

Open Charm Production in Deep
Inelastic Diffractive ep Scattering at
HERA

Joanne Elise Cole

High Energy Physics Group
Blackett Laboratory
Imperial College of Science, Technology and Medicine
University of London

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

January 1999

Abstract

A study of open charm production in deep inelastic diffractive scattering at HERA, using 43.3 pb^{-1} of data taken with the ZEUS detector, is presented. Diffractive interactions in which the virtual photon dissociates into a hadronic system X and the proton remains intact or dissociates into a low mass state, Y , are identified by the presence of a large rapidity gap in the final state. Charm production is identified by the reconstruction of $D^{*\pm}(2010)$ mesons in the decay channel $D^{*+} \rightarrow (D^0 \rightarrow K^-\pi^+)\pi^+ + (\text{c.c.})$.

The accessible kinematic region for this measurement is $6 < Q^2 < 400 \text{ GeV}^2$, $0.02 < y < 0.7$, $p_T(D^{*\pm}) > 1.5 \text{ GeV}$, $|\eta(D^{*\pm})| < 1.5$, $0.001 < x_P < 0.016$ and $\beta < 0.8$, in which the cross section for diffractive $D^{*\pm}$ production is measured to be $245 \pm 39(\text{stat})_{-60}^{+66}(\text{syst}) \text{ pb}$. The ratio of diffractive to total $D^{*\pm}$ production in this region is also measured and found to be $6.3 \pm 1.1(\text{stat})_{-1.4}^{+1.6}(\text{syst}) \%$, which is consistent with the fraction of diffractively-produced events in inclusive deep inelastic scattering.

Differential $D^{*\pm}$ cross sections are also measured as a function of Q^2 , W , x_P , $p_T(D^{*\pm})$ and $\eta(D^{*\pm})$ and are compared to expectations from different phenomenological models. The data clearly show that diffractive heavy flavour production is not suppressed relative to light flavours and favour a substantial hard gluon component in the diffractive exchange. Limited statistics and significant theoretical uncertainties prevent further discrimination between different theoretical models.

Contents

1	Introduction	23
1.1	Deep Inelastic Scattering	24
1.2	Proton Structure Functions	26
1.2.1	The Quark Parton Model	27
1.2.2	Quantum Chromodynamics	28
1.3	The Structure Functions in QCD	29
1.4	Factorisation	33
1.5	Charm Production at HERA	33
1.5.1	Charm Evolution	36
1.5.2	Charm Hadronisation	39
1.6	Parametrisations of the Parton Distribution Functions	40
1.7	Summary	42
2	Diffraction and Open Charm Production	43
2.1	Regge Theory and the Soft Pomeron	45
2.2	Regge Theory at HERA	47
2.3	Models of hard diffractive scattering at HERA	49
2.3.1	The Non-perturbative Factorisable Pomeron Model	51

2.3.2	Perturbative QCD Pomeron Models	54
2.3.3	Soft Colour Interaction Models	58
2.4	Charm Production in Diffractive Scattering	59
2.5	Summary	63
3	HERA and the ZEUS Detector	65
3.1	The HERA Collider	65
3.2	HERA Performance 1995 - 1997	67
3.3	The ZEUS Detector	67
3.3.1	The Central Tracking Detector	71
3.3.2	The Calorimeter	72
3.3.3	Preshower Detectors	74
3.3.4	Luminosity Measurement	76
4	Data Preselection	79
4.1	The Data Acquisition System	79
4.1.1	Backgrounds	80
4.1.2	The ZEUS Trigger System	81
4.1.3	Simulation of the Detector and DAQ System	85
4.2	Offline Reconstruction	86
4.2.1	Calorimeter Clustering	86
4.2.2	Track and Vertex Reconstruction	87
4.3	Post-Reconstruction Data Analysis	89
4.3.1	Noise Suppression	91
4.3.2	Positron Identification	92

4.3.3	Positron Position Reconstruction	94
4.3.4	Positron Energy Loss Corrections	96
4.3.5	Calorimeter Energy Scale	97
4.3.6	Hadronic Energy Flow Reconstruction	98
4.4	Kinematic Variable Reconstruction	100
4.4.1	The Electron Method	100
4.4.2	The Jacquet-Blondel Method	103
4.4.3	The Double Angle Method	104
4.4.4	Radiative Events	106
4.5	Summary	106
5	SRTD Energy Corrections	107
5.1	Introduction	107
5.2	The Small Angle Rear Tracking Detector	109
5.3	Selection of Data Samples	110
5.3.1	Kinematic Peak Events	111
5.3.2	DIS ρ^0 Events	112
5.3.3	Elastic QED Compton Events	114
5.4	Methods of Energy Prediction	116
5.4.1	Kinematic Peak Sample	116
5.4.2	DIS ρ^0 Sample	117
5.4.3	Elastic QED Compton Sample	122
5.5	Dead Material Distributions	122
5.6	Positron Energy Correction Results	124

5.6.1	Results from Method I	126
5.6.2	Results from Method II	133
5.6.3	Comparison of Method Results	133
5.7	Conclusions	137
6	Selection of Candidates	141
6.1	Monte Carlo Simulation	141
6.2	Trigger Strategy	143
6.2.1	FLT Strategy	144
6.2.2	TLT Strategy	144
6.3	DIS Event Selection	148
6.4	$D^{*\pm}(2010)$ Candidate Selection	153
6.4.1	Track Momentum Requirements	153
6.4.2	Invariant Mass Distributions	159
6.5	Diffractive Event Selection	161
6.6	Characteristics of the Data Sample	164
6.7	Summary	168
7	Extraction of Cross Sections	171
7.1	The Kinematic Region	171
7.2	Data — Monte Carlo Comparisons	175
7.3	Bin Selection	177
7.4	Backgrounds	179
7.5	Cross Section Extraction	183
7.6	Systematic Error Evaluation	187

7.7	Ratio Extraction	194
7.8	Comparison with Theoretical Predictions	197
8	Conclusions and Outlook	207

List of Figures

1.1	Diagrams for DIS via a neutral current interaction (left) and a charged current interaction (right).	25
1.2	Lepton-parton scattering in the Quark Parton Model, where the struck parton carries away a fraction ϵ of the proton's momentum.	27
1.3	The structure function $F_2(x, Q^2)$ versus Q^2 at various fixed values of Bjorken x . Results from both HERA and fixed target experiments are shown. The evidence for scaling violations at low x are clear from the HERA data, while there is no significant Q^2 dependence in the data from the fixed target experiments at higher x , ie. Bjorken scaling is observed.	31
1.4	The Altarelli-Parisi splitting functions $P_{jk}(x/y)$ (for further explanation see text).	32
1.5	An example of a ladder diagram in QCD, showing parton evolution in the proton.	32
1.6	The leading order diagrams for boson gluon fusion (left) and flavour excitation (right).	34
1.7	The x_D distribution extracted from 1994 H1 data. The black dots are from $D^{*\pm}$ data, while the open dots are from D^0 data. The shaded histogram indicates the expected x_D distribution from the BGF process, while the dotted line indicates that expected from flavour excitation. The solid curve represents measurements of charm production in νN scattering made by the CDHS and E531 experiments [14]. The H1 data clearly favour the BGF mechanism while the neutrino scattering results favour flavour excitation.	35

1.8	The gluon distribution extracted from H1 charm data as a function of x . Results from both DIS and photoproduction data are shown and compared to the gluon density extracted using QCD fits to measurements of $F_2(x, Q^2)$ (grey band). The prediction of the CTEQ parton distribution function parametrisation [23] is also shown (solid line).	36
1.9	The structure function $F_2^{c\bar{c}}$ as a function of x at various fixed values of Q^2 . The latest ZEUS and H1 results are compared to a NLO QCD calculation which uses GRV as the input gluon density (see section 1.7 for more information). The width of the band reflects a variation in charm mass from 1.2 to 1.6 GeV.	37
1.10	The predictions for F_2 and $F_2^{c\bar{c}}$ from three different parametrisations (see text) as a function of x for three different values of Q^2	42
2.1	The three different types of interaction in which a colourless exchange can occur: (a) elastic scattering, (b) single diffractive and (c) double diffractive.	44
2.2	The spin versus t plot made by Chew and Frautschi. The line indicates the fit that they performed, which was found to be $\alpha = 0.55 + 0.86t$	45
2.3	Fits performed by Donnachie and Landshoff to pp , $p\bar{p}$ and $\pi^\pm p$ data [38]. All the fits clearly indicate the pomeron intercept to be $\alpha_P(0) = 1.0808$	47
2.4	The Donnachie-Landshoff fit to γp data [38, 18]. The measurements of the total cross section at HERA are shown at $\sqrt{s} \simeq 200$ GeV, where \sqrt{s} is in this case equivalent to W . The fixed target measurements which were used in the original fits are also shown at low \sqrt{s}	48
2.5	A generic diffractive interaction at HERA in which the proton remains intact and the photon dissociates into a hadronic system X	50
2.6	The simplest leading order interpretation of the perturbative two-gluon exchange model. There are three other diagrams which contribute at this order.	54

2.7	Two examples of higher order real gluon corrections. On the left, the case where the gluon is softer than either of the quarks is shown. On the right, one of the quarks is the softest parton.	56
2.8	Two examples of higher order virtual gluon corrections, whose contribution to the diffractive cross section is estimated in the LMRT model. The dotted lines are the virtual gluon loops.	57
2.9	The ratio of diffractively produced charm at $x_p = 0.005$ as a function of Q^2 at fixed β and β at fixed Q^2 . Various different models are compared, the details of which are given in the text.	61
3.1	A Schematic of the HERA ring, including the pre-accelerator chain.	66
3.2	The HERA luminosity delivered to the ZEUS experiment since the start of data-taking in 1992.	67
3.3	2-d cross section of ZEUS in the $r - \phi$ plane.	68
3.4	2-d cross-section parallel to the beam pipe of the ZEUS detector.	68
3.5	An octant of the CTD. The larger dots indicate the sense wires	72
3.6	The FCAL as seen from the interaction point	74
3.7	One of the FCAL modules in the ZEUS detector.	75
3.8	The Layout of the ZEUS Luminosity Monitor and paths of the different particles from the Interaction Point (IP).	76
4.1	The topology expected from (a) a standard ep event and (b) a proton beam-gas interaction occurring prior to the entry of the proton beam into the main detector volume.	81
4.2	The different methods of clustering used in the positron finders implemented in the Third Level Trigger. (a) shows the cone method of clustering, while (b) shows the island method.	84

4.3	Integrated luminosity for high quality data versus day of running for all the positron-proton datasets. EVTAKA is part of the data quality monitoring procedure.	90
4.4	The Cell energy imbalance as a function of cell energy in (a) data, and (b) Monte Carlo.	92
4.5	SINISTRA efficiency estimated from QED Compton data and Monte Carlo taken from [8].	94
4.6	The layout of the 3×3 array of towers used by ELECPO to determine the position of the scattered positron. “Emax” indicates the highest energy cell and the strips and module labels are indicated.	96
4.7	The different types of energy flow objects which can be formed using a combination of calorimeter and CTD information.	99
4.8	Schematic representing a standard ep collision, indicating the main quantities which can be used to reconstruct the event kinematic variables x , y and Q^2	101
4.9	Isolines in the x — Q^2 plane of (a) positron energy, (b) positron angle, (c) struck quark energy and (d) struck quark angle.	102
4.10	The fractional difference between the reconstructed and true values of Q^2 , y and x , using the electron method, from a sample of DIS Monte Carlo as a function of the true variable. The error bars are the RMS widths of the fractional differences.	103
4.11	The fractional difference between the reconstructed and true values of Q^2 , y and x using the Jacquet-Blondel method from a sample of DIS Monte Carlo as a function of the true variables. The error bars are the RMS widths of the fractional difference distributions.	104
4.12	The fractional difference between the reconstructed and true values of Q^2 , y and x using the Double Angle method, from a sample of DIS Monte Carlo as a function of the true variables. The error bars are the RMS widths of the fractional difference distributions.	105

5.1	Scattered positron energy distributions for 1995 data (black dots) and Monte Carlo (histogram) where the positron is detected in the SRTD region. Standard DIS cuts have been applied to obtain the distributions. Energy loss by the scattered positron is indicated by the shift of the peak to a value which is lower than the nominal positron beam energy (see text). There is also a noticeable discrepancy between data and Monte Carlo, indicating that the energy loss is inaccurately simulated in Monte Carlo.	108
5.2	Layout for the SRTD as in the 1995 Data Taking Period. The grey hatched areas indicate the four SRTD quadrants, which are numbered 1 to 4 starting in the top right-hand corner and going round anti-clockwise. The RCAL cells are also indicated and are numbered for use in the comparison of data and Monte Carlo kinematic peak events, as performed in section 5.6.3. . . .	109
5.3	Isolines of Scattered Positron Energy (GeV) as seen in the $x - Q^2$ plane. . .	111
5.4	The Feynman diagrams for the elastic QED Compton process.	114
5.5	The “True” Kinematic Peak energy distribution, taken from Monte Carlo. The line indicates a fit to the peak to obtain the predicted energy for this sample.	117
5.6	a) Comparison of the predicted energy and the “true” energy from the Monte Carlo generator level, without taking into account any initial state radiation in the prediction method, b) the same comparison, but taking into account initial state radiation and c) comparison of the predicted and true energies as a function of proton p_T	119
5.7	a) Comparison of $E_{pred} - E_{true}$ as a function of the p_T -corrected $E - P_z$. The tail to low values of $E - P_z$ is predominantly due to ISR events. b) Fraction 1 (as described in the main text) as a function of $(E - P_z)_{corr}$ cut. c) Fraction 2 (also described in the main text) as a function of $(E - P_z)_{corr}$ cut.	120
5.8	a) The effect of requiring $(E - P_z)_{corr} > 50$ GeV on the absolute difference between the predicted and true energies. b) The resolution of the ρ^0 energy prediction method as a function of the predicted energy.	121

5.9	a) Comparison of the predicted energy and the “true” energy from the Monte Carlo generator level, without taking into account any initial state radiation and under the assumption that proton p_T is small compared to p_L . b) The same comparison, taking into account ISR. c) Comparison of the predicted and “true” energies as a function of proton p_T	123
5.10	a) Comparison of the reconstructed predicted energy and the generator level energy as a function of the predicted energy. The agreement is good within the range $6 < E_{pred} < 23$ GeV (as shown by the vertical lines). b) The resolution of the elastic QED Compton energy prediction method as a function of predicted energy. This is fitted to determine the parametrisation of the resolution used in the correction determination.	124
5.11	a) The SRTD hit distribution in the Monte Carlo for hits with $E_{SRTD}/2 < 10$ MIPS. b) The hit distribution for the same sample, but for hits with $E_{SRTD}/2 > 60$ MIPS. c) The SRTD hit distribution in the data kinematic peak sample for hits with $E_{SRTD}/2 < 10$ MIPS. d) The hit distribution for the same sample, but for hits with $E_{SRTD}/2 > 60$ MIPS. e) A Comparison of $E_{SRTD}/2$ versus radius for data and Monte Carlo.	125
5.12	Uncorrected positron energy, E_{CAL} , versus SRTD energy (in MIPS) for each of the three data samples. The elastic QED Compton sample has been divided up into four bins in predicted positron energy to gain wider positron energy coverage. The behaviour is linear in all bins and the number given on each plot is the slope from a linear fit.	128
5.13	Uncorrected positron energy, E_{CAL} , versus SRTD energy for each of the three Monte Carlo samples. The elastic QED Compton sample has been divided into four bins in predicted positron energy. The curves on the plots are the results of fitting a second order polynomial and the fit parameters are summarised in table 5.1.	129

5.14	The energy dependence of the correction parameters for both Data and Monte Carlo. The linear and quadratic fit parameters are plotted versus the mean corrected energy of the sample. The filled circles indicate the elastic QED Compton sample, the open circles the DIS ρ sample and the open triangles the kinematic peak sample. The dotted lines indicate the weighted mean values of the correction parameter in each plot.	131
5.15	Comparison between data and Monte Carlo scattered positron energy spectra, E_{corr} , after the dead material correction has been applied. There is a clear shift between the two distributions even after the SRTD corrections have been applied.	132
5.16	The distribution of the shifts to the predicted energy in data and Monte Carlo as a function of corrected energy (before the shift is applied).	133
5.17	Comparison of the predicted energy to the corrected energy in Monte Carlo for the three samples as a function of SRTD energy, using the two different methods (see text). The error bars indicate the RMS in each bin of SRTD energy. A clear discrepancy can be seen in the kinematic peak sample corrected using method I.	135
5.18	Comparison of scattered positron variables, before and after correction, between Data and Monte Carlo. The Monte Carlo is weighted by luminosity and is re-weighted to the structure function CTEQ4D.	136
5.19	Cell-by-cell comparison of the data and Monte Carlo kinematic peak samples. There are clear cell-to-cell variations, which could be accounted for by differences in cell-by-cell calibration between data and Monte Carlo.	138
6.1	The efficiencies for FLT46 (see text) after the final DIS $D^{*\pm}$ cuts in both data (closed circles) and Monte Carlo (open triangles) as a function of Q^2 , W , $D^{*\pm}$ transverse momentum and $D^{*\pm}$ pseudorapidity.	145
6.2	The efficiencies for DIS13 (see text) after the final DIS $D^{*\pm}$ cuts in both data (closed circles) and Monte Carlo (open triangles) as a function of Q^2 , W , $D^{*\pm}$ transverse momentum and $D^{*\pm}$ pseudorapidity.	147

6.3	The SINISTRA probability distribution from non-diffractive RAPGAP Monte Carlo. The shaded area indicates those candidates which are kept in the final $D^{*\pm}$ sample. Note that the peak near 1 is several orders of magnitude larger than the tail.	149
6.4	The y_e distribution from the data, after all DIS selection requirements have been applied. The peak at very high y comes from “fake” positron events. The shaded region indicates the events which are excluded by the $y_{DA} < 0.7$ requirement. This clearly excludes the “fake” positron background.	150
6.5	The $Q^2 - x$ distribution for data which satisfy all DIS requirements and $D^{*\pm}$ preselection. The data clearly lie predominantly above $Q^2 = 1 \text{ GeV}^2$	152
6.6	The generated transverse momentum distribution of the slow pion taken from the non-diffractive RAPGAP sample.	155
6.7	The relevant distributions for the selection of the $p_T(K, \pi)$ cut. Figure (a) shows the efficiency of $D^{*\pm}$ reconstruction versus cut value, while figure (b) shows the purity versus cut value. The definitions of efficiency and purity are given in the text. Figures (c) and (d) show the p_T distributions for K and π candidates from Monte Carlo (solid line) and from the wrong charge combinations in the data, as a background example (dotted line).	156
6.8	The relevant distributions for the selection of the $ p_T(K) + p_T(\pi) $ cut. Figures (a) and (b) show the efficiency and purity respectively versus cut value. The definitions of efficiency and purity are similar to those used in the $p_T(K, \pi)$ cuts. Figures (c) and (d) show $p_T(K)$ versus $p_T(\pi)$ in the Monte Carlo and in the background from the data respectively. The diagonal line indicates the cut which is used in this analysis. The size of each box is proportional to the density of the points on the original scatter plot.	157
6.9	The ΔM distribution in data after the application of all cuts except the $p(D^0)/p(\pi_s)$ cut (solid line). The dashed line shows the effect of applying the $p(D^0)/p(\pi_s)$ cut.	158

6.10	(a) The ΔM distribution after the application of all DIS and $D^{*\pm}$ selection cuts. The data exhibit a clear peak (black dots). The solid line is a fit performed to the distribution (see text) and the dashed line is an estimate of the combinatorial background using wrong charge combinations from the data. (b) The $M(D^0)$ distribution after the application of all DIS and $D^{*\pm}$ selection cuts. The distribution has been fitted (see text) and the results are indicated by the solid line. The expected two-peak structure is observed (see text for details).	160
6.11	The η_{MAX} distribution for data $D^{*\pm}$ candidates which lie in both the $M(D^0)$ and ΔM signal windows. The shaded area indicates the candidates which are kept for the measurement of the cross section (see text).	162
6.12	The θ_H versus η_{MAX} distribution for (a) data and (b) - (f) Monte Carlo. Figure (b) shows the distribution for non-diffractive $D^{*\pm}$ Monte Carlo, while (c) and (d) are from the two different diffractive samples. Figures (e) and (f) both come from the inclusive DIS Monte Carlo sample, but the latter has $D^{*\pm}$ finding performed, while the former does not.	163
6.13	(a) The ΔM distribution for events which satisfy all diffractive selection requirements, as well as the DIS and $D^{*\pm}$ requirements. Once again, the solid line is a fit to the distribution from which the number of events has been extracted and the dashed line is an estimate of the combinatorial background made using wrong-charge combinations. (b) The $M(D^0)$ distribution for events which satisfy all selection requirements fitted in the same manner as the distribution in figure 6.10. The expected two-peak structure is not as clear due to the limited statistics available.	164
6.14	The average multiplicity versus η_{MAX} in data, calculated using ZUFOs. The expected increase in multiplicity with increasing η_{MAX} is observed. . .	166
6.15	The $\ln M_X^2$ distribution and the M_X versus W distribution from the final DIS $D^{*\pm}$ data sample. The candidates which satisfy the $\eta_{MAX} < 2$ requirement and which lie in the kinematic region chosen for cross section measurements in the next chapter are marked by the hatched area in the left-hand plot and the larger triangles in the right-hand plot.	167

6.16	The $Q^2 - x$ distribution from the final DIS $D^{*\pm}$ sample. Those marked by triangles are the candidates which satisfy the diffractive requirements. . . .	167
6.17	One of the diffractive $D^{*\pm}$ candidates selected from 1995 data. The $r-\phi$ view (left) shows the small number of tracks identified in the event, three of which make up the $D^{*\pm}$ candidate. The $r-z$ view (right) shows that the event clearly contains a rapidity gap in the forward region of the detector. The calorimeter cluster identified as the scattered positron candidate is also indicated.	169
7.1	The acceptances versus $p_T(D^{*\pm})$ and $\eta(D^{*\pm})$ from the diffractive RAPGAP sample. The plots indicate that $p_T(D^{*\pm}) > 1.5$ GeV and $ \eta(D^{*\pm}) < 1.5$ is an acceptable kinematic region for the cross section measurements.	173
7.2	The $x_P - \beta$ distributions from the diffractive RAPGAP and RIDI Monte Carlo samples, before and after the application of the η_{MAX} cut. There is a substantial effect on the accessible x_P range. It should also be noted that x_{pom} is equivalent to x_P in all subsequent figures.	174
7.3	(a) The $x_P - \beta$ distribution from the data, indicating that the majority of candidates lie within the chosen kinematic region, (b) the efficiency versus x_P cut determined using the diffractive RAPGAP sample (see text for details).	175
7.4	Comparisons between data (black dots), RAPGAP (solid lines) and RIDI (dotted lines) Monte Carlo for a wide variety of DIS and $D^{*\pm}$ quantities. . .	178
7.5	A comparison of the β and x_P distributions in data and Monte Carlo. The data is indicated by black dots, while RAPGAP is represented by the solid lines and RIDI by the dotted lines. See text for discussion of the distribution shapes.	179
7.6	The ΔM distributions in each of the cross sections bins. The bin numbers increase with the bin quantity. The curves indicate maximum likelihood fits to each distribution. The numbers of $D^{*\pm}$ candidates are given in table 7.1.	180

7.7	The resolutions and purities in each cross section bin. The left-hand column of plots shows $(X_{rec} - X_{true})/X_{true}$ in each of the cross section bins. The error bars are the width of the distribution in each bin, giving an estimate of the resolution for each bin. The central point gives an estimate of the accuracy with which the quantity is reconstructed. The right-hand column of plots shows the purity of each cross section bin.	181
7.8	The reconstructed η_{MAX} distribution from the non-diffractive RAPGAP Monte Carlo sample.	182
7.9	The three regions into which the face of the RCAL is divided for the simulation of the TLT trigger strategy in Monte Carlo.	184
7.10	The acceptances in the different cross section bins calculated using the diffractive RAPGAP Monte Carlo sample.	185
7.11	The cross sections resulting from each systematic error check versus check number for the total cross section (top) and in the Q^2 and W bins.	190
7.12	The cross sections resulting from each systematic error check versus check number for the x_P , $p_T(D^{*\pm})$ and $\eta(D^{*\pm})$ bins.	191
7.13	The differential cross sections calculated from the diffractive $D^{*\pm}$ data sample, using diffractive RAPGAP Monte Carlo to unfold from the measured number of candidates. The black dots are the points from this analysis, while the open triangles are the preliminary ZEUS results shown at ICHEP98. The inner error bars indicate the statistical error, while the outer bars are the statistical and systematic errors added in quadrature. The systematic shift between the two sets of results is caused by the kinematic region in which the cross sections are measured being different. . . .	193
7.14	The DIS ΔM distributions in the Q^2 bins (top row) and W bins (bottom row) used in the calculation of the $D^{*\pm}$ ratios.	195

7.15	The differential ratios of diffractive $D^{*\pm}$ production to DIS $D^{*\pm}$ production as a function of Q^2 and W . The inner error bars indicate the statistical errors, while the outer bars indicate the statistical and systematic errors added in quadrature. The solid horizontal line indicates the total ratio, while the shaded band indicates its statistical and systematic errors added in quadrature.	196
7.16	The final differential cross sections compared to the RAPGAP implementation of the Ingelman-Schlein model using hard (solid line) and soft (dashed line) parton distributions, appropriately normalised (see text).	199
7.17	A comparison of the final differential cross sections with the five different fits made by Alvero et al., the details of which are in the text. The data appear to favour fits B and D . It should be noted that the upper solid line is fit SG and is in some cases the only solid line visible.	201
7.18	The theoretical uncertainties associated with the Ingelman-Schlein model. The solid line shows the “nominal value”, ie. that which is implemented in the RAPGAP Monte Carlo samples, while the dotted lines indicate the effect of varying each parameter. In the final plot the upper dotted line shows the combined effect of $m_c = 1.1$ GeV and $\alpha_P(0) = 1.157$, while the lower dotted line shows the combined effect of $m_c = 1.4$ GeV and $\alpha_P(0) = 1.0850$	203
7.19	A comparison of the final differential cross sections with the predictions from the model of Buchmüller et al. (“BHM”) and from the RIDI Monte Carlo samples (“RIDI”). Further details of the models are given in the text.	204

List of Tables

3.1	Performance of the HERA Machine for the period 1995 — 1997	69
3.2	The dimensions of the CTD Superlayers.	73
4.1	A Summary of the good data available for analysis for 1995 —1997.	91
5.1	The parameters of the fit to each of the Monte Carlo Samples.	127
5.2	Comparison of the results obtained from correction methods I and II for data and Monte Carlo	134
6.1	A Summary of availability of the different TLT triggers in the three different years of data-taking.	148
6.2	A Summary of the cumulative selection efficiency from each Monte Carlo sample separately and combined. The errors given are statistical.	159
6.3	A Summary of fit parameters determined from the DIS and diffractive DIS ΔM and $M(D^0)$ distributions. The values of ΔM and $M(D^0)$ are those given by the Particle Data Group [18], while the resolutions are those estimated from the different Monte Carlo samples.	165
7.1	The cross sections in the different bins chosen for the differential distributions.	187
7.2	The mean values of Q^2 and W in the cross section bins from data and Monte Carlo before and after acceptance correction.	188
7.3	A summary of the information relating to the inclusive $D^{*\pm}$ cross sections.	197

7.4	The fit parameters from the Alvero et al. calculations.	200
-----	---	-----

Chapter 1

Introduction

HERA, the world's first electron-proton collider, has opened up a new regime for the study of proton structure, by allowing proton constituents with momentum fraction x as low as $\sim 10^{-5}$ to be resolved. In opening up the low x region, it has become possible to study hadronic interactions which were previously only understood in the context of phenomenological models (so-called “soft” processes), within the same experiment as those which can be explained using perturbative Quantum Chromodynamics (pQCD) — the “hard” processes.

This capability is particularly important in the area of diffractive physics. Diffractive events are distinguished by the presence of a large phase-space gap in which no particles are produced. Prior to the advent of HERA, such interactions were investigated at both fixed target experiments and pp colliders, where in the latter case they are governed predominantly by non-perturbative physics. However at HERA, diffractive interactions in which a highly virtual photon is exchanged can also be studied. The photon virtuality provides a hard scale and hence perturbative QCD is applicable.

One particularly important process within diffraction is the production of open charm, which provides an alternative hard scale to the boson virtuality, Q^2 , in the form of the large charm quark mass. Also, as we shall see, charm production is highly sensitive to the role of gluons in both diffractive and non-diffractive interactions. Owing to these facts, the production of open charm in diffraction provides a key discriminator between the different theoretical models available to describe this process, given the very different roles which gluons play in the different theoretical predictions.

This thesis presents an analysis of diffractive open charm production in deep inelastic scattering, performed on the combined 1995-97 ZEUS dataset, which corresponds to an integrated luminosity of 43.3 pb^{-1} . The charm is tagged by reconstructing $D^{*\pm}(2010)$ mesons which have decayed in the $D^{*+} \rightarrow D^0\pi_s^+ \rightarrow (K^-\pi^+)\pi_s^+$ (+c.c.) decay channel. This first chapter introduces deep inelastic scattering, the theoretical background and the role of open charm production in understanding proton structure. Chapter two then gives a review of diffraction, as well as a detailed summary of the different theoretical models currently available. The role of open charm production in diffractive physics is also discussed.

Chapter three describes HERA and the ZEUS detector, emphasising the components important for this analysis. Chapter four then describes how the data are selected online and reconstructed offline. Chapter five describes one particular aspect of the offline reconstruction, namely the understanding of energy loss from the scattered positron and how it may be corrected. Chapters six and seven then describe the selection of diffractive $D^{*\pm}$ candidates in DIS events and present measurements of the $D^{*\pm}$ cross sections and the fraction of DIS $D^{*\pm}$ events which are diffractively produced. Finally, chapter eight provides a brief summary and some thoughts for the future of this particular measurement using the ZEUS detector at HERA.

1.1 Deep Inelastic Scattering

Deep inelastic lepton nucleon scattering is experimentally investigated through the scattering of a high energy lepton off a hadronic target or beam. Deep implies that the exchanged boson is highly virtual, while inelastic indicates that the initial state hadron disintegrates into a hadronic system Y . At HERA such DIS processes are observed and can be divided into two basic classes according to the charge of the exchanged boson, as shown in figure 1.1.

DIS processes in which a $\gamma^{(*)}$ or Z^0 is exchanged are known as neutral current (NC) interactions, while those mediated by W^\pm exchange are referred to as charged current (CC) interactions. In the latter case, the incoming charged lepton changes into the corresponding neutrino in the final state. This thesis is concerned primarily with NC interactions mediated by a virtual photon. Interactions in which an almost real photon is exchanged

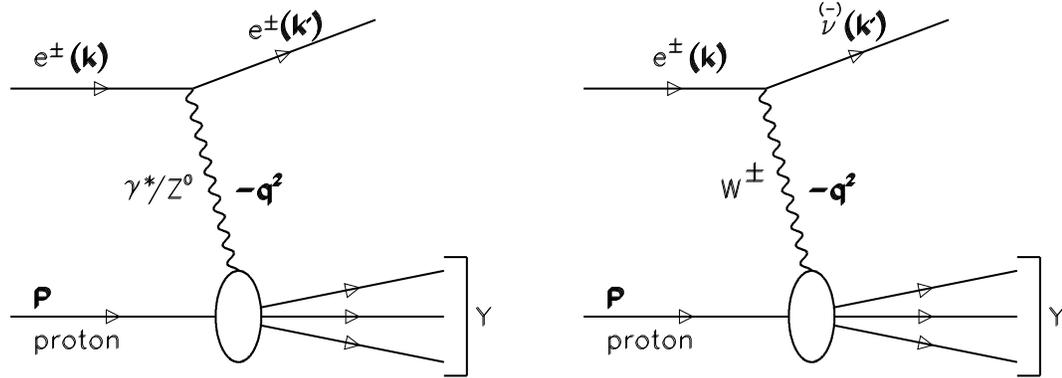


Figure 1.1: Diagrams for DIS via a neutral current interaction (left) and a charged current interaction (right).

are also observed. These are known as photoproduction processes.

It is customary to describe DIS processes in terms of the three Lorentz-invariant kinematic variables Q^2 , x and y . These are defined as

$$Q^2 = -\mathbf{q}^2 = -(\mathbf{k} - \mathbf{k}')^2 > 0$$

$$x = \frac{Q^2}{2\mathbf{P} \cdot \mathbf{q}} \quad 0 < x < 1$$

$$y = \frac{\mathbf{P} \cdot \mathbf{q}}{\mathbf{P} \cdot \mathbf{k}} \quad 0 < y < 1 \quad (1.1)$$

Neglecting the particle masses, Q^2 , x and y are related by $Q^2 = xys$, where $s = (\mathbf{P} + \mathbf{k})^2$ is the centre of mass energy squared.

Q^2 is the four-momentum transfer squared between lepton and proton and ranges from 0 to s . It represents the virtuality of the exchanged boson, and in photoproduction is close to zero, while in DIS, Q^2 is significantly larger than zero. When the photon is real it can only be transversely polarised, while virtual photons can be both transversely and longitudinally polarised. For $Q^2 \ll M_{Z^0}^2$, neutral current interactions are dominated by virtual photon exchange. The virtuality of the exchanged boson is inversely proportional to the transverse distance which can be probed within the proton, indicating that as Q^2 increases, finer and finer structure can be discerned.

y is the fractional energy loss of the lepton in the proton rest frame, while x is the Bjorken scaling variable [1, 2]. In the Quark Parton Model (see section 1.2.1) x can be interpreted as the fraction of the proton's momentum carried by the struck quark. Both x and y lie in the range 0 to 1.

The invariant mass of the hadronic final state (indicated by Y in figure 1.1) is known as W , which is simply the photon-proton centre of mass energy. It is related to the other kinematic variables via

$$W^2 = (\mathbf{P} + \mathbf{q})^2 \approx \frac{Q^2(1-x)}{x} \stackrel{x \ll 1}{\approx} \frac{Q^2}{x} \quad (1.2)$$

which arises from conservation of momentum at the hadronic vertex. Once again the particle masses have been neglected.

1.2 Proton Structure Functions

The neutral current ep interaction shown in figure 1.1 can be viewed in terms of a lepton-photon interaction followed by a photon-proton interaction. The former can be calculated precisely in Quantum Electrodynamics (QED). However, the latter is not so well understood, due to our lack of exact knowledge of proton structure. It can be described through the proton structure functions, F_1 , F_2 and F_3 . The differential $e^\pm p$ cross sections in x and Q^2 can be written in terms of these structure functions as

$$\frac{d^2\sigma^{NC}(e^\pm p)}{dx dQ^2} = \frac{4\pi\alpha^2}{xQ^4} \left[\frac{y^2}{2} 2xF_1(x, Q^2) + (1-y)F_2(x, Q^2) \mp (y - \frac{y^2}{2})xF_3(x, Q^2) \right] \quad (1.3)$$

xF_3 only becomes significant at high Q^2 ($Q^2 \sim M_Z^2$) when Z^0 exchange becomes important. Thus, in the context of this thesis it may safely be neglected as only the low Q^2 region ($Q^2 < 400 \text{ GeV}^2$) is considered here. Taking this into account, and using the relation $F_L = F_2 - 2xF_1$, where F_L is the longitudinal structure function, we can therefore write

$$\frac{d^2\sigma^{NC}(e^\pm p)}{dx dQ^2} = \frac{2\pi\alpha^2}{xQ^4} [(1 + (1-y)^2)F_2(x, Q^2) - y^2 F_L] \quad (1.4)$$

Fixed target experiments performed at SLAC in the late 1960s indicated that the proton had a partonic structure [3] and also found that the structure functions were more or less independent of Q^2 but were dependent on x [1], as had already been predicted by Bjorken [2] (the so-called ‘‘Bjorken scaling’’).

Later the quarks and gluons of QCD were associated with the partons which had been previously used to explain DIS processes [4], indicating lepton-quark scattering as the way to understand ep scattering.

1.2.1 The Quark Parton Model

The quark parton model (QPM) was developed by Feynman [4] in order to explain the scaling predicted by Bjorken and experimentally measured at SLAC. The QPM treats the proton as a group of point-like, non-interacting, spin-1/2 constituents (partons). The ep cross section can therefore be written as the incoherent sum of the elastic lepton-parton interactions (see figure 1.2). Scaling is clearly a property of such a model as a point-like particle will look the same regardless of the scale at which the proton is probed.

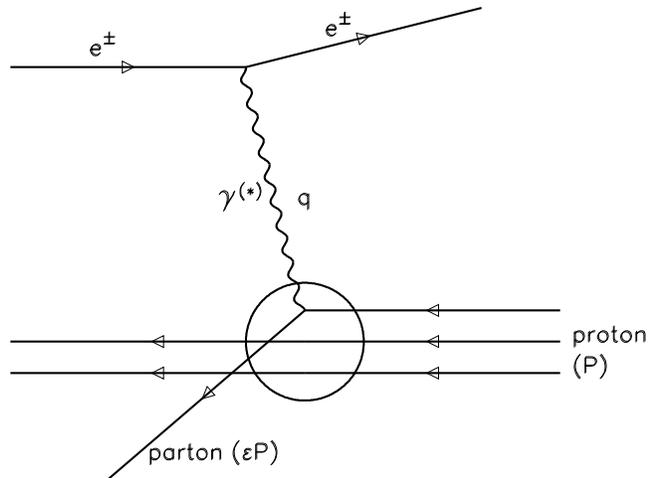


Figure 1.2: Lepton-parton scattering in the Quark Parton Model, where the struck parton carries away a fraction ϵ of the proton's momentum.

In the infinite proton momentum frame (the frame in which the photon four-momentum is entirely timelike), all parton transverse momenta are negligible and the partons appear to move and interact very slowly compared to the time taken for the lepton-proton interaction. Thus, they can be considered as non-interacting free particles. The Bjorken scaling variable, x , takes on a very simple physical interpretation within the context of this model, as can be seen by considering momentum conservation in the photon-parton

interaction, with the parton carrying a fraction ϵ of the proton's momentum:

$$(\epsilon\mathbf{P} + \mathbf{q})^2 = \epsilon^2\mathbf{P}^2 + \mathbf{q}^2 + 2\epsilon\mathbf{P} \cdot \mathbf{q} = m_i^2 \quad (1.5)$$

where m_i is the mass of parton i in the proton. Neglecting the parton and proton masses ($\epsilon^2\mathbf{P}^2 \approx \epsilon^2M^2 = 0$) equation 1.5 reduces to

$$\mathbf{q}^2 + 2\epsilon\mathbf{P} \cdot \mathbf{q} = 0 \Rightarrow \epsilon = \frac{Q^2}{2\mathbf{P} \cdot \mathbf{q}} \equiv x \quad (1.6)$$

Thus in the QPM x can be interpreted as the fraction of the proton's momentum carried by the struck parton.

Within the QPM, the structure functions can be shown to be [4]

$$\begin{aligned} F_2(x) &= \sum_i e_i^2 x f_i(x) \\ F_1(x) &= \frac{1}{2} \sum_i e_i^2 f_i(x) \end{aligned} \quad (1.7)$$

where e_i are the parton charges and $f_i(x)$ are the parton density functions which can be interpreted simply as the probability of finding parton i with a momentum fraction between x and $x + dx$ in the proton, subject to the constraint

$$\sum_i \int x f_i(x) dx = 1 \quad (1.8)$$

which is known as the momentum sum rule. F_2 and F_1 are then connected by the Callan-Gross relation which states that

$$2xF_1(x) = F_2(x) \quad (1.9)$$

and is a direct consequence of the assumption that partons are massless, spin-1/2, non-interacting particles and implies that $F_L = 0$ (see equation 1.4).

Through measurements made at SLAC and in νN [3] scattering, these partons were associated with the three valence quarks of the proton, which at that time were believed to be the only constituents of the proton, giving this model its name.

1.2.2 Quantum Chromodynamics

When the momentum sum rule (equation 1.8) was first measured experimentally [3], it was found that the momentum of the quarks could only account for about 50 % of the

proton's momentum. The remaining 50 % was attributed to neutral partons which have since been associated with the gluons of Quantum Chromodynamics (QCD). Gluons are the gauge bosons of the strong interaction which acts on particles that carry colour charge, ie. quarks. There are eight gluons due to the fact that they themselves are coloured. This means that the gluons can couple to each other — one of the major differences between QCD and QED.

The coupling constant of QCD, α_s , depends on Q^2 and is found to be large at small Q^2 (large distances) and small at large Q^2 (small distances). At leading order, α_s is inversely proportional to $\ln Q^2$, indicating that as Q^2 becomes large α_s vanishes logarithmically. This phenomenon is known as asymptotic freedom.

The consequences of the “running” of α_s are that when Q^2 is large, α_s is small enough to allow the use of perturbation theory. However, as α_s becomes large, perturbation theory breaks down and a more phenomenological approach to QCD must be taken. It is important to note that Q^2 is not the only hard scale available. Examples of the alternatives include large quark masses in heavy flavour production and high transverse momentum in jet production.

1.3 The Structure Functions in QCD

The quark parton model has to be significantly modified in order to accommodate the presence of gluons in the proton. It has already been shown that gluons cause the momentum sum rule to be broken, and the fact that the proton can no longer be considered to be dominated by massless, spin-1/2 partons results in the violation of the Callan-Gross relation, implying $F_L \neq 0$.

The presence of gluons is also the source of scaling violations (the structure functions' weak dependence on Q^2). This is because the struck quark could have radiated a gluon, which would cause it to lose momentum, and hence to be probed at a lower value of x . Alternatively, the quark could have originated from a gluon and hence have a lower value of x than the parent gluon. It is important to note that the presence of gluons in the proton can only be detected through processes such as these, as the exchanged photon only couples to quarks.

In short, the parton being probed could be the result of strong interactions within the proton rather than one of the “original” constituents (the so-called “valence” quarks). The higher the virtuality of the probe, the finer the substructure it can resolve; thus as Q^2 increases, so the number of quark-antiquark pairs and gluons in the proton “sea” will also increase. This leads to a softening of the valence quark x distribution with increasing Q^2 . In terms of F_2 , this indicates that at low x , F_2 will rise with Q^2 , while at high x , F_2 will fall with increasing Q^2 . All the expected effects of gluons in the proton have been observed at HERA (see figure 1.3 [5]).

In order to describe the evolution of the quark ($q^i(x, Q^2)$, where i indicates the quark flavour under consideration) and gluon ($g(x, Q^2)$) distributions with Q^2 , the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [6, 7] can be used. These equations contain terms which depend on the quark and gluon distributions and the Altarelli-Parisi splitting functions, $P_{jk}(\frac{x}{y})$, which are illustrated in figure 1.4. These functions represent the probability that parton k with momentum fraction y emitted parton j with momentum fraction x in the interval $Q^2 \rightarrow Q^2 + d \ln Q^2$. The splitting functions shown in figure 1.4 are the $\mathcal{O}(\alpha_s)$ contributions. The DGLAP equations are used to evolve the parton distributions with Q^2 from a starting scale Q_0^2 , where the initial form of the parton distributions must be specified.

The calculation of the amplitude for inelastic ep scattering is then performed by summing the contributions from ladder diagrams (see figure 1.5 for an example), where parton evolution is described as a series of gluon and/or quark emissions. These diagrams lead to terms in $\ln Q^2$ and $\ln(1/x)$ in the perturbation series.

In the DGLAP formalism the dominant terms which must be summed are those of the form $\alpha_s(Q^2) \ln Q^2$ (the so-called “leading logs”). This is known as the Leading Log Approximation (LLA). The $\ln(1/x)$ terms are neglected. In figure 1.5, the longitudinal momentum fractions will be ordered $x_1 > x_2 > \dots > x_n$ and the transverse momenta are strongly ordered ie. $k_{t_1}^2 \ll k_{t_2}^2 \dots \ll k_{t_n}^2$. The numbering scheme is indicated in the figure.

The case where $\ln Q^2$ terms dominate and $\ln(1/x)$ terms are neglected is valid in the large x , Q^2 region, that is, $\alpha_s(Q^2) \ln Q^2 \sim 1$ and $\alpha_s(Q^2) \ln(1/x) \ll 1$. However, when x becomes smaller and we enter the moderate Q^2 regime, it would seem reasonable to assume that terms in $\ln(1/x)$ should also become important. This means that the cross sections would be dominated by terms of the form $\alpha_s(Q^2) \ln Q^2 \ln(1/x)$, ie. $\alpha_s(Q^2) \ln Q^2 \ll 1$ and

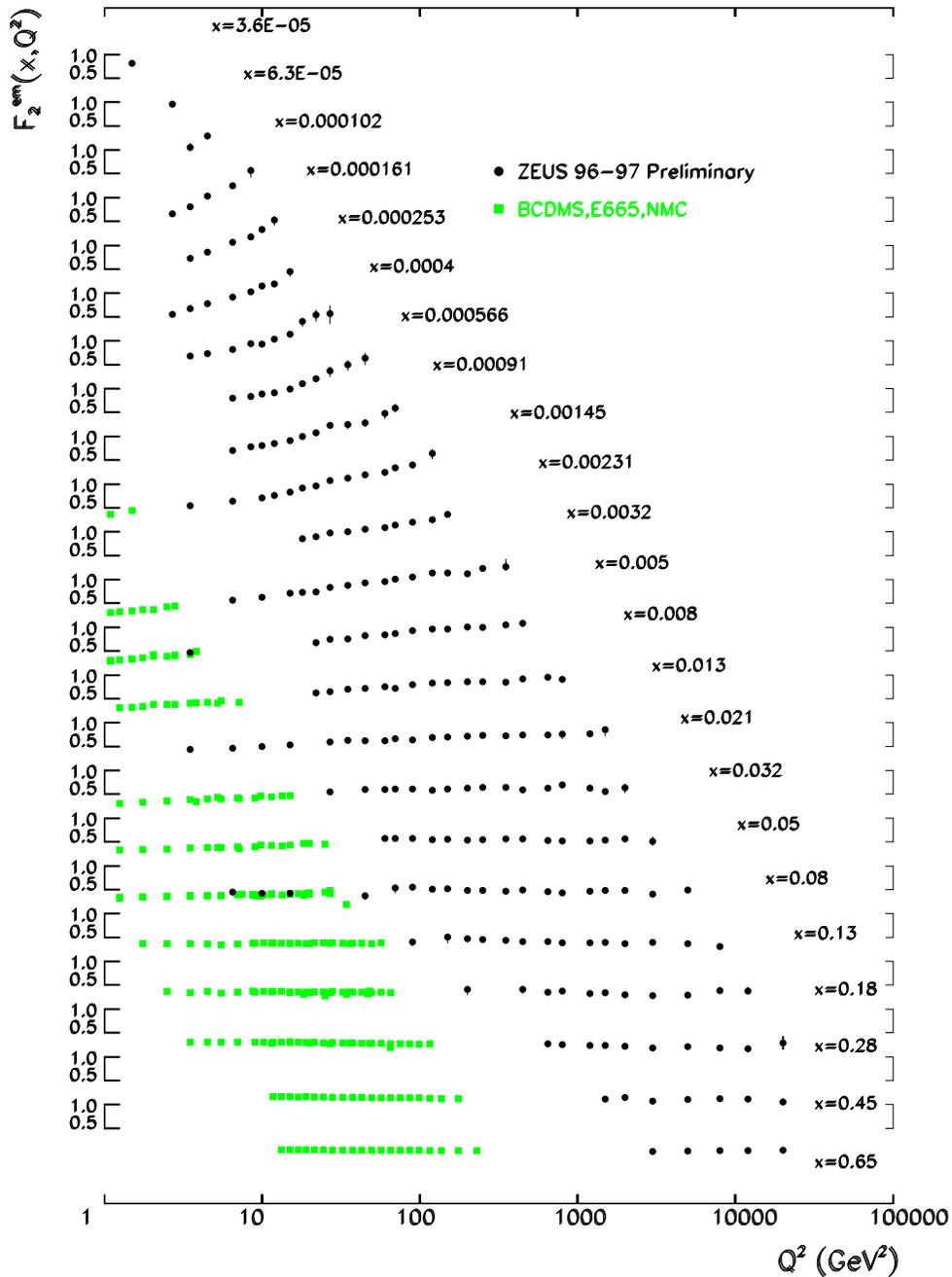


Figure 1.3: The structure function $F_2(x, Q^2)$ versus Q^2 at various fixed values of Bjorken x . Results from both HERA and fixed target experiments are shown. The evidence for scaling violations at low x are clear from the HERA data, while there is no significant Q^2 dependence in the data from the fixed target experiments at higher x , i.e. Bjorken scaling is observed.

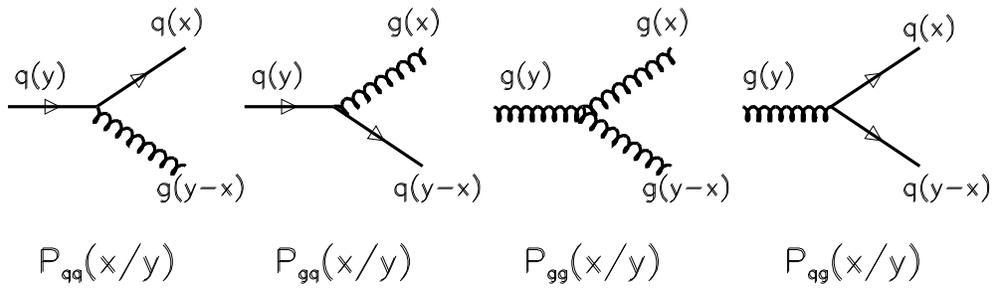


Figure 1.4: The Altarelli-Parisi splitting functions $P_{jk}(x/y)$ (for further explanation see text).

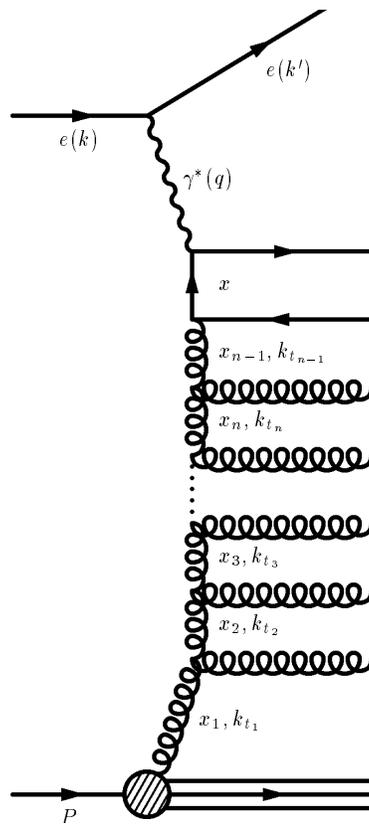


Figure 1.5: An example of a ladder diagram in QCD, showing parton evolution in the proton.

$\alpha_s(Q^2) \ln(1/x) \ll 1$, but $\alpha_s(Q^2) \ln Q^2 \ln(1/x) \sim 1$. In such a case the rungs of the gluon ladder in figure 1.5 would also become strongly ordered in x , as well as in k_t^2 .

At very low x , it has been suggested that the $\ln(1/x)$ terms should dominate and the $\ln Q^2$ terms should be neglected, that is, $\alpha_s(Q^2) \ln Q^2 \ll 1$ and $\alpha_s(Q^2) \ln(1/x) \sim 1$. This means that there is no longer a strong ordering in k_t^2 , but that the strong ordering in x remains. This approach has been used in the development of the Balitsky-Fadin-Kuraev-Lipatov (BFKL) evolution equation [9]. The gluon ladders of the BFKL formalism are sometimes referred to as the BFKL Pomeron (see the next chapter for further discussion).

So far, the DGLAP equations have been found to give good agreement over the full kinematic range in which F_2 has been measured.

1.4 Factorisation

In order to use QCD to describe DIS processes, it must be possible to “embed” the hard lepton-quark scattering (the hard scale being given by Q^2) into the observed ep interaction. Given that extended objects, such as hadrons, are composite particles which must be treated non-perturbatively, it is assumed that the interaction can be separated into the hard lepton-quark scatter and the non-perturbative hadronisation process. The former takes place over a very short timescale and can be calculated in perturbative QCD, while the latter occurs over a relatively long timescale. It is assumed that due to the different timescales, the hadronisation process cannot influence the hard scatter. This separation of the interaction according to timescale is known as factorisation [10].

The perturbative phase of the interaction is clearly independent of the incoming hadron beam type, which is not the case for the hadronisation process. The latter, however, does not depend on the actual interaction. This means that the proton parton distribution functions extracted at HERA should be equally applicable to proton-proton collisions at, for example, the Tevatron or the LHC.

1.5 Charm Production at HERA

There are several possible mechanisms which could account for charm production in DIS at HERA. The two most significant possibilities are Boson Gluon Fusion (BGF) and flavour

excitation, both of which are illustrated in figure 1.6. It is also possible for charm to be produced in beauty decays or during fragmentation, but the cross sections for these mechanisms are thought to be small [11].

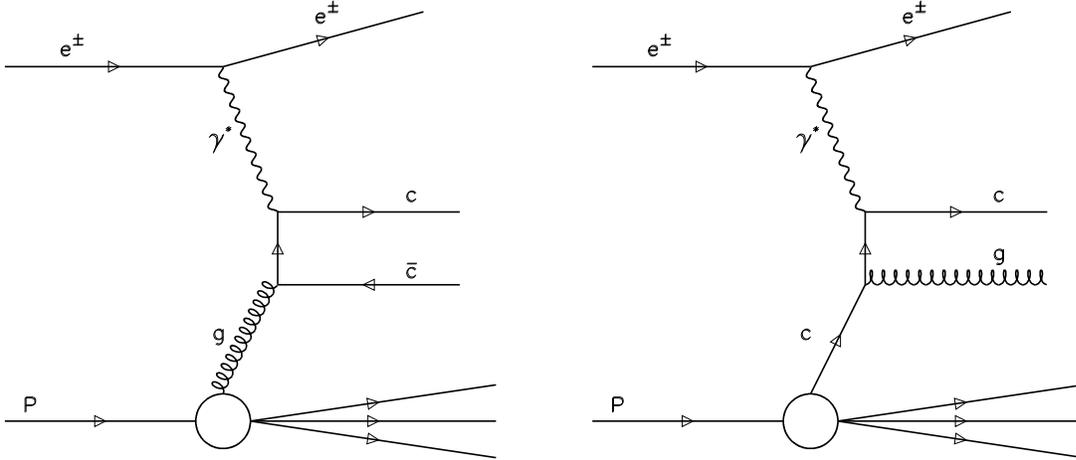


Figure 1.6: The leading order diagrams for boson gluon fusion (left) and flavour excitation (right).

It has also been proposed that there could be intrinsic charm in the proton, ie. a non-perturbative charm component [12]. However, this is only expected to become significant at much higher values of x than are currently reached by HERA charm data.

Measurements of charm production in DIS were first made by the European Muon Collaboration (EMC), from which a study of the production mechanism was made. The data were found to be well described by boson gluon fusion, but not so well by flavour excitation [13]. From this same data sample intrinsic charm was excluded down to the level of 0.6 % at 90 % confidence level in the kinematic region defined by $Q^2 > 1 \text{ GeV}^2$ and $60 < \nu < 220 \text{ GeV}$. $\nu = P \cdot q / M_p$ and is defined as the energy lost by the positron and transferred to the hadronic system in the proton rest frame. It is related to x and Q^2 via $x = Q^2 / (2M_p \nu)$.

The production mechanism has also been investigated by both the ZEUS and H1 collaborations using $D^{*\pm}(2010)$ and D^0 mesons [11, 14]. This is done by studying the fraction of the γ^*p centre of mass energy carried by the D meson using $x_D = \frac{2|p_D^*|}{W}$, where p_D^* is the momentum of the D meson in the γ^*p frame.

In the γ^*p frame, the $c\bar{c}$ pair produced in boson gluon fusion recoils against the proton

remnant, resulting in a falling x_D distribution which peaks at very low values. In contrast, only one charm quark recoils against the proton remnant in flavour excitation, leading to an x_D distribution peaked at larger values. Such a measurement can be extracted from D meson data as the meson generally carries a high fraction on the charm quark momentum [15].

The x_D distribution measured from H1 data [14] is shown in figure 1.7. The data, which are produced predominantly near the kinematic threshold ($W^2 \geq 4m_c^2$), clearly favour boson gluon fusion as the dominant mechanism. However, more recent measurements made using ZEUS data [16] are not as well described by BGF alone, but are found to have a softer x_D distribution, the opposite to the expected effect of a flavour excitation contribution. The explanation for this effect is not clear, but could be a result of the hadronisation model used. However, the agreement is sufficiently good for the BGF process to be considered as a reasonable approximation.

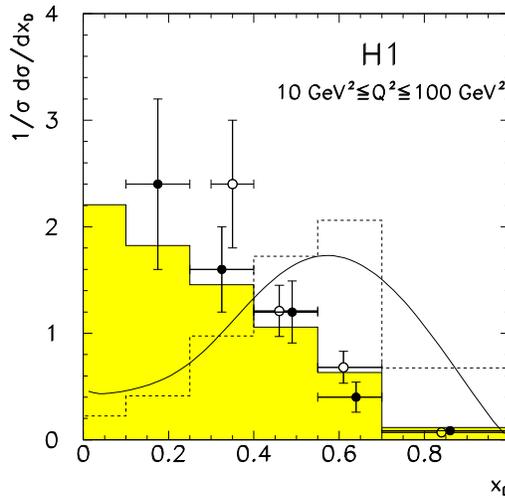


Figure 1.7: The x_D distribution extracted from 1994 H1 data. The black dots are from $D^{*\pm}$ data, while the open dots are from D^0 data. The shaded histogram indicates the expected x_D distribution from the BGF process, while the dotted line indicates that expected from flavour excitation. The solid curve represents measurements of charm production in νN scattering made by the CDHS and E531 experiments [14]. The H1 data clearly favour the BGF mechanism while the neutrino scattering results favour flavour excitation.

Given the clear evidence that charm production in DIS is driven principally by boson gluon fusion, the measurement of charm production via the tagging of D mesons should

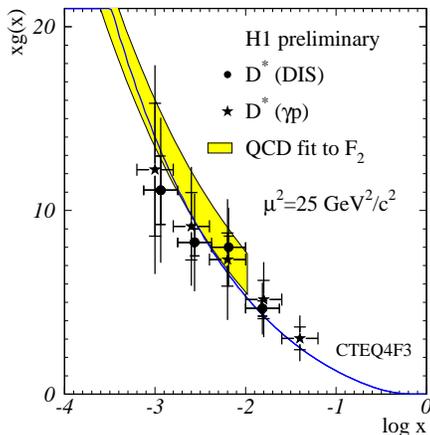


Figure 1.8: The gluon distribution extracted from H1 charm data as a function of x . Results from both DIS and photoproduction data are shown and compared to the gluon density extracted using QCD fits to measurements of $F_2(x, Q^2)$ (grey band). The prediction of the CTEQ parton distribution function parametrisation [23] is also shown (solid line).

provide a good experimental handle on the gluon density of the proton. Measurements of the charm contribution to the proton structure function, $F_2^{c\bar{c}}$, have already been made, as shown in figure 1.9 [16]. Both collaborations have used the available charm data to make an extraction of the gluon density using different techniques [17]. The gluon density extracted from H1 data is shown in figure 1.8.

There are, however, several issues which must be resolved in order to improve the accuracy of the theoretical treatment of heavy flavours. One such issue is the uncertainty on the charm mass which is currently restricted to the range 1.1 — 1.4 GeV [18]. Previously, the uncertainty on the charm mass was 1.2 — 1.6 GeV, as indicated in figure 1.9. Another key issue which must be addressed is the treatment of mass in the evolution equations, which is the subject of the next section.

1.5.1 Charm Evolution

There are two basic methods of treating charm in the evolution equations based on the factorisation equation [23]

$$\sigma_{p \rightarrow X}(s, Q^2) = \sum_i q^i(x, \mu^2) \otimes \hat{\sigma}_{i \rightarrow X}(\hat{s}, Q^2, \mu^2) \quad (1.10)$$

where i is the sum over all the different flavours which can actively participate in the interaction at the energy scale Q^2 . $\hat{\sigma}$ is the cross section for the hard scatter which is

convoluted with the parton distribution functions $q^i(x, \mu^2)$, where μ is the factorisation scale. Charm can either be included in the sum because it can be considered as an active participant if the energy scale is high enough or it can be excluded from the initial state and dealt with separately.

The former is known as a variable flavour number (VFN) scheme and has been used in several different global analyses (see section 1.7) with the simplification that above the threshold for charm production $Q^2 = \mu^2$ ($\mu_c = \mathcal{O}(m_c)$), $q_p^c(x, \mu^2)$ can be generated through $g \rightarrow c\bar{c}$ in the standard massless NLO DGLAP equations. This version is known as a zero mass VFN scheme and was developed because the massive coefficient functions required to account for the non-zero charm mass were not calculated until recently [19]. $q_p^c(x, \mu^2)$ is set to zero below threshold. At leading order the charm contribution to F_2 in this scheme is

$$F_2^{c\bar{c}} = \frac{8}{9} x q_p^c(x, Q^2) \quad (1.11)$$

The zero mass VFN scheme works well away from the threshold region ie. for $Q^2 \gg m_c^2$. However, it is clear that charm can be produced via boson gluon fusion for $W^2 \geq 4m_c^2$, which can be well below the Q^2 threshold at low x .

Thus an alternative approach to charm evolution was developed by treating it as a heavy quark and not as an active initial state parton, thus doing away with the idea of the charmed parton distribution. This approach is known as a fixed flavour number (FFN) scheme. In the case where the number of flavours is set to $n = 3$, charm is produced exclusively via boson gluon fusion and the leading order charm contribution to F_2 is given by

$$F_2^{c\bar{c}} = \int dz C_g(z, Q^2, \mu^2) \frac{x}{z} g\left(\frac{x}{z}, \mu^2\right) \quad (1.12)$$

where the effect of the non-zero charm mass is taken into account through the use of the massive coefficient function C_g . This works well in the threshold region, but is expected to break down in the asymptotic region due to the large log contributions of the form $\alpha_s(\mu^2) \ln\left(\frac{Q^2}{m_c^2}\right)^n$ [19]. In the zero mass VFN scheme these are the logs which are resummed by the DGLAP equations, indicating the potential validity of treating charm as a massless parton in this region.

Clearly a consistent treatment of charm production from threshold through to the asymptotic region is required in the evolution schemes. These schemes use the principals of the

FFN scheme near threshold and switch to a VFN scheme in the asymptotic region. Several different implementations of such generalised VFN schemes exist, which differ in the way the contributions from the two different schemes are matched in the Q^2 threshold region.

Aivazis, Collins, Olness and Tung (ACOT) [20] use the recently calculated massive coefficient functions in the standard NLO DGLAP equations, while Martin, Roberts, Ryskin and Stirling (MRRS) [21] have calculated massive splitting functions instead. Further studies have also been made by Roberts and Thorne [22]. The different implementations have been found to give results which lie within 5 % of each other.

As an alternative approach, Lai and Tung [23] have tried a FFN scheme where $n = 4$, based on the premise that charm is a “borderline” case for heavy quark treatment. The results of this approach are very similar to those achieved by the VFN schemes.

1.5.2 Charm Hadronisation

Charm production at HERA can be studied through the production of the J/ψ meson (or other $c\bar{c}$ bound states) or through open charmed meson production. It is the investigation of the latter which is of primary interest in this thesis, in particular the production of $D^{*\pm}$ mesons which have decayed through the

$$D^{*+} \rightarrow (D^0 \rightarrow K^- \pi^+) \pi_s^+ + \text{c.c.} \quad (1.13)$$

channel. This channel is useful as the $D^* \rightarrow D^0$ decay is near threshold, producing a low energy pion (the so-called “slow” pion, indicated by the subscript). The kinematic constraints on this stage of the decay allow the mass difference $\Delta M = M(D^*) - M(D^0)$ to be measured more accurately than the masses of the individual D mesons. This means that the measurement of ΔM yields a prominent signal in a region of phase space where the combinatorial background is heavily suppressed [24].

However, in order to make such an investigation, it is necessary to know how charm quarks fragment into D mesons. There are several models available to describe the fragmentation process [25], the most popular of which was proposed by Peterson et al. [26]. In this model, quantum-mechanical methods are used to determine the amplitude for the fragmentation of a heavy quark Q into a hadron $H = Q\bar{q}$ and a light quark q . This is found to be inversely proportional to the energy transferred during the fragmentation process, leading

to a fragmentation function of the form

$$D_Q^H(z) = \frac{N}{z[1 - (1/z) - \varepsilon_Q/(1-z)]^2} \quad (1.14)$$

where $z = \frac{x}{y}$ (as defined in the DGLAP equations), $\varepsilon_Q \sim m_q^2/m_Q^2$ and N is the normalisation which is fixed by

$$\sum \int dz D_Q^H(z) = 1 \quad (1.15)$$

where the sum runs over all hadrons containing Q . Fits performed by the OPAL Collaboration on $D^{*\pm}$ data yield $\varepsilon_c = 0.035 \pm 0.007 \pm 0.006$ [25]. More recently they have determined the probability of a charm quark fragmenting to a $D^{*\pm}$, P_c , to be $0.222 \pm 0.014(stat) \pm 0.014(syst) \pm 0.004$ [27]. The final error is that due to externally measured branching ratios. This is in good agreement with the measurement of P_c from lower energy data [25].

1.6 Parametrisations of the Parton Distribution Functions

In section 1.2 the parton distribution functions (pdfs) and their role in the proton structure functions was described. Given that it is not possible at present to determine the pdfs theoretically, it is important to use structure function measurements to obtain parametrisations of the pdfs. Several different groups have performed global analyses of structure function data, the most widely used of which are described here. The procedure for obtaining the parametrisations requires the assumption of a particular form for the pdfs at some starting scale Q_0^2 . These starting values are then evolved using the NLO DGLAP equations. The parameters involved in the pdfs are extracted from fits to data from a wide range of experiments, such as the fixed target electron data from SLAC and the muon data from EMC, BCDMS and NMC. Neutrino data from experiments such as CCFR and the latest HERA results are also used. The predictions for F_2 and $F_2^{c\bar{c}}$ from the three different global analyses discussed in the following sections are illustrated in figure 1.10.

Martin, Roberts and Stirling (MRS)

MRS parametrize the pdfs, $f_i(X, Q_0^2)$ as [8, 28]

$$xf_i(x, Q_0^2) = A_i x^{-\lambda_i} (1-x)^{\eta_i} (1 + \varepsilon_i \sqrt{x} + \gamma_i x) \quad (1.16)$$

where A_i , λ_i , η_i and γ_i are free parameters in the fit to the data (although limits are set on some of the A parameters through momentum sum rules etc.). They originally chose $Q_0^2 = 4 \text{ GeV}^2$, but more recently moved to $Q_0^2 = 1 \text{ GeV}^2$ [28]. They previously treated charm as a massless parton with the threshold at $\mu_c^2 = 2.7 \text{ GeV}^2$ (as obtained from EMC F_2^c measurements). However, in their latest pdf parametrisations [29] they have implemented the generalised variable flavour number scheme proposed by Roberts and Thorne [22, 21].

CTEQ Collaboration

The pdf parametrisations determined by the CTEQ Collaboration are very similar to those used by MRS [8, 19], although they originally used $Q_0^2 = 2.56 \text{ GeV}^2$. The other main difference is the form of the gluon density at Q_0^2 , which is given by

$$xg(x, Q_0^2) = A_0 x^{A_1} (1-x)^{A_2} (1 + A_3 x^{A_4}) \quad (1.17)$$

They also treated charm as a massless parton. However, more recently [23] they have lowered Q_0^2 to 1 GeV^2 and have implemented the ACOT variable flavour number scheme [20] to describe the charm contribution.

Glück, Reya and Vogt (GRV)

GRV originally assumed that at a very low starting scale ($Q_0^2 = 0.34 \text{ GeV}^2$) the proton consisted of only valence quark distributions, the evolution of which would generate the sea quark and gluon distributions [30]. However, this did not agree well with data, so they introduced valence-like gluon and sea quark densities of the form

$$\begin{aligned} xg(x, Q_0^2) &= Ax^\alpha(1-x)^\beta \\ x\bar{q}(x, Q_0^2) &= A'x^{\alpha'}(1-x)^{\beta'} \end{aligned} \quad (1.18)$$

where α , $\alpha' > 0$ and are (along with β , β' , A and A') free parameters of the fit. However, in contrast to MRS and CTEQ, they treat charm as a heavy quark in their parametrisation [31].

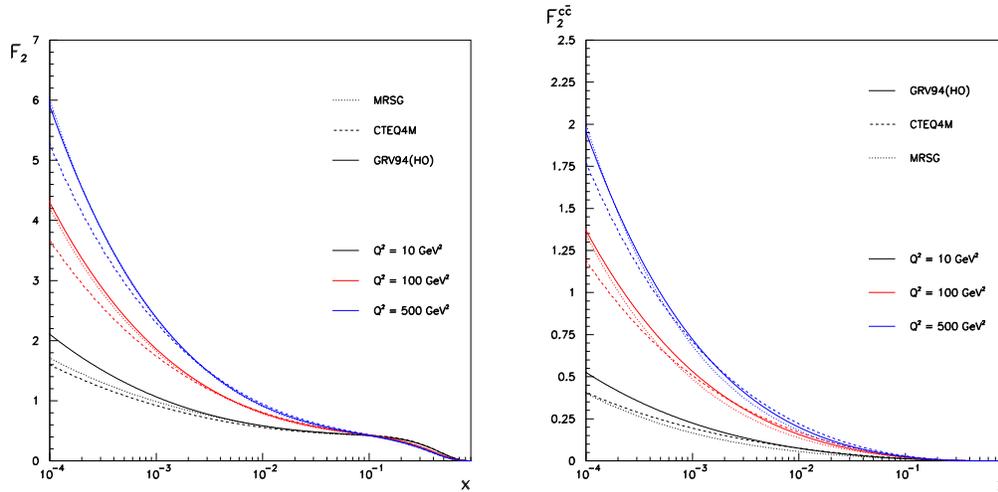


Figure 1.10: The predictions for F_2 and $F_2^{c\bar{c}}$ from three different parametrisations (see text) as a function of x for three different values of Q^2 .

1.7 Summary

In this chapter the concepts and variables of deep inelastic lepton nucleon scattering have been introduced. The development of the theoretical treatment of proton substructure has also been discussed, including the parametrisations of the parton distributions within the proton and the evolution equations needed to describe the experimentally-observed Q^2 dependence. The theoretical description of charm production in DIS has also been discussed given its important role in the measurement of the gluon distribution of the proton and the understanding of the flavour decomposition of F_2 .

The contents of this chapter provides important background to the discussion of diffractive interactions in the next chapter.

Chapter 2

Diffraction and Open Charm Production

When the early deep inelastic scattering (DIS) data taken at HERA were first analysed, it was found that approximately 10 % of the events in the sample contained a large rapidity gap ie. a region at large rapidity¹ within the detector in which no activity was observed.

In an ordinary DIS event, such rapidity gaps are suppressed by an exponential factor which depends on the width of the rapidity gap, as the colour field between the struck quark and the proton remnant after the hard scatter will cause particles to be produced in the gap during hadronisation. However, it is proposed that a colourless exchange occurs in rapidity gap events, preventing the formation of a colour field between the hadronic final state from the hard scatter and the proton remnant. Hence no particles will be produced to fill the gap during hadronisation.

The mediator of this colourless exchange is assumed to have the quantum numbers of the vacuum and is generically referred to as the Pomeron (\mathbb{P}). Similar large rapidity gap events are observed in hadron-hadron scattering and are known as *diffractive* events. The name derives from comparing the scattering of a beam of hadrons off a target with diffraction in optics. In the case of a broad beam of plane polarised light hitting a small

¹The rapidity of a particle with energy E and momentum p (p_z being the momentum component along the beam direction) is defined by $\mathcal{Y} = 1/2 \ln(\frac{E+p_z}{E-p_z})$. It can be reduced to pseudorapidity, $\eta = -\ln(\tan \theta/2)$ in the high energy limit where masses can be neglected. θ is the polar angle measured with respect to the proton beam direction.

piece of polaroid, parallels can be drawn between elastic, inelastic and diffractive scattering and the different effects of the polaroid on the light beam. The optical diffraction process which takes place effectively creates a new state with a different polarisation to that of the original beam. In the particle physics interpretation, it is said that a new state has been “diffracted into existence” [32].

The three different types of event in which a colourless exchange can occur are illustrated in figure 2.1. The single diffractive events are of principal interest here; at HERA they are known as photon dissociative events, as the photon dissociates into a hadronic system X while the proton remains intact.

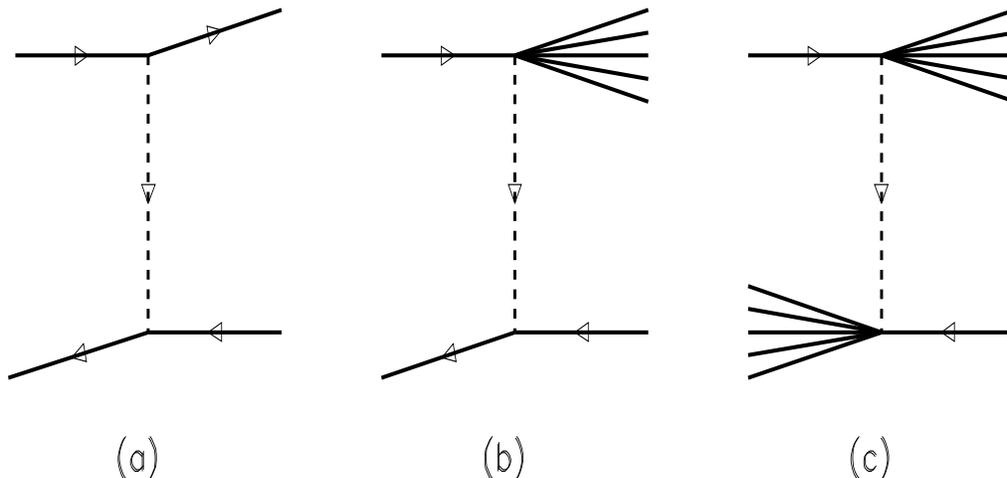


Figure 2.1: The three different types of interaction in which a colourless exchange can occur: (a) elastic scattering, (b) single diffractive and (c) double diffractive.

The idea of Pomeron exchange was first suggested in order to describe t -channel² hadron-hadron interactions as part of Regge theory. This will be discussed in the next section as it still plays a major role in the description of diffractive physics within the context of QCD. The different QCD models and their consequences for open charm production will also be discussed.

² t is the Mandelstam variable, given by $t = (P - P')^2$. In hadron-hadron scattering P and P' are the four-momenta of one of the hadrons before and after the interaction. In diffractive DIS they refer to the four-momenta of the incoming and outgoing proton.

2.1 Regge Theory and the Soft Pomeron

Regge theory was developed during the late 1960s as a precursor to QCD [33] in order to explain the soft t -channel exchange process in hadron-hadron collisions. Regge theory describes this process by colourless particle exchange (known as *poles*) which can be any of an infinite series of integer-spin, colourless particles ranging from the relatively light ρ vector meson through to the heavier a_2 and beyond. The heavier the particle the shorter the range of the force. This series of particles was found to lie in a straight line in the spin (α) — mass squared ($M^2 = t$) plane by Chew and Frautschi [34] as shown in figure 2.2. The line is referred to as a Regge trajectory.

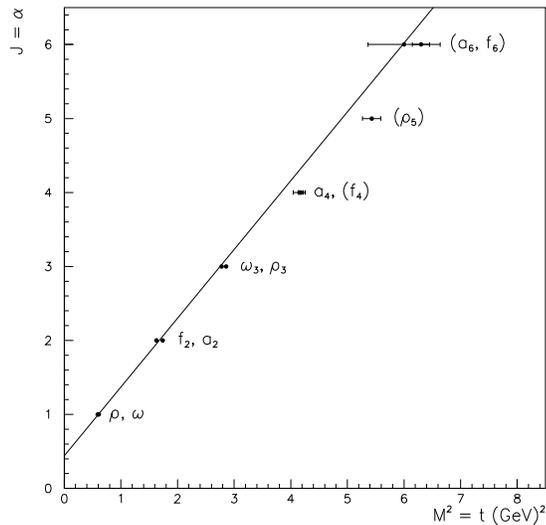


Figure 2.2: The spin versus t plot made by Chew and Frautschi. The line indicates the fit that they performed, which was found to be $\alpha = 0.55 + 0.86t$.

In order to calculate the amplitude for a colourless exchange, the contributions from all possible exchanges must be summed. Although the only physically possible states are those with integer spin, the entire trajectory is required to calculate the amplitude. This is done by treating spin as a continuous complex variable. Chew and Frautschi observed that the Regge trajectories can be described in just such a way by $\alpha(t) = \alpha_0 + \alpha't$. This can then be used to determine the scattering amplitude, A , which is a function of both s and t . For $s \rightarrow \infty$, A is given by

$$A(s, t) \sim \beta(t)s^{\alpha(t)} \quad (2.1)$$

where $\beta(t)$ accounts for the coupling of the trajectory to the external particles. The differential elastic cross section for $s \rightarrow \infty$ is therefore given by

$$\frac{d\sigma}{dt} \propto \left(\frac{s}{s_0}\right)^{2\alpha(t)-2} \quad (2.2)$$

where $s_0 \approx 1 \text{ GeV}^2$ is the hadronic mass scale. Using the optical theorem [32], the total cross section, σ_{TOT} , can be related to the forward ($t = 0$) elastic cross section from which it can be shown that

$$\sigma_{TOT} \propto \left(\frac{s}{s_0}\right)^{\alpha(0)-1} \quad (2.3)$$

However, the value of $\alpha(0)$ measured by Chew and Frautschi did not account for the experimental data, in which the total cross section was found to rise slowly with \sqrt{s} . It was suggested that a trajectory with $\alpha(0) \geq 1$ would describe the slow increase of the total cross section.

Theorems were then put forward by Pomeranchuk [35] which showed that if the exchange was charged, the total cross section would vanish asymptotically, while Foldy and Peierls [36] showed that if the total cross section did not fall with increasing s , the process must be dominated by the exchange of an object carrying the vacuum quantum numbers³. The trajectory which would fulfil these requirements was suggested by Gribov [37] and was named the Pomeron.

Donnachie and Landshoff took these ideas and fitted a wide range of total cross section data with a function of the form

$$\sigma_{TOT} = X s^{\alpha_P(0)-1} + Y s^{\alpha_R(0)-1} \quad (2.4)$$

The first term is the contribution from pomeron exchange and the second is the contribution from so-called ‘‘Reggeon’’ exchange — the trajectory originally observed by Chew and Frautschi. Some of the fits performed by Donnachie and Landshoff are shown in figure 2.3 and are summarised in [38]. They found the pomeron intercept to be 1.0808 in all the fits they made. The slope α'_P was extracted from $d\sigma/dt$ data [38] and found to have a value of 0.25.

Just as the Regge trajectory observed by Chew and Frautschi has physically observable poles, it would seem reasonable to imagine that the pomeron trajectory should have them

³Zero isospin and even under the charge conjugation operation.

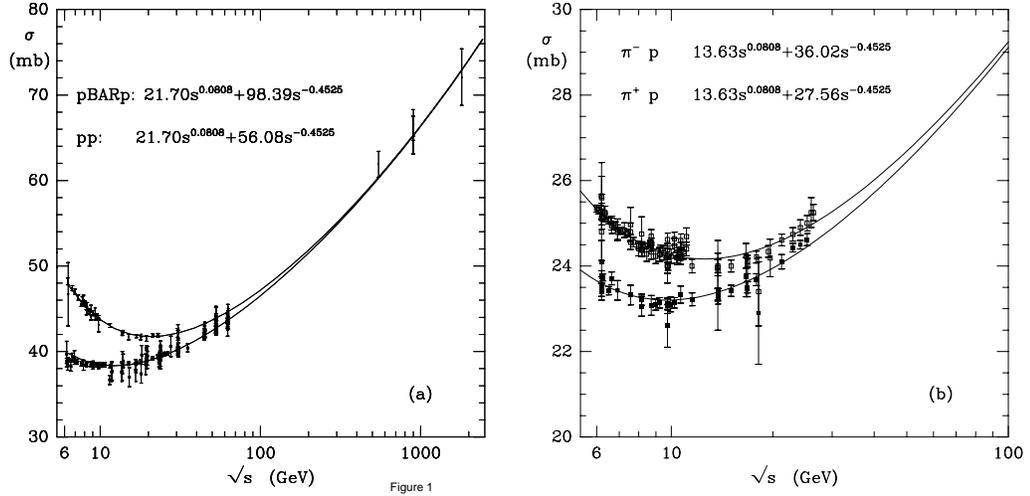


Figure 2.3: Fits performed by Donnachie and Landshoff to pp , $p\bar{p}$ and $\pi^\pm p$ data [38]. All the fits clearly indicate the pomeron intercept to be $\alpha_P(0) = 1.0808$.

too. A glueball candidate has been observed by the WA91 collaboration [39] with $\alpha = 2$ and a mass of 1926 ± 12 MeV, making it possibly the first observed pole of the pomeron trajectory.

After the advent of QCD, the pomeron of Regge theory became known as the “soft” pomeron due to its success in describing predominantly non-perturbative processes. However, as will become apparent in the next section, this soft pomeron is insufficient to describe interactions involving a hard scale, indicating that a QCD-inspired hard or perturbative pomeron is required.

2.2 Regge Theory at HERA

When viewed in the proton rest frame, deep inelastic scattering can be interpreted as a two stage, ie. factorisable, process in which the virtual photon fluctuates into a hadronic system which then scatters off the stationary proton [40]. This interpretation is an important component of many of the different theoretical models available to describe diffraction and will be discussed further in section 2.3. It also indicates that DIS can be considered as a hadron-hadron interaction to which Regge theory can be applied. Donnachie and Landshoff found that the $\sigma_{TOT}^{\gamma p}$ data from fixed target experiments and HERA exhibited the expected Regge behaviour, as shown in figure 2.4 [38].

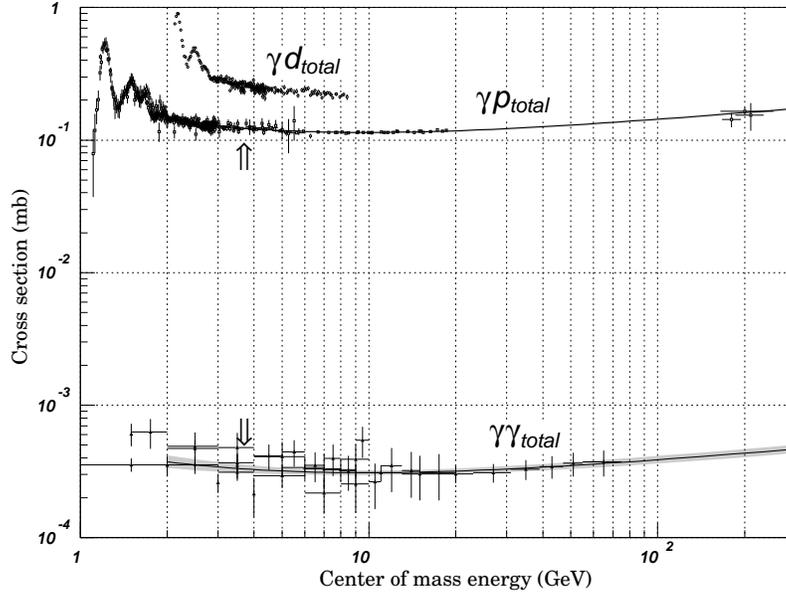


Figure 2.4: The Donnachie-Landshoff fit to γp data [38, 18]. The measurements of the total cross section at HERA are shown at $\sqrt{s} \simeq 200$ GeV, where \sqrt{s} is in this case equivalent to W . The fixed target measurements which were used in the original fits are also shown at low \sqrt{s} .

Given that Regge theory describes the s -dependence of $\sigma_{TOT}^{\gamma p}$, it was expected that it could also describe the W -dependence of $\sigma_{TOT}^{\gamma p}$ at HERA, through

$$\sigma_{TOT}^{\gamma p} \propto W^{2\alpha_P(0)-2} \quad (2.5)$$

This was found to be the case [41]. However, given that Regge theory is valid provided $s \gg |t|$ or equivalently $s \gg Q^2$ (which corresponds to the low x region), it was expected that Regge theory should also be able to reproduce the low x , Q^2 behaviour of $F_2(x, Q^2)$, given that

$$\frac{4\pi^2\alpha_{em}}{Q^2} F_2(x, Q^2) \approx \sigma_{TOT}^{\gamma^* p}(W^2, Q^2) \quad (2.6)$$

in the Regge limit. However, it was found that $\alpha_P(0) = 1.0808$ simply could not reproduce the steep rise of $F_2(x, Q^2)$ at low x . As was discussed in chapter 1, this rise can be accounted for within QCD by the presence of gluons in the proton. It was found, however, that increasing $\alpha_P(0)$ to about 1.5 would reproduce the rise, leading to the suggestion that this “hard” pomeron intercept could be linked specifically to the steep rise predicted by the BFKL equations. Consequently this “hard” pomeron is referred to by some as the BFKL pomeron [42].

Similar effects are observed in the W dependence of vector meson production cross sections. The photoproduction cross sections for light vector mesons (ρ , ω and ϕ) all exhibit a soft W dependence, that is, $\sigma \propto W^{0.16}$ from equation 2.5. However, as soon as a hard scale is introduced, such as Q^2 in the DIS production of ρ mesons [43] or the large charm mass in photoproduction of J/ψ mesons [44], a hard W dependence is observed, ie. $\sigma \propto W^{0.8}$, indicating that $\alpha_P(0)$ must be significantly larger than 1.0808.

It is important to stress that the pomeron is not necessarily a real particle and that its exact nature is not known. Clearly interactions which exhibit the characteristics of a colourless exchange and also involve a hard scale are important tools for discerning the nature of the pomeron within the framework of perturbative QCD.

The determination of the internal structure of the proton discussed in chapter 1 clearly establishes inclusive DIS as a probe of the quark parton distribution functions, the hard scale being provided by Q^2 . Equally, charm production provides a good experimental handle on the gluon density — here the hard scale comes from the large charm mass, as well as from Q^2 .

Given that diffractive interactions at HERA exhibit the characteristics of a colourless exchange, it seems reasonable to assume that they are caused by pomeron exchange and that measurements of diffractive DIS and diffractive charm production are therefore key processes through which the precise nature of the pomeron can be probed.

The models available to describe hard diffractive scattering and their consequences for open charm production are covered in the remaining sections of this chapter.

2.3 Models of hard diffractive scattering at HERA

Many different models are available to describe hard diffractive processes and they can be broadly divided into three categories. The first two treat the pomeron as an object which is either a non-perturbative system whose roots are in Regge theory or as a perturbative two-gluon exchange mechanism. The third does not assume the existence of the pomeron, but attempts to explain diffraction through soft interactions in the colour field of the proton.

Before these models are discussed in detail, the variables used to describe the diffractive

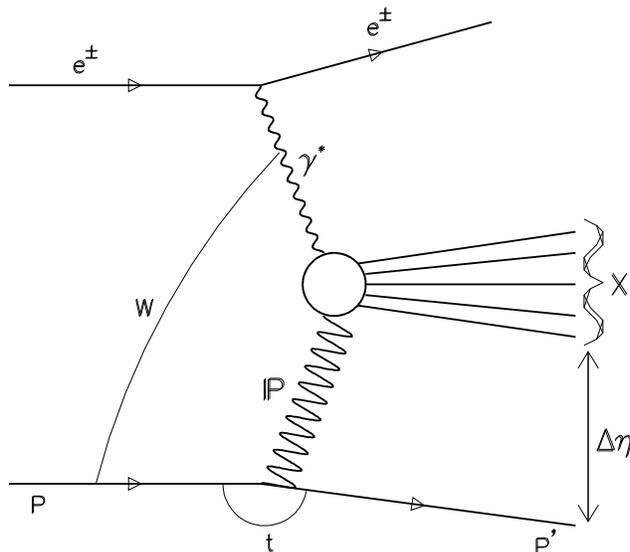


Figure 2.5: A generic diffractive interaction at HERA in which the proton remains intact and the photon dissociates into a hadronic system X .

process (in addition to x , Q^2 and W) must be defined. Figure 2.5 shows a generic photon dissociative diffractive interaction where the hadronic system into which the photon dissociates is referred to as X and has an invariant mass of M_X and the proton remains intact scattering through a small angle, implying that $t = (P - P')^2$ is small. $\Delta\eta$ is the size of the rapidity gap between the hadronic system X and the scattered proton.

Two variables are used to describe the pomeron itself: x_P , given by

$$x_P = \frac{(P - P') \cdot q}{P \cdot q} = \frac{M_X^2 + Q^2 - t}{W^2 + Q^2 - M_p^2} \quad (2.7)$$

which is the fraction of the proton's momentum carried by the pomeron, and β , given by

$$\beta = \frac{Q^2}{2(P - P') \cdot q} = \frac{Q^2}{M_X^2 + Q^2 - t} \quad (2.8)$$

which is the fraction of the pomeron's momentum carried by the struck parton in the pomeron. Given that t is assumed to be small and neglecting the proton mass, M_p , these variables reduce to

$$x_P = \frac{M_X^2 + Q^2}{W^2 + Q^2} \quad (2.9)$$

and

$$\beta = \frac{Q^2}{M_X^2 + Q^2}. \quad (2.10)$$

β and x_P are then simply related to x via $\beta = x/x_P$ in this approximation.

Clearly the definitions for x_P and β given here are inspired by the Regge theory approach to diffraction. These variables have different interpretations in the other available models, but have been implemented in such a way as to make them directly comparable to each other.

Using these variables, the diffractive DIS cross section can be written in the form

$$\frac{d\sigma^D}{dx_P dt dx dQ^2} = \frac{4\pi\alpha^2}{xQ^4} \left[(1 - y + \frac{y^2}{2}) F_2^D(x, Q^2, x_P, t) - \frac{y^2}{2} F_L^D(x, Q^2, x_P, t) \right] \quad (2.11)$$

in analogy with the DIS cross section given in the previous chapter (equation 1.4).

2.3.1 The Non-perturbative Factorisable Pomeron Model

This is a model of diffractive exchange, first proposed by Ingelman and Schlein [45], in which the pomeron is treated as an extended object with partonic structure. Studies of jet production in diffractive $p\bar{p}$ scattering made by the UA8 collaboration [46] indicated that the pomeron was not only partonic, but that the data favoured a hard partonic structure, that is, the parton distributions in the pomeron were of the form $6\beta(1 - \beta)$. Ingelman and Schlein also suggested (along with Donnachie and Landshoff [47]) that the diffractive structure function could be factorised into the emission of the pomeron and the hard scatter involving one of the partons in the pomeron, implying that the structure function, F_2^D , can be written as

$$F_2^D(x, Q^2, x_P, t) = f(x_P, t) F_2^P(\beta, Q^2, t) \quad (2.12)$$

where $f(x_P, t)$ can be interpreted as the probability for the proton to emit a pomeron at a particular value of x_P and t . This is clearly a non-perturbative process and must therefore be deduced from data. Given the assumption of factorisation, the form and parameters of $f(x_P, t)$ used successfully by Donnachie and Landshoff are assumed to be applicable here. The Donnachie-Landshoff form of $f(x_P, t)$ is given by [38]

$$f(x_P, t) = \frac{9b^2}{4\pi^2} [F_1(t)]^2 x_P^{1-2\alpha(t)} \quad (2.13)$$

where $b \approx 1.8 \text{ GeV}^{-1}$ is the pomeron coupling strength and $\alpha(t)$ is the pomeron trajectory discussed in section 2.1. $F_1(t)$ is the Dirac form factor of the proton which has been measured to high accuracy in low energy ep scattering [48]. Other parametrisations of $f(x_P, t)$ also exist, such as that by Streng [49], which is given by

$$f(x_P, t) = \frac{\beta^2(0)}{16\pi} x_P^{1-2\alpha_P(t)} e^{-b_0|t|} \quad (2.14)$$

where $\beta^2(0) = 58.74 \text{ GeV}^2$. Ingelman and Bruni also suggested a form for $f(x_P, t)$ which is discussed in [50].

$F_2^P(\beta, Q^2, t)$ in equation 2.12 is the pomeron structure function which can be treated much like $F_2(x, Q^2)$ of the proton, ie. one can assume initial forms for the parton distributions and evolve them using the DGLAP equations.

Ingelman and Schlein assumed that the pomeron consisted entirely of gluons and proposed two different possibilities: a hard gluon distribution

$$\beta g(\beta) = 6\beta(1 - \beta) \quad (2.15)$$

or a soft gluon distribution

$$\beta g(\beta) = 6(1 - \beta)^5 \quad (2.16)$$

The latter is similar in form to the proton gluon density, but the shape of the former was favoured by the UA8 data [46]. Neither distribution could reproduce the normalisation of the UA8 results.

Gehrmann and Stirling [51] developed the Ingelman-Schlein model into two distinct possibilities. In the first one, they assume the pomeron consists of hard gluon and quark distributions, but with the requirement that the momentum of the pomeron is carried predominantly by gluons, ie. $f_g \gg f_q$, where f_g, f_q are the momentum fractions of the gluons and quarks respectively.

In the second model, originally proposed by Kniehl, Kohrs and Kramer [52] to describe diffractive jet production, a direct pointlike coupling to quarks is allowed in addition to the resolved component described in the first model. This leads to a significant enhancement in the cross section at high β compared to the first model. Gehrmann and Stirling introduced this second model in an attempt to account for the fact that the Ingelman-Schlein model did not reproduce the normalisation of the UA8 jet data. The only alternative solution

was to introduce a process-dependent normalisation factor, which in turn would imply that factorisation did not hold.

Early ZEUS measurements of the diffractive structure function and diffractive photoproduction of jets [53] were interpreted within the context of the Ingelman-Schlein model with a pomeron consisting only of hard quark and gluon distributions. It was found that the fraction of the pomeron's momentum carried by hard partons, \sum_P , lay in the range $0.4 < \sum_P < 1.6$ and that the fraction of these hard partons which were gluons, f_g , lay in the range $0.3 < f_g < 0.8$. The relatively wide ranges measured reflect the statistical and systematic errors from both sets of measurements, as well as the theoretical uncertainties involved. These results indicate that the pomeron probed at ZEUS is made up predominantly of hard partons, a large fraction of which are gluons. More recent H1 results on the QCD fitting of $F_2^{D(3)}$ [54] also found that a substantial hard gluon contribution is required to reproduce their results.

Similar results were also found by Alvero, Collins, Terron and Whitmore (ACTW) [55] who performed fits to a range of results from both ZEUS and H1. They then attempted to predict cross sections for diffractive jet and weak boson production at the Tevatron. Their predictions substantially over-estimated the results produced by CDF, possibly indicating the breakdown of factorisation. Further evidence for this breakdown was presented by the CDF Collaboration, who extracted values of \sum_P and f_g using their results on diffractive jet production and diffractive W^\pm boson production [56, 57]. They measured $\sum_P = 0.18 \pm 0.04$ and f_g to be 0.7 ± 0.2 . While there is clearly good agreement in the value of f_g , the ZEUS measurement of \sum_P lies substantially above the $p\bar{p}$ results. If it is assumed that the momentum sum rule holds, then the fact that the CDF value of \sum_P deviates significantly from unity implies a different pomeron flux normalisation to that observed by ZEUS. If factorisation holds, such a mismatch should not exist. It has been proven theoretically that factorisation holds for diffractive deep inelastic scattering [58]. However, in hadron-hadron interactions, diagrams have been found where factorisation is broken, giving a possible explanation for the differences between HERA and Tevatron data.

Despite the apparent breaking of factorisation observed in $p\bar{p}$ interactions, this type of model has been found to describe a wide range of HERA results, including the early measurements of diffractive open charm production made by the ZEUS and H1 Collaborations [59, 60].

2.3.2 Perturbative QCD Pomeron Models

The simplest description of the diffractive exchange in perturbative QCD is by the exchange of two t -channel gluons in a colour singlet state, originally proposed by Low and Nussinov [61] (see figure 2.6). This leads to a diffractive structure function which is proportional to the square of the proton gluon density.

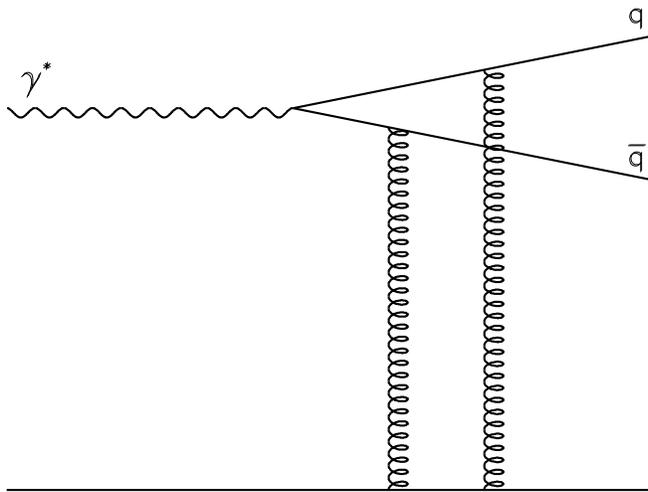


Figure 2.6: The simplest leading order interpretation of the perturbative two-gluon exchange model. There are three other diagrams which contribute at this order.

As already briefly mentioned in section 2.2, one of the key elements of these models (as well as those in the next section) is the fluctuation of the virtual photon into a hadronic system, which at leading order is a $q\bar{q}$ pair. In the proton rest frame the fluctuation occurs within a distance $l_c = 1/2M_p x$ of the proton. At low x , l_c can be as large as 1000 fm, which implies that DIS can be interpreted as the scattering of a $q\bar{q}$ pair off the proton target in this frame. The transverse size of this pair when it reaches the proton is $\rho^2 \approx 1/k_T^2$, where k_T is the relative transverse momentum of the $q\bar{q}$ pair.

At low x , the virtual photon tends to fluctuate into a configuration where its momentum is split asymmetrically between the two quarks, resulting in a large transverse separation. Confinement places an upper bound on this separation, referred to as the confinement radius, R_c . The original aligned jet model based on the Quark-Parton model [40] assumes that this is the only possible configuration for the $q\bar{q}$ pair, resulting in a transverse cross

section which varies as $1/Q^2$. This then gives a form for the proton structure function, F_2 in non-diffractive DIS which is independent of Q^2 . There is also a contribution to the interaction cross section at low x from the more symmetric configurations, which results in a small transverse separation. This contribution is small compared to that from the asymmetric configurations.

Nikolaev and Zakharov [62] view the diffractive interaction as the exchange of two non-interacting and apparently uncorrelated gluons which couple to the $q\bar{q}$ pair from the photon, as shown in figure 2.6. They predict that the interaction cross section is as described above, that is, dominated by asymmetric configurations, but with a small contribution from more symmetric configurations. This can be seen from the interaction cross section, which is given by

$$\sigma(x, \rho) = \frac{4\pi\alpha_s}{3} \int \frac{d^2\mathbf{k}}{\mathbf{k}^4} \mathcal{F}(x, \mathbf{k})(1 - e^{i\mathbf{k}\cdot\rho}) \quad (2.17)$$

where \mathbf{k} is the transverse momentum of the exchanged gluons and $\mathcal{F}(x, \mathbf{k})$ is the unintegrated gluon density. Performing the integration for $\rho^2 < R_c^2$, the interaction cross section becomes

$$\sigma(\rho) \propto \rho^2 \quad (2.18)$$

This behaviour, along with the small absolute size of $\sigma(\rho)$ is known as colour transparency [63], resulting from the fact that the quarks screen each other when ρ is small. One of the consequences of colour transparency is the suppression of initial and final state interactions.

Nikolaev and Zakharov also predict that for relatively large transverse separations

$$\sigma(\rho) \propto R_c^2 \quad (2.19)$$

indicating the saturation of the cross section at the confinement limit. They also predict that $\rho \sim 1/m_q$, where m_q is the mass of the quark produced in the photon fluctuation, implying that light flavours are produced through large size configurations, while heavy flavours are produced entirely through colour transparency. Nikolaev and Zakharov state that the presence of colour transparency leads to factorisation breaking.

The higher order process $\gamma^* \rightarrow q\bar{q}g$ was also considered by Nikolaev and Zakharov, but it was concluded that this process is essentially driven by the leading order $\gamma^* \rightarrow q\bar{q}$ process, resulting in similar cross section behaviour to that expected from leading order calculations.

Levin, Ryskin, Martin and Teubner (LMRT) [64] have developed a similar model within the context of charm production, based on the exchange of a gluon ladder. This was considered originally by Ryskin alone [65]. Despite the differences between LMRT and Nikolaev and Zakharov, LMRT derive very similar formulae for the transverse and longitudinal cross sections for the leading order $\gamma^* \rightarrow q\bar{q}$ process. Higher order corrections are then also considered, but in much more detail than by Nikolaev and Zakharov. First, real gluon emission corrections are evaluated, examples of which are shown in figure 2.7.

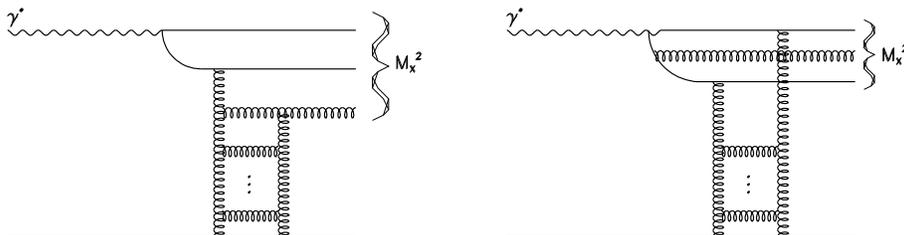


Figure 2.7: Two examples of higher order real gluon corrections. On the left, the case where the gluon is softer than either of the quarks is shown. On the right, one of the quarks is the softest parton.

Only real gluon emission configurations where the transverse momenta of the partons are strongly ordered are considered, as the necessary formulae only exist for such configurations. This is, however, expected to give a good estimate for the purpose of quantifying these corrections.

The configuration where the gluon transverse momentum is smaller than either of the quark transverse momenta gives a $1/M_X^2$ behaviour for $M_X^2 \gg Q^2, m^2$. When one of the quarks has the smallest transverse momentum the behaviour is found to be $1/M_X^4$. When combined with the leading order cross section, it is found that at low M_X^2 the $\gamma^* \rightarrow q\bar{q}$ process dominates, but that at larger M_X^2 , the $\gamma^* \rightarrow q\bar{q}g$ process noticeably enhances the cross section.

Virtual gluon corrections are also considered, examples of which are shown in figure 2.8. The full $\mathcal{O}(\alpha_s)$ calculations for such corrections do not exist for diffractive processes, but they have been investigated for the Drell-Yan process, $q\bar{q} \rightarrow \gamma^*$ [66]. In this case it was found that virtual loop corrections can enhance cross sections involving this process by a factor of at least two.

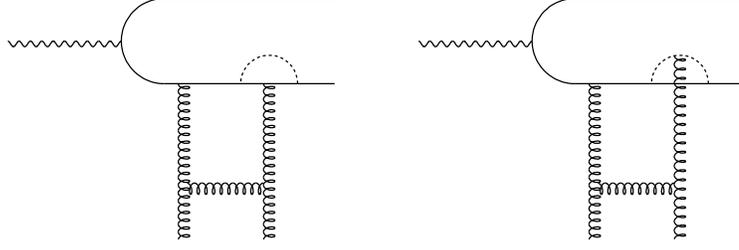


Figure 2.8: Two examples of higher order virtual gluon corrections, whose contribution to the diffractive cross section is estimated in the LMRT model. The dotted lines are the virtual gluon loops.

The corrections shown in figure 2.8 are those which provide an enhancement of the cross section of the form $\sigma^{q\bar{q}} = \sigma_0^{q\bar{q}} \exp(\alpha_s C_F \pi)$, where $\sigma_0^{q\bar{q}}$ is the leading order cross section and C_F is a colour factor. The other possible virtual loop corrections do not cause cross section enhancements. Similar effects are found for the higher order $\gamma^* \rightarrow q\bar{q}g$ process from virtual gluon emissions. The exponential enhancement factor is estimated to lie in the range 2.7 - 4.0 and is applicable over the full kinematic region. It is also found that performing the integration explicitly over the exchanged gluon transverse momentum enhances the cross section by a factor of two compared to using the leading log approximation. It should also be noted that factorisation is predicted to hold for all the higher order processes as well as at leading order and that the MRS(A') gluon density is used to make numerical predictions.

Ryskin and Solano [67, 65] have implemented a very similar model based on the same equations as were used by LMRT, but using the leading log approximation. GRV is used to describe the proton gluon density and a simpler version of the higher order corrections is implemented.

Similar calculations have been made by Bartels et al. [68], Bialas and Peschanski [69] and Gotsman, Levin and Maor [70], all based on the same equations as the Nikolaev and Zakharov model.

An attempt to account for the soft pomeron phenomena using QCD is made by Diehl [71], based on the Landshoff-Nachtmann model [72], in which the pomeron is treated as the exchange of two non-perturbative gluons. The calculations yield significantly different

forms for the cross sections compared to the perturbative gluon models, but result in very similar shapes for the spectra of the different observables. The principal difference observed is in the normalisation compared to the perturbative gluon models, which will typically be lower. This is due to the difference in the energy dependencies implemented in the perturbative and non-perturbative cases.

Comparisons of such models with experimental results from HERA indicate that both the Nikolaev-Zakharov model and the Ryskin and Solano Monte Carlo provide a reasonable description of the data. However, the first observations of diffractive charm production by both ZEUS and H1 are at too high a level to be reasonably described by the original Nikolaev and Zakharov model. A more recent implementation, however, provides a more reasonable charm contribution, as will be discussed in section 2.4.

2.3.3 Soft Colour Interaction Models

In the soft colour interaction (SCI) models diffractive interactions are not considered as being mediated by a colourless object, but instead as the consequences of the propagation of the $q\bar{q}$ pair in the colour field of the proton are investigated.

Edin, Ingelman and Rathsman [73] base their model on the boson gluon fusion (BGF) process which generates a substantial part of the total cross section at low x . Conventionally, the colour structure formed during the hard scatter passes directly to the hadronisation stage without any alteration. However, Edin et al. suggest that soft interactions may occur as the $q\bar{q}$ system produced in the hard scatter propagates through the colour field of the proton. Such interactions would not alter the four-momenta of the $q\bar{q}$ pair, but could change the colour topology prior to hadronisation. Thus the $q\bar{q}$ pair could emerge as a colour singlet state, causing rapidity gaps to occur.

Buchmüller and Hebecker developed a similar model [74] also based on the BGF process. It is assumed that the $q\bar{q}$ pair is produced in a colour octet state, which then evolves into a parton cluster as it propagates through the colour field of the proton, where its colour is changed randomly through soft interactions. The probability of producing a colour singlet state in this manner is approximately 1/9.

More recently, Buchmüller, Hebecker and McDermott [75] have developed a semi-classical approach to diffraction in which the proton (at low x) is treated as a classical colour field

localised within a radius $1/\Lambda_{QCD}$. In the proton rest frame, the $q\bar{q}$ pair produced by photon dissociation then scatters off this colour field. If the $q\bar{q}$ pair is in a colour singlet state after the scatter, then a large rapidity gap is produced. As with the two-gluon exchange models, it is found that the asymmetric configurations of the $q\bar{q}$ pair dominate and that heavy flavour production is suppressed at leading order.

However, Buchmüller et al. also consider the higher order process $\gamma^* \rightarrow q\bar{q}g$. It is found that only the configurations where both the quark and the antiquark have high transverse momenta and the gluon is soft can enhance the cross section. This is because the two high transverse momentum partons are considered to be sufficiently close together to be treated as one. Hence when the gluon is soft, an asymmetric configuration results regardless of the flavour of the $q\bar{q}$ pair. The case where either the quark or antiquark is soft will give similar results to the leading order $\gamma^* \rightarrow q\bar{q}$ process. It is important to note that a $q\bar{q}g$ state can only be produced at relatively large M_X implying (as was found in the LMRT model) that this higher order process only enhances the cross section in that region.

As with the two-gluon exchange models, the diffractive structure function is dependent on the proton gluon density. Factorisation is also predicted to hold for the same reasons cited in the LMRT model.

2.4 Charm Production in Diffractive Scattering

It is well known that non-diffractive charm production is regarded as a good experimental handle on the gluon density of the proton, as was discussed in section 2.2. Given this fact it is clear that not only should charm production in diffraction provide a good method of determining the gluon density of the pomeron, it should also provide a valuable probe of the nature of the diffractive exchange.

Clearly, if hard QCD dominates in diffractive processes, it would be expected that the fraction of diffractive interactions which contain charm, $R(\text{charm})$ should be given by

$$R(\text{charm}) \approx \frac{e_c^2}{\sum_q e_q^2} = 0.4 \quad (2.20)$$

where q indicates the sum over the quark flavours up to and including charm. Clearly, this value can only be achieved above the threshold for $c\bar{c}$ production, that is, $Q^2 > 4m_c^2$. In the factorisable pomeron models, charm is produced by boson gluon fusion involving a gluon

from the pomeron. Given the high gluon fraction already measured within the context of these models [53, 54, 56, 57, 46], a large fraction of diffractively produced charm is predicted. In contrast, the leading order perturbative two-gluon exchange models predict the suppression of charm production, although it is clear that higher order corrections could have a substantial impact on charm production in these models.

A comparison of the different models and their predictions for $R(\text{charm})$ was performed in [76]. Figure 2.9 is taken from this study and shows the β and Q^2 dependencