21 GeV.

The luminosity measurement must take into account several sources of background, the most important of which comes from electron beam-gas bremsstrahlung, $eZ \to e\gamma Z$. This process can mimic a Bethe-Heitler process and its contribution is subtracted statistically using electron pilot bunches. Also the synchrotron radiation from the electron beam that passes the absorbers can cause pile up in the photon counter which must be corrected off-line. The systematic uncertainties, taking these and other factors into account, result in uncertainties in the measurement of the integrated luminosity between 1.34% and 2.25% in the years 1997 to 2000.

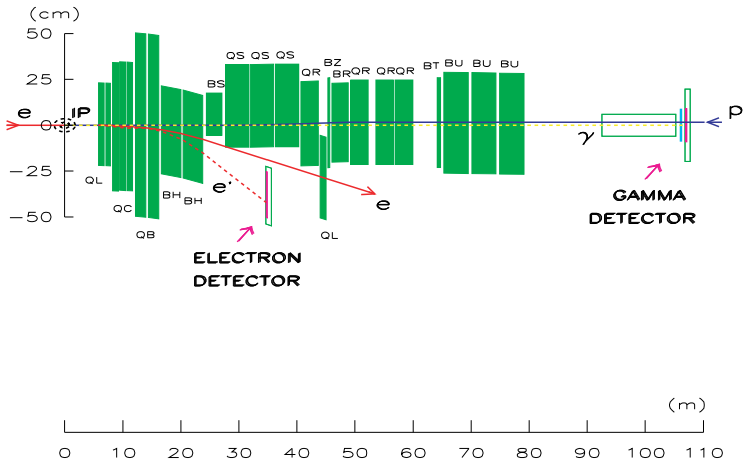

Figure 3.10: Layout of the luminosity detectors and collimators [48].

3.2.5 Other detector components

In addition to the detector components described above, two other components are relevant to this analysis through their role in the trigger: the Veto Wall and the C5 counter.

The Veto Wall [45] is an iron wall placed at around 7.5 m upstream from the interaction point in the incoming proton beam direction. It protects the detector by stopping particles created by interactions of the beam halo with beam line material. The iron wall is 87 cm thick and is instrumented with scintillators on both sides. It acts as a veto against events produced by particles passing through the wall.

The C5 detector [67] consists of two planes of scintillator layers interleaved with layers of tungsten and placed around the beampipe at z = -1.2 m. It detects particles from beam-gas events and its signal is used as a veto in the trigger. It also provides rate information to the ZEUS and HERA shift crews.

3.2.6 The ZEUS trigger system

The ZEUS trigger system [68] was designed to filter an input event rate of 10.42 MHz to an output rate of interesting physics events of ∼10 Hz. In addition to the high rate of events, the volume of data is also significant, with an event size of up to 500 kBytes (average 80 kBytes [53]). While not every bunch crossing results in a collision, the task is still made difficult by the high rate of beam-gas background events (up to ∼500 kHz). This background can leave energy deposits in the detector that approximate those of physics events and must be rejected with a very good efficiency. A small fraction of "passthrough" events are kept, regardless of trigger decision, for monitoring purposes.

The system is divided into three trigger levels. This allows more processing time for the more complex algorithms used at the upper levels while keeping the system essentially dead time free. A simplified diagram of the trigger chain is shown in figure 3.11.

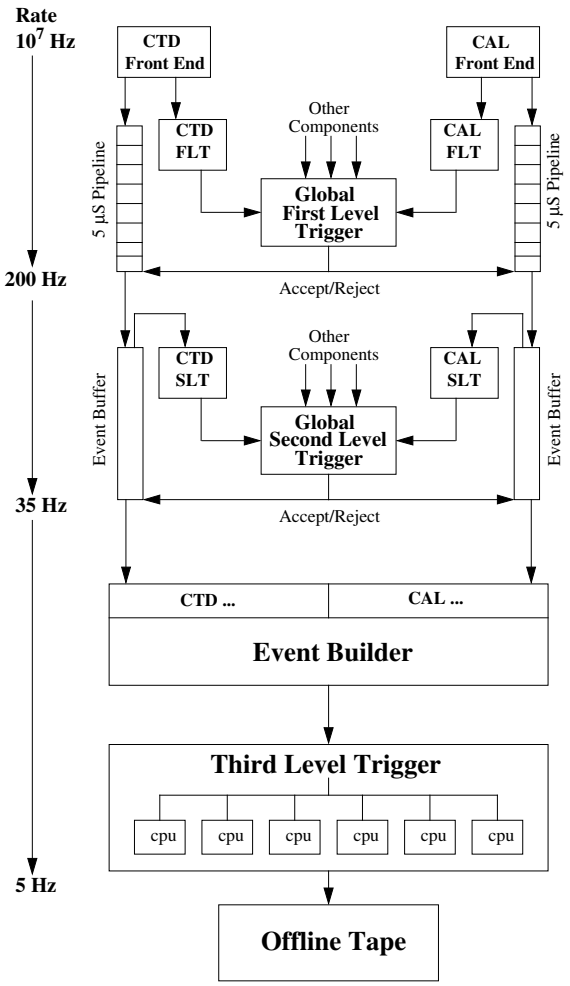

Figure 3.11: Schematic of the ZEUS trigger system [69]. The drawing is simplified by showing only the CTD and calorimeter processing chains.

The first level trigger

Due to the very short interval between events, the first level trigger (FLT) is implemented in hardware. All the data and the electronics must be pipelined to allow time for trigger processing. Each step must be repeated every 96 ns. Each detector component does part of the processing of its own data. A global first level trigger (GFLT) combines the information

from the component FLTs and takes a decision exactly 46 beam crossings (4.4 μs) after the event took place. The logic circuits in the component FLTs are stable, whereas the GFLT is implemented through the use of memory lookup tables and can be easily re-programmed. This gives a good degree of flexibility to the trigger, accommodating the changing requirements and running conditions through the lifetime of the experiment [68]. The processing time assigned to the GFLT is 20 beam crossings, or 1.92 μs, with the different components passing their information to the GFLT between 1.0 and 2.5 μs after each beam crossing. The result of the GFLT is a logical OR of 64 trigger slots which can be individually enabled, or prescaled to take only a predefined fraction of the events accepted by the slot, to avoid flooding the data stream. The GFLT accept rate is kept below 1 kHz, which is the maximum input rate of the second level trigger.

Individual components store the event data in either analogue or digital pipelines. The information used at the FLT is usually less precise than the final information used in the reconstructed event. The CTD, for example, uses the faster but less precise z-by-timing system to provide information about the event vertex to the FLT. At the FLT, the detector components that are most relevant to neutral current analysis are the CTD and the CAL, combined with the various veto detectors.

The main purpose of the tracking first level trigger is to reject events arising from interactions outside of the detector volume. To do this, the tracking FLT performs a first determination of the event vertex. The CTD z-by-timing system is used together with the forward tracking detectors to perform this task. Although the maximum drift time of the ionisation charge in a CTD cell is longer than 5 bunch crossings, the tilt of the cell planes with respect to the radius ensures that the arrival times are distributed evenly in this interval, and that hits can be attributed to a specific bunch crossing. The inputs to the tracking first level trigger are coarse three dimensional points measured in superlayers 1, 3, and 5. These are used to search for track segments in each superlayer and these segments are then used to reconstruct tracks. The vertex resolution is around 20 cm if only hits from superlayer 1 are used. This number improves to better than 5 cm when hits from either superlayers 3 or 5 or the forward tracking detectors are used.

In the calorimeter FLT (CFLT), calorimeter cells are grouped together in trigger towers, with a coarser granularity than is available off-line. These are usually composed of four (two) EMC sections and two (one) HAC sections in the FCAL and BCAL (RCAL). These towers combine EMC cells and the HAC cells behind them in groups that better approximate a projective geometry. A fixed fraction (5%) of the PMT charge signal from each EMC or HAC section of trigger towers is passed to the trigger electronics cards. These signals are pedestal corrected, summed and zero suppressed if the measured energy in a section is less than a few hundred MeV. Signals from the EMC and HAC (combining

HAC1 and HAC2) sections of each tower are summed separately and binned in a rough logarithmic scale according to 6 threshold values [69]. The surviving cells are used by the CFLT to calculate global sums of energy and transverse energy in the calorimeter. They are also used to search for patterns consistent with the detection of isolated electrons or muons, by comparing energy deposits in groups of towers to predefined patterns. The fact that only a coarse energy resolution and spatial granularity of the calorimeter are used at this level, combined with the assumption that the event vertex is at the nominal interaction point, lead to a relatively poor accuracy in the measured quantities [64].

In addition to the calorimeter trigger, and placed between the FLT and SLT, further processing of calorimeter data is done by the fast clear processor (FC) [70]. The FC is a special purpose digital processor which examines the calorimeter data to search for clusters. Once a first level trigger is issued by the GFLT, the individual component triggers start digitizing the event data to be sent to the second level trigger. This step takes up to 100 μs. The FC makes use of this time to further analyse calorimeter data and abort the event, if it is consistent with background, before the event is passed to the second level trigger. The FC takes, on average, 20 μs to perform this task, depending on the event complexity. When it is in operation, the FC aborts enough events to accomodate a GFLT output rate of 5-10 kHz without significant dead time.

Signals from the Veto Wall [71] and the C5 detectors are also important at the FLT level. Their function is to veto beam-gas and beam-halo events by detecting the energy deposits left by these events upstream of the CAL.

The second level trigger

The second level trigger (SLT) reduces the event rate further to around 100 Hz. Information from the various detector components is combined in the global second level trigger (GSLT) to implement simplified physics filters. The GSLT and the component SLTs run in a transputer network where fast parallel processing is easily implemented.

At this level there is sufficient time to use the full spatial and energy resolution of the calorimeter and to reconstruct tracks, although the information from the stereo layers in the CTD is not used [53]. A better estimate of quantities such as the vertex position, the number of clusters and the timing in the calorimeter, etc, is available at this point and can be used to set more stringent cuts against background events.

The third level trigger

Once a positive decision is taken by the GSLT, data from each component are sent to another transputer network, the event builder (EVB). The EVB combines information from all detector components and sends the complete event to the third level trigger

(TLT) [72]. This last trigger level is implemented in a farm of workstations. The TLT runs simplified versions of off-line reconstruction and physics-based filter code on complete events. On acceptance, events are written to magnetic tape via an optical fiber link (FLINK), at a rate of ∼10 Hz, for off-line analysis. Figure 3.12 shows the rate of accepted events at the various trigger levels for the data taking period of 2000, plotted against the instantaneous luminosity. Background rates were as high as 500 kHz (not shown). As expected, the accept rates show an approximately linear rise with the luminosity.

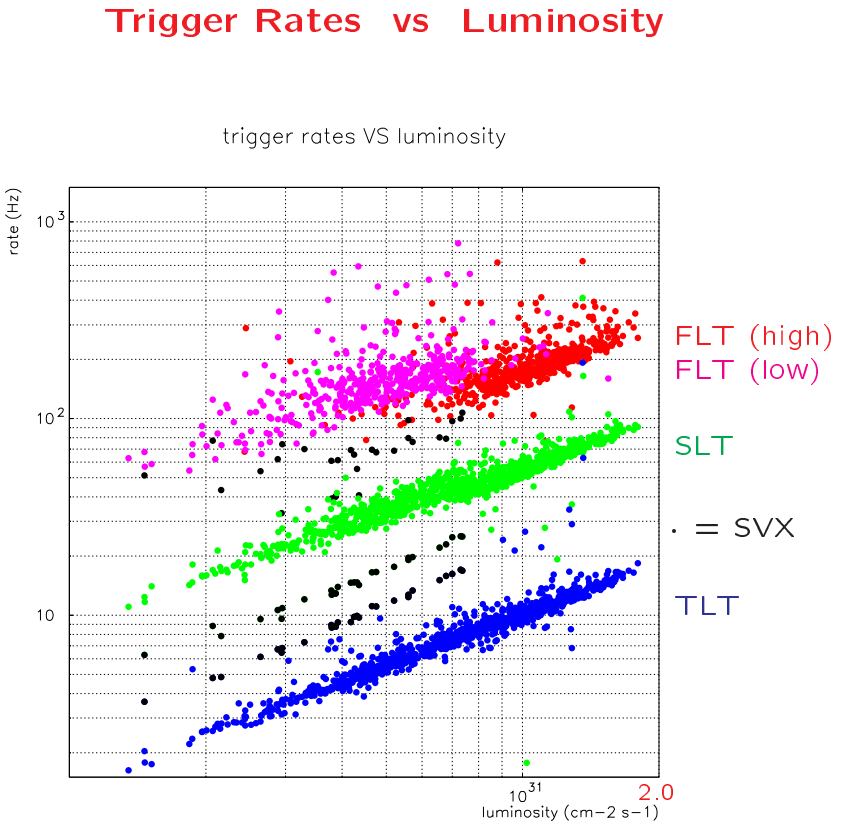

Figure 3.12: Output rates of the ZEUS trigger system during the 2000 data taking period [48]. points lebelled SVX ("shifted vertex") correspond to a small set of runs where the beams were made to interact closer to the FCAL, to access a region of lower momentum transfer between the electrons and the protons. Points marked FLT (high) and FLT (low) correspond to different configurations of the FLT trigger logic used under high and low luminosity conditions (i.e. generally at the beginning and at the end of each fill, respectively).

3.3 Detector simulation

In the analysis of events measured in the ZEUS detector, one must take into account the imperfect detector acceptance arising from factors such as the incomplete solid angle coverage or the inactive material in the detector. In order to estimate this quantity, as

well as possible biases in the quantities measured by such a complex system, the use of simulated events is mandatory.

The event simulations consists of two steps. In a first step, the known physical processes arising from the $e^{\pm}p$ collision are simulated. This must include the "hard scattering" process, in which a virtual gauge boson is exchanged between the interacting particles, and also processes such as the emission of photons by the scattered or incoming lepton or the hadronisation of the struck quark. These processes are generated by several programs which will be discussed in chapter 4. The second step consists of simulating the interactions of particles emerging from the collision with the detector and inactive material inside it. Absorption of particles in inactive material or the production of additional particles may arise in this step. Both steps in the event simulation use Monte Carlo techniques.

In practice, event generators relevant for ZEUS are collected under a common interface called AMADEUS (see figure 3.13) to produce events in a common format that can be passed to the detector simulation package. The ZEUS detector simulation is done by the MOZART program, which is based on the GEANT 3.13 package [73]. The output of MOZART is passed further to the CZAR package (also known as ZGANA), which simulates the trigger chain, and to ZEPHYR, which does event reconstruction for both measured and simulated data. The output from this process is a set of events that has the same format as the real data obtained with the ZEUS detector but has, additionally, information about the "true" quantities of the simulated events, such as the momenta of particles and their identity.

The processing of the detector simulation stage is done by the FUNNEL facility, which distributes the simulated events to several computer farms located in 14 different institutes of the ZEUS collaboration. FUNNEL processes in excess of 2-3 million events per week.

The natural radioactivity of the uranium in the calorimeter was studied in trigger "passthrough" events, where the energy deposits in the calorimeter originate entirely from calorimeter noise. The rate and energy distribution of the noise signal was simulated in MOZART and added to the generated events.

3.3.1 The longitudinal vertex distribution

Another addition to the simulation is the longitudinal vertex distribution. The spatial distribution of the interaction point is very well constrained in the x and y directions, due to the small transverse size of the beams. The vertex distribution along z, on the other hand, has a complicated structure with a central peak a few centimeters wide, and satellites that extend to more than 100 cm on either side of the nominal interaction point. This distribution depends on the HERA operating parameters and conditions and thus varies from run to run and certainly from year to year. A wide vertex distribution affects the measurement of experimental quantities, and must be well reproduced by the MC

simulation. Thus, an unbiased vertex distribution must be measured from data in each running period and constitutes an input to the event simulation performed by MOZART. The determination of the longitudinal vertex distribution is discussed in chapter 7.

3.4 The software environment

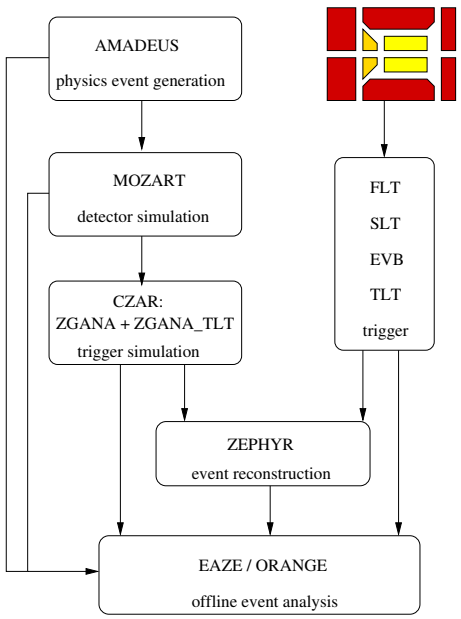

Figure 3.13: Structure of the ZEUS software for Monte Carlo production and data analysis.

Off-line analysis in ZEUS is usually done with the help of some common tools. The analysis code usually consists of FORTRAN77 user code which is linked to a set of libraries of precompiled routines. These deal with the more computing specific aspects of retrieving the data and running through large numbers of events. The library commonly used for this purpose is called EAZE for Effortless Analysis of ZEUS Events (a clear overstatement). In addition to EAZE, other libraries can be linked to the user code, which contain either analysis routines or running-period specific routines.

Recently, a new package, ORANGE, was developed within the collaboration that constitutes an extra layer of complexity. The ORANGE package aims to gather the accumulated knowledge of the various physics groups in the form of analysis and correction routines and their default parameters. ORANGE can be linked to a minimal EAZE job setup to call, in the right order and with sensible default parameters, several kinds of analysis routines. The analysis and correction routines called, as well as the form of the output ntuple, can be specified by the user through a series of control cards. User code can be linked to the ORANGE library at compile time.

Both EAZE and ORANGE can be directed to select events accepted by specific sets of trigger slots, or they can be requested to accept events based on physical quantities. The latter is achieved through the use of a tag database, ZES (for ZEUS Event Store), based on Objectivity. In this database, many physical or detector related quantities serve as tags that qualify each event. Users can base their event selection on these quantities, avoiding the time consuming task of recalculating every variable with their code only to reject, a large fraction of the analysed events.

The structure of ZEUS software is shown schematically in figure 3.13. The analysis work reported in this thesis was based on modifications to ORANGE library v2001a.2.

Chapter 4

Monte Carlo simulation

The MC simulation of events is one of the most powerful tools used in the analysis of the recorded data. The complexity of both the detectors used and of the physics events produced in current high-energy-physics experiments, makes analytic calculations of the experimental acceptance and resolution unfeasible. Instead, Monte Carlo (MC) techniques are used to simulate both the physics events and the response of the detector. Monte Carlo simulated events are used extensively to study systematic effects that affect the measurement and also in the unfolding of the cross sections from the observed data.

The simulation of the detector response was mentioned in section 3.3. The present chapter deals with the simulation of the physics processes that are the object of this analysis. This task is accomplished by programs known as Monte Carlo generators.

Monte Carlo generators implement theoretical models of the physical interactions to be simulated. Several approximations are usually employed to minimise the processing time necessary for the event generation. Additionally, the theoretical description of the physics cannot, usually, be implemented to all orders, and phenomenological models are often used as a suitable approximation.

The output from MC generators consists, for each event, of a list of initial and final state particles and their four-momenta. This is then used as an input to the simulation of the detector in order to obtain quantities which can be directly compared to the observed data.

4.1 DIS Monte Carlo

At the Born level, NC DIS is described through the exchange of a single virtual photon or Z^0 between the incoming electron and a free quark in the proton. Other, higher order, diagrams are necessary to calculate the cross section. Even though the aim of this analysis is to measure the NC DIS cross section at Born level, the MC simulation must contain

these effects if it is to be compared to real data. Figure 4.1 shows some of the higher order corrections.

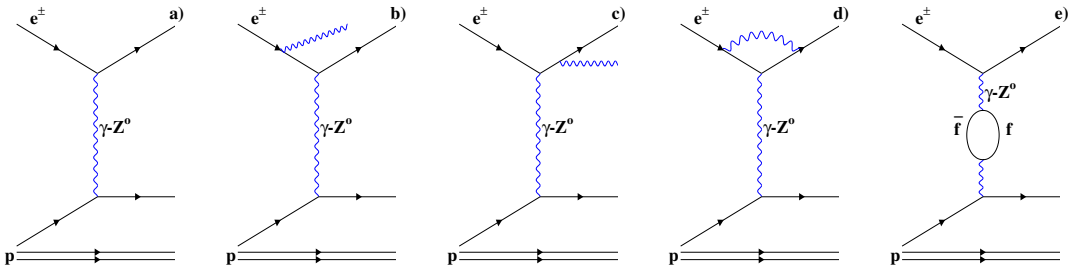

Figure 4.1: Feynman diagrams for **a)**: the Born level NC DIS process and for some electroweak corrections: ISR **b)**; FSR **c)**; vertex correction **d)**; and self-energy **e)**.

The emission of photons from the initial or final state electron, known as initial and final state radiation (ISR and FSR, respectively) changes the magnitude of the cross section and also affects the kinematics of the event. ISR photons are emitted collinearly with the electron beam and lost in the beampipe, but have the effect of lowering the centre-of-mass energy of the collision. An FSR photon can lower the energy of the detected electron, but is often detected as part of the same calorimeter cluster as the scattered electron, in which case its energy is added to the electron energy. Figure 4.2 shows histograms of the scattered electron energy in MC events generated with $Q^2 > 100$ GeV^2 before and after photon emission, as well as the photon energy.

An additional sub-process, QED Compton scattering (QEDC), should be mentioned here. The Feynman diagrams for this process are shown in figure 4.3 and are very similar to the ISR and FSR diagrams shown in figure 4.1. The difference between the ISR or FSR and QEDC diagrams is that the intermediate electron, between the two photon vertices, is off-shell. The process results in the emission of a hard photon not collinear with the intial or the final state electrons. The QED Compton process can be further divided into elastic and inelastic QEDC scattering. In the former, the proton is not dissociated. In the latter, the proton breaks up, as in a normal DIS process.

The kinematic reconstruction methods used in the present analysis result in dramatic migrations in the case of QEDC events because the hard photon emitted from the electron line is considered as part of the hadronic final state. These migrations are not a problem on their own if they are correctly modelled by the MC simulation. A study has shown that this is the case for inelastic QEDC [74]. Elastic QEDC processes, on the other hand, are not simulated by the event generators used and must be rejected by the off-line selection cuts (see sections 6.2.5 and6.5).

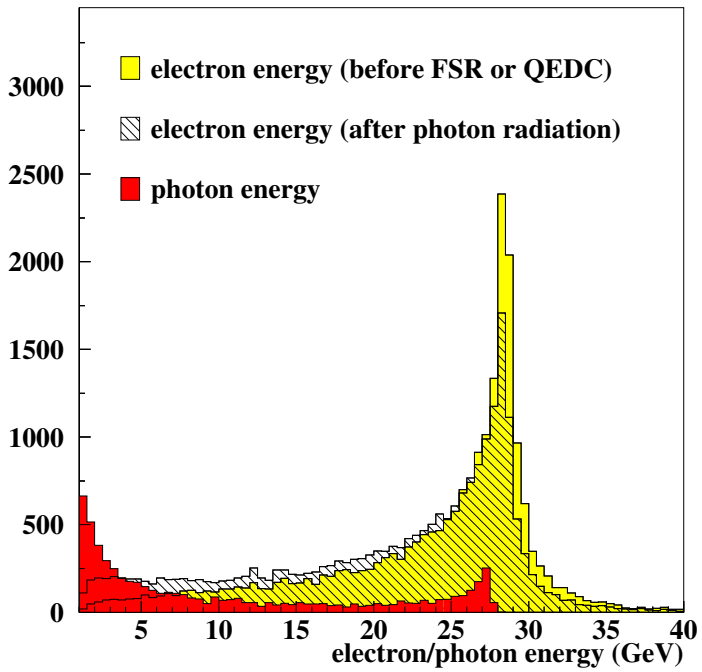

Figure 4.2: Effect of FSR and QEDC emission of a single photon on the energy distribution of the scattered electron. The energy distribution of the emitted photon is shown as a dark-shaded histogram.

The DJANGOH 1.1 Monte Carlo (MC) program [75] was used to generate NC DIS events. The CTEQ5D [25] parton density function set was used in the generation of the DIS Monte Carlo events. DJANGOH is an interface to two programs: LEPTO 6.5.1 [76], which simulates the hard scattering process, and HERACLES 4.6.1, which provides the electroweak radiative corrections [77]. The corrections provided by HERACLES include single photon emission (ISR and FSR), as well as the complete set of one loop weak corrections described by self-energy and two boson exchange diagrams.

4.1.1 The parton cascade and hadronisation

On the hadronic side of the event, at Born level, the cross section is given by the diagram shown in figure 4.4 **a)**. As in the case of electroweak corrections, other diagrams have to be used to obtain a realistic simulation. Unlike the former case, though, the interacting quarks cannot be detected or considered to be free. They are connected by the colour field to the rest of the proton and only the hadrons resulting from the non-perturbative hadronisation process are visible in the detector. Figures 4.4 **b)** to **d)** show Feynman diagrams corresponding to corrections of order α_s. These corrections do not modify the cross section. Instead, they are taken into account through the structure functions.

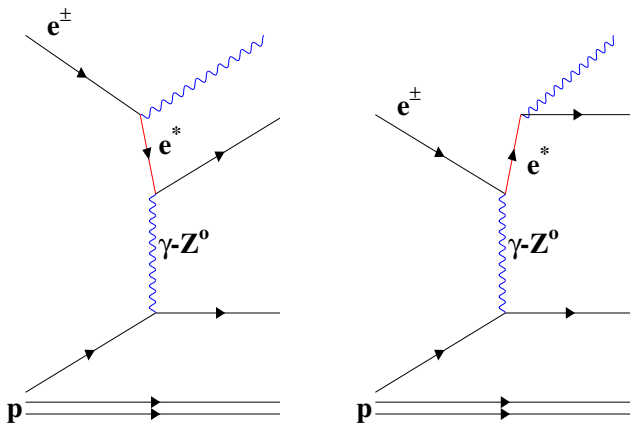

Figure 4.3: Feynman diagrams for the QED Compton process. The intermediate electron is off-shell.

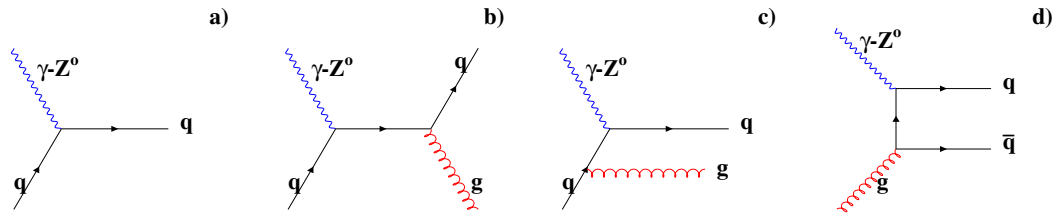

Figure 4.4: Feynman diagrams for: **a)** the leading order process; **b)** and **c)** gluon emission; and **d)** the boson-gluon fusion processes.

The LEPTO program, interfaced to DJANGOH, implements first order QCD corrections to the cross section through exact calculations of their QCD matrix elements (ME). Higher order corrections are implemented approximately by LEPTO through the parton shower (PS) model [78]. In this model, the struck quark may be part of a QCD cascade (shower) initiated by another parton in the proton. The interaction with the gauge boson results in a final state quark. This quark also originates a parton shower by branching into daughter particles. The evolution of initial and final state parton showers is determined by the DGLAP equations (see section 2.4.1). This is the matrix elements and parton showers (MEPS) approach to the simulation of the hadronic side of the event.

An alternative model exists, in which quark-antiquark pairs ($q\overline{q}$), for example, are seen as a colour dipole. The emission of a gluon by one of the quarks in the pair can be treated as radiation emitted by the colour dipole. As the gluon itself has colour, the resulting $q\overline{q}g$ combination can be treated as two independent dipoles, one formed by the quark and the gluon and the other by the gluon and the antiquark. The emission of a second gluon is

then treated approximately as radiation from two independent dipoles, etc. This is known as the colour dipole model (CDM).

The colour dipole model is implemented in the ARIADNE 4.10 program [79]. This generator simulates only the QCD cascade and must be interfaced to other programs that deal with the hard interactions, hadronisation and particle decays. In DIS, the struck quark and the proton remnant form a colour dipole which is used by ARIADNE to simulate the initial stage of the QCD cascade. The first emission produced by this dipole is matched to the $\mathcal{O}(\alpha_s)$ matrix elements. To simulate the boson-gluon (BGF) process (figure 4.4 **d)**), which is not described by the CDM, the struck-quark-remnant dipole emits the anti-partner of the struck quark according to the BGF matrix element. The emission of partons in the gap between the struck quark direction and the proton remnant direction is determined by the dipole connecting quark and remnant.

The result of the QCD parton cascade, as implemented either in the MEPS or the ARIADNE models is a set of coloured quarks and gluons. These are still not the final state colourless hadrons that are detected in the experiment. One last step is necessary, usually known as fragmentation. This is achieved by the JETSET 7.410 [80] program, distributed together with PYTHIA 5.724, which implements the Lund string fragmentation algorithm [80]. In this algorithm, each pair of partons is connected by a string representing the colour force field. As the partons move apart, the potential energy between the partons increases until enough energy exists for the breaking of the string and the formation of a new $q\overline{q}$ pair. The procedure is repeated and groups of quarks combined to form hadrons. The process is stopped when the available energy is contained in the mass and kinetic energies of hadrons. The decay of unstable hadrons can, at this stage, be taken into account.

The DJANGOH program, with QCD cascades produced by ARIADNE 4.10 [79], has been used as the default Monte Carlo generator in the analysis of high-Q^2 NC DIS events, as it shows good agreement with the data [81] (see section 7.2). Changes to the CDM implementation in ARIADNE were introduced in 1999 to improve the treatment of high-Q^2 DIS events [82, 83]. These modifications were used in the present analysis.

4.1.2 The DJANGOH sample

The NC DIS sample generated with DJANGOH consists of 832151 e^-p and 1.3 million e^+p NC DIS events in 10 different sets generated with minimum Q^2 values ranging from 100 GeV^2 to 50 000 GeV^2. Table 4.1 gives a summary of the sample used.

A different sample of around one million events was generated using the MEPS model implemented in LEPTO, instead of the CDM model from ARIADNE. This sample was used to study the agreement between the MC description of the hadronic final state and the

DJANGOH	e^-p			e^+p		
Q^2_{min} (GeV^2)	nr. events	σ (pb)	$\mathcal{L}$ (pb^{-1})	nr. events	σ (pb)	$\mathcal{L}$ (pb^{-1})
25	164670	4.6×10^4	3.6	84251	4.6×10^4	1.8
40	199955	2.6×10^4	7.6	84251	2.6×10^4	2.9
100	318758	8.2×10^3	39	909791	8.1×10^4	112
400	57818	1.2×10^3	48	98755	1.2×10^3	84
1250	25000	2.2×10^2	115	47989	2.0×10^2	242
2500	12000	7.2×10^2	167	23995	5.9×10^2	407
5000	12000	2.2×10^1	554	23990	1.5×10^1	1617
10000	11960	5.4×10^0	2231	17108	2.7×10^0	6290
20000	11997	8.5×10^{-1}	14164	16252	3.1×10^{-1}	52426
30000	5995	1.9×10^{-1}	32352	11990	5.4×10^{-1}	220404
40000	5998	4.3×10^{-2}	140798	12000	1.1×10^{-2}	1103955
50000	6000	9.2×10^{-3}	652173	11990	2.1×10^{-3}	5661000

Table 4.1:
Summary of non-diffractive Monte Carlo data used in the analysis of e^-p and e^+p events.

observed data, and to estimate the systematic uncertainty produced by our incomplete knowledge of the hadronisation process.

Combining MC samples

To compare distributions obtained in real and simulated data, MC events must be properly normalised. This is achieved by normalising the simulated events to the measured integrated luminosity, $\mathcal{L}^{data}$. To do this, an equivalent integrated luminosity, $\mathcal{L}^{MC}$, is defined for the MC sets according to the expression:

$$m = \mathcal{L}^{MC} \cdot \sigma_{MC} \Leftrightarrow \mathcal{L}^{MC} = m/\sigma_{MC} \tag{4.1}$$

where m is the number of MC generated events, σ_{MC} is the total cross section to which the MC sample corresponds, and $\mathcal{L}^{MC}$ is the equivalent integrated luminosity of the sample. A weight, w, can then be assigned to each MC event according to:

$$w = \mathcal{L}^{data}/\mathcal{L}^{MC} \tag{4.2}$$

As the NC DIS cross section falls roughly as $1/Q^4$, using sets of MC events generated with different minimum Q^2 avoids the generation of very large numbers of events. To combine the various sets, different weights, w_i are given to MC events in different Q^2 regions. As the MC sets partially overlap in Q^2, the expression used to calculate the weight, w_i, of an event with $Q^2_{event} > Q^2_i$ was:

$$w_i = \frac{\mathcal{L}^{data}}{\sum_{j \leq i} \mathcal{L}^{MC}_j} \tag{4.3}$$

where i denotes the index of the sample with highest minimum Q^2 such that $Q^2_i < Q^2_{event}$.

RAPGAP	e^-p			e^+p		
Q^2_{min} (GeV^2)	nr. events	σ (pb)	$\mathcal{L}$ (pb^{-1})	nr. events	σ (pb)	$\mathcal{L}$ (pb^{-1})
100	37757	8.1×10^3	4.6	99998	8.1×10^2	12
400	30000	1.2×10^3	25	113618	1.2×10^2	97
1250	19995	2.2×10^2	92	66940	1.9×10^2	344

Table 4.2:
Summary of diffractive Monte Carlo data used in the analysis of e^-p and e^+p events.

4.1.3 Diffractive Monte Carlo

One of the first results from the HERA experiments was the observation of a class of events in which no hadronic activity exists in the rapidity gap between the current jet and the proton remnant [84, 85]. This kind of event is known as *diffractive*. As the proton and struck quark are coloured objects, they are connected by colour flow in non-diffractive events, which results in the emission of particles in the gap between remnant and current jet. This colour flow does not exist in diffractive events, which can be described by the exchange of a colourless object between the virtual photon and the proton. The difference between the two classes of events can be seen in the distribution of the maximum pseudo-rapidity in DIS events, η_{max}, defined as:

$$\eta_{max} = -\ln\left(\tan\frac{\theta_c}{2}\right) \tag{4.4}$$

where θ_c is the polar angle of the most forward energy cluster (or "condensate", see section 6.4.1) found in the calorimeter with energy above 400 MeV (to reject noise). Diffractive events typically have lower values of η_{max}, which define the measured rapidity gap between proton remnant and current jet. The addition of diffractive events is necessary to describe the η_{max} distribution found in the data (see figure 4.5 below).

Diffractive events were generated with RAPGAP 2.08/06 [86]. This program uses HERACLES for the calculation of radiative corrections and ARIADNE to simulate the parton cascade. JETSET simulates the fragmentation of the parton cascade to produce the hadronic final state. RAPGAP was modified to generate exclusively diffractive events and to reproduce the NC DIS cross section as given by DJANGOH. The CTEQ5D parton density function set was used for the generation of diffractive events.

Table 4.2 gives a summary of the MC sets generated with RAPGAP. Samples of 87752 e^-p and 280556 e^+p diffractive events were generated.

Combining diffractive and non-diffractive MC samples

The diffractive MC events were combined with non-diffractive events to obtain an optimal description of the hadronic final state observed in real data. To do this, a weight factor

w_{diff} was attributed to each diffractive MC event, with a corresponding factor $w_{n-diff} = 1 - w_{diff}$ attributed to non-diffractive events. The values of these weights were determined by fitting the η_{max} distribution found in the combined MC sample to the one found in real data. The weight factors determined from the fit (w_{diff} and w_{n-diff} for diffractive and non-diffractive events, respectively) are multiplied by other weight factors, such as the weight attributed to each MC set discussed in section 4.1.2, to obtain the total weight of each MC event. It should be noted here that this procedure is not intended as a measure of the diffractive DIS cross section, which is beyond the scope of the present analysis. What is intended, is to obtain a good description of the topology of the hadronic final states found in real events. This procedure is necessary to avoid systematic effects which would arise from a poor description of the event topologies.

The fit was done by adding the η_{max} distributions found in diffractive and non-diffractive MC samples (both normalized to $\mathcal{L}^{data}$), weighted by w_{diff} and w_{n-diff}, respectively. The weights were varied in small steps and the χ^2 of the agreement between the combined MC distribution and the η_{max} distribution was minimised. The distributions up to $\eta_{max} = 3$ were used to calculate the χ^2, as higher values of η_{max} correspond to a region where the beampipe and dead material may bias the measurement. The procedure used all data and MC events that pass the selection procedure. The η_{max} distributions found in real data and in the diffractive and non-diffractive MC samples are shown in figure 4.5. The χ^2 of the fit is also shown in the figure as a function of w_{n-diff}. From this fit,

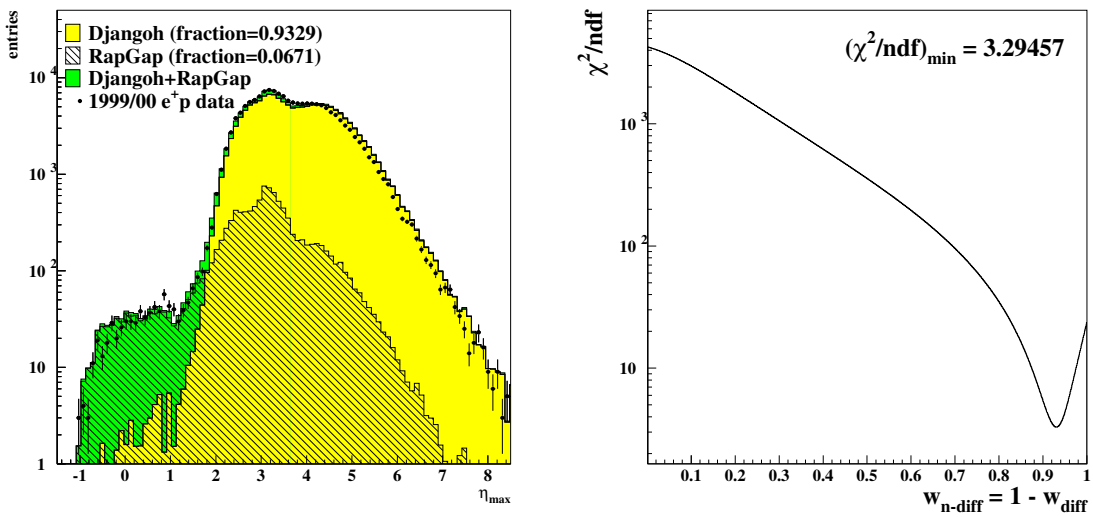

Figure 4.5: Left: distributions of maximum pseudo-rapidity found in data and Monte Carlo, and combination of diffractive and non-diffractive Monte Carlo samples. Right: ratio of the χ^2 of the fit to the number of degrees of freedom as a function of the non-diffractive MC weight factor w_{n-diff}.

4.1.4 Reweighting to the cross section including F_L

While indirect measurements of the F_L structure function at $Q^2 \leq 35$ GeV have been published by the H1 collaboration [87, 88], no direct measurements of this structure function have been made at high Q^2. The F_L term in the NC DIS cross section is suppressed by a factor of y^2, and only becomes significant at high y.

Instead of including the calculation of F_L in the Monte Carlo simulation, events were generated with $F_L = 0$ and reweighted according to the ratio of the cross sections calculated with and without this term. Figure 4.6 shows lego plots of the reduced cross section, $\tilde{\sigma}(x, Q^2)$ (see section 8.2.4 for the definition of the reduced cross section), in the plane of the kinematic variables x and Q^2. The effect of the F_L term can be seen at the boundary of the accessible phase space, defined by $y = Q^2/(s \cdot x) = 1$, i.e. at the low x limit for each value of Q^2. The F_L term has the effect of lowering the value of the cross section at high y, as it has a negative contribution to the cross section.

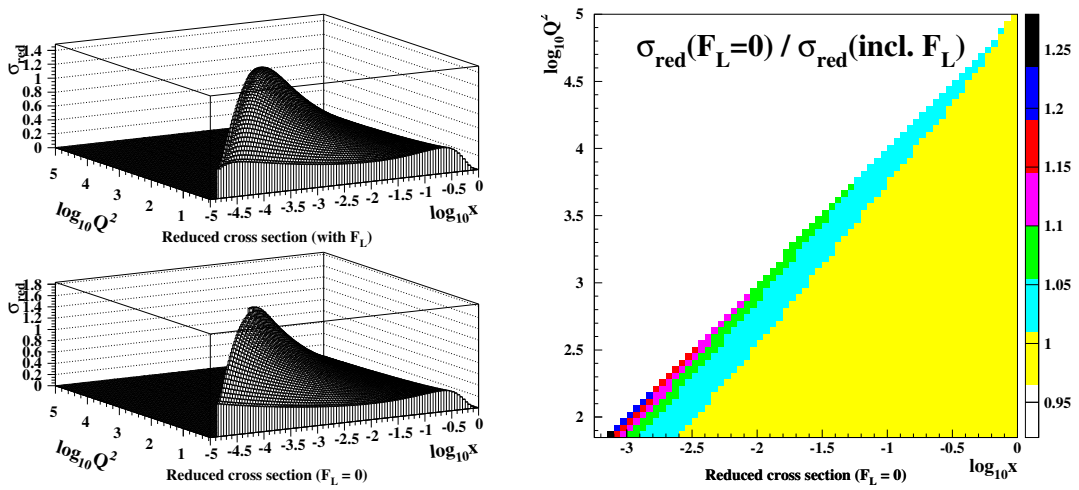

Figure 4.6: Left: effect of the F_L term on the reduced cross section (noted σ_{red} on the figure). The difference between the cross sections calculated with and without the F_L term can be seen in the low-x (high-y) edge of the lego plots. Right: ratio of the Born level reduced cross sections calculated with and without considering the F_L term, shown versus x and Q^2.

To reweight the MC events, a 2-dimensional grid was built with the ratio of the values of the NC DIS cross sections calculated with and without the F_L term. A weight factor was assigned to each MC event which was equal to the value of this ratio at the x and Q^2 where the event was generated. Linear interpolation between grid points was used to obtain the particular values of the weights given to each event. The validity of this procedure was verified by using a different program which calculated the Born level cross section for each MC event. Again, the cross section including F_L divided by the cross

section excluding this structure function were used as weights, only this time avoiding the use of a grid with its inherent interpolation errors. The values of the measured cross sections obtained with both methods were compared, and their differences were found to be negligible.

Also shown in figure 4.6, is the ratio $\tilde{\sigma}_{F_L=0}/\tilde{\sigma}_{F_L}$, versus $\log_{10} x$ and $\log_{10} Q^2$. This ratio corresponds to the inverse of the weights assigned to MC events as a function of x and Q^2. The region of phase space showing the largest effect of the F_L term, with differences of up to 20% between cross sections with and without F_L, is located at low x and Q^2 (see figure 4.6). Events in this region are eliminated from the final event sample by the analysis cuts (see figure 6.25).

4.2 Photoproduction background

The main source of background in the present analysis is the misidentification of photoproduction (PHP) events as NC DIS events. In the photoproduction process, a quasi-real photon is exchanged between the electron and the proton (i.e. the process has $Q^2 \sim 0$). As the exchanged momentum is very small, the electron goes practically undeflected and is lost in the rear beamhole. In the present case, a working definition of photoproduction as events in which the scattered electron is not detected in the calorimeter is adopted.

In PHP events, a particle belonging to the hadronic final state may be misidentified as the scattered electron. This can occur, for example, in the case of photons emitted in the decay of a π^0. These photons will produce an electromagnetic energy cluster in the calorimeter wich is hard to distinguish from that of the scattered electron. As the PHP cross section is much larger than that of NC DIS, the background contamination from this source can be significant.

PHP may be separated into two subprocesses, called direct and resolved photoproduction. In direct photoproduction, the exchanged photon participates in the hard scatter. In resolved PHP, the photon fluctuates into a hadronic system, which acts as a source of partons that interact with the proton. In the latter case, only a fraction of the photon momentum will participate in the hard scattering process. Feynman diagrams for the two subprocesses are shown in figure 4.7.

Photoproduction events were simulated using the HERWIG 5.9 generator [89]. The CTEQ5-LO [25] set of proton parton density functions was used, together with the GRV-LO [43] set of photonic parton density functions.

To limit the number of generated events, kinematic cuts were implemented at the generator level: the transverse momentum of the hard scatter was required to be above 4 GeV; either the transverse energy, E_T, was required be greater than 30 GeV, or the total

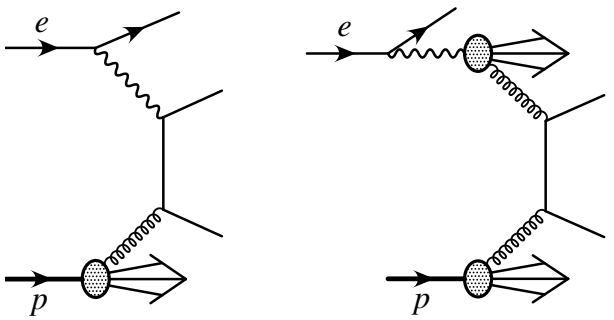

Figure 4.7: Feynman diagrams representing direct and resolved photoproduction.

HERWIG	direct			resolved		
Q^2_{max} (GeV^2)	nr. events	σ (pb)	$\mathcal{L}$ (pb^{-1})	nr. events	σ (pb)	$\mathcal{L}$ (pb^{-1})
4	1090000	6.5×10^4	681	3600000	2.9×10^5	470

Table 4.3: Summary of photoproduction Monte Carlo samples used in the analysis of both the e^-p and e^+p observed events. The equivalent MC integrated luminosity indicated in the table includes corrections for the generator-level cuts and for the uncertainty in the calculated PHP cross section.

transverse momentum was required to be greater than 6 GeV. Particles emitted in the solid angles corresponding to the beamholes were excluded from the calculation of these quantities.

Table 4.3 gives a summary of the MC sets used in the present analysis. The same sets, generated with an e^- initial state lepton, were used to estimate the background contamination in the e^+p and the e^-p real data samples. This is allowed because the photoproduction cross section is the same for both charged leptons. The trigger and off-line analysis cuts, described in chapter 6, reject 98.6% and 99.4% of the direct and the resolved simulated PHP events, respectively.

4.2.1 Normalisation of the photoproduction Monte Carlo

The equivalent luminosity, quoted in table 4.3, and used to normalise the PHP MC samples, is *not* calculated directly from the number of events and the cross sections quoted in the same table. Two factors must be taken into account: the generator level cuts and the uncertainty in the calculated cross section.

The generator level cuts reject a large fraction of the generated events. An *effective* cross section can be defined, which is obtained by multiplying the true cross section by the selection efficiency. The selection efficiency is given by the ratio of the number of events that passed the generator level cuts to the total number of events that were generated. The effective cross sections of the direct and resolved PHP samples were found to be 2.8×10^3 pb and 13.4×10^3 pb, respectively.

A further correction must be applied to obtain the equivalent luminosity of PHP MC sets. The photoproduction cross section is not precisely known. Detailed studies [90, 91] have shown that the cross section estimated by the MC generator was too high. A correction to the normalisation of the PHP Monte Carlo data was estimated by tagging a set of photoproduction events using one of the electron calorimeters in the luminosity system. In a fraction of the photoproduction events, the scattered electron, deflected by the beam optics, was detected downstream in the calorimeter (Lumi-e tagger) placed at $z = -35$ m (see section 3.2.4). This provides a means to tag PHP events. Figure 4.8 shows the distribution of energy found in the Lumi-e tagger for real data and PHP MC data. Due to the characteristics of the beam optics, this detector has good acceptance for electrons scattered from the interaction region with energies between 5 GeV and 15 GeV. The energy distribution found in real data is not described by the PHP MC alone. Electron-beam bremsstrahlung events constitute a source of background to the tagging of photoproduction, and must be added to the PHP MC distribution to obtain a good description of the observed data. The bremsstrahlung events shown in the figure were selected using a random trigger. The figure shows that the acceptance of the Lumi-e detector is well described by the MC simulation.

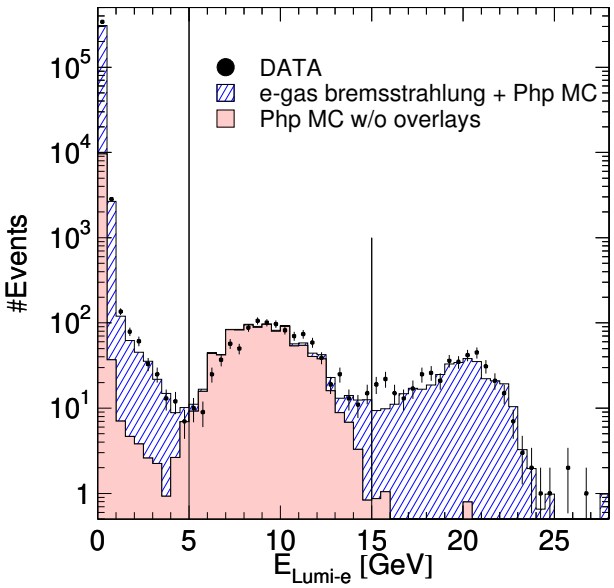

Figure 4.8: Energy deposited in the Lumi-e tagger by the scattered electron in PHP events. Overlay electron-beam bremsstrahlung events selected by a random trigger are added to the PHP MC distribution. From reference [91].

The ratio of the number of tagged photoproduction events to the total number of (NC DIS plus PHP) events selected with an appropriate set of cuts should be the same in real and simulated data, since the tagging efficiency is well modeled by the MC simulation. A

complication arises due to *overlay* events in real data, in which a NC DIS event and a PHP event occur simultaneously. This can be overcome by comparing the behaviour of real and simulated data, as the simulation does not contain overlaid events. Differences between the ratios found in real and simulated data were attributed to an innacurate estimation of the PHP cross section by the simulation, and allowed the estimation of correction factors.

The size of the discrepancy between the estimated cross section and the true one was determined to be of a factor of 2 in the 1998/99 data sample [90] and 1.75 in the 1999/00 sample [91]. These factors are in agreement within the errors of the method used. The value of 1.75 was used in the present analysis and the normalisation was varied by +100% and -50% to estimate the size of the systematic error due to the uncertainty in the normalisation of the PHP background. The equivalent integrated luminosity of the direct- and resolved-PHP MC sets, indicated in table 4.3, was obtained by multiplying the integrated luminosity after generator-level cuts by the factor 1.75 .

Chapter 5

Measurement of Kinematic Variables

An accurate measurement of the Lorentz-invariant kinematic variables x, y, and Q^2 is essential for the measurement of NC DIS cross sections and structure functions. In this section, the methods used in the measurement of the kinematic variables are reviewed and their resolutions compared.

The near hermeticity and good resolution of the ZEUS detector can be exploited to measure four variables that characterise the events. These are the angles and energies of the final state electron and hadronic system. Figure 5.1 illustrates the NC DIS process. The four momentum of the incoming and scattered electron are given by k and k', respectively. P is the four momentum of the initial proton and P' is the four momentum of the struck quark. In the ZEUS reference frame, neglecting particle masses, these four momenta can be expressed as:

$$k = \begin{pmatrix} E_e \\ 0 \\ 0 \\ -E_e \end{pmatrix}, \quad k' = \begin{pmatrix} E'_e \\ E'_e \sin\theta\cos\phi \\ E'_e \sin\theta\sin\phi \\ E'_e \cos\theta \end{pmatrix}, \quad P = \begin{pmatrix} E_p \\ 0 \\ 0 \\ E_p \end{pmatrix}, \quad P' = \begin{pmatrix} E_h \\ P_{x,h} \\ P_{y,h} \\ P_{z,h} \end{pmatrix}, \qquad (5.1)$$

where θ and ϕ are, respectively, the polar and azimuthal angles of the scattered electron, E_e is the electron beam energy, E'_e is the energy of the scattered electron, E_p is the energy of the incoming proton. The initial momentum of the struck quark is xP. On the hadronic side, the picture is complicated by the proton remnant, which is lost in the forward beampipe, and by the colour flow between the remnant and current jet. Assuming the naive quark parton model, without colour flow between remnant and current jet, the four momentum of the struck quark, P', can be reconstructed by a sum of the four momenta of all particles in the current jet.

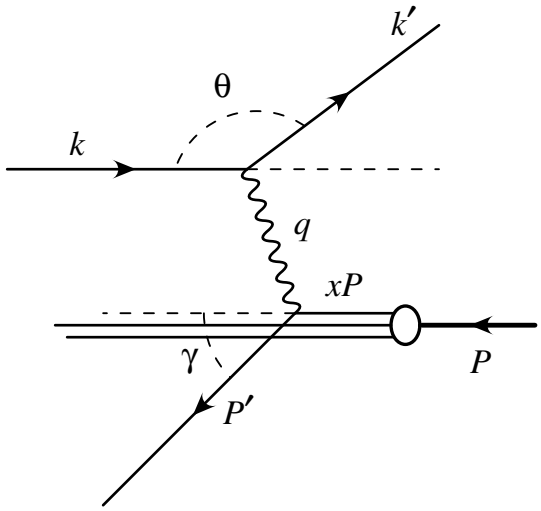

Figure 5.1: Feynman diagram of deep inelastic scattering at Born level. The incoming and scattered lepton has momenta k and k', respectively. The interaction is described at Born level by the exchange of a gauge boson of four momentum q between the lepton and a quark in the proton. The interacting quark carries a fraction x of the proton's momentum, denoted by P. The scattering angles of the final state lepton, θ, and struck quark, γ, are also represented in the diagram.

Figure 5.2 illustrates the topology of DIS events in different regions of phase space. The arrows represent the momenta of the scattered lepton (thin arrows) and current jet (thick arrows) at polar angles θ and γ_h. At low Q^2 the polar angle of the scattered electron is close to 180°. The angle of the current jet varies from very low values at low y (high x), corresponding to a very energetic forward going jet which is detected in the FCAL, to backward going jets at low x (high y). As Q^2 increases the final state lepton is scattered at increasingly large polar angles until, at the highest Q^2, it has very high energy and low polar angle.

In order to reconstruct the energy and angle of the current jet, the influence of the details of the fragmentation process (number of jets, spatial distribution of final state particles in the ZEUS reference frame, etc.), as well as the influence of the loss of particles through the forward beam hole must be minimised. This is achieved through the use of quantities that have little sensitivity to these factors: the hadronic transverse momentum, $p_{T,h}$, and the difference between the energy and longitudinal momentum of the hadronic final state, δ_h. These are given by:

$$p_{T,h}^2 = p_{x,h}^2 + p_{y,h}^2 = (\sum_i^h E_i \sin\theta_i \cos\phi_i)^2 + (\sum_i^h E_i \sin\theta_i \sin\phi_i)^2, \qquad (5.2)$$

$$\text{and} \quad \delta_h = (E - p_z)_h = \sum_i^h (E_i - p_{z,i}), \qquad (5.3)$$

where the sums run over all particles in the hadronic system. Neither of these quantities is significantly affected by the loss of particles through the forward beam hole, because such

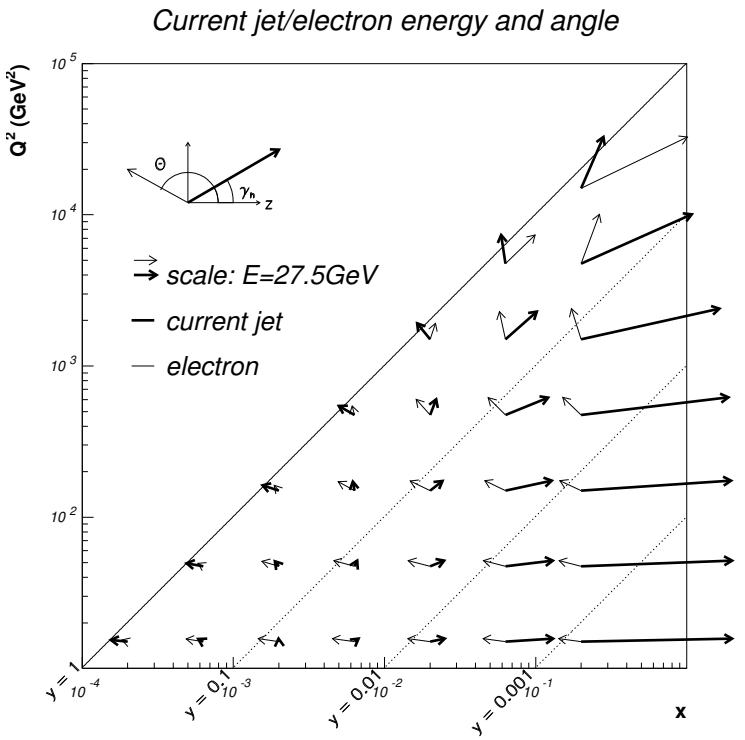

Figure 5.2: Schematic representation of DIS event topology. Thick arrows represent the angle and energy of the current jet. Thin arrows correspond to the scattered lepton.

particles have small transverse momentum and their longitudinal momentum is nearly the same as their energy.

The $p_{T,h}$ and δ_h variables can be used to calculate an angle, γ_h, according to:

$$\cos\gamma_h = \frac{p_{T,h}^2 - \delta_h^2}{p_{T,h}^2 + \delta_h^2}, \tag{5.4}$$

which, in the naive quark parton model, is the scattering angle of the struck quark. The current jet may be approximated by a massless object with energy E_h and scattering angle, in the ZEUS reference frame, γ_h. Two caveats should be noted here. The first is that the measurement of $p_{T,h}$ and δ_h, and therefore of γ_h, become difficult and imprecise at very low y where most of the hadronic system is lost through the forward beampipe. The second is that, for non-diffractive events, the measured γ_h is affected by the radiation emitted due to the colour flow between the current jet and the proton remnant. γ_h corresponds to the polar angle of the scattered quark only in the naive quark parton model.

The variables E'_e, θ, E_h, and γ_h can now be used to reconstruct the Lorentz-invariant kinematic variables that characterise the event. Since there are only two independent variables (x and Q^2, for example), six combinations of the measured quantities can, in principle, be used [92]. The ones that were used in this analysis are described in the next three sections.

5.1 The electron reconstruction method

This method uses the scattered electron variables only. The kinematic variables are given by:

$$Q^2_{el} = 2\,E_e E'_e(1+\cos\theta) = 4\,E_e\,E'_e\cos^2\frac{\theta}{2} \tag{5.5}$$

$$y_{el} = 1 - \frac{E'_e}{2\,E_e}(1-\cos\theta) = 1 - \frac{E'_e}{E_e}\sin^2\frac{\theta}{2} \tag{5.6}$$

$$x_{el} = \frac{Q^2_{el}}{s\,y_{el}} = \frac{E'_e\cos^2\frac{\theta}{2}}{E_p(1-\frac{E'_e}{E_e}\sin^2\frac{\theta}{2}),} \tag{5.7}$$

where s is the square of the centre of mass energy.

Figure 5.3 shows lines of constant electron polar angle (left) and energy (right) in the kinematic plane. In the figure, it can be seen that for low y and low Q^2, the energy of the scattered electron becomes approximately independent of the event kinematics and is approximately the same as the beam energy (27.5 GeV). This region of phase space is known as the *kinematic peak* region. In the neighbourhood of the kinematic peak region, the resolution of the electron method is poor, since a small error in the measurement of the electron energy leads to a large migration in the reconstructed kinematic variables (the interval 27 GeV $< \mathrm{E}'_\mathrm{e} <$ 28 GeV in measured energy corresponds to a x range as large as $0.005 < x < 1$ for $Q^2 \sim 10$ GeV^2).

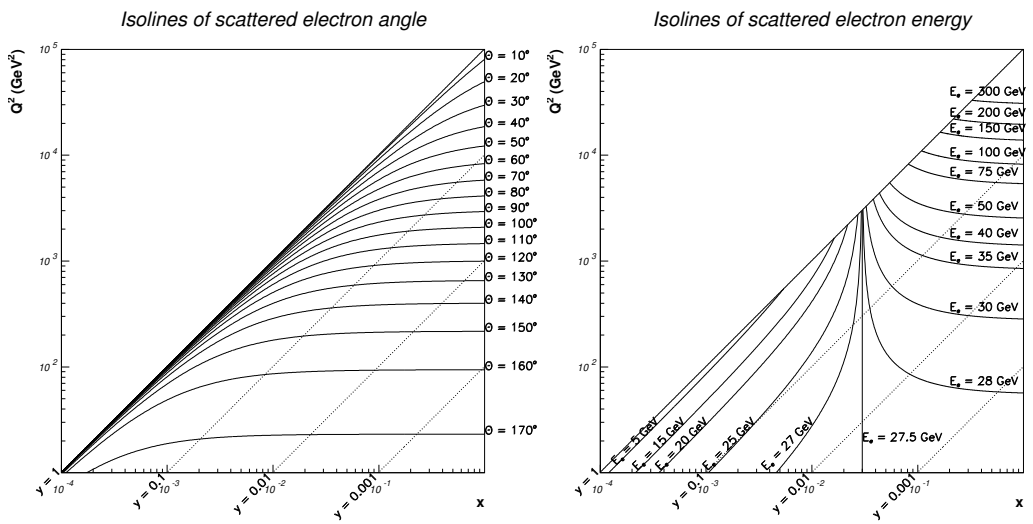

Figure 5.3: Lines of constant electron angle and energy. As a reference, the calorimeter angular limits from the nominal interaction point are around 2.2° for the FCAL beam hole and 176.5° for the RCAL beamhole. "Kinematic peak" events have a scattered electron energy close to 27.5 GeV.

5.2 The Jacquet-Blondel reconstruction method

This method uses only quantities measured from the hadronic final state [93]. It is the only method available for the kinematic reconstruction of charged current DIS events, where the final state neutrino goes undetected leaving only an imbalance in the transverse momentum as a signature. The kinematic variables are given by:

$$y_{JB} = \frac{\sum_i^h (E_i - p_{z,i})}{2\,E_e} = \frac{\delta_h}{2\,E_e} \tag{5.8}$$

$$Q^2_{JB} = \frac{\sum_i^h \left(p^2_{x,i} + p^2_{y,i}\right)}{1 - y_{JB}} = \frac{p^2_{T,h}}{1 - y_{JB}} \tag{5.9}$$

$$x_{JB} = \frac{Q^2_{JB}}{s\,y_{JB}} \tag{5.10}$$

where the sums run over all energy clusters measured in the calorimeter, except that of the scattered electron. Figure 5.4 shows lines of constant transverse momentum and hadronic $E - p_Z$ (left) and of the energy of the struck quark (right).

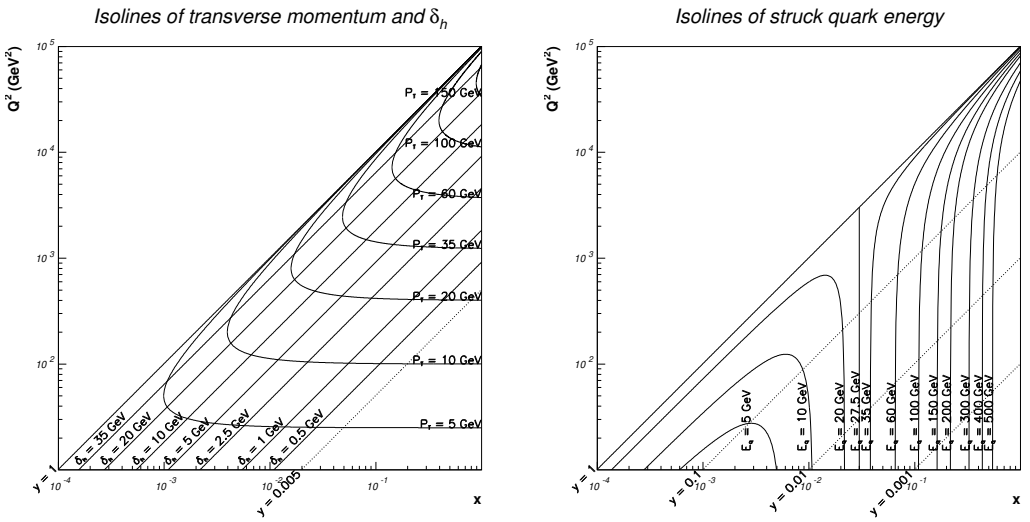

Figure 5.4: Isolines of the struck quark transverse momentum and $\delta_h = (E - p_z)_h$ (left), and of the struck quark energy (right).

The accuracy with which the hadronic final state particles can be measured is relatively poor, due to energy loss in inactive material and other effects such as backscattering in the calorimeter. Therefore, the measurement of $p_{T,h}$ is not very accurate, resulting in a large uncertainty in Q^2. On the other hand, this method provides a good estimate of y in the low-y region, where the electron angle and energy carry little kinematic information.

5.3 The double angle reconstruction method

The double-angle method [92] uses the polar angles of the scattered electron, θ, and of the hadronic system, γ_h. The measurement of these angles is, in general, less affected by detector efects and resolution than measurements of energy which makes this method better suited to measure the kinematic variables in most of the available phase space. In particular, this method is, to a good approximation, independent of the calorimeter energy scale. It is, however, sensitive to the emission of initial and final state radiation, as is discussed in chapter 8. The double angle method was used in the present analysis analysis for the extraction of the kinematic variables.

The kinematic variables are given by the formulae:

$$Q^2_{DA} = 4\,E_e^2 \frac{\sin\gamma_h\,(1+\cos\theta)}{\sin\gamma_h + \sin\theta - \sin\,(\theta+\gamma_h)} \tag{5.11}$$

$$y_{DA} = \frac{\sin\theta\,(1-\cos\gamma_h)}{\sin\gamma_h + \sin\theta - \sin\,(\theta+\gamma_h)} \tag{5.12}$$

$$x_{DA} = \frac{E_e}{E_p}\frac{\sin\gamma_h + \sin\theta + \sin\theta + \gamma_h}{\sin\gamma_h + \sin\theta - \sin\,(\theta+\gamma_h)} \tag{5.13}$$

Figure 5.5 shows isolines of scattered electron angle, θ, and of the hadronic angle, γ_h, in the kinematic plane. As can be seen in the figure, the electron is scattered to high polar angles in low-Q^2 events, and to progressively lower angles as Q^2 grows. The hadronic angle is low for low-y (high-x) events and grows to higher values as the inelasticity increases.

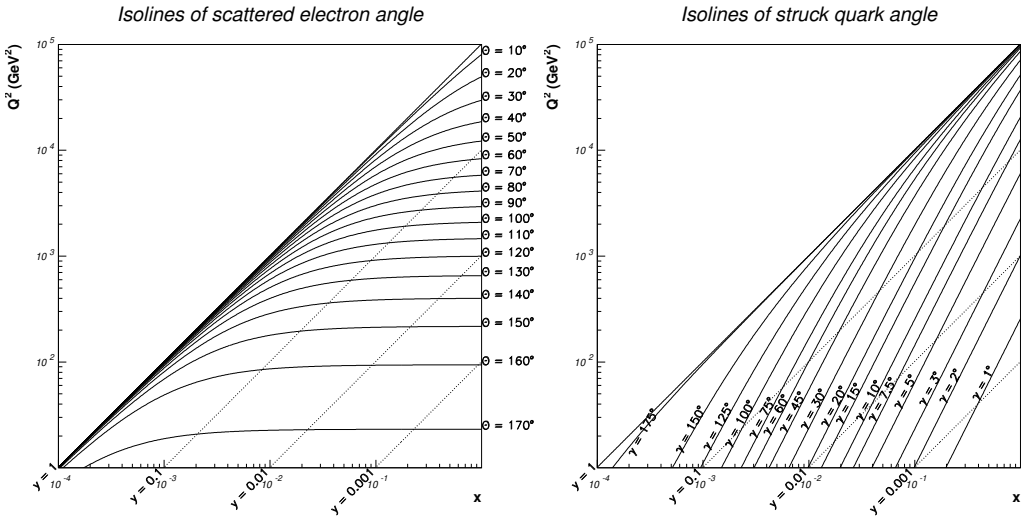

Figure 5.5: Lines of constant scattered electron polar angle, θ, and of the hadronic angle, γ_h.

A useful feature of this method is that an expression for the scattered electron energy, E'_e, is obtained from the formulae 5.11 to 5.13 given above, in which E'_e (usually written

$E_{e,DA}$ in this context) is a function of the measured angles θ and γ_h. The expression is:

$$E_{e,DA} = \frac{1}{2\,E_e}\,\frac{Q^2}{1+\cos\theta} = \frac{2\,E_e\,\sin\gamma_h}{\sin\gamma_h + \sin\theta - \sin\left(\theta+\gamma_h\right)} \tag{5.14}$$

This allows an estimate of variables such as the transverse momentum of the scattered electron, p_T^e, and of the hadronic system, p_T^h, that is in principle independent of the measured energies. The transverse momenta are given by the expressions:

$$p_T^h \equiv p_{T,DA}^e = E_{e,DA} \times \sin\theta = \tfrac{2\,E_e\,\sin\gamma_h\,\sin\theta}{\sin\gamma_h+\sin\theta-\sin\left(\theta+\gamma_h\right)} \tag{5.15}$$

$$\Leftrightarrow \qquad p_{T,DA} = \sqrt{Q^2_{DA} \times (1-y_{DA})} \tag{5.16}$$

The transverse momentum given by the double angle method was useful in studies of systematic effects as a minimum bias estimate of the true transverse momentum.

5.4 Resolution of the reconstruction methods

Figures 5.6 to 5.8 show the resolution and bias in the measured Q^2, x, and y, obtained with the three reconstruction methods, in bins of x and Q^2. The resolution of a measured kinematic variable was determined from Monte Carlo events generated with known values of the kinematic variables Q^2_{gen}, x_{gen}, and y_{gen}. It should be noted that the definition of these values is sometimes the object of some confusion. In radiative events, the presence of an ISR photon emitted from the electron line changes the centre-of-mass energy of the collision, $\sqrt{s}$. This should be taken into account. Also, the four momentum of FSR photons should be added to the electron four momentum when calculating the values of the generated values of the kinematic variables, x_{gen}, y_{gen} and Q^2_{gen}, in MC events. Alternatively, the kinematic variables should be calculated at the hadronic vertex. The quantity plotted in each histogram in figures 5.6 to 5.8 is the relative difference between the generated values of the kinematic variables and those measured with each reconstruction method after the MC events were passed through the full detector simulation. In the case of Q^2, x, and y, respectively, this can be written as:

$$\frac{Q^2_{gen}-Q^2_{DA/el/JB}}{Q^2_{gen}}\;;\;\frac{x_{gen}-x_{DA/el/JB}}{x_{gen}}\;;\;\frac{y_{gen}-y_{DA/el/JB}}{y_{gen}} \tag{5.17}$$

No explicit cuts on the kinematic variables were applied to the MC events used to determine the resolution. Other off-line selection cuts, described in section 6.5, were applied to the MC events. The resolution was determined in bins of x and Q^2 (limits marked on the plots) which are larger than the ones used for the extraction of the cross section, as only a general comparison between methods is intended. Histograms are displayed in the

approximate positions of the bins in the $x - Q^2$ plane. Events were assigned to each bin according to the values of their generated kinematic variables.

The shaded histograms correspond to the double angle method. The mean and rms width of the double angle histogram are indicated in each plot. The mean value, μ_{DA}, corresponds to the bias of the double angle estimate of the variable. This is in general quite small, being of the order of a few percent or less. The rms width, σ_{DA}, gives an indication of the resolution of the double angle method. A light solid line represents the electron method, and a dashed line represents the Jacquet-Blondel resolution. The normalisation of the three histograms is arbitrary. The double angle method can be seen to offer the best overall resolution and bias. In some parts of phase space, the other methods show comparable resolution. At high-y, the electron method shows comparable (and, in a few bins, better) resolution and bias than that of the double angle method. At low-y, the Jacquet-Blondel is, in general, the second best method to estimate x and y.

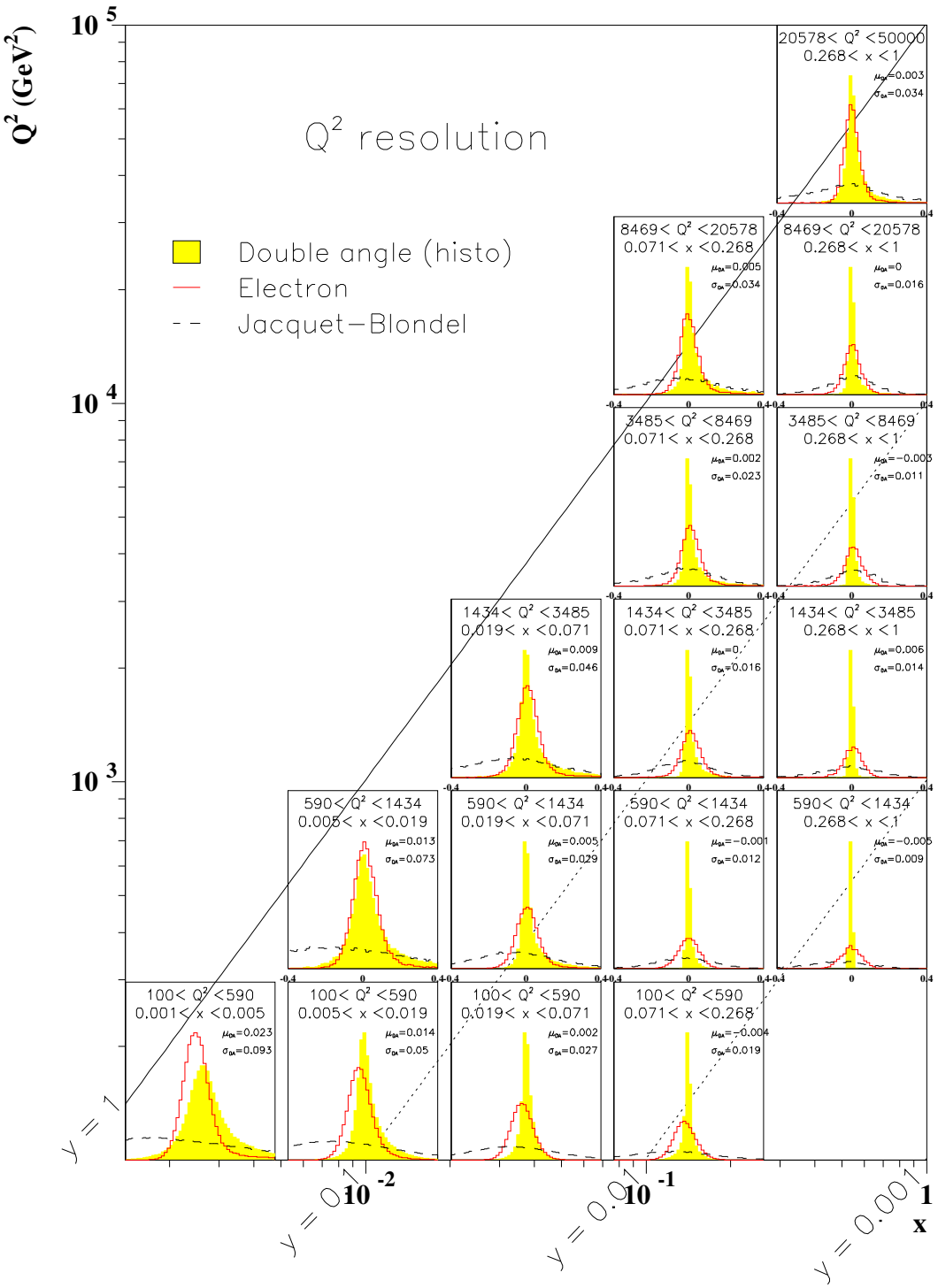

Figure 5.6: Resolution of the double angle, electron, and Jacquet-Blondel reconstruction methods in the measurement of Q^2. The double angle method can be seen to offer the best overall resolution and bias.

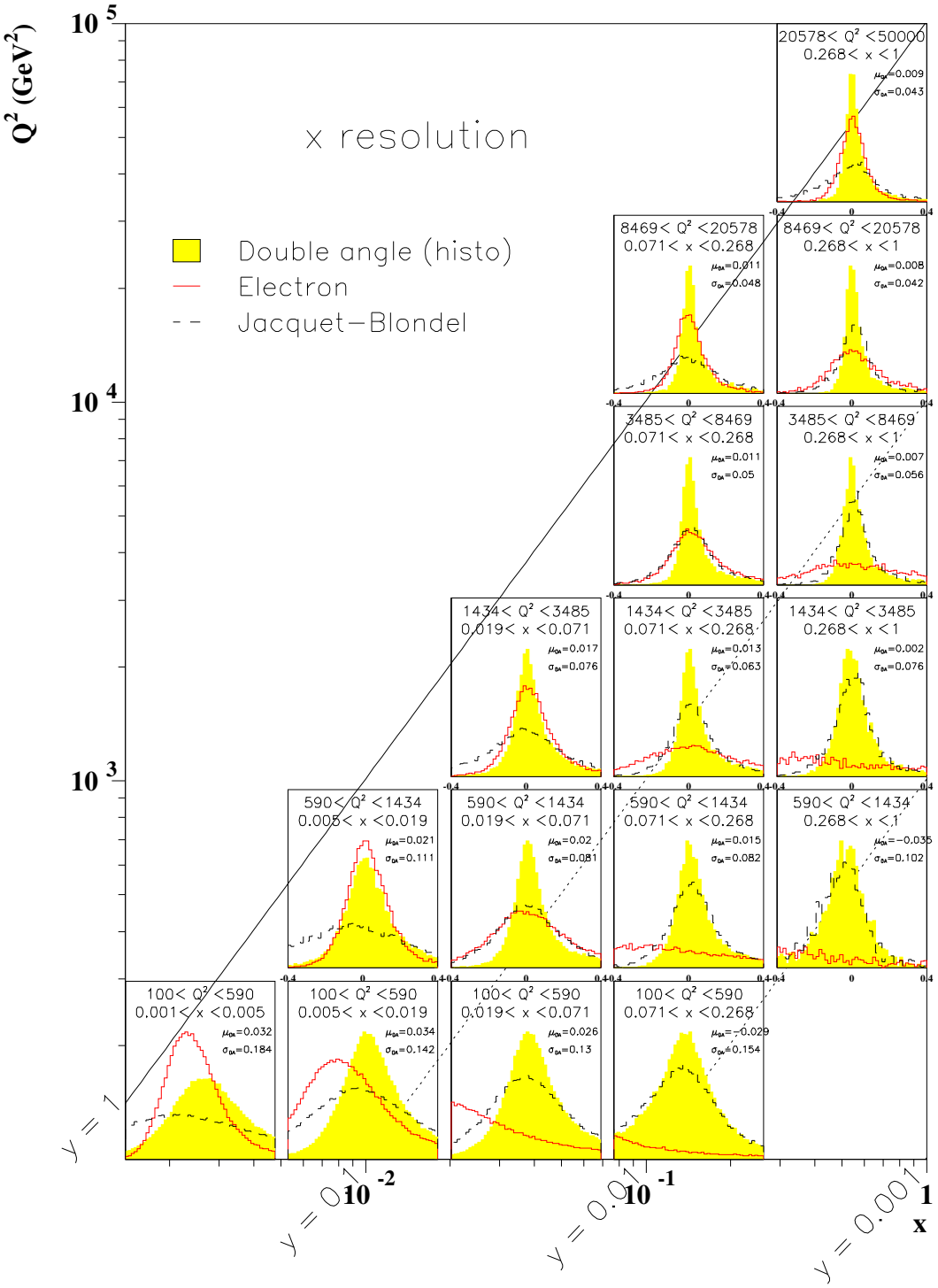

Figure 5.7: Resolution of the double angle, electron, and Jacquet-Blondel reconstruction methods in x. The double angle method can be seen to offer the best overall resolution and bias. The electron method shows a comparable resolution at high y, while at low y the Jacquet-Blondel method is second best.

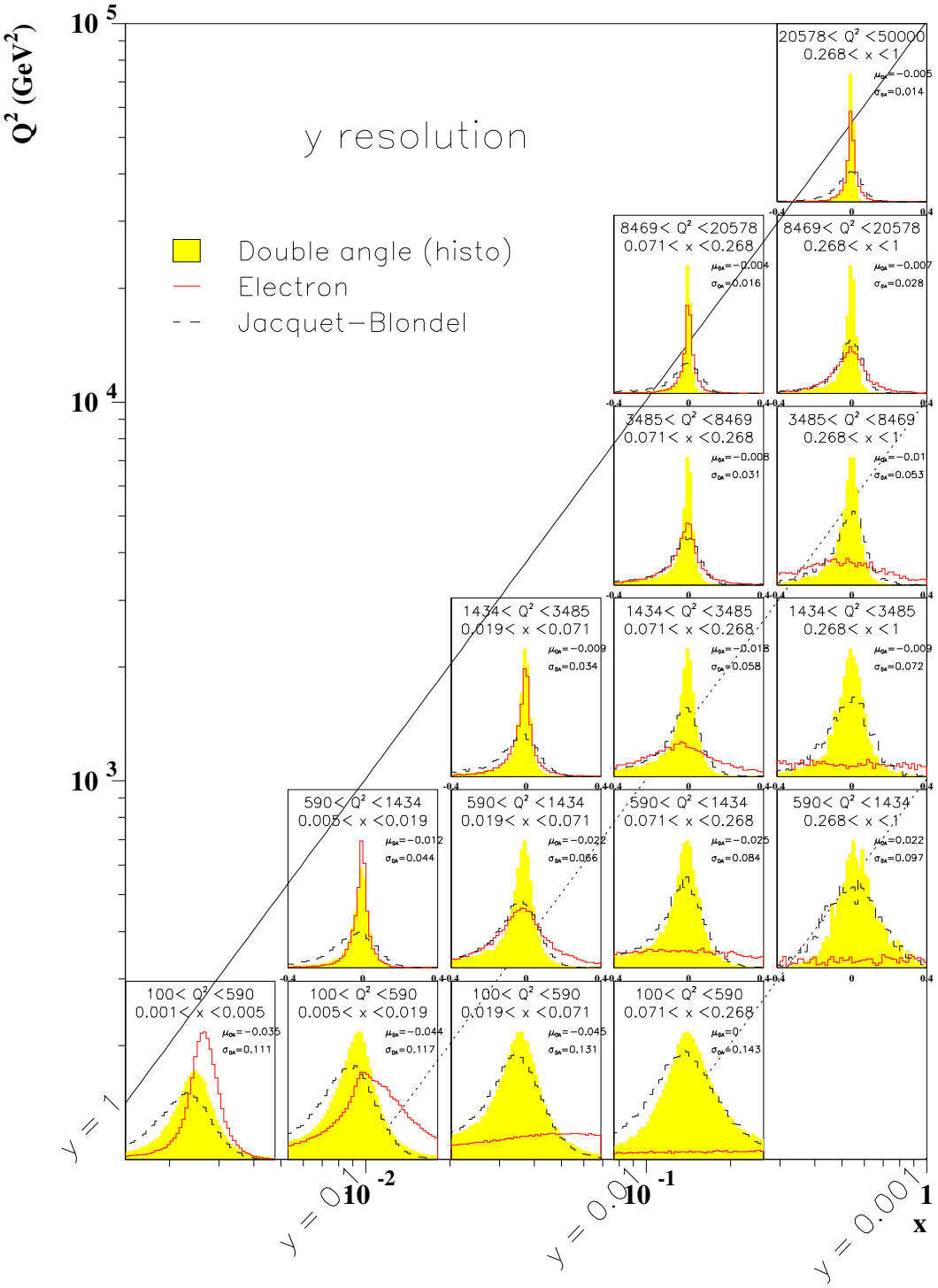

Figure 5.8: Resolution of the double angle, electron, and Jacquet-Blondel reconstruction methods in y. The double angle method, again, offers the best overall resolution and bias. The electron method shows better resolution in some bins at high y, while at low y the Jacquet-Blondel method is second best.

Chapter 6

Event Selection and Reconstruction

The signature of neutral current DIS events is the presence of a scattered electron in the detector. This electron must be correctly and efficiently identified among the many final state particles. NC DIS events must also be separated from much more frequent background processes which may resemble true NC DIS to a large extent. This chapter deals with the strategies used to select and reconstruct NC DIS events, both at the pre-selection level, performed online by the trigger, and in the off-line analysis.

6.1 Event characteristics

In low- and medium-Q^2 events, approximately in the range of 1 GeV^2 < Q^2 < 500 GeV^2, the scattered electron emerges from the interaction point at relatively high polar angles, and is detected in the RCAL. In a large fraction of these events the energy of the scattered electron is close to the beam energy. These are referred to as kinematic peak events. For higher-Q^2 events, the electron is scattered at lower polar angles and at higher energies, being detected more frequently in the BCAL and the FCAL. The transverse momentum of the electron is balanced by that of the hadronic final state. The proton remnant may carry a significant fraction $(1-x)$ of the original proton momentum and is lost through the forward beamhole, however, it carries negligible transverse momentum.

A high-Q^2 NC DIS event is shown in figure 6.1. In this event, the scattered electron is detected in the BCAL while particles from the hadronic final state are detected in the BCAL and the FCAL.

Several quantities can be used in the selection of events. The variable δ_{tot} proves very useful to this effect. It is defined as:

$$\delta_{tot} = \sum_i (E - p_Z)_i = \sum_i (E_i - E_i \cos\theta_i), \tag{6.1}$$

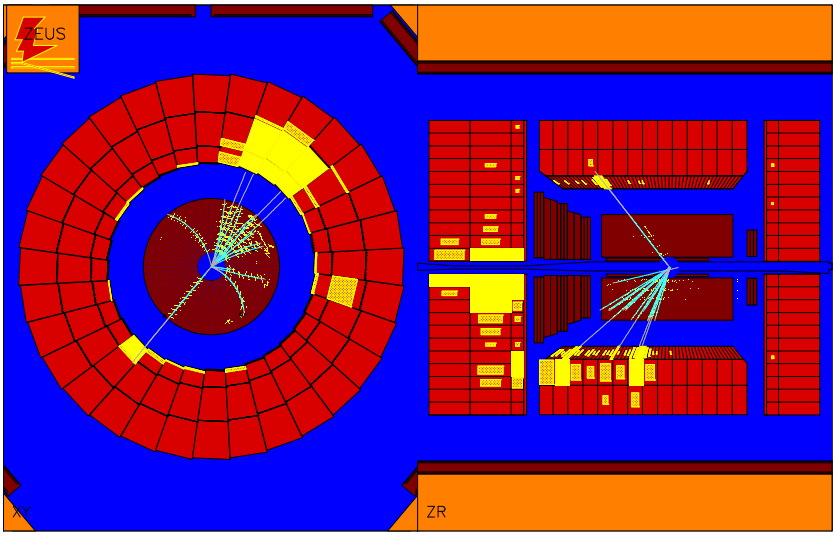

Figure 6.1: An example of a high-Q^2 event. The scattered electron leaves a track in the CTD and an energy cluster mainly in the EMC section of BCAL. Particles belonging to the hadronic final state balance the transverse momentum of the scattered electron and are detected in the BCAL and FCAL, leaving several tracks in the CTD.

where the sum runs over all energy deposits, E_i, in the calorimeter, with polar angles θ_i measured from the position of the calorimeter clusters and the interaction vertex. From momentum conservation, and neglecting the electron and proton masses, δ_{tot} must be conserved in the interaction and is equal to twice the energy of the electron beam, i.e.:

$$\delta_{tot} = E_e - (-E_e) + E_p - E_p = 2E_e \tag{6.2}$$

Particles that escape undetected through the forward beamhole contribute negligibly to δ_{tot}. The longitudinal momentum of such particles nearly cancels their energy in equation 6.1. This is the case for the proton remnant. Particles that are lost through the rear beamhole, contribute most strongly to the measured δ_{tot}, reducing it by twice the particle energy. This is the case both for photoproduction events (see section 6.2.1), in which the electron is scattered at polar angles very close to 180°, and for events in which the incoming electron emits a collinear energetic photon before the interaction (initial state radiation).

The net transverse momentum, p_T, and the transverse energy, E_T are also used in the event selection. These are defined as:

$$p_T^2 = p_x^2 + p_y^2 = (\textstyle\sum_i E_i \sin\theta_i \cos\phi_i)^2 + (\sum_i E_i \sin\theta_i \sin\phi_i)^2, \tag{6.3}$$

and

$$E_T = \textstyle\sum_i E_i \sin\theta_i, \tag{6.4}$$

where ϕ_i is the azimuthal angle and the sums, again, run over all energy deposits in the calorimeter. The net p_T of a neutral current event should be close to zero, since

the transverse momentum of particles escaping through the beamholes is small and the remaining particles have zero net transverse momentum due to momentum conservation.

6.2 Characteristics of the background

Background events can be mistaken for neutral current events during reconstruction. These events can originate from several physical processes. The most important of these are described below.

6.2.1 Photoproduction

While photoproduction (PHP) constitutes an important class of physics process, such events can sometimes be misidentified as high-Q^2 neutral current if one particle in the hadronic final state leaves a signal consistent with that of a scattered electron. Several variables can be used to reject PHP interactions. For example, for PHP events the distribution of δ_{tot} should peak close to zero. A lower cut on δ_{tot}, rejecting events with $\delta_{tot} < 38$ GeV, helps to eliminate photoproduction events. Also a cut on the scattered electron energy improves the purity of the electron selection and therefore contributes to the suppression of the PHP background. A third method consists of rejecting events in which the fake electron is surrounded by hadronic clusters in the calorimeter. The presence of such clusters indicates that the electron candidate may belong to the hadronic final state. Figure 6.2 shows an example of a photoproduction event.

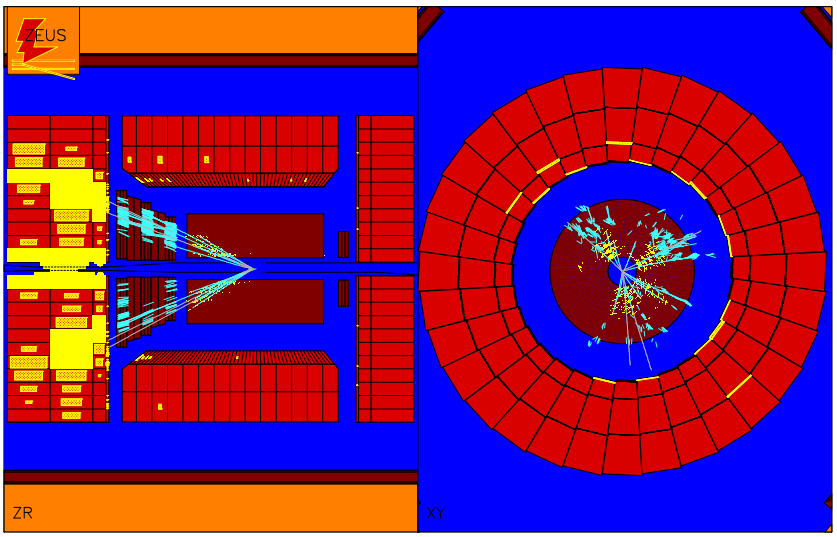

Figure 6.2: A photoproduction event showing a three-jet structure. The electron escapes through the rear beam hole.

Not all photoproduction events can be rejected, though, as some are indistinguishable

from DIS events. The number of background events remaining after all selection cuts have been applied must be estimated and subtracted from the selected sample.

6.2.2 Beam-gas interactions

Protons in the beam can interact with residual gas in the beampipe and produce events that are sometimes misidentified as NC DIS. If a beam-gas interaction occurs upstream of the detector the event can be rejected by the Veto Wall or by calorimeter timing cuts. Most particles in proton beam-gas events emerge from the interaction at low polar angles, leading to a small measured δ_{tot}. The lower cut on the measured δ_{tot}, used to reject photoproduction, is therefore also effective in rejecting beam-gas events.

Off-beam axis protons interacting with the beampipe itself can also be placed in this category. Unlike NC DIS, such events will have a large imbalance in the total transverse momentum and can be rejected by a cut on the total transverse momentum.

Electron beam-gas events can be indistinguishable from real ep events, since a scattered electron and an hadronic final state may be measured in the detector. The rate of such events must be estimated and subtracted from the selected sample. In the present analysis, this source of background was found to be negligible (see section 6.5.1).

Finally, due to the very high rate of beam-gas compared to the rate of physics events, proton or electron beam-gas events can occur at the same time as a NC DIS event. Such events, known as overlay events, can be rejected with an upper cut on the measured δ_{tot} at 65 GeV.

6.2.3 Halo and cosmic muons

The interaction of particles from the proton beam halo with residual gas or elements of the accelerator optics upstream of the detector creates muons through the decay of charged pions. The muons can pass through the veto wall and the RCAL, almost parallel to the beam axis and create a fake electron cluster if they interact in the EMC sections of the calorimeter. Such events will often have a large imbalance in the total transverse momentum if the muons leave energy deposits mainly in one side of the calorimeter. Energy deposits from these events will also arrive early at the RCAL and late at the FCAL compared to events originating inside the detector which reach both calorimeter sections at roughly the same time. The calorimeter timing can thus be used to reject such events.

High energy showers are produced in the interaction of cosmic rays with gas in the upper atmosphere. These constitute another significant background at HERA. Although the ZEUS detector is placed around 30 m underground, many muons from energetic showers can still reach it. Cosmic showers can also be rejected by the calorimeter timing, since

the muons will reach the upper part of BCAL before reaching the lower BCAL. Again, the total transverse momentum of these events will often be unbalanced if more energy is deposited in one side of the detector. Figure 6.3 shows a cosmic shower detected in ZEUS. Several high energy muons can be seen to go through the detector.

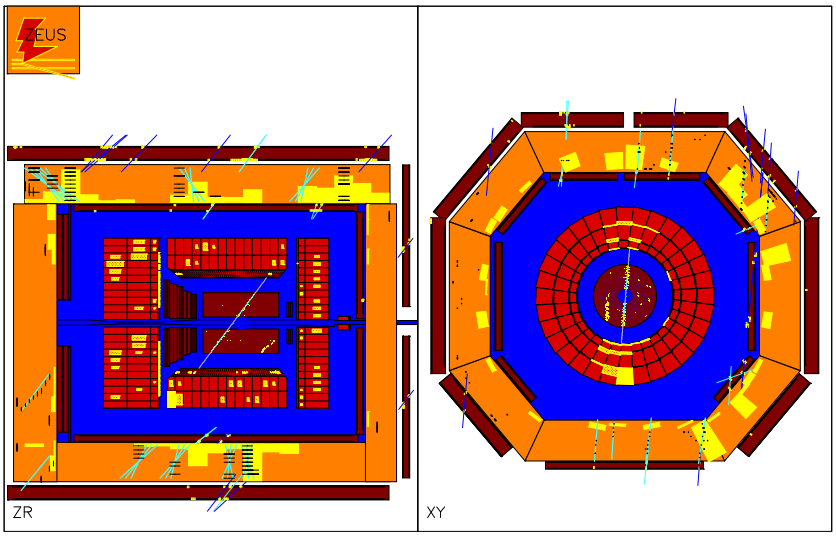

Figure 6.3: A cosmic shower, whith one muon track in the CTD. Muons are detected in the BAC and muon chambers as well as in the CAL and CTD.

6.2.4 Calorimeter sparks

Electrical discharges sometimes occur between the photocathode and the shielding of the calorimeter photomultipliers (PMTs) or in unstable or defective PMT bases. These bases contain a Cockroft-Walton generator that supplies the PMT high voltage. This is referred to as a calorimeter spark. Such sparks add random noise to an event and, if they occur in a calorimeter EMC cell, can result in the mis-identification of the scattered electron. The signature of such events is a large energy deposit in only one of the two PMTs that read each calorimeter cell. The large difference between the signals from the two PMTs in the cell, together with information from surrounding cells and the history of the PMT, are used by the trigger to reject such events. Figure 6.4 shows an event where a large spark in a BCAL EMC cell fakes an electron signal. In this case, the measured electron would have an energy of 1 TeV, which is kinematically impossible for a NC DIS event in HERA. The particular event shown here has since been corrected, during the re-processing of the 1998 data sample.

6.2.5 Elastic QED Compton events

Elastic Compton scattering corresponds to the reaction: $ep \rightarrow e'\gamma p$, where the initial or scattered electron radiates a hard photon (see chapter 4). The photon is included in

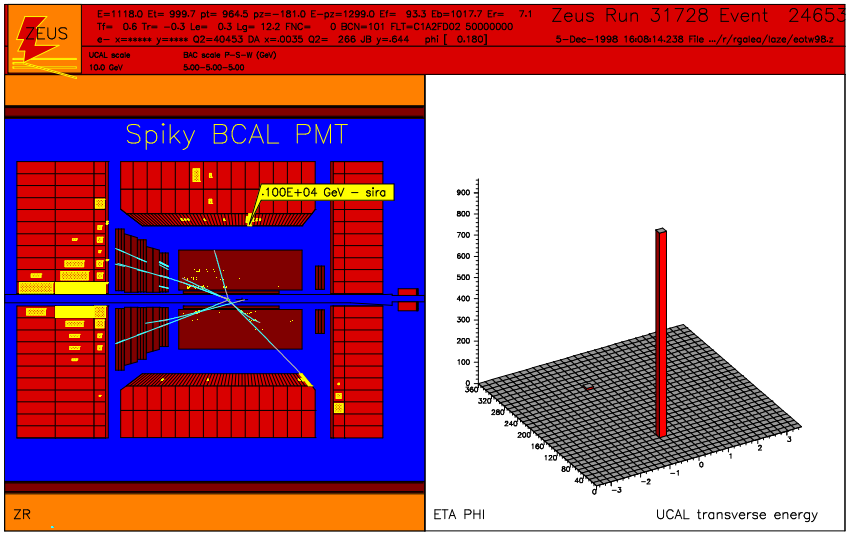

Figure 6.4: An event in which a calorimeter spark produces a fake signal corresponding to an electron with the (un-physical) energy of 1 TeV. All other energy deposits are small compared to this. The scattered electron probably corresponds to the energy cluster seen in the lower right-hand side of the BCAL.

the hadronic final state during the event reconstruction and measurement, resulting in a severely biased measurement of the kinematic variables. In *inelastic* Compton scattering, the proton breaks up resulting in a final state containing hadrons in addition to the scattered electron and photon.

This process has a well known cross section which is included in the Monte Carlo simulation for the inelastic case. Inelastic Compton events are, therefore, not rejected in the analysis. Elastic Compton events are not well modelled in the Monte Carlo and must be rejected. Their very clean topology makes this rejection straightforward, as the detected electron and photon must contain virtually all of the energy in the event. Figure 6.5 shows an elastic Compton event in which both the electron and the photon are clearly observed and there is little additional activity in the detector.

6.3 Data Preselection

The preselection performed by the ZEUS trigger aims to identify events where an isolated scattered electron is present while rejecting as much background as possible. It must combine the, sometimes conflicting, requirements of selecting a very pure sample of NC DIS events and of selecting events as efficiently as possible. The neutral current cross section varies by around six orders of magnitude in the Q^2 interval that is examined in this analysis. This makes the requirement for an efficient trigger very important at high Q^2, where the rate of events is very small. Previous measurements of NC DIS cross sections [33] were already limited by systematic errors up to a Q^2 of about 800 GeV^2, which enhances the interest in very high-Q^2 and the importance of an efficient trigger.

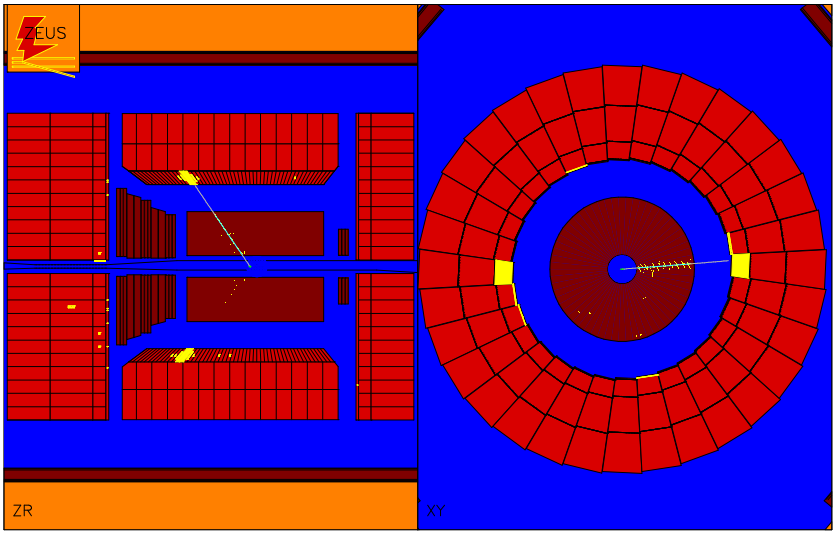

Figure 6.5: Elastic Compton event with a well measured electron leaving a track in the CTD and a photon measured in the lower BCAL.

6.3.1 First level trigger

The first level trigger provides several logical "bits" used in the event selection. The FLT decision is based on information from the CAL and the CTD. The calorimeter FLT information is generated by trigger towers, as described in section 3.2.6. Different requirements are applied in each CAL section, to allow for the varying topology of NC DIS events in different regions of phase space. Additional information is supplied by the CTD, which uses the fast z-by-timing system to search for charged tracks. Some events, known as "passthrough" events, are recorded irrespective of what trigger branches may accept them. These are randomly selected and used for trigger and detector studies.

An event was accepted by the FLT if one of the following conditions was satisfied:

- **REMC:** the total energy measured in the RCAL EMC section, excluding the first ring of towers around the beampipe ($\sim 60 \times 60$ cm^2), is greater than 3.4 GeV;
- **REMC-ISOe $\times$ REMCth sum:** a pattern of up to four trigger towers are found to have energy deposits consistent with either an electromagnetic shower (i.e. the energy in the EMC section is above 2.08 GeV and is at least 75% of the energy in the tower). This group of towers must be surrounded by quiet towers, where the measured energy is below 2.08 GeV and 0.95 GeV in the EMC and HAC sections, respectively. This condition, known as **REMC-ISOe** is determined by the CFLT according to the scheme shown in figure 6.6. In addition, it is required that the summed energies in RCAL towers with EMC energy above 625 MeV be greater than 3.75 GeV (**REMCth sum** condition). The combination of requirements allows the

use of a low energy threshold in the **REMC-ISOe** selection while efficiently rejecting background [94];

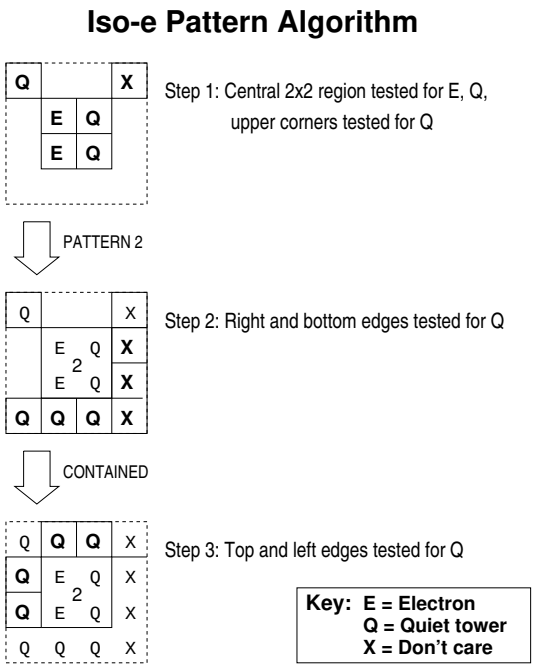

Figure 6.6: Algorithm used by the CFLT to identify isolated electrons ("ISO-e" trigger) [69]. Sets of 4×4 trigger towers are examined. The pattern recognition searches for one or more towers containing energy deposits consistent with the detection of an electron (**E**), and which are surrounded by "quiet" towers (**Q**).

- **BEMC:** The total energy detected in the BCAL EMC section is greater than 4.78 GeV. As electrons detected in the BCAL must cross the CTD, the existence of a reconstructed track is also demanded. This requirement reduces the background contamination from halo muons [95], that go through the BCAL almost parallel to the beam. No attempt to match the track to the energy deposits is made at this stage. Tracks are reconstructed using the z-by-timing system;

- **FEMC:** In very high-Q^2 events, the scattered electron is detected in the FCAL. In these cases, energy deposits from particles belonging to the hadronic final state make the selection of electrons more difficult. For such events, the hadronic system is either scattered at high angles or has high transverse momentum, $p_{T,h}$. Events were selected if one of the following conditions are met [95]: the transverse energy, E_T, is greater than 30 GeV (to select high-Q^2, high-$p_{T,h}$ events); the total energy in the EMC section of FCAL is greater than 10 GeV and a track is measured in the CTD (to select high-Q^2 events with a high hadronic angle); or E_T is greater than 11.6 GeV and a track is reconstructed in the CTD (an intermediate condition).

6.3.2 Second level trigger

At the second level trigger, more time is available to calculate event quantities than is available at the FLT. The SLT makes use of the calorimeter timing to reject background events. The timing is defined such that a particle emerging from the nominal interaction point, at $(x, y, z) = (0, 0, 0)$, travelling at the speed of light, is measured in the calorimeter at $t_{CAL} = 0$ ns. To do this, the time of signals from each calorimeter channel is corrected for the time of flight of relativistic particles emerging from the nominal vertex, as well as for all the delays connected to the development of showers, transmission of light pulses and of electronic signals, etc [61]. The timing of signals is calculated for individual calorimeter sections with a resolution of 1 to 2 ns.

The SLT also uses the energy information from the photon calorimeter in the luminosity monitor (Lumi-γ) to calculate δ_{tot}. Owing to time constraints, it is assumed at this level that the event vertex was at the nominal interaction point. This assumption results in a relatively poor resolution of the δ_{tot} measurement at the SLT stage.

Events are accepted by the SLT if all of the following requirements are satisfied [96]:

- $|t_{RCAL}| < 8$ ns: Particles created in proton beam-gas events occurring upstream from the detector travel together with the proton beam. These particles leave energy deposits in the RCAL which are early with respect to particles coming from the nominal interaction point by $\sim 2 \times 1.5$ m/c ~ 10 ns, where 1.5 m corresponds to the approximate distance from the nominal interaction point to the RCAL. The cut on $t_{RCAL} > -8$ ns removes such events. Events triggered by cosmic showers occur randomly. A cut on $t_{RCAL} < 8$ ns helps to reject these events. Figure 6.7 **a)** shows the RCAL time distribution measured with SLT passthrough events. The peak at -10 ns corresponds to proton beam-gas events;

- $|t_{FCAL}| < 8$ ns: the same arguments apply when electron beam-gas events and the FCAL are considered. With good beam conditions, many fewer electron beam-gas events are observed than proton beam-gas events (see figure 6.7 **b)**);

- $|t_{FCAL} - t_{RCAL}| < 8$ ns: events with a vertex close to the nominal interaction point should have $t_{FCAL} - t_{RCAL}$ close to zero. This cut helps to reject beam-gas as well as cosmic and halo muon background (see figure 6.7 **c)**). Historically, this cut was motivated by the very long (~ 40 cm) proton bunches accelerated by HERA in early running, due to a lower RF frequency than is used at present. This caused a broadening of the t_{RCAL} distribution which was compensated by using $t_{FCAL} - t_{RCAL}$ [61];

- $t^{up}_{BCAL} - t^{down}_{BCAL} > -10$ ns: cosmic ray muons go through the upper BCAL approximately 12 ns before crossing the lower BCAL. This cut removes such events (see figure 6.7 **d)**);
- $\delta_{tot} + 2 \times E_{Lumi-\gamma} > 29$ GeV: this cut selects NC DIS events. The cut is placed at a reasonably low value because the SLT estimate of δ_{tot} has limited accuracy. This cut is tightened in the TLT and in the off-line analysis. Events with a hard ISR photon pass this cut if the radiated photon is detected in the photon calorimeter of the luminosity system (Lumi-γ). In this case, the $E - p_z$ of the photon is included in the sum;
- $\delta_{tot} < 100$ GeV: very high $E - p_Z$ indicates the presence of overlaid events. The overall conserved δ_{tot} of two NC DIS events occurring in coincidence is 110 GeV.

In addition to these cuts, the second level trigger also rejects empty events, triggered only by the CFLT but having little energy in the calorimeter and searches the calorimeter for sparks.

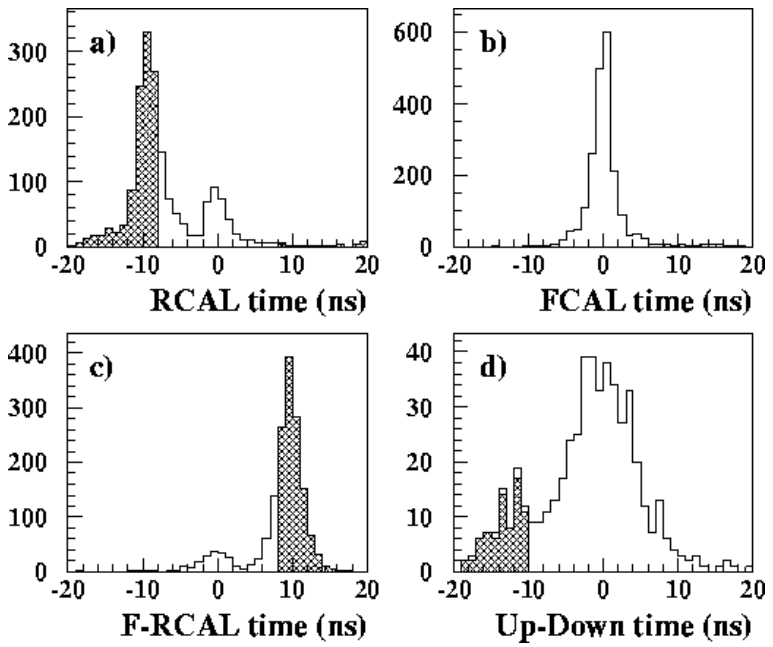

Figure 6.7: Calorimeter time distributions from second level trigger "passthrough" events [96]. The SLT timing cuts are illustrated by the hatched areas. The histograms shown were obtained from passthrough events and thus show events which are rejected by normal SLT cuts (hatched areas).

6.3.3 Third level trigger

Information from all components is available to the TLT. The longer processing time available also allows more sophisticated algorithms to be used. At the TLT a simplified

version of the off-line tracking package, VCTRAK [97, 72], is run. As a result, the z position of the vertex is known and can be used to perform a better calculation of δ_{tot}. Various electron finder algorithms are also run, to try to identify the scattered electron and classify the event as NC DIS. In addition, information from the muon chambers is used together with CTD tracks and CAL information to search for halo and cosmic muon events [72].

At the TLT, different sets of cuts exist which are tuned to select different physics channels. NC DIS events may thus be selected by a low-Q^2 NC DIS trigger bit with some cuts on the electron candidate energy or position. Such events fail to pass a high-Q^2 trigger selection. Typically, a logical *or* of several TLT bits, together with additional cuts on measured quantities, is available off-line as a single "data-summary-tape bit" (DST bit).

The DST selection required for the event pre-selection was:

- An electron candidate is identified by any of five different electron finder algorithms, with an energy greater than 4 GeV;
- $\delta_{tot} + 2 \times E_{Lumi-\gamma} > 30$ GeV using the measured vertex position;
- $Q^2_{DA} > 80$ GeV^2, where Q^2_{DA} is reconstructed with the Double Angle method (see section 5.3). The TLT bits used in this selection actually apply fiducial cuts on the position of the electron candidate. These cuts are less restrictive than the requirement of $Q^2_{DA} > 80$ GeV^2 applied by the DST selection;
- $y_{el} < 0.95$, where y_{el} is reconstructed with the electron method (see section 5.1). The aim of this cut is to eliminate photoproduction (see section 6.5).

In addition to these cuts, the TLT vetoes beam-gas and cosmic-shower events based on the calorimeter timing, using similar cuts to those implemented by the SLT, and on the information from the muon chambers. It also rejects events where sparks are found in the calorimeter with little additional deposited energy.

Events that pass the TLT selection are stored on tape to be analysed off-line.

6.3.4 Data quality

Additional criteria were used in the selection of the data used in the present analysis. The purpose of these was to ensure that the used events were recorded with all essential parts of the detector in operation.

A minimal set of requirements for data to be used in the analysis is defined, known as the EVTAKE selection. These requirements are:

- The CTD must be at full operating voltage and with no large dead regions;

- The luminosity monitor must be fully functional;
- The magnetic field must be on;
- The CAL must be operational without a significant number of dead channels.

Data were used in the present analysis only if it passed the EVTAKE selection. Additionally, monitoring information for each sub-detector is used to reject runs where detector problems may affect the quality of the data.

6.4 Off-line event reconstruction

Events that survive the trigger preselection and data quality monitoring process must still be corrected for a number of detector effects. This is done in the off-line analysis, where the most precise methods can be used to obtain the best possible resolution that the ZEUS detector can offer. Correction factors such as the calorimeter energy scale or alignment can be most accurately determined from large samples of physics data. After some corrections have been applied, and local energy clusters reconstructed, the scattered electron is again searched for. Finally, the hadronic final state characteristics are measured. This section describes the methods used to correct the most important systematic effects and measure the event variables with the best achievable accuracy.

The final event selection is done using the corrected variables. The selection cuts are described in section 6.5.

6.4.1 Calorimeter corrections

Noise, hot cells and holes

Spurious signals in calorimeter cells can severely affect the measurement of kinematic quantities, since most types of analysis rely heavily on calorimeter measurements. This is especially true for events where little hadronic activity is observed.

A first calorimeter noise suppression is done during the event reconstruction by the ZEPHYR package (see section 3.3). At this stage, cells are eliminated from the event if the measured energy is less than 60 MeV for EMC cells and 110 MeV for the (larger and therefore more noisy) HAC cells. These cuts correspond to a cutoff at around four standard deviations of the noise distribution [98]. Stricter cuts of 100 MeV and 150 MeV for EMC and HAC cells, respectively, were applied in the present analysis to isolated cells where an energy deposited was measured.

Noise signals can originate not only from the uranium radioactivity but also from the PMTs that read each cell or from the signal processing electronics. In such cases, one of the

PMTs or electronics channels will usually fire while the other does not. Studies [99, 100] of the calorimeter noise were done using passthrough events and empty Monte Carlo events. These studies have shown that the imbalance, E_{L-R}, in the energy readings from the two PMTs that read each cell can be used to eliminate noisy cells from further processing. The imbalance cut that was used can be expressed as $|E_{L-R}| = |E_L - E_R| < 0.49 \times E_{L+R} + 0.03$ (GeV), where E_L and E_R are the energies measured by the left and right PMTs, respectively, and E_{L+R} is the total energy measured in the cell.

After this cut is applied, some discrepancies still remain between data and Monte Carlo. These are due to calorimeter "hot" cells, which fire more frequently and measure more energy than the average for the calorimeter section to which they belong. Such cells are identified by the calorimeter group during the data quality assessment. To this end, the average energy and frequency of firing of calorimeter cells in randomly selected events is calculated and cells are marked as hot if their behaviour is significantly different from the average. The list of hot cells is used in the noise suppression algorithm either to eliminate the cell information or to assign to it twice the energy of the operational PMT.

The inverse problem also occurs; one or both of the PMTs that read out a cell may be dead. If only one channel is not working, then twice the energy of the second PMT is assigned to the cell. This cell becomes a "bad cell". If both channels are dead (this is known as a "hole") the information from the cell is lost. Bad channels and holes can only be repaired during shutdowns but they account, in general, for a very small fraction of the 11836 calorimeter channels, being of the order of a few holes and a few tens of bad cells.

If one of the cells that is attributed to the scattered electron cluster is a known bad channel, a more sophisticated procedure is used which is explained in section 6.4.3.

Calorimeter alignment

The position of the calorimeter relative to the ZEUS reference frame must be known to a very good accuracy. As described below, the polar angle of the scattered electron can either be given by the reconstructed electron track or determined using the event vertex together with the reconstructed position of the electron in the calorimeter. This angle enters directly into the calculation of the kinematic variables, thus making a good alignment of the calorimeter very important. Also, the matching of a reconstructed track to the calorimeter cluster position, which relies on a good knowledge of the calorimeter position, is used off-line to reject background.

The three calorimeter sections, being physically independent, must be aligned separately. The natural way to do the alignment is to use the electron track reconstructed in the CTD and projected to the face of the calorimeters, since the CTD sets the ZEUS reference frame. The endpoint of the projected track can then be compared to the calorimeter

reconstructed position of the electron. Any differences found can be used to determine the relative alignment of CTD and calorimeter.

In practice, too few scattered electrons are detected in the FCAL to allow a precise alignment. In the BCAL, two degrees of freedom were studied [90]: a shift of BCAL along z and a rotation around the z axis. These were chosen because they contribute more strongly to the measurement of the kinematic variables, either through the measurement of the electron polar angle or due to a cut on the distance between track endpoint and cluster position (track-cluster matching). In the RCAL, both halves of the calorimeter section must be studied separately, since they are physically separate.

The determination of the BCAL position in z makes use of the projective geometry of the BCAL EMC section. While this feature of the BCAL can lead to some mismeasurement of the electron energy, which is corrected later (see section 6.4.3), it can show clearly where the cracks between cells are located. The scattered-electron energy can be calculated from the electron and hadronic final-state angles, θ and γ_h, respectively, according to equation 5.14. Figure 6.8 shows the ratio of the measured electron energy, E_{cal}, to the calculated one, E_{DA}, for both real data and Monte Carlo simulated data. The ratio is plotted against the z position of the electron track extrapolated to the face of BCAL. The structure evident in the plots comes from the gaps between EMC cells, where some energy is lost in interactions with dead material. An inverted gaussian is fitted to each dip both in data (plot **a)**) and Monte Carlo (**b)**). The average of the distributions is compared by plotting their difference as a function of the crack number in plot **c)**. The average offset is -0.03 ± 0.03 cm, which is consistent with a perfect agreement between the z measured position of the BCAL and the value used in the simulation of the detector. Events from the 1998/99 data sample, that passed all neutral current selection cuts, were used for this study.

A similar exercise is done to align the BCAL in azimuthal angle, ϕ. The calorimeter geometry is non-projective in ϕ, as the 32 wedge shaped BCAL modules are rotated by 2.5° around an axis close to the module centre and parallel to the z axis. This, together with the module side cover, suppresses the production of Cherenkov light by showers developing in the wavelength shifters, which would create very large non-uniformities of response for electrons reaching the calorimeter close to the intermodule gap [58]. Some non-uniformity can still be detected, where the measured electron energy is a few percent higher close to the intermodule gap. This was used to verify the alignment of the BCAL in ϕ which was found to be the same in data and Monte Carlo to an uncertainty of 0.6 mrad [90]. Table 6.1 shows the results of the alignment of BCAL and RCAL.

Several factors make the alignment of the RCAL more difficult. Due to the limited CTD acceptance in the area covered by the RCAL, only electrons that reached the calorimeter

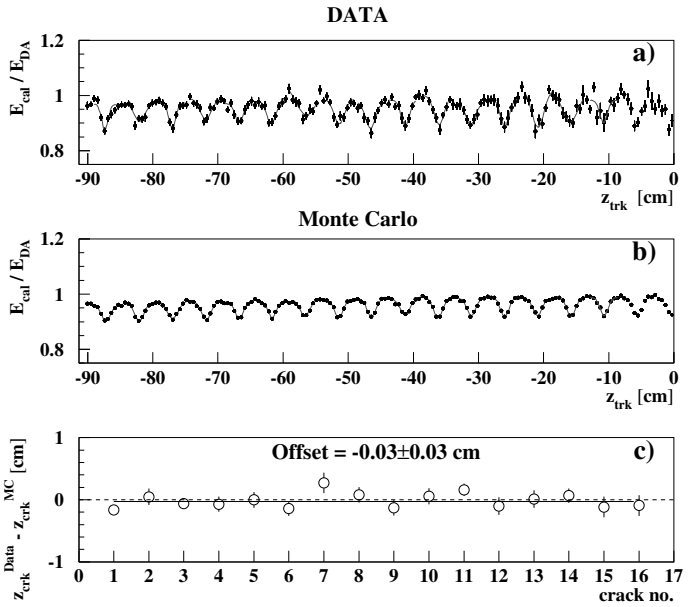

Figure 6.8: BCAL alignment along z. The ratio of the scattered electron energy measured in the BCAL, E_{cal}, and calculated energy, E_{DA}, versus the z-coordinate of the extrapolated electron track is plotted for data (**a**)) and Monte Carlo (**b**)). The position of the observed dips is used to measure the offset between the z position of the real calorimeter and the simulation (**c**)). Reproduced from reference [90].

BCAL		RCAL	left half	RCAL	right half
δz	-0.3 ± 0.3 mm	δz	-2.2 ± 0.4 mm	δz	0.3 ± 0.7 mm
$\delta \phi$	0.0 ± 0.6 mrad	$\delta \phi$	1.7 ± 0.9 mrad	$\delta \phi$	0.0 ± 0.8 mrad
		δx	1.6 ± 0.2 mm	δx	0.9 ± 0.3 mm

Table 6.1: Shifts and rotations found in real data with respect to the MC simulation for the barrel and rear calorimeter sections [90].

at a radius greater than about 70 cm could be used for this study. In addition, especially for these high radii, the non-projective geometry of the RCAL suppresses the structure that one might expect from the inter-cell gaps. Also, any small bias in the RCAL electron position reconstruction would distort the measurement. The electron position given by the HES was used instead and compared to the electron track endpoint which was extrapolated to the HES plane, 7.1 cm inside the RCAL. The z position of the HES is more accurately known than the reconstructed position of the electron cluster. Each HES "ski" is inserted vertically in a calorimeter module, so a possible vertical shift must be determined beforehand for each module. This makes the determination of the RCAL position in the y direction unreliable. On the other hand, shifts in the calorimeter position in x, z and in the azimuthal angle ϕ were determined with accuracies better than 1 mm and 1 mrad [90]. These corrections, listed in table 6.1, are applied to the measured positions of the scattered electron in real data. Figure 6.9 shows the difference in the electron position along x measured in the HES and extrapolated from the electron track reconstructed in

the CTD, $x_{HES} - x_{trk}$, as a function of x_{HES}, after all corrections were applied.

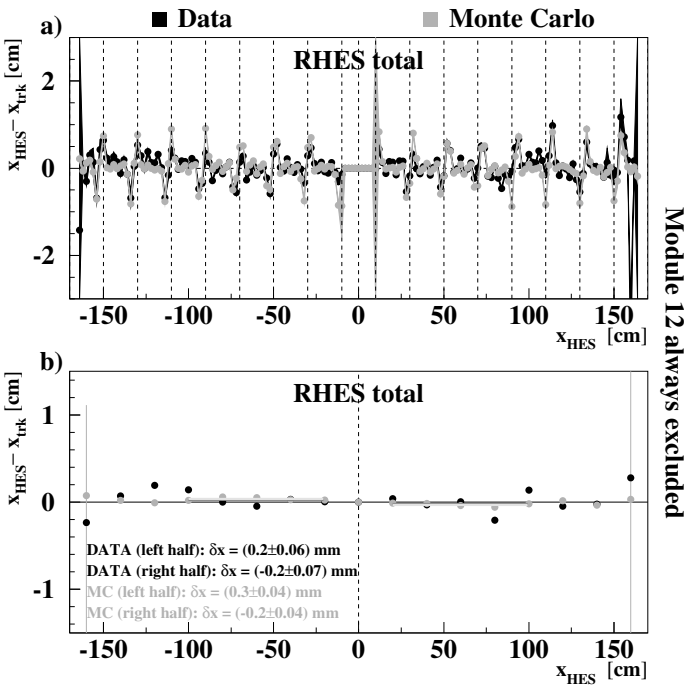

Figure 6.9: RCAL alignment using the HES after all corrections have been applied. The difference in the x position of the scattered electron measured with the HES (x_{HES}) and extrapolated from the electron track measured in the CTD (x_{trk}) is plotted versus x_{HES}. The dashed limes represent the boundaries between calorimeter modules. The bottom plot shows the average shift in x measured in each module. Module 12 is the central module of the calorimeter. Its position was verified independently. Reproduced from reference [90].

Local clustering

The aim of local clustering is to determine groups of calorimeter cells that may be attributed to the showering of one or as few as possible particles. Building clusters compensates for the finite granularity of the calorimeter by averaging information from several cells to obtain a good position resolution.

At the reconstruction phase, in the ZEPHYR package, local clusters of cells, called *condensates* [98, 101], are formed by ordering calorimeter cells in decreasing energy and using the highest energy cell as a seed for a condensate. Adjacent cells (i.e. cells that share a common side) are added to the condensate and erased from the list of available cells if their measured energy is above a predefined cut. The condensate grows until no other adjacent cell has energy above the threshold. The process is repeated until no cells are left in the list. The sum of the constituent cell energies is attributed to the condensates. The position of the condensates is simply the energy weighted average of the cell positions. Condensates do not take into consideration the possible showering of

particles in neighbouring calorimeter sections. Also, no attempt is made to join diagonally adjacent cells.

A refinement of the condensates algorithm is the formation of *islands* [102]. These form a local maximum of measured energy surrounded by lower energy deposits. To build an island, the energy measured in each calorimeter tower is compared to the energy measured in its immediate neighbours. Both towers that share a side and diagonally adjacent towers, which share only an edge, are considered in this step. If a higher energy neighbour is found, a link is established to that neighbour and the cell is included in the island, otherwise the tower will be associated with itself. The process is repeated until all towers have been included in an island or form an island by themselves. Islands formed using this algorithm are used, for example, in the first stage of the SINISTRA electron finder [103] (see section 6.4.3). Thus, while condensates may show two energy peaks separated in space, such a group of cells would be divided into two islands (see figure 6.10). Unlike condensates, islands may contain towers from different sections of the CAL, spanning the super-cracks. Showers developing in the super-crack regions between FCAL and BCAL or between BCAL and RCAL are thus correctly reconstructed.

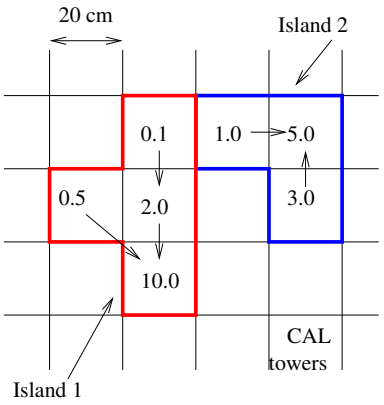

Figure 6.10: Schematic representation of tower island clustering. The arrows illustrate a link from each calorimeter tower to its highest energy immediate neighbour. The numbers in each tower represent the measured energy. In the present case, two islands are formed with two local maxima and a valley in the middle. Note that one single condensate would be formed in this situation. The same process can be done using cells instead of islands.

The same procedure can be applied separately to calorimeter cells in each of the EMC/HAC1/HAC2 layers of the calorimeter. Islands formed from cells in EMC, HAC1 or HAC2 are referred to as *cell islands* [102].

Once cell islands are formed, they can be linked to islands in other layers of the calorimeter to form *cone islands* [102]. The matching starts from HAC2 and works inwards, towards the measured event vertex, matching HAC2 cell islands to HAC1 or EMC islands, then HAC1 to EMC, and finally EMC islands to other EMC islands. The angular

separation between cell islands in different layers is translated into a probability that they came from the showering of a single particle. The parametrisation of the probability was determined from Monte Carlo simulations of single pions showering in the calorimeter. Cell islands in different layers are combined if the probability obtained from their angular separation exceeds a predefined cut.

To determine the position of the either cell or cone islands, a cell "centre" is first determined by using the signal measured in each of the two PMTs reading each cell, with a parametrisation obtained from test beam results. The position of the island is then calculated as a weighted average of the centres of the constituent cells where the weights depend logarithmically on the cell energy according to [104]:

$$\vec{r}_{isl} = \frac{\sum_i w_i \vec{r}_i}{\sum_i w_i} \tag{6.5}$$

$$w_i = \max(0, W_0 + \ln \frac{E_i}{\sum_j E_j}) \tag{6.6}$$

where $\vec{r}_{isl}$ is the position assigned to the island, w_i are the weights attributed to each cell where the energy E_i was measured and the sums run over all the cells in the cluster. The logarithmic dependence of the weights takes into account the radial exponential falloff of the shower energy. A tunable parameter is used, W_0, which was adjusted with different values for EMC and HAC cells. This parameter sets a threshold on the minimum fraction of the shower energy in the cells used in the position determination and it determines the relative weight of the tails of the shower in the position weighting. Values of $W_0^e = 4$ and $W_0^h = 2$ were found to be appropriate for EMC and HAC cells, respectively [102]. The minimum fraction of the shower energy measured in used cells is given by e^{-W_0}, resulting in 1.8% and 13% for EMC and HAC cells, respectively.

6.4.2 Track and vertex reconstruction

The ZEUS off-line track reconstruction is performed by the VCTRAK package and is described in detail in elsewhere [97, 56]. A brief outline is given in this section.

Charged tracks can be reconstructed either with detected hits in the CTD only or using additional hits from the forward and rear tracking detectors. In the present analysis, the "CTD-only" mode was used, which avoided the need for additional alignment between detectors.

The reconstruction of tracks in the CTD starts by using track segments formed by three hits in one of the outer axial superlayers. CTD hit positions are determined from the digitized times of signals from each wire. The determination of the position of the hit takes into account the drift velocity and the Lorentz angle and also of the time, t_0, when tracks from the vertex should cross the plane of the wire where the drift occurs. The

track seed is extrapolated inwards, towards the beampipe, with the help of a virtual hit at $x = y = 0$. The track candidate picks up hits as it advances to the centre of the CTD, and the measured parameters become more precise. In each CTD cell, the position of hits is ambiguous because the ionisation charge can drift from either direction to the wire (the cell measures true hits in one direction and *ghost* hits in the other). This left-right ambiguity is lifted as more hits are added to the tracks, since the hits in the ghost track will not be aligned. The tilt in the cell plane with respect to the CTD radius makes an average track cross two cells in each superlayer, which helps eliminate the ghost hits. The hits from axial superlayers form an arc in the $x - y$ plane.

Information from axial superlayers equipped with z-by-timing is then used, if available, to limit the region in which the search for stereo hits is performed. Hits in the stereo layers assigned to the track are used together with the z-by-timing information to determine the z-position and polar angle of the track. A five-parameter helix fit is used to determine the parameters of each track. The pattern recognition has two iterations. In the first, only tracks that reach down to superlayer 1 are accepted. In a second pass, other tracks are reconstructed if they cross at least two axial and one stereo superlayers. The tracks found by VCTRAK are extrapolated outwards to the face of the calorimeter with the help of a detailed map of the magnetic field.

Reconstructed tracks are used to find the event vertex. Since 1998, the default vertex finding algorithm searches for secondary vertices in addition to searching for the primary event vertex. A fraction of the tracks observed in the detector arise from particle decay or secondary interactions with material such as the beampipe. The multi-vertex algorithm avoids biases in the vertex reconstruction by assigning such tracks to secondary vertices.

Tracks are first filtered to eliminate those which are not compatible with the beamline and to group tracks in clusters which are loosely consistent with a common vertex. Each track is, at this stage, assigned to as many vertices as possible. Next, tracks in the resulting clusters are extrapolated through the magnetic field to perigee with the beamline. A fit is then performed on each cluster to determine the common vertex. Tracks are dropped if they contribute too much to the vertex χ^2. The vertices of surviving track clusters are potential primary or secondary vertices. At this point, in the standard multi-vertex reconstruction, a loop is performed over all tracks in which each of them can be the seed of the primary vertex. The seed track is assumed to originate in the primary vertex and the whole set of reconstructed tracks are analysed together with a diffuse pseudo-beam (centered at the beam spot in the $x - y$ plane and with $\sigma_x = \sigma_y = 0.7$ cm). Tracks are dropped from vertices until no track belongs to more than one vertex and the tracks assigned to secondary vertices point away from the primary vertex. The result of the loop is a set of vertices to which a figure of merit has been given based on criteria such as the

χ^2, number of degrees of freedom and the number and length of tracks that originate in each vertex [56]. Other criteria, such as the dimensions of the beampipe, for example, are taken into account. The best candidate is selected as the primary vertex.

Each vertex, primary or secondary, and its associated tracks is input to a fast "full fit" algorithm [105] where an accurate vertex position is evaluated and tracks are constrained to emerge from it. Track momenta are re-evaluated in this process.

The resolution of the primary vertex in x and y was determined from Monte Carlo to be of the order of 1 mm. This is larger than both the known transverse size of the beams and the run-to-run variations of the mean beam position in the $x - y$ plane. As the coordinates of the vertex in x and y have negligible errors in the determination of the kinematic variables, these were set to zero in the present analysis.

"Electron-seed" vertex

In the present analysis, a further improvement to the multi-vertex algorithm is achieved by using the identified scattered electron track as the seed to the primary vertex finding [95]. The electron track is required to belong to the primary vertex if a well measured track is found that matches the electron cluster in the calorimeter (distance of closest approach less than 20 cm) and has momentum greater than 5 GeV. A lower limit of 8 hits in stereo-wire layers is required to ensure that the extrapolation of the electron track to the vertex is accurate. The χ^2 returned by the vertex finding algorithm is required to be less than 5 times the number of degrees of freedom for the electron-seed vertex to be used. Otherwise the standard multi-vertex result is used. The use of the electron-seed vertex requires VCTRACK to be called again after the scattered electron has been identified.

As the scattered electron must come from the primary vertex in NC DIS events, at sufficiently high Q^2 the ambiguity as to which vertex is the primary is therefore resolved. This method also results in a better resolution on the vertex position than simply using the point of closest approach between the beamline and the electron track, because the extra tracks from the hadronic final state help in reducing the uncertainty on the vertex position [95] (see figure 6.11 **d)**). A comparison between the standard multi-vertex reconstruction done by VCTRAK and the electron-seed multi-vertex algorithm is shown in figure 6.11 **a)**. This comparison was performed using Monte Carlo generated events without any analysis selection cuts applied. The electron-seed method has a slightly better resolution than the normal algorithm when a good electron track exists. Figure 6.11 **b)** shows the vertex resolution determined using MC events that passed all off-line analysis cuts (see section 6.5). Here, events where a good electron track was used to determine the primary vertex are combined with events where the vertex was determined by the standard algorithm. The latter (scaled up by a factor of 5) is also shown as a darker

shaded histogram. The vertex resolution for all selected events is around 3 mm and has a negligible bias.

A small fraction of the selected events are still reconstructed far from the true generated vertex. These are shown in the figure 6.11 **c)** as a probability for reconstructed vertices to be a certain distance from the true vertex. Figure **d)** shows the resolution of the electron-seed vertex finding method as a function of the number of tracks assigned to the primary vertex.

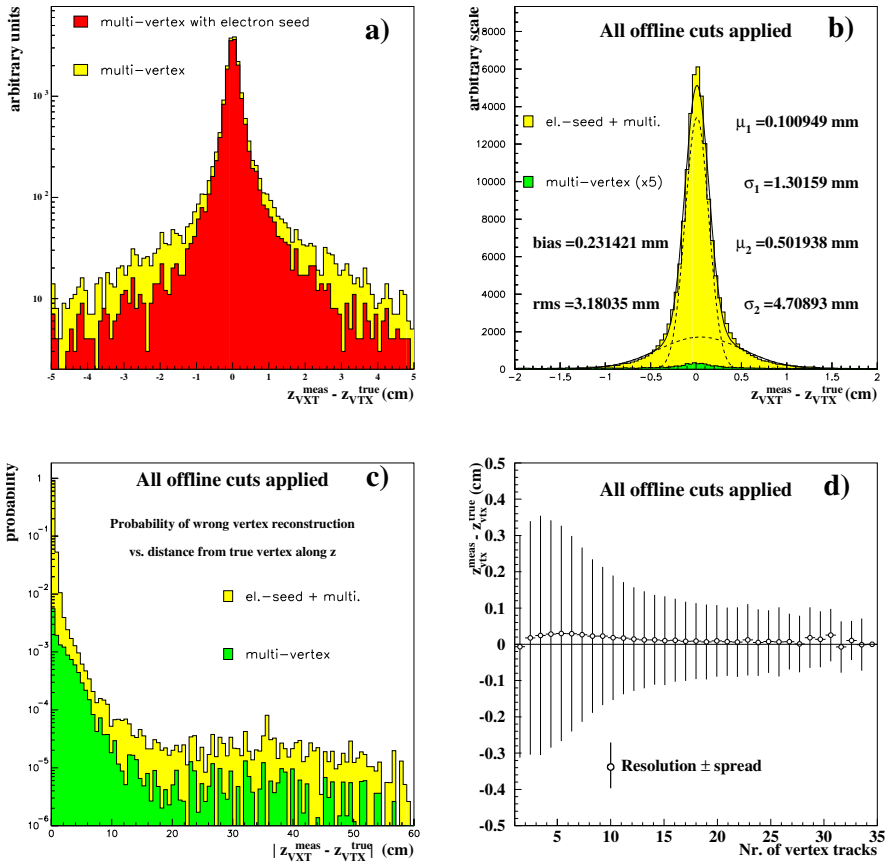

Figure 6.11: a) z vertex resolution of the standard multi-vertex reconstruction algorithm compared to the electron-seed multi-vertex algorithm used in the present analysis. **b)** resolution of the combined electron-seed/standard longitudinal vertex. The histogram corresponding to the standard algorithm has been scaled up to be visible. The off-line analysis cuts have been applied. **c)** probability for the reconstructed vertex, $z_{VTX}^{meas.}$, to be measured at a distance $|z_{VTX}^{measured} - z_{VTX}^{true}|$ from the generated value, z_{VTX}^{true}. **d)** profile plot showing the resolution of the electron-seed method versus the number of tracks emerging from the primary vertex.

6.4.3 Electron Identification and Corrections

NC events were selected by identifying events where an isolated high energy electromagnetic cluster was found in the calorimeter and attributed to the scattered electron. The

routines used to determine wether a scattered electron is present in the event are called electron finders. Once the presence of an electron candidate in an event has been determined, its position and energy must be measured. Several detector effects can bias the measurement and require appropriate corrections. This section gives an overview of the electron finding algorithm and of the corrections used.

Electron Finder

The electron finder employed in this analysis, SINISTRA95 [106], is based on a neural network. This approach is well suited to differentiate between complex classes of patterns, described by many variables. The basic idea is to identify the scattered electron based on the shape of the shower it induces in the CAL.

The scattered electron generally leaves energy deposits in several calorimeter cells, due to showering in the inactive material between the vertex and the calorimeter. The process of finding the scattered electron starts by grouping calorimeter cells into tower islands (defined in section 6.4.1). An electron scattered into the RCAL, for example, typically leaves energy deposits in three to four EMC cells. Each island is then analysed by SINISTRA95, which determines a probability, P_{Si}, that it corresponds to the scattered electron.

Cells in the identified tower islands, are submitted to a pre-processing stage to obtain a fixed number of input quantities that are later used as the input to the neural network. The geometric centre of each cell is first corrected along one direction according to the energies measured by its two PMTs. The energy measured in the cell is considered to be deposited in the corrected cell centre. In a second step, the shower dimensions are transformed into a coordinate space in which the calorimeter is projective with respect to the vertex position. This minimises the distorting effect of the varying detector geometry, which is especially severe in the super-crack regions between RCAL and BCAL or FCAL and BCAL. In this step, the centre of gravity of the shower is calculated, and the line joining this point to the event vertex is taken as the shower axis.

The shower is, at this stage, represented by a set of points (cells) with energy deposits assigned to them. The number of points varies from event to event. A final pre-processing step is performed, where the distribution of points is represented by (a few terms of) an infinite series of orthogonal polynomials. The angular distribution of energy deposits around the shower axis is described by a set of moments of the Zernike polynomials [107, 108], while the longitudinal energy distribution is described by moments of the Legendre polynomials. A total of 16 moments were used to describe the shape of the showers and used as input parameters for the neural network [106]. The total energy of the island being analysed was used as an additional input.

SINISTRA95 was optimized for the detection of leptons at low and medium Q^2 and trained on a sample of low and medium-Q^2 simulated events. Some contamination of the selected sample should be expected, especially at high Q^2 or at high y, where clusters originated by photons produced in the decay of neutral pions can simulate the showering of an electron. Events were required to have at least one calorimeter cluster with a SINISTRA95 probability $P_{Si} \geq 0.9$.

Monte Carlo studies [109, 110] indicate that the selection efficiency and the purity of the event samples selected by SINISTRA95 depends on the region being investigated. Typically, the efficiency is around 95% for electrons scattered into the RCAL and rear BCAL, decreasing steadily towards the forward region, from the middle of BCAL to the FCAL. The efficiency decreases slightly in the region of the super-cracks between calorimeter sections.

The position of the scattered electron is determined with a similar algorithm [111] to the one described in section 6.4.1. If one of the cells attributed to the electron is bad (one dead PMT), the imbalance information from neighbouring cells, in z for the BCAL and y for the F/RCAL, is used to obtain an estimate of the correct position in the bad cell [112]. The energy of the scattered electron is measured in the calorimeter, which has a better resolution than the CTD for energies above 10 GeV [113].

Electron angle measurement & CTD acceptance

The polar angle of the scattered electron is one of the variables used in the determination of the kinematic variables. An accurate determination of the angle is therefore essential. The best resolution is achieved by determining the electron angle from the reconstructed electron track in the CTD. This is the method employed if a well measured track is matched to the electron candidate. Otherwise, a linear interpolation between the primary vertex and the identified electron is used. As the magnetic field has only a small transverse component, near the ends of the solenoid, this is a good approximation as long as the electron cluster position is accurately identified.

To determine which method should be used and to further reject background and incorrectly identified electrons, a track matching the electron cluster is required if the scattered electron is within the CTD acceptance, i.e. if the position of the electron cluster in the calorimeter is such that the electron must leave a track in the CTD. The definition of the CTD acceptance region depends on the position of the primary vertex and on the polar angle of the track. The electron track is considered to be within the CTD acceptance region if the CTD exit radius of the track is greater than 45 cm. This implies that an electron within the CTD acceptance region reaches at least superlayer 4, leaving hits in two stereo superlayers which ensures a good polar angle measurement. This region

corresponds approximately to the angular region $23^\circ < \theta < 156^\circ$ for vertices close to the nominal interaction point.

Electron specific corrections

Although the scattered electron energy does not enter directly in the determination of the kinematic variables it is used in the event selection through a cut on the value of the energy. The measurement of the electron energy in the calorimeter, on the other hand, is biased by several detector effects. The scattered electron crosses a significant amount of inactive material before reaching the calorimeter where some of its energy is lost. The calorimeter itself, not being homogeneous, will yield slightly different measurements depending upon the position of incidence. Dead material between cells and showering in wavelength shifter material give rise to these non-uniformities in response. Finally, unavoidable shortcomings in the Monte Carlo simulation of the detector result in small differences between the results of the simulation and the measured energies.

Electromagnetic energy scale correction: Before any other corrections are made, global energy scale corrections are performed by the RCALCORR routine. This routine is called in EAZE (see section 3.4) before the electron or the hadronic final state reconstruction. In the BCAL, the measured energy is scaled up by 5%. In the RCAL inner cells, correction factors for individual cells were determined from 1996/97 data [46]. These factors vary between -5% and 13%. The correction factors were estimated by comparing the ratio of measured electron energy in data and MC to either the energy predicted by the double angle method, $E_{e,DA}$ (see equation 5.14), or to the expected energy in kinematic peak events. The scattered electrons in the events used to estimate the correction factors were required to be detected more than 2 cm away from the cell boundaries, so that corrections for the calorimeter non-uniformity can be performed independently (section 6.4.3). RCAL cells outside a radius of 70 cm were given an overall correction factor of 2.2% as not enough electrons were detected in this region to enable cell-by-cell corrections. This correction improves the global agreement between real and simulated data.

The small number of electrons detected in the FCAL forbids a detailed cell-by-cell study of the energy scale. The determination of an overall correction was still possible by comparing the raw energy of the measured electrons to the energy calculated with the double angle method, $E_{e,DA}$. As a result of this study [74], the energy of electrons detected in the FCAL was scaled up by 8.2% in data to match the distribution found in Monte Carlo (see figure 6.12).

After all corrections, the uncertainty in the electromagnetic energy scale was found to be 2% at an energy of 8 GeV, decreasing linearly to 1% at energies of 15 GeV and

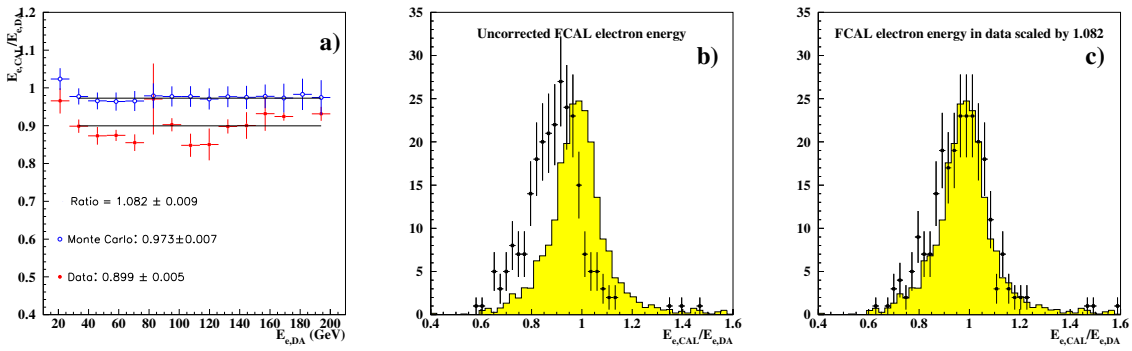

Figure 6.12: Electron energy correction in the FCAL. **a)** shows profile histograms of the ratio between the measured and predicted electron energy in the FCAL. In **b)**, this ratio is compared in data and MC for the uncorrected electron energy. **c)** shows the effect of scaling the measured electron energy by 8.2%.

above [33]. This uncertainty is taken as one of the sources of systematic errors to be considered in the extraction of cross sections (see section 8.1.4).

Dead PMT correction: The energy measured in bad calorimeter cells is by default assumed to be twice the value measured in the working PMT. In this case the imbalance information is set to zero. A further improvement on this algorithm was used in the present analysis for electrons detected in the FCAL or BCAL. If the most energetic cell in the electron cluster is bad, then the imbalance, and thus the cell energy, can be determined by using either information from neighbouring cells or from the reconstructed electron track, when one is available [114, 115]. In the latter case, the correlation between the distance from the endpoint of the extrapolated electron track and the cell centre and the energy imbalance is used.

Dead material correction: The correction for energy loss in inactive material was done differently for electrons detected in the RCAL than for electrons found in the FCAL or BCAL. In the RCAL, information from the presampler was used to estimate, for each event, a correction to the measured energy. The amount of dead material in front of the BCAL is more uniform and better known than is the case for the RCAL (see figure 3.9). The energy correction was, in this case, based on a parametrisation obtained in test beam studies of the calorimeter [59], used together with the map of dead material obtained from a simulation of the detector. The same correction was used for electrons scattered into the FCAL, since the number of high-Q^2 events, with the electron detected in the FCAL was too low to allow the calibration of the forward presampler.

The electron crosses inactive material before reaching the face of the RCAL and so may give rise to a shower. The energy measured in the presampler is a reflection of the

number of particles in this shower and fluctuates from event to event. The presampler energy, measured in number of minimum ionising particles (MIPs) crossing this detector, is added to the electron energy measured in the RCAL according to:

$$E_e^{corr} = E_e^{cal} + E_{pres} \times C_{pres}(r_e^{cal}), \tag{6.7}$$

where E_e^{cal} is the electron energy measured in the RCAL (corrected only for the calorimeter energy scale), E_{pres} is the number of MIPs measured in the presampler, and $C_{pres}(r_e^{cal})$ is a factor determined as a function of radius of the electron cluster on the face of RCAL, r_e^{cal}. The correction factor C_{pres} was determined in a previous analysis [46] using kinematic peak events, where the energy of the scattered electron is close to the beam energy and by comparing the measured energy to the energy predicted by the double angle method. No significant energy dependence was found for the necessary correction (see figure 6.13 **c)**). A small dependence of the correction factor $C_{pres}(r_e^{cal})$ on the electron radius, on the other hand, originates from the electron energy loss going through different amounts of dead material. Different correction factors were determined for real and simulated data. This may be attributed to an incomplete simulation of the amount of dead material in this region of the detector. Figure 6.13 (**a)** and **b)**) shows the energy loss versus the presampler energy in two bins of electron energy measured in data from the 1999/00 period with $r_e^{cal} > 80$ cm. Figure 6.13 (**c)** shows the electron energy loss per MIP measured in the presampler.

Calorimeter non-uniformity corrections: A different correction is necessary to account for electron energy loss or overestimation caused by the non-uniformities of the calorimeter. The energy of electrons entering the calorimeter very near to the cell boundaries is underestimated due to losses in the inactive material between cells. If part of the electron shower develops in the wavelength shifter (WLS) plates which run between calorimeter modules (see figure 3.8) its energy is, on the contrary, overestimated due to the emission of Cherenkov light in the WLS.

In the RCAL, the measured electron energy and the double angle energy, $E_{e,DA}$, were used to determine the appropriate correction [116]. Energy loss is more pronounced in cells close to the RCAL beamhole or cells with boundaries that are nearly projective with respect to the beampipe. In the latter case, electrons enter the calorimeter almost parallel to the cell gaps, loosing a greater fraction of their energy in inactive material.

Both the energy loss in gaps between BCAL cells and energy overestimation in the space between BCAL modules were mentioned in 6.4.1. Appropriate corrections were determined for this calorimeter through the use of the measured and double-angle energy while determining the electron position from the extrapolated electron tracks [117]. The

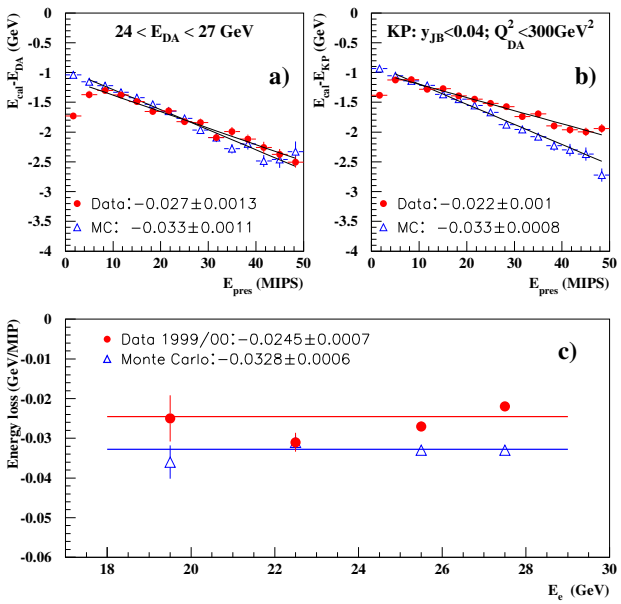

Figure 6.13: Dead material correction for the energy of electrons scattered into the RCAL. Plot **a)** shows the electron energy loss in dead material obtained from the double angle method prediction, E_{DA}, for the bin 24 GeV $<$ E_{DA} $<$ 27 GeV, versus the presampler signal. Plot **b)** shows the same quantities obtained using kinematic peak events. **c)** shows the ratio of energy loss to the presampler signal versus electron energy. This figure corresponds to events from the 1999/00 data sample, with $r_e^{cal} > 80$ cm.

ratio of uncorrected electron energy to the double-angle predicted energy, and the effect of the correction (together with the dead material correction), can be seen in figure 6.14.

Figure 6.15 shows the effect of the dead material and non-uniform corrections on the overall agreement of the electron energy between data and Monte Carlo.

Smearing of the Monte Carlo electron energy: After the corrections described above were applied, a difference remained between the energy resolution in real data and Monte Carlo. This can be attributed to an incomplete simulation of the detector material and possibly to low energy processes not modelled in the MC. The electron energy obtained in simulated events must therefore be smeared. This is done by multiplying the energy in MC events by a random Gaussian-distributed factor according to:

$$E_{e,MC}^{corr} = (1 + 0.03\,G) \times E_{e,MC}, \tag{6.8}$$

where G is a random variable, generated according to a Gaussian distribution of zero mean and unit variance. The factor of 0.03 was determined to give good agreement between data and MC.

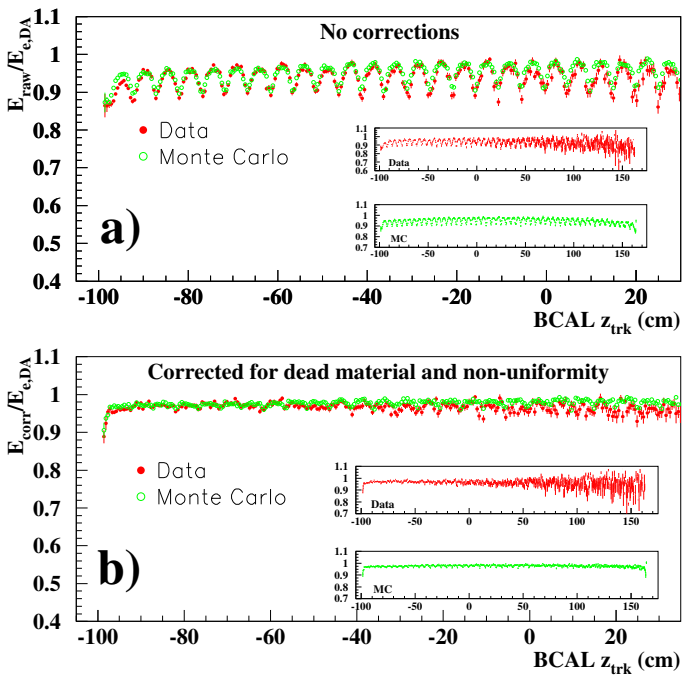

Figure 6.14: Non-uniformity and dead material corrections to the electron energy in the BCAL. The figures show the ratio of uncorrected and corrected measured electron energy to predicted energy as a function of z on the face of the BCAL given by the extrapolated electron tracks.

Figure 6.15 shows the combined effect of the corrections to the measured electron energy. The top plot shows the un-corrected, un-smeared energy, measured in the calorimeter for real and simulated data. The corrected energy distribution doesn't agree with the MC distribution (middle plot). The effect of the smearing of the electron energy in the MC data is clearly visible in the bottom plot.

6.4.4 Reconstruction of the Hadronic Final State

Once an electron candidate is identified, the calorimeter cells corresponding to the electron cluster can be subtracted from the event and the energy deposits in remaining cells attributed to the hadronic final state. In the analysis of inclusive neutral current DIS events, one is interested in measuring the polar angle and energy that characterises the entire hadronic system. These are approximately the same as the scattering angle and energy of the struck quark. These quantities can be obtained using only the energy deposits in calorimeter cells after excluding the electron cluster. However, several effects can systematically affect this measurement and methods to obtain a more accurate estimate of these quantities have been developed. These are performed by the CORANDCUT package [118].

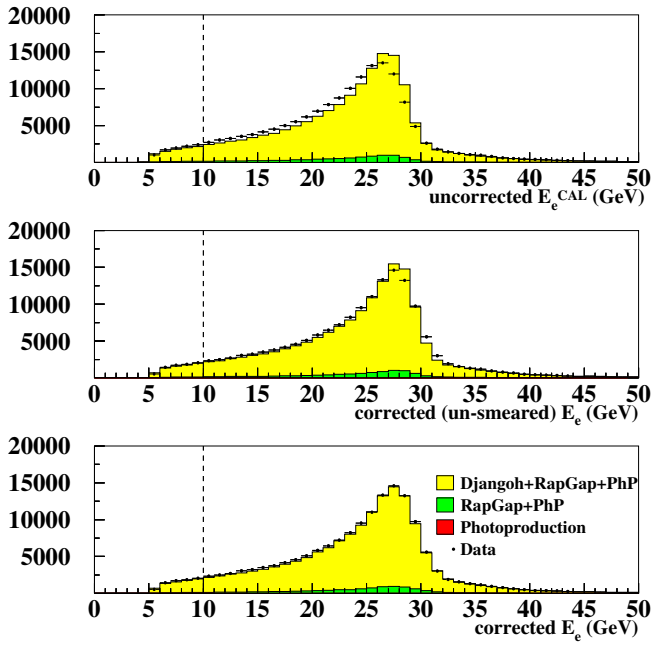

Figure 6.15: Effect of electron energy corrections. The figure shows the un-corrected, un-smeared energy, measured in the calorimeter for real and simulated data. The smeared MC electron energy distribution shows good agreement with the corrected energy distribution found in real data. The vertical dashed lines indicate the off-line cut of $E_e > 10$ GeV.

Energy loss in dead material: The energy loss of hadronic final state particles showering in inactive material depends on the amount of material between the interaction point and the calorimeter. As the amount of material crossed by final state particles depends on the origin and angle of their trajectories, and because of the spatial spread of particles in the final state, the measured hadronic angle can be severely biased. To compensate for this bias, the energy of reconstructed cone islands is corrected according to the dead material map obtained from Monte Carlo [119]. This correction is based on a detailed parametrisation of the energy loss in simulated events, studying separately the cases of electromagnetic and hadronic islands (classified according to the fraction of the cluster's energy found in the EMC and HAC sections of the calorimeter). It was found that the same correction could be applied to both types of islands for island energies above 7 GeV where radiative losses dominate. Below 7 GeV, the energy in hadronic islands was found to be overestimated. This effect is not related to the dead material, but rather to ionisation losses of low energy hadrons in the calorimeter as was observed in beam tests [120] and to the different zero suppression cuts in the EMC and HAC sections [118]. A correction for this effect was applied that was based on a parametrisation of the energy loss as a function of both the island total energy and on the fraction of that energy measured in EMC cells.

Backsplash: Another important effect is the redirection of energy by particle scattering in either the inactive material or the calorimeter itself (*albedo* effect, or *backsplash*). This effect is most important at low y, where most of the hadronic energy is deposited in the forward direction. In this case, a small energy deposit in the rear of the calorimeter, caused by the backscattering of a particle from the forward direction, can create a significant bias in the measured hadronic $E - p_Z$, and therefore also in the hadronic angle γ_h (see equation 5.4).

To limit the effect of the redirection of particles in the calorimeter or inactive material, low energy islands are rejected according to the polar angle of the islands with respect to the vertex [118]. The polar angle of the current jet, γ_h, is determined iteratively, rejecting at each step low energy islands ($E_{isl} < 3$ GeV) with angles above a threshold γ_{max}. The threshold value, γ_{max}, is a function of γ_h which was determined from a set of neutral current Monte Carlo events by requiring that not more than 1% of islands not originating from backscattering are removed. The procedure typically converges after 2 or 3 iterations. Figure 6.16 shows γ_{max} as a function of the hadronic angle.

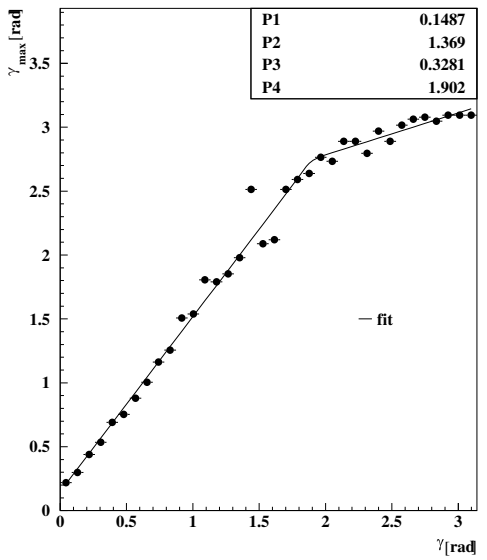

Figure 6.16: The γ_{max}function used in the CORANDCUT package to eliminate islands created by the backsplash effect [118].

Super-crack corrections: Particles crossing the "super-cracks" between the BCAL and the FCAL or the RCAL lose energy in various support structures and other inactive material. A further correction to the island energy was applied in these cases. The correction was determined separately for real data and Monte Carlo using events with a

single hadronic jet to scan the gaps between calorimeter sections. To ensure the measured jet energy could give a good estimate of losses in the super-cracks, the hadronic jets in the selected sample were required to be narrow. To this end, the maximum difference in polar angle between the measured hadronic angle and the islands included in the jet was required to be less than 0.25 radians. The correction was derived by comparing the measured transverse momentum of the hadronic system, p_T^h, to the prediction from the double angle method, $p_{T,DA}$ (see formula 5.15). The ratio $p_T^h/p_{T,DA}$ was plotted against the hadronic angle, γ^h, and inverted gaussians were fitted to the super-crack regions. The result of the fits was used to parametrise the losses in the super-crack regions according to the longitudinal vertex position.

The calorimeter hadronic energy scale

After noise suppression, the RCALCORR energy scale correction, and the dead material corrections, the distributions of the hadronic final state energy measured by the calorimeter is still different in data and MC. The hadronic energy scale and its uncertainty was investigated using NC DIS events [121, 122]. A sample of one-jet events was used to determine the energy scale in the FCAL and BCAL. This sample was selected using the cone-jet algorithm since the objective of the selection was to consider events in which most hadronic energy had been deposited in a limited region of the calorimeter. The FCAL and BCAL were studied separately by requiring that most of the jet energy was detected in one or the other calorimeter sections. The ratios $p_T^h/p_{T,DA}$ and p_T^h/p_T^e, of the measured hadronic transverse momentum to both the double angle and the measured electron transverse momentum, respectively, were calculated as a function of the fraction of energy measured in EMC cells. Comparing these ratios in data and MC, correction factors were obtained for the response of the HAC and EMC sections separately, by extrapolating to EMC fractions of 0 and 1, in both the FCAL and BCAL. The resulting correction factors are shown in table 6.2. After applying the hadronic energy scale corrections, the remaining uncertainty for the BCAL and the FCAL is 1%, as can be seen in figure 6.17.

The determination the hadronic energy scale in the RCAL could not rely on the same algorithm. The average hadronic energy measured in the RCAL is very low. This makes the transverse momentum of the proton remnant and of hadrons produced by the colour flow between current jet and remnant of comparable size to that of the current jet itself. Thus, diffractive events, which have no net colour flow between current jet and proton remnant, were used instead. These were selected by requiring a large rapidity gap between current jet and proton remnant so that all of the hadronic energy was detected in the RCAL with no appreciable energy deposits elsewhere in the calorimeter. This sample was used, as explained above, to calculate the response of the EMC and HAC sections of RCAL and

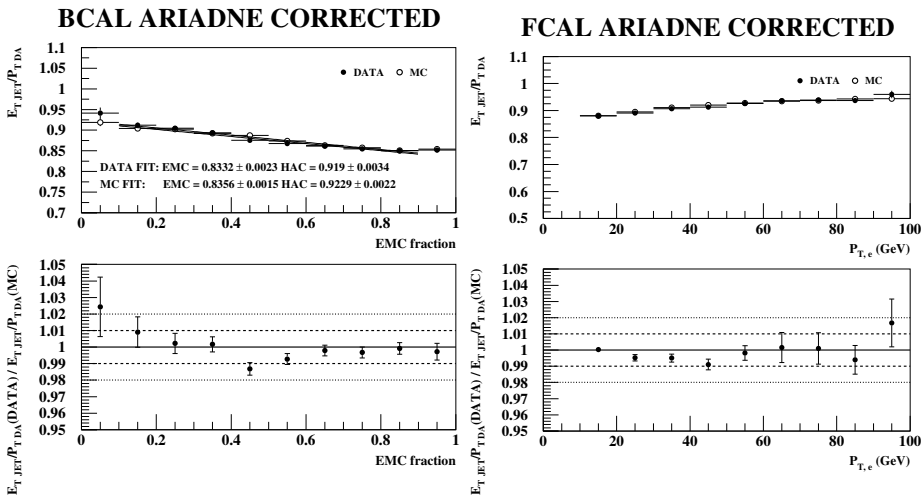

Figure 6.17: Calorimeter hadronic energy scale determination for the BCAL and FCAL. The top plots show the ratios of the transverse momentum of the jet to the prediction from the double angle method, for both real and simulated data. These values are shown against the fraction of the island's energy measured in the EMC and HAC sections of each calorimeter. This allowed the correction factors for these sections to be obtained by extrapolation. The bottom plots show the ratio of the values obtained for real data and MC events. Reproduced from reference [121].

	RCAL	BCAL	FCAL
EMC	1.000 ± 0.02	1.003 ± 0.01	1.024 ± 0.01
HAC	1.000 ± 0.02	1.044 ± 0.01	0.941 ± 0.01

Table 6.2: Correction factors for the hadronic energy scale of the different calorimeter sections.

the corresponding correction factors. These are listed in table 6.2. The small number of events that passed the selection cuts allowed the determination of the hadronic scale of the RCAL to an uncertainty of 2%.

6.5 Off-line event selection

After all corrections have been applied to the measured quantities, events where an electron has been found by SINISTRA95 are selected if they satisfy several criteria. The description and motivation for each of the cuts applied is given in this section. Several figures illustrate the applied cuts. These were obtained by selecting events with the entire set of analysis cuts except for the particular cut that the figure refers to. In this way, the limits imposed by each cut can be clearly illustrated. The distributions found in real data are, in these figures, compared to the MC prediction including diffractive DIS and photoproduction background. The different contributions are shown by histograms in different colours/shades. Figure 6.18 indicates the colour code.

The following cuts are applied to the identified electron candidate:

- **Probability cut:** $P_{Si} > 0.9$. The probability value returned by SINISTRA95 takes values between 0, for hadron-like energy islands, to 1 for electron-like islands. To ensure a good purity of the selected sample, the probability is required to be large;

- **Electron energy cut:** $E'_e > 10$ GeV. The photoproduction background is greater at low E'_e, where relatively low-energy neutral pions decay into photons, faking an electron. The purity of the selected sample is enhanced by requiring higher electron energy. This, on the other hand, excludes a fraction of the available phase space at low Q^2 (see figure 6.25);

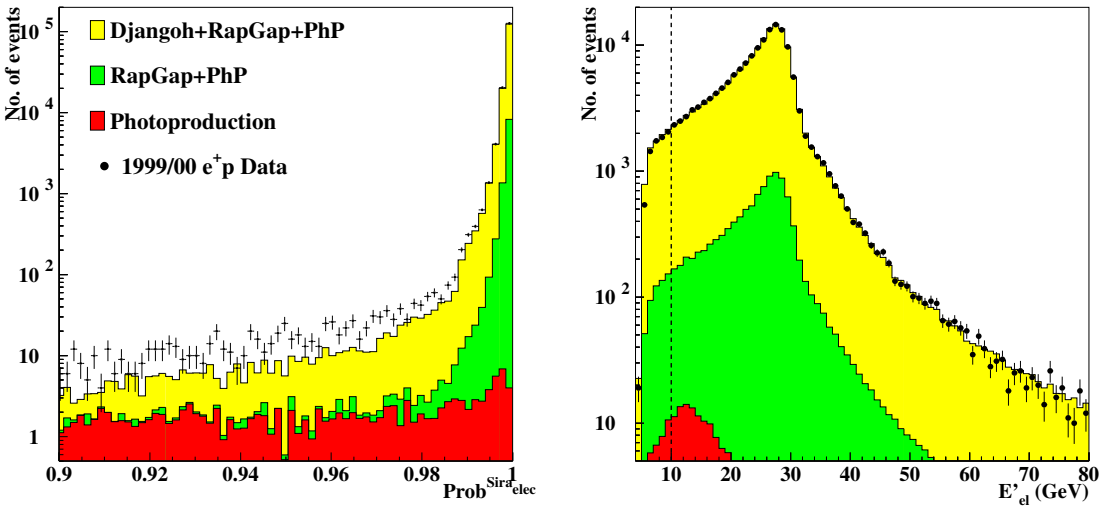

Figure 6.18: Distributions of the electron probability returned by SINISTRA95 (left), and of the electron energy (right). The right histogram was obtained by selecting events with the whole set of analysis cuts except for the cut on the electron energy, which is indicated by the dashed line.

- **Isolation cut:** $E^{cone}_{notelec.} < 5$ GeV. The calorimeter energy measured in a cone around the electron position but not assigned to the electron cluster is required to be less than 5 GeV. This cone is defined by $R_{cone} = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$, where ϕ is the azimuthal angle around the trajectory of the scattered electron and η is the pseudo-rapidity, $\eta = -\ln(\tan\theta/2)$. This cut improves the purity of the selected sample by requiring the detected electrons to be isolated, thus contributing to the rejection of events where an hadronic cluster fakes the scattered electron. The distribution of $E^{cone}_{notelec.}$ is shown in figure 6.19;

- **Track matching:** if the electron is within the CTD acceptance, as defined in section 6.4.3, then the existence of a matching track is required. The distance of closest

approach (DCA) between the calorimeter cluster and the endpoint of the extrapolated CTD track is required to be less than 10 cm. This rejects events where a photon shower is wrongly selected as the electron cluster. Figure 6.19 shows the distributions of the DCA in real and simulated data. Some disagreement can be seen between the real data and the simulation, where the MC simulation seems to result in a better track-cluster matching;

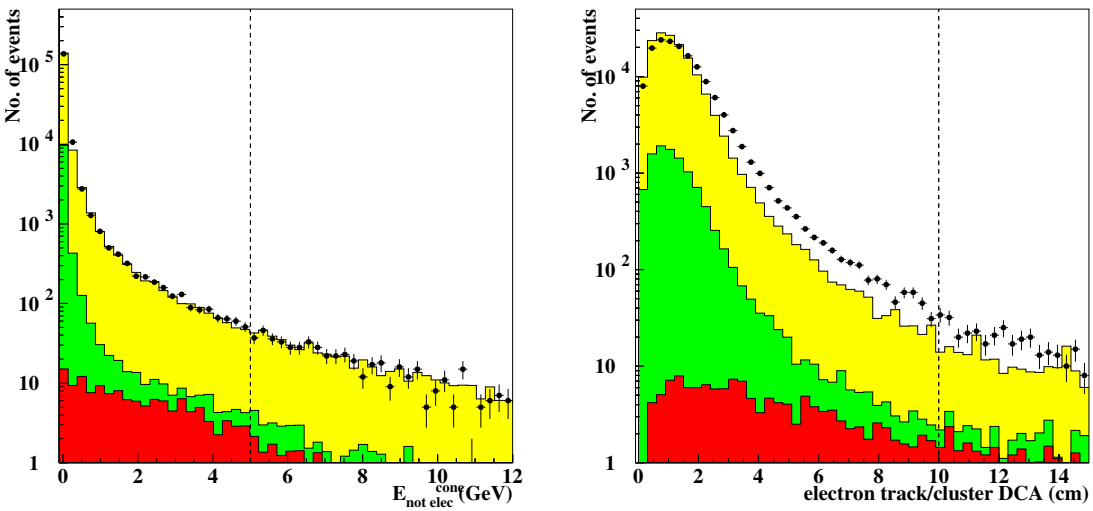

Figure 6.19: Distributions of the energy not assigned to the scattered electron in a cone around the electron position, $E^{cone}_{notelec.}$ (left), and of the distance of closest approach (DCA) between the extrapolated electron track and the electron cluster (right). The dashed lines indicate the respective analysis cuts.

- **Electron track momentum:** $p_e^{trk} > 5$ GeV. If the electron is within the CTD acceptance, the electron momentum measured from its track (see figure 6.20) is required to be high, to enhance the purity of the selected sample. As the electron angle is calculated from the matched track, it is important that the matching is done correctly;

- **Forward transverse momentum cut:** $p_{T,e} > 30$ GeV. For electrons scattered outside the forward acceptance of the CTD no matching track is required. Instead, the previous cuts are replaced by a cut on the electron transverse momentum of the electron candidate. The distribution of the electron transverse momentum for such events is shown in figure 6.20;

- **Super-crack cuts:** events are rejected if the scattered electron cluster or track endpoint is reconstructed in the edges of BCAL due to the uncertainty in the energy measurements and because of the large amounts of dead material in this region. The

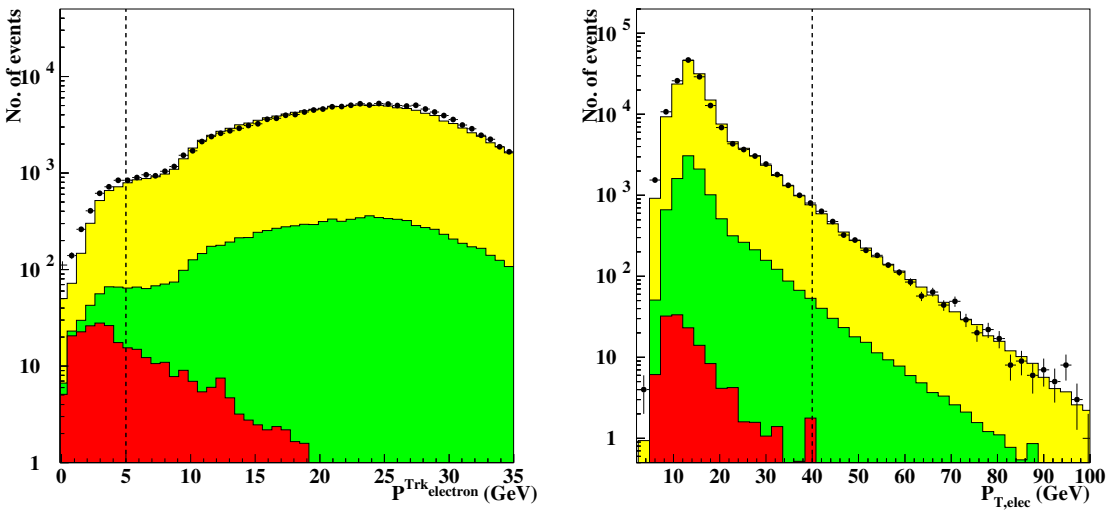

Figure 6.20: Distributions of the momentum of the electron track, p_e^{trk} (left), and of the measured transverse momentum of the electron candidate for electrons outside of the forward CTD acceptance (right). The dashed lines indicate the respective analysis cuts.

excluded areas of BCAL correspond to -98.5 cm $<$ z $<$ -104 cm and 164 cm $<$ z $<$ 174 cm. The effect of this cut may be observed in figure 6.25, as a depletion in the density of events in a region of phase space where the final state electron is scattered into the super-crack between RCAL and BCAL;

- **Cryogenic supply region:** the tubes supplying gaseous and liquid helium [45] to the superconducting solenoid partially occlude the face of the RCAL where, additionally, EMC cells do not exist to allow enough space for the tubes. This region is avoided by rejecting events if the electron candidate cluster or extrapolated track endpoint is found in the RCAL with $|x| < 10$ cm and $y > 80$ cm. Figure 6.21 shows a scatter plot of the electron position reconstructed in the RCAL. The region eliminated in this cut can be seen in the figure, highlighted by a rectangle.

- **Electron radius in RCAL:** $R_e < 175$ cm. The simulation of the RCAL outer region, which is occluded by the BCAL is flawed, leading to a poor simulation of the trigger acceptance. This problem is avoided by excluding electrons detected in the outer RCAL [74]. The effect of this cut on the electron position in the RCAL for accepted events can be seen in figure 6.21.

The cuts listed above aim at selecting events with a well identified scattered electron. A different class of cuts was applied which aim at reducing the background contamination and migrations in kinematic space due to ill-measured event quantities.

- **Vertex cut:** -50 cm $< z_{\mathrm{VTX}} <$ 50 cm. The vertex distribution covers a wide range of z, with a pronounced maximum close to the nominal interaction point at $z = 0$ cm (a more detailed discussion of the vertex distribution can be found in section 7.1). This cut is justified by both the better understanding of the CTD and calorimeter for events occurring close to the centre of the CTD and by the more accurate knowledge of the vertex distribution itself in that region. Figure 6.21 shows the measured longitudinal vertex distribution;

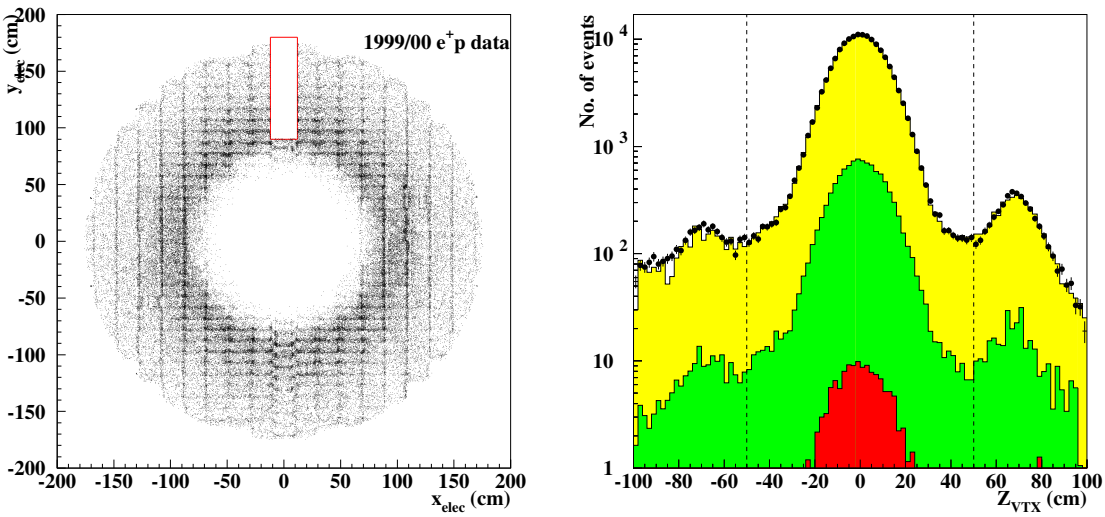

Figure 6.21: Illustration of the cuts on the electron position in the RCAL (left). Only events that passed all analysis cuts contribute to this plot. Right: distribution of the event vertex along z. The dashed lines indicate the off-line cuts applied on the z vertex position.

- **Cuts on δ_{tot}.** A perfectly measured NC DIS event has $\delta_{tot} = (E - p_z)_{tot} = 2\,E_e = 55$ GeV, with E_e the electron beam energy. To suppress the photoproduction background, where the scattered electron escapes detection through the rear beampipe, a lower cut of $\delta_{tot} > 38$ GeV was used. This cut is also helpful in rejecting events with hard initial state radiation. For events where the scattered electron is outside the forward limit of the CTD acceptance region this cut is tightened to $\delta_{tot} > 44$ GeV. An upper cut of $\delta_{tot} < 65$ GeV removes overlaid events where a normal DIS reaction occurs in the same bunch crossing as another process. Figure 6.22 shows the distributions of δ_{tot} observed in real and simulated data;

- **Further photoproduction rejection:** $y_e < 0.95$. In photoproduction events, where no electron can be found in the calorimeter, photons from the decay of a π^0 from the hadronic final state may be mistaken for the electron candidate. This

occurs more frequently in the FCAL where the tracking cuts cannot be used to reject the event. In this case, y_e, given by:

$$y_e = 1 - \frac{E'_e}{E_e}(1 - \cos\theta) \tag{6.9}$$

becomes large due to both the relatively small energy of decay photons and the low angle (see figure 6.22);

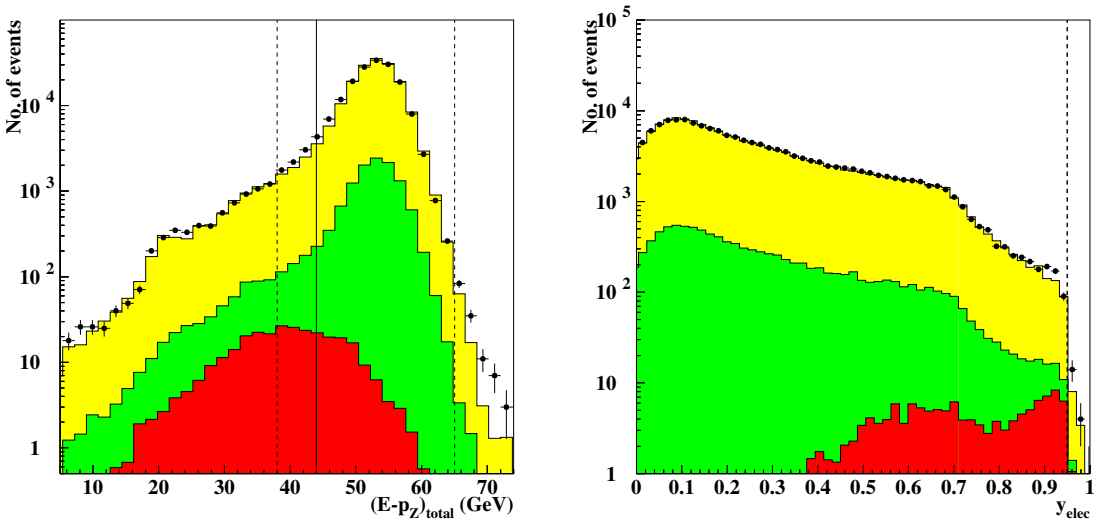

Figure 6.22: Distributions of the measured δ_{tot} (left) and y_{el} obtained in real and simulated data. The dashed lines indicate the cut values. A vertical solid line indicates the value of the cut on δ_{tot} for events where the scattered electron is detected outside the forward CTD acceptance.

- Transverse momentum balance: $p_T/\sqrt{E_T} < 4$ GeV$^{1/2}$ and $p_T/E_T < 0.7$. The total transverse momentum, p_T, and transverse energy, E_T, are given by equations 6.3. The transverse momentum of a perfectly measured NC DIS event should be zero. But, with a real detector, this quantity is measured with a finite resolution which depends approximately on $\sqrt{E_T}$ as can be seen from the spread of the points in figure 6.23. At low E_T, the cut on $p_T/\sqrt{E_T}$ is in an un-physical region, since p_T must always be smaller than E_T. Some events are measure at low E_T with an unusually high p_T, these do not appear in signal Monte Carlo and are eliminated by the cut on p_T/E_T. The solid line in figure 6.23 shows the combined cut. This cut rejects mainly background from cosmic ray events that survive the timing cuts in the trigger and show an unbalanced total transverse momentum by, for example, depositing more energy in one side of the calorimeter. The p_T cuts also reject beam-related background;

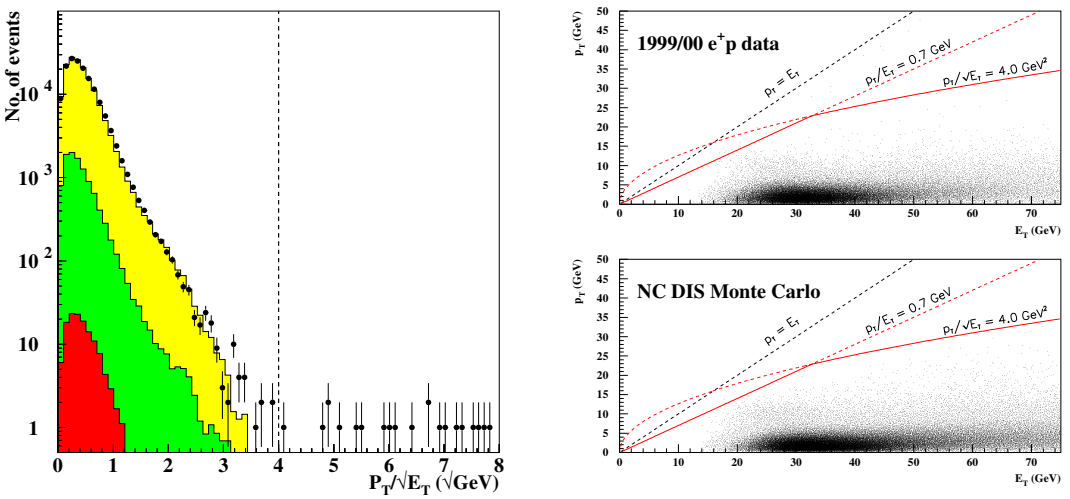

Figure 6.23: Cuts on p_T/E_T and $p_T/\sqrt{E_T}$. The dashed lines show the $p_T = E_T$ kinematic limit and also the individual cuts. The solid line shows the combined cut. The real and simulated data were selected with all the off-line cuts but relaxed p_T/E_T and $p_T/\sqrt{E_T}$ cuts. Some real events shown in the upper plot with high transverse momentum are most likely cosmic rays or beam-related background.

- **Cut on the FCAL projection of γ_h:** $R_{\gamma_h}^{FCAL} < 20$ cm. At very low values of the hadronic angle, γ_h, many events show little hadronic activity in the calorimeter except for energy deposits around the FCAL beamhole. The energy deposits measured in the cells around the beamhole is generally a small fraction of the energy of the hadronic system due to the presence of the beamhole. Studies of the hadronic energy flow in this region (see section 7.2) have shown that the detector and hadronisation simulation in this region is somewhat uncertain. The problem is aggravated by the presence of the FPC and the leakage of energy from showers developing in this detector into the FCAL. As a result of this, the measurement of γ_h becomes severely biased. To avoid the problematic region while maintaining a large acceptance, a cut is performed on the projection of the measured hadronic angle on the face of FCAL, rejecting events where the radius of the projection falls within 20 cm of the centre of the beampipe (see figure 6.24);

- **Monte Carlo validity cut:** $y_{JB}(1 - x_{DA})^2 > 0.004$. The Monte Carlo simulation is not valid at very low y and high x [75]. This cut eliminates the region of very low invariant mass of the hadronic final state, W, which is dominated by nucleon resonances and is not simulated in the Monte Carlo (see figure 6.24);

- **QED Compton rejection.** Elastic QED Compton events are not simulated by the Monte Carlo generator and must be rejected. This is done by searching for

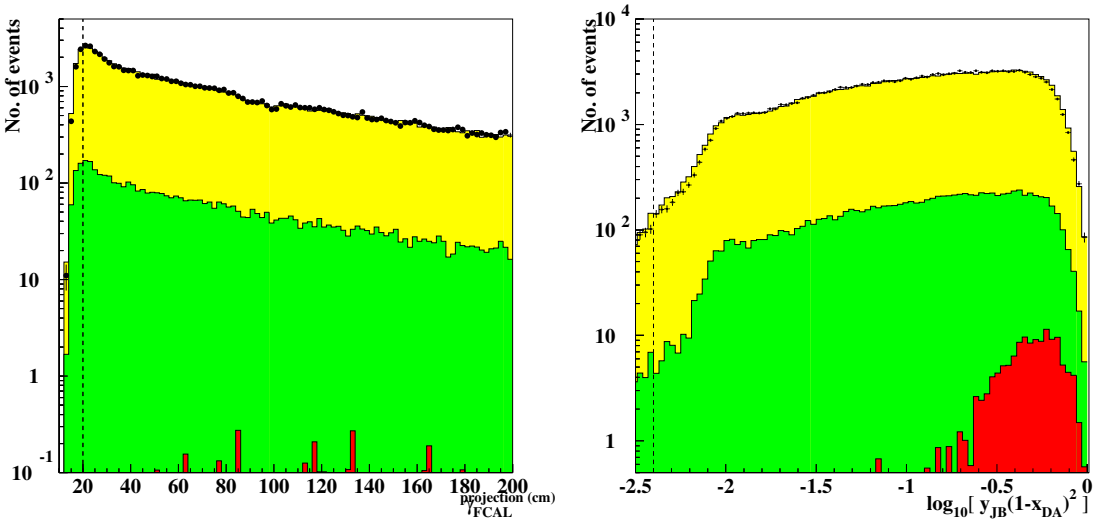

Figure 6.24: Dstributions of the radius of the projection of the hadronic angle on the face of the FCAL, $R_{\gamma_h}^{FCAL}$ (left), and of the quantity $y_{JB}(1-x_{DA})^2$. The cut on $R_{\gamma_h}^{FCAL}$, indicated by a dashed line, corresponds to different values of γ_h, depending on the vertex position. The MC validity cut is indicated by a dashed line at $y_{JB}(1-x_{DA})^2 > 0.004$.

an additional electromagnetic cluster in the calorimeter, with no associated track, and rejecting the event if the summed energies of electron and photon candidate was within 2 GeV of the total energy measured in the calorimeter. A fraction of the QED Compton events observed are inelastic, where the proton dissociates and leaves some energy deposit in the calorimeter making these events harder to reject. In such cases, the energy deposit due to the photon is attributed to the hadronic final state, causing migrations in phase space. On the other hand, it was found that the inelastic process is correctly simulated [74] and it is not necessary to reject them.

The effect of some of the cuts described above is to limit the region of phase space where the analysis is performed. These are shown in figure 6.25.

6.5.1 Final data samples

The sets of data that passed all the trigger and analysis selection cuts consisted of 221240 e^-p events and of 867132 e^+p events, corresponding to integrated luminosities of 15.87 pb^{-1} and 62.53 pb^{-1}, respectively.

Figures 6.26 and 6.27 show histograms of the kinematic variables Q^2, x, and y measured with the double angle method, of the scattered electron and hadronic angles, and of the electron and total transverse momentum. Only events accepted in the final sample have contributed to the histograms. The distributions observed in real data are compared to the

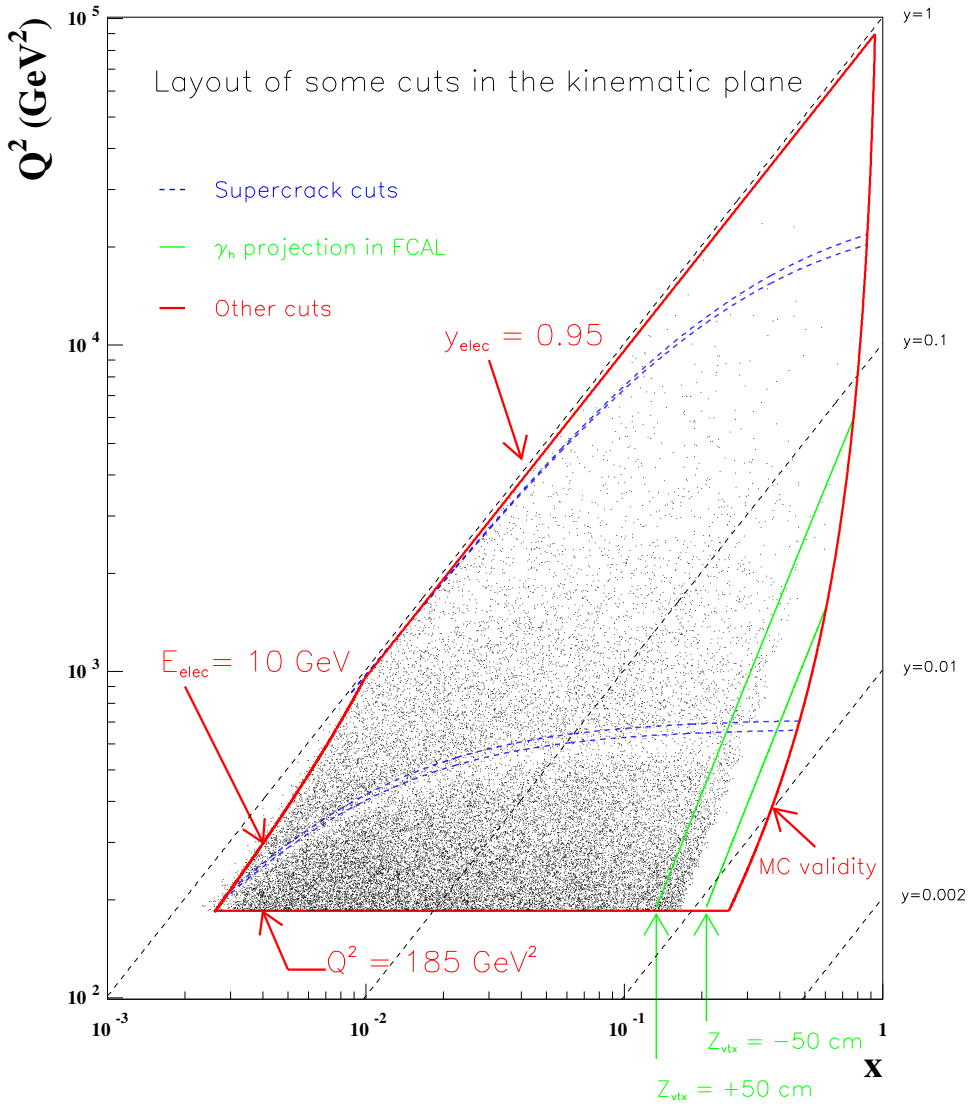

Figure 6.25: Layout of some off-line cuts in the kinematic plane. The position of the cut on the FCAL projection of γ_h at low y depends on the z position of the vertex. Extreme values for this cut are $z = -50$ cm and $z = 50$ cm, and are illustrated in the figure. Only a fraction of the selected events are shown, to avoid making the figure unclear.

MC prediction, including photoproduction, non-diffractive and diffractive DIS MC events. The figures show good agreement between the observed data and the MC simulation.

Figure 6.28 shows the number of accepted events per unit of integrated luminosity as a function of the run number in both the e^-p and the e^+p data samples. If experimental conditions are stable, this number depends simply on the integrated cross section of the selected events. This is thus an indication of the quality of the data taking process for each run. This quantity is, within statistical errors, spread around a constant value close

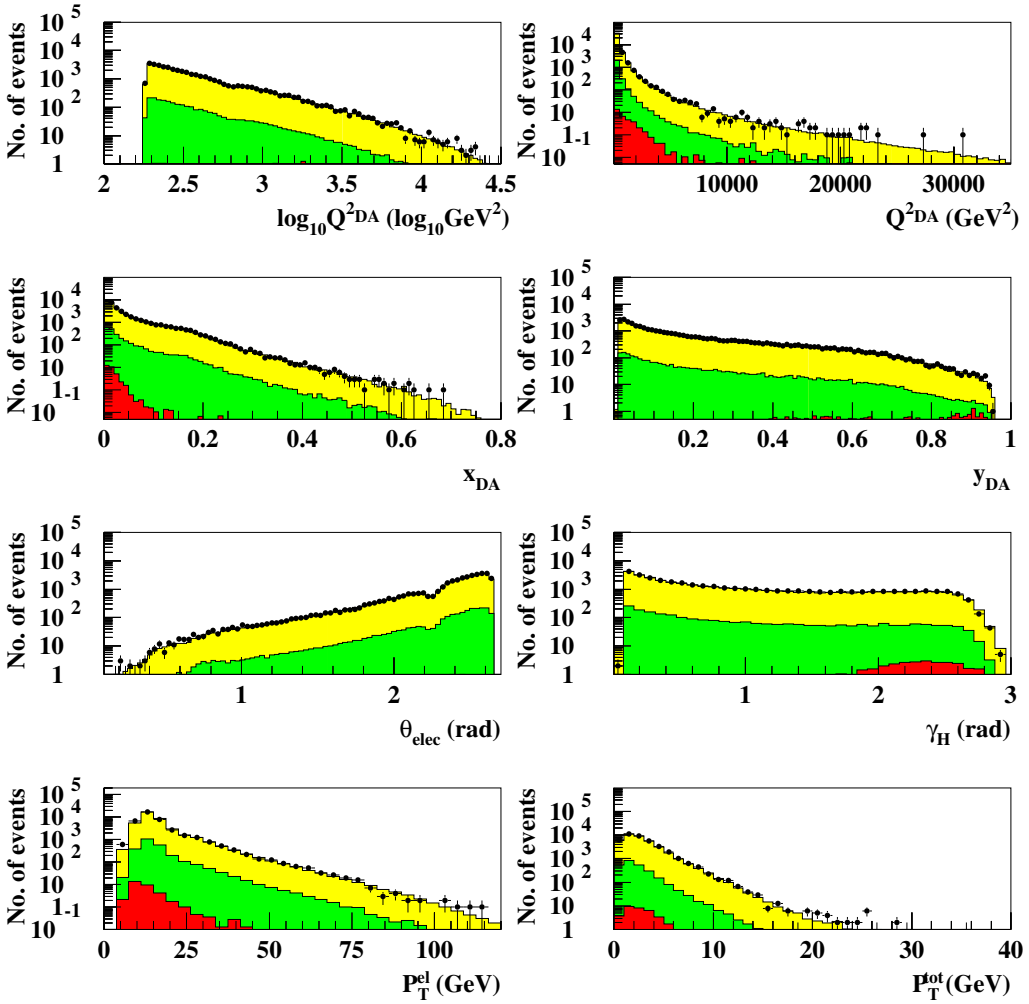

Figure 6.26: Control plots showing a comparison between some variables in Monte Carlo and real data accepted events for the 1998/99 data samples.

to 2.5 events/nb^{-1}.

Beam-gas subtraction

As described in section 3.1.1, the background from beam-gas events may be estimated using recorded events from collisions of unpaired electron or proton bunches. For the 1998/99 period, 20 e-pilot events passed the trigger and analysis selection. 53 events were recorded in 1999/00 in the same category. In both cases, the events were closely examined and were found to be normal NC DIS events from $e^{\pm}p$ scattering with the correct beam energies. The supposedly empty proton bunches, in fact, contain some residual charge

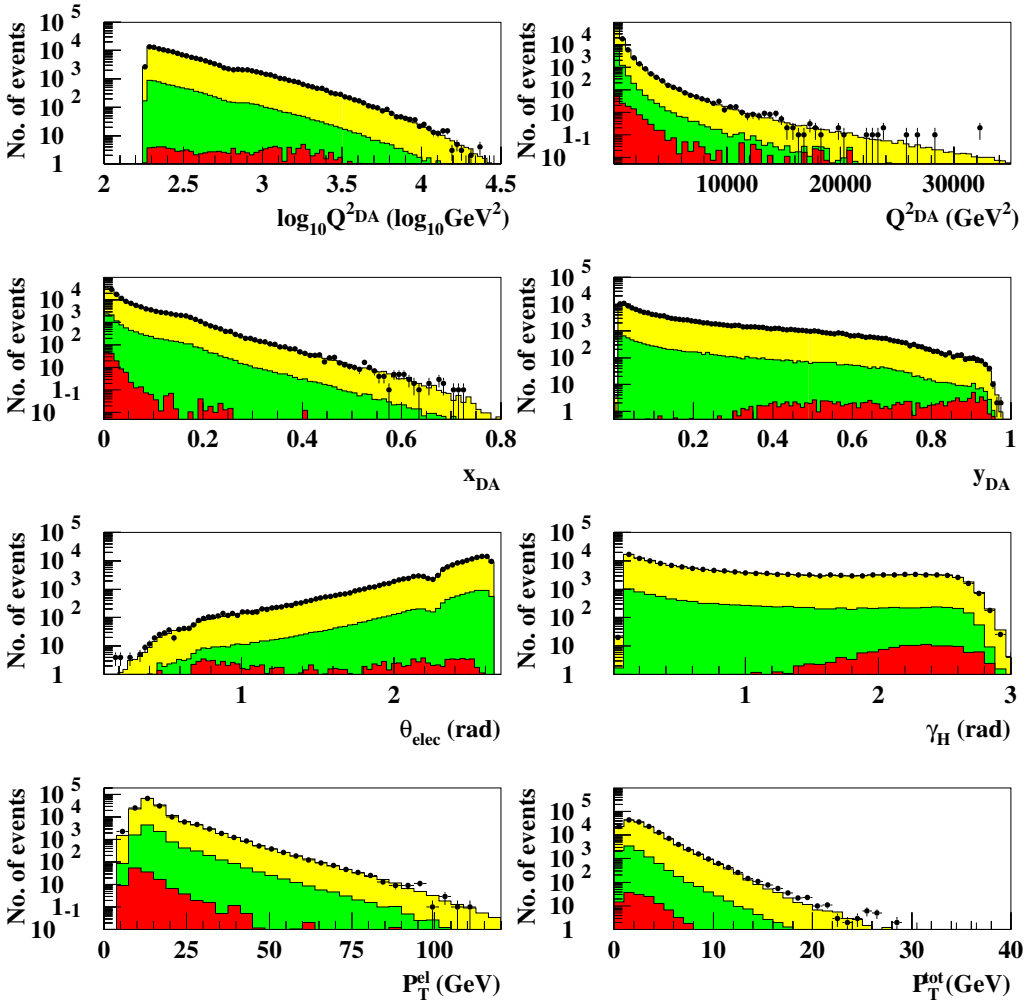

Figure 6.27: Control plots showing a comparison between some variables in Monte Carlo and real data accepted events for the 1999/00 data samples.

which gives rise to these events. Typical ratios of the measured currents in filled and empty bunches are around 10 for electron bunches and 30 for proton bunches. Figure 6.29 shows the measured ratios of currents in colliding and non-colliding bunches for one data collection period. The e-pilot events that passed all analysis cuts were rejected, as they are not counted in the calculation of the integrated luminosity.

In the 1998/99 sample, 47 p-pilot events were accepted by the trigger selection. The corresponding number for the 1999/00 data sample was 81 events. No p-pilot events passed the analysis selection cuts for either data sample. Finally, 12 and 29 events from both electron and proton empty bunches, passed the trigger selection. None of these was

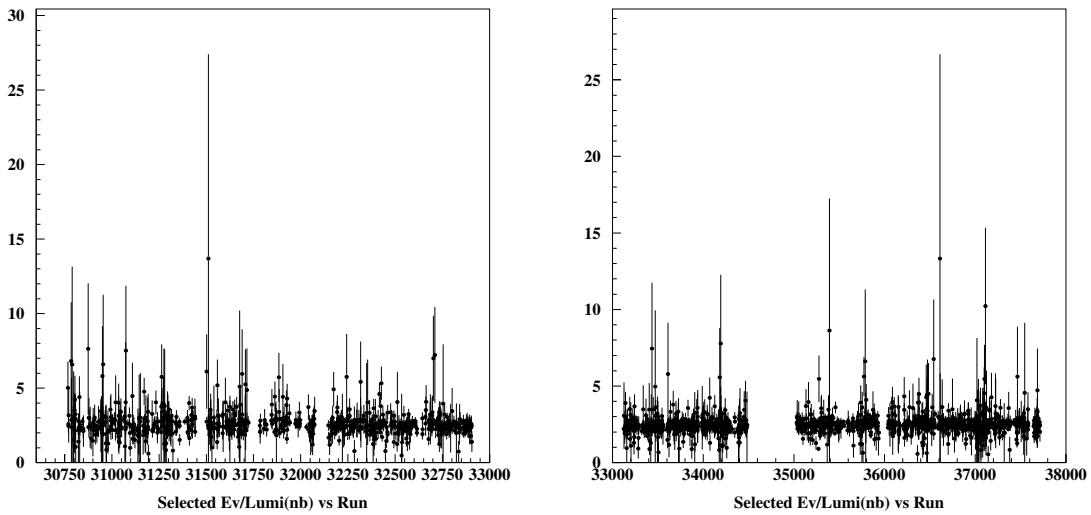

Figure 6.28: Number of accepted events per unit of integrated luminosity as a function of the run number for the the 1998/99 (left) and the 1999/00 (right) data samples.

accepted by the analysis cuts. It was concluded that beam-gas events were efficiently rejected by trigger and analysis cuts combined, having a negigible contribution to the background contamination of the selected sample.

6.5.2 Event excess at high y and Q^2

A small but non-negligible excess of data events over the MC prediction was found in 1999/00 data at $Q^2 \gtrsim 5000$ GeV^2 and $y \gtrsim 0.65$. No clear cause for this deficit was found in spite of a detailed study [123, 124]. Figure 6.30 shows histograms of some variables measured in the region where the excess was found.

Similar results were obtained when using either the electron or the Jacquet-Blondel methods. This rules out a deficiency in the reconstruction of either the electron or the hadronic final state as the only source of the problem. Different electron-finder algorithms were used, as well as both the ARIADNE and the MEPS hadronisation models, without visible changes to the observed excess.

The set of parton density functions (PDFs) that were used in the generation of Monte Carlo data were obtained using older high-Q^2 data from the H1 and ZEUS experiments [25]. This left the possibility that the observed deficit of MC events was due to an underestimation of the PDFs in the high-Q^2 region. To exclude this possibility, the simulated data was reweighted to a recent fit which included the latest F_2 results from ZEUS [24]. No significant changes were observed in the region of the excess.

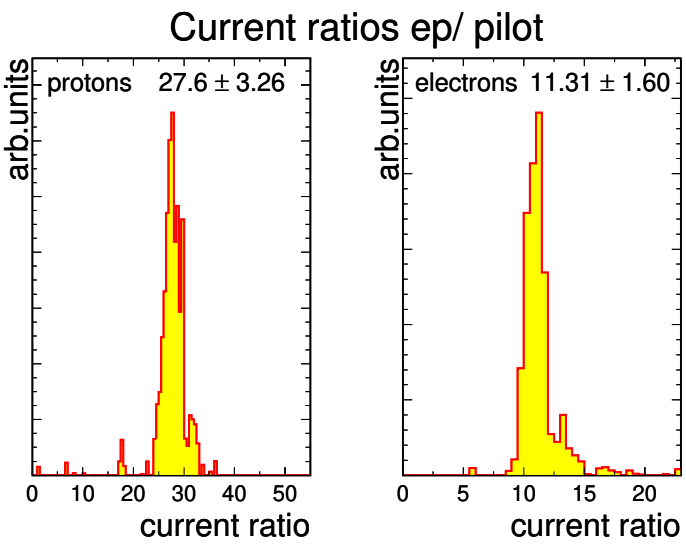

Figure 6.29: Ratios of currents in colliding and pilot bunches measured in the 1999 $e^- p$ running period.

The events in the kinematic region of the observed excess have scattered electrons detected mainly in the FCAL, with a few in the forward BCAL. The detection of the scattered electron in this region of the detector is less well understood than in the rear part of the ZEUS detector where the extensive data sets collected to date have resulted in a good understanding of the experimental acceptance. Nevertheless, the use of different algorithms in the identification of the scattered electron gives some confidence that the events were correctly identified as NC DIS.

Finally, around 50 events in the excess region were scanned with the aid of the ZEUS event visualisation package. These were selected with cuts on y, Q^2 and γ_h to obtain a set of events from the very region of the excess. Most of the events could be clearly identified as NC DIS, with no clear source of background present.

The observed excess may arise due to a combination of detector effects which cause the migration of events in phase-space. Another possibility would be an incorrect modelling of the dead material in the region between BCAL and FCAL leading to lower acceptance for MC events than exists for real events. This excess of events is currently under study.

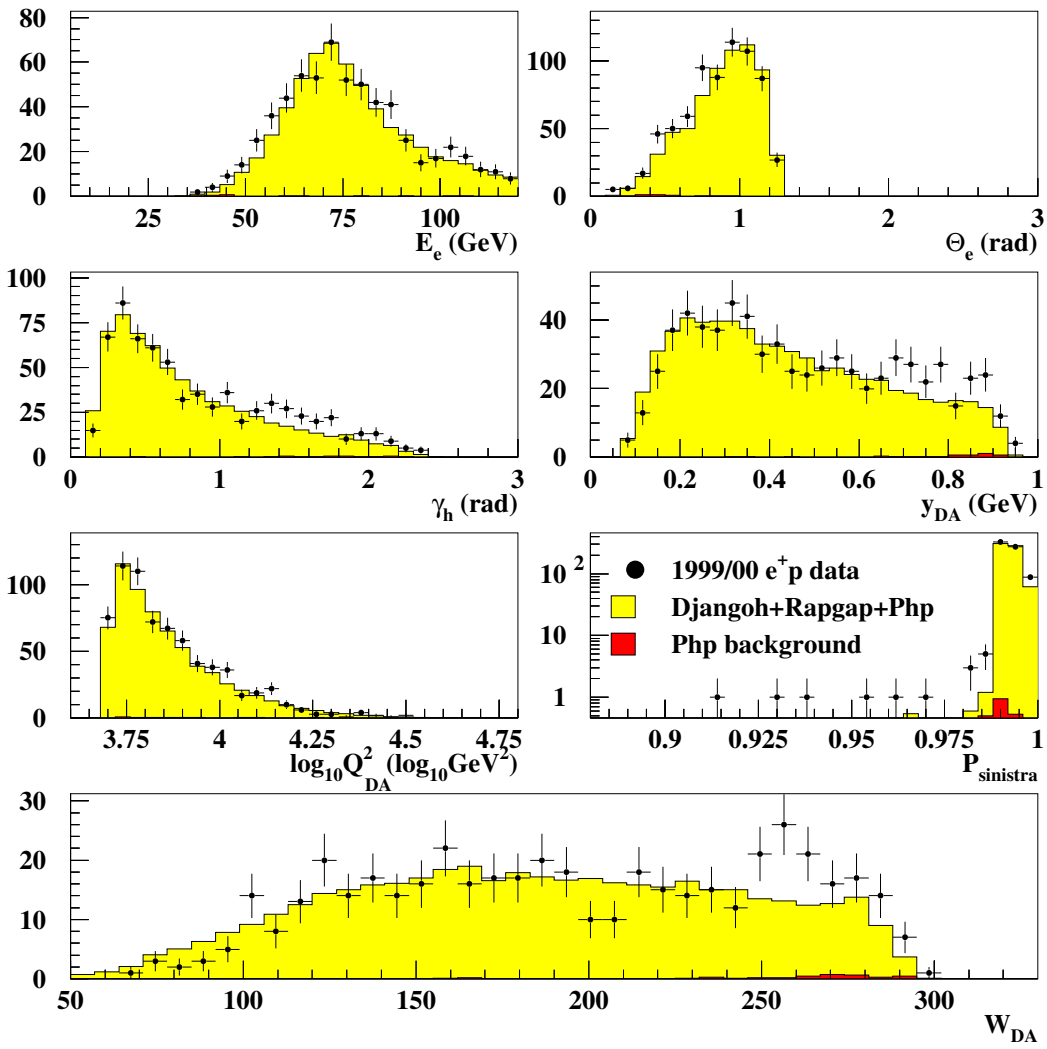

Figure 6.30: Excess of observed events over the MC prediction, at $Q^2 \gtrsim 5000$ GeV^2 and $y \gtrsim 0.65$, in the 1999/00 data sample. The solid line and shaded histograms correspond to MC events. The histograms represent the distributions of the scattered electron energy and polar angle, the hadronic angle, the kinematic variables y_{DA} and Q^2_{DA}, and the electron identification probability. Events the 1999/00 data sample were selected by requiring $Q^2_{DA} > 5000$ GeV^2 and all the normal analysis cuts.

Chapter 7

Systematic studies: vertex distribution and QCD cascade

In this chapter, two studies of systematic effects are described which were performed in the context of the present analysis. These are the determination of the longitudinal vertex distribution for the 1998/99 e^-p data sample [125], and a study of the models of the QCD cascade implemented in the Monte Carlo software [126]. The determination of the vertex distribution for the 1999/00 e^+p data sample is also described.

7.1 Longitudinal vertex distribution

The distribution of the z-coordinate of the vertex, z_{VTX}, is an essential input to the Monte Carlo software. This distribution is used by the MOZART package (see section 3.3) to attribute a position in space to the primary vertices of generated events. As the acceptances of the ZEUS detector and trigger depend on the position of the vertex, it becomes important that the distribution used in MOZART corresponds closely to the true distribution of the vertex in real data.

Although most events originate within ~ 10 cm of the nominal interaction point, which reflects the length of the proton bunches accelerated by HERA, several effects contribute to make the tails of the longitudinal vertex distribution very long. The overall length of the distribution is close to the length of the CTD. A further complication arises from the necessary tuning of the accelerator parameters, which creates a slightly different vertex distribution for each fill. The distribution that must be used in the generation of events is, therefore, characteristic of the total data set that is being analysed.

Due to the dependence of the acceptance on the vertex position, the vertex distribution that is used in the generation of events cannot simply be obtained from the use of physics oriented cuts.

The off-line cuts on $|z_{VTX}| < 50$ cm do not eliminate the impact of the vertex position dependence of the detector and trigger acceptance. Different vertex distribution shapes correspond to a different fraction of events passing the off-line vertex cut in real data and MC data. As the integrated luminosity is used to normalise the MC sample to the data this leads to different numbers of events being accepted in real and simulated data, because the measurement of the luminosity does not take the vertex position into account.

The methods used in the determination of the vertex distribution for the 1998/99 e^-p data are described below. One of these was used to obtain the final minimum bias vertex distribution, and the other as an independent verification of the results. A modified version of the latter was used in a different analysis for the measurement of the vertex distribution of the 1999/00 e^+p data sample [74]. These methods correspond to using different off-line cuts than are described in chapter 6 and they have required the use of different Monte Carlo samples.

The method used for the actual measurement of the vertex distribution will be referred to as the Q&R method, after the authors of reference [127], on which it is based. This method was modified in this work to eliminate some remaining bias. The second method is based on reference [95] and will be referred to as the W method, after its author.

The vertex finding algorithm used in this study was the standard multi-vertex algorithm implemented in VCTRAK (see section 6.4.2). At the time when this study was performed, the electron-seed multi-vertex algorithm was not implemented. This led to some bias in the W method, as is explained below. No noticeable effects were seen in the Q&R method.

The aim of both methods is to select a large a sample of NC DIS events with the minimum possible bias from the trigger and detector. Therefore, the selected sample must obey the following requirements:

- The selection efficiency must not depend on the vertex position;
- The vertex should be accurately measured for all positions of the vertex.

7.1.1 The minimum bias methods

The Q&R method

The main limitation to the detection of the scattered electron in low-Q^2 events comes from the rear beam hole. An electron originating from an event with $z_{VTX} = -50$ cm with a polar angle $\theta_e \sim 175^\circ$ may be lost in the rear beam hole. On the other hand, an electron originating at $z_{VTX} = +50$ cm with the same polar angle would be correctly detected in the calorimeter.

The dependence of the trigger selection efficiency on z_{VTX} can be removed by requiring that all electrons have a polar angle less than a critical value. This value can be determined geometrically so that electrons from the end of the CTD closer to the RCAL are still accepted. However, if such a cut is applied to the whole CTD range, many events with a vertex near the centre of the CTD or closer to the FCAL, and a correctly detected electron, will be rejected needlessly. A higher polar angle cut is sufficient in the centre and FCAL side of the CTD.

The hadronic tracks must also be correctly reconstructed in the CTD, to satisfy the requirement that the vertex position must be correctly determined. As mentioned in section 6.4.2, the resolution of the vertex measurement depends strongly on the number of tracks emerging from the primary vertex. To satisfy this requirement, it is necessary that the hadronic tracks leave hits in several stereo superlayers of the CTD. To achieved this, the CTD was divided into three regions along z and an optimum hadronic track angle attributed to each of these regions. Then, in each event, the track which is closest to this angle is selected and cuts are applied to its polar angle (θ_{had}). Note that this is not a statement on the hadronic angle, γ_h, but only a requirement that some hadronic track will be close to an optimum angle. The chosen CTD regions are:

Region 1 (RCAL side):	-100 cm	$< z_{VTX} <$	-40 cm
Region 2 (central):	-50 cm	$< z_{VTX} <$	50 cm
Region 3 (FCAL side):	40 cm	$< z_{VTX} <$	100 cm

The optimum angles are then: 90° for the central region, 45° for region 1, and 135° for region 3. Straight tracks with these angles would leave hits in all 9 superlayers of the CTD. Figure 7.1 illustrates the topology of such events.

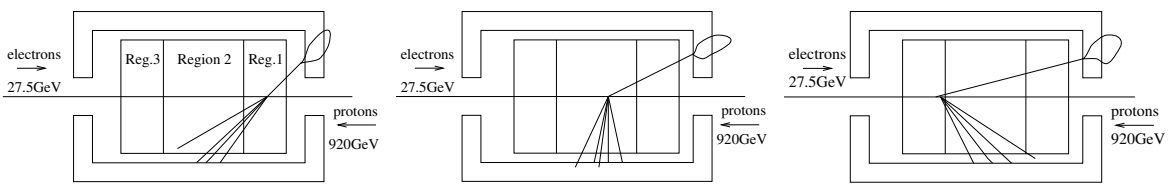

Figure 7.1: Topology of events selected by the Q&R method. The calorimeter and CTD are represented schematically, as are the scattered electron and hadronic final state particles.

For each region a maximum electron angle was determined geometrically in order to maximize the acceptance in each region while making sure that electrons in selected events hit the RCAL sufficiently far from the beamhole. The relative normalization of the distributions found in each region was calculated from Monte Carlo events.

Event Selection The events used in this study were selected by requiring a low-Q^2 NC DIS trigger bit. Two kinds of cuts were applied off-line: cuts on the detected electron and event kinematics and the angle cuts referred to above on the electron and hadronic final state particles.

- $E_e > 15$ GeV and $P(Si) > 0.9$. These cuts ensure that the electron is detected with very good efficiency and purity. As before, E_e is the scattered electron energy measured in the calorimeter and P_{Si} is the electron candidate probability returned by SINISTRA95;

- $\delta_{tot} = (E - P_z)_{total} \in [35\ \text{GeV}, 65\ \text{GeV}]$. This cut rejects photoproduction and hard initial state radiation events as well as beam-gas events, as described in section 6.5;

- $\theta_e \leq \theta_{max}^{region}$. This cut is trigger dependent. In 1998/99 the third level trigger DIS03 (a medium-Q^2 neutral-current trigger slot) applied a 25 cm radius cut around the rear beam hole. The maximum values of θ_e in each region were chosen such that an electron would be detected in the RCAL outside the 25 cm limit plus a safety margin of some centimeters. Thus the critical angle for region 1 (RCAL side) is smaller than the critical angle in region 3 (FCAL side). Cut values are listed in table 7.1;

  Other constraints on the limits of the electron angle were found. The selection method implies an effective lower-Q^2 cut (see formulas 5.5 and table 7.1). This cut was lowest in region 3, where a higher θ_e cut can be applied to the scattered electron. To avoid generating a very large sample of Monte Carlo events, the lower Q^2 nominal limit in region 3 was increased by lowering the θ_e cut.

  Some bias in the selection efficiency was observed in region 3. Events generated with $z_{VTX} \sim 50$ cm were detected with greater efficiency than events generated with a larger z_{VTX}. This was found to be caused by the dead material in the solenoid. Events generated with higher z_{VTX} were more likely to have an electron crossing the solenoid where it lost part of its energy and failed to pass the selection cuts. A lower θ limit was imposed on events in this region such that they would not cross the volume of the solenoid. This, on the other hand, lowered the detection efficiency in that region. In other regions most of the electrons were still detected in the RCAL and a lower θ_e cut was found unnecessary.

- $\theta_{had} \in \left[\theta_{had,min}^{region}, \theta_{had,max}^{region}\right]$. In order to ensure a precise and efficient vertex reconstruction, at least one track was required to be reasonably close to the optimum polar angle in the region to which the event belongs. The optimum angles are 45° in region 1, 90° in region 2 and 135° in region 3. Tracks with such angles and reasonably high

	θ_{min}	θ_{max}	$\theta_{had,min}$	$\theta_{had,min}$	Q^2_{min}
Region 1	—	152°	29°	90°	98 GeV^2
Region 2	—	161°	49°	129°	45 GeV^2
Region 3	154°	166°	90°	150°	26 GeV^2

Table 7.1: Angular cuts applied to events in the three CTD regions and effective Q^2 limits.

transverse momenta will cross all the CTD superlayers and be well reconstructed. The range $\left[\theta^{region}_{had,min}, \theta^{region}_{had,max}\right]$ was chosen such that tracks with polar angles in those ranges will cross several stereo CTD superlayers. Their polar angle, and therefore the z position of the vertex, will then be correctly measured. The values of the θ_{had} cuts are listed in table 7.1.

The W method

An independent method was used to verify the results of the Q&R method. The former is based on reference [95] but was used in the present study with an important simplification. In the original method the vertex reconstruction was done using the electron-seed algorithm, while the standard multi-vertex algorithm was used in this study. As a result, it was found that the measured vertex distribution was biased in the z_{VTX} range of [50,100] cm. The method is, nevertheless, essentially unbiased in the remaining z_{VXT} range of [-100,50].

The idea underlying the W method is that well measured tracks should always be present in the selected events. To achieve this goal, NC DIS events were selected with the same trigger bit as above and using the same cuts on the electron energy, SINISTRA95 probability and δ_{tot} as for the Q&R method. The following additional cuts were applied on the scattered electron polar angle, θ_e, and on the hadronic angle, γ_h: $90° < \theta_e < 151°$ and $27.5° < \gamma_h < 90°$. Thus, an event originating near the FCAL end of the CTD has a good electron track even if no hadronic tracks were reconstructed. In the same way, an event near the RCAL end of the CTD will have well measured hadronic tracks emerging from the primary vertex even if the electron leaves no track in the CTD. Figure 7.2 illustrates the topology of accepted events that occur near either end of the CTD. The fact that the electron-seed vertex finding algorithm was not employed causes the reconstructed vertex position to be biased for $z_{VTX} > 50$ cm due to hadronic tracks emerging from secondary vertices.

From the cut on the electron polar angle, $\theta_e < 151°$ (which is close to the value used in region 1 in the Q&R method), it can be seen that this method is restricted to higher-Q^2 events than the previous method. The additional restrictive cuts on the hadronic polar angle contribute further to reduce the acceptance of the method in the low-Q^2 region.

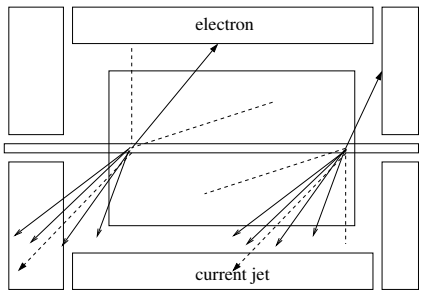

Figure 7.2: Topology of events selected by the W method. The calorimeter and CTD are represented schematically. The two events shown represent extreme cases of events happening near either end of the CTD.

7.1.2 Measurements and tests using Monte Carlo

A Monte Carlo sample was generated with a flat vertex distribution over the whole z_{VTX} range of [-100,100] cm. It consisted of around one million NC DIS events, generated with DJANGOH [75], with $Q^2 > 5$ GeV^2. This sample was used to check that the selection methods were indeed unbiased and also to establish the relative normalisation of the Q&R method in each z region.

The normalisation of the vertex distributions measured with both methods relied on the flat vertex Monte Carlo sample. This allowed the measurement of what was called the *selection efficiency*, ε, of the methods, which was simply the ratio of selected to generated events in bins of the generated z_{VTX} (figure 7.3). This quantity was then used to normalise the three regions of the vertex distribution measured in data and MC with the Q&R method and also for the relative normalisation of the distributions found with both methods. This was done by scaling the distributions in each region by the inverse of the calculated efficiency, $(1/\varepsilon_i)$.

As can be seen in figure 7.3, the Q&R method is essentially unbiased in each z_{VTX} region. The selection efficiency of the W method shows some z dependence for $z_{VTX} >$ 50 cm.

A fraction of the photoproduction events pass the cuts imposed by the minimum bias methods. This is not a problem in itself as long as the selection efficiency is unbiased as the true vertex distribution is independent of the physics process being observed. Photoproduction Monte Carlo events were used to determine the influence of the PHP background on the vertex distributions obtained using both methods. Figure 7.4 **a)** shows the true vertex distribution found in a PHP MC sample, along with the distributions of events accepted by each method. Figure 7.4 **b)** and **c)** show the ratios of the vertex distributions found by each method (normalized by $1/\varepsilon_i$ as given by the flat vertex MC), to the generated vertex distribution. Few events are selected by the W method. The Q&R

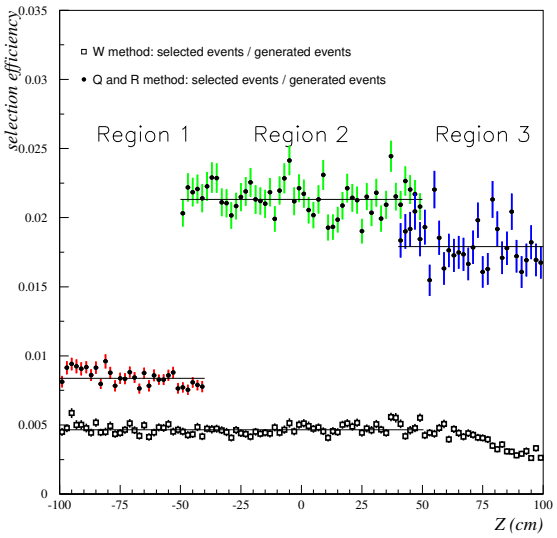

Figure 7.3: Measurement of the selection efficiency in the Q&R and the W methods with the help of the flat vertex MC sample.

method is more affected by this background but is essentially unbiased, except in the region $z_{VTX} \in [-100, -75]$cm.

7.1.3 Measurement of the 1998/99 vertex distribution

The final sample corresponded to an integrated luminosity of 12.23 pb^{-1}. The vertex distributions measured for the 1998/99 data with the Q&R and the W methods are compared in figure 7.5. The results of the W method are presented in the form of a band. The two distributions show good agreement, except in the region of $z_{VTX} \lesssim -80$ cm, which was attributed to the bias due to photoproduction background. In the region of $z_{VTX} > 50$ cm, no significant deviation was found between the distributions, possibly because of the few events selected by the W method and therefore of the large errors attributed to this distribution.

Fit of analytical function to the vertex

To obtain a smoothly varying vertex distribution and thus to avoid the effect of statistical errors in scarcely populated bins, the shape of the vertex distribution obtained with he Q&R method was fitted using an analytical function. The sum of five gaussian distributions (equation 7.1) was used in the fit. The result of the fit is shown in figure 7.6. The central peak was acceptably modeled by one gaussian. Satellites at $z_{VTX} \approx \pm 70$ cm were well described by two smaller gaussians. The two remaining gaussians were necessary to

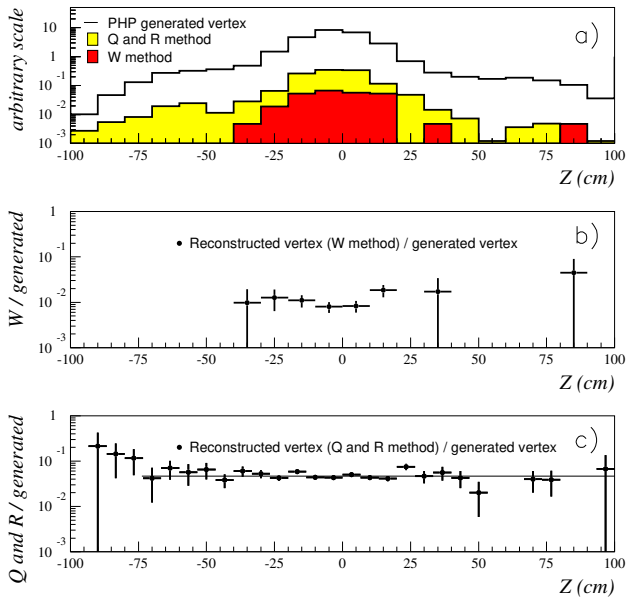

Figure 7.4: Vertex distributions determined in a photoproduction Monte Carlo sample using both methods. **a)** Generated and selected photoproduction MC events. **b)** and **c)** Photoproduction MC events selected by the each method, scaled by $1/\varepsilon$ and divided by the generated vertex distribution. A fit to the ratio in the range $[-73, 100]$ cm is also shown in plot **c)**.

fit the shape of the vertex distribution in the transition region between the main peak and the satellites. The region of $z_{VTX} \in [85, 100]$ cm was omitted from the fit.

$$\begin{aligned} f(Z) &= A_1 \exp -\frac{(z-\overline{Z}_1)^2}{2\sigma_1^2} + A_2 \exp -\frac{(z-(\overline{Z}_2-\overline{Z}_1))^2}{2\sigma_{23}^2} \\ &+ A_3 \exp -\frac{(z-(\overline{Z}_3-\overline{Z}_1))^2}{2\sigma_{23}^2} + A_4 \exp -\frac{(z-(\overline{Z}_4-\overline{Z}_1))^2}{2\sigma_{45}^2} \\ &+ A_5 \exp -\frac{(z-(\overline{Z}_5-\overline{Z}_1))^2}{2\sigma_{45}^2} \end{aligned} \tag{7.1}$$

The parameters $Z_i, (i = 2, 3, 4, 5)$ refer to the distance from the main peak to the position of secondary vertices. The results of the fit to the measured vertex distribution (normalized to unit area) are summarized in table 7.2.

The use of this function in the fit can be related to the physical origin of the longitudinal vertex distribution. The width of the main peak and of the satellites is a result of the lengths of the proton bunches ($\sigma_Z^p = 8.3$ cm) [16], which are much longer than the electron beam bunches ($\sigma_Z^e = 8.5$ mm). This width also reflects changes in the collider parameter tuning, which causes the vertex average position to change from run to run.

The proton and electron beams have a "bucket" structure given by the RF accelerating frequencies $\nu_p = 208.13$ MHz for the protons and $\nu_e = 499.776$ MHz for the electrons [44].

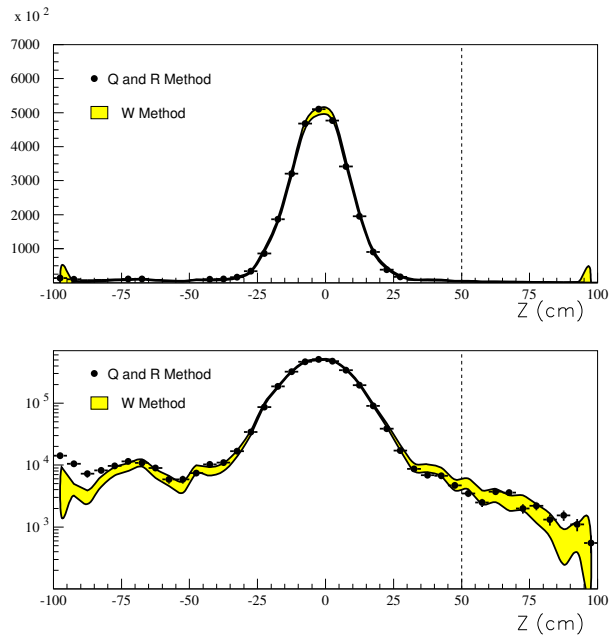

Figure 7.5: Comparison between the two methods found in 98/99 data. The same plot is shown with a linear and a logarithmic scale, so that all the features can be appreciated.

peak	$\overline{Z}(cm)$	$\sigma_Z(cm)$	height
main peak (1)	-2.23	10.5	0.071
satellite 2	33	10	9.82×10^{-3}
satellite 3	-36	10	1.25×10^{-3}
satellite 4	64	12	4.094×10^{-4}
satellite 5	-71	12	1.483×10^{-3}

Table 7.2: Parameters obtained from the fit to the measured vertex distribution. The $\overline{Z}$ value quoted for satellites corresponds to their distance from the centroid of the main peak, not to the ZEUS coordinate system.

The interval between buckets is then 4.8 ns or 144.14 cm in the proton beam and 2.0 ns or 60.03 cm in the electron beam.

As the trigger time window is finite, not only events happening at $t = 0$ ns will be taken, but also events taking place somewhat sooner or later than the nominal time. Buckets neighboring the nominal proton and electron bunches (known as satellite bunches) are partially filled due to unavoidable imperfections in the injection of the beams into HERA. A fraction of the recorded interactions correspond to the nominal electron bunch and an early or late satellite proton bunch, and vice versa. Events involving the nominal electron bunch and a satellite proton bunch happen at a time $\Delta t = 4.8/2$ ns $= 2.4$ ns and a distance of around 72 cm with respect to the nominal time and vertex. Events involving the nominal proton bunch and a satellite electron bunch happen at a distance of around 30 cm from the main peak, corresponding to the time $\Delta t = 2.0/2$ ns $= 1.0$ ns. These

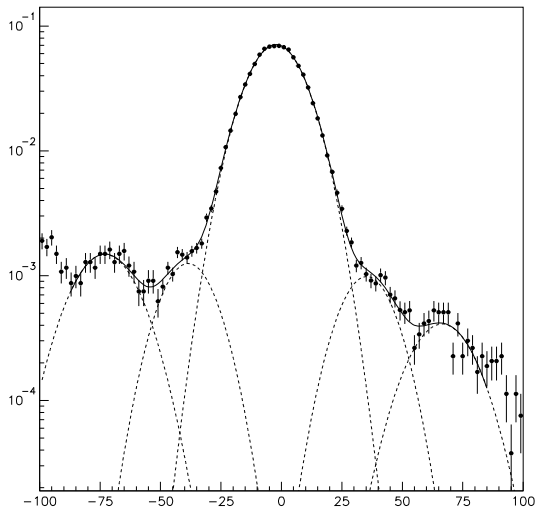

Figure 7.6: Fit to the measured minimum bias vertex distribution (normalized to unit area) found in 1998/99 e^-p data. Individual gaussian functions are represented as dashed lines.

values agree reasonably well with the values of the fit parameters. Both the proton bunch-length and the separation between satellites and the main peak are close to the predicted values. The differences could be attributed to variation of HERA tuning parameters.

The vertex distribution that resulted from this study was integrated in the Monte Carlo production software and used for the generation of events for the analysis of 1998/99 data.

The process of normalisation of the three z regions and of fitting an analytical function to the resulting distribution yields a natural way to determine the systematic uncertainties arising from the vertex distribution determination. The normalisation of regions 1 and 3 was first varied by the standard deviation of the respective selection efficiencies. The normalisation of region 2 is constrained by the requirement that the integral of the distribution must remain constant. The resulting distributions were then re-fitted to the same analytical function, yielding new vertex distribution functions. The ratios of these to the nominal vertex distribution function, shown in figure 7.7, were used to reweight the vertex distribution found in the neutral current analysis. The process was repeated varying the normalisation of regions 1 and 3 in a correlated and anti-correlated way. The maximum deviations found in the extracted cross sections were taken to correspond to the systematic uncertainty in the simulation of the vertex distribution.

7.1.4 The minimum bias vertex distribution for 1999/00

The W method was used in a different analysis [74] to determine the minimum bias vertex distribution for 1999/00. In this case, the electron-seed vertex finding algorithm was used, which corrected for the z dependence found in the simplified algorithm at $z > 50$ cm. The

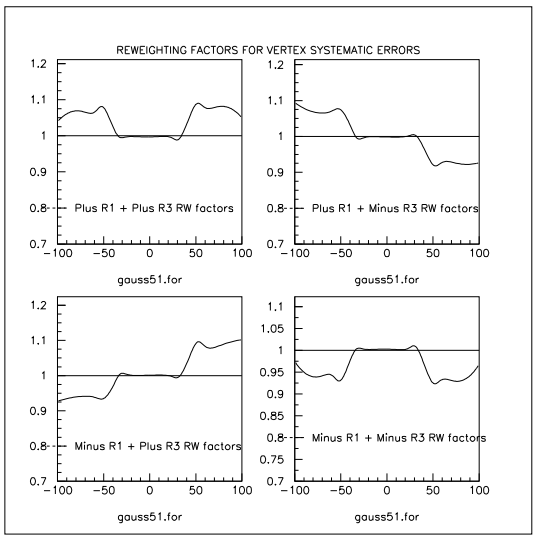

Figure 7.7: Reweighting factors used in the determination of the systematic uncertainty arising from the simulation of the vertex distribution for 1998/99.

increased statistics collected during the 1999/00 run period made the use of this method more convenient than the Q&R method. In this case, the measured vertex distribution was fitted to a function consisting of the sum of 5 Gaussian distributions plus a constant term. The distribution and the fitted function can be seen in figure 7.8.

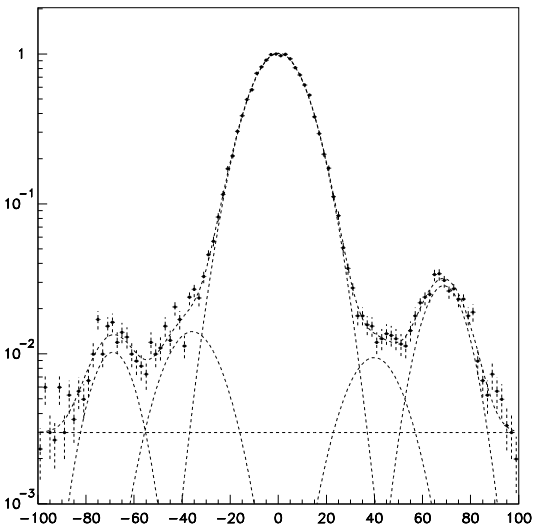

Figure 7.8: Minimum bias vertex distribution for the 1999/00 data sample. Reproduced from reference [74].

The systematic uncertainty associated with the determination of the vertex distribution was, in this case, obtained by varying the height of the outer satellites by three standard deviations of their nominal height. The ratios of the new distributions to the nominal

one are, as before, used to reweight the vertex distribution found in the neutral current analysis. The maximum deviations were taken as the systematic uncertainties associated with the simulation of the vertex.

7.2 Simulation of the QCD cascade

The double-angle method that was used to reconstruct the kinematic variables (see chapter 5) relies on the accurate and unbiased reconstruction of the hadronic angle, γ_h. This angle, on the other hand, is reconstructed from the spatial position and value of energy clusters left in the calorimeter by hadronic final state particles. The topology of the final state is determined by the complex processes of the QCD cascade and hadronisation, involving many partons, and by the non-perturbative processes that lead to the formation of final state hadrons.

Phenomenological models are employed in the Monte Carlo generators to simulate the complicated hadronisation process. The details of the simulated hadronic final state can affect both the measurement of the hadronic angle and the selection of MC events. A poor simulation of the hadronic final state can combine with other sources of systematic uncertainty such as the dead material map, backscattering of energetic particles in the calorimeter and the finite granularity of the calorimeter and result in a bias in the measured γ_h.

To investigate these effects, the hadronic final state in high-Q^2 NC e^+p DIS events was simulated with ARIADNE and with LEPTO MEPS and the observed hadronic energy-flow was compared to 1999/00 e^+p data. The definition of the quantities measured are given in section 7.2.1. These quantities, being essentially the ratios of the energy deposited in a region of space to the number of events, are independent of the value of the NC DIS cross section.

7.2.1 The hadronic energy-flow – definitions

The hadronic energy-flow was defined as the average energy deposited in the calorimeter, by hadronic final-state particles in bins of the polar angle, θ, radius (on the faces of the FCAL and RCAL) or z (on the face of the BCAL). It was calculated, for a given bin, as the ratio of the total hadronic energy deposit in the bin to the number of events in the data sample. For real data:

$$\overline{E}_j^D = \frac{\sum_{i=1}^{N_D} E_i^j}{N_D} \tag{7.2}$$

where E_i^j is the energy found in bin j in event i, and N_D is the total number of accepted events in the data sample. The energy cluster associated with the scattered electron is

excluded from the sum. The hadronic energy-flow quantities that will be presented are: $\frac{1}{N}\frac{dE}{d\theta}$, $\frac{1}{N}\frac{dE}{dR}$, or $\frac{1}{N}\frac{dE}{dz}$, where N denotes the number of events, and R is the radius in the face of the FCAL and the RCAL.

Studies of the transverse energy flow can be found in the literature [128, 129]. This quantity corresponds to weighting the energy in formula 7.2 by $\sin\theta_i$, where θ_i is the polar angle of the i^{th} energy deposit with respect to the vertex. For the purpose of the present study, the definition in formula 7.2 was found to be more interesting as it gives a more direct picture of the energy deposits in the calorimeter. For the same reason, the energy flow was not plotted against pseudo-rapidity but against polar angle or calorimeter dimensions instead, and the laboratory frame was used instead of the Breit frame.

A similar expression was used for Monte Carlo:

$$\overline{E}_j^{MC} = \frac{\sum_{i=1}^{N_{Pass}} \omega_i E_i^j}{\mathcal{L}_{data} \sum_{i=1}^{N_{Pass}} \omega_i} \tag{7.3}$$

where ω_i is the MC event weight, used to combine different MC sets, N_{Pass} is the number of MC events that pass the selection cuts, and $\mathcal{L}_{data}$ is the data sample luminosity.

The error found in the hadronic energy-flow measured in real data was calculated as:

$$\delta_{\overline{E}_j^D} = \frac{\sqrt{\overline{E_j^2} - \overline{E}_j^2}}{\sqrt{N_D}} \tag{7.4}$$

The error bars plotted with the hadronic energy-flow distributions correspond to the uncertainty in the value of the mean hadronic energy-flow, rather than the spread of the variable around the mean.

For Monte Carlo events, an equivalent formula can be obtained by replacing N_D by the equivalent number of events $\tilde{n}_{MC}$, defined as:

$$\tilde{n}_{MC} = \frac{(\sum^N \omega_i)^2}{\sum^N \omega_i^2} \tag{7.5}$$

But, in most instances, the uncertainty in the hadronic energy-flow measured in MC events is assumed to be much smaller than the uncertainty in real data and is not calculated.

The hadronic energy-flow as defined above is sometimes inadequate to observe small differences between real and simulated data. The reason is that the hadronic energy-flow can vary significantly as a function of polar angle, forcing the use of logarithmic scales. A different definition was also used where the energy flow is calculated in each bin by counting only events with some energy deposit in that bin (equation 7.6). The denominator $N_{E>0}$ in equation 7.6 becomes smaller than N_D from equation 7.2 and the details of the distribution can be seen more clearly. This new definition can be more sensitive to the details of the hadronic final state. This algorithm will be referred to as the *hadronic energy-flow in fired bins.*

$$\overline{E}_j^D = \frac{\sum_{i=1}^{N_{E>0}} E_i^j}{N_{E>0}} \tag{7.6}$$

7.2.2 Data and Monte Carlo samples

The data set employed in this study corresponds to the entire 1999/00 e^+p sample. As in the neutral current analysis, events were selected by tagging the scattered positron using the SINISTRA95 electron finder, and the corrections described in chapter 6 were applied. The calorimeter cells associated with the electron found by SINISTRA95 were excluded from the calculation of the hadronic energy-flow. The selection was done as in the neutral current analysis. The list of off-line cuts used for this study corresponds to the list given in section 6.5 excluding the MC validity cut ($y_{JB}(1-x_{DA})^2 > 0.004$) and the cut on the radius of the hadronic angle projection on the FCAL ($R_{\gamma_h}^{FCAL} > 20$ cm). The latter cuts were adopted after the present study was finished. The resulting sample is a rather pure set of 162011 NC DIS events with a measured Q^2 greater than 185 GeV^2 and $1.7 \times 10^{-3} < x < 1.0$.

The Monte Carlo sample generated with ARIADNE consists of 1.2 million NC DIS events in 10 samples generated with minimum Q^2 values ranging from 100 GeV^2 to 50 000 GeV^2. A different sample of around one million events was generated using LEPTO MEPS, instead of ARIADNE, in the same Q^2 ranges.

These non-diffractive samples were combined with 3 sets of diffractive events generated using RAPGAP with minimum Q^2 values of 100 GeV^2, 400 GeV^2, and 1250 GeV^2, amounting to around 300 000 events. The two MC samples were combined so that 6.7% of the final MC sample corresponded to diffractive events. It should be noted that the necessary inclusion of diffractive MC has a direct influence on the hadronic energy-flow, especially at low polar angles where diffractive events have little or no hadronic activity.

The Calorimeter energy scale

The hadronic energy scale affects the measurement of the hadronic energy-flow directly, resulting in differences between real and simulated data which do not originate in the simulation of the QCD cascade. The energy scale corrections described in section 6.4.4 were not available at the time of this study. A previous, less precise energy scale correction was used instead.

To verify the effect of the hadronic energy scale corrections, the ratio of the measured and calculated transverse momenta, P_T^{had}/P_T^{DA} (see sections 6.4.4 and 5.3), was compared in real data and MC events, and is shown in figure 7.9 versus the hadronic angle, γ_h. The effect of the CORANDCUT correction of losses in dead material and in the super-cracks between calorimeter sections can be clearly seen by comparing the two upper plots. The first γ_h bin shows a ratio of P_T^{had}/P_T^{DA} which is significantly different from 1 for both real and simulated data, due to the loss of particles through the forward beamhole. The

arrows show the approximate angular limits of the BCAL as viewed from the nominal vertex. In the lower plots, it can be seen that most points agree in data and MC events corrected by CORANDCUT to better than 1%. Events with low measured γ_h, where the current jet is close to the FCAL beamhole, show a somewhat larger difference between real and simulated data. It is not clear whether this disagreement is caused by a wrong energy scale correction factor, by a poor simulation of the hadronic flow or by a combination of these factors. Excluding the part of the FCAL closest to the beampipe, these plots show that the calorimeter energy scale can be considered well determined to within 1.5%. Thus, it can be disregarded as a possible cause for any substantial disagreement between the hadronic energy-flow found in data and MC events, especially at large hadronic angles.

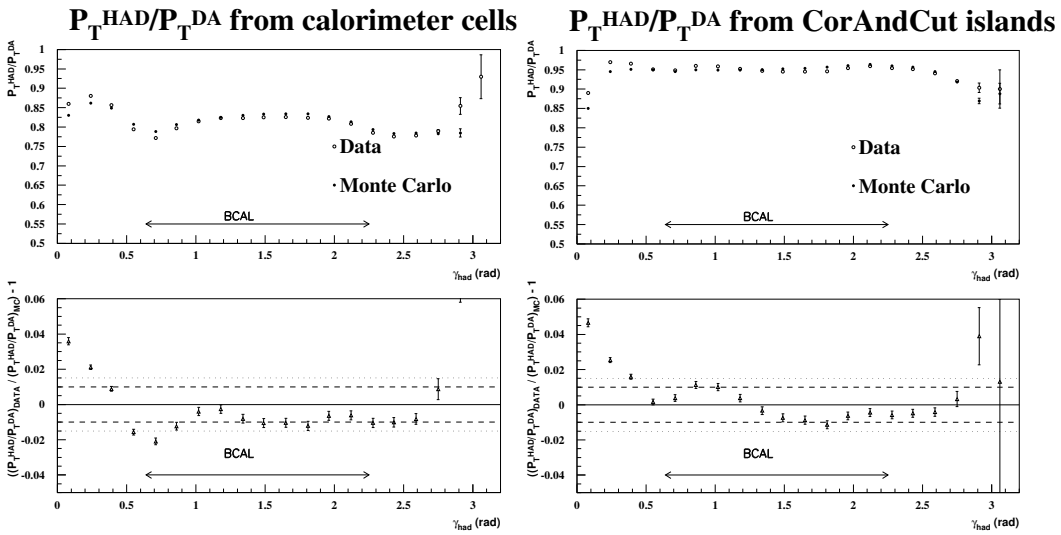

Figure 7.9: Comparison of ratio P_T^{had}/P_T^{DA} in data and MC, versus the hadronic angle, for raw calorimeter cell information (left) and after CORANDCUT corrections (right). The dashed and dotted lines mark, respectively, the 1% and 1.5% levels of agreement between data and ARIADNE MC.

7.2.3 Results

First impressions

Figure 7.10 shows the energy-flow distributions found in data and in both MC samples in bins of the measured hadronic angle γ_h. It can be seen that events with a low hadronic angle have energy deposits in the very forward direction (i.e. at small polar angles) which fall rapidly with increasing polar angle θ. In particular, the energy deposits corresponding to the current jet can only be clearly distinguished from the proton remnant in the third γ_h bin ($0.314 < \gamma_h < 0.471$ rad). For higher γ_h the current jet, transition region, and the vicinity of the proton remnant become clearly identifiable. The angular position of the current jet can be seen to evolve towards higher polar angles according to the hadronic angle range of the bin.

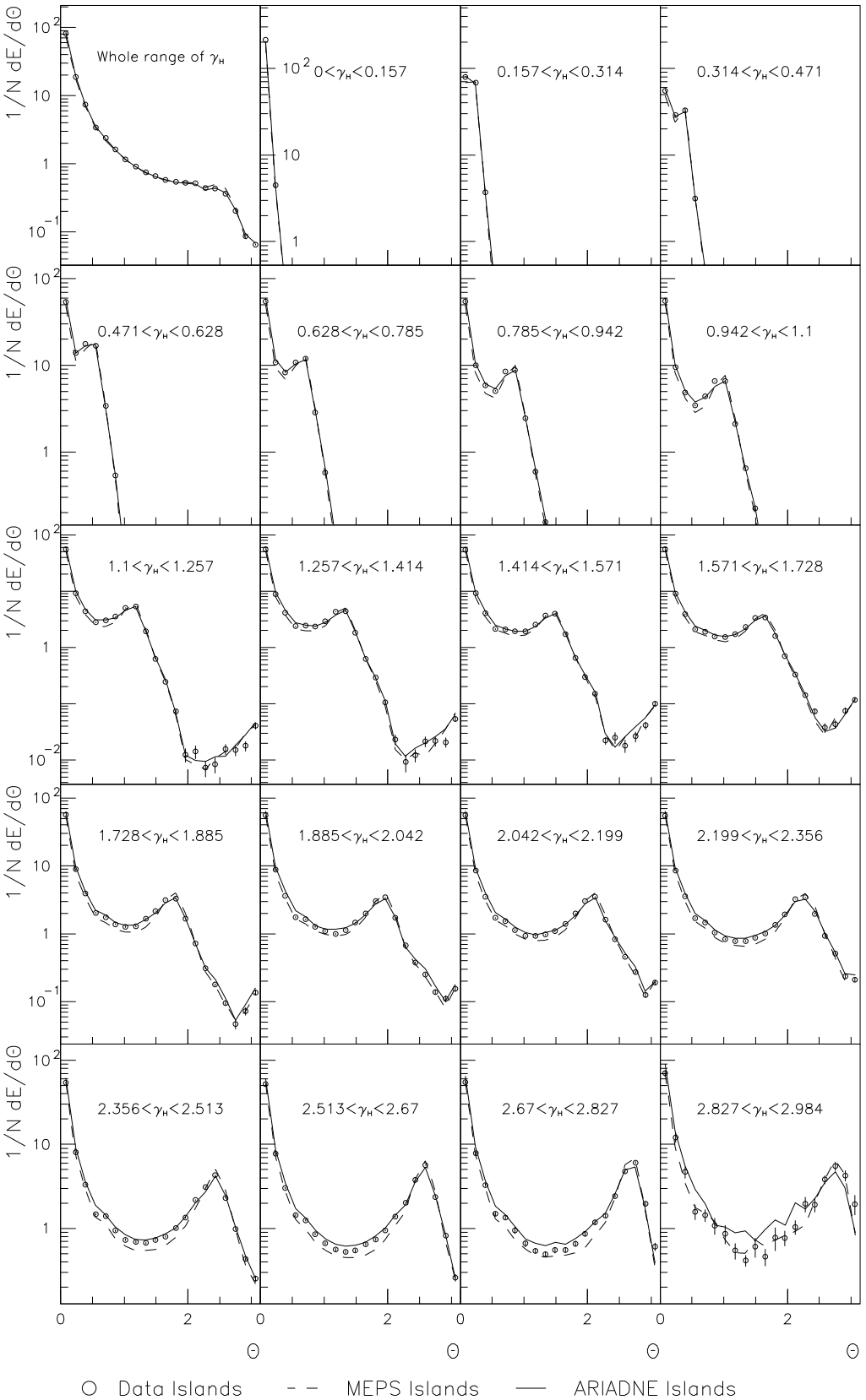

Figure 7.10: Distributions of energy-flow in CORANDCUT islands versus polar angle θ in bins of hadronic angle γ_h. ARIADNE and LEPTO MEPS MC are compared to data points.

Both Monte Carlo generators reproduce the data in its broad features, but several problems become evident when looking at the details. Figure 7.11 shows the hadronic energy-flow in the BCAL, in selected bins of γ_h, plotted against the z-coordinate of impact on the face of the calorimeter. Here again the peak corresponding to the energy deposits from the current jet can be seen to evolve from the front of the BCAL to the rear direction as γ_h increases. Here it becomes clear that ARIADNE reproduces the data well, whereas LEPTO MEPS shows a worse level of agreement. In comparison to data, LEPTO MEPS is seen to produce more energy deposits in the peak and rear direction (lower z, to the left of the peak in the plots) and less energy in the transition region between current jet and proton remnant (to the right of the current jet peak in the plots). The error bars in the hadronic energy-flow measured in data are, in general, smaller than the circles marking the points. This shows that the disagreement between data and MC can be significant.

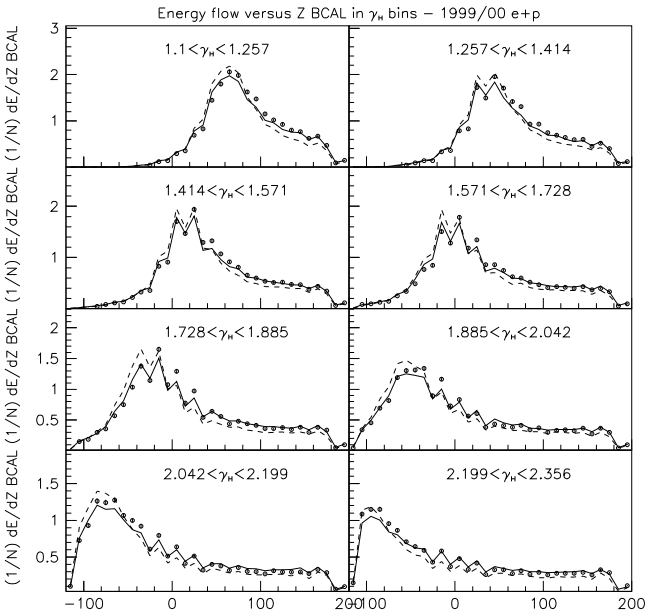

Figure 7.11: Distributions of energy-flow in CORANDCUT islands found in selected bins of γ_h in the BCAL versus z (in cm). Both ARIADNE (solid line) and LEPTO MEPS (dashed line) MC are compared to data (open circles).

Another interesting measurement is that of the energy-flow distribution in fired bins as a function of radius in the FCAL, shown in figure 7.12 for some bins of γ_h. Here the simulation of the proton remnant and the colour flow between current jet and proton remnant are, again, clearly visible. As figure 7.12 shows, there is a noticeable disagreement between data and both MC samples in the region of small FCAL radius. These plots were obtained using the definition of the hadronic energy-flow from equation 7.6, i.e., counting in each FCAL radial bin only those events which left an energy deposit in the bin greater than the noise cuts. The bins are 10 cm wide, which means that the first bin accounts

for islands reconstructed in the space occupied by the beamhole and the Forward Plug Calorimeter (a small calorimeter located in the FCAL beamhole, see below). The second and third bins correspond roughly to the FCAL inner ring of towers. This disagreement persists when the current jet is out of the FCAL angular range, as can be seen from the last two γ_h bins shown in the figure.

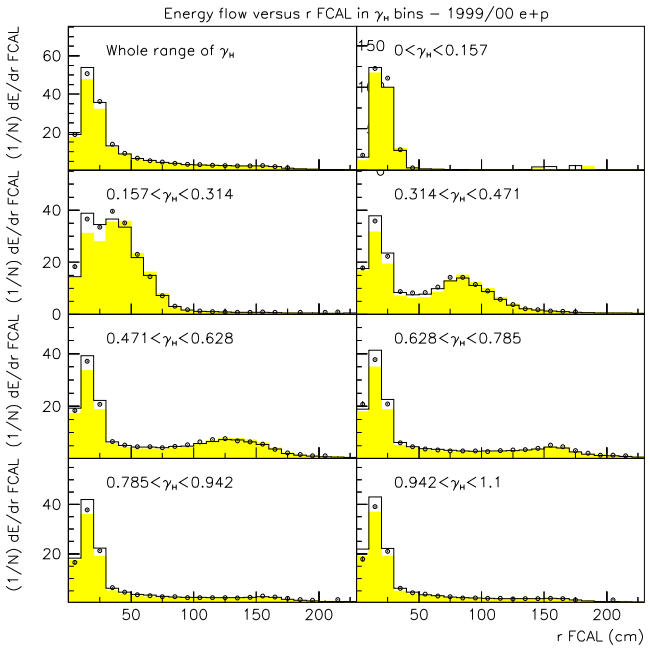

Figure 7.12: Distribution of energy flow in fired bins in CORANDCUT islands versus FCAL radius. Both ARIADNE (solid line) and LEPTO MEPS (shaded histograms) are compared to real data.

From the difference between the data distribution and both MC histograms in figure 7.12, it seems clear that there is a problem with the simulation of the energy flow in the vicinity of the proton remnant.

Hadronic energy-flow in the forward direction

Two different problems have arisen from the previous discussion: the energy flow in the forward direction, close to the FCAL beamhole, and the MC description of the hadronisation in the bulk of the detector, i.e. at larger hadronic angles. This section describes the attempts made in this study to solve the disagreement in the forward region. The level of agreement between real and simulated data at higher polar angles is described in section 7.2.3 along with the method used to estimate the systematic uncertainty associated with the QCD-cascade models implemented in the Monte Carlo generators.

Varying the fraction of diffractive Monte Carlo: As has been noted before, the fraction of diffractive MC events which is mixed with the non-diffractive MC sample has

a direct influence on the hadronic energy flow, especially in the forward direction. This fraction was varied until a reasonable agreement was found, in the hadronic energy-flow, between real data and the ARIADNE MC sample. It was found that a fraction of around 27% of RAPGAP was necessary to describe the data. This is clearly too great a fraction, when compared to the previously established 6.7% (see section 4.1.3). Even with this increased fraction of diffractive MC events, the distributions found in data and MC events are somewhat different.

Modifying MC parameters: The disagreement found between the energy flow measured in real data and Monte Carlo may be removed through the use of different parameters. While a complete tuning of the MC generators was beyond the scope of the present study, the values of some parameters were changed in an attempt to improve the agreement. Small samples of MC events (around 100 thousand events each) were generated with modified values of the proton intrinsic transverse momentum, k_T, and the transverse momentum used in the simulation of the proton remnant [76], referred to here by k_T^{rem}.

For the generation of these events, the QCD-cascade stage was performed by ARIADNE. The width of the (gaussian distributed) k_T and k_T^{rem} was modified from the default values of 0.44 GeV and 0.35 GeV, respectively. In the case of the k_T parameter, the alternative values of 0.01 GeV and 1.66 GeV were used, while for k_T^{rem} the values 0.05 GeV and 0.8 GeV were used.

Small differences were found in the hadronic energy-flow measured in the new MC samples, but the agreement with data did not improve. No further attempts to tune the Monte Carlo parameters were made.

Energy flow in the FPC: A small calorimeter located in the FCAL beamhole, the Forward Plug Calorimeter (FPC) can, in principle, provide valuable information on the hadronic activity at very small angles in the forward direction. The FPC is a lead-scintillator sampling calorimeter with a total depth corresponding to 5.4λ, located around the beampipe inside of the FCAL beamhole [130]. Figure 7.13 shows a schematic representation of the FPC. The information provided by this detector would allow the differences in hadronic energy-flow between data and MC events in the very forward direction to be corrected. Using the FPC in the analysis of neutral current data could also provide a more accurate reconstruction of γ_h at low y, extending the accessible phase space. The presence of the FPC in the FCAL beamhole, on the other hand, causes the existence of cross-talk or "energy leakage" between the FPC and the inner FCAL cells. Showers developing primarily in either calorimeter may be partially detected in adjacent cells of the other, thus biasing the measurement of energy deposits.

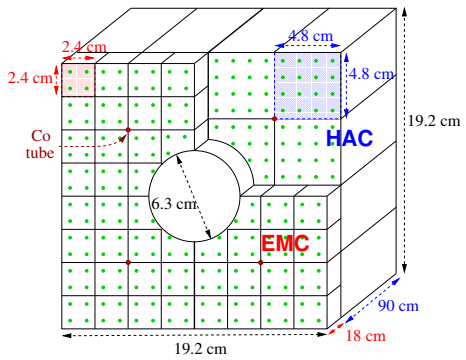

Figure 7.13: Schematic representation of the Forward Plug Calorimeter. Some cells in the electromagnetic (EMC) section are omitted to allow a view of four cells from the hadronic section (HAC). The beampipe is located in the central hole in the FPC.

The energy flow in the FPC was measured in real and simulated data. The two measurements were found to disagree severely. This was attributed to a known bug in the treatment of the energy deposits in the simulated FPC that resulted in a wrong distribution of energy deposits along the x direction. No further investigation of the FPC was undertaken.

Island → cells splitting in the FCAL: Plotting the energy flow in calorimeter cells (without dead material or backscattering corrections), it was found that the MC does not describe the data well. In fact, both MC generators show a lack of energy both in the current jet and in the transition region, when analysed with calorimeter cells. This can be attributed to an imprecise parametrisation of the dead material in the detector or to limitations in the simulation of the QCD cascade. On the other hand, the simulation of energy deposits in FCAL cells close to the beampipe was found to be reasonably accurate when the current jet is far from the forward region (note the first few bins, up to a radius of around 40 cm in figure 7.14).

The acceptable agreement between MC and real data in the energy flow measured in calorimeter cells near the centre of the FCAL motivated an attempt to quantify the systematic uncertainty asssociated with the simulation of the QCD cascade in the forward direction. Islands formed within 30 cm of the centre of the FCAL beamhole were separated into their constituent cells. The cell energy and position information was combined with the islands measured elsewhere in the calorimeter and used in the reconstruction of the hadronic angle.

In figure 7.15 the resulting measured hadronic angle is plotted against the hadronic angle obtained from the generated MC variables, γ_{true}. This variable was obtained from

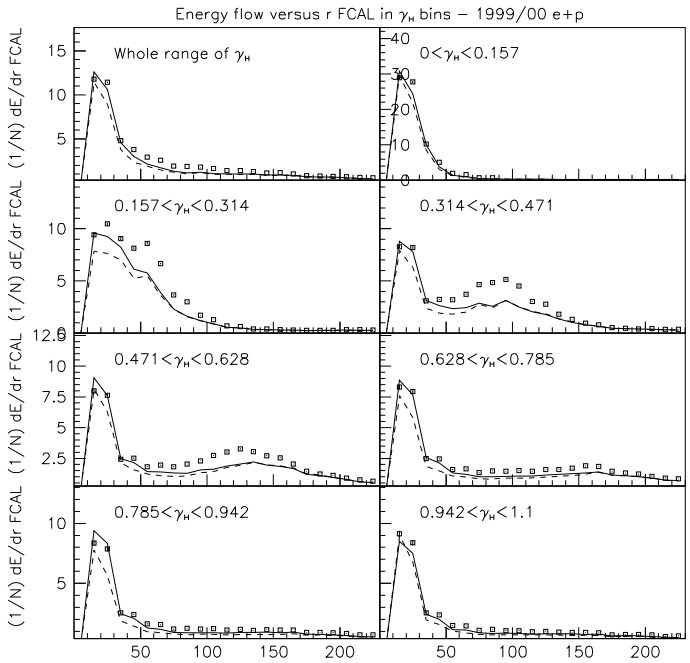

Figure 7.14: Hadronic energy-flow in fired bins measured from the un-corrected information from calorimeter cells plotted versus FCAL radius (in cm). Although the MC description of the measurements is generally poor, the first four radial bins (i.e., up to a FCAL radius $R_{FCAL} < 40$ cm) are, in general, acceptably described by ARIADNE. The first radial bin corresponds to an area with no calorimeter cells (beampipe and FPC). The solid line corresponds to ARIADNE and the dashed line to LEPTO MEPS. The data points are represented by open squares.

the definition of γ_h (formula 5.4) and the formulas 5.8 for Q^2 and y from the Jacquet-Blondel reconstruction method [93], resulting in the expression:

$$\cos\gamma_h^{true} = \frac{Q^2_{MC}(1-y_{MC}) - 4E^2_{e,beam}y^2_{MC}}{Q^2_{MC}(1-y_{MC}) + 4E^2_{e,beam}y^2_{MC}} \tag{7.7}$$

where Q^2_{MC} and y_{MC} stand for the generated Q^2 and y.

The upper plot in figure 7.15 shows the γ_h angle, measured in the default way and with island $\rightarrow$ cells splitting, plotted against γ_{true}. A straight line is also shown which is defined by $\gamma_h = \gamma_{true}$. The island $\rightarrow$ cells splitting algorithm closely follows the default algorithm but deviates from the generated value of γ_{true} at low angles. A very pronounced deviation in measured γ_h exists at high γ_{true}. This is a region of very high y and of little statistical significance. Backscattering of particles from the rear calorimeter towards the FCAL, which is not corrected by CORANDCUT, may be one source of this deviation.

The effect of the modified algorithm is small for high hadronic angles, but is significant when the current jet is close to the FCAL beamhole. The application of this algorithm to the measurement of γ_h in the neutral current analysis gives an upper limit on the size of migrations in the kinematic plane caused by the description of the hadronisation in the

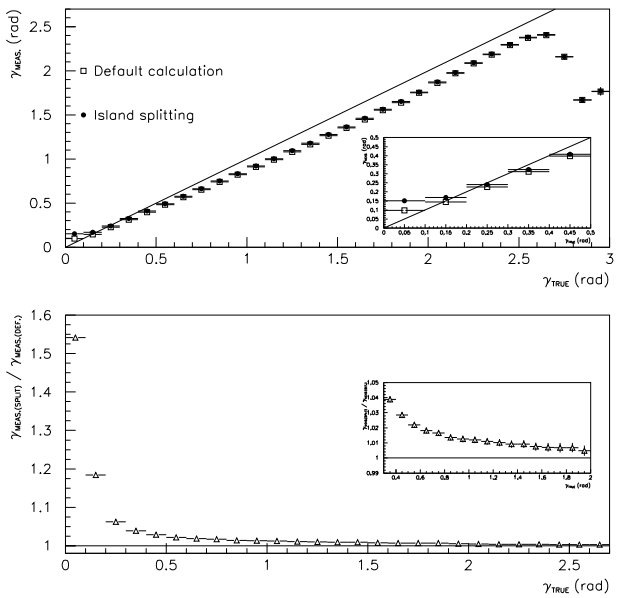

Figure 7.15: Hadronic angle, γ_h, measured with and without island → cells splitting in the FCAL and plotted against the "true" hadronic angle γ_{true} obtained from MC variables. The deviation caused by the splitting of FCAL islands reconstructed near the beampipe is quite high at very low angles, but decreases rapidly for higher hadronic angles (lower plot).

MC. The two extra analysis cuts, not used at the time of this study, contribute to avoid the region where the measurement of the hadronic angle can be significantly biased. As can be observed in figure 6.25, the Monte Carlo validity cut and the cut on the radius of the projected hadronic angle in the FCAL help to eliminate the region of very low hadronic angle, where the agreement between simulation and real data is poor.

Hadronic energy-flow away from the proton remnant

The hadronic energy-flow in fired bins provides a valuable clue to the systematic differences in the hadronisation between the Monte Carlo simulation and what is found in data. When plotting the energy flow in fired bins versus the z-coordinate in the BCAL, it was found that the data showed higher values than either MC model. The influence of the hadronic scale of the calorimeter was ruled out as the origin of this discrepancy both because of the ratios of p_T^{had}/p_T^{DA} shown in figure 7.9 and because the energy flow calculated using all events (formula 7.2) showed good agreement between data and ARIADNE MC.

Figures 7.16 and 7.17, show ratios of the average energy deposit in data and ARIADNE MC events, using the whole data and MC samples, in radial bins of FCAL and RCAL and in bins along z in the BCAL. The upper plots show the ratios of the average energy found in data and MC events. The lower plots show the same quantity, but now calculated by

counting, in each bin, only those events that leave some energy deposit in that bin. This is analogous to the calculation of energy flows in fired bins. The upper plots, then, reflect the total hadronic activity, integrated in phase space, and are not sensitive to the topology of the hadronic final state. The lower plots, on the other hand, are sensitive to the spatial distribution of the particles in the hadronic final state, as the number of calorimeter bins that are "fired" (i.e., which have some energy deposit), depend on the spatial distribution of the hadronic final state.

The error bars in figures 7.16 and 7.17 were obtained by propagating the uncertainty in the hadronic energy-flow found in real and simulated data (given by equations 7.4 and 7.5), considering the two to be uncorrelated. This may lead to an overestimation of the errors on the ratio, which is noticeable in the figures, as the simulated hadronic energy-flow is in fact correlated to that found in real data. The degree of correlation depends on the agreement between the simulation and the real data, so the error bars in the figures correspond, in fact, to an upper limit of the estimated uncertainty.

The upper plots seem to confirm that the energy scale correction employed is correct (see section 7.2.2). In figure 7.16, some structure is visible in the upper left plot. This can be attributed to deficiencies in the dead material map implemented in the MC simulation. The first three radial bins in this plot have been neglected because, as was found previously, the MC models do not describe the data close to the FCAL beamhole. The upper plots in figure 7.16 show very good agreement between data and MC events, and in figure 7.17 the upper plot shows a bias of around 0.5%, which is within the error attributed to the BCAL hadronic energy scale correction employed in this study.

The lower plots in these figures, on the other hand, show a bias of around 3% to 4% for all the calorimeter sections. This seems to indicate that the QCD-cascade process in the MC simulation leads to the energy in the hadronic final state being distributed among a greater number of bins, on average, than in data.

To compare the spatial distribution of the hadronic final state particles in data and MC events, the hadronic energy-flow in fired bins was plotted as a function of the fraction θ/γ_h, where θ is the polar angle of the energy deposits and γ_h is the measured hadronic angle. This is shown in the plots in the left column in figure 7.18. The transition region between proton remnant and current jet is clearly distinguishable at $\theta/\gamma_h < 1$. The energy deposits corresponding to the current jet form the peak close to $\theta/\gamma_h = 1$. Plots in the right column show the ratios of energy flow found in data and both ARIADNE and LEPTO MEPS.

From figure 7.18, it becomes clear that ARIADNE tends to have larger energy deposits in the transition region than those measured in the data. This seems to be the cause for the bias found in figures 7.16 and 7.17. In the transition region, the behaviour of MEPS

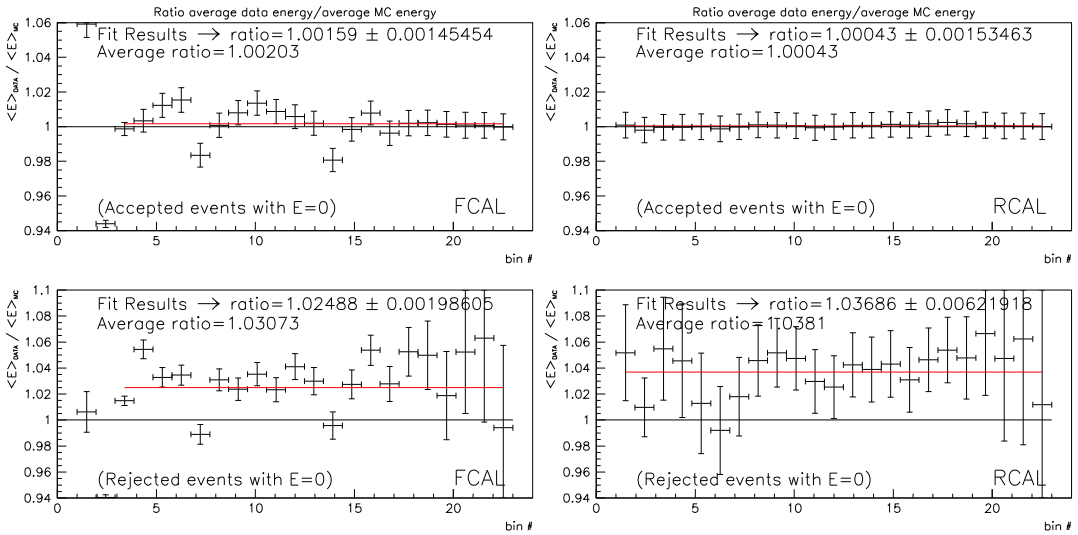

Figure 7.16: Ratio of average deposited energy in data and MC events in 10 cm radial bins in the FCAL and RCAL sections of the calorimeter. In the lower plots the average energies are calculated excluding data and MC events with no energy deposits in each bin. Some fluctuations are visible in the FCAL which are probably caused by dead material. The first two FCAL bins are excluded from the average ratio calculation.

seems to be the opposite to the one found in ARIADNE, with the hadronic energy-flow in the transition region being smaller than that found in data. The behaviour of MEPS, with respect to real data, in the current jet region is similar to the one found in the ARIADNE sample, but the trend with respect to ARIADNE is inverted with MEPS producing more hadronic energy-flow for $\theta/\gamma_h \gtrsim 0.9$.

The different behaviour observed in the two MC samples provide a natural way to determine the systematic uncertainty associated with the QCD-cascade models implemented in the MC generators. This subject is developed in the next section.

It should be noted that the differences found between data and MC may not be due to problems in the simulation of the hadronisation process alone. These differences may also be caused by the way in which the simulated final state particles interact with dead material, in the detector simulation, or combine with the effects of detector granularity and non-uniformity, etc. A systematic error that characterises the disagreement between the energy flow found in data and MC will be a measure of these other effects as well.

7.2.4 The hadronisation systematic uncertainty

Figure 7.18 shows that the hadronic energy-flow in the transition region between current jet and proton remnant is overestimated by ARIADNE and underestimated by MEPS. The hadronic energy-flow measured in real data is, in this region, generally between the results obtained from the two simulations. In a small region near the current jet, the energy flow

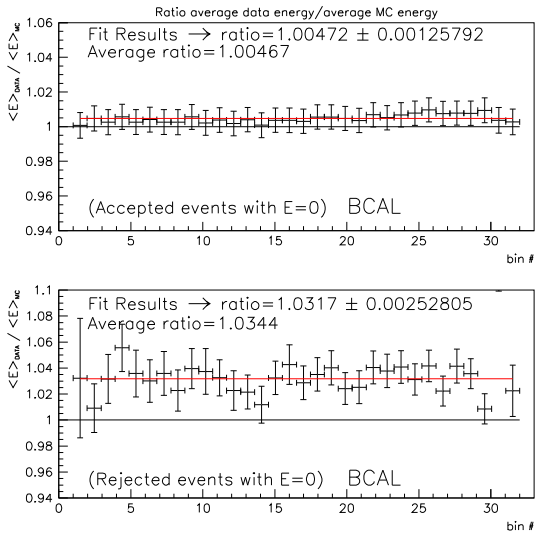

Figure 7.17: Ratio of average deposited energy in data and MC in 10 cm bins along BCAL z (lower edge of bin 1 corresponds to z = -110 cm; upper edge of bin 32 to Z=220 cm). In the lower plot the average energies are calculated excluding data and MC events with no energy deposits in each bin. In the upper plot, a bias of 0.5% is visible which may correspond to a small bias in the energy scale correction factors. This bias is within the error attributed to correction factors.

is underestimated by both ARIADNE and MEPS, and for $\theta/\gamma_h \gtrsim 0.9$ ARIADNE produces less hadronic energy-flow than MEPS. Other bins of the measured hadronic angle, not shown in the figure, show similar trends.

In general, the difference between the hadronic energy-flow given by the two MC generators is enough to span the energy flow results found in the data. This observation was used to ascribe a systematic uncertainty to the measured neutral current cross sections, arising from the limited knowledge of the hadronisation process. To do this, as a first approximation, the cross sections can be extracted with both Monte Carlo models and their difference used as an estimate of the systematic error.

Figure 7.19 shows the hadronic energy-flow measured in the data, in bins of the reconstructed kinematic variable y_{DA}, compared to the energy flow found in Monte Carlo events. The hadronic energy-flow given by ARIADNE is represented by a solid line. A shaded band around this line corresponds to:

$$\overline{E}_j^{Ariadne} \pm \mid \overline{E}_j^{Ariadne} - \overline{E}_j^{MEPS} \mid \tag{7.8}$$

where $\overline{E}_j$ stands for the hadronic energy-flow, given by the two QCD-cascade models, found in polar angle bin j for each bin of the hadronic angle, γ_h.

A measured cross section, σ, can be assigned a systematic uncertainty, Δ_σ, associated with the model of the QCD cascade used in the unfolding, which is given by:

$$\sigma = \sigma_{Ariadne} \pm \Delta_\sigma \,, \text{ with } \Delta_\sigma = \mid \sigma_{\text{Ariadne}} - \sigma_{\text{MEPS}} \mid \tag{7.9}$$

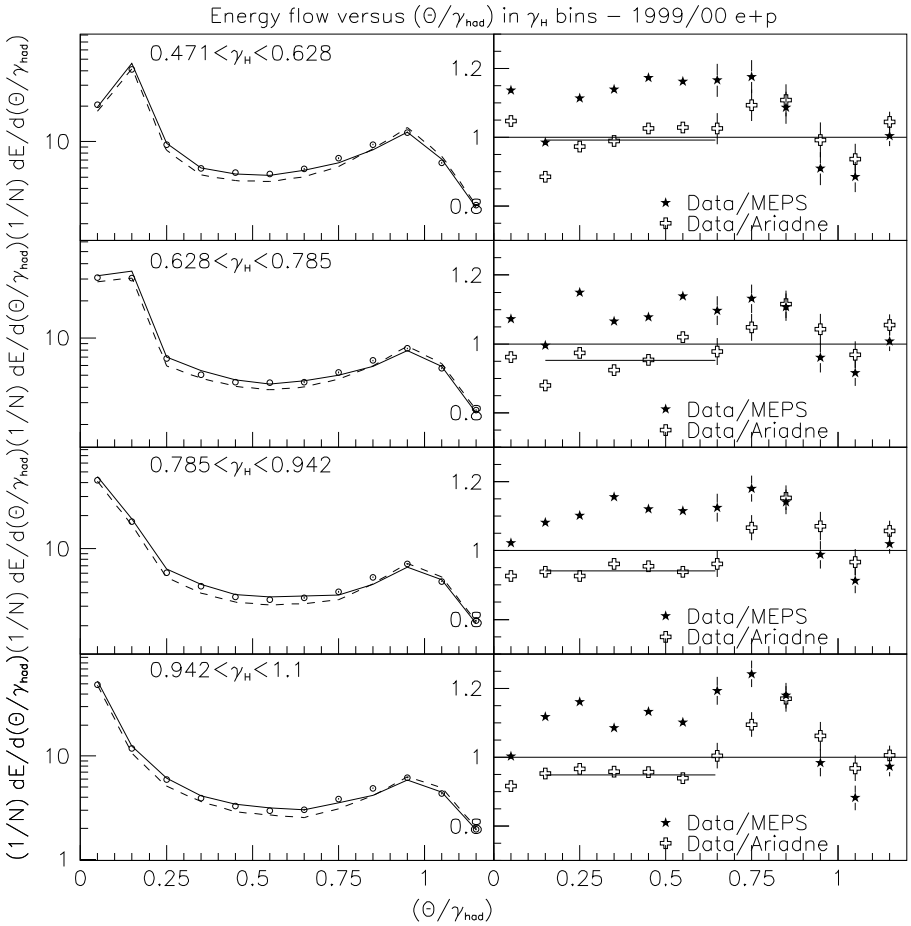

Figure 7.18: Energy flow in fired bins plotted versus θ/γ_h to highlight the difference between the current jet and transition regions. Plots in the right hand column show the ratio of energy flow in data and in ARIADNE MC. On the left hand column plots, ARIADNE is plotted as a solid line and LEPTO MEPS as a dashed line. The open circles correspond to the the energy flow measured in data.

where $\sigma_{Ariadne}$ and σ_{MEPS} are the cross section measured using MC samples generated with ARIADNE and MEPS, respectively, and σ can be any single or double differential cross section with respect to any kinematic variables.

As can be seen in the figure, in the majority of polar angle bins, the hadronic energy flow measured in data is contained within the band. This is indicated in each γ_h bin by the "accept fraction", which corresponds to the fraction of the 20 polar angle bins in which either the hadronic energy-flow measured in data or its statistical error are within the uncertainty band given by the MC models. The "accept fraction" is, then, a measure of the adequacy of the estimated systematic uncertainty.

While the method just described seems justified, it leads to an overestimation of the systematic uncertainty associated with the simulation of the hadronisation process. This can be seen from the values of the "accept fraction" noted in figure 7.19. These mean that, in between 85% and 100% of the 20 polar angle bins, the assigned uncertainty is

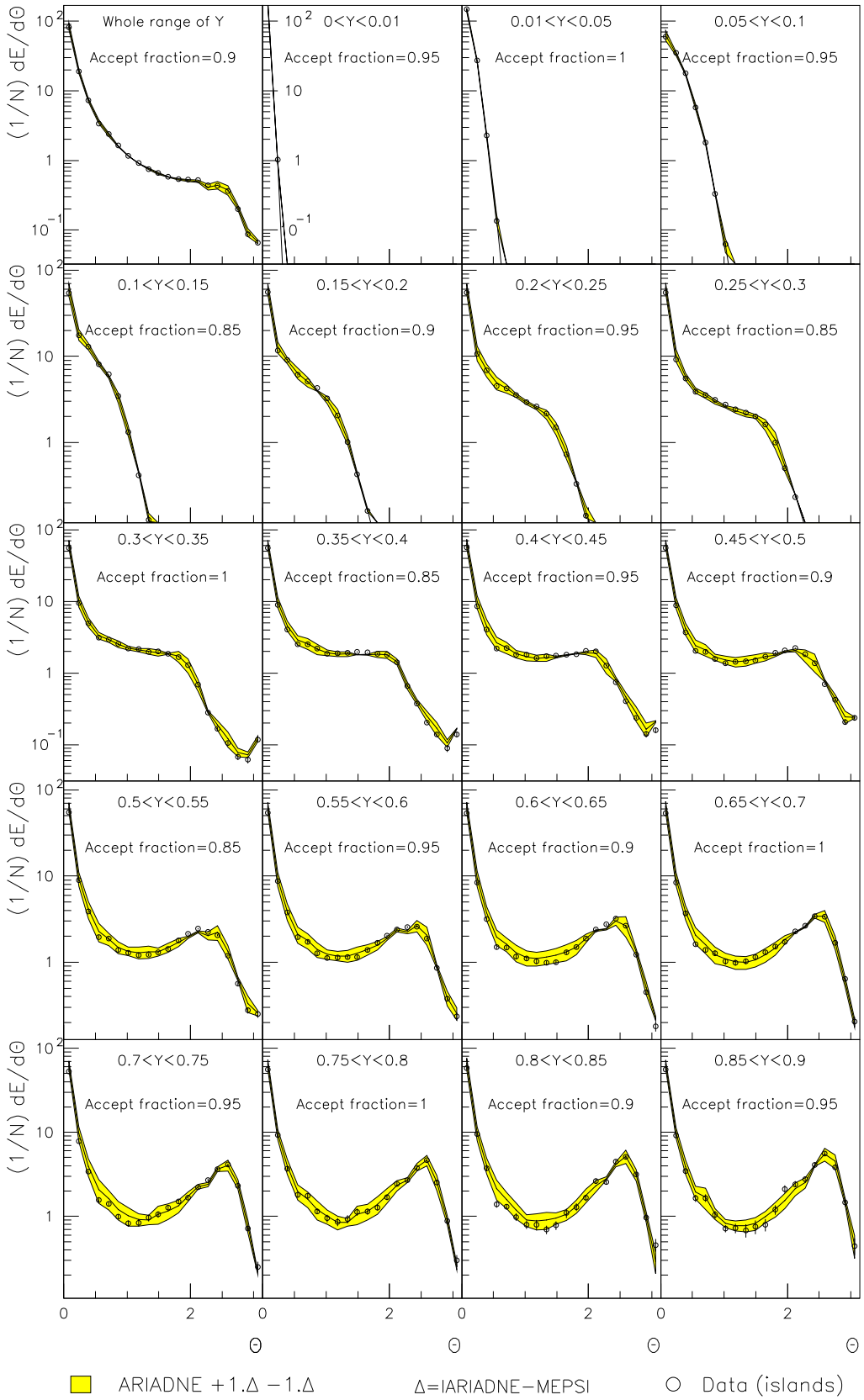

Figure 7.19: Comparison of energy flow in data (open dots) to ARIADNE MC in some bins of the hadronic angle. A region around the ARIADNE results (shadowed band) corresponds to the difference between ARIADNE and LEPTO MEPS points (i.e., $\overline{E}_{Ariadne} \pm \mid \overline{E}_{Ariadne} - \overline{E}_{MEPS} \mid$). For each γ_h bin, the fraction of data points that falls within the band was calculated and noted in the figure. To calculate this fraction, the first bin (which contains the FCAL beam hole) was ignored.

larger than the difference in hadronic energy-flow between real data and the ARIADNE MC sample. Also, the energy flow measured in real data is usually below the value given by ARIADNE.

As is noted above, ARIADNE yields a more accurate description of the energy flow, and is used in the extraction of the nominal values of the cross sections in the neutral current analysis. Also, as was seen in figure 7.18, ARIADNE usually overestimates the energy flow in the transition region between proton remnant and current jet. These observations should be used in a more realistic estimate of the systematic uncertainty associated with the QCD-cascade model. Such an estimate should include the deviation of the energy flow observed in ARIADNE from what is observed in data. The difference between the hadronic energy-flow observed in ARIADNE and MEPS events, on the other hand, provides a scale to the uncertainty which is a measure of the variations seen in the QCD-cascade models.

The differences in energy flow measured in bins of the reconstructed kinematic variable y_{DA}, between real data and ARIADNE and between ARIADNE and MEPS, is shown in figure 7.20 (left). The average differences can be used to define an intermediate uncertainty band of the hadronic energy-flow, which is asymmetric with respect to the results of the ARIADNE model, to account for the overestimation of the hadronic activity given by this model in the transition region. The new hadronic energy flow uncertainty band is given by $\mu \pm \Delta$, where μ, the new centre of the band, is equal to the mean of the $\overline{E}_{Ariadne} - \overline{E}_{data}$ distribution and is a measure of the systematic deviation of the simulation from the measured energy flow. The size of the uncertainty, Δ, is given by the mean of the $\overline{E}_{Ariadne} - \overline{E}_{MEPS}$ distribution and is a reflection of the differences between the phenomenological models of the hadronisation process. The average fractions of this uncertainty band F^+ and F^-, that lie respectively above and below the hadronic energy-flow obtained from ARIADNE, are given by:

$$F^+ = \frac{\mu - \overline{\Delta}}{2\,\overline{\Delta}} \text{ , and } F^- = \frac{\mu + \overline{\Delta}}{2\,\overline{\Delta}} \tag{7.10}$$

where $\overline{\Delta}$ is the mean of the $\overline{E}_{Ariadne} - \overline{E}_{MEPS}$ distribution (see figure 7.18 left). The values F^+ and F^- give a measure of the asymmetry of the uncertainty band with respect to the ARIADNE results. Slightly different values of F^+ and F^- were found, depending on which variable was used in the measurements. The values of $F^+ = 0.29$ and $F^- = 0.71$ were obtained when measuring the hadronic energy-flow residuals in bins of γ_h, while the values of $F^+ = 0.33$ and $F^- = 0.67$ were obtained with bins of y_{DA}.

The hadronic energy-flow measured in the majority (75% to 85%) of real data bins is within the limits of this uncertainty band (see figure 7.20, right), indicating an over-estimation of the uncertainty. This allows a further step to reduce the size of the error

band by using the average "accept fraction". The size of the band, $2\,\Delta$, is reduced until only 68% of the data hadronic energy-flow values are within the uncertainty band, while maintaining the asymmetry of the band with respect to the ARIADNE results. Although this is not a statistically sound argument but rather a rule of thumb, it still results in a realistic estimate of the systematic uncertainty. The particular value of 68.3% was chosen simply by analogy to the standard deviation of a normal distribution, even though the errors in the simulation of the hadronic energy-flow are not of a random nature. The determination of the fraction of $2\,\Delta$ to be used as the size of the final uncertainty band was done using measurements of the energy flow and its residuals in bins of Q^2, y and γ_h. To this end, a fraction F of the width of the uncertainty, $2\,\Delta$, was varied between 0.4 and 1. Then, the average fraction of the real data hadronic energy-flow values that fell within the uncertainty band was calculated. The result is shown in figure 7.20. A value of $F = 0.60$ was estimated from the figure by choosing the largest fraction returned by any of the three variables used: γ_h, y_{DA} and Q^2_{DA}.

A final hadronic energy-flow uncertainty band was defined as the interval $[\delta^-, \delta+]$, with the boundary values given by:

$$\delta^- = \overline{E}_{Ariadne} - F \times F^- \times \Delta\,, \text{ and } \delta^+ = \overline{\mathrm{E}}_{\mathrm{Ariadne}} + \mathrm{F} \times \mathrm{F}^+ \times \Delta, \tag{7.11}$$

where $\Delta = \overline{E}_{Ariadne} - \overline{E}_{MEPS}$. This uncertainty band is asymmetric with respect to ARIADNE and can be calculated in terms of ARIADNE and MEPS results. The values of $\delta^+ = 18\%$ and $\delta^- = 42\%$ were found and the uncertainty band was, finally:

$$\overline{E}^{Ariadne}\ {}^{+0.18\times(\overline{\mathrm{E}}_{\mathrm{Ariadne}}-\overline{\mathrm{E}}_{\mathrm{MEPS}})}_{-0.42\times(\overline{\mathrm{E}}_{\mathrm{Ariadne}}-\overline{\mathrm{E}}_{\mathrm{MEPS}})} \tag{7.12}$$

In this way, all the available knowledge obtained from the study of the hadronic energy flow can be combined to obtain a realistic estimate of the systematic uncertainties associated with the simulation of the hadronic final state. Figure 7.21 shows the hadronic energy-flow measured in real data as well as the uncertainty band associated with the MC description of the hadronisation process. As can be seen from the figure (see the fraction of data points within the band, noted in each plot), the uncertainty band determined as described above is a realistic estimate of the systematic differences between the Monte Carlo simulation and the data.

The algorithm was applied directly to the determination of the systematic errors in the reduced cross section unfolding. The "recipe" used to estimate the systematic errors associated with the simulation of the hadronisation process consists of: extracting the cross sections with both the default MC generator (ARIADNE) and with MEPS, and using +18% and -42% of the difference between the extracted cross sections to ascribe upper and

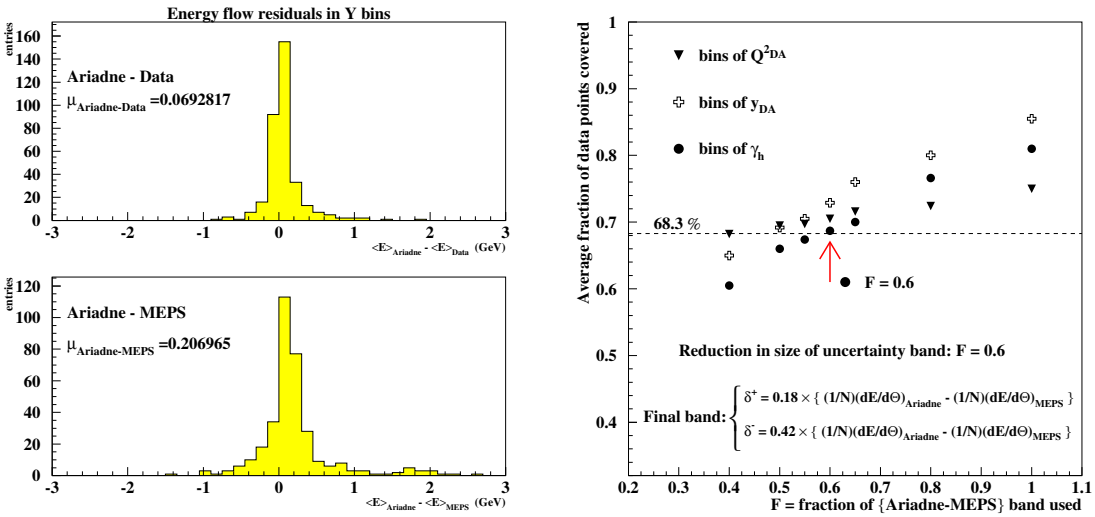

Figure 7.20: Determination of the best fraction of the difference between ARIADNE and MEPS simulated data that should be used to estimate the hadronisation systematic uncertainty. The mean differences between the hadronic energy-flow measured in data and MC events simulated with ARIADNE and MEPS (histograms on the left), are used to obtain the factors F^+ and F^-. These factors correspond to the fraction of the energy flow uncertainty band that should lie, respectively, above and below the central values given by ARIADNE. The size of the uncertainty band is reduced in such a way that it covers 68% of the energy flow values measure in real data (right plot).

lower estimates of the systematic uncertainty to the default measurement. This systematic uncertainty can be expressed, in an analogous way to equation 7.9, by:

$$\sigma = \sigma_{Ariadne}{}^{+\delta_\sigma^+}_{-\delta_\sigma^-} \tag{7.13}$$

$$\text{with} \quad \delta_\sigma^+ = 0.18 \times (\sigma_{Ariadne} - \sigma_{MEPS}) \tag{7.14}$$

$$\text{and} \quad \delta_\sigma^- = 0.42 \times (\sigma_{Ariadne} - \sigma_{MEPS}) \tag{7.15}$$

As in equation 7.9, $\sigma_{Ariadne}$ and σ_{MEPS} are the cross sections unfolded using MC events generated with ARIADNE and MEPS, respectively. It should be noted that in the improved method *Delta* is a signed value. This means that the largest uncertainty is assigned to the side of the cross section unfolded using LEPTO MEPS with respect to that unfolded using ARIADNE. Thus if, in a certain bin, σ_{MEPS} is larger than $\sigma_{Ariadne}$ then also $\delta_\sigma^+ > \delta_\sigma^-$, and vice versa.

The systematic errors found in double differential bins using this process are of the order of 2.5% or less for most bins, reaching around 10% in some high Q^2 bins.

Discussion

The process of determining the hadronisation systematic uncertainty, described above, is to some degree a compromise between a mathematically sound algorithm and a practical one.

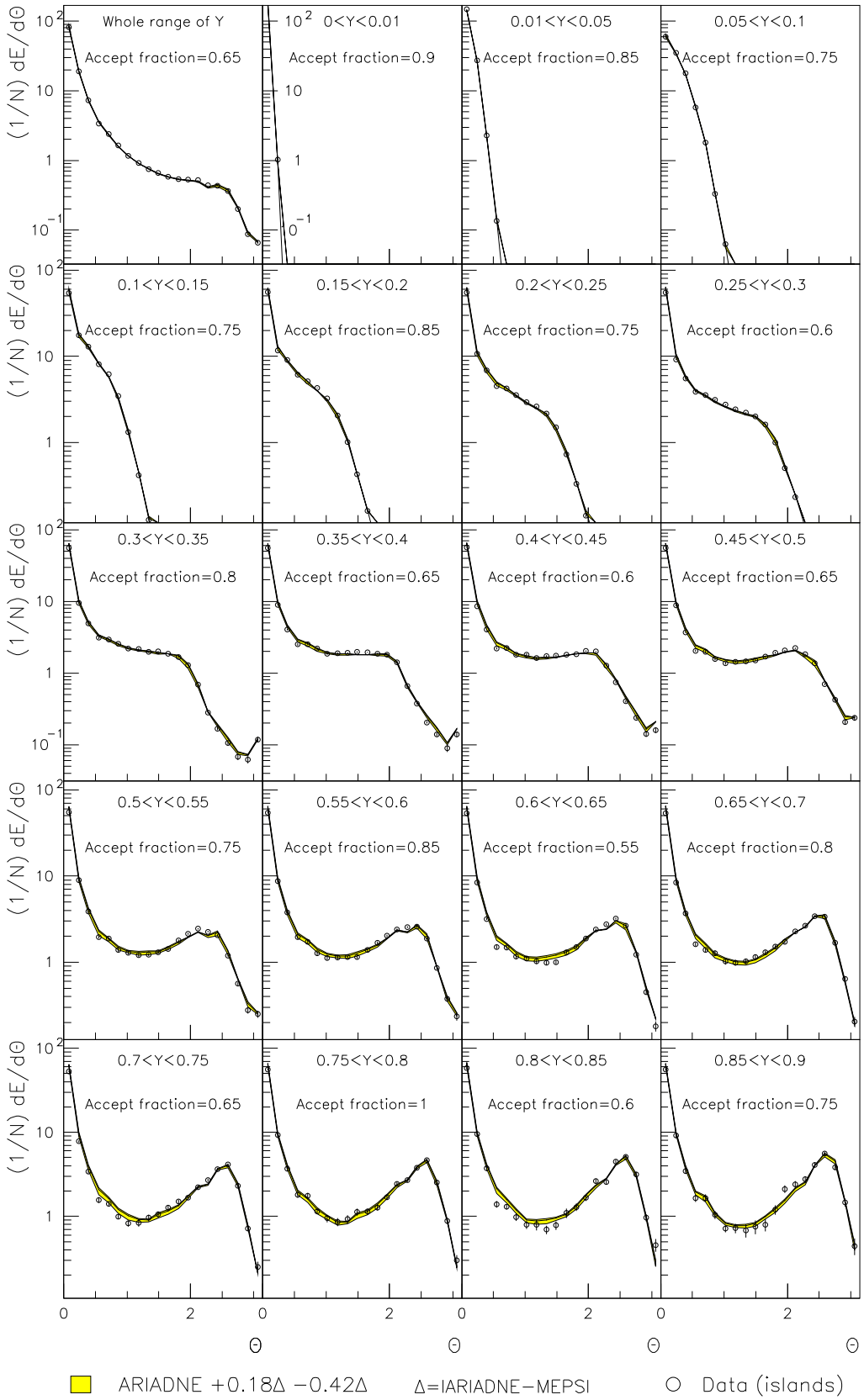

Figure 7.21: Analogous to figure 7.19 but plotted in bins of y_{DA}. Only part of the absolute difference between ARIADNE and MEPS is used to build the band around ARIADNE and calculate the fraction of data points inside the band.

The hadronic energy-flow provides a good way to examine the adequacy of the simulation. It reflects in a natural way the bias that may arise in the reconstructed hadronic angle due to the details of the QCD-cascade simulation. In an ideal situation, one would have a large set of different models available, from which to determine the variability of the simulated hadronic final state more accurately. With only two different models it would seem more reasonable to use the simple difference between the two cross section values obtained in some bin as the systematic uncertainty. However, from the study of the hadronic energy-flow, it becomes clear that QCD-cascade simulation in ARIADNE is closer to reality than the MEPS model. This knowledge should be used to obtain a realistic estimate of the systematic uncertainty, since the colour-dipole model, implemented in ARIADNE, is the default used in the analysis. The hadronic energy-flow found in the MEPS model serves as a scale against which the agreement between ARIADNE and real data can be measured.

The results obtained from the study of the hadronic final state are used directly in the determination of the systematic uncertainty in the cross sections. The justification for this relies on the observations that the hadronic energy-flow found in real data is, in general, in between that found in each of the two models, and in the close connection between the hadronic energy-flow and the hadronic angle, γ_h. Any bias in the QCD-cascade simulation will filter into the analysis through a bias in the reconstructed hadronic angle. Again, the difference in the cross sections obtained from the use of the two models sets the scale of the uncertainty.

Retrospectively, a different approach could have been followed which seems more straightforward and justifiable. The events simulated using the two different models could have been combined according to the results of the study of the hadronic energy-flow, in an analogous way to the combination of diffractive and non-diffractive MC samples described in section 4.1.3. Such a combination would yield an optimum description of the hadronic final state by the simulation. The systematic uncertainty could then have been determined by varying the weights of the two MC samples in the combination in some way dictated by the study of the hadronic energy-flow.

Chapter 8

Extraction of the cross sections

This chapter presents the results of the present analysis. Several quantities were measured, including single and double differential NC DIS cross sections as well as the structure functions F_2 and xF_3.

The method used for the unfolding of the NC DIS cross sections is discussed. The bins used in the measurement and the criteria for their selection are described in section 8.1.3, followed by a discussion of the sources of systematic uncertainty in section 8.1.4. The results of the cross section measurement are presented in section 8.2. Sections 8.3 and 8.4 describe the extraction of the structure functions F_2 and xF_3, respectively, and the results of the extraction are presented.

8.1 Unfolding of the cross sections

In a perfect detector, with full solid angle coverage, the measurement of a cross section σ would be performed simply by dividing the number of detected events, N, by the integrated luminosity of the accumulated data set, $\mathcal{L}$:

$$\sigma = \frac{N}{\mathcal{L}} \tag{8.1}$$

In a real experiment, however, the detector and trigger acceptances are finite, as is the detector resolution. Several types of effects can be identified that constitute a departure from ideal conditions. Background events, wrongly identified as NC DIS, can lead to an overestimation of the cross sections. The limited acceptance of detector and trigger, as well as the analysis cuts, contribute to the observation of only a subset of the scattering events taking place in the experiment. The finite resolution of the detector contributes to a smearing of the extracted kinematic variables, and therefore of the cross sections. Finally, systematic effects arising either from the detector of from processes such as initial or final state radiation (ISR/FSR), can lead to the migration of events between different regions

of phase space. The cross sections must then be estimated from the number of observed events taking into account the acceptance and smearing effects of detector, trigger and analysis cuts as realistically as possible. The process of estimating the true quantities from the observed data is known as unfolding [131, 132]. Due to the complexity of high energy physics experiments, the usage of simulated data in the unfolding procedure constitutes, in general, the only feasible way of correcting for these effects.

Considering the case of a probability density function $f(x)$ of a variable x that is measured using a detector with finite resolution, the probability $f'(x')$ of measuring a value x' of the variable is given by [132]:

$$f'(x') = \int_{-\infty}^{\infty} t(x, x') f(x) dx \tag{8.2}$$

where $t(x, x')$ is called the transfer function and includes both migration and smearing effects as well as the effects of finite acceptance of detector, trigger and the analysis cuts. In practice, the accessible phase space is divided into bins where events are accumulated. If $f(x)$ represents a cross section to be measured, the integral in equation 8.2 is transformed into a sum over the bins, according to:

$$d'_\mu = \sum_\nu T_{\mu\nu} d_\nu \quad \text{and} \quad m'_\mu = \sum_\nu T_{\mu\nu} m_\nu \tag{8.3}$$

where d_ν and m_ν are the true number of data events and the number of Monte Carlo events generated in a bin ν, and d'_μ and m'_μ are the number of data and MC events reconstructed in bin μ that pass the event selection. These numbers of real and simulated events correspond to the integral of the true and measured distributions, $f(x)$ and $f'(x')$, over each bin multiplied by the integrated luminosity. In the presence of a known background, the number of data events can be corrected by subtracting the estimated number of background events. The transfer function $t(x, x')$ is transformed into the transfer matrix $T_{\mu\nu}$, given by:

$$T_{\mu\nu} = \frac{\int_\mu \int_\nu t(x, x')\, f(x)\, dx'\, dx}{\int_\nu f(x)\, dx} \tag{8.4}$$

If the used bins are small, compared to the experimental resolution in the measured variable, the smearing effects of the detector lead to a number of events being generated in a certain bin and detected in a different bin. This creates large correlations between neighbouring bins and thus large off-diagonal terms in the transfer matrix. For large enough bins, on the other hand, most events generated in a certain bin will be reconstructed in that bin. The transfer matrix becomes essentially diagonal with small values for the off-diagonal elements. It should be noted, however, that large bins imply a loss of information on the shape of the cross sections being measured. The diagonal elements of the matrix are less than one, reflecting the finite acceptance of detector and trigger.

An estimate, $\hat{T}_{\mu\nu}$, of the transfer matrix can be obtained from Monte Carlo simulated events by replacing the integrals over density functions $f(x)$ and $f'(x')$ by the number of generated and reconstructed (and accepted) MC events. The transfer matrix is then given by:

$$\hat{T}_{\mu\nu} = \frac{m'_{\mu\nu}}{m_\nu} \tag{8.5}$$

In this equation, $m'_{\mu\nu}$ corresponds to the number of MC events generated in bin ν that were reconstructed in bin μ and passed the analysis cuts, while m_ν corresponds to the number of MC events generated in bin ν.

The estimated transfer matrix can, in principle, be inverted. This would allow an estimate of the true number of events occurring in the experiment according to the equation:

$$d_\nu = \sum_\mu \left(\hat{T}^{-1}\right)_{\nu\mu} d'_\mu \tag{8.6}$$

In this case, equation 8.1 would be used to extract the cross sections in each bin. The extracted cross section in each bin would then be correlated with the values found in other bins, the correlation being given by the non-diagonal elements of the transfer matrix. This method is mathematically precise, but difficulties arise if the simulation does not precisely describe the acceptance, migrations and resolutions. Apart from errors in the simulation, the inversion of the transfer matrix may yield nonsensical results due to statistical fluctuations in the observed data [131]. Also, the transfer matrix may not even be regular. Due to these difficulties, the matrix inversion method is not usually employed.

8.1.1 Bin-by-bin unfolding

A simpler method was used in the present analysis, which is equivalent to considering the transfer matrix to be diagonal. This is often referred to as bin-by-bin unfolding. In this method, only the net migration of events into and out of each bin is considered, and small correlations between bins are disregarded. A correction to the number of real events measured in each bin is obtained from the ratio of the number of measured MC events observed in the bin to the number of simulated events generated in the bin:

$$d_\mu = \hat{T}^{-1}_{\mu\mu} d'_\mu = \frac{m_\mu}{m'_\mu} d'_\mu \tag{8.7}$$

This correction thus accounts for the limited acceptance of detector, trigger and the effect of the analysis cuts.

Using equation 8.1 one now obtains the integrated cross section in the bin. To make the notation clear, it helps to consider a specific cross section. The double differential cross section, $d^2\sigma/dxdQ^2$ will be considered, where the bins will be defined in the $x - Q^2$

plane as the intervals $\text{bin}_\mu \equiv \Delta x \times \Delta Q^2$, but the expressions are readily applicable to other cross sections:

$$\sigma_\mu = \int_{\Delta x} \int_{\Delta Q^2} \frac{d^2\sigma}{dxdQ^2} dxdQ^2 = \frac{d'_\mu}{\mathcal{L}} \frac{m_\mu}{m'_\mu} \tag{8.8}$$

The measured cross section is to be quoted at a certain point inside the bin, the bin centre. To obtain this value, the integrated cross section in the bin must be multiplied by a factor $\mathcal{C}_{centre}$. This is the theoretical ratio of the (Born level) differential cross section at the bin centre to the integrated cross section in the bin. By using the Born-level cross section, one automatically obtains a measurement of the cross section corrected for higher order electroweak and radiative diagrams. To maintain consistency, this should be calculated with the same parton density functions (PDFs) as are used in the generation of the Monte Carlo simulated data. As is shown below, the bin-by-bin unfolding method automatically does this correction. For the double differential cross section, the bin centering correction factor is given by:

$$\mathcal{C}_{centre} = \frac{\left(\frac{d^2\sigma}{dxdQ^2}\right)^{Born}_{theory}}{\int_{\Delta x} \int_{\Delta Q^2} \left(\frac{d^2\sigma}{dxdQ^2}\right)^{Born}_{theory} dxdQ^2} \tag{8.9}$$

But the integral in the denominator is just the number of MC events generated in the bin, m_μ, divided by the integrated luminosity, $\mathcal{L}$, i.e.:

$$\mathcal{C}_{centre} = \frac{\left(\frac{d^2\sigma}{dxdQ^2}\right)^{Born}_{theory} \mathcal{L}}{m_\mu} \tag{8.10}$$

Using the above equation, the measured double differential cross section becomes:

$$\frac{d^2\sigma}{dxdQ^2} = \mathcal{C}_{centre} \frac{d'_\mu}{\mathcal{L}} \frac{m_\mu}{m'_\mu} = \frac{d'_\mu}{m'_\mu} \left(\frac{d^2\sigma}{dxdQ^2}\right)^{Born}_{theory} \tag{8.11}$$

The event selection cuts described in chapter 6 greatly reduce the number of background events contaminating the final sample. The only important remaining source of contamination comes from photoproduction events which are simulated and statistically subtracted from the selected sample of real events. This leads finally to the expression used to unfold the cross section:

$$\frac{d^2\sigma}{dxdQ^2} = \frac{d'_\mu - b'_\mu}{m'_\mu} \left(\frac{d^2\sigma}{dxdQ^2}\right)^{Born}_{theory} \tag{8.12}$$

where b'_μ is the number of background events estimated from the photoproduction Monte Carlo simulation and properly normalised.

The extraction of the cross sections by this method neglects correlations between bins. These have to be estimated in the determination of systematic errors by identifying which sources of uncertainty show a non-negligible correlation between neighbouring bins.

In principle, both the acceptance correction and the bin centering correction (equations 8.7 and 8.9, respectively), rely on a perfect description by the simulation of both the observed and the true distributions in real data. This can be circumvented by using an iterative unfolding procedure. In each iteration step, the measured cross sections are used to obtain new parton density functions or structure functions which are then used to obtain new simulated data, usually by reweighting the existing simulation. The reweighted simulated data is again used to extract new measured cross sections and the process is repeated until the measurement converges. Such a process was used in previous ZEUS measurements [27]. This typically resulted in two or three iterations to reach an agreement of $\sim 0.5\%$ between successive iterations. Early structure function measurements from HERA experiments contributed to an improved knowledge of the parton density functions in the HERA kinematic range. In later measurements, due to the improved agreement between the MC simulation and the observed data, the iterative process was found to be unnecessary [117, 33]. This was also true in the present analysis.

8.1.2 Statistical uncertainties

The observed number of events in each bin follow the Poisson distribution. One complication arises from the necessary weighting of MC events. Each simulated event is given a weight ω_i corresponding to the normalisation to the integrated luminosity, diffractive MC mixing, reweighting for the appropriate vertex distribution, or photoproduction background normalisation, etc. (cf. chapter 4). The number of signal and background MC events in a bin μ, with weights ω_μ^{DIS} and ω_μ^{PHP} respectively, can then be written as the sums, $m'_\mu = \sum \omega_i^{DIS}$ and $b'_\mu = \sum \omega_i^{PHP}$ and are in general not integer. The errors on the number of signal and background MC events are given by:

$$\delta m'_\mu = \sqrt{\sum \left(\omega_i^{DIS}\right)^2} \quad and \quad \delta b'_\mu = \sqrt{\sum \left(\omega_i^{PHP}\right)^2} \tag{8.13}$$

The number of real events observed in each bin, d'_μ are integers and their uncertainties are given by $\delta d'_\mu = \sqrt{d'_\mu}$. In bins with fewer than 12 observed events, the statistical uncertainty was given the Poisson 68.3% confidence interval limits assuming no background [133]. The problem of bins with very few events does not apply to MC data, as the simulated events greatly outnumber those observed. The relative statistical uncertainty of the unfolded cross section is obtained from equation 8.12, through error propagation:

$$\frac{\delta\sigma_\mu^{stat}}{\sigma_\mu} = \sqrt{\left(\frac{\sqrt{\left(\delta d'_\mu\right)^2 + \left(\delta b'_\mu\right)^2}}{d'_\mu - b'_\mu}\right)^2 + \left(\frac{\delta m'_\mu}{m'_\mu}\right)^2} \tag{8.14}$$

where σ_μ represents any cross section measurement in bin μ.

8.1.3 Bin selection

The first step in the unfolding process is to divide the accessible phase space into a set of bins in which the cross sections are to be measured. As becomes clear from the discussion of the unfolding method, it is important that migrations between neighboring bins are minimized. This can be achieved by choosing bins that are sufficiently large with respect to the resolution of the variable being measured. Figures 5.6 to 5.8 provide an overview of the resolution of kinematic variables measured with the double angle method in different regions of phase space. Large bins, on the other hand, correspond to some loss of information on the shape of the extracted cross sections. Another determining factor in choosing the size of the bins is that a reasonable number of events should be observed in the bin to minimise the statistical error of the measured cross sections. As the neutral current DIS cross section falls roughly with $1/Q^4$, this means that larger bins must be chosen at large-Q^2. Thus a balance must be reached between the measurement of the cross sections in many bins with the consequent large statistical and systematic uncertainties and bin-to-bin correlations and the measurement of few bins with smaller uncertainties but also with limited knowledge of the shape of the unfolded cross sections.

For ease of comparison with previous measurements, the same double differential bins in the x and Q^2 kinematic variables were used in the present analysis as were used in the extraction of the F_2 structure function from ZEUS data from the 1996/97 running period [33]. For the measurement of the single differential cross sections, $d\sigma/dQ^2$, $d\sigma/dx$, and $d\sigma/dy$, a similar selection of bins to those used in a previous measurement of high-Q^2 NC cross sections [117] was used. The size of the bins was found to be adequate in the present analysis and further criteria were applied to determine which bins were to be used. These are discussed in the next section.

Efficiency and Purity

A selection of bins were found to be suitable in the present analysis based on criteria of efficiency, $\mathcal{E}_\mu$, and purity, $\mathcal{P}_\mu$, that characterise each bin μ. These quantities were estimated from MC data according to:

$$\mathcal{E}_\mu = \frac{\#\text{events generated } \textit{and} \text{ accepted in bin } \mu}{\#\text{events generated in bin } \mu} \tag{8.15}$$

$$\mathcal{P}_\mu = \frac{\#\text{events generated } \textit{and} \text{ accepted in bin } \mu}{\#\text{events accepted in bin } \mu} \tag{8.16}$$

Simulated events (properly normalised and weighted) were counted as *generated* in a bin if their generated kinematic variables fell within bin boundaries and passed the cuts of $Q^2_{true} > 185$ GeV^2, $y_{true} < 0.95$, $y_{true} \times (1 - x_{true})^2 > 0.004$, and $E_{e,true} > 10$ GeV.

Events are counted as *accepted* in a bin if their *reconstructed* kinematic variables have values inside the bin boundaries and if the events pass all the trigger and analysis cuts.

Efficiency values below 50% indicate that more than half of the events generated in a bin, either did not pass the trigger or analysis selection, or migrated to other bins. This is, in principle, corrected for by the unfolding method as long as the MC simulation of the detector and trigger acceptances is accurate. Bins were used in the present analysis if their estimated efficiency was larger than 30%.

The bin purity constitutes an estimate of the fraction of events recorded in a bin for which the values of the true kinematic variables lie in that bin. Thus, purity values below 50% indicate that the bin is dominated by events generated elsewhere in the kinematic plane. Bins were used in the present analysis if their estimated purity was above 50%.

Values of efficiency and purity are plotted in figures 8.1 for single differential bins in Q^2. Both efficiency and purity values are quite high in most bins. The lower efficiency values at $Q^2 \approx 700$ GeV^2 are due to the supercrack cuts (cf. section 6.5) which are not taken into account in the determination of the number of generated events in the corresponding bins. The step-like increase of efficiency and purity at $Q^2 \approx 4000$ GeV^2 is caused by the increased bin size, which limits the effect of migrations.

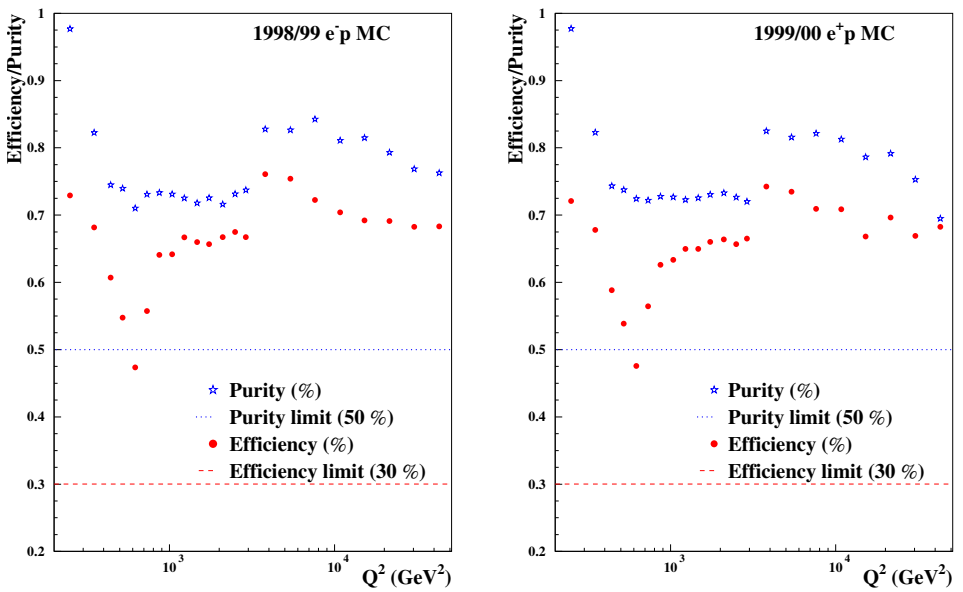

Figure 8.1: Efficiency and purity estimates for the bins used for the measurement of the single differential cross section $d\sigma/dQ^2$. The values represented here were obtained with simulated data corresponding to the 1999/00 e^+p data set.

Figure 8.2 shows the estimates of efficiency and purity in single differential bins x, and y. Purity values for the bins used in the cross section measurement are above 50%.

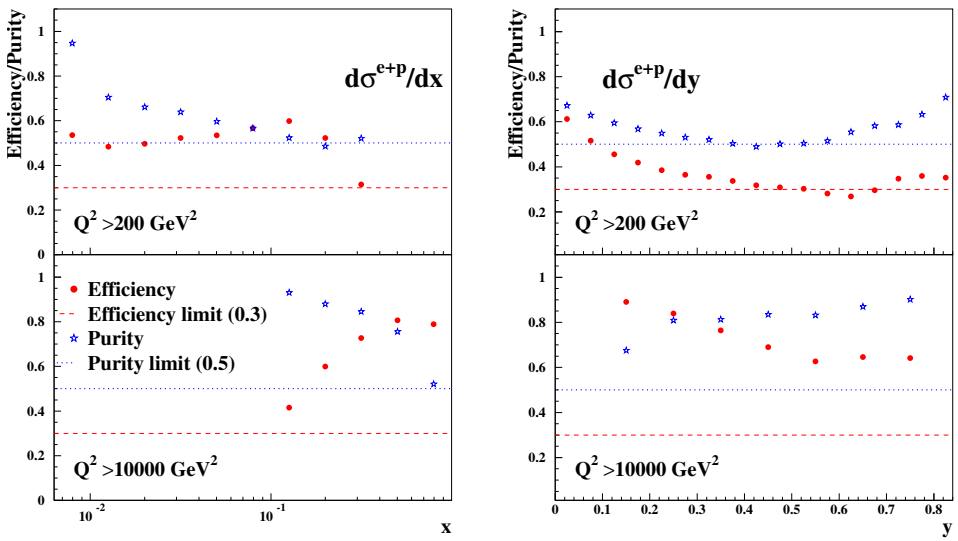

Figure 8.2: Efficiency and purity estimates for the bins used for the measurement of the single differential cross sections $d\sigma/dx$ (left column) and $d\sigma/dy$ (right column). The values represented represented in the figure were obtained with simulated data corresponding to the 1999/00 e^+p data set.

Efficiency values are generally above 30% except for one bin at high y, with efficiency of 28%, which was still used in the analysis.

Figure 8.3 shows the estimated efficiency and purity in double differential bins in x and Q^2. These were obtained from simulated data samples corresponding to the 1999/00 running period. Similar values were obtained for the 1998/99 running period.

Migration

As discussed above, the true values of the kinematic variables of recorded events can be significantly different from the estimated variables. Detector effects, such as the backsplash of hadronic final state particles or scattering of particles in inactive material, can bias the measurement of the kinematic variables. Also certain classes of events, such as QED Compton scattering or events in which the initial or final state electron radiates a photon (ISR/FSR), can severely bias the estimate of x and Q^2. Figure 8.4 illustrates the migration of events in the kinematic plane. The arrows in the figure are drawn from the average of the true x and Q^2 variables of the events generated in each bin to the average position where these events were reconstructed. For radiative events, the calculation of the true x and Q^2 must take into account the lower centre of mass energy of the hard scattering in ISR events and the emission of a final state photon in FSR events. In the present case, the true event quantities were calculated at the electron vertex by adding the four momentum of the final state photon to that of the scattered electron in FSR events.

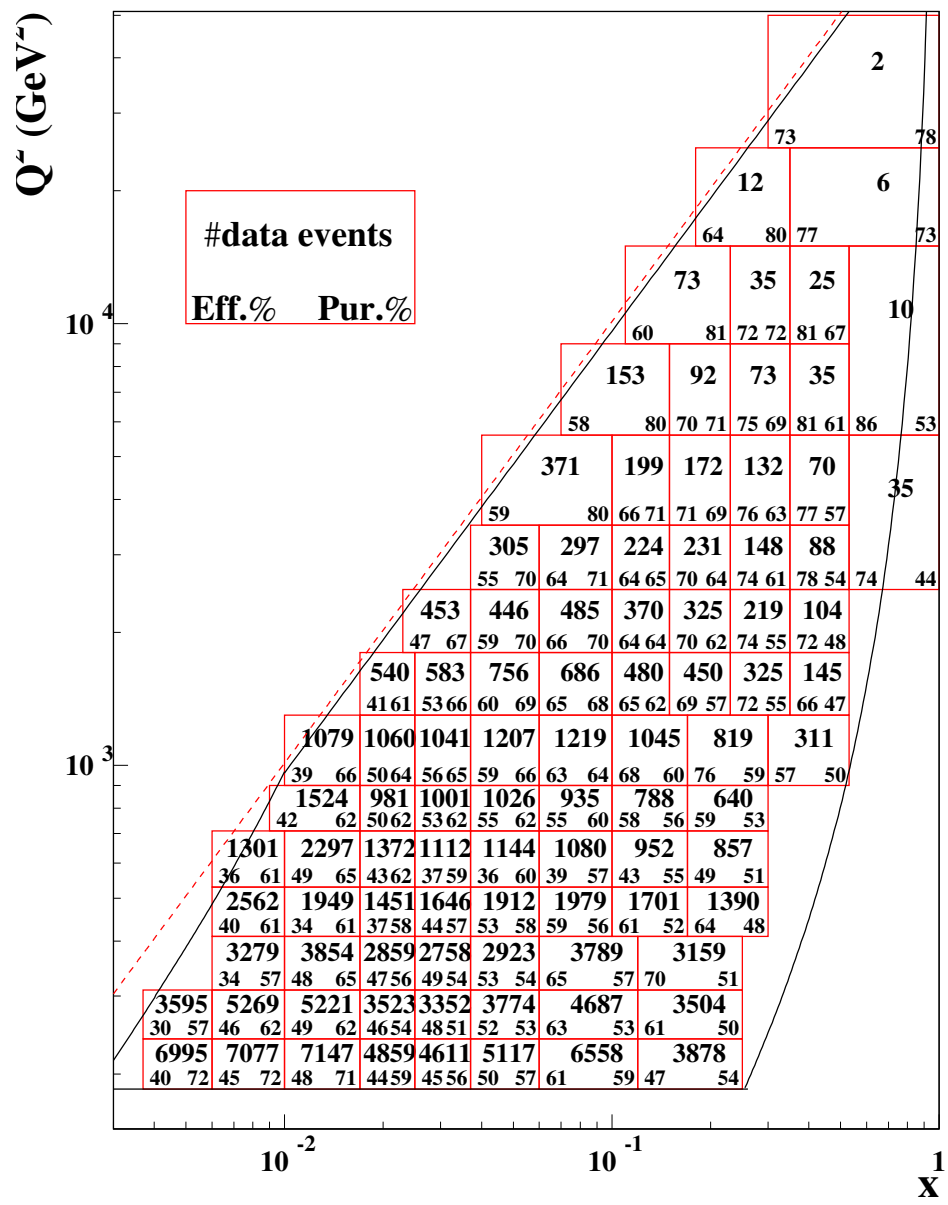

Figure 8.3: Efficiency and purity estimates for the bins used in the measurement of the double differential cross section $d\sigma/dxdQ^2$.

The migrations shown by the arrows in figure 8.4 are seen to be small. This is a further indication that the bins have adequate size compared to the resolution of the reconstruction methods and that small correlations should be expected between neighbouring bins

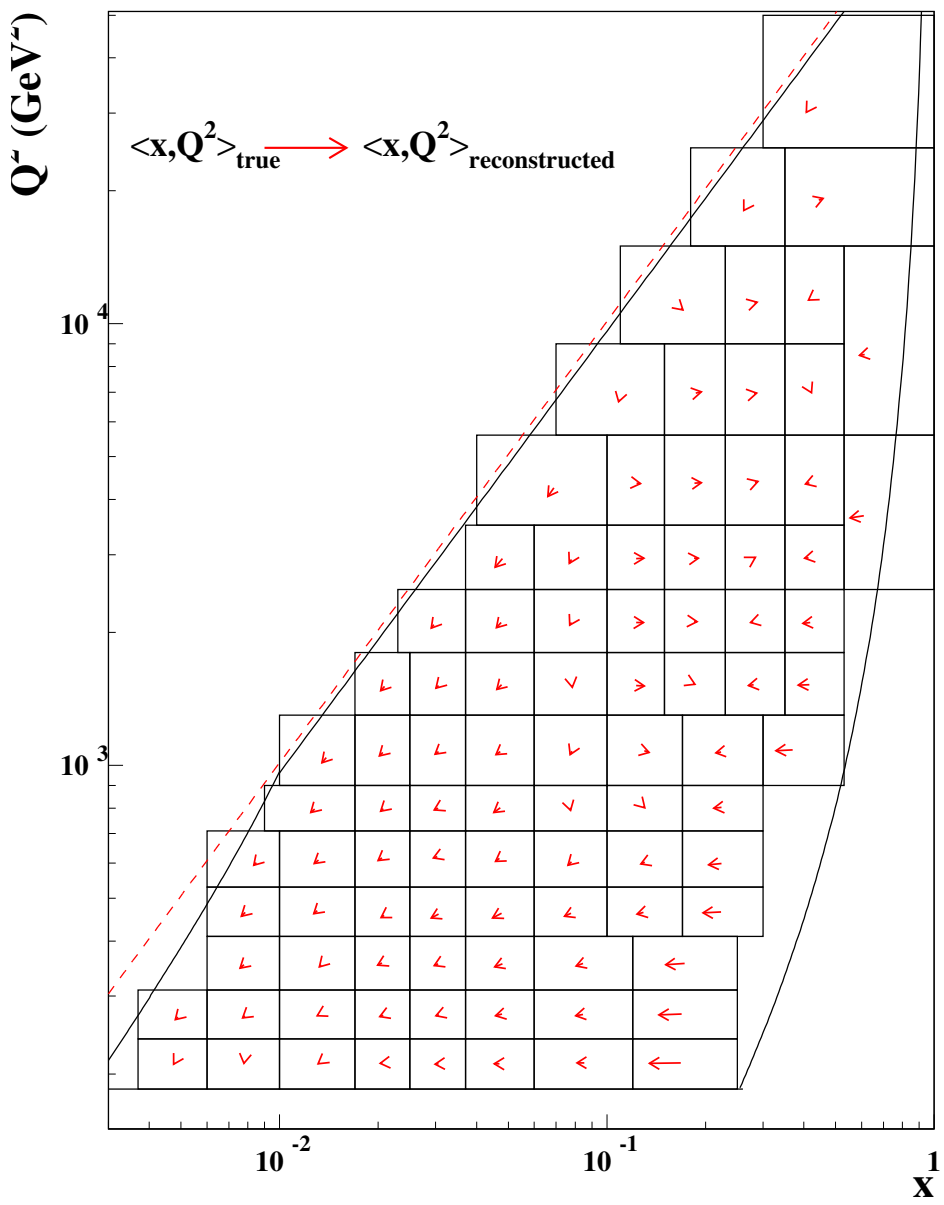

Figure 8.4: Migration of events in phase space. The arrows in the figure are drawn from the average x and Q^2 of the events gnerated in each bin to the average position where these events were reconstructed.

8.1.4 Systematic uncertainties

The evaluation of systematic uncertainties of the measured cross sections and structure functions is essential to the understanding of the measurements. The large data sets available since 1996 have contributed to a very good understanding of the ZEUS detector and its systematic effects. This has contributed to the refinement of the correction methods

employed. As a result, the latest ZEUS measurement of the F_2 structure function has statistical and systematic uncertainties below 2% in most of the accessible phase space [33]. The total uncertainty of the measured F_2 was dominated by systematic errors in many bins, up to $Q^2 \approx 800$ GeV^2 [46].

A realistic estimation of the systematic uncertainties associated with the measured cross sections and structure functions is essential to the measurements. Systematic uncertainties reflect the limit to which a measurement can be made using the available tools. These include both the detector and analysis tools such as the MC simulation and corrections to the measured quantities. Unlike statistical errors, which can be determined in a general way, the evaluation of systematic uncertainties relies sometimes on a subjective understanding of the possible error sources. Uncertainties and biases may be caused by detector effects or their inadequate parametrisation in the MC simulation, or by the physics inputs in the MC simulation, such as the parton density functions. Once the most important sources of systematic errors have been identified, they must be quantified and their values combined in a total uncertainty. This step is complicated by the fact that systematic errors are not necessarily well behaved functions of some variable, and are often correlated in unknown ways.

Sources of systematic uncertainty

Several sources of systematic uncertainties have been identified and are described below. The effect of each contribution i was evaluated by varying a corresponding analysis cut or reconstruction parameter, such as the calorimeter energy scale. The cross sections are then re-evaluated. The relative difference between the re-evaluated cross sections, σ_i, and the nominal value, σ, was taken as the relative systematic uncertainty contribution from that error source, noted here as δ_i. The relative uncertainties are combined in each bin, separately for contributions which increase the measured cross section and the ones that result in a smaller measured cross section. The resulting total systematic uncertainty for each bin, δ_{syst}, may be asymmetric with respect to the nominal cross sections and is given by:

$$\frac{\delta^+_{syst.}}{\sigma} = \sqrt{\sum_i \delta_i^2} = \sqrt{\sum_i \left(\frac{\sigma_i - \sigma}{\sigma}\right)^2} \quad , \sigma_i - \sigma > 0 \tag{8.17}$$

$$\frac{\delta^-_{syst.}}{\sigma} = \sqrt{\sum_i \delta_i^2} = \sqrt{\sum_i \left(\frac{\sigma - \sigma_i}{\sigma}\right)^2} \quad , \sigma_i - \sigma < 0 \tag{8.18}$$

The contributions to the systematic uncertainty can be divided into cut variations and reconstruction parameter variations. In each of the former, the analysis cut on a

certain variable is varied according to the resolution of the measured variable and the cross section is re-evaluated. The cuts are varied for both real and simulated data, as this type of variation aims at testing the accuracy of the simulation near the cut values. In the variation of reconstruction parameters, constants such as the calorimeter energy scale or the photoproduction background normalisation are varied and the event quantities re-measured. As an example, figures 8.5 and 8.6 show the estimated relative systematic uncertainties from all the sources considered. The complete set of relative systematic uncertainties on measured quantities is shown in appendix A.

In total, 18 sources of uncertainty were identified. Eleven of these were found either to be small or to show negligible bin-to-bin correlation. These sources of error were considered uncorrelated and are denoted by δ_i^u below. The remaining seven sources of systematic error were considered bin-to-bin correlated and are denoted δ_i^c. Bin-to-bin correlated systematic error sources create regular patterns of the estimated uncertainty in neighbouring bins, such as can be seen in figure 8.6. Some degree of correlation between neighbouring bins is visible in some error sources shown in figure 8.5. These were neglected if the size of the uncertainty was small. In those cases, the error source was classified as uncorrelated.

- δ_1^u – **variation of the $(\mathbf{E - p_Z})_{\mathbf{total}}$ cut**: the upper cut on the measured $(E-p_Z)_{total}$ is varied by ± 2 GeV^2. This cut aims at the elimination of overlaid events. Varying the cut simultaneously changes the level of background contamination and tests the Monte Carlo description of the data (see figure 6.22).

- δ_2^u – **variation of the track matching cut**: the distance of closest approach between the electron positions given by the calorimeter and by the reconstructed electron track is tightened from 10 cm to 8 cm.

- δ_3^u – **variation of the $\mathbf{y_e}$ cut**: this cut is tightened from $y_e < 0.95$ to $y_e < 0.9$. Photoproduction background contamination is most significant at high y (see figure 6.22). This cut variation, as δ_2, probes the amount of background that is allowed into the selected sample. In bins with few events, at high Q^2, the effects of this cut variation can overestimate the systematic error. This is caused by the statistical effect of a few events being removed from a sparsely populated bin. To obtain a better estimate of the systematic error, the number of events in real and simulated data with $Q^2_{DA} > 3200$ GeV^2 for single differential cross sections and $Q^2_{DA} > 3500$ GeV^2 for double differential cross sections were counted. The relative systematic error was estimated from these numbers of events to be -1% for the single differential and -0.3% for the double differential cross sections.

- δ_4^u – **variation of the $p_T/\sqrt{E_T} < 4$ GeV$^{1/2}$ cut**: the cut on the transverse momentum balance is varied by ± 1 GeV^2 (see figure 6.23). This cut variation influences the amount of background from cosmic rays that is allowed into the sample. This systematic uncertainty is generally very small, except in a few bins with very few events, where the effect of one event being included or eliminated leads to large deviations.

- δ_5^u – **variation of the $R_{\gamma_h}^{FCAL}$ cut**: the radius of the hadronic angle projection on the face of the FCAL is varied by ± 3 cm

- δ_6^u – **electron angle resolution**: θ_e is varied according to its resolution. The variation is ± 3 mrad if the angle of the scattered electron is determined from a matched track and, if not, by ± 5 mrad (± 2 mrad) for electrons detected in the RCAL (FCAL).

- δ_7^u – **hadronic energy scale**: the relative uncertainty on the hadronic energy scale was found to be 2% for the RCAL and 1% for the FCAL and BCAL [122]. To estimate the uncertainty due to the hadronic energy scale, the energy measured in each calorimeter section was rescaled both in a correlated and in an anti-correlated way to the other two sections and, for each section, in an anti-correlated way for the EMC and HAC parts of the calorimeter. This resulted in 10 different combinations of rescaled energy values. For each combination, the cross sections were re-evaluated. The largest variation of the cross sections in each bin was taken as the systematic uncertainty associated with the hadronic energy scale. The hadronic energy scale uncertainty in the single differential cross section $d\sigma/dQ^2$ is shown in figure 8.5. The apparent bin-to-bin correlation from this source of uncertainty is an effect of the representation (the maximum positive and negative values of the uncertainty obtained from the 10 separate variations of the hadronic energy scale are represented by different symbols).

- δ_8^u – **electron energy resolution**: the electron energy resolution obtained in Monte Carlo events is better than that observed in real data. To account for this difference, the energy reconstructed in MC events is smeared as described in section 6.4.3. To account for the systematic uncertainty from the different resolutions, the smearing was turned off.

- δ_9^u – **diffractive contribution**: the fraction of simulated diffractive NC DIS events that is mixed with non-diffractive MC events was varied by $\pm 50\%$.

- δ^u_{10} – **variation of γ_{max}**: the parametrisation of γ_{max} used in the correction for the backscattering of in the calorimeter (see section 6.4.4) was changed in a range where the reconstruction of the hadronic angle, γ_h, stays close to optimal.

- δ^u_{11} – **Island → cells splitting**: as described in section 7.2, the MC description of the energy flow near the proton remnant is not perfect. In addition to a cut on the radius of the projection of γ_h on the FCAL face, the use of calorimeter cells instead of the reconstructed islands near the FCAL beamhole was used as a further estimate of the systematic error associated with the simulation of the hadronisation process.

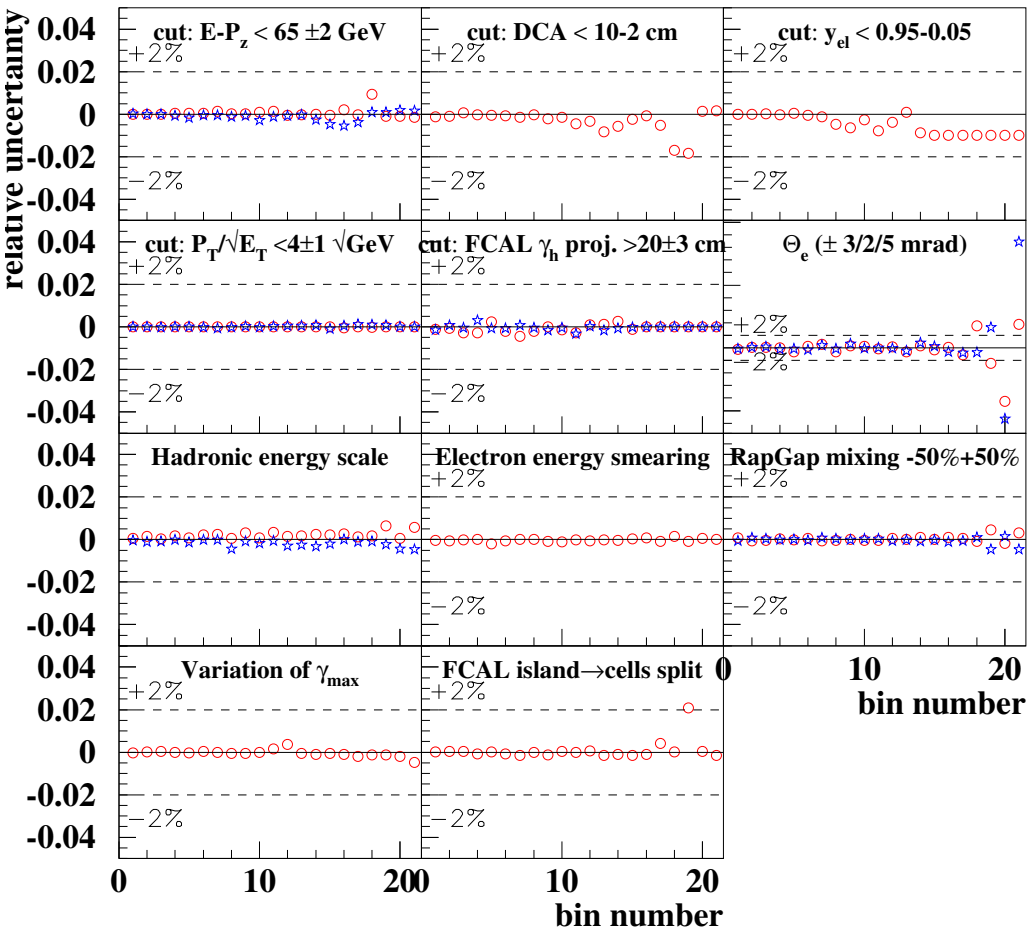

Figure 8.5: Bin-to-bin uncorrelated systematic uncertainties δ^u_1 to δ^u_{11} in the measured single differential cross section $d\sigma/dQ^2$. The systematic uncertainty arising from the $y_e < 0.95$ cut is averaged over several high-Q^2 bins to avoid statistical effects (see text). The apparent bin-to-bin correlation in the hadronic energy scale systematic uncertainty is an effect of the representation (see text).

The unfolding method used in the present analysis gives no information on the correlations existing between bins. Bin-to-bin correlations in the systematic uncertainties must therefore be estimated from the behaviour observed in adjacent bins for the various error

sources. The sources of error that were found to be significant and to exhibit bin-to-bin correlations are listed below. The remaining sources were considered as uncorrelated. Figure 8.7 shows the relative systematic uncertainty of the $d\sigma/dQ^2$ relative error cross section compared to the uncorrelated and to each correlated source.

- δ_1^c – **reweighting to ZEUS NLO QCD fit**: the unfolded cross sections should not depend on the input PDFs to the MC simulation. As an estimate of the effect of using a certain PDF set in the simulation, the MC data was reweighted to the PDFs obtained from a next-to-leading-order QCD fit to previous ZEUS data. The ZEUS-S fit [24] included the measurement of F_2 from 1996/97 data [33] and data from fixed target experiments, and did not include any ZEUS data from 1998 to 2000. Cross sections were then recalculated and the relative difference to the nominal values measured using the CTEQ5D PDF set [25] was taken as one estimate of the systematic error.

- δ_2^c – **variation of the electron energy scale**: the relative uncertainty on the scale of the measured electron energy was found to be 2% at an energy of 8 GeV, decreasing linearly to 1% at energies of 15 GeV and above [33]. The measured electron energy was varied by this amount.

- δ_3^c – **photoproduction normalisation**: the normalisation of the simulated photoproduction background is somewhat uncertain as is described in section 4.2.1. This normalisation was varied by -50% and +100%.

- δ_4^c – **variation of the isolation cut**: the cut on the energy measured in the calorimeter in a cone around the electron position and not attributed to the scattered electron is varied by ± 2 GeV^2.

- δ_5^c – **variation of the $\mathbf{p_e^{trk}}$ cut**: the cut on the electron momentum measured from the identified track is raised to 10 GeV and lowered to 0 GeV. Again, this cut variation tests the MC description of the data and also the background contamination, due to low momentum hadronic particles which are misidentified as the scattered electron (see figure 6.20).

- δ_6^c – **longitudinal vertex-distribution**: the Monte Carlo events were reweighted as described in section 7.1.

- δ_7^c – **simulation of the hadronisation**: the cross sections were unfolded using a MC sample generated with the Lepto MEPS hadronisation model. Relative systematic uncertainties were determined acording to section 7.2.4. The correlation between neighbouring bins, not apparent in figure 8.6, is very noticeable in the $d\sigma/dx$ cross section (not shown).

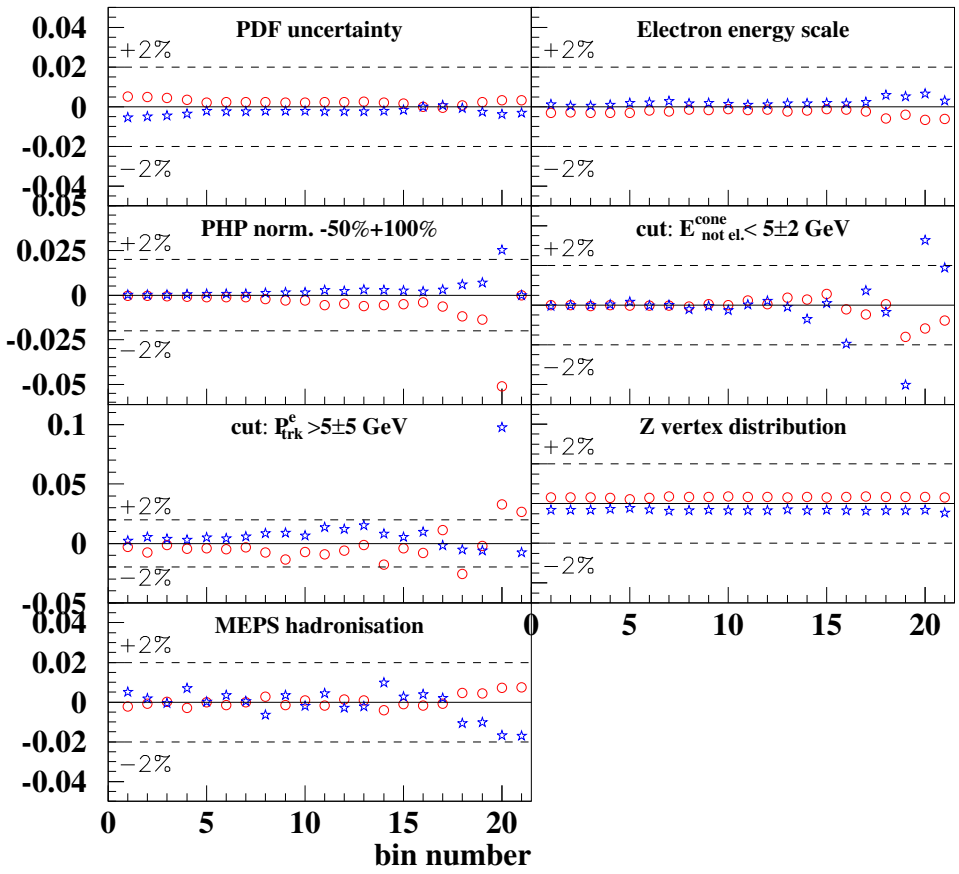

Figure 8.6: Bin-to-bin correlated systematic errors δ_1^c to δ_9^c of the single differential cross section $d\sigma/dQ^2$.

In addition to the sources described above, the uncertainty in the measurement of the integrated luminosity contributed with relative normalisation uncertainties of 1.8% and 2.25% for the 1998/99 e^-p and the 1999/00 e^+p, respectively. These uncertainties are not included in the calculated systematic errors displayed in the figures, except in the case of the xF_3 structure function, as its measurement requires the use of both the e^-p and e^+p data sets.

Note on the PDF uncertainty

The measured cross sections and structure functions must not depend on the set of parton density functions that is input into the Monte Carlo generation programs. The same parton distributions are used to generate the Monte Carlo data and to calculate the theoretical cross section value that is used in the unfolding. This should cancel the possible differences that may arise from the use of different PDFs in the MC simulation. Problems may arise in this picture, though, if the experimental acceptance is not constant in a bin and in the case of migrations across the bin boundaries.

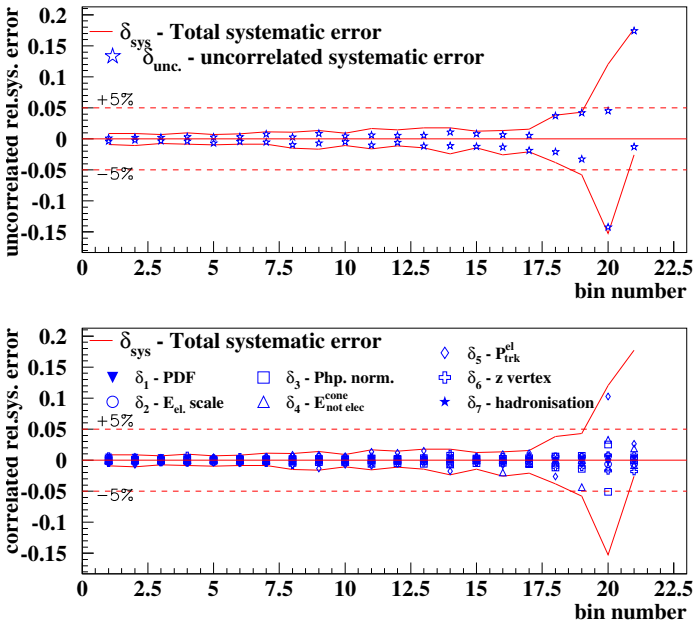

Figure 8.7: Total systematic uncertainty in the measured single differential cross section $d\sigma/dQ^2$ showing separately the total uncorrelated systematic uncertainty (top) and as the estimated uncertainties from individual bin-to-bin correlated sources (bottom).

The cross section is influenced by the PDFs through the structure functions. Both the shape and the integral of the cross section inside some bin will depend on the PDF set used in the calculation. If the experimental acceptance can be considered constant inside a bin, then the acceptance correction, m/m' (see equation 8.11), correctly accounts for different integrals of the cross section in the bin. On the other hand, if the experimental acceptance varies inside the bin, then its effect on the number of MC events counted in the bin depends on the shape of the cross section.

This is illustrated by figure 8.8. Consider two cross sections $\sigma_1(x)$ and $\sigma_2(x)$, calculated from two different PDF sets 1 and 2. The observed cross sections are $\sigma_1'(x) = \sigma_1(x)\mathcal{A}(x)$ and $\sigma_2'(x) = \sigma_2(x)\mathcal{A}(x)$. For simplicity, the bin centre will be considered to be at point $\mathcal{C}$, where $\sigma_1(x) = \sigma_2(x)$. The ratios of the cross sections to the central value for the bin can be different if the experimental acceptance, $\mathcal{A}(x)$, is not constant inside the bin. In the case illustrated in the figure, the observed cross section $\sigma_1'(x)$ is larger than the integral of $\sigma_2'(x)$ even though the integral of $\sigma_1(x)$ is smaller than that of $\sigma_2(x)$.

Thus, the interplay between the shape of the calculated cross section and the unknown experimental acceptance is a source of systematic uncertainty which must be estimated. It should be noted that this systematic uncertainty does not arise due to the different numbers of MC events generated in some region of phase space through the use of each

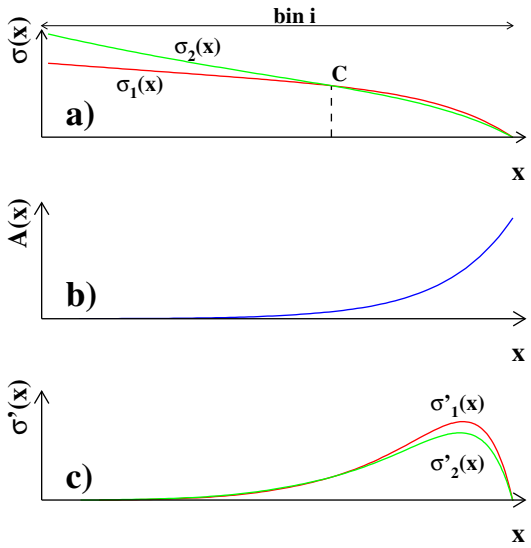

Figure 8.8: a): schematic representation of cross sections, $\sigma_1(x)$ and $\sigma_2(x)$, in some bin i, obtained from different sets of parton density functions. The experimental acceptance, $\mathcal{A}$, may vary inside a bin (**b**)). The observed cross sections, $\sigma_1'(x)$ and $\sigma_2'(x)$, are affected by the varying experimental acceptance (**c**)). **C** corresponds to the point where the cross section is to be unfolded (bin centre).

possible set of PDFs. It should only account for the combined effect of the shapes of the cross sections and the experimental acceptance.

To estimate this uncertainty, the Monte Carlo generation may be re-tuned through the use of different input PDFs, and the cross sections measured using the new MC data. To avoid re-generating large amounts of simulated data, the sample generated with CTEQ5D was reweighted using the ratios of cross sections obtained from different PDFs. The systematic uncertainty was obtained from the difference between the cross sections σ_i', measured using the MC events reweighted to the PDF set i, and the nominal measured cross section values, σ_0', unfolded using MC events generated with the CTEQ5D set. The unfolding of the systematically modified cross sections, σ_i', was done using the cross section at the bin centre, σ_i^{centre}, calculated from PDF set i. This leads to the estimated relative systematic uncertainties, δ_i:

$$\delta_i = \frac{\sigma_0' - \sigma_i'}{\sigma_0'} = \frac{\frac{d'}{m_0'}\sigma_0^{centre} - \frac{d'}{m_i'}\sigma_i^{centre}}{\frac{d'}{m_0'}\sigma_0^{centre}} = \tag{8.19}$$

$$1 - \frac{m_0'}{m_i'}\frac{\sigma_i^{centre}}{\sigma_0^{centre}} = 1 - \frac{\int_{bin}\mathcal{A}(x)\sigma_0(x)/\sigma_0^{centre}dx}{\int_{bin}\mathcal{A}(x)\sigma_i(x)/\sigma_i^{centre}dx} =$$

$$\frac{\int_{bin}\mathcal{A}(x)\sigma_0(x)/\sigma_0^{centre}dx - \int_{bin}\mathcal{A}(x)\sigma_i(x)/\sigma_i^{centre}dx}{\int_{bin}\mathcal{A}(x)\sigma_i(x)/\sigma_i^{centre}dx}$$

which contain information from both the cross sections calculated with each PDF set, σ_i ($i = 0, 1..$), and the experimental acceptance $\mathcal{A}$. In the special case where the acceptance

$\mathcal{A}(x)$ is constant throughout the bin, this reduces to the the relative difference of the bin centre corrections obtained from each set of PDFs. This reflects the current level of understanding of the proton structure, as similar PDFs will lead to small uncertainties.

In the present analysis, the ZEUS-S set of PDFs was available with its own uncertainties. These are a reflection not only of the errors in the data sets used in the fit, but also from uncertainties related to the input parameters and the functional form of the PDFs [24]. These uncertainties provided a convenient way to vary the input PDFs in the simulated data that would correspond to realistic uncertainties. To this end, the MC data were first reweighted to the cross sections obtained from the nominal ZEUS-S fit results and then to the cross sections obtained from the PDFs corresponding to the total (positive and negative) uncertainty from the fit. In the notation of equation 8.20, σ_0 corresponded to the ZEUS-S nominal values and σ_i to the fit uncertainties. The resulting relative uncertainties were applied to the nominal cross sections extracted with CTEQ5D cross sections.

8.2 Cross section results

This section presents the results obtained in the measurement of cross sections.

8.2.1 Single differential cross section $\mathbf{d\sigma/dQ^2}$

The single differential cross section $\frac{d\sigma}{dQ^2}$ is obtained by integrating the expression for $\frac{d^2\sigma}{dxdQ^2}$ (equation 2.10) over x. This leaves essentially a $1/Q^4$ dependence of the cross section with additional terms corresponding to the weak neutral current. In practice, the integration is performed implicitly in the measurement when events are accumulated in bins of Q^2. Some cuts imposed in the event selection, such as the MC validity cut, limit the integration boundaries (see figure 6.25).

Figure 8.9 shows the measured cross section in the interval 185 GeV^2 < Q2 < 36200 GeV^2 for both the e^+p and e^-p data samples. The measured values are given in tables A.1 and A.6. The separation between the two cross sections shown in the figure is due to the violation of parity by the weak neutral current, which is parametrised by the xF_3 term of the cross section. Thus, the e^-p cross section is enhanced by the positive contribution of the xF_3 term, while this contribution enters with a negative sign in the e^+p cross section. The cross section falls by six orders of magnitude in the observed interval of Q^2. The highest-Q^2 events found in each sample had measured values of 37000 GeV^2 in the e^-p sample and 32000 GeV^2 in the e^+p sample.

The measured cross sections are well described by the Standard Model prediction, as can be seen by comparing the curves shown in figure 8.9. Several curves are shown

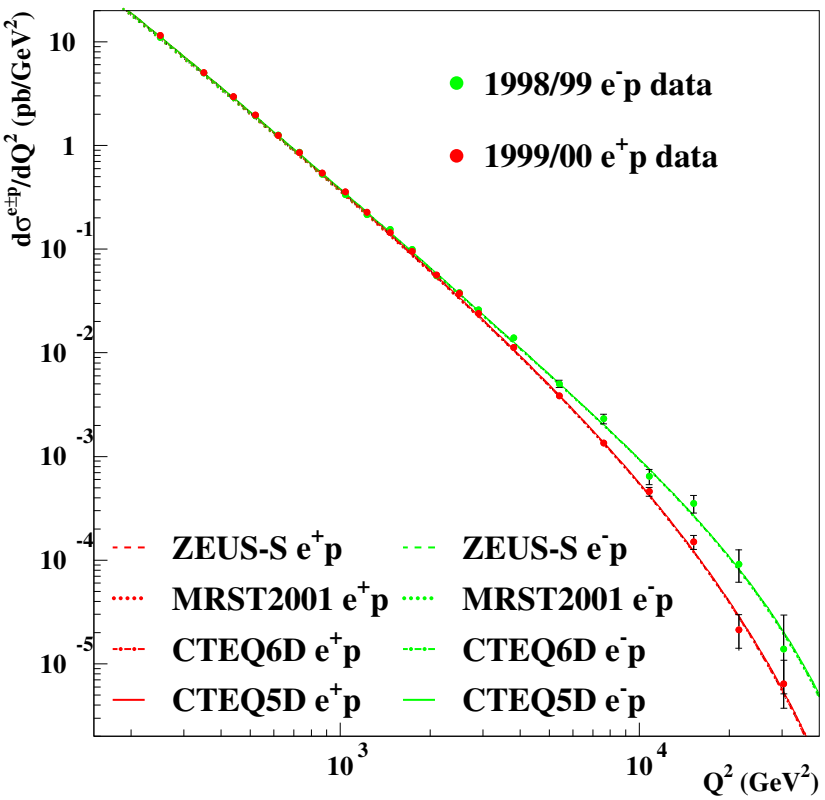

Figure 8.9: Single differential cross section $d\sigma/dQ^2$ measured for e^+p and e^-p scattering. The effect of the Z^0 exchange is visible at high Q^2 by the separation of the e^+p and e^-p cross sections. Measured cross section points are plotted together with the cross sections calculated using several sets of PDFs. Statistical errors are shown as terminated error bars. The total size of the error bars corresponds to the total uncertainty, obtained by combining statistical and systematic errors in quadrature.

(overlapped) in the figure. These correspond to the predictions calculated using different sets of PDFs, including the CTEQ5D set, used in the MC simulation, the ZEUS-S set, MRST(2001), and CTEQ6D.

The agreement between the measurement and the theoretical prediction is better appreciated in figure 8.10. In this figure, the ratios of the measured cross section to the calculations are shown. The shaded band in the figure corresponds to the ratios of the $d\sigma^{e^\pm p}/dQ^2$ obtained from the upper and lower uncertainties in the ZEUS-S PDF set. Also shown are the ratios of the ZEUS-S and MRST(2001) predictions to the CTEQ5D.

Within errors, the theoretical calculations are in agreement. It should be noted that the uncertainty in the measured luminosity reflects into a normalisation uncertainty of 2.25% for the $d\sigma^{e^+p}/dQ^2$ cross section and 1.8% for $d\sigma^{e^-p}/dQ^2$.

Figure 8.11 again shows the measured neutral current $d\sigma^{e^\pm p}/dQ^2$ cross sections, but

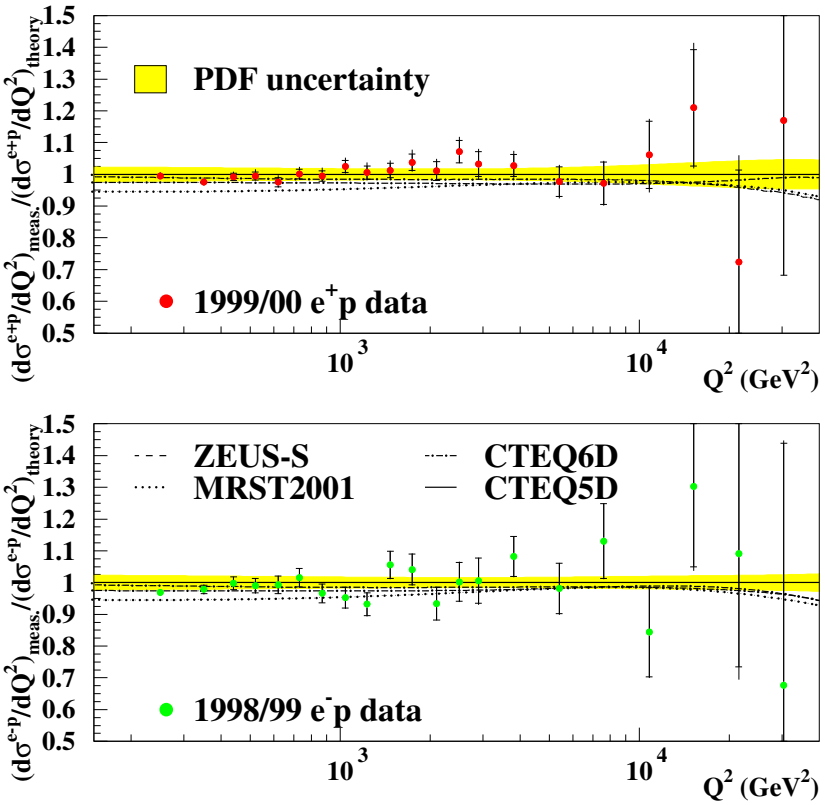

Figure 8.10: Ratios of the single differential cross sections $d\sigma^{e^-p}/dQ^2$ and $d\sigma^{e^+p}/dQ^2$ to the Standard Model predictions. The shaded band represents the uncertainty associated with the PDFs. Also shown are the ratios of the SM predictions using the CTEQ6D, MRST(2001) and the ZEUS-S fit PDF sets to that of CTEQ5D. Inner error bars correspond to the statistical uncertainty. The outer error bars correspond to statistical and systematic errors added in quadrature.

additionally it shows previous ZEUS measurements of the charged current cross section [38, 122]. As can be seen in the figure, although for low Q^2 the charged current cross section is much smaller than the neutral current one, at high Q^2 they become similar, reflecting the unification of the weak and electromagnetic interactions.

8.2.2 Single differential cross section dσ/dx

The single differential cross section $d\sigma/dx$ was measured in two ranges of Q^2: $Q^2 > 200$ GeV^2 and $Q^2 > 10000$ GeV^2. The measurements are shown in figure 8.12 for both the e^+p and the e^-p data samples and given in tables A.4 and A.9 for $Q^2 > 200$ GeV^2 and tables A.5 and A.10 for $Q^2 > 10000$ GeV^2 .

As can be seen in the figure, the accessible range of x becomes smaller as Q^2 grows. The measurements also become dominated by statistical errors, while for the $Q^2 > 200$ GeV^2

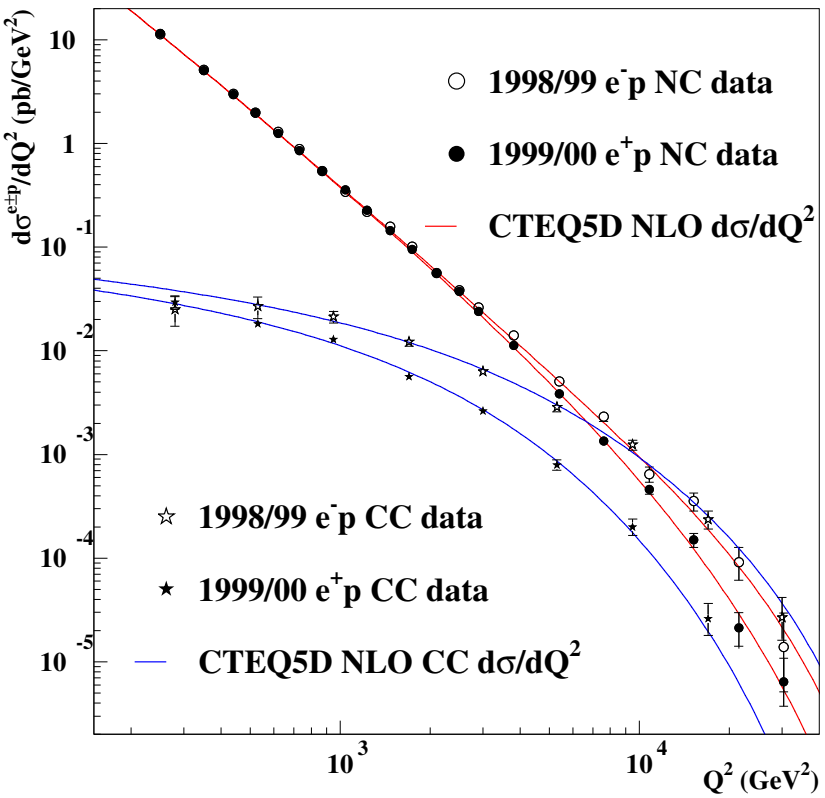

Figure 8.11: Comparison of the single differential cross section $d\sigma/dQ^2$ for e^-p and e^+p neutral current and charged current DIS. The solid lines correspond to the Standard Model predictions calculated using the CTEQ5D PDF set. Inner error bars correspond to statistical errors only, while the outer error bars correspond to the statistical and systematic errors added in quadrature.

range, statistical and systematic errors are comparable. In the highest Q^2 range, the effect of the neutral weak current is clearly visible as a separation of the $d\sigma^{e+p}/dx$ and $d\sigma^{e-p}/dx$ cross sections.

The measured cross sections are, within errors, in agreement with the theoretical predictions from the Standard Model. Figures 8.13 **a)** to **d)** show the ratios of the measured cross sections to the theoretical values. The shaded bands correspond to the fractional errors, in the calculated cross section, obtained from the ZEUS-S PDF errors as described in section 8.1.4. The width of these bands can be seen, in figures 8.13 **a)** and **b)**, to be smallest at values of x of around 0.06. This is an artificial effect, due to the sum rules constraining the QCD fits to experimental data. The effect of these constraints is such that if, for example, the valence u-quark density is systematically increased at high x, enhancing the NC DIS cross section in this region, then it must decrease at low x, to keep a constant valence u-quark density in the proton. Thus, a turning point is created at a

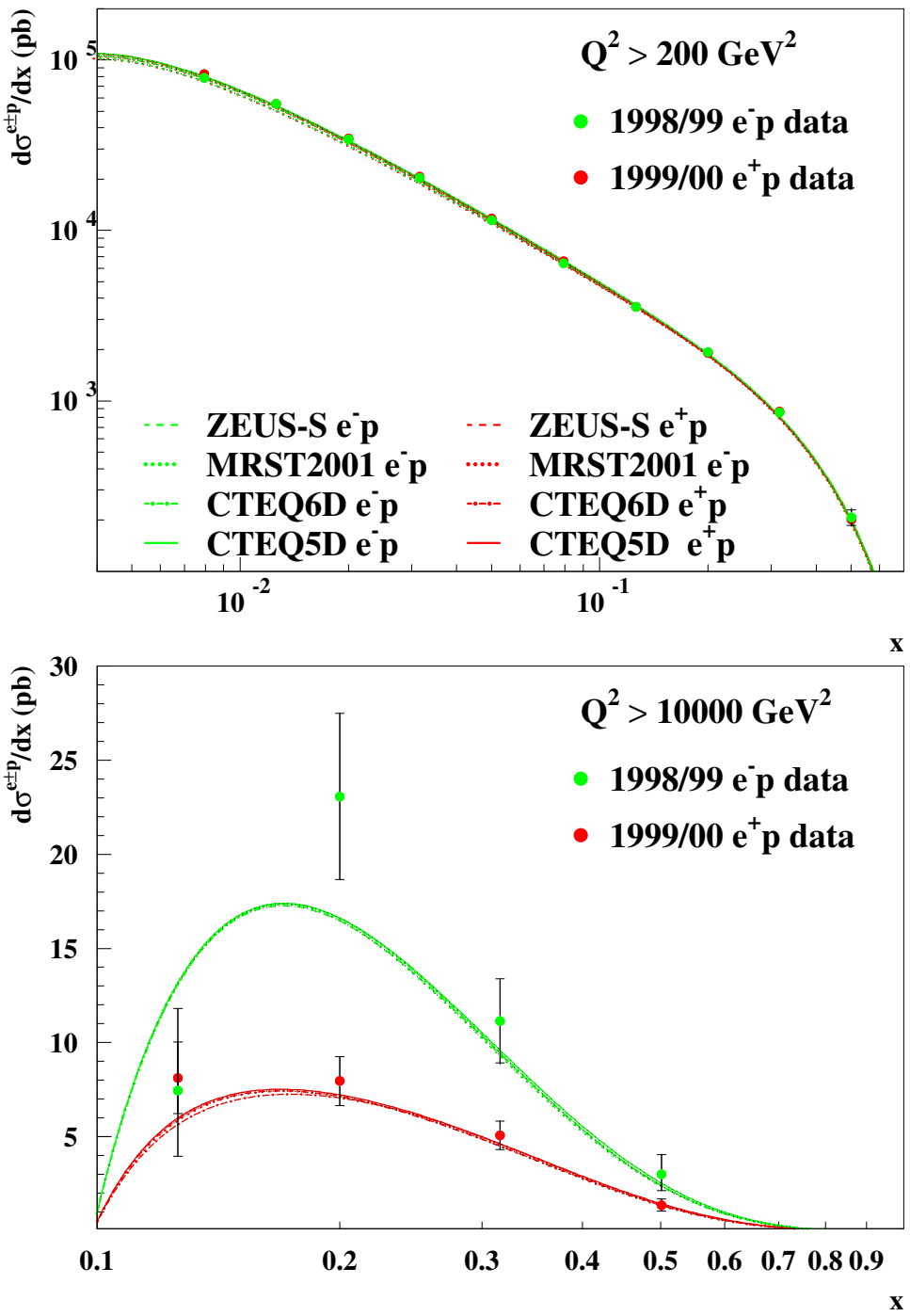

Figure 8.12: Single differential cross section $d\sigma^{e^\pm p}/dx$ for $Q^2 > 200$ GeV^2 (top), and $Q^2 > 10000$ GeV^2 (bottom). The inner error bars correspond to statistical uncertainties. Outer error bars correspond to the statistical and systematic uncertainties added in quadrature. The measurements are compared to the Standard Model predictions calculated using the MRST(2001), CTEQ5D, CTEQ6D and the ZEUS-S fit PDF sets. While the e^-p and e^+p cross sections are almost indistinguishable for $Q^2 > 200$ GeV^2, a clear separation can be seen for $Q^2 > 10000$ GeV^2.

medium value of x, where the calculated uncertainties are smallest.

The curves drawn in each figure show the relative difference between the cross sections calculated with CTEQ5D and the ZEUS-S, MRST(2001) and CTEQ6D PDF sets. The largest differences come from the use of the MRST(2001) PDF set. As figures 8.13 **a)** and **b)** show, the measurement uncertainties at low-Q^2 are now comparable to the differences in the calculated cross sections from the use of different PDF sets. This situation is reversed at high-Q^2, where the statistical uncertainties clearly dominate the measurements.

8.2.3 Single differential cross section dσ/dy

The single differential cross section $d\sigma/dy$ for both e^-p and e^+p NC DIS was measured for the Q^2 ranges: $Q^2 > 200$ GeV^2, and $Q^2 > 10000$ GeV^2. The measurements are shown in figure 8.14. The results are given in tables A.2 and A.7 for $Q^2 > 200$ GeV^2 and in tables A.3 and A.8 for $Q^2 > 10000$ GeV^2. As in the case of the $d\sigma/dx$ cross section, in the high-Q^2 range, the $d\sigma^{e-p}/dy$ cross section becomes higher than $d\sigma^{e+p}/dy$ due to the effect of the Z^0 exchange and $\gamma - Z^0$ interference. The measured cross sections are in good agreement with Standard Model expectation.

Figures 8.15 **a)** to **d)** show the ratios of the measured cross sections to the theoretical values. Again, the differences between cross sections calculated with different PDF sets are, at low-Q^2, comparable to the experimental uncertainty. The statistical uncertainty is dominant in the high-Q^2 range.

The measurements of the single differential cross sections, apart for their intrinsic interest, probe the available data in different kinematic variables. They show a good level of agreement between the data and the Standard Model.

8.2.4 Reduced cross section $\tilde{\sigma}^{e^{\pm}p}$

The reduced cross section at Born level, $\tilde{\sigma}^{\pm}_{Born}$, is obtained by dividing the expression of the double differential cross section by the factors affecting F_2:

$$\tilde{\sigma}^{e^{\pm}p}_{Born}(x,Q^2) = \left(\frac{2\pi\alpha^2}{x\,Q^4}Y_+\right)^{-1} \frac{d^2\sigma^{\pm}_{Born}(x,Q^2)}{dx\,dQ^2} = F_2 - \frac{y^2}{Y_+}F_L \mp \frac{Y_-}{Y_+}xF_3 \qquad (8.20)$$

where $Y_+ = 1 + (1-y)^2$ and $Y_- = 1 - (1-y)^2$.

The unfolding of the reduced cross section was performed as for the single differential cross sections shown above. Figure 8.16 shows the reduced cross section plotted as a function of x in bins of constant Q^2. The values measured at the bin centres, represented by the dots, are compared to the theoretical reduced cross section curves calculated with the CTEQ5D PDF set, used in the MC simulation, as well as with the CTEQ6D and MRST(2001) sets. These results are also given in appendix B.

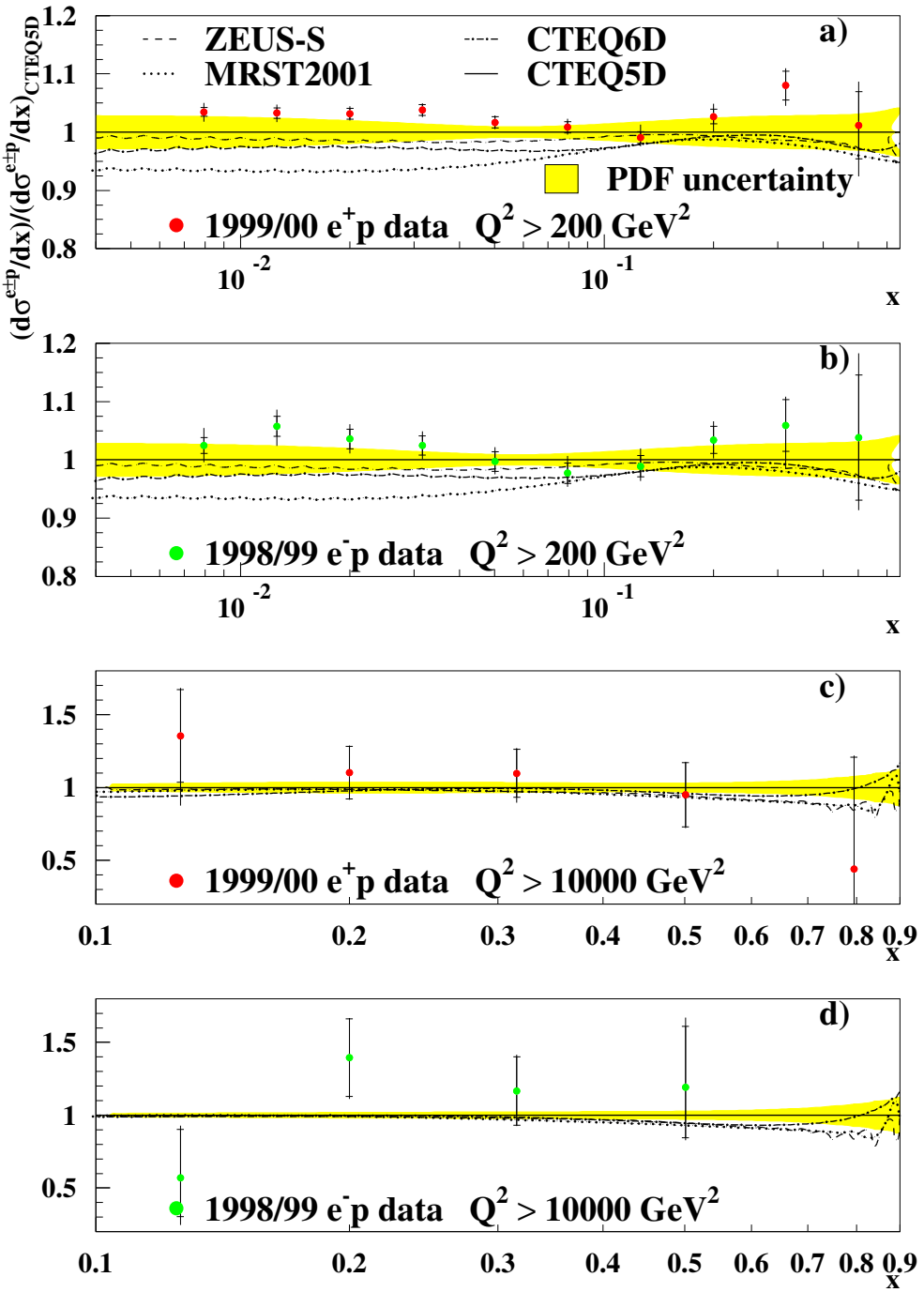

Figure 8.13: Ratios of the single differential cross section $d\sigma/dx$ measured for e^+p (**a**) and **c**)) and e^-p (**b**) and **d**)) for $Q^2 > 200\text{GeV}^2$ and $Q^2 > 10000\text{GeV}^2$. The shaded band corresponds to the uncertainty associated with the PDFs. Also shown are the ratios of the SM predictions using the CTEQ6D, MRST(2001) and the the ZEUS-S fit PDF sets to that of CTEQ5D. Inner error bars correspond to the statistical uncertainty. The outer error bars correspond to statistical and systematic errors added in quadrature.

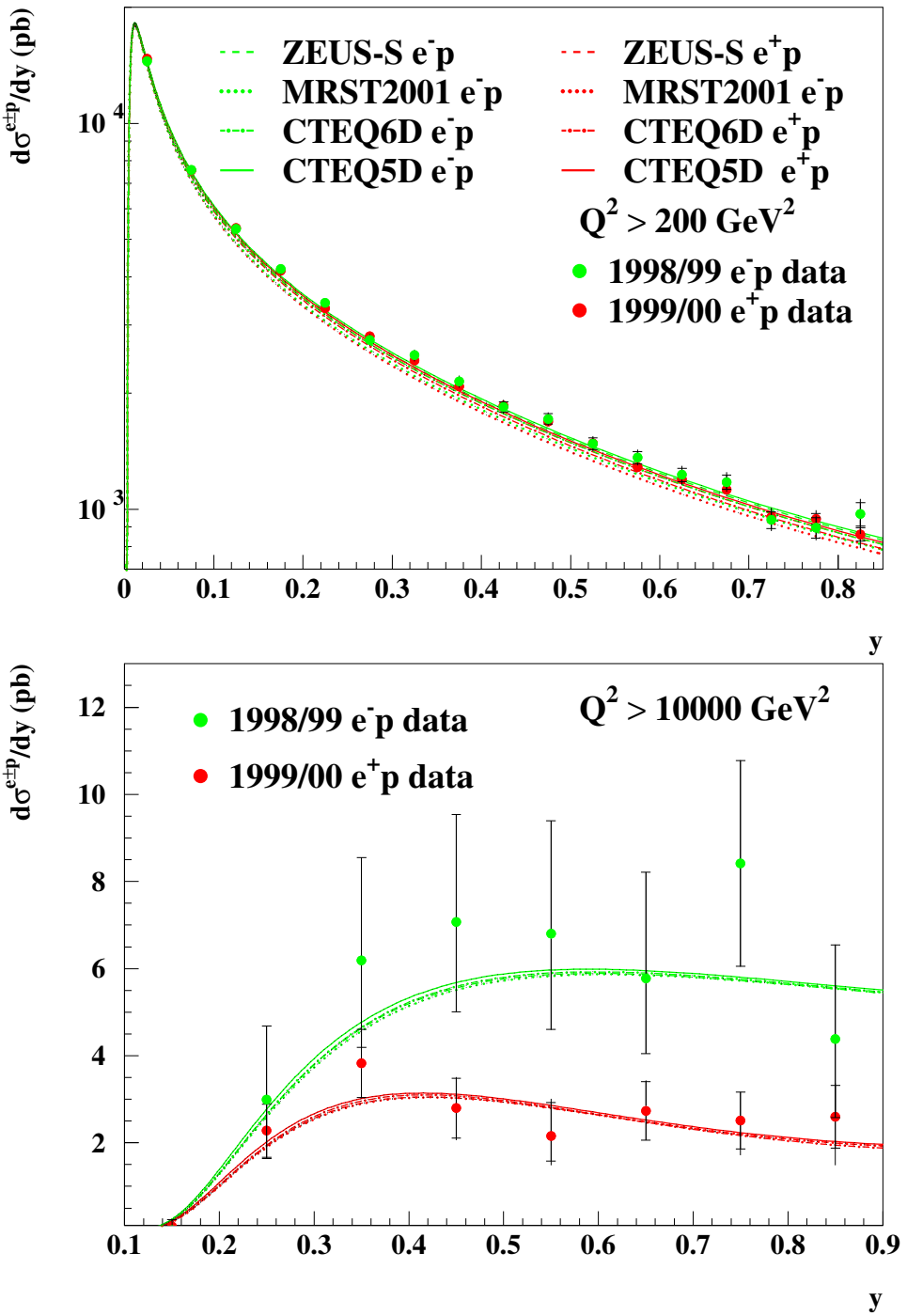

Figure 8.14: Single differential cross section $d\sigma/dy$ for, and $Q^2 > 10000$ GeV^2. The lines correspond to the Standard Model predictions calculated using the MRST(2001), CTEQ5D, CTEQ6D and the ZEUS-S fit PDF sets. The inner error bars correspond to statistical uncertainties. Outer error bars correspond to the statistical and systematic uncertainties added in quadrature. A clear separation between the e^-p and e^+p cross sections can be seen for $Q^2 > 10000$ GeV^2.

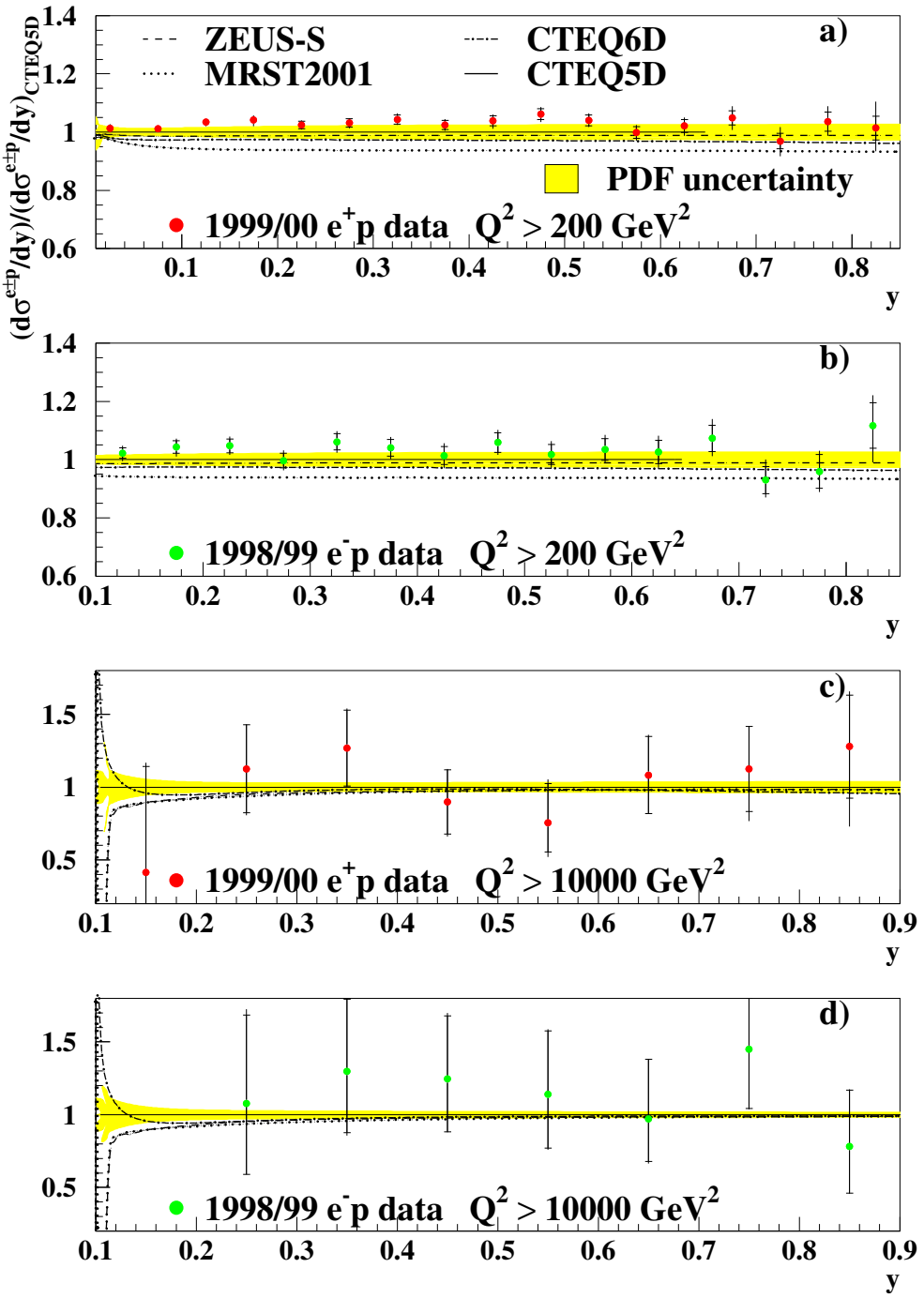

Figure 8.15: Ratios of the single differential cross section $d\sigma/dy$ measured for e^+p (**a**) and **c**)) and e^-p (**b**) and **d**)) for $Q^2 > 200\text{GeV}^2$ and $Q^2 > 10000\text{GeV}^2$. The shaded band corresponds to the uncertainty associated with the PDFs. Also shown are the ratios of the SM predictions using the CTEQ6D, MRST(2001) and the the ZEUS-S fit PDF sets to that of CTEQ5D. Inner error bars correspond to the statistical uncertainty. The outer error bars correspond to statistical and systematic errors added in quadrature.

Reduced cross sections for the e^+p and the e^-p processes, $\tilde{\sigma}^+_{Born}(x,Q^2)$ and $\tilde{\sigma}^-_{Born}(x,Q^2)$, respectively, are shown in the figure. The effect of the Z^0 exchange, and its interference with photon exchange, becomes visible at Q^2 values of around 3000 GeV^2 and above.

8.3 The extraction of F_2

The F_2^{em} structure function was obtained by correcting the unfolded reduced cross section for contributions from F_L and xF_3, as well as for the Z^0 exchange, F_2^Z, and $\gamma - Z^0$ interference, $F_2^{\gamma-Z}$, contributions to the generalised F_2 function. The correction was done according to expression 8.21, where the correction factors were calculated with the CTEQ5D PDF set. The sign of the correction factor for the xF_3 contribution is reversed when determining F_2^{em} from e^+p and e^-p.

$$\begin{aligned} F_2^{em}(x,Q^2) &= \tilde{\sigma}^{e\pm p}(x,Q^2) \times \\ &\quad \left(1 - \frac{Q^2}{Q^2+M_Z^2}\frac{F_2^{\gamma-Z}}{F_2^{em}} - \frac{Q^4}{(Q^2+M_Z^2)^2}\frac{F_2^Z}{F_2^{em}} + \frac{y^2}{Y_+}\frac{F_L}{F_2^{em}} \mp \frac{Y_-}{Y_+}\frac{xF_3}{F_2^{em}}\right) \\ &= \tilde{\sigma}^{e\pm p}(x,Q^2) \times (1 - \Delta_{F_2} + \Delta_{F_L} \mp \Delta_{xF_3}) \end{aligned} \tag{8.21}$$

Figure 8.17 shows the measurements obtained with expression 8.21 from the e^+p and e^-p data samples. These results are also given in appendix C.

8.3.1 Combined results on F_2

The two F_2^{em} determinations were combined into a single measurement. The e^-p sample corresponds roughly to 1/4 of the integrated luminosity of the e^+p sample. This is partially compensated at high Q^2 by the higher NC DIS cross section for e^-p scattering, which enhances the statistical significance of this sample. The combination of the two measurements benefits from the fact that both analyses were identical. Systematic errors were therefore essentially the same. The ratio of the relative systematic uncertainties in the two samples is shown in figure 8.18. As can be seen in the figure, for low-Q^2 bins, the estimates of the relative systematic uncertainties are essentially the same for both measurements. For higher-Q^2 bins it can be expected that the estimated systematic uncertainties are somewhat affected by statistical fluctuations in the data, due to the relatively small numbers of events used in the estimation. As a result, the ratios show a larger spread but are still close to unity. It was assumed that the relative systematic uncertainty values estimated using the larger e^+p data sample were a good approximation to the systematic uncertainties affecting the combined measurement. This is justified by the ratios in figure 8.18 and avoids double-counting.

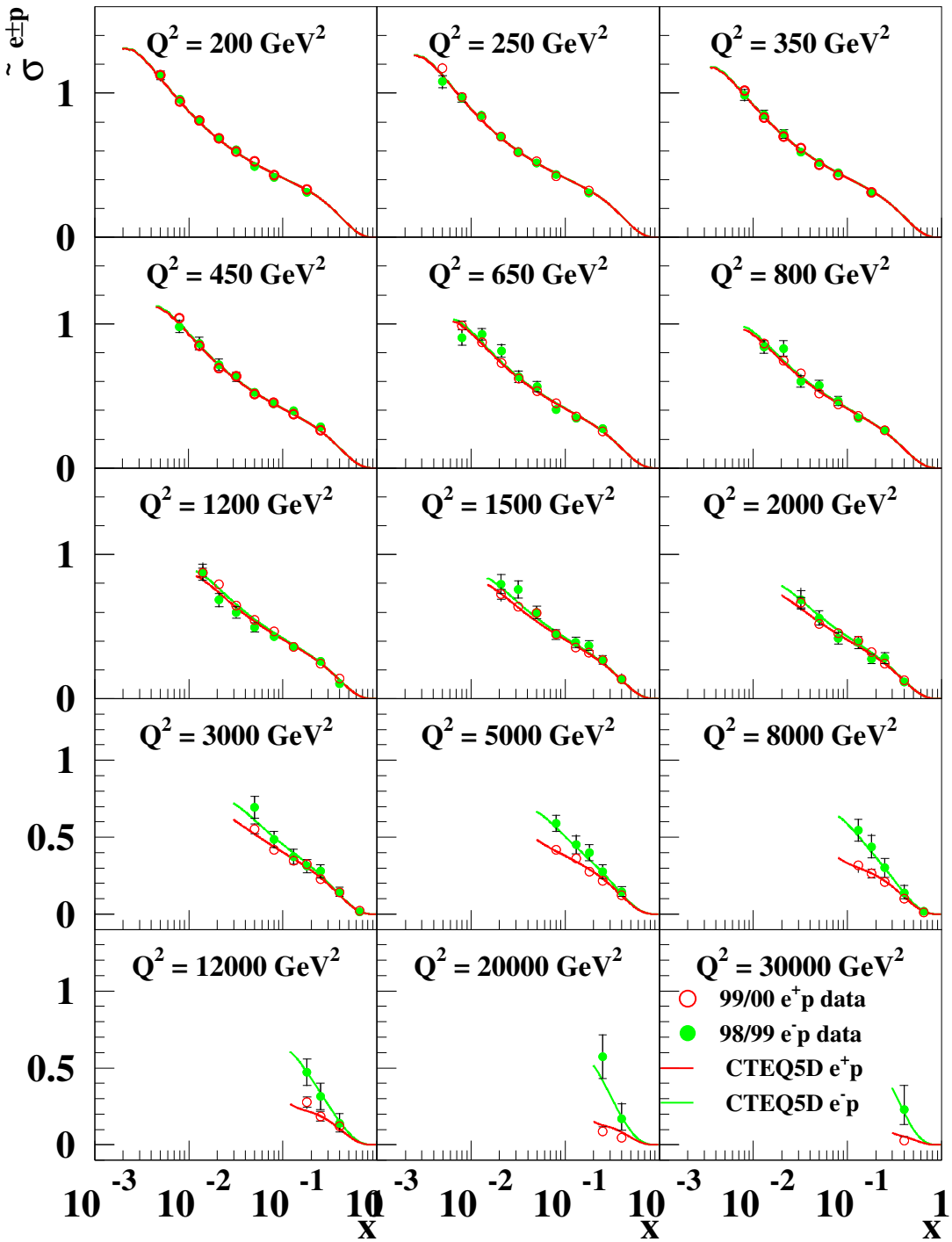

Figure 8.16: Double differential reduced cross sections $\tilde{\sigma}^{e^+p}$ (open circles) and $\tilde{\sigma}^{e^-p}$ (filled circles). The theory curves plotted together with the measured values were obtained with the CTE5D, CTEQ6D, MRST(2001) and ZEUS-S sets of PDFs.

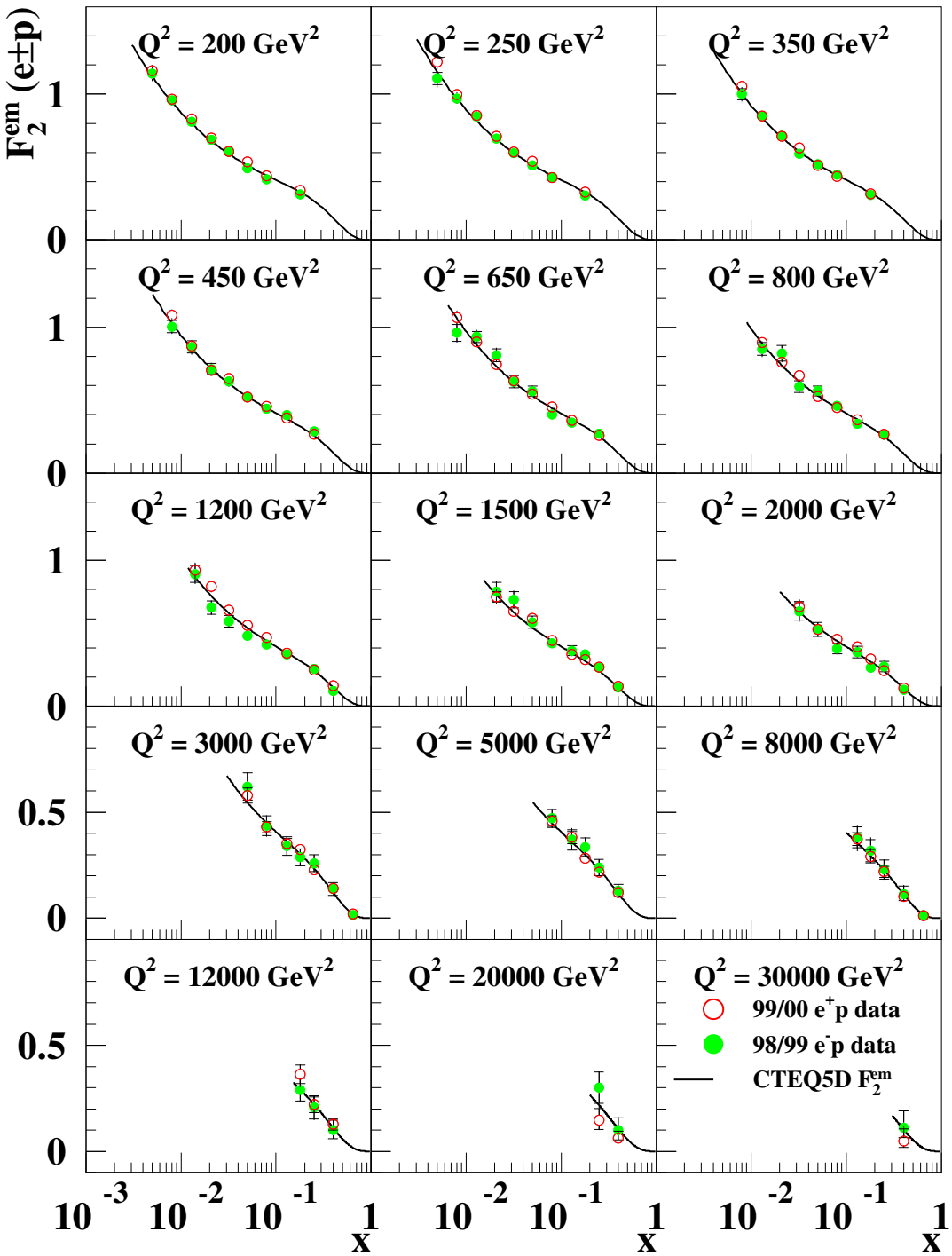

Figure 8.17: The F_2^{em} structure function measurements from the e^+p and e^-p data samples.

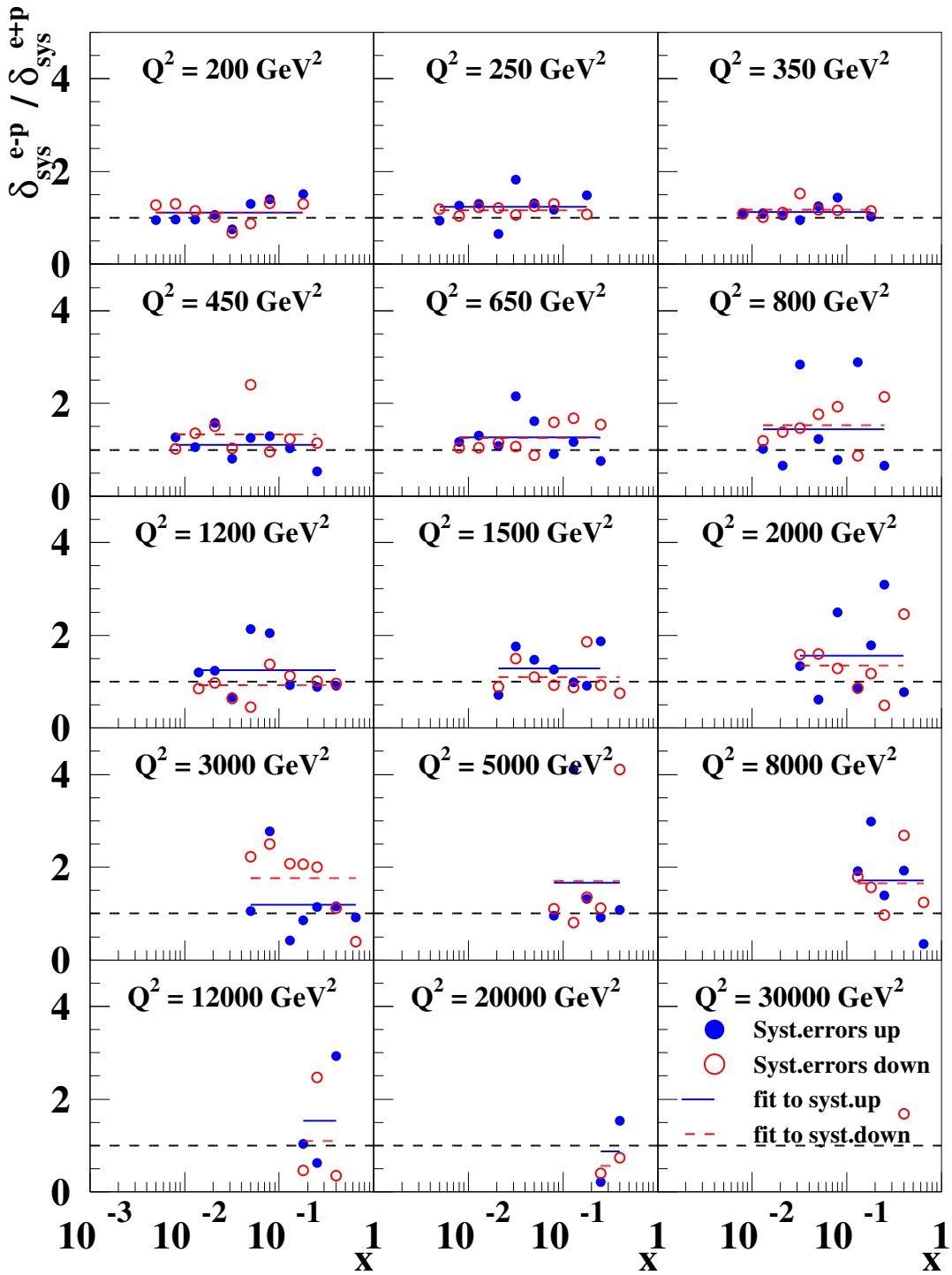

Figure 8.18: Ratios of relative systematic uncertainties in the F_2 measurements from the e^-p and the e^+p data samples. Positive and negative uncertainties are shown by solid and empty circles, respectively. The ratios are approximately 1 in low-Q^2 bins, where statistical fluctuations that affect the determination of systematic uncertainties at higher-Q^2 bins are negligible.

The uncertainty in the integrated luminosity must be treated differently, because it contributes with an unknown absolute normalisation factor that applies to each data sample. The *relative* normalisation of the two data samples, though, can be determined from the data. The ratio of the $d\sigma^{e-p}/dQ^2$ and $d\sigma^{e+p}/dQ^2$ cross sections was used to determine the relative normalisation, i.e. the bias in the measurement of the integrated luminosity of one data sample with respect to the integrated luminosity of the other. The ratio of the cross sections must be measured at relatively low Q^2, where the effect of the xF_3 term is negligible. Figure 8.19 shows the measurement of this ratio which indicates a normalisation factor of 1.005 ± 0.009 between the samples which is consistent with zero relative bias. A luminosity uncertainty of 2.25% was considered to apply to the combined sample.

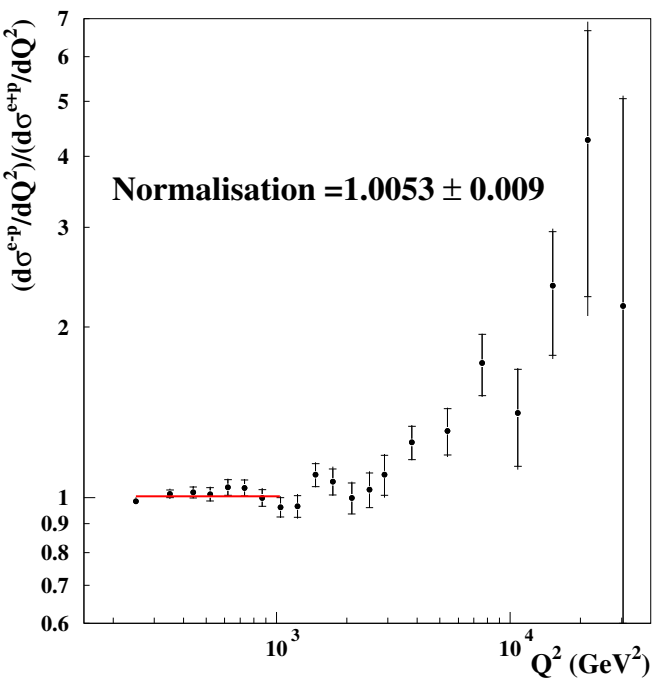

Figure 8.19: Ratio of the single differential cross sections $d\sigma^{e-p}/dQ^2$ and $d\sigma^{e+p}/dQ^2$. The cross section ratio at low Q^2, where electroweak effects are negligible, constitutes a measurement of the relative normalisation of the 1998/99 e^-p and the 1999/00 e^+p data sets. The ratio is consistent with zero normalisation offset between the two data sets.

The statistical errors in the two measurements of F_2, Δ^{e-p}_{stat} and Δ^{e+p}_{stat}, were used to obtain a combined measurement, according to the maximum likelihood principle:

$$F_2^{combined} = \frac{\frac{F_2^{e-p}}{(\Delta^{e-p}_{stat})^2} + \frac{F_2^{e+p}}{(\Delta^{e+p}_{stat})^2}}{\frac{1}{(\Delta^{e-p}_{stat})^2} + \frac{1}{(\Delta^{e+p}_{stat})^2}} \tag{8.22}$$

In most of the bins used, more than 12 events were observed in the data. This allowed the use of symmetric Gaussian errors obtained from the square root of the number of events. In these, expression 8.22 was used without modification. In only 4 and 5 bins for the e^+p and the e^-p data samples, respectively, were Poisson confidence intervals employed to estimate the errors. In these bins, half of the difference between the upper and lower boundary of the Poisson 68.3% confidence interval was used in expression 8.22.

For bins where the number of observed events was larger than 12 in both measurements, the statistical errors of the combined measurement were estimated as:

$$\Delta_{F_2}^{combined} = \sqrt{\frac{1}{\frac{1}{(\Delta_{stat}^{e-p})^2} + \frac{1}{(\Delta_{stat}^{e+p})^2}}} \tag{8.23}$$

Figure 8.20 shows the combined measurement of F_2 using the measurements obtained with both the e^-p and the e^+p samples. The combined F_2 is given in tables C.1 and C.2.

8.4 The extraction of xF$_3$

The parity violating part of the NC DIS cross section is contained in the xF_3 structure function. The reduced cross sections for these processes can be written as:

$$\tilde{\sigma}^{e+p} = F_2 - \frac{y^2}{Y_+}F_L - \frac{Y_-}{Y_+}xF_3 \tag{8.24}$$

$$\tilde{\sigma}^{e-p} = F_2 - \frac{y^2}{Y_+}F_L + \frac{Y_-}{Y_+}xF_3 \tag{8.25}$$

So, xF_3 can be extracted from measurements of the reduced cross sections for the e^-p and e^+p according to the expression:

$$xF_3 = \frac{1}{2}\frac{Y_+}{Y_-}\left(\tilde{\sigma}^{e-p} - \tilde{\sigma}^{e+p}\right) \tag{8.26}$$

A first measurement of xF_3 from ZEUS [134] was made using e^+p data collected in the 1996/97 period and e^-p data collected in the 1998/99 period. The proton beam energy during 1996 and 1997 was 820 GeV and the centre of mass energy, $\sqrt{s}$, was therefore lower than in the 1998/99 period. This made necessary the use of additional kinematic correction factors in the measurement that are not necessary in the present analysis. On the other hand, the two analyses use a common data sample and are not, therefore, independent. Additionally, the limiting factor in the precision of this measurement is the common e^-p sample, so the gain in precision is not substantial. The measurement presented here can, nevertheless, be seen as a verification of the previous determination of xF_3.

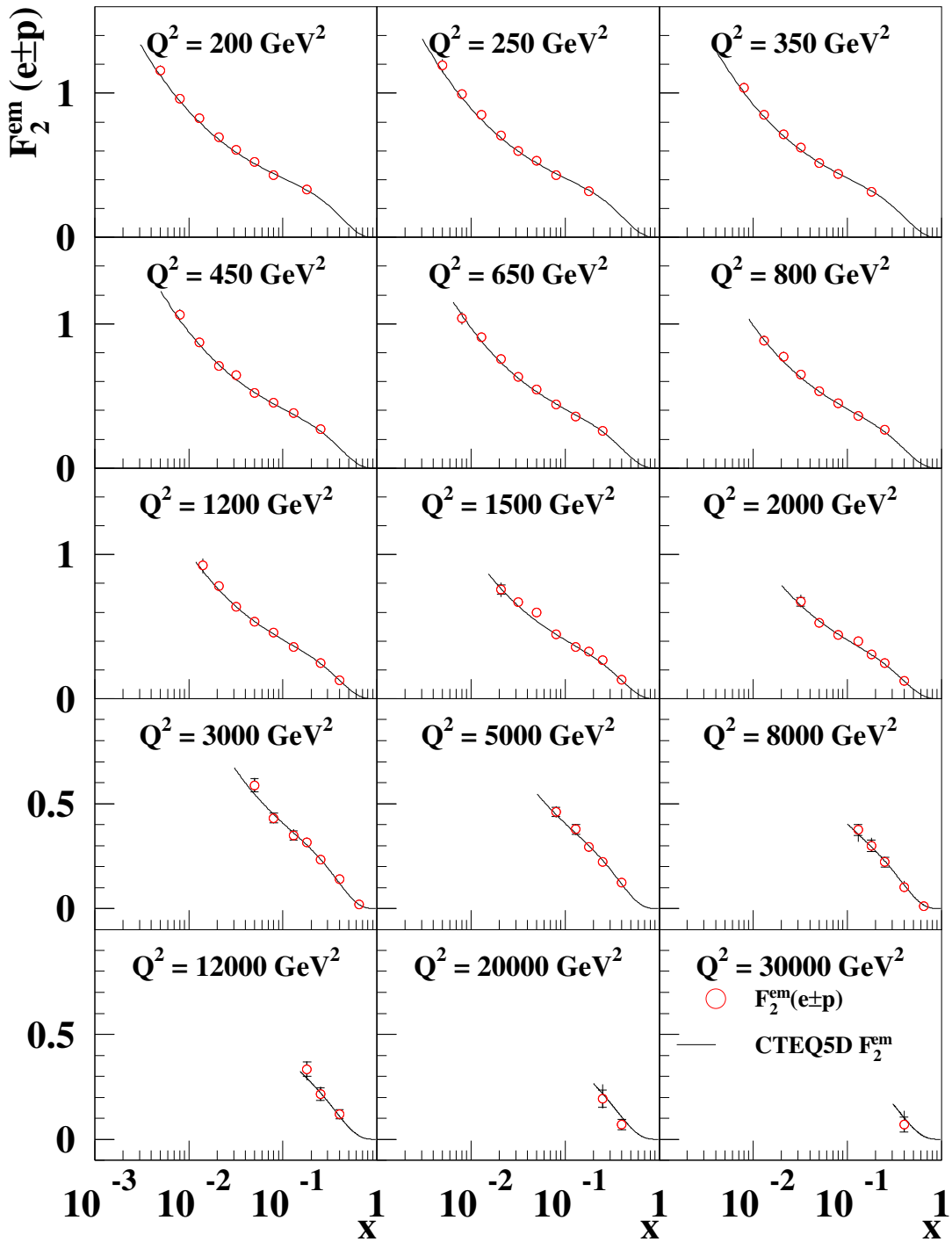

Figure 8.20: The F_2^{em} structure function measurement from the combined e^+p and e^-p data samples.

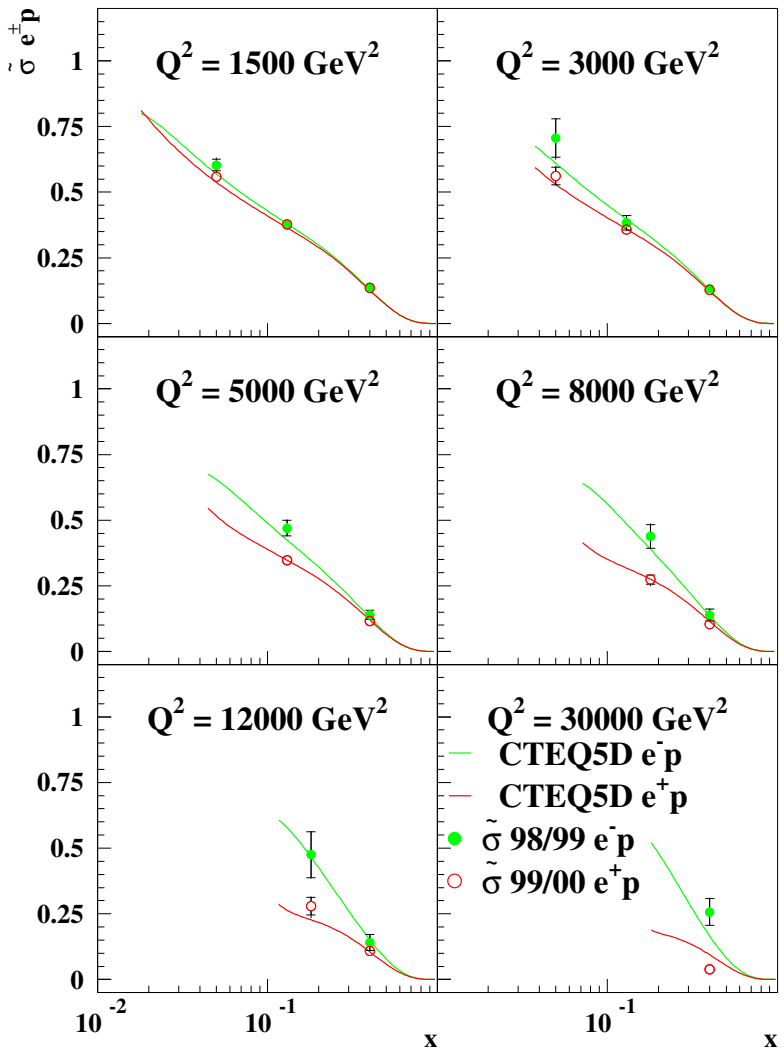

Figure 8.21: Double differential reduced cross sections $\tilde{\sigma}^{e^-p}$ and $\tilde{\sigma}^{e^+p}$ measured in the bins used for the extraction of xF_3. The solid lines correspond to xF_3 calculated using the CTEQ5D PDF set.

Due to the limited number of events in the high-Q^2 region where sensitivity to xF_3 exists, several bins used in the measurement of F_2 were combined. Figure 8.21 shows the reduced cross section measured in those bins for both the e^+ and e^- data sets.

Figure 8.22 shows the measured xF_3, and the values of are given in table C.3. In chapter 9, these results are compared to measurements published by the H1 and ZEUS collaborations.

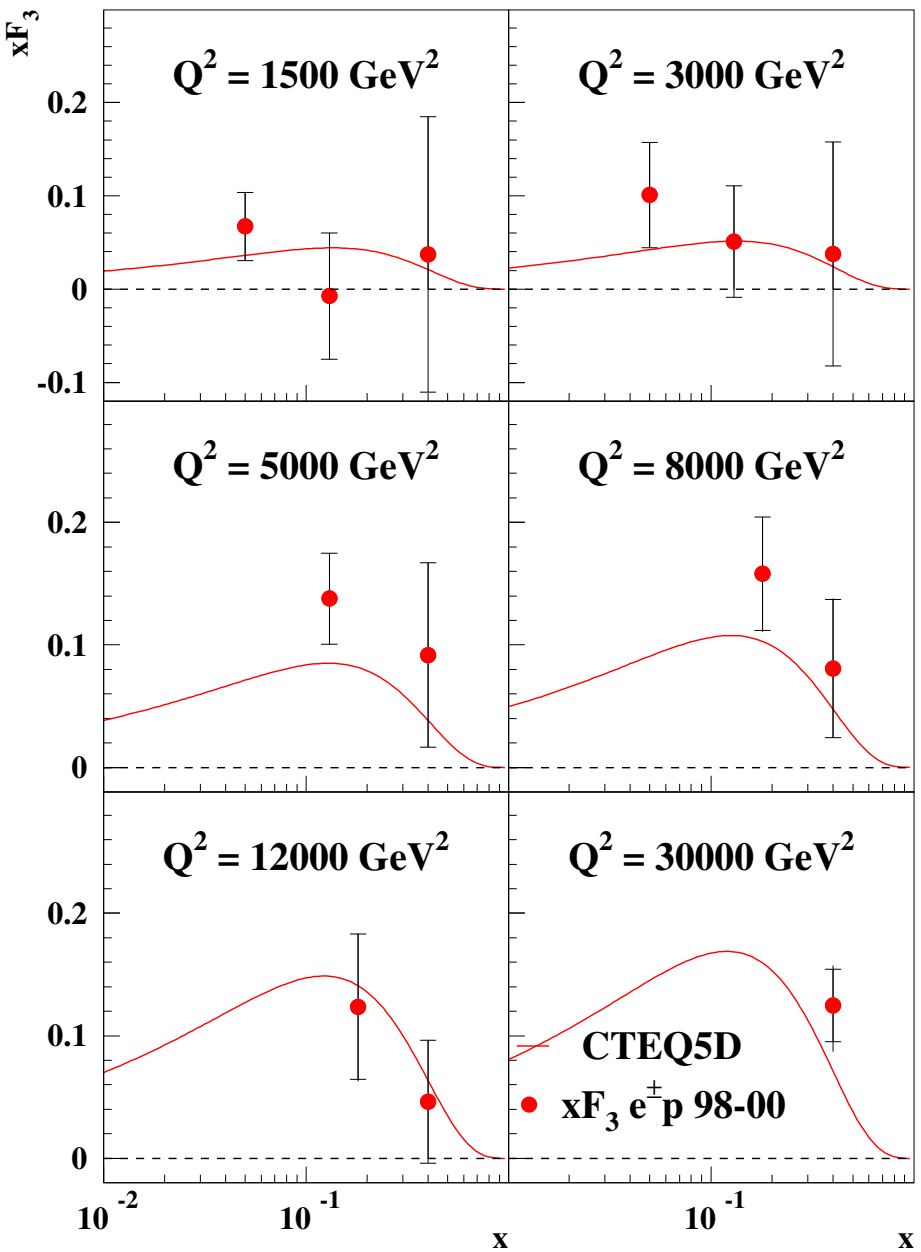

Figure 8.22: Measurement of the xF_3 structure function compared to the Standard Model prediction calculated using the CTEQ5D parton density functions.

Chapter 9

Conclusions and outlook

In this chapter, some of the results obtained in the present work are reviewed and compared to published results. In the section 9.2, the results on F_2 are compared to measurements from both fixed-target experiments and to HERA results. In section 9.1.2, a comparison is made between the results obtained on xF_3 and ZEUS and H1 results. The last section gives an outlook on what results can be expected from the new HERA running period for inclusive NC DIS.

9.1 Comparison to published results

9.1.1 Comparison of F_2 with published results

A great wealth of measurements of the electromagnetic structure function F_2 have been accumulated by several experiments over the years. These span a very large region of phase space. Figure 9.1 shows a comparison between the results obtained in the present analysis and published results from the HERA experiments H1 [34] and ZEUS [33], as well as results from several fixed-target experiments. Results from the HERA experiments do not include measurements done with the 1999/00 data sets, which have not been published at the time of writing. The fixed target measurements shown here were published by the NMC [135], BCDMS [136] and E665 [137] collaborations. F_2 is represented in the figure as a function of Q^2 in bins of x. Only a selection of bins of x are shown in the figure, to allow a close look at the level of agreement between measurements. The figure shows the approximate scaling behavior of F_2 for x close to 0.2, and shows clearly the logarithmic scaling violations for other values of x. All experimental points are in good agreement within errors. The curves shown in the figure correspond to the expectations from the Standard Model calculated with different sets of parton density functions. The CTEQ6D set of PDFs supersedes the CTEQ5D set. The latter is shown for completeness, as it was used in the generation of the NC DIS Monte Carlo data.

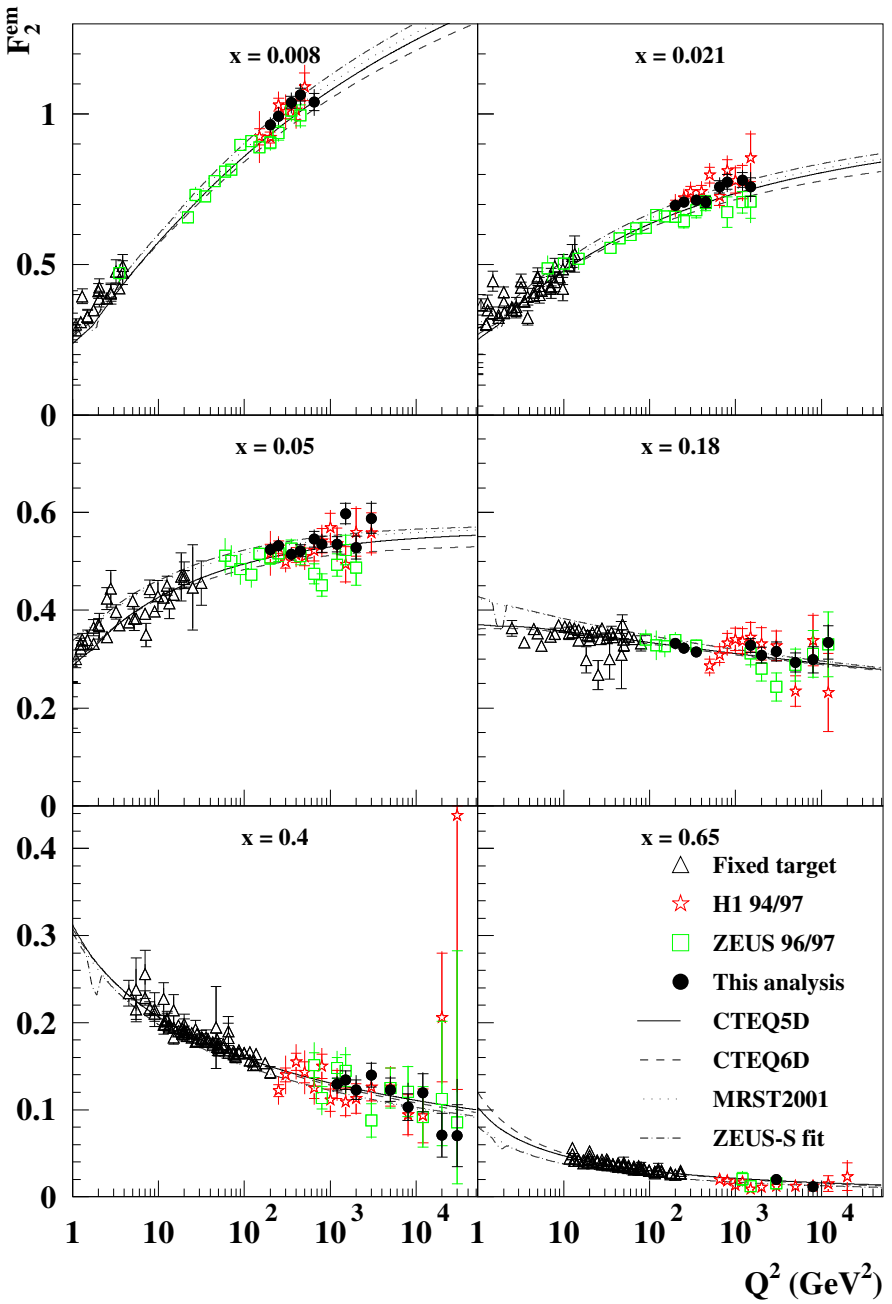

Figure 9.1: Results for the F_2 structure function measurement (filled circles) compared to measurements from the fixed-target experiments NMC [135], BCDMS [136] and E665 [137] (empty triangles), and to recent results from H1 (stars) and ZEUS (empty squares). The H1 [34] measurement was obtained using e^+p data collected from 1994 to 1997. The ZEUS measurement of F_2 [33] was performed using e^+p data from 1996 and 1997 and e^-p data from 1998 and 1999. The e^-p data set used in the ZEUS measurement is, therefore, the same as used in the present analysis. The measurements are compared to the Standard Model prediction calculated using the CTEQ5D, CTEQ6D, MRST2001 and ZEUS-S sets of parton density functions.

Figure 9.2 shows the same experimental results as figure 9.1 but expands the abcissae scale by restricting Q^2 to $Q^2 > 100$ GeV^2. The level of precision of the determinations of F_2 at high Q^2 and of the agreement between different measurements is more clearly visible in figure 9.2. As can be seen, the experimental points are in agreement within the respective uncertainties. The calculations of the structure function F_2 that use different sets of PDFs show significant differences in some regions of phase space. The calculations performed using the ZEUS-S and the MRST2001 seem to be similar and to differ from the calculations performed using CTEQ PDFs. At high Q^2 and low x, the experimental uncertainty of the present measurement is comparable to the differences between the calculations obtained from different PDF sets. The HERA experimental results including data from 1999/00 can thus be expected to have some impact on the next iteration of the parton density functions.

A measurement of the structure function F_2 combining data from the 1999/00 running period with data from 1996/97 instead of that of the 1998/99 period will be published in a forthcoming ZEUS paper [138]. The 1996/97 data set corresponds to 30 pb^{-1} [33], around twice the size of the e^-p data set used in the present analysis. The larger integrated luminosity should lead to smaller statistical errors in the ZEUS measurement than those achieved in the present work. Only in three bins at the highest Q^2 should the measurement presented here have smaller errors than those of the ZEUS analysis, due to the enhanced e^-p cross section.

9.1.2 Comparison of xF_3 with published results

The measurement of the structure function xF_3 is shown in figure 9.3 together with the recent measurement from the H1 [37] experiment and ZEUS [134]. The H1 and ZEUS measurements were the first determination of xF_3 at high Q^2. These measurements have used e^+p data samples collected at a centre-of-mass energy of $\sqrt{s} = 300$ GeV and e^-p data at $\sqrt{s} = 318$ GeV. Corrections were therefore necessary to account for the differences in the kinematic factors multiplying xF_3 in the expression for the cross section (see equation 2.10 in chapter 2). Due to the different centre-of-mass energies at which the cross sections were extracted, the F_L term does not cancel in the subtraction of the e^-p and e^+p cross sections and the difference must be taken into account. These difficulties were avoided in the present analysis by using data samples collected at the same centre-of-mass energy. The e^-p sample used in the present analysis is the same as was used in the ZEUS measurement. The two extractions of xF_3 are therefore not independent.

The three measurements show good agreement and are well described by the Standard Model prediction calculated with the CTEQ5D PDF set. The data may indicate a somewhat higher value of xF_3 than given by the calculation at $Q^2 \sim 5000$ GeV and $x \lesssim 0.2$, but the large statistical errors preclude any definite statement.

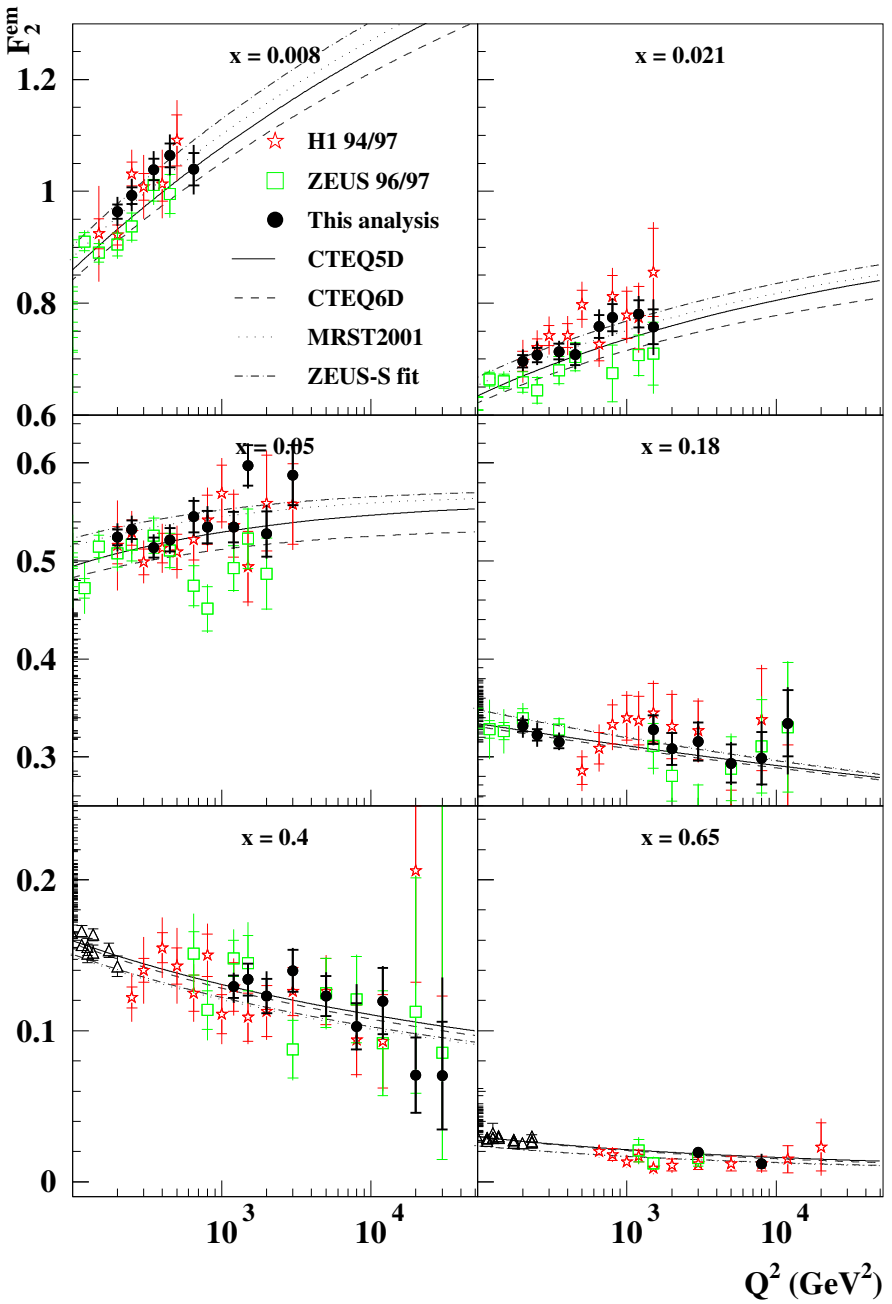

Figure 9.2: Results for the structure function F_2 (filled circles) compared to published results from H1 [34] and ZEUS [33] and some experimental points from fixed-target experiments as in figure 9.1. This figure shows the same bins as figure 9.1 but restricted to $Q^2 > 100$ GeV^2.

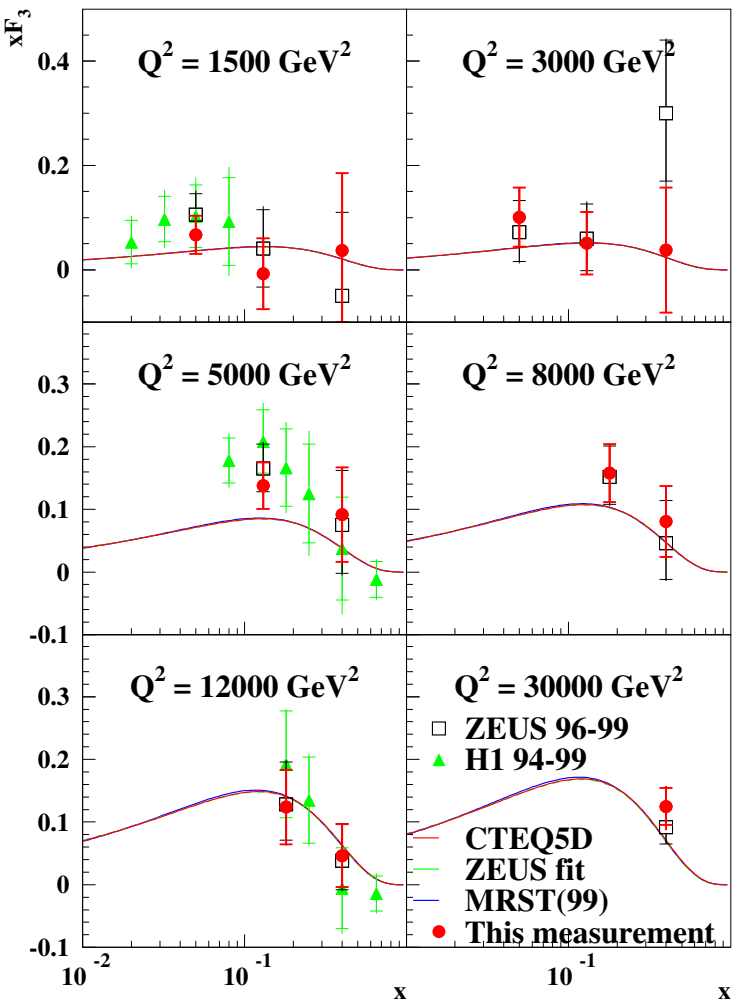

Figure 9.3: Measurement of the xF_3 structure function compared to the most recent measurements from H1 and ZEUS. The H1 [37] measurement was obtained using e^+p data collected from 1994 to 1997 and e^-p data from 1998 and 1999. The ZEUS measurement of xF_3 [134] was performed using e^+p data from 1996 and 1997 and e^-p data from 1998 and 1999. The e^-p data set used in the ZEUS measurement is, therefore, the same as used in the present analysis. The measurements are compared to the Standard Model prediction calculated using the CTEQ5D parton density functions.

9.2 Outlook: HERA-II

The HERA collider has recently started its operation after a very comprehensive upgrade. As a result of the upgrade, the instantaneous luminosity is expected to increase substan-

tially. The goal of the upgrade is that HERA will collect a data set with integrated luminosity of around 1 fb^{-1} by 2006, when the collider will reach the end of its life.

Along with the upgrade to the collider, both the H1 and the ZEUS experiments have been extensively improved. In the ZEUS detector, a new silicon-strip vertex detector and a new forward tracking system composed of several layers of straw-tube drift chambers have been installed. These will improve the detector capabilities especially in the tagging and study of heavy particles such as the $D^{*\pm}$ mesons, by improving the accuracy of the vertex finding, the momentum determination and the solid angular coverage of the tracking detectors. Improving the tagging capabilities for charmed mesons, on the other hand, will result in better determinations the $F_2^{c\overline{c}}$ structure function and of the charm content of the proton.

Another important feature of the HERA upgrade will be the longitudinal polarisation of the lepton beam. Polarisations of up to ~ 60 % are expected to be achieved, where the polarisation, $\mathcal{P}$, is defined by:

$$\mathcal{P} = \frac{N_R - N_L}{N_R + N_L} \tag{9.1}$$

where N_R and N_L are the number of right- and left-handed leptons in the beam, respectively. For polarised-lepton proton NC DIS the Born-level cross section given in chapter 2 must be modified. Neglecting the contribution of F_L, which is small at high-Q^2, the NC DIS cross section can be written as [139]:

$$\frac{d^2\sigma^{e^{\pm}p}}{dxdQ^2} = \frac{2\pi\alpha^2}{xQ^4}\left[Y_+F_2 \mp Y_-xF_3 + \mathcal{P}\left(Y_+F_2^{\mathcal{P}} \mp Y_-xF_3^{\mathcal{P}}\right)\right] \tag{9.2}$$

where $Y_\pm = 1 \pm (1-y)^2$, and F_2 and xF_3 are the usual structure functions measured in the present work. The new structure functions $F_2^{\mathcal{P}}$ and $xF_3^{\mathcal{P}}$ parametrise the polarisation dependence of the cross section.

At high Q^2, where the Z^0 exchange and the virtual photon-Z^0 interference become significant contributions to the NC DIS process, the cross section depends, not only on whether the lepton beam consists of electrons or positrons, but also on the polarisation of the beam. This is shown in figure 9.4. In the figure, the upper plot shows the ratios of the reduced cross section to the electromagnetic reduced cross section, calculated assuming no weak neutral current, plotted versus Q^2 for $x = 0.25$. The experimental points shown were measured in the present analysis. The bottom plot shows the same ratios for both left- and right-polarised electron and positron beams with $\mathcal{P} = \pm 70$ %. The points shown correspond to Monte Carlo predictions assuming an integrated luminosity of 250 pb^{-1} per beam for each polarisation. The polarised lepton beams will allow a much more complete study of the helicity structure of the weak contribution to the NC cross section. The different cross sections obtained for electron and positron beams and for positive

and negative beam polarisations may be used to extract the electroweak vector and axial couplings, v_q and a_q, of the light quarks u and d, in a model independent way [139]. The relative uncertainties on the measured axial and vector couplings v_u,a_u,v_d and v_d, were estimated to be 13%, 6%, 17% and 17%, respectively. This assumes an integrated luminosity of 1 pb^{-1} divided equally between the four lepton/polarisation combinations, with $\mathcal{P} = 70$ %.

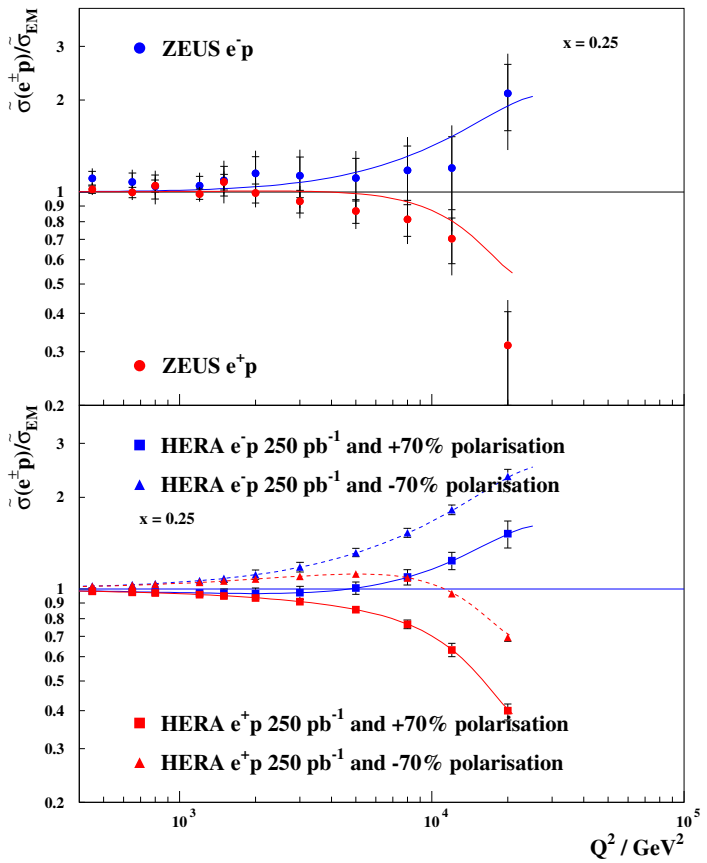

Figure 9.4: Double differential reduced cross section $\tilde{\sigma}$ plotted as a function of Q^2 for $x = 0.25$ and the effect of lepton beam polarisation. Top: measured reduced cross section for unpolarised $e^{\pm}p$ NC DIS at $x = 0.25$. Bottom: reduced cross sections for $e^{\pm}p$ with left- and right-polarised lepton beams.

References

[1] Francis Halzen and Alan D. Martin. *Quarks and Leptons: An introductory course in modern particle physics.* John Wiley & sons, 1984.

[2] L3 The ALEPH, DELPHI and LEP Working Group for Higgs Boson Searches OPAL Collaborations. Search for the Standard Model Higgs Boson at LEP. LHWG Note/ 2002-01, CERN, July 2001. Unpublished.

[3] Y. Fukuda et al. Evidence for Oscillation of Atmospheric Neutrinos. *Phys. Rev. Lett.*, 81(24):1562–1567, 1998.

[4] Q.R. Ahmad et al. Direct Evidence for Neutrino Flavour Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 89(1):011301, 2002.

[5] G. Jungman, M. Kamionkowski, and K. Griest. Supersymmetric dark matter. *Phys.Rep.*, 267:195–373, 1996.

[6] E.D. Bloom et al. High-energy inelastic $e-p$ scattering at 6° and 10°. *Phys. Rev. Lett.*, 23(16):930–934, 1969.

[7] M. Breidenbach et al. Observed behavior of highly inelastic electron-proton scattering. *Phys. Rev. Lett.*, 23(16):935–939, 1969.

[8] M. Gell-Mann. A schematic model of baryons and mesons. *Phys. Lett.*, 8(3):214 – 215, 1964.

[9] R.P. Feynman. The Behavior of Hadron Collisions at Extreme Energies. Talk given at the Third International Conference on High Energy Collisions, Sate University of New York, Stony Brook, 1969. In L.M. Brown, editor, *Selected papers of Richard Feynman*, number 27 in Series on 20^{th} Century Physics, pages 497–518. World Scientific, Singapore, 2000.

[10] The ZEUS Collaboration. Search for contact interactions, large extra dimensions and finite quark radius in ep collisions at HERA. Article in preparation.

[11] H1 Collaboration. Search for compositeness, leptoquarks and large extra dimensions in eq contact interactions at HERA. *Phys. Lett.*, B(479):358–370, 2000. DESY Report: DESY-00-027.

[12] J.D. Bjorken and E.A. Paschos. Inelastic Electron-Proton and γ-Proton Scattering and the Structure of the Nucleon. *Phys. Rev.*, 185(5):1975–1982, 1969.

[13] J.D. Bjorken. Asymptotic Sum Rules at Infinite Momentum. *Phys. Rev.*, 179(5):1548–1553, 1969.

[14] A.M. Cooper-Sarkar, R.C. Devenish, and A. De Roeck. Structure functions of the nucleon and their interpretation. *International Journal of Modern Physics*, A 13(20):3385, August 1998.

[15] Ken Long. A pedagogical introduction to DIS. Talk given at the ZEUS Collaboration meeting, Hamburg, 23-27 October 2000. Unpublished.

[16] Particle Data Group. Review of Particle Physics. *Eur. Phys. J.*, C 15:1–878, 2000.

[17] ZEUS Collaboration. Measurement of the proton structure function F_2 in ep scattering at HERA. *Phys. Lett.*, B(316):412 – 426, 1993.

[18] V.N. Gribov and L.N. Lipatov. Deep inelastic ep scattering in perturbation theory. *Sov. J. Nucl. Phys.*, 15 no. 4:438–450, 1972.

[19] L.N. Lipatov. The parton model and perturbation theory. *Sov. J. Nucl. Phys.*, 20 no. 1:94–102, 1975.

[20] Yu.L. Dokshitser. Calculation of structure functions of deep-inelastic scattering and e^+e^- anihilation by perturbation theory in quantum chromodynamics. *Sov. Phys. JETP*, 46(4):641–653, 1977.

[21] G. Altarelli and G. Parisi. Asymptotic freedom in parton language. *Nucl. Phys.*, B 126:298–318, 1977.

[22] David J. Gross and C.H. Llewelyn-Smith. High-energy neutrino-nucleon scattering, current algebra and partons. *Nucl. Phys.*, B 14:337–347, 1969.

[23] E.A. Hawker et al. Measurement of the Light Antiquark Flavor Asymmetry in the Nucleon Sea. *Phys. Rev. Lett.*, 80(17):3715–3718, 1998.

[24] ZEUS Collaboration. A ZEUS Next-to-Leading-Order QCD Analysis of Data on Deep Inelastic Scattering. DESY Report DESY 02-105, DESY, August 2002. Accepted by Phys. Rev. D. Available at: hep-ex/0208023.

[25] H.L. Lai et al. Global QCD Analysis of Parton Structure of the Nucleon: CTEQ5 Parton Distributions. *Eur. Phys. J*, C 12:375–392, 2000.

[26] H1 Collaboration. A Measurement and QCD Analysis of the Proton Structure Function $F_2(x, Q^2)$ at HERA. *Nucl. Phys.*, B 470:3–38, 1996. DESY Report: DESY-96-039.

[27] ZEUS Collaboration. Measurement of the F_2 structure function in deep inelastic e^+p scattering using 1994 data from the ZEUS detector at HERA. *Z. Phys.*, C 72:399–424, 1996. DESY Report: DESY-96-076.

[28] ZEUS Collaboration. Measurement of the proton structure function F_2 at low x and low Q^2 at HERA. *Z. Phys.*, C 69:609–620, 1996. DESY Report: DESY-95-193.

[29] H1 Collaboration. A Measurement of the Proton Structure Function $F_2(x, Q^2)$ at low x and low Q^2 at HERA. *Nucl. Phys.*, B 497:3–28, 1997. DESY Report: DESY-97-042.

[30] The ZEUS Collaboration. ZEUS results on the measurement and phenomenology of F_2 at low x and low Q^2. *Eur. Phys. J.*, C 7:609–630, 1999. DESY Report: DESY-98-121.

[31] S.D. Drell and T-M. Yan. Massive lepton-pair production in hadron-hadron collisions at high energies. *Phys. Rev. Lett.*, 25(5):316–320, 1970.

[32] J. Pumplin et al. New generation of parton distributions with uncertainties from global QCD analysis. *J. High Energy Phys. JHEP*, 07:012, 2002.

[33] The ZEUS Collaboration. Measurement of the neutral current cross section and F_2 structure function for deep inelastic e^+p scattering at HERA. *Eur. Phys. J.*, C 21:443–471, 2001.

[34] H1 Collaboration. Measurement of Neutral and Charged Current Cross Sections in Positron-Proton Collisions at Large Momentum Transfer. *Eur. Phys. J*, C 13:609–639, 2000. DESY Report: DESY-99-107.

[35] The DØ Collaboration. Ratio of Jet Cross Sections at $\sqrt{s} = 630$ GeV and 1800 GeV. *Phys. Rev. Lett.*, 86(12):2523–2527, 2001.

[36] A.D. Martin et al. MRST2001: partons and α_S from precise deep inelastic scattering and Tevatron jet data. *Eur. Phys. J*, C 23:73 – 87, 2002.

[37] H1 Collaboration. Measurement of Neutral and Charged Current Cross Sections in Electron-Proton Collisions at High-Q^2. *Eur. Phys. J*, C 19:269–288, 2001.

[38] The ZEUS Collaboration. Measurement of the high-Q^2 charged-current e^+p deep inelastic scattering cross sections at HERA. *Eur. Phys. J.*, C 12:411–428, 2000.

[39] H1 Collaboration. Measurement of $D^{*\pm}$ meson production and F_2^c in deep-inelastic scattering at HERA. *Phys. Lett.*, B(528):199–214, 2002. DESY Report: DESY-01-105.

[40] The ZEUS Collaboration. Measurement of $D^{*\pm}$ production and of the charm contribution to F_2 in deep inelastic scattering at HERA. *Eur. Phys. J.*, C 12:35–52, 2000.

[41] CDF Collaboration. Measurement of the inclusive jet cross section in $p\overline{p}$ collisions at $\sqrt{s} = 1.8$ TeV. *Phys. Rev.*, D 64:032001, 2001.

[42] A.D. Martin et al. Uncertainties of predictions from parton distributions i: Experimental errors. hep-ph/0211080, November 2002. See also: http://durpdg.dur.ac.uk, viewed on 14 February 2003.

[43] M. Glück, E. Reya, and A. Vogt. Photonic parton distributions. *Phys. Rev.*, D 46(5):1973–1979, 1992.

[44] B.H. Wiik. HERA status. In Buchmüller and Ingelman [140], pages 1–16.

[45] ZEUS Collaboration. The ZEUS detector, status report. Technical report, DESY, 1993. Unpublished.

[46] Niels Tuning. *Proton Structure Functions at HERA*. PhD thesis, Amsterdam University, Amsterdam, September 2001.

[47] ZEUS Luminosity Group. Luminosity Measurement in the ZEUS Experiment. ZEUS-Note 01-004, DESY, March 2001.

[48] The ZEUS experiment web page. available at: http://www-zeus.desy.de.

[49] B. Foster et al. The design and construction of the ZEUS central tracking detector. *Nucl. Instr. Meth.*, A 338:254–283, 1994.

[50] G.P. Heath et al. The ZEUS first level tracking trigger. *Nucl. Instr. Meth.*, A 315:431–435, 1992.

[51] N. Harnew et al. Vertex triggering using time difference measurements in the ZEUS central tracking detector. *Nucl. Instr. Meth.*, A 279:290–296, 1989.

[52] Arnulf Quadt and Olaf Ruske. A New Method to Measure Vertex Distributions. ZEUS Note 98-036, DESY, June 1998.

[53] A. Quadt et al. The design and performance of the ZEUS central tracking detector second-level trigger. *Nucl. Instr. Meth.*, A 438:472–501, 1999.

[54] R. Hall-Wilton et al. The CTD Tracking Resolution. ZEUS-Note 99-024, DESY, May 1999.

[55] Corinna Caterall and Frédérique Pelucchi. Comparison of the CTD and Calorimeter Measurement of the DIS Positron Angle in High Q^2 Events. ZEUS-Note 97-056, DESY, July 1997.

[56] G.F. Hartner. VCTRAK Briefing: Program and Math. ZEUS-Note 98-058, DESY, August 1998.

[57] A. Andersen et al. Construction and beam test of the ZEUS forward and rear calorimeter. *Nucl. Instr. Meth.*, A 309:101–142, 1991.

[58] M. Derrick et al. Design and construction of the ZEUS barrel calorimeter. *Nucl. Instr. Meth.*, A 309:77–100, 1991.

[59] B. Bernstein et al. Beam tests of the ZEUS barrel calorimeter. *Nucl. Instr. Meth.*, A 336:23–52, 1993.

[60] Fabiola Gianotti. Issues in calorimetry. In Prosper and Danilov [141], pages 313–387.

[61] A. Caldwell et al. The time calibration of the ZEUS Calorimeter. ZEUS Note 93-021, DESY, March 1993.

[62] A. Dwuraźni et al. Experimental study of electron-hadron separation in calorimeters using silicon diodes. *Nucl. Instr. Meth.*, A 277:176–186, 1989.

[63] ZEUS Collaboration. Measurement of elastic ω photoproduction at HERA. *Z. Phys.*, C 73:73–84, 1996.

[64] A.D. Tapper. *Measurement of charged current deep inelastic scattering cross sections using the ZEUS detector at HERA*. PhD thesis, Imperial College of Science, Technology and Medicine, University of London, London, 2001.

[65] A. Bamberger et al. The presampler for the forward and rear calorimeter in the ZEUS detector. *Nucl. Instr. Meth.*, A 382:419–429, 1996.

[66] Krystina Olkiewicz and Andrzej Eskreys. Luminosity Off-line Luminosity Calculation in the ZEUS Experiment in 1997, 1998 and 1999. ZEUS Note 99-044, DESY, August 1999.

[67] J.A. Crittenden et al. The C5 Upgrade. ZEUS Note 01-002, DESY, January 2001.

[68] R. Carlin et al. Experience with the ZEUS Trigger System. *Nucl. Phys.*, B Proc. Suppl. 44:430–434, 1995.

[69] W.A. Smith et al. The ZEUS Calorimeter First Level Trigger. *Nucl. Instr. Meth.*, A 355:278–294, 1995.

[70] B.G. Bylsma et al. A fast cluster finding processor for the ZEUS calorimeter. *Nucl. Instr. Meth.*, A 337:512–520, 1994.

[71] Michał Kasprzak and Radosław Pawlak. Muon Detection Efficiency of Veto Wall. ZEUS Note 94-108, DESY, September 1994.

[72] Frédéric Bénard, Robert S. Orr, Stefan Polenz, and David Simmons. The 1995 TLT Filter Software. ZEUS Note 95-164, DESY, December 1995.

[73] R. Brun et al. GEANT3. Technical Report CERN-DD/EE/84-1, CERN, 1987.

[74] Matthias Moritz. *Measurement of High Q^2 Neutral Current DIS Cross Section at HERA*. PhD thesis, Universität Hamburg, Hamburg, December 2001.

[75] G.A. Schuler and H. Spiesberger. DJANGO: The interface for the event generators HERACLES and LEPTO. In Buchmüller and Ingelman [140], pages 1419–1432.

[76] G. Ingelman, A. Edin, and J. Rathsman. LEPTO 6.5: A Monte Carlo generator for deep inelastic lepton-nucleon scattering. *Comp. Phys. Comm.*, 101:108, 1997.

[77] A. Kwiatkowski, H. Spiesberger, and Möring H.-J. HERACLES: An Event Generator for ep Interactions at HERA Including Radiative Processes - VERSION 4.1. In Buchmüller and Ingelman [140], pages 1294–1310.

[78] M. Bengtsson, G. Ingelman, and T. Sjöstrand. Parton Cascade Evolution and Event Structure at HERA. In R.D. Peccei, editor, *Proceedings of the HERA Workshop*, volume 1, pages 149–165. DESY, 1987.

[79] L. Lönnblad. ARIADNE version 4 – a program for simulation of QCD cascades implementing the colour dipole model. *Comp. Phys. Comm.*, 71:15, 1992.

[80] Torbjörn Sjöstrand. High-energy-physics event generation with PYTHIA 5.7 and JETSET 7.4. *Comp. Phys. Comm.*, 82:74–89, 1994.

[81] G. Grindhammer, D. Krücker, and G. Nellen. A Comparison of Energy Flows and Particle Spectra in Deep-Inelastic Scattering at HERA with Monte Carlo Models. In *Proceedings of the Workshop on Monte-Carlo generators for HERA physics* [142]. DESY-PROC-1999-12, available at: http://www.desy.de/∼heramc.

[82] Leif Lönnblad. The Colour-Dipole model and the ARIADNE program at high Q^2. In *Proceedings of the Workshop on Monte-Carlo generators for HERA physics* [142]. DESY-PROC-1999-12, available at: http://www.desy.de/∼heramc.

[83] N.H. Brook et al. A Comparison of Deep Inelastic Scattering Monte-Carlo Event Generators to HERA Data. In *Proceedings of the Workshop on Monte-Carlo generators for HERA physics* [142]. DESY-PROC-1999-12, available at: http://www.desy.de/∼heramc.

[84] ZEUS Collaboration. Observation of events with a large rapidity gap in deep inelastic scattering at HERA. *Phys. Lett.*, B 315:481–493, 1993.

[85] H1 Collaboration. Deep inelastic scattering events with a large rapidity gap at HERA. *Nucl. Phys.*, B 429:477–502, 1994.

[86] H. Jung. Hard diffractive scattering in high energy ep collisions and the Monte-Carlo generator RAPGAP. *Comp. Phys. Comm.*, 39:347, 1986.

[87] H1 Collaboration. Determination of the longitudinal proton structure function $F_L(x, Q^2)$ at low x. *Phys. Lett.*, B 393:452–464, 1997.

[88] H1 Collaboration. Deep-Inelastic Inclusive ep Scattering at Low x and a Determination of α_s. *Eur. Phys. J*, C 21:33–61, 2001.

[89] G. Marchesini, B.R. Webber, G. Abbiendi, I.G. Knowles, M.H. Seymour, and L. Stanco. Herwig 5.1 – a Monte Carlo event generator for simulating hadron emission reactions with interfering gluons. *Comp. Phys. Comm.*, 67:465 – 508, 1992.

[90] Alexander Kappes. *Measurement of $e^- p \to e^- X$ differential cross sections at high Q^2 and of the Structure Function xF_3 with ZEUS at HERA*. PhD thesis, Universität Bonn, Bonn, December 2001. BONN-IR-2001-16.

[91] Henning Schnurbusch. *Measurement of Proton Structure from high-Q^2 Neutral Current Events in Deep Inelastic Scattering at HERA*. PhD thesis, Universität Bonn, Bonn, September 2002.

[92] Stan Bentvelsen, Jos Engelen, and Paul Kooijman. Reconstruction of (x, Q^2) and extraction of structure functions in neutral current scattering at HERA. In Buchmüller and Ingelman [140], pages 23–41.

[93] F. Jacquet and A. Blondel. Proceedings of the study of an *ep* facility for Europe. In U. Amaldi, editor, *Proceedings of the study of an ep facility for Europe*, pages 391–394, 1979.

[94] Michael Wodarczyk. Calorimeter First Level Trigger Studies for a Lower Electron Threshold in 1996. ZEUS Note 95-163, DESY, December 1995.

[95] Michael Wodarczyk. *Measurement of the proton structure function F_2 at HERA using the 1996 and 1997 ZEUS data.* PhD thesis, University of Wisconsin, Madison, 1999.

[96] F. Chlebana. Description of the ZEUS Global Second Level Trigger in 1994. ZEUS Note 94-102, DESY, October 1995.

[97] G.F. Hartner et al. VCTRAK (3.07/04): Offline Output Information. ZEUS-Note 97-064, DESY, September 1997.

[98] Paul de Jong. Status of the Uranium Calorimeter Reconstruction Software, User Guide version 2.00. ZEUS Note 92-019, DESY, February 1992.

[99] A. Savin. Study of the Calorimeter Noise in the 1996 data. ZEUS Note 98-07, DESY, 1998.

[100] Mario Martinez. Calorimeter Noisy Cells Study for 1998-2000 ZEUS Data. ZEUS Note 00-016, DESY, 2000.

[101] Marc de Kamps. Changes and extensions of the calorimeter reconstruction programme. ZEUS Note 94-014, DESY, February 1994.

[102] Gennady Briskin. *Diffractive Dissociation in e^-p Deep Inelastic Scattering.* PhD thesis, Tel Aviv University, Tel Aviv, June 1998.

[103] Halina Habramowicz, Allen Caldwell, and Ralph Sinkus. Neural network based electron identification in the ZEUS calorimeter. *Nucl. Instr. Meth.*, A 365:508–517, 1995.

[104] T.C. Awes et al. A simple method of shower localization and identification in laterally segmented showers. *Nucl. Instr. Meth.*, A 311:130–138, 1992.

[105] P. Billoir and S. Qian. Fast vertex fitting with a local parametrization of tracks. *Nucl. Instr. Meth.*, A 311:139–150, 1992.

[106] R. Sinkus and T. Voss. Particle identification with neural networks using a rotational invariant moment representation. *Nucl. Instr. Meth.*, A 391:360–368, 1997.

[107] Max Born and Emil Wolf. *Principles of Optics.* Cambridge University Press, Cambridge, seventh (expanded) edition, 1991.

[108] F. Zernike. Beugungstheorie des Schneidenverfahrens und seiner verbesserten Form, der Phasenkontrastmethode. *Physica*, 1:689–704, 1943.

[109] A. López-Durán Viani and S. Schlensted. Electron finder efficiencies and impurities. a comparison between SINISTRA95, EM and EMNET. ZEUS-Note 99-077, DESY, January 1999.

[110] Amaya López-Durán Viani. *Measurement of the Proton Structure Function xF_3 in High Q^2 DIS Events Using ZEUS Data.* PhD thesis, Humboldt-Universität, Berlin, November 2001. DESY-THESIS-2001-056.

[111] Ch. Amelung. Electron Position Reconstruction in ZEUS: Another Update of the ELECPO Package (Based on 1994/96 Data). ZEUS Note 97-053, DESY, June 1997.

[112] Ch. Amelung. Electron Position Reconstruction in ZEUS: Further Update of the ELECPO Package (Based on 1995 Data). ZEUS Note 96-093, DESY, October 1996.

[113] Niels Tuning. ZUFOs: Hadronic final state reconstruction with calorimeter, tracking and backsplash correction. ZEUS Note 01-021, DESY, June 2001.

[114] Wuyi Liu. *Search for Resonances Decaying to e+jet and Limits on First Generation Leptoquarks in e^+p Interactions at HERA.* PhD thesis, Columbia University, New York, 2001.

[115] A. Caldwell, W. Liu, B. Mellado, and B. Straub. BCAL Electron Studies - Part II: Energy Scale Calibration. ZEUS Note 98-018, DESY, March 1998.

[116] Rolf Deffner. *Proton Structure Functions at HERA.* PhD thesis, Universität Bonn, Bonn, December 1999. BONN-IR-99-20.

[117] The ZEUS Collaboration. Measurement of the high-Q^2 neutral current e^+p deep inlastic scattering cross-sections at HERA. *Eur. Phys. J.*, C 11:427–445, 1999.

[118] J. Grosse-Knetter. Corrections for the Hadronic Final State. ZEUS Note 98-031, DESY, April 1998.

[119] J. Grosse-Knetter. Energy Correction for Islands. ZEUS Note 97-039, DESY, July 1997.

[120] The ZEUS Calorimeter Group. Response of a uranium-scintillator calorimeter to electrons, pions and protons in the momentum range 0.5-10 GeV/c. *Nucl. Instr. Meth.*, A 290:95–108, 1990.

[121] Alex Tapper. Calibration of the ZEUS Calorimeter for hadrons and jets. talk given at the Workshop on Energy Calibration of the ATLAS Calorimeter, 21-24 July 2002. Unpublished.

[122] ZEUS Collaboration. Measurement of the high-Q^2 charged current cross sections in e^-p deep inelastic scattering at HERA. *Phys. Lett.*, B 539:197–217, 2002. DESY Report: DESY-02-064.

[123] José Ricardo Gonçalo. Neutral current analysis status. Talk given at the ZEUS Collaboration meeting, London, June 2002. Unpublished.

[124] Matthias Moritz. Study: high-y and high-Q^2 (neutral current 99/00). Private communication.

[125] José Ricardo Gonçalo and Kenneth Long. A minimum bias vertex distribution for the 1998/1999 e^-p data. ZEUS Note 00-027, DESY, September 2000.

[126] José Ricardo Gonçalo. Simulation of the hadronic final state and systematic uncertainties in high-Q^2 NC DIS. ZEUS Note 02-004, DESY, March 2002.

[127] Arnulf Quadt and Olaf Ruske. A New Method to Measure Vertex Distributions. ZEUS-Note 98-036, DESY, June 1998.

[128] The H1 Collaboration. Energy flow in the hadronic final state of diffractive and non-diffractive deep-inelastic scattering at HERA. *Z. Phys.*, C 70:609–620, 1996.

[129] H1 Collaboration. Measurement of transverse energy flow in deep-inelastic scattering at HERA. *Eur. Phys. J.*, C 12:595–607, 2000.

[130] A. Bamberger et al. The ZEUS forward plug calorimeter with lead-scintillator plates and WLS fibel readout. *Nucl. Instr. Meth.*, A 450:235–252, 2000. DESY Report: DESY-99-194.

[131] Glen Cowan. *Statistical data analysis.* Oxford University Press, 1998.

[132] G. Zech. Comparing statistical data to monte carlo simulation – parameter fitting and unfolding. DESY Report DESY 95-113, DESY, June 1995.

[133] Gary J. Feldman and Robert D. Cousins. Unified approach to the classical statistical analysis of small signals. *Phys. Rev.*, D 57(7):3873–3889, 1998.

[134] ZEUS Collaboration. Measurement of the high-Q^2 neutral current cross sections in e^-p deep inelastic scattering at HERA. DESY report 02-113, DESY, August 2002. Available at: hep-ex/0208040. Accepted by Eur. Phys. J.

[135] NMC Collaboration, M. Arneodo, et al. Measurement of the proton and deuteron structure functions, F_2^p and F_2^d, and of the ratio σ_L/σ_T. *Nucl. Phys.*, B 483:3–43, 1997.

[136] BCDMS Collaboration, A.C. Benvenuti, et al. A high statistics measurement of the proton structure functions $F_2(x, Q^2)$ and R from deep inelastic scattering at high Q^2. *Phys. Lett.*, B 223:485–489, 1989.

[137] E665 Collaboration, M.R. Adams, et al. Proton and deuteron structure functions in muon scattering at 470 GeV. *Phys. Rev.*, D 54:3006, 1996.

[138] ZEUS Collaboration. Measurement of high-Q^2 neutral current cross sections in e^+p deep inelastic scattering at HERA. Article in preparation.

[139] R.J. Cashmore et al. Measurement of Weak Neutral Current Couplings of Quarks at HERA. In G. Ingelman, A. De Roeck, and R. Klanner, editors, *Proceedings of the Workshop on Future Physics at HERA*, volume 1. DESY, 1995/96.

[140] W. Buchmüller and G. Ingelman, editors. *Proceedings of the Workshop on Physics at HERA*. DESY, 1991.

[141] Harrison Prosper and Michael Danilov, editors. *Proceedings of the NATO Advanced Study Institute on Techniques and Concepts of High-Energy Physics*, number 566 in NATO Science Series *C*. Kluwer Academic Publishers, 2001.

[142] *Proceedings of the Workshop on Monte-Carlo generators for HERA physics*. DESY, 1998/99. DESY-PROC-1999-12, available at: http://www.desy.de/~heramc.

Appendix A

Tables of single-differential cross sections

Q^2 range (GeV^2)	Q_c^2 (GeV^2)	$d\sigma^{e^-p}/dQ^2$ (pb/GeV^2) measured	$d\sigma^{e^-p}/dQ^2$ (pb/GeV^2) SM	N_{obs}	N_{bg}	$\mathcal{P}$	$\mathcal{E}$
200. – 300.	250	$11.340 \pm 0.103\,^{+0.208}_{-0.198}$	11.220	15670	5.8	0.98	0.73
300. – 400.	350	$5.118 \pm 0.071\,^{+0.098}_{-0.091}$	5.022	6757	2.9	0.82	0.68
400. – 476.	440	$3.002 \pm 0.060\,^{+0.053}_{-0.065}$	2.890	2970	1.6	0.74	0.61
476. – 566.	520	$1.983 \pm 0.046\,^{+0.035}_{-0.042}$	1.924	2120	1.7	0.74	0.55
566. – 673.	620	$1.294 \pm 0.036\,^{+0.025}_{-0.031}$	1.251	1504	1.9	0.71	0.47
673. – 800.	730	$(8.83 \pm 0.25\,^{+0.21}_{-0.16}) \cdot 10^{-1}$	$8.36 \cdot 10^{-1}$	1413	1.8	0.73	0.56
800. – 951.	870	$(5.42 \pm 0.16\,^{+0.14}_{-0.12}) \cdot 10^{-1}$	$5.41 \cdot 10^{-1}$	1254	1.6	0.73	0.64
951. – 1131.	1040	$(3.42 \pm 0.12\,^{+0.06}_{-0.08}) \cdot 10^{-1}$	$3.47 \cdot 10^{-1}$	978	2.3	0.73	0.64
1131. – 1345.	1230	$(2.19 \pm 0.08\,^{+0.05}_{-0.04}) \cdot 10^{-1}$	$2.28 \cdot 10^{-1}$	754	2.3	0.73	0.67
1345. – 1600.	1470	$(1.58 \pm 0.06\,^{+0.03}_{-0.03}) \cdot 10^{-1}$	$1.45 \cdot 10^{-1}$	646	1.6	0.72	0.66
1600. – 1903.	1740	$(1.01 \pm 0.05\,^{+0.03}_{-0.02}) \cdot 10^{-1}$	$0.95 \cdot 10^{-1}$	496	2.6	0.73	0.66
1903. – 2263.	2100	$(5.60 \pm 0.31\,^{+0.19}_{-0.15}) \cdot 10^{-2}$	$5.86 \cdot 10^{-2}$	346	1.6	0.72	0.67
2263. – 2691.	2500	$(3.85 \pm 0.24\,^{+0.13}_{-0.08}) \cdot 10^{-2}$	$3.75 \cdot 10^{-2}$	281	1.7	0.73	0.67
2691. – 3200.	2900	$(2.63 \pm 0.19\,^{+0.05}_{-0.12}) \cdot 10^{-2}$	$2.56 \cdot 10^{-2}$	208	1.2	0.74	0.67
3200. – 4525.	3800	$(1.41 \pm 0.08\,^{+0.04}_{-0.04}) \cdot 10^{-2}$	$1.27 \cdot 10^{-2}$	312	1.3	0.83	0.76
4525. – 6400.	5400	$(5.06 \pm 0.41\,^{+0.14}_{-0.16}) \cdot 10^{-3}$	$5.08 \cdot 10^{-3}$	157	0.4	0.83	0.75
6400. – 9051.	7600	$(2.32 \pm 0.24\,^{+0.09}_{-0.09}) \cdot 10^{-3}$	$2.03 \cdot 10^{-3}$	95	0.4	0.84	0.72
9051. – 12800.	10800	$(6.5 \pm 1.1\,^{+0.4}_{-0.2}) \cdot 10^{-4}$	$7.6 \cdot 10^{-4}$	37	0.3	0.81	0.70
12800. – 18100.	15200	$(3.5 \pm 0.7\,^{+0.1}_{-0.2}) \cdot 10^{-4}$	$2.7 \cdot 10^{-4}$	27	0.1	0.81	0.69
18100. – 25600.	21500	$(9.1\,^{+3.5}_{-3.0}\,^{+1.3}_{-0.5}) \cdot 10^{-5}$	$8.3 \cdot 10^{-5}$	10	0.1	0.79	0.69
25600. – 36200.	30400	$(1.4\,^{+1.6}_{-0.9}\,^{+0.1}_{-0.1}) \cdot 10^{-5}$	$2.0 \cdot 10^{-5}$	2	0.0	0.77	0.68
36200. – 51200.	43100	$(1.1\,^{+1.3}_{-0.7}\,^{+0.0}_{-0.0}) \cdot 10^{-5}$	$0.3 \cdot 10^{-5}$	2	0.0	0.76	0.68

Table A.1: Single differential cross section $d\sigma^{e^-p}/dQ^2$. The following quantities are given for each bin: the lower and upper limits of Q^2, the bin centre, Q_c^2, the measured cross section, $d\sigma^{e^-p}/dQ^2$, at the Born level, the SM prediction calculated using CTEQ5D PDFs, the number of events observed, N_{obs}, the estimated number of background events, N_{bg}, the purity, $\mathcal{P}$, and the efficiency, $\mathcal{E}$, estimated for the bin. The first error of the measured cross section corresponds to the statistical uncertainty and the second one to the systematic uncertainty.

y range	y	$d\sigma^{e^-p}/dy$ (pb) , $Q^2 > 200\text{GeV}^2$		N_{obs}	N_{bg}	$\mathcal{P}$	$\mathcal{E}$
		measured	SM				
0.0500 – 0.1000	0.0750	$(7.569 \pm 0.110\,^{+0.120}_{-0.122}) \cdot 10^3$	$7.502 \cdot 10^3$	5869	0.3	0.67	0.62
0.1000 – 0.1500	0.1250	$(5.327 \pm 0.094\,^{+0.095}_{-0.104}) \cdot 10^3$	$5.212 \cdot 10^3$	4012	0.5	0.63	0.52
0.1500 – 0.2000	0.1750	$(4.197 \pm 0.085\,^{+0.081}_{-0.102}) \cdot 10^3$	$4.022 \cdot 10^3$	3041	0.5	0.60	0.46
0.2000 – 0.2500	0.2250	$(3.429 \pm 0.077\,^{+0.072}_{-0.069}) \cdot 10^3$	$3.275 \cdot 10^3$	2471	0.4	0.57	0.42
0.2500 – 0.3000	0.2750	$(2.743 \pm 0.068\,^{+0.067}_{-0.062}) \cdot 10^3$	$2.753 \cdot 10^3$	1986	0.5	0.55	0.40
0.3000 – 0.3500	0.3250	$(2.507 \pm 0.066\,^{+0.056}_{-0.061}) \cdot 10^3$	$2.363 \cdot 10^3$	1795	1.3	0.52	0.37
0.3500 – 0.4000	0.3750	$(2.144 \pm 0.060\,^{+0.049}_{-0.052}) \cdot 10^3$	$2.060 \cdot 10^3$	1582	1.7	0.52	0.36
0.4000 – 0.4500	0.4250	$(1.843 \pm 0.056\,^{+0.056}_{-0.056}) \cdot 10^3$	$1.818 \cdot 10^3$	1327	2.1	0.51	0.34
0.4500 – 0.5000	0.4750	$(1.714 \pm 0.054\,^{+0.045}_{-0.044}) \cdot 10^3$	$1.619 \cdot 10^3$	1246	2.3	0.50	0.33
0.5000 – 0.5500	0.5250	$(1.479 \pm 0.050\,^{+0.046}_{-0.041}) \cdot 10^3$	$1.454 \cdot 10^3$	1048	2.6	0.51	0.31
0.5500 – 0.6000	0.5750	$(1.361 \pm 0.049\,^{+0.044}_{-0.038}) \cdot 10^3$	$1.314 \cdot 10^3$	942	2.4	0.51	0.31
0.6000 – 0.6500	0.6250	$(1.227 \pm 0.049\,^{+0.046}_{-0.048}) \cdot 10^3$	$1.196 \cdot 10^3$	766	2.5	0.53	0.29
0.6500 – 0.7000	0.6750	$(1.175 \pm 0.050\,^{+0.053}_{-0.045}) \cdot 10^3$	$1.095 \cdot 10^3$	669	1.9	0.56	0.27
0.7000 – 0.7500	0.7250	$(9.376 \pm 0.470\,^{+0.532}_{-0.373}) \cdot 10^2$	$10.080 \cdot 10^2$	462	2.6	0.59	0.31
0.7500 – 0.8000	0.7750	$(8.959 \pm 0.542\,^{+0.381}_{-0.375}) \cdot 10^2$	$9.340 \cdot 10^2$	310	1.8	0.59	0.36
0.8000 – 0.8500	0.8250	$(9.718 \pm 0.683\,^{+0.594}_{-0.862}) \cdot 10^2$	$8.699 \cdot 10^2$	230	2.4	0.63	0.37
0.8500 – 0.9000	0.8750	$(7.810 \pm 0.713\,^{+0.000}_{-0.000}) \cdot 10^2$	$8.147 \cdot 10^2$	135	3.4	0.70	0.38

Table A.2: Single differential cross section $d\sigma^{e^-p}/dy$ for $Q^2 > 200$ GeV^2. The following quantities are given for each bin: the lower and upper limits of y, the bin centre, y_c, the measured cross section, $d\sigma^{e^-p}/dy$, at the Born level, the SM prediction calculated using CTEQ5D PDFs, the number of events observed, N_{obs}, the estimated number of background events, N_{bg}, the purity, $\mathcal{P}$, and the efficiency, $\mathcal{E}$, estimated for the bin. The first error of the measured cross section corresponds to the statistical uncertainty and the second one to the systematic uncertainty.

y range	y	$d\sigma^{e^+p}/dy$ (pb) , $Q^2 > 10000$ GeV^2		N_{obs}	N_{bg}	$\mathcal{P}$	$\mathcal{E}$
		measured	SM				
0.2000 – 0.3000	0.2500	$2.991\,^{+1.685}_{-1.351}\,^{+0.608}_{-0.105}$	2.775	5	0.0	0.78	0.89
0.3000 – 0.4000	0.3500	$6.187\,^{+2.364}_{-2.000}\,^{+0.241}_{-0.620}$	4.770	10	0.0	0.81	0.78
0.4000 – 0.5000	0.4500	$7.072\,^{+2.468}_{-2.063}\,^{+0.711}_{-0.278}$	5.679	11	0.0	0.83	0.69
0.5000 – 0.6000	0.5500	$6.799\,^{+2.596}_{-2.196}\,^{+0.625}_{-0.572}$	5.967	10	0.0	0.84	0.65
0.6000 – 0.7000	0.6500	$5.772\,^{+2.441}_{-1.723}\,^{+0.287}_{-0.581}$	5.952	9	0.0	0.85	0.68
0.7000 – 0.8000	0.7500	$8.415 \pm 2.362\,^{+0.858}_{-0.369}$	5.803	13	0.1	0.89	0.67
0.8000 – 0.9000	0.8500	$4.385\,^{+2.162}_{-1.803}\,^{+0.427}_{-0.260}$	5.607	7	0.3	0.93	0.63
0.9000 – 1.0000	0.9500	$8.789\,^{+6.119}_{-3.666}\,^{+0.713}_{-8.789}$	5.407	4	0.0	0.99	0.36

Table A.3: Single differential cross section $d\sigma^{e^-p}/dy$ for $Q^2 > 10000$ GeV^2. The following quantities are given for each bin: the lower and upper limits of y, the bin centre, y_c, the measured cross section, $d\sigma^{e^-p}/dy$, at the Born level, the SM prediction calculated using CTEQ5D PDFs, the number of events observed, N_{obs}, the estimated number of background events, N_{bg}, the purity, $\mathcal{P}$, and the efficiency, $\mathcal{E}$, estimated for the bin. The first error of the measured cross section corresponds to the statistical uncertainty and the second one to the systematic uncertainty.

x range	x	$d\sigma^{e-p}/dx$ (pb) , $Q^2 > 200$ GeV^2		N_{obs}	N_{bg}	$\mathcal{P}$	$\mathcal{E}$
		measured	SM				
0.0006 – 0.0010	0.0079	$(8.217 \pm 0.108\,^{+0.216}_{-0.204}) \cdot 10^4$	$8.019 \cdot 10^4$	7378	11.1	0.95	0.54
0.0100 – 0.0160	0.0126	$(5.708 \pm 0.092\,^{+0.127}_{-0.153}) \cdot 10^4$	$5.399 \cdot 10^4$	4778	7.1	0.71	0.49
0.0160 – 0.0250	0.0200	$(3.500 \pm 0.057\,^{+0.065}_{-0.062}) \cdot 10^4$	$3.379 \cdot 10^4$	4616	6.4	0.66	0.51
0.0250 – 0.0400	0.0316	$(2.069 \pm 0.034\,^{+0.035}_{-0.033}) \cdot 10^4$	$2.019 \cdot 10^4$	4653	4.4	0.64	0.53
0.0400 – 0.0630	0.0501	$(1.170 \pm 0.020\,^{+0.020}_{-0.016}) \cdot 10^4$	$1.173 \cdot 10^4$	4119	2.3	0.60	0.55
0.0630 – 0.1000	0.0794	$(6.501 \pm 0.116\,^{+0.156}_{-0.106}) \cdot 10^3$	$6.653 \cdot 10^3$	3806	0.9	0.57	0.58
0.1000 – 0.1600	0.1260	$(3.623 \pm 0.067\,^{+0.090}_{-0.062}) \cdot 10^3$	$3.664 \cdot 10^3$	3575	0.4	0.53	0.61
0.1600 – 0.2500	0.2000	$(1.956 \pm 0.044\,^{+0.042}_{-0.041}) \cdot 10^3$	$1.892 \cdot 10^3$	2389	0.3	0.50	0.54
0.2500 – 0.4000	0.3160	$(8.635 \pm 0.363\,^{+0.182}_{-0.484}) \cdot 10^2$	$8.155 \cdot 10^2$	629	0.0	0.53	0.34
0.4000 – 0.6300	0.5010	$(2.086 \pm 0.216\,^{+0.195}_{-0.126}) \cdot 10^2$	$2.010 \cdot 10^2$	98	0.0	0.51	0.41

Table A.4: Single differential cross section $d\sigma^{e-p}/dx$ for $Q^2 > 200$ GeV^2. The following quantities are given for each bin: the lower and upper limits of x, the bin centre, x_c, the measured cross section, $d\sigma^{e-p}/dx$, at the Born level, the SM prediction calculated using CTEQ5D PDFs, the number of events observed, N_{obs}, the estimated number of background events, N_{bg}, the purity, $\mathcal{P}$, and the efficiency, $\mathcal{E}$, estimated for the bin. The first error of the measured cross section corresponds to the statistical uncertainty and the second one to the systematic uncertainty.

x range	x	$d\sigma^{e-p}/dx$ (pb) , $Q^2 > 10000$ GeV^2		N_{obs}	N_{bg}	$\mathcal{P}$	$\mathcal{E}$
		measured	SM				
0.1000 – 0.1600	0.1260	$7.520\ ^{+4.393}_{-3.521}\ ^{+1.730}_{-2.407}$	13.200	5	0.2	0.93	0.47
0.1600 – 0.2500	0.2000	$(2.319 \pm 0.444\,^{+0.089}_{-0.154}) \cdot 10^1$	$1.662 \cdot 10^1$	28	0.2	0.88	0.62
0.2500 – 0.4000	0.3160	$(1.118 \pm 0.224\,^{+0.083}_{-0.048}) \cdot 10^1$	$0.959 \cdot 10^1$	25	0.0	0.85	0.73
0.4000 – 0.6300	0.5010	$3.011\ ^{+1.051}_{-0.879}\ ^{+0.598}_{-0.271}$	2.525	11	0.0	0.75	0.80

Table A.5: Single differential cross section $d\sigma^{e-p}/dx$ for $Q^2 > 10000$ GeV^2. The following quantities are given for each bin: the lower and upper limits of x, the bin centre, x_c, the measured cross section, $d\sigma^{e-p}/dx$, at the Born level, the SM prediction calculated using CTEQ5D PDFs, the number of events observed, N_{obs}, the estimated number of background events, N_{bg}, the purity, $\mathcal{P}$, and the efficiency, $\mathcal{E}$, estimated for the bin. The first error of the measured cross section corresponds to the statistical uncertainty and the second one to the systematic uncertainty.

Q^2 range (GeV^2)	Q_c^2 (GeV^2)	$d\sigma^{e+p}/dQ^2$ ($\mathrm{pb}/\mathrm{GeV}^2$) measured	$d\sigma^{e+p}/dQ^2$ ($\mathrm{pb}/\mathrm{GeV}^2$) SM	N_{obs}	N_{bg}	$\mathcal{P}$	$\mathcal{E}$
200. – 300.	250	$11.540 \pm 0.058\,^{+0.100}_{-0.104}$	11.200	61820	21.8	0.98	0.72
300. – 400.	350	$5.052 \pm 0.038\,^{+0.043}_{-0.053}$	5.006	26110	11.6	0.82	0.68
400. – 476.	440	$2.945 \pm 0.032\,^{+0.021}_{-0.022}$	2.876	11140	6.4	0.74	0.59
476. – 566.	520	$1.960 \pm 0.025\,^{+0.019}_{-0.017}$	1.912	8154	6.9	0.74	0.54
566. – 673.	620	$1.244 \pm 0.019\,^{+0.009}_{-0.012}$	1.240	5715	7.3	0.72	0.48
673. – 800.	730	$(8.51 \pm 0.13\,^{+0.07}_{-0.07}) \cdot 10^{-1}$	$8.27 \cdot 10^{-1}$	5489	6.8	0.72	0.56
800. – 951.	870	$(5.44 \pm 0.09\,^{+0.06}_{-0.05}) \cdot 10^{-1}$	$5.33 \cdot 10^{-1}$	4805	6.0	0.73	0.63
951. – 1131.	1040	$(3.56 \pm 0.07\,^{+0.04}_{-0.05}) \cdot 10^{-1}$	$3.39 \cdot 10^{-1}$	3875	8.7	0.73	0.63
1131. – 1345.	1230	$(2.27 \pm 0.05\,^{+0.03}_{-0.04}) \cdot 10^{-1}$	$2.21 \cdot 10^{-1}$	2967	8.9	0.72	0.65
1345. – 1600.	1470	$(1.44 \pm 0.03\,^{+0.01}_{-0.02}) \cdot 10^{-1}$	$1.40 \cdot 10^{-1}$	2288	6.5	0.73	0.65
1600. – 1903.	1740	$(9.50 \pm 0.24\,^{+0.15}_{-0.15}) \cdot 10^{-2}$	$9.00 \cdot 10^{-2}$	1792	10.1	0.73	0.66
1903. – 2263.	2100	$(5.62 \pm 0.16\,^{+0.08}_{-0.06}) \cdot 10^{-2}$	$5.48 \cdot 10^{-2}$	1327	6.2	0.73	0.66
2263. – 2691.	2500	$(3.73 \pm 0.12\,^{+0.06}_{-0.05}) \cdot 10^{-2}$	$3.43 \cdot 10^{-2}$	1046	6.4	0.73	0.66
2691. – 3200.	2900	$(2.40 \pm 0.09\,^{+0.04}_{-0.06}) \cdot 10^{-2}$	$2.29 \cdot 10^{-2}$	752	4.3	0.72	0.67
3200. – 4525.	3800	$(1.13 \pm 0.04\,^{+0.01}_{-0.02}) \cdot 10^{-2}$	$1.08 \cdot 10^{-2}$	948	4.7	0.82	0.74
4525. – 6400.	5400	$(3.86 \pm 0.18\,^{+0.05}_{-0.10}) \cdot 10^{-3}$	$3.91 \cdot 10^{-3}$	468	1.9	0.82	0.73
6400. – 9051.	7600	$(1.34 \pm 0.09\,^{+0.02}_{-0.03}) \cdot 10^{-3}$	$1.37 \cdot 10^{-3}$	220	1.4	0.82	0.71
9051. – 12800.	10800	$(4.58 \pm 0.46\,^{+0.17}_{-0.17}) \cdot 10^{-4}$	$4.28 \cdot 10^{-4}$	104	1.2	0.81	0.71
12800. – 18100.	15200	$(1.50 \pm 0.23\,^{+0.06}_{-0.09}) \cdot 10^{-4}$	$1.23 \cdot 10^{-4}$	45	0.6	0.79	0.67
18100. – 25600.	21500	$(2.14\,^{+0.86}_{-0.73}\,^{+0.24}_{-0.33}) \cdot 10^{-5}$	$2.92 \cdot 10^{-5}$	10	0.5	0.79	0.70
25600. – 36200.	30400	$(6.4\,^{+4.4}_{-2.7}\,^{+1.1}_{-0.2}) \cdot 10^{-6}$	$5.4 \cdot 10^{-6}$	4	0.0	0.75	0.67

Table A.6: Single differential cross section $d\sigma^{e+p}/dQ^2$. The following quantities are given for each bin: the lower and upper limits of Q^2, the bin centre, Q_c^2, the measured cross section, $d\sigma^{e+p}/dQ^2$, at the Born level, the SM prediction calculated using CTEQ5D PDFs, the number of events observed, N_{obs}, the estimated number of background events, N_{bg}, the purity, $\mathcal{P}$, and the efficiency, $\mathcal{E}$, estimated for the bin. The first error of the measured cross section corresponds to the statistical uncertainty and the second one to the systematic uncertainty.

y range	y	$d\sigma^{e+p}/dy\,(\mathrm{pb})$, $\mathrm{Q}^2 > 200$ GeV^2		N_{obs}	N_{bg}	$\mathcal{P}$	$\mathcal{E}$
		measured	SM				
0.0500 – 0.1000	0.0750	$(7.554 \pm 0.060\,^{+0.092}_{-0.071}) \cdot 10^3$	$7.465 \cdot 10^3$	22830	1.2	0.67	0.61
0.1000 – 0.1500	0.1250	$(5.354 \pm 0.052\,^{+0.034}_{-0.026}) \cdot 10^3$	$5.174 \cdot 10^3$	15580	1.8	0.63	0.52
0.1500 – 0.2000	0.1750	$(4.149 \pm 0.046\,^{+0.027}_{-0.040}) \cdot 10^3$	$3.985 \cdot 10^3$	11830	1.7	0.59	0.46
0.2000 – 0.2500	0.2250	$(3.319 \pm 0.041\,^{+0.034}_{-0.057}) \cdot 10^3$	$3.239 \cdot 10^3$	9313	1.6	0.57	0.42
0.2500 – 0.3000	0.2750	$(2.803 \pm 0.038\,^{+0.024}_{-0.025}) \cdot 10^3$	$2.717 \cdot 10^3$	7691	1.8	0.55	0.38
0.3000 – 0.3500	0.3250	$(2.429 \pm 0.035\,^{+0.019}_{-0.029}) \cdot 10^3$	$2.329 \cdot 10^3$	6717	5.0	0.53	0.37
0.3500 – 0.4000	0.3750	$(2.078 \pm 0.032\,^{+0.029}_{-0.023}) \cdot 10^3$	$2.028 \cdot 10^3$	5918	6.5	0.52	0.36
0.4000 – 0.4500	0.4250	$(1.857 \pm 0.030\,^{+0.021}_{-0.025}) \cdot 10^3$	$1.787 \cdot 10^3$	5249	8.0	0.50	0.34
0.4500 – 0.5000	0.4750	$(1.686 \pm 0.029\,^{+0.021}_{-0.034}) \cdot 10^3$	$1.589 \cdot 10^3$	4676	8.9	0.49	0.32
0.5000 – 0.5500	0.5250	$(1.482 \pm 0.027\,^{+0.023}_{-0.027}) \cdot 10^3$	$1.425 \cdot 10^3$	4064	10.3	0.50	0.31
0.5500 – 0.6000	0.5750	$(1.285 \pm 0.026\,^{+0.017}_{-0.016}) \cdot 10^3$	$1.287 \cdot 10^3$	3458	9.2	0.50	0.30
0.6000 – 0.6500	0.6250	$(1.195 \pm 0.026\,^{+0.021}_{-0.025}) \cdot 10^3$	$1.170 \cdot 10^3$	2939	10.0	0.51	0.28
0.6500 – 0.7000	0.6750	$(1.122 \pm 0.026\,^{+0.021}_{-0.029}) \cdot 10^3$	$1.071 \cdot 10^3$	2546	7.8	0.55	0.27
0.7000 – 0.7500	0.7250	$(9.543 \pm 0.259\,^{+0.297}_{-0.295}) \cdot 10^2$	$9.849 \cdot 10^2$	1788	10.3	0.58	0.30
0.7500 – 0.8000	0.7750	$(9.445 \pm 0.305\,^{+0.400}_{-0.422}) \cdot 10^2$	$9.116 \cdot 10^2$	1230	6.9	0.59	0.35
0.8000 – 0.8500	0.8250	$(8.605 \pm 0.350\,^{+0.343}_{-0.295}) \cdot 10^2$	$8.487 \cdot 10^2$	749	9.7	0.63	0.36
0.8500 – 0.9000	0.8750	$(8.330 \pm 0.411\,^{+0.592}_{-0.551}) \cdot 10^2$	$7.945 \cdot 10^2$	508	13.1	0.71	0.35

Table A.7: Single differential cross section $d\sigma^{e+p}/dy$ for $Q^2 > 200$ GeV^2. The following quantities are given for each bin: the lower and upper limits of y, the bin centre, y_c, the measured cross section, $d\sigma^{e+p}/dy$, at the Born level, the SM prediction calculated using CTEQ5D PDFs, the number of events observed, N_{obs}, the estimated number of background events, N_{bg}, the purity, $\mathcal{P}$, and the efficiency, $\mathcal{E}$, estimated for the bin. The first error of the measured cross section corresponds to the statistical uncertainty and the second one to the systematic uncertainty.

y range	y	$d\sigma^{e+p}/dy\,(\mathrm{pb})$, $\mathrm{Q}^2 > 10000$ GeV^2		N_{obs}	N_{bg}	$\mathcal{P}$	$\mathcal{E}$
		measured	SM				
0.1000 – 0.2000	0.1500	$0.09\,^{+0.15}_{-0.05}\,^{+0.00}_{-0.00}$	0.21	1	0.0	0.68	0.89
0.2000 – 0.3000	0.2500	$2.28 \pm 0.61\,^{+0.04}_{-0.21}$	2.02	14	0.0	0.81	0.84
0.3000 – 0.4000	0.3500	$3.82 \pm 0.79\,^{+0.04}_{-0.16}$	3.01	24	0.0	0.81	0.76
0.4000 – 0.5000	0.4500	$2.80 \pm 0.69\,^{+0.09}_{-0.28}$	3.11	17	0.1	0.83	0.69
0.5000 – 0.6000	0.5500	$2.15\,^{+0.77}_{-0.57}\,^{+0.26}_{-0.35}$	2.86	12	0.0	0.83	0.63
0.6000 – 0.7000	0.6500	$2.73 \pm 0.67\,^{+0.15}_{-0.09}$	2.53	17	0.1	0.87	0.65
0.7000 – 0.8000	0.7500	$2.51 \pm 0.65\,^{+0.06}_{-0.34}$	2.23	15	0.5	0.90	0.64
0.8000 – 0.9000	0.8500	$2.60 \pm 0.72\,^{+0.27}_{-0.36}$	2.03	15	1.4	0.93	0.54
0.9000 – 1.0000	0.9500	$4.26\,^{+1.94}_{-1.46}\,^{+3.62}_{-0.30}$	1.92	9	0.0	0.97	0.28

Table A.8: Single differential cross section $d\sigma^{e+p}/dy$ for $Q^2 > 10000$ GeV^2. The following quantities are given for each bin: the lower and upper limits of y, the bin centre, y_c, the measured cross section, $d\sigma^{e+p}/dy$, at the Born level, the SM prediction calculated using CTEQ5D PDFs, the number of events observed, N_{obs}, the estimated number of background events, N_{bg}, the purity, $\mathcal{P}$, and the efficiency, $\mathcal{E}$, estimated for the bin. The first error of the measured cross section corresponds to the statistical uncertainty and the second one to the systematic uncertainty.

x range	x	$d\sigma^{e+p}/dx$ (pb) , $Q^2 > 200$ GeV^2		$N_{\rm obs}$	$N_{\rm bg}$	$\mathcal{P}$	$\mathcal{E}$
		measured	SM				
0.0063 – 0.0010	0.0079	$(8.255 \pm 0.059\,^{+0.106}_{-0.122}) \cdot 10^4$	$7.978 \cdot 10^4$	28920	43.3	0.95	0.54
0.0100 – 0.0160	0.0126	$(5.537 \pm 0.049\,^{+0.062}_{-0.067}) \cdot 10^4$	$5.362 \cdot 10^4$	18200	27.5	0.70	0.48
0.0160 – 0.0250	0.0200	$(3.455 \pm 0.032\,^{+0.028}_{-0.023}) \cdot 10^4$	$3.349 \cdot 10^4$	17630	25.2	0.66	0.50
0.0250 – 0.0400	0.0316	$(2.072 \pm 0.019\,^{+0.012}_{-0.017}) \cdot 10^4$	$1.996 \cdot 10^4$	18000	16.6	0.64	0.52
0.0400 – 0.0630	0.0501	$(1.176 \pm 0.011\,^{+0.008}_{-0.008}) \cdot 10^4$	$1.156 \cdot 10^4$	15820	8.4	0.60	0.53
0.0630 – 0.1000	0.0794	$(6.596 \pm 0.064\,^{+0.071}_{-0.048}) \cdot 10^3$	$6.542 \cdot 10^3$	15080	3.6	0.57	0.57
0.1000 – 0.1600	0.1260	$(3.563 \pm 0.036\,^{+0.073}_{-0.039}) \cdot 10^3$	$3.598 \cdot 10^3$	13770	1.9	0.52	0.60
0.1600 – 0.2500	0.2000	$(1.909 \pm 0.024\,^{+0.034}_{-0.041}) \cdot 10^3$	$1.859 \cdot 10^3$	8942	1.3	0.49	0.52
0.2500 – 0.4000	0.3160	$(8.679 \pm 0.197\,^{+0.130}_{-0.197}) \cdot 10^2$	$8.035 \cdot 10^2$	2356	0.2	0.52	0.31
0.4000 – 0.6300	0.5010	$(2.011 \pm 0.115\,^{+0.041}_{-0.131}) \cdot 10^2$	$1.988 \cdot 10^2$	339	0.1	0.50	0.38

Table A.9: Single differential cross section $d\sigma^{e+p}/dx$ for $Q^2 > 200$ GeV^2. The following quantities are given for each bin: the lower and upper limits of x, the bin centre, x_c, the measured cross section, $d\sigma^{e+p}/dx$, at the Born level, the SM prediction calculated using CTEQ5D PDFs, the number of events observed, $N_{\rm obs}$, the estimated number of background events, $N_{\rm bg}$, the purity, $\mathcal{P}$, and the efficiency, $\mathcal{E}$, estimated for the bin. The first error of the measured cross section corresponds to the statistical uncertainty and the second one to the systematic uncertainty.

x range	x	$d\sigma^{e+p}/dx$ (pb) , $Q^2 > 10000$ GeV^2		$N_{\rm obs}$	$N_{\rm bg}$	$\mathcal{P}$	$\mathcal{E}$
		measured	SM				
0.1000 – 0.1600	0.1260	$8.122 \pm 1.907\,^{+0.362}_{-1.003}$	5.993	20	0.9	0.93	0.42
0.1600 – 0.2500	0.2000	$7.951 \pm 1.298\,^{+0.301}_{-0.256}$	7.220	39	1.0	0.88	0.60
0.2500 – 0.4000	0.3160	$5.071 \pm 0.761\,^{+0.121}_{-0.541}$	4.621	45	0.1	0.85	0.73
0.4000 – 0.6300	0.5010	$1.360 \pm 0.315\,^{+0.018}_{-0.095}$	1.431	19	0.1	0.76	0.81
0.6300 – 1.0000	0.7940	$0.018\,^{+0.031}_{-0.011}\,^{+0.001}_{-0.001}$	0.040	1	0.0	0.52	0.79

Table A.10: Single differential cross section $d\sigma^{e+p}/dx$ for $Q^2 > 10000$ GeV^2. The following quantities are given for each bin: the lower and upper limits of x, the bin centre, x_c, the measured cross section, $d\sigma^{e+p}/dx$, at the Born level, the SM prediction calculated using CTEQ5D PDFs, the number of events observed, $N_{\rm obs}$, the estimated number of background events, $N_{\rm bg}$, the purity, $\mathcal{P}$, and the efficiency, $\mathcal{E}$, estimated for the bin. The first error of the measured cross section corresponds to the statistical uncertainty and the second one to the systematic uncertainty.

Appendix B

Tables of double-differential reduced cross sections

Q^2 range (GeV^2)	x range	Q_c^2 (GeV^2)	x_c	$\tilde{\sigma}(e^-p)$ measured	$\tilde{\sigma}(e^-p)$ SM	$N_{\rm obs}$	$N_{\rm bg}$	$\mathcal{P}$	$\mathcal{E}$
185 – 240	$(0.37 - 0.60)\cdot 10^{-2}$	200	$0.50\cdot 10^{-2}$	$1.123 \pm 0.030\,^{+0.015}_{-0.029}$	1.105	1817	2.45	0.73	0.41
	$(0.60 - 1.00)\cdot 10^{-2}$		$0.80\cdot 10^{-2}$	$(9.57 \pm 0.25\,^{+0.12}_{-0.23})\cdot 10^{-1}$	$9.43\cdot 10^{-1}$	1884	0.61	0.72	0.46
	$(0.10 - 0.17)\cdot 10^{-1}$		$0.13\cdot 10^{-1}$	$(8.11 \pm 0.22\,^{+0.06}_{-0.12})\cdot 10^{-1}$	$7.99\cdot 10^{-1}$	1807	0.23	0.71	0.49
	$(0.17 - 0.25)\cdot 10^{-1}$		$0.21\cdot 10^{-1}$	$(6.87 \pm 0.22\,^{+0.16}_{-0.10})\cdot 10^{-1}$	$6.78\cdot 10^{-1}$	1233	0.09	0.59	0.43
	$(0.25 - 0.37)\cdot 10^{-1}$		$0.32\cdot 10^{-1}$	$(6.06 \pm 0.20\,^{+0.07}_{-0.11})\cdot 10^{-1}$	$5.90\cdot 10^{-1}$	1173	0.05	0.55	0.45
	$(0.37 - 0.60)\cdot 10^{-1}$		$0.50\cdot 10^{-1}$	$(4.92 \pm 0.16\,^{+0.13}_{-0.07})\cdot 10^{-1}$	$5.12\cdot 10^{-1}$	1259	0.02	0.57	0.49
	$(0.60 - 1.20)\cdot 10^{-1}$		$0.80\cdot 10^{-1}$	$(4.15 \pm 0.12\,^{+0.14}_{-0.07})\cdot 10^{-1}$	$4.42\cdot 10^{-1}$	1627	0.03	0.58	0.62
	$0.12 - 0.25$		0.18	$(3.14 \pm 0.11\,^{+0.12}_{-0.14})\cdot 10^{-1}$	$3.28\cdot 10^{-1}$	968	0.00	0.53	0.48
240 – 310	$(0.37 - 0.60)\cdot 10^{-2}$	250	$0.50\cdot 10^{-2}$	$1.08 \pm 0.04\,^{+0.02}_{-0.04}$	1.12	874	1.65	0.58	0.30
	$(0.60 - 1.00)\cdot 10^{-2}$		$0.80\cdot 10^{-2}$	$(9.66 \pm 0.30\,^{+0.25}_{-0.14})\cdot 10^{-1}$	$9.64\cdot 10^{-1}$	1385	1.00	0.62	0.48
	$(0.10 - 0.17)\cdot 10^{-1}$		$0.13\cdot 10^{-1}$	$(8.45 \pm 0.26\,^{+0.18}_{-0.19})\cdot 10^{-1}$	$8.15\cdot 10^{-1}$	1357	0.18	0.63	0.50
	$(0.17 - 0.25)\cdot 10^{-1}$		$0.21\cdot 10^{-1}$	$(6.95 \pm 0.26\,^{+0.08}_{-0.10})\cdot 10^{-1}$	$6.89\cdot 10^{-1}$	917	0.10	0.54	0.47
	$(0.25 - 0.37)\cdot 10^{-1}$		$0.32\cdot 10^{-1}$	$(5.96 \pm 0.23\,^{+0.18}_{-0.07})\cdot 10^{-1}$	$5.97\cdot 10^{-1}$	872	0.05	0.51	0.48
	$(0.37 - 0.60)\cdot 10^{-1}$		$0.50\cdot 10^{-1}$	$(5.15 \pm 0.19\,^{+0.12}_{-0.10})\cdot 10^{-1}$	$5.16\cdot 10^{-1}$	965	0.03	0.51	0.52
	$(0.60 - 1.20)\cdot 10^{-1}$		$0.80\cdot 10^{-1}$	$(4.34 \pm 0.14\,^{+0.13}_{-0.08})\cdot 10^{-1}$	$4.44\cdot 10^{-1}$	1213	0.00	0.53	0.62
	$0.12 - 0.25$		0.18	$(3.07 \pm 0.12\,^{+0.16}_{-0.11})\cdot 10^{-1}$	$3.26\cdot 10^{-1}$	862	0.00	0.50	0.61
310 – 410	$(0.60 - 1.00)\cdot 10^{-2}$	350	$0.80\cdot 10^{-2}$	$(9.87 \pm 0.39\,^{+0.15}_{-0.25})\cdot 10^{-1}$	$9.90\cdot 10^{-1}$	822	1.60	0.59	0.34
	$(0.10 - 0.17)\cdot 10^{-1}$		$0.13\cdot 10^{-1}$	$(8.51 \pm 0.30\,^{+0.11}_{-0.13})\cdot 10^{-1}$	$8.38\cdot 10^{-1}$	1022	0.14	0.67	0.50
	$(0.17 - 0.25)\cdot 10^{-1}$		$0.21\cdot 10^{-1}$	$(7.18 \pm 0.30\,^{+0.08}_{-0.13})\cdot 10^{-1}$	$7.06\cdot 10^{-1}$	752	0.09	0.57	0.48
	$(0.25 - 0.37)\cdot 10^{-1}$		$0.32\cdot 10^{-1}$	$(5.93 \pm 0.26\,^{+0.04}_{-0.17})\cdot 10^{-1}$	$6.08\cdot 10^{-1}$	690	0.12	0.54	0.50
	$(0.37 - 0.60)\cdot 10^{-1}$		$0.50\cdot 10^{-1}$	$(5.20 \pm 0.21\,^{+0.10}_{-0.07})\cdot 10^{-1}$	$5.23\cdot 10^{-1}$	770	0.02	0.55	0.55
	$(0.60 - 1.20)\cdot 10^{-1}$		$0.80\cdot 10^{-1}$	$(4.48 \pm 0.16\,^{+0.08}_{-0.06})\cdot 10^{-1}$	$4.46\cdot 10^{-1}$	996	0.02	0.57	0.66
	$0.12 - 0.25$		0.18	$(3.08 \pm 0.12\,^{+0.08}_{-0.06})\cdot 10^{-1}$	$3.23\cdot 10^{-1}$	806	0.00	0.52	0.72
410 – 530	$(0.60 - 1.00)\cdot 10^{-2}$	450	$0.80\cdot 10^{-2}$	$(9.81 \pm 0.42\,^{+0.27}_{-0.14})\cdot 10^{-1}$	$10.03\cdot 10^{-1}$	645	1.61	0.61	0.40
	$(0.10 - 0.17)\cdot 10^{-1}$		$0.13\cdot 10^{-1}$	$(8.64 \pm 0.42\,^{+0.12}_{-0.28})\cdot 10^{-1}$	$8.54\cdot 10^{-1}$	485	0.55	0.63	0.34
	$(0.17 - 0.25)\cdot 10^{-1}$		$0.21\cdot 10^{-1}$	$(7.17 \pm 0.39\,^{+0.18}_{-0.24})\cdot 10^{-1}$	$7.18\cdot 10^{-1}$	398	0.05	0.59	0.39
	$(0.25 - 0.37)\cdot 10^{-1}$		$0.32\cdot 10^{-1}$	$(6.32 \pm 0.33\,^{+0.05}_{-0.12})\cdot 10^{-1}$	$6.16\cdot 10^{-1}$	436	0.00	0.58	0.46
	$(0.37 - 0.60)\cdot 10^{-1}$		$0.50\cdot 10^{-1}$	$(5.25 \pm 0.24\,^{+0.12}_{-0.18})\cdot 10^{-1}$	$5.27\cdot 10^{-1}$	534	0.03	0.59	0.55
	$(0.60 - 1.00)\cdot 10^{-1}$		$0.80\cdot 10^{-1}$	$(4.46 \pm 0.21\,^{+0.06}_{-0.03})\cdot 10^{-1}$	$4.47\cdot 10^{-1}$	500	0.00	0.56	0.61
	$0.10 - 0.17$		0.13	$(3.99 \pm 0.20\,^{+0.07}_{-0.11})\cdot 10^{-1}$	$3.73\cdot 10^{-1}$	473	0.00	0.54	0.64
	$0.17 - 0.30$		0.25	$(2.86 \pm 0.16\,^{+0.05}_{-0.14})\cdot 10^{-1}$	$2.59\cdot 10^{-1}$	386	0.00	0.50	0.66
530 – 710	$(0.60 - 1.00)\cdot 10^{-2}$	650	$0.80\cdot 10^{-2}$	$(9.05 \pm 0.54\,^{+0.24}_{-0.24})\cdot 10^{-1}$	$10.00\cdot 10^{-1}$	324	0.95	0.61	0.37
	$(0.10 - 0.17)\cdot 10^{-1}$		$0.13\cdot 10^{-1}$	$(9.27 \pm 0.39\,^{+0.23}_{-0.15})\cdot 10^{-1}$	$8.73\cdot 10^{-1}$	652	1.52	0.67	0.51
	$(0.17 - 0.25)\cdot 10^{-1}$		$0.21\cdot 10^{-1}$	$(8.13 \pm 0.44\,^{+0.23}_{-0.11})\cdot 10^{-1}$	$7.35\cdot 10^{-1}$	394	0.36	0.63	0.43
	$(0.25 - 0.37)\cdot 10^{-1}$		$0.32\cdot 10^{-1}$	$(6.33 \pm 0.41\,^{+0.21}_{-0.11})\cdot 10^{-1}$	$6.29\cdot 10^{-1}$	272	0.07	0.61	0.36
	$(0.37 - 0.60)\cdot 10^{-1}$		$0.50\cdot 10^{-1}$	$(5.66 \pm 0.36\,^{+0.18}_{-0.08})\cdot 10^{-1}$	$5.35\cdot 10^{-1}$	293	0.03	0.60	0.35
	$(0.60 - 1.00)\cdot 10^{-1}$		$0.80\cdot 10^{-1}$	$(4.05 \pm 0.28\,^{+0.08}_{-0.10})\cdot 10^{-1}$	$4.51\cdot 10^{-1}$	236	0.00	0.59	0.38
	$0.10 - 0.17$		0.13	$(3.47 \pm 0.24\,^{+0.09}_{-0.08})\cdot 10^{-1}$	$3.72\cdot 10^{-1}$	242	0.00	0.54	0.43
	$0.17 - 0.30$		0.25	$(2.74 \pm 0.19\,^{+0.06}_{-0.08})\cdot 10^{-1}$	$2.55\cdot 10^{-1}$	241	0.00	0.50	0.50
710 – 900	$(0.90 - 1.70)\cdot 10^{-2}$	800	$1.30\cdot 10^{-2}$	$(8.40 \pm 0.45\,^{+0.14}_{-0.21})\cdot 10^{-1}$	$8.80\cdot 10^{-1}$	398	1.54	0.64	0.43
	$(0.17 - 0.25)\cdot 10^{-1}$		$0.21\cdot 10^{-1}$	$(8.29 \pm 0.53\,^{+0.07}_{-0.31})\cdot 10^{-1}$	$7.45\cdot 10^{-1}$	283	0.34	0.60	0.50
	$(0.25 - 0.37)\cdot 10^{-1}$		$0.32\cdot 10^{-1}$	$(6.01 \pm 0.41\,^{+0.31}_{-0.14})\cdot 10^{-1}$	$6.36\cdot 10^{-1}$	240	0.07	0.60	0.53
	$(0.37 - 0.60)\cdot 10^{-1}$		$0.50\cdot 10^{-1}$	$(5.72 \pm 0.36\,^{+0.17}_{-0.14})\cdot 10^{-1}$	$5.40\cdot 10^{-1}$	298	0.02	0.64	0.56
	$(0.60 - 1.00)\cdot 10^{-1}$		$0.80\cdot 10^{-1}$	$(4.66 \pm 0.32\,^{+0.06}_{-0.14})\cdot 10^{-1}$	$4.53\cdot 10^{-1}$	244	0.10	0.61	0.55
	$0.10 - 0.17$		0.13	$(3.44 \pm 0.26\,^{+0.14}_{-0.08})\cdot 10^{-1}$	$3.73\cdot 10^{-1}$	192	0.00	0.57	0.57
	$0.17 - 0.30$		0.25	$(2.63 \pm 0.24\,^{+0.06}_{-0.14})\cdot 10^{-1}$	$2.53\cdot 10^{-1}$	143	0.00	0.52	0.56

Table B.1: Double differential reduced cross section $\tilde{\sigma}^{e-p}(x, Q^2)$ for the process $e^-p \to e^-X$. The following quantities are given for each bin: the lower and upper limits of x and Q^2, the bin centre, (x_c, Q_c^2) , the measured cross section, $\tilde{\sigma}_{e-p}$, at the Born level, the SM prediction calculated using CTEQ5D PDFs, the number of events observed, $N_{\rm obs}$, the estimated number of background events, $N_{\rm bg}$, the purity, $\mathcal{P}$, and the efficiency, $\mathcal{E}$, estimated for the bin. The first error of the measured cross section corresponds to the statistical uncertainty and the second one to the systematic uncertainty.

Q^2 range (GeV^2)	x range	Q_c^2 (GeV^2)	x_c	$\tilde{\sigma}(e^-p)$		$N_{\rm obs}$	$N_{\rm bg}$	$\mathcal{P}$	$\mathcal{E}$
				measured	SM				
900 – 1300	$(0.10-0.17)\cdot10^{-1}$	1200	$0.14\cdot10^{-1}$	$(8.78\pm0.55^{+0.37}_{-0.35})\cdot10^{-1}$	$8.61\cdot10^{-1}$	291	3.32	0.68	0.40
	$(0.17-0.25)\cdot10^{-1}$		$0.21\cdot10^{-1}$	$(6.85\pm0.46^{+0.19}_{-0.17})\cdot10^{-1}$	$7.63\cdot10^{-1}$	248	1.04	0.66	0.52
	$(0.25-0.37)\cdot10^{-1}$		$0.32\cdot10^{-1}$	$(5.98\pm0.40^{+0.04}_{-0.07})\cdot10^{-1}$	$6.53\cdot10^{-1}$	252	0.20	0.64	0.55
	$(0.37-0.60)\cdot10^{-1}$		$0.50\cdot10^{-1}$	$(4.95\pm0.30^{+0.15}_{-0.04})\cdot10^{-1}$	$5.51\cdot10^{-1}$	302	0.20	0.66	0.62
	$(0.60-1.00)\cdot10^{-1}$		$0.80\cdot10^{-1}$	$(4.33\pm0.27^{+0.13}_{-0.08})\cdot10^{-1}$	$4.59\cdot10^{-1}$	297	0.05	0.64	0.66
	0.10 – 0.17		0.13	$(3.63\pm0.24^{+0.08}_{-0.05})\cdot10^{-1}$	$3.74\cdot10^{-1}$	266	0.08	0.62	0.70
	0.17 – 0.30		0.25	$(2.61\pm0.19^{+0.07}_{-0.05})\cdot10^{-1}$	$2.50\cdot10^{-1}$	221	0.00	0.57	0.74
	0.30 – 0.53		0.40	$(1.06\pm0.14^{+0.04}_{-0.06})\cdot10^{-1}$	$1.31\cdot10^{-1}$	66	0.00	0.46	0.57
1300 – 1800	$(0.17-0.25)\cdot10^{-1}$	1500	$0.21\cdot10^{-1}$	$(7.95\pm0.67^{+0.23}_{-0.34})\cdot10^{-1}$	$7.71\cdot10^{-1}$	156	2.19	0.63	0.44
	$(0.25-0.37)\cdot10^{-1}$		$0.32\cdot10^{-1}$	$(7.57\pm0.59^{+0.30}_{-0.19})\cdot10^{-1}$	$6.65\cdot10^{-1}$	182	0.31	0.66	0.53
	$(0.37-0.60)\cdot10^{-1}$		$0.50\cdot10^{-1}$	$(5.98\pm0.43^{+0.12}_{-0.11})\cdot10^{-1}$	$5.60\cdot10^{-1}$	208	0.13	0.68	0.62
	$(0.60-1.00)\cdot10^{-1}$		$0.80\cdot10^{-1}$	$(4.47\pm0.35^{+0.06}_{-0.05})\cdot10^{-1}$	$4.64\cdot10^{-1}$	178	0.02	0.69	0.66
	0.10 – 0.15		0.13	$(3.93\pm0.35^{+0.05}_{-0.08})\cdot10^{-1}$	$3.77\cdot10^{-1}$	134	0.06	0.62	0.68
	0.15 – 0.23		0.18	$(3.68\pm0.34^{+0.05}_{-0.14})\cdot10^{-1}$	$3.18\cdot10^{-1}$	127	0.00	0.57	0.68
	0.23 – 0.35		0.25	$(2.70\pm0.30^{+0.12}_{-0.03})\cdot10^{-1}$	$2.50\cdot10^{-1}$	85	0.02	0.53	0.71
	0.35 – 0.53		0.40	$(1.32\pm0.23^{+0.36}_{-0.08})\cdot10^{-1}$	$1.29\cdot10^{-1}$	37	0.00	0.43	0.68
1800 – 2500	$(0.23-0.37)\cdot10^{-1}$	2000	$0.32\cdot10^{-1}$	$(6.87\pm0.63^{+0.52}_{-0.43})\cdot10^{-1}$	$6.83\cdot10^{-1}$	129	2.52	0.69	0.49
	$(0.37-0.60)\cdot10^{-1}$		$0.50\cdot10^{-1}$	$(5.60\pm0.51^{+0.10}_{-0.15})\cdot10^{-1}$	$5.76\cdot10^{-1}$	129	0.25	0.70	0.60
	$(0.60-1.00)\cdot10^{-1}$		$0.80\cdot10^{-1}$	$(4.20\pm0.40^{+0.23}_{-0.07})\cdot10^{-1}$	$4.75\cdot10^{-1}$	117	0.05	0.68	0.66
	0.10 – 0.15		0.13	$(3.90\pm0.42^{+0.07}_{-0.05})\cdot10^{-1}$	$3.82\cdot10^{-1}$	92	0.00	0.63	0.65
	0.15 – 0.23		0.18	$(2.76\pm0.33^{+0.08}_{-0.04})\cdot10^{-1}$	$3.20\cdot10^{-1}$	75	0.00	0.60	0.72
	0.23 – 0.35		0.25	$(2.83\pm0.38^{+0.17}_{-0.04})\cdot10^{-1}$	$2.50\cdot10^{-1}$	61	0.00	0.57	0.72
	0.35 – 0.53		0.40	$(1.16\pm0.23^{+0.09}_{-0.07})\cdot10^{-1}$	$1.28\cdot10^{-1}$	27	0.00	0.49	0.75
2500 – 3500	$(0.37-0.60)\cdot10^{-1}$	3000	$0.50\cdot10^{-1}$	$(6.95\pm0.72^{+0.20}_{-0.37})\cdot10^{-1}$	$6.11\cdot10^{-1}$	98	0.87	0.70	0.55
	$(0.60-1.00)\cdot10^{-1}$		$0.80\cdot10^{-1}$	$(4.86\pm0.52^{+0.19}_{-0.23})\cdot10^{-1}$	$5.01\cdot10^{-1}$	91	0.06	0.71	0.66
	0.10 – 0.15		0.13	$(3.74\pm0.49^{+0.07}_{-0.12})\cdot10^{-1}$	$3.97\cdot10^{-1}$	61	0.03	0.67	0.67
	0.15 – 0.23		0.18	$(3.13\pm0.42^{+0.04}_{-0.12})\cdot10^{-1}$	$3.29\cdot10^{-1}$	57	0.00	0.65	0.69
	0.23 – 0.35		0.25	$(2.78\pm0.42^{+0.04}_{-0.12})\cdot10^{-1}$	$2.54\cdot10^{-1}$	45	0.00	0.57	0.73
	0.35 – 0.53		0.40	$(1.46\pm0.31^{+0.07}_{-0.07})\cdot10^{-1}$	$1.28\cdot10^{-1}$	23	0.00	0.55	0.78
	0.53 – 1.00		0.65	$(1.59^{+0.76}_{-0.63}\ ^{+0.13}_{-0.08})\cdot10^{-2}$	$1.94\cdot10^{-2}$	7	0.00	0.37	0.73
3500 – 5600	$(0.40-1.00)\cdot10^{-1}$	5000	$0.80\cdot10^{-1}$	$(5.90\pm0.54^{+0.18}_{-0.14})\cdot10^{-1}$	$5.62\cdot10^{-1}$	128	1.17	0.81	0.61
	0.10 – 0.15		0.13	$(4.52\pm0.58^{+0.36}_{-0.06})\cdot10^{-1}$	$4.37\cdot10^{-1}$	63	0.00	0.71	0.66
	0.15 – 0.23		0.18	$(4.00\pm0.51^{+0.14}_{-0.17})\cdot10^{-1}$	$3.55\cdot10^{-1}$	63	0.00	0.69	0.72
	0.23 – 0.35		0.25	$(2.77\pm0.44^{+0.06}_{-0.07})\cdot10^{-1}$	$2.68\cdot10^{-1}$	41	0.02	0.63	0.79
	0.35 – 0.53		0.40	$(1.47\pm0.32^{+0.02}_{-0.17})\cdot10^{-1}$	$1.31\cdot10^{-1}$	22	0.00	0.61	0.80
5600 – 9000	$(0.70-1.50)\cdot10^{-1}$	8000	$1.30\cdot10^{-1}$	$(5.45\pm0.71^{+0.28}_{-0.27})\cdot10^{-1}$	$5.07\cdot10^{-1}$	61	0.36	0.82	0.60
	0.15 – 0.23		0.18	$(4.38\pm0.73^{+0.46}_{-0.18})\cdot10^{-1}$	$4.04\cdot10^{-1}$	37	0.02	0.73	0.72
	0.23 – 0.35		0.25	$(3.01\pm0.61^{+0.16}_{-0.03})\cdot10^{-1}$	$2.97\cdot10^{-1}$	25	0.02	0.69	0.77
	0.35 – 0.53		0.40	$(1.38^{+0.50}_{-0.37}\ ^{+0.23}_{-0.14})\cdot10^{-1}$	$1.39\cdot10^{-1}$	12	0.00	0.60	0.81
	0.53 – 1.00		0.65	$(1.8^{+1.2}_{-0.7}\ ^{+0.1}_{-0.1})\cdot10^{-2}$	$1.9\cdot10^{-2}$	4	0.00	0.52	0.81
9000 – 15000	0.11 – 0.23	12000	0.18	$(4.7\ \pm0.9\ ^{+0.2}_{-0.1})\cdot10^{-1}$	$4.7\cdot10^{-1}$	30	0.23	0.83	0.62
	0.23 – 0.35		0.25	$(3.2\ \pm0.8\ ^{+0.0}_{-0.3})\cdot10^{-1}$	$3.4\cdot10^{-1}$	14	0.00	0.72	0.72
	0.35 – 0.53		0.40	$(1.4^{+0.7}_{-0.5}\ ^{+0.2}_{-0.0})\cdot10^{-1}$	$1.5\cdot10^{-1}$	7	0.00	0.64	0.82
15000 – 25000	0.18 – 0.35	20000	0.25	$(5.7\ \pm1.4\ ^{+0.3}_{-0.2})\cdot10^{-1}$	$4.2\cdot10^{-1}$	17	0.20	0.81	0.67
	0.35 – 1.00		0.40	$(1.7^{+1.0}_{-0.8}\ ^{+0.4}_{-0.0})\cdot10^{-1}$	$1.8\cdot10^{-1}$	5	0.02	0.73	0.78
25000 – 50000	0.30 – 1.00	30000	0.40	$(2.3^{+1.6}_{-1.0}\ ^{+0.1}_{-0.2})\cdot10^{-1}$	$2.1\cdot10^{-1}$	4	0.00	0.81	0.74

Table B.2: Double differential reduced cross section $\tilde{\sigma}^{e^-p}(x,Q^2)$ for the process $e^-p \to e^-X$. The following quantities are given for each bin: the lower and upper limits of x and Q^2, the bin centre, (x_c, Q_c^2) , the measured cross section, $\tilde{\sigma}_{e-p}$, at the Born level, the SM prediction calculated using CTEQ5D PDFs, the number of events observed, $N_{\rm obs}$, the estimated number of background events, $N_{\rm bg}$, the purity, $\mathcal{P}$, and the efficiency, $\mathcal{E}$, estimated for the bin. The first error of the measured cross section corresponds to the statistical uncertainty and the second one to the systematic uncertainty.

Q^2 range (GeV^2)	x range	Q_c^2 (GeV^2)	x_c	$\tilde{\sigma}(e^+p)$		$N_{\rm obs}$	$N_{\rm bg}$	$\mathcal{P}$	$\mathcal{E}$
				measured	SM				
185 – 240	$(0.37 - 0.60)\cdot 10^{-2}$	200	$0.50\cdot 10^{-2}$	$1.124 \pm 0.017\,^{+0.018}_{-0.017}$	1.104	6986	9.31	0.72	0.40
	$(0.60 - 1.00)\cdot 10^{-2}$		$0.80\cdot 10^{-2}$	$(9.40 \pm 0.14\,^{+0.13}_{-0.13})\cdot 10^{-1}$	$9.42\cdot 10^{-1}$	7075	2.33	0.72	0.45
	$(0.10 - 0.17)\cdot 10^{-1}$		$0.13\cdot 10^{-1}$	$(8.11 \pm 0.12\,^{+0.08}_{-0.08})\cdot 10^{-1}$	$7.98\cdot 10^{-1}$	7146	0.95	0.71	0.48
	$(0.17 - 0.25)\cdot 10^{-1}$		$0.21\cdot 10^{-1}$	$(6.87 \pm 0.13\,^{+0.14}_{-0.09})\cdot 10^{-1}$	$6.77\cdot 10^{-1}$	4859	0.37	0.59	0.44
	$(0.25 - 0.37)\cdot 10^{-1}$		$0.32\cdot 10^{-1}$	$(5.95 \pm 0.11\,^{+0.11}_{-0.18})\cdot 10^{-1}$	$5.89\cdot 10^{-1}$	4611	0.15	0.56	0.45
	$(0.37 - 0.60)\cdot 10^{-1}$		$0.50\cdot 10^{-1}$	$(5.26 \pm 0.09\,^{+0.10}_{-0.10})\cdot 10^{-1}$	$5.12\cdot 10^{-1}$	5117	0.08	0.57	0.50
	$(0.60 - 1.20)\cdot 10^{-1}$		$0.80\cdot 10^{-1}$	$(4.32 \pm 0.07\,^{+0.09}_{-0.05})\cdot 10^{-1}$	$4.42\cdot 10^{-1}$	6558	0.11	0.59	0.61
	0.12 – 0.25		0.18	$(3.33 \pm 0.07\,^{+0.08}_{-0.11})\cdot 10^{-1}$	$3.28\cdot 10^{-1}$	3878	0.00	0.54	0.47
240 – 310	$(0.37 - 0.60)\cdot 10^{-2}$	250	$0.50\cdot 10^{-2}$	$1.17 \pm 0.02\,^{+0.02}_{-0.03}$	1.12	3588	6.58	0.57	0.30
	$(0.60 - 1.00)\cdot 10^{-2}$		$0.80\cdot 10^{-2}$	$(9.73 \pm 0.17\,^{+0.15}_{-0.13})\cdot 10^{-1}$	$9.62\cdot 10^{-1}$	5265	3.76	0.62	0.47
	$(0.10 - 0.17)\cdot 10^{-1}$		$0.13\cdot 10^{-1}$	$(8.36 \pm 0.14\,^{+0.10}_{-0.13})\cdot 10^{-1}$	$8.14\cdot 10^{-1}$	5220	0.61	0.62	0.49
	$(0.17 - 0.25)\cdot 10^{-1}$		$0.21\cdot 10^{-1}$	$(6.98 \pm 0.15\,^{+0.18}_{-0.06})\cdot 10^{-1}$	$6.88\cdot 10^{-1}$	3523	0.33	0.54	0.46
	$(0.25 - 0.37)\cdot 10^{-1}$		$0.32\cdot 10^{-1}$	$(5.92 \pm 0.13\,^{+0.06}_{-0.07})\cdot 10^{-1}$	$5.96\cdot 10^{-1}$	3352	0.16	0.51	0.48
	$(0.37 - 0.60)\cdot 10^{-1}$		$0.50\cdot 10^{-1}$	$(5.29 \pm 0.11\,^{+0.08}_{-0.07})\cdot 10^{-1}$	$5.16\cdot 10^{-1}$	3774	0.11	0.53	0.52
	$(0.60 - 1.20)\cdot 10^{-1}$		$0.80\cdot 10^{-1}$	$(4.23 \pm 0.08\,^{+0.10}_{-0.05})\cdot 10^{-1}$	$4.43\cdot 10^{-1}$	4687	0.00	0.53	0.63
	0.12 – 0.25		0.18	$(3.23 \pm 0.07\,^{+0.11}_{-0.10})\cdot 10^{-1}$	$3.26\cdot 10^{-1}$	3504	0.00	0.50	0.61
310 – 410	$(0.60 - 1.00)\cdot 10^{-2}$	350	$0.80\cdot 10^{-2}$	$1.02 \pm 0.02\,^{+0.01}_{-0.02}$	0.99	3273	6.20	0.57	0.34
	$(0.10 - 0.17)\cdot 10^{-1}$		$0.13\cdot 10^{-1}$	$(8.31 \pm 0.16\,^{+0.08}_{-0.12})\cdot 10^{-1}$	$8.35\cdot 10^{-1}$	3854	0.47	0.65	0.48
	$(0.17 - 0.25)\cdot 10^{-1}$		$0.21\cdot 10^{-1}$	$(6.99 \pm 0.16\,^{+0.06}_{-0.10})\cdot 10^{-1}$	$7.03\cdot 10^{-1}$	2859	0.31	0.56	0.47
	$(0.25 - 0.37)\cdot 10^{-1}$		$0.32\cdot 10^{-1}$	$(6.21 \pm 0.14\,^{+0.06}_{-0.09})\cdot 10^{-1}$	$6.06\cdot 10^{-1}$	2758	0.43	0.54	0.49
	$(0.37 - 0.60)\cdot 10^{-1}$		$0.50\cdot 10^{-1}$	$(5.04 \pm 0.11\,^{+0.06}_{-0.05})\cdot 10^{-1}$	$5.21\cdot 10^{-1}$	2923	0.07	0.54	0.53
	$(0.60 - 1.20)\cdot 10^{-1}$		$0.80\cdot 10^{-1}$	$(4.31 \pm 0.08\,^{+0.04}_{-0.05})\cdot 10^{-1}$	$4.45\cdot 10^{-1}$	3789	0.08	0.57	0.65
	0.12 – 0.25		0.18	$(3.13 \pm 0.07\,^{+0.08}_{-0.06})\cdot 10^{-1}$	$3.23\cdot 10^{-1}$	3159	0.00	0.51	0.70
410 – 530	$(0.60 - 1.00)\cdot 10^{-2}$	450	$0.80\cdot 10^{-2}$	$1.04 \pm 0.02\,^{+0.02}_{-0.01}$	1.00	2556	6.32	0.61	0.40
	$(0.10 - 0.17)\cdot 10^{-1}$		$0.13\cdot 10^{-1}$	$(8.49 \pm 0.22\,^{+0.09}_{-0.17})\cdot 10^{-1}$	$8.49\cdot 10^{-1}$	1947	2.44	0.61	0.34
	$(0.17 - 0.25)\cdot 10^{-1}$		$0.21\cdot 10^{-1}$	$(6.92 \pm 0.21\,^{+0.07}_{-0.13})\cdot 10^{-1}$	$7.14\cdot 10^{-1}$	1451	0.18	0.58	0.37
	$(0.25 - 0.37)\cdot 10^{-1}$		$0.32\cdot 10^{-1}$	$(6.39 \pm 0.18\,^{+0.10}_{-0.12})\cdot 10^{-1}$	$6.13\cdot 10^{-1}$	1646	0.00	0.57	0.44
	$(0.37 - 0.60)\cdot 10^{-1}$		$0.50\cdot 10^{-1}$	$(5.13 \pm 0.13\,^{+0.09}_{-0.05})\cdot 10^{-1}$	$5.25\cdot 10^{-1}$	1912	0.10	0.58	0.53
	$(0.60 - 1.00)\cdot 10^{-1}$		$0.80\cdot 10^{-1}$	$(4.52 \pm 0.12\,^{+0.03}_{-0.04})\cdot 10^{-1}$	$4.46\cdot 10^{-1}$	1979	0.00	0.56	0.59
	0.10 – 0.17		0.13	$(3.74 \pm 0.10\,^{+0.07}_{-0.08})\cdot 10^{-1}$	$3.72\cdot 10^{-1}$	1701	0.00	0.52	0.61
	0.17 – 0.30		0.25	$(2.64 \pm 0.08\,^{+0.08}_{-0.11})\cdot 10^{-1}$	$2.58\cdot 10^{-1}$	1390	0.00	0.48	0.64
530 – 710	$(0.60 - 1.00)\cdot 10^{-2}$	650	$0.80\cdot 10^{-2}$	$(9.89 \pm 0.31\,^{+0.18}_{-0.24})\cdot 10^{-1}$	$9.85\cdot 10^{-1}$	1294	3.76	0.61	0.36
	$(0.10 - 0.17)\cdot 10^{-1}$		$0.13\cdot 10^{-1}$	$(8.71 \pm 0.21\,^{+0.12}_{-0.13})\cdot 10^{-1}$	$8.62\cdot 10^{-1}$	2291	5.92	0.65	0.49
	$(0.17 - 0.25)\cdot 10^{-1}$		$0.21\cdot 10^{-1}$	$(7.28 \pm 0.22\,^{+0.19}_{-0.06})\cdot 10^{-1}$	$7.27\cdot 10^{-1}$	1370	1.62	0.62	0.43
	$(0.25 - 0.37)\cdot 10^{-1}$		$0.32\cdot 10^{-1}$	$(6.22 \pm 0.21\,^{+0.05}_{-0.09})\cdot 10^{-1}$	$6.22\cdot 10^{-1}$	1112	0.30	0.59	0.37
	$(0.37 - 0.60)\cdot 10^{-1}$		$0.50\cdot 10^{-1}$	$(5.33 \pm 0.18\,^{+0.09}_{-0.09})\cdot 10^{-1}$	$5.30\cdot 10^{-1}$	1144	0.10	0.60	0.36
	$(0.60 - 1.00)\cdot 10^{-1}$		$0.80\cdot 10^{-1}$	$(4.49 \pm 0.16\,^{+0.10}_{-0.05})\cdot 10^{-1}$	$4.47\cdot 10^{-1}$	1080	0.00	0.57	0.39
	0.10 – 0.17		0.13	$(3.58 \pm 0.13\,^{+0.08}_{-0.03})\cdot 10^{-1}$	$3.70\cdot 10^{-1}$	952	0.00	0.55	0.43
	0.17 – 0.30		0.25	$(2.53 \pm 0.10\,^{+0.08}_{-0.05})\cdot 10^{-1}$	$2.54\cdot 10^{-1}$	857	0.00	0.51	0.49
710 – 900	$(0.90 - 1.70)\cdot 10^{-2}$	800	$1.30\cdot 10^{-2}$	$(8.62 \pm 0.25\,^{+0.14}_{-0.15})\cdot 10^{-1}$	$8.63\cdot 10^{-1}$	1517	5.67	0.62	0.42
	$(0.17 - 0.25)\cdot 10^{-1}$		$0.21\cdot 10^{-1}$	$(7.44 \pm 0.27\,^{+0.18}_{-0.18})\cdot 10^{-1}$	$7.32\cdot 10^{-1}$	980	1.37	0.62	0.50
	$(0.25 - 0.37)\cdot 10^{-1}$		$0.32\cdot 10^{-1}$	$(6.58 \pm 0.24\,^{+0.08}_{-0.08})\cdot 10^{-1}$	$6.27\cdot 10^{-1}$	1001	0.24	0.62	0.53
	$(0.37 - 0.60)\cdot 10^{-1}$		$0.50\cdot 10^{-1}$	$(5.19 \pm 0.18\,^{+0.12}_{-0.06})\cdot 10^{-1}$	$5.33\cdot 10^{-1}$	1026	0.09	0.62	0.55
	$(0.60 - 1.00)\cdot 10^{-1}$		$0.80\cdot 10^{-1}$	$(4.43 \pm 0.16\,^{+0.09}_{-0.05})\cdot 10^{-1}$	$4.48\cdot 10^{-1}$	935	0.35	0.60	0.55
	0.10 – 0.17		0.13	$(3.62 \pm 0.15\,^{+0.03}_{-0.10})\cdot 10^{-1}$	$3.69\cdot 10^{-1}$	788	0.00	0.56	0.58
	0.17 – 0.30		0.25	$(2.64 \pm 0.12\,^{+0.10}_{-0.07})\cdot 10^{-1}$	$2.52\cdot 10^{-1}$	640	0.00	0.53	0.59

Table B.3: Double differential reduced cross section $\tilde{\sigma}^{e+p}(x, Q^2)$ for the process $e^+p \to e^+X$. The following quantities are given for each bin: the lower and upper limits of x and Q^2, the bin centre, (x_c, Q_c^2), the measured cross section, $\tilde{\sigma}_{e+p}$, at the Born level, the SM prediction calculated using CTEQ5D PDFs, the number of events observed, $N_{\rm obs}$, the estimated number of background events, $N_{\rm bg}$, the purity, $\mathcal{P}$, and the efficiency, $\mathcal{E}$, estimated for the bin. The first error of the measured cross section corresponds to the statistical uncertainty and the second one to the systematic uncertainty.

Q^2 range (GeV^2)	x range	Q_c^2 (GeV^2)	x_c	$\tilde{\sigma}(e^+p)$		$N_{\rm obs}$	$N_{\rm bg}$	$\mathcal{P}$	$\mathcal{E}$
				measured	SM				
900 – 1300	$(0.10 - 0.17)\cdot 10^{-1}$	1200	$0.14\cdot 10^{-1}$	$(8.74 \pm 0.30^{+0.28}_{-0.43})\cdot 10^{-1}$	$8.26\cdot 10^{-1}$	1057	12.7	0.66	0.39
	$(0.17 - 0.25)\cdot 10^{-1}$		$0.21\cdot 10^{-1}$	$(7.95 \pm 0.28^{+0.16}_{-0.21})\cdot 10^{-1}$	$7.34\cdot 10^{-1}$	1056	3.97	0.64	0.50
	$(0.25 - 0.37)\cdot 10^{-1}$		$0.32\cdot 10^{-1}$	$(6.47 \pm 0.23^{+0.13}_{-0.16})\cdot 10^{-1}$	$6.32\cdot 10^{-1}$	1040	0.72	0.65	0.56
	$(0.37 - 0.60)\cdot 10^{-1}$		$0.50\cdot 10^{-1}$	$(5.46 \pm 0.18^{+0.06}_{-0.16})\cdot 10^{-1}$	$5.36\cdot 10^{-1}$	1206	0.67	0.66	0.59
	$(0.60 - 1.00)\cdot 10^{-1}$		$0.80\cdot 10^{-1}$	$(4.65 \pm 0.15^{+0.05}_{-0.05})\cdot 10^{-1}$	$4.48\cdot 10^{-1}$	1219	0.17	0.64	0.63
	0.10 – 0.17		0.13	$(3.58 \pm 0.12^{+0.09}_{-0.04})\cdot 10^{-1}$	$3.67\cdot 10^{-1}$	1045	0.35	0.60	0.68
	0.17 – 0.30		0.25	$(2.45 \pm 0.10^{+0.07}_{-0.05})\cdot 10^{-1}$	$2.47\cdot 10^{-1}$	819	0.00	0.59	0.76
	0.30 – 0.53		0.40	$(1.40 \pm 0.09^{+0.06}_{-0.08})\cdot 10^{-1}$	$1.30\cdot 10^{-1}$	311	0.00	0.50	0.57
1300 – 1800	$(0.17 - 0.25)\cdot 10^{-1}$	1500	$0.21\cdot 10^{-1}$	$(7.24 \pm 0.34^{+0.33}_{-0.36})\cdot 10^{-1}$	$7.27\cdot 10^{-1}$	527	8.89	0.61	0.41
	$(0.25 - 0.37)\cdot 10^{-1}$		$0.32\cdot 10^{-1}$	$(6.39 \pm 0.28^{+0.13}_{-0.09})\cdot 10^{-1}$	$6.31\cdot 10^{-1}$	582	1.23	0.66	0.53
	$(0.37 - 0.60)\cdot 10^{-1}$		$0.50\cdot 10^{-1}$	$(5.95 \pm 0.23^{+0.06}_{-0.10})\cdot 10^{-1}$	$5.36\cdot 10^{-1}$	756	0.46	0.69	0.60
	$(0.60 - 1.00)\cdot 10^{-1}$		$0.80\cdot 10^{-1}$	$(4.46 \pm 0.18^{+0.04}_{-0.06})\cdot 10^{-1}$	$4.48\cdot 10^{-1}$	686	0.08	0.68	0.65
	0.10 – 0.15		0.13	$(3.54 \pm 0.17^{+0.05}_{-0.09})\cdot 10^{-1}$	$3.66\cdot 10^{-1}$	480	0.23	0.62	0.65
	0.15 – 0.23		0.18	$(3.19 \pm 0.16^{+0.05}_{-0.06})\cdot 10^{-1}$	$3.10\cdot 10^{-1}$	450	0.00	0.57	0.69
	0.23 – 0.35		0.25	$(2.67 \pm 0.16^{+0.06}_{-0.04})\cdot 10^{-1}$	$2.45\cdot 10^{-1}$	325	0.08	0.55	0.72
	0.35 – 0.53		0.40	$(1.35 \pm 0.12^{+0.02}_{-0.11})\cdot 10^{-1}$	$1.28\cdot 10^{-1}$	145	0.00	0.47	0.66
1800 – 2500	$(0.23 - 0.37)\cdot 10^{-1}$	2000	$0.32\cdot 10^{-1}$	$(6.66 \pm 0.34^{+0.37}_{-0.25})\cdot 10^{-1}$	$6.24\cdot 10^{-1}$	441	9.73	0.67	0.47
	$(0.37 - 0.60)\cdot 10^{-1}$		$0.50\cdot 10^{-1}$	$(5.18 \pm 0.26^{+0.17}_{-0.08})\cdot 10^{-1}$	$5.34\cdot 10^{-1}$	445	0.97	0.70	0.59
	$(0.60 - 1.00)\cdot 10^{-1}$		$0.80\cdot 10^{-1}$	$(4.54 \pm 0.22^{+0.09}_{-0.05})\cdot 10^{-1}$	$4.46\cdot 10^{-1}$	485	0.16	0.70	0.66
	0.10 – 0.15		0.13	$(4.04 \pm 0.22^{+0.10}_{-0.07})\cdot 10^{-1}$	$3.64\cdot 10^{-1}$	370	0.00	0.64	0.64
	0.15 – 0.23		0.18	$(3.25 \pm 0.19^{+0.04}_{-0.04})\cdot 10^{-1}$	$3.08\cdot 10^{-1}$	325	0.00	0.62	0.70
	0.23 – 0.35		0.25	$(2.44 \pm 0.18^{+0.03}_{-0.10})\cdot 10^{-1}$	$2.42\cdot 10^{-1}$	219	0.00	0.55	0.74
	0.35 – 0.53		0.40	$(1.27 \pm 0.13^{+0.12}_{-0.03})\cdot 10^{-1}$	$1.25\cdot 10^{-1}$	104	0.00	0.48	0.72
2500 – 3500	$(0.37 - 0.60)\cdot 10^{-1}$	3000	$0.50\cdot 10^{-1}$	$(5.54 \pm 0.33^{+0.15}_{-0.12})\cdot 10^{-1}$	$5.20\cdot 10^{-1}$	301	3.26	0.70	0.55
	$(0.60 - 1.00)\cdot 10^{-1}$		$0.80\cdot 10^{-1}$	$(4.18 \pm 0.25^{+0.03}_{-0.07})\cdot 10^{-1}$	$4.38\cdot 10^{-1}$	297	0.24	0.71	0.64
	0.10 – 0.15		0.13	$(3.48 \pm 0.24^{+0.19}_{-0.03})\cdot 10^{-1}$	$3.59\cdot 10^{-1}$	224	0.11	0.65	0.64
	0.15 – 0.23		0.18	$(3.24 \pm 0.22^{+0.06}_{-0.04})\cdot 10^{-1}$	$3.03\cdot 10^{-1}$	231	0.00	0.64	0.70
	0.23 – 0.35		0.25	$(2.29 \pm 0.20^{+0.02}_{-0.03})\cdot 10^{-1}$	$2.38\cdot 10^{-1}$	148	0.00	0.61	0.75
	0.35 – 0.53		0.40	$(1.42 \pm 0.16^{+0.06}_{-0.06})\cdot 10^{-1}$	$1.22\cdot 10^{-1}$	88	0.00	0.54	0.78
	0.53 – 1.00		0.65	$(2.13 \pm 0.37^{+0.20}_{-0.27})\cdot 10^{-2}$	$1.89\cdot 10^{-2}$	35	0.00	0.44	0.74
3500 – 5600	$(0.40 - 1.00)\cdot 10^{-1}$	5000	$0.80\cdot 10^{-1}$	$(4.20 \pm 0.23^{+0.13}_{-0.09})\cdot 10^{-1}$	$4.09\cdot 10^{-1}$	363	4.45	0.80	0.59
	0.10 – 0.15		0.13	$(3.62 \pm 0.26^{+0.03}_{-0.08})\cdot 10^{-1}$	$3.42\cdot 10^{-1}$	199	0.00	0.71	0.66
	0.15 – 0.23		0.18	$(2.75 \pm 0.22^{+0.06}_{-0.08})\cdot 10^{-1}$	$2.92\cdot 10^{-1}$	172	0.00	0.69	0.71
	0.23 – 0.35		0.25	$(2.16 \pm 0.19^{+0.06}_{-0.04})\cdot 10^{-1}$	$2.30\cdot 10^{-1}$	132	0.08	0.63	0.76
	0.35 – 0.53		0.40	$(1.23 \pm 0.15^{+0.02}_{-0.02})\cdot 10^{-1}$	$1.18\cdot 10^{-1}$	70	0.00	0.57	0.77
5600 – 9000	$(0.70 - 1.50)\cdot 10^{-1}$	8000	$1.30\cdot 10^{-1}$	$(3.18 \pm 0.26^{+0.06}_{-0.07})\cdot 10^{-1}$	$3.02\cdot 10^{-1}$	151	1.43	0.80	0.58
	0.15 – 0.23		0.18	$(2.63 \pm 0.28^{+0.06}_{-0.05})\cdot 10^{-1}$	$2.65\cdot 10^{-1}$	92	0.10	0.71	0.70
	0.23 – 0.35		0.25	$(2.08 \pm 0.25^{+0.07}_{-0.02})\cdot 10^{-1}$	$2.15\cdot 10^{-1}$	73	0.10	0.69	0.75
	0.35 – 0.53		0.40	$(1.01 \pm 0.17^{+0.08}_{-0.03})\cdot 10^{-1}$	$1.13\cdot 10^{-1}$	35	0.00	0.61	0.81
	0.53 – 1.00		0.65	$(1.15^{+0.44}_{-0.37} \pm 0.17^{+0.17}_{-0.02})\cdot 10^{-2}$	$1.71\cdot 10^{-2}$	10	0.00	0.53	0.86
9000 – 15000	0.11 – 0.23	12000	0.18	$(2.77 \pm 0.33^{+0.08}_{-0.20})\cdot 10^{-1}$	$2.20\cdot 10^{-1}$	72	1.15	0.81	0.60
	0.23 – 0.35		0.25	$(1.85 \pm 0.32^{+0.08}_{-0.04})\cdot 10^{-1}$	$1.87\cdot 10^{-1}$	35	0.00	0.72	0.72
	0.35 – 0.53		0.40	$(1.20 \pm 0.24^{+0.05}_{-0.12})\cdot 10^{-1}$	$1.04\cdot 10^{-1}$	25	0.00	0.67	0.81
15000 – 25000	0.18 – 0.35	20000	0.25	$(8.6^{+3.3}_{-2.5}{}^{+2.5}_{-0.9})\cdot 10^{-2}$	$12.8\cdot 10^{-2}$	11	0.87	0.80	0.64
	0.35 – 1.00		0.40	$(4.7^{+2.7}_{-1.8}{}^{+0.8}_{-0.2})\cdot 10^{-2}$	$8.2\cdot 10^{-2}$	6	0.12	0.73	0.77
25000 – 50000	0.30 – 1.00	30000	0.40	$(2.6^{+2.9}_{-1.6}{}^{+2.0}_{-0.1})\cdot 10^{-2}$	$5.4\cdot 10^{-2}$	2	0.00	0.78	0.73

Table B.4: Double differential reduced cross section $\tilde{\sigma}^{e+p}(x, Q^2)$ for the process $e^+p \to e^+X$. The following quantities are given for each bin: the lower and upper limits of x and Q^2, the bin centre, (x_c, Q_c^2), the measured cross section, $\tilde{\sigma}_{e+p}$, at the Born level, the SM prediction calculated using CTEQ5D PDFs, the number of events observed, $N_{\rm obs}$, the estimated number of background events, $N_{\rm bg}$, the purity, $\mathcal{P}$, and the efficiency, $\mathcal{E}$, estimated for the bin. The first error of the measured cross section corresponds to the statistical uncertainty and the second one to the systematic uncertainty.

Appendix C

Tables of structure functions

Q^2 range (GeV^2)	x range	Q_c^2 (GeV^2)	x_c	F_2^{em} measured	F_2^{em} SM	$F_2^{e^-p}$	$F_2^{e^+p}$
185 – 240	$(0.37 - 0.60) \cdot 10^{-2}$	200	$0.50 \cdot 10^{-2}$	$1.157 \pm 0.015\ ^{+0.033}_{-0.031}$	1.122	1.140 ± 0.03	1.162 ± 0.02
	$(0.60 - 1.00) \cdot 10^{-2}$		$0.80 \cdot 10^{-2}$	$(9.64 \pm 0.12\ ^{+0.23}_{-0.22}) \cdot 10^{-1}$	$9.46 \cdot 10^{-1}$	0.960 ± 0.03	0.965 ± 0.01
	$(0.10 - 0.17) \cdot 10^{-1}$		$0.13 \cdot 10^{-1}$	$(8.25 \pm 0.11\ ^{+0.16}_{-0.15}) \cdot 10^{-1}$	$7.99 \cdot 10^{-1}$	0.811 ± 0.02	0.829 ± 0.01
	$(0.17 - 0.25) \cdot 10^{-1}$		$0.21 \cdot 10^{-1}$	$(6.96 \pm 0.11\ ^{+0.18}_{-0.14}) \cdot 10^{-1}$	$6.77 \cdot 10^{-1}$	0.686 ± 0.02	0.700 ± 0.01
	$(0.25 - 0.37) \cdot 10^{-1}$		$0.32 \cdot 10^{-1}$	$(6.06 \pm 0.10\ ^{+0.14}_{-0.20}) \cdot 10^{-1}$	$5.89 \cdot 10^{-1}$	0.605 ± 0.02	0.606 ± 0.01
	$(0.37 - 0.60) \cdot 10^{-1}$		$0.50 \cdot 10^{-1}$	$(5.24 \pm 0.08\ ^{+0.12}_{-0.12}) \cdot 10^{-1}$	$5.11 \cdot 10^{-1}$	0.491 ± 0.02	0.535 ± 0.01
	$(0.60 - 1.20) \cdot 10^{-1}$		$0.80 \cdot 10^{-1}$	$(4.33 \pm 0.06\ ^{+0.11}_{-0.07}) \cdot 10^{-1}$	$4.41 \cdot 10^{-1}$	0.415 ± 0.01	0.439 ± 0.01
	0.12 – 0.25		0.18	$(3.32 \pm 0.06\ ^{+0.09}_{-0.11}) \cdot 10^{-1}$	$3.28 \cdot 10^{-1}$	0.313 ± 0.01	0.338 ± 0.01
240 – 310	$(0.37 - 0.60) \cdot 10^{-2}$	250	$0.50 \cdot 10^{-2}$	$1.19 \pm 0.02\ ^{+0.04}_{-0.04}$	1.15	1.107 ± 0.04	1.221 ± 0.03
	$(0.60 - 1.00) \cdot 10^{-2}$		$0.80 \cdot 10^{-2}$	$(9.92 \pm 0.15\ ^{+0.26}_{-0.24}) \cdot 10^{-1}$	$9.69 \cdot 10^{-1}$	0.971 ± 0.03	0.999 ± 0.02
	$(0.10 - 0.17) \cdot 10^{-1}$		$0.13 \cdot 10^{-1}$	$(8.52 \pm 0.13\ ^{+0.18}_{-0.19}) \cdot 10^{-1}$	$8.15 \cdot 10^{-1}$	0.845 ± 0.03	0.854 ± 0.01
	$(0.17 - 0.25) \cdot 10^{-1}$		$0.21 \cdot 10^{-1}$	$(7.07 \pm 0.13\ ^{+0.21}_{-0.12}) \cdot 10^{-1}$	$6.88 \cdot 10^{-1}$	0.694 ± 0.03	0.712 ± 0.02
	$(0.25 - 0.37) \cdot 10^{-1}$		$0.32 \cdot 10^{-1}$	$(6.01 \pm 0.11\ ^{+0.11}_{-0.11}) \cdot 10^{-1}$	$5.95 \cdot 10^{-1}$	0.595 ± 0.02	0.603 ± 0.01
	$(0.37 - 0.60) \cdot 10^{-1}$		$0.50 \cdot 10^{-1}$	$(5.32 \pm 0.09\ ^{+0.11}_{-0.10}) \cdot 10^{-1}$	$5.15 \cdot 10^{-1}$	0.513 ± 0.02	0.538 ± 0.01
	$(0.60 - 1.20) \cdot 10^{-1}$		$0.80 \cdot 10^{-1}$	$(4.31 \pm 0.07\ ^{+0.12}_{-0.07}) \cdot 10^{-1}$	$4.43 \cdot 10^{-1}$	0.433 ± 0.01	0.430 ± 0.01
	0.12 – 0.25		0.18	$(3.22 \pm 0.06\ ^{+0.11}_{-0.10}) \cdot 10^{-1}$	$3.25 \cdot 10^{-1}$	0.306 ± 0.01	0.328 ± 0.01
310 – 410	$(0.60 - 1.00) \cdot 10^{-2}$	350	$0.80 \cdot 10^{-2}$	$1.04 \pm 0.02\ ^{+0.03}_{-0.03}$	1.00	1.000 ± 0.04	1.052 ± 0.02
	$(0.10 - 0.17) \cdot 10^{-1}$		$0.13 \cdot 10^{-1}$	$(8.51 \pm 0.15\ ^{+0.17}_{-0.20}) \cdot 10^{-1}$	$8.39 \cdot 10^{-1}$	0.852 ± 0.03	0.851 ± 0.02
	$(0.17 - 0.25) \cdot 10^{-1}$		$0.21 \cdot 10^{-1}$	$(7.14 \pm 0.14\ ^{+0.13}_{-0.15}) \cdot 10^{-1}$	$7.04 \cdot 10^{-1}$	0.716 ± 0.03	0.713 ± 0.02
	$(0.25 - 0.37) \cdot 10^{-1}$		$0.32 \cdot 10^{-1}$	$(6.22 \pm 0.13\ ^{+0.10}_{-0.13}) \cdot 10^{-1}$	$6.06 \cdot 10^{-1}$	0.591 ± 0.03	0.632 ± 0.01
	$(0.37 - 0.60) \cdot 10^{-1}$		$0.50 \cdot 10^{-1}$	$(5.14 \pm 0.10\ ^{+0.09}_{-0.08}) \cdot 10^{-1}$	$5.21 \cdot 10^{-1}$	0.518 ± 0.02	0.512 ± 0.01
	$(0.60 - 1.20) \cdot 10^{-1}$		$0.80 \cdot 10^{-1}$	$(4.39 \pm 0.08\ ^{+0.06}_{-0.07}) \cdot 10^{-1}$	$4.44 \cdot 10^{-1}$	0.446 ± 0.02	0.437 ± 0.01
	0.12 – 0.25		0.18	$(3.15 \pm 0.06\ ^{+0.08}_{-0.05}) \cdot 10^{-1}$	$3.22 \cdot 10^{-1}$	0.307 ± 0.01	0.317 ± 0.01
410 – 530	$(0.60 - 1.00) \cdot 10^{-2}$	450	$0.80 \cdot 10^{-2}$	$1.06 \pm 0.02\ ^{+0.03}_{-0.03}$	1.03	1.005 ± 0.04	1.084 ± 0.02
	$(0.10 - 0.17) \cdot 10^{-1}$		$0.13 \cdot 10^{-1}$	$(8.70 \pm 0.20\ ^{+0.20}_{-0.24}) \cdot 10^{-1}$	$8.56 \cdot 10^{-1}$	0.865 ± 0.04	0.871 ± 0.02
	$(0.17 - 0.25) \cdot 10^{-1}$		$0.21 \cdot 10^{-1}$	$(7.08 \pm 0.19\ ^{+0.13}_{-0.18}) \cdot 10^{-1}$	$7.15 \cdot 10^{-1}$	0.714 ± 0.04	0.706 ± 0.02
	$(0.25 - 0.37) \cdot 10^{-1}$		$0.32 \cdot 10^{-1}$	$(6.45 \pm 0.16\ ^{+0.13}_{-0.15}) \cdot 10^{-1}$	$6.13 \cdot 10^{-1}$	0.629 ± 0.03	0.650 ± 0.02
	$(0.37 - 0.60) \cdot 10^{-1}$		$0.50 \cdot 10^{-1}$	$(5.21 \pm 0.12\ ^{+0.11}_{-0.08}) \cdot 10^{-1}$	$5.24 \cdot 10^{-1}$	0.522 ± 0.02	0.521 ± 0.01
	$(0.60 - 1.00) \cdot 10^{-1}$		$0.80 \cdot 10^{-1}$	$(4.55 \pm 0.10\ ^{+0.06}_{-0.06}) \cdot 10^{-1}$	$4.45 \cdot 10^{-1}$	0.443 ± 0.02	0.459 ± 0.01
	0.10 – 0.17		0.13	$(3.83 \pm 0.09\ ^{+0.08}_{-0.09}) \cdot 10^{-1}$	$3.71 \cdot 10^{-1}$	0.397 ± 0.02	0.379 ± 0.01
	0.17 – 0.30		0.25	$(2.71 \pm 0.07\ ^{+0.08}_{-0.11}) \cdot 10^{-1}$	$2.57 \cdot 10^{-1}$	0.285 ± 0.02	0.267 ± 0.01
530 – 710	$(0.60 - 1.00) \cdot 10^{-2}$	650	$0.80 \cdot 10^{-2}$	$1.04 \pm 0.03\ ^{+0.03}_{-0.04}$	1.06	0.961 ± 0.06	1.067 ± 0.03
	$(0.10 - 0.17) \cdot 10^{-1}$		$0.13 \cdot 10^{-1}$	$(9.09 \pm 0.19\ ^{+0.23}_{-0.23}) \cdot 10^{-1}$	$8.80 \cdot 10^{-1}$	0.934 ± 0.04	0.901 ± 0.02
	$(0.17 - 0.25) \cdot 10^{-1}$		$0.21 \cdot 10^{-1}$	$(7.58 \pm 0.20\ ^{+0.23}_{-0.14}) \cdot 10^{-1}$	$7.30 \cdot 10^{-1}$	0.808 ± 0.04	0.744 ± 0.02
	$(0.25 - 0.37) \cdot 10^{-1}$		$0.32 \cdot 10^{-1}$	$(6.32 \pm 0.19\ ^{+0.10}_{-0.13}) \cdot 10^{-1}$	$6.23 \cdot 10^{-1}$	0.627 ± 0.04	0.633 ± 0.02
	$(0.37 - 0.60) \cdot 10^{-1}$		$0.50 \cdot 10^{-1}$	$(5.45 \pm 0.16\ ^{+0.11}_{-0.11}) \cdot 10^{-1}$	$5.29 \cdot 10^{-1}$	0.560 ± 0.04	0.541 ± 0.02
	$(0.60 - 1.00) \cdot 10^{-1}$		$0.80 \cdot 10^{-1}$	$(4.42 \pm 0.14\ ^{+0.12}_{-0.07}) \cdot 10^{-1}$	$4.46 \cdot 10^{-1}$	0.401 ± 0.03	0.455 ± 0.02
	0.10 – 0.17		0.13	$(3.58 \pm 0.12\ ^{+0.09}_{-0.05}) \cdot 10^{-1}$	$3.69 \cdot 10^{-1}$	0.345 ± 0.02	0.363 ± 0.01
	0.17 – 0.30		0.25	$(2.60 \pm 0.09\ ^{+0.08}_{-0.05}) \cdot 10^{-1}$	$2.53 \cdot 10^{-1}$	0.272 ± 0.02	0.257 ± 0.01
710 – 900	$(0.90 - 1.70) \cdot 10^{-2}$	800	$1.30 \cdot 10^{-2}$	$(8.84 \pm 0.23\ ^{+0.25}_{-0.25}) \cdot 10^{-1}$	$8.93 \cdot 10^{-1}$	0.852 ± 0.05	0.895 ± 0.03
	$(0.17 - 0.25) \cdot 10^{-1}$		$0.21 \cdot 10^{-1}$	$(7.74 \pm 0.24\ ^{+0.24}_{-0.23}) \cdot 10^{-1}$	$7.39 \cdot 10^{-1}$	0.822 ± 0.05	0.761 ± 0.03
	$(0.25 - 0.37) \cdot 10^{-1}$		$0.32 \cdot 10^{-1}$	$(6.50 \pm 0.21\ ^{+0.12}_{-0.12}) \cdot 10^{-1}$	$6.28 \cdot 10^{-1}$	0.593 ± 0.04	0.670 ± 0.02
	$(0.37 - 0.60) \cdot 10^{-1}$		$0.50 \cdot 10^{-1}$	$(5.35 \pm 0.16\ ^{+0.14}_{-0.08}) \cdot 10^{-1}$	$5.32 \cdot 10^{-1}$	0.563 ± 0.04	0.526 ± 0.02
	$(0.60 - 1.00) \cdot 10^{-1}$		$0.80 \cdot 10^{-1}$	$(4.51 \pm 0.15\ ^{+0.11}_{-0.08}) \cdot 10^{-1}$	$4.47 \cdot 10^{-1}$	0.460 ± 0.03	0.449 ± 0.02
	0.10 – 0.17		0.13	$(3.60 \pm 0.13\ ^{+0.05}_{-0.11}) \cdot 10^{-1}$	$3.68 \cdot 10^{-1}$	0.340 ± 0.03	0.367 ± 0.01
	0.17 – 0.30		0.25	$(2.66 \pm 0.11\ ^{+0.10}_{-0.07}) \cdot 10^{-1}$	$2.50 \cdot 10^{-1}$	0.261 ± 0.02	0.267 ± 0.01

Table C.1: Electromagnetic structure function F_2^{em} measured from the combined e^-p and e^+p data samples. The following quantities are given for each bin: the lower and upper limits of x and Q^2, the bin centre, (x_c, Q_c^2), the value of the combined F_2^{em} measurement with statistical and systematic uncertainties, the SM prediction calculated using CTEQ5D PDFs, and the values of F_2^{em} extracted using only the e^-p sample ($F_2^{e^-p}$) and the e^+p sample ($F_2^{e^+p}$). The first error of the combined measurement corresponds to the statistical uncertainty and the second one to the systematic uncertainty. Only statistical uncertainties are given for $F_2^{e^-p}$ and $F_2^{e^+p}$.

Q^2 range (GeV^2)	x range	Q_c^2 (GeV^2)	x_c	F_2^{em}			
				measured	SM	F_2^{e-p}	F_2^{e+p}
900 – 1300	$(0.10 - 0.17)\cdot 10^{-1}$	1200	$0.14\cdot 10^{-1}$	$(9.27 \pm 0.28\,^{+0.37}_{-0.50})\cdot 10^{-1}$	$8.89\cdot 10^{-1}$	0.906 ± 0.06	0.934 ± 0.03
	$(0.17 - 0.25)\cdot 10^{-1}$		$0.21\cdot 10^{-1}$	$(7.81 \pm 0.24\,^{+0.21}_{-0.25})\cdot 10^{-1}$	$7.54\cdot 10^{-1}$	0.677 ± 0.05	0.822 ± 0.03
	$(0.25 - 0.37)\cdot 10^{-1}$		$0.32\cdot 10^{-1}$	$(6.40 \pm 0.20\,^{+0.15}_{-0.18})\cdot 10^{-1}$	$6.37\cdot 10^{-1}$	0.583 ± 0.04	0.660 ± 0.02
	$(0.37 - 0.60)\cdot 10^{-1}$		$0.50\cdot 10^{-1}$	$(5.35 \pm 0.15\,^{+0.08}_{-0.17})\cdot 10^{-1}$	$5.36\cdot 10^{-1}$	0.481 ± 0.03	0.554 ± 0.02
	$(0.60 - 1.00)\cdot 10^{-1}$		$0.80\cdot 10^{-1}$	$(4.59 \pm 0.13\,^{+0.07}_{-0.08})\cdot 10^{-1}$	$4.47\cdot 10^{-1}$	0.422 ± 0.03	0.471 ± 0.02
	$0.10 - 0.17$		0.13	$(3.60 \pm 0.11\,^{+0.10}_{-0.06})\cdot 10^{-1}$	$3.66\cdot 10^{-1}$	0.355 ± 0.02	0.362 ± 0.01
	$0.17 - 0.30$		0.25	$(2.49 \pm 0.09\,^{+0.08}_{-0.05})\cdot 10^{-1}$	$2.46\cdot 10^{-1}$	0.257 ± 0.02	0.247 ± 0.01
	$0.30 - 0.53$		0.40	$(1.29 \pm 0.07\,^{+0.05}_{-0.07})\cdot 10^{-1}$	$1.29\cdot 10^{-1}$	0.104 ± 0.01	0.140 ± 0.01
1300 – 1800	$(0.17 - 0.25)\cdot 10^{-1}$	1500	$0.21\cdot 10^{-1}$	$(7.58 \pm 0.31\,^{+0.38}_{-0.40})\cdot 10^{-1}$	$7.62\cdot 10^{-1}$	0.786 ± 0.07	0.750 ± 0.04
	$(0.25 - 0.37)\cdot 10^{-1}$		$0.32\cdot 10^{-1}$	$(6.68 \pm 0.26\,^{+0.16}_{-0.13})\cdot 10^{-1}$	$6.42\cdot 10^{-1}$	0.731 ± 0.06	0.652 ± 0.03
	$(0.37 - 0.60)\cdot 10^{-1}$		$0.50\cdot 10^{-1}$	$(5.98 \pm 0.21\,^{+0.10}_{-0.12})\cdot 10^{-1}$	$5.39\cdot 10^{-1}$	0.575 ± 0.04	0.605 ± 0.02
	$(0.60 - 1.00)\cdot 10^{-1}$		$0.80\cdot 10^{-1}$	$(4.46 \pm 0.16\,^{+0.07}_{-0.08})\cdot 10^{-1}$	$4.47\cdot 10^{-1}$	0.430 ± 0.03	0.451 ± 0.02
	$0.10 - 0.15$		0.13	$(3.61 \pm 0.15\,^{+0.07}_{-0.10})\cdot 10^{-1}$	$3.65\cdot 10^{-1}$	0.380 ± 0.03	0.356 ± 0.02
	$0.15 - 0.23$		0.18	$(3.28 \pm 0.15\,^{+0.07}_{-0.07})\cdot 10^{-1}$	$3.08\cdot 10^{-1}$	0.358 ± 0.03	0.321 ± 0.02
	$0.23 - 0.35$		0.25	$(2.67 \pm 0.14\,^{+0.07}_{-0.05})\cdot 10^{-1}$	$2.43\cdot 10^{-1}$	0.263 ± 0.03	0.268 ± 0.02
	$0.35 - 0.53$		0.40	$(1.34 \pm 0.11\,^{+0.02}_{-0.11})\cdot 10^{-1}$	$1.27\cdot 10^{-1}$	0.130 ± 0.02	0.135 ± 0.01
1800 – 2500	$(0.23 - 0.37)\cdot 10^{-1}$	2000	$0.32\cdot 10^{-1}$	$(6.74 \pm 0.30\,^{+0.39}_{-0.28})\cdot 10^{-1}$	$6.48\cdot 10^{-1}$	0.652 ± 0.06	0.682 ± 0.03
	$(0.37 - 0.60)\cdot 10^{-1}$		$0.50\cdot 10^{-1}$	$(5.28 \pm 0.23\,^{+0.19}_{-0.10})\cdot 10^{-1}$	$5.42\cdot 10^{-1}$	0.526 ± 0.05	0.528 ± 0.03
	$(0.60 - 1.00)\cdot 10^{-1}$		$0.80\cdot 10^{-1}$	$(4.44 \pm 0.19\,^{+0.10}_{-0.07})\cdot 10^{-1}$	$4.48\cdot 10^{-1}$	0.396 ± 0.04	0.461 ± 0.02
	$0.10 - 0.15$		0.13	$(3.98 \pm 0.20\,^{+0.11}_{-0.09})\cdot 10^{-1}$	$3.63\cdot 10^{-1}$	0.370 ± 0.04	0.407 ± 0.02
	$0.15 - 0.23$		0.18	$(3.08 \pm 0.16\,^{+0.06}_{-0.06})\cdot 10^{-1}$	$3.06\cdot 10^{-1}$	0.263 ± 0.03	0.325 ± 0.02
	$0.23 - 0.35$		0.25	$(2.50 \pm 0.16\,^{+0.05}_{-0.11})\cdot 10^{-1}$	$2.40\cdot 10^{-1}$	0.272 ± 0.04	0.244 ± 0.02
	$0.35 - 0.53$		0.40	$(1.23 \pm 0.11\,^{+0.12}_{-0.03})\cdot 10^{-1}$	$1.24\cdot 10^{-1}$	0.112 ± 0.02	0.127 ± 0.01
2500 – 3500	$(0.37 - 0.60)\cdot 10^{-1}$	3000	$0.50\cdot 10^{-1}$	$(5.88 \pm 0.31\,^{+0.17}_{-0.14})\cdot 10^{-1}$	$5.45\cdot 10^{-1}$	0.620 ± 0.06	0.578 ± 0.03
	$(0.60 - 1.00)\cdot 10^{-1}$		$0.80\cdot 10^{-1}$	$(4.31 \pm 0.23\,^{+0.06}_{-0.09})\cdot 10^{-1}$	$4.48\cdot 10^{-1}$	0.435 ± 0.05	0.430 ± 0.03
	$0.10 - 0.15$		0.13	$(3.49 \pm 0.21\,^{+0.20}_{-0.06})\cdot 10^{-1}$	$3.61\cdot 10^{-1}$	0.340 ± 0.04	0.352 ± 0.02
	$0.15 - 0.23$		0.18	$(3.16 \pm 0.19\,^{+0.08}_{-0.06})\cdot 10^{-1}$	$3.02\cdot 10^{-1}$	0.287 ± 0.04	0.325 ± 0.02
	$0.23 - 0.35$		0.25	$(2.35 \pm 0.17\,^{+0.05}_{-0.05})\cdot 10^{-1}$	$2.36\cdot 10^{-1}$	0.258 ± 0.04	0.229 ± 0.02
	$0.35 - 0.53$		0.40	$(1.40 \pm 0.14\,^{+0.06}_{-0.06})\cdot 10^{-1}$	$1.20\cdot 10^{-1}$	0.137 ± 0.03	0.140 ± 0.02
	$0.53 - 1.00$		0.65	$(1.95 \pm 0.32\,^{+0.18}_{-0.25})\cdot 10^{-2}$	$1.85\cdot 10^{-2}$	0.015 ± 0.01	0.021 ± 0.00
3500 – 5600	$(0.40 - 1.00)\cdot 10^{-1}$	5000	$0.80\cdot 10^{-1}$	$(4.61 \pm 0.22\,^{+0.16}_{-0.11})\cdot 10^{-1}$	$4.48\cdot 10^{-1}$	0.471 ± 0.04	0.458 ± 0.02
	$0.10 - 0.15$		0.13	$(3.77 \pm 0.24\,^{+0.08}_{-0.10})\cdot 10^{-1}$	$3.58\cdot 10^{-1}$	0.369 ± 0.05	0.380 ± 0.03
	$0.15 - 0.23$		0.18	$(2.93 \pm 0.20\,^{+0.09}_{-0.10})\cdot 10^{-1}$	$2.98\cdot 10^{-1}$	0.335 ± 0.04	0.282 ± 0.02
	$0.23 - 0.35$		0.25	$(2.21 \pm 0.17\,^{+0.08}_{-0.07})\cdot 10^{-1}$	$2.31\cdot 10^{-1}$	0.238 ± 0.04	0.217 ± 0.02
	$0.35 - 0.53$		0.40	$(1.23 \pm 0.13\,^{+0.03}_{-0.04})\cdot 10^{-1}$	$1.16\cdot 10^{-1}$	0.130 ± 0.03	0.121 ± 0.01
5600 – 9000	$(0.70 - 1.50)\cdot 10^{-1}$	8000	$1.30\cdot 10^{-1}$	$(3.75 \pm 0.26\,^{+0.11}_{-0.11})\cdot 10^{-1}$	$3.55\cdot 10^{-1}$	0.381 ± 0.05	0.373 ± 0.03
	$0.15 - 0.23$		0.18	$(2.99 \pm 0.27\,^{+0.11}_{-0.09})\cdot 10^{-1}$	$2.94\cdot 10^{-1}$	0.318 ± 0.05	0.292 ± 0.03
	$0.23 - 0.35$		0.25	$(2.21 \pm 0.23\,^{+0.09}_{-0.06})\cdot 10^{-1}$	$2.26\cdot 10^{-1}$	0.229 ± 0.05	0.219 ± 0.03
	$0.35 - 0.53$		0.40	$(1.03 \pm 0.15\,^{+0.09}_{-0.04})\cdot 10^{-1}$	$1.12\cdot 10^{-1}$	0.112 ± 0.04	0.101 ± 0.02
	$0.53 - 1.00$		0.65	$(1.18 \pm 0.36\,^{+0.18}_{-0.04})\cdot 10^{-2}$	$1.65\cdot 10^{-2}$	0.015 ± 0.01	0.011 ± 0.00
9000 – 15000	$0.11 - 0.23$	12000	0.18	$(3.34 \pm 0.34\,^{+0.14}_{-0.26})\cdot 10^{-1}$	$2.91\cdot 10^{-1}$	0.290 ± 0.05	0.364 ± 0.04
	$0.23 - 0.35$		0.25	$(2.16 \pm 0.31\,^{+0.11}_{-0.08})\cdot 10^{-1}$	$2.22\cdot 10^{-1}$	0.207 ± 0.06	0.220 ± 0.04
	$0.35 - 0.53$		0.40	$(1.20 \pm 0.22\,^{+0.06}_{-0.13})\cdot 10^{-1}$	$1.09\cdot 10^{-1}$	0.100 ± 0.04	0.126 ± 0.03
15000 – 25000	$0.18 - 0.35$	20000	0.25	$(1.93 \pm 0.41\,^{+0.56}_{-0.21})\cdot 10^{-1}$	$2.18\cdot 10^{-1}$	0.300 ± 0.07	0.146 ± 0.05
	$0.35 - 1.00$		0.40	$(7.1 \pm 2.5\,^{+1.2}_{-0.4})\cdot 10^{-2}$	$10.6\cdot 10^{-2}$	0.100 ± 0.05	0.061 ± 0.03
25000 – 50000	$0.30 - 1.00$	30000	0.40	$(7.0 \pm 3.6\,^{+5.4}_{-0.3})\cdot 10^{-2}$	$10.3\cdot 10^{-2}$	0.112 ± 0.06	0.050 ± 0.04

Table C.2: Electromagnetic structure function F_2^{em} measured from the combined e^-p and e^+p data samples. The following quantities are given for each bin: the lower and upper limits of x and Q^2, the bin centre, (x_c, Q_c^2), the value of the combined F_2^{em} measurement with statistical and systematic uncertainties, the SM prediction calculated using CTEQ5D PDFs, and the values of F_2^{em} extracted using only the e^-p sample (F_2^{e-p}) and the e^+p sample (F_2^{e+p}). The first error of the combined measurement corresponds to the statistical uncertainty and the second one to the systematic uncertainty. Only statistical uncertainties are given for F_2^{e-p} and F_2^{e+p}.

Q^2 range (Gev^2)	x range	Q_c^2 (Gev^2)	x_c	xF_3		N^{e-p}	N^{e+p}
				measured	SM		
$1300 - 2500$	$0.017 - 0.060$	1500	0.050	$(6.72 \pm 3.64^{+0.18}_{-0.16}) \cdot 10^{-2}$	$3.7 \cdot 10^{-2}$	833	2819
	$0.060 - 0.230$		0.130	$(-0.7 \pm 6.7 \pm 0.01) \cdot 10^{-2}$	$4.418 \cdot 10^{-2}$	723	2796
	$0.230 - 0.530$		0.400	$(3.72 \pm 14.8 \pm 0.07) \cdot 10^{-2}$	$2.089 \cdot 10^{-2}$	210	793
$2500 - 3500$	$0.037 - 0.060$	3000	0.050	$(10.1 \pm 5.6 \, ^{+0.03}_{-0.02}) \cdot 10^{-2}$	$6.351 \cdot 10^{-2}$	99	301
	$0.060 - 0.230$		0.130	$(5.09 \pm 6.00 \, ^{+0.10}_{-0.08}) \cdot 10^{-2}$	$7.621 \cdot 10^{-2}$	209	752
	$0.230 - 0.750$		0.400	$(3.78 \pm 12.0 \, ^{+0.11}_{-0.10}) \cdot 10^{-2}$	$3.458 \cdot 10^{-2}$	68	253
$3500 - 5600$	$0.040 - 0.230$	5000	0.130	$(13.8 \pm 3.7 \pm 0.30) \cdot 10^{-2}$	$10.761 \cdot 10^{-2}$	255	734
	$0.230 - 0.530$		0.400	$(9.17 \pm 7.5 \pm 0.26) \cdot 10^{-2}$	$4.746 \cdot 10^{-2}$	63	202
$5600 - 9000$	$0.070 - 0.230$	8000	0.180	$(15.8 \pm 4.6 \pm 0.42) \cdot 10^{-2}$	$13.319 \cdot 10^{-2}$	98	242
	$0.230 - 0.750$		0.400	$(8.07 \pm 5.6 \, ^{+0.34}_{-0.23}) \cdot 10^{-2}$	$6.032 \cdot 10^{-2}$	39	114
$9000 - 15000$	$0.110 - 0.230$	12000	0.180	$(12.4 \pm 5.9 \, ^{+0.60}_{-0.80}) \cdot 10^{-2}$	$15.942 \cdot 10^{-2}$	30	72
	$0.230 - 0.530$		0.400	$(4.63 \pm 5.02 \, ^{+0.17}_{-0.24}) \cdot 10^{-2}$	$7.103 \cdot 10^{-2}$	21	60
$15000 - 50000$	$0.180 - 0.750$	30000	0.400	$(12.5 \pm 2.9 \, ^{+1.32}_{-2.36}) \cdot 10^{-2}$	$8.946 \cdot 10^{-2}$	26	22

Table C.3: Table of measured xF_3 values. For each bin, the lowest and highest x and Q^2 are given in the table. The centre of each bin is denoted as x_c and Q_c^2. The measured values of xF_3 are shown together with the statistical (first) and systematic (second) uncertainties. The value of xF_3 predicted by the SM using CTEQ5D PDFs is also shown. The last two columns show the number of events found in each bin for the e^-p and the e^+p data samples, respectively. The estimated number of background events was found to be negligible.

Appendix D

List of acronyms

BCAL	–	Barrel Calorimeter
BGF	–	Boson Gluon Fusion
BPC	–	Beampipe Calorimeter
CAL	–	Calorimeter
CC	–	Charged Current
CDM	–	Colour Dipole Model
CERN	–	European organisation for nuclear research
CFLT	–	Calorimeter First Level Trigger
CTD	–	Central Tracking Detector
CTEQ	–	Coordinated Theoretical-Experimental Project on QCD
DESY	–	Deutsches Elektronen Synchrotron
DGLAP	–	Dokshitzer-Gribov-Lipatov-Altarelli-Parisi evolution equations
DIS	–	Deep Inelastic Scattering
EAZE	–	Effortless Analysis of ZEUS Events
EMC	–	Electromagnetic Calorimeter
EVB	–	Event Builder
FC	–	Fast Clear processor
FCAL	–	Forward Calorimeter
FLT	–	First Level Trigger
FPC	–	Forward Plug Calorimeter
FSR	–	Final-State Radiation
GFLT	–	Global First Level Trigger
GSLT	–	Global Second Level Trigger
GRV	–	Glück-Reya-Vogt
HAC	–	Hadronic Calorimeter
HERA	–	Hadron Electron Ring Anlage
HES	–	Hadron-Electron Separator
ISR	–	Initial-State Radiation
LEP	–	Large Electron-Positron collider
LHC	–	Large Hadron Collider

LO	–	Leading Order
MC	–	Monte Carlo simulation
MEPS	–	Matrix Elements and Parton Showers model of QCD cascade
MRST	–	Martin-Roberts-Stirling-Thorne
NC	–	Neutral Current
NLO	–	Next-to-Leading Order
ORANGE	–	Overlying Routine for Analysis Ntuple Generation
PDF	–	Parton Density Function
PHP	–	Photoproduction process
PMT	–	Photomultiplier Tube
QED	–	Quantum Electrodynamics
QEDC	–	Quantum Electrodynamics Compton process
QCD	–	Quantum Chromodynamics
QPM	–	Quark-Parton Model
RCAL	–	Rear Calorimeter
SLAC	–	Stanford Linear Accelerator Center
SLT	–	Second Level Trigger
SM	–	Standard Model of particle physics
TLT	–	Third Level Trigger
WLS	–	Wavelength Shifter
ZES	–	ZEUS Event Store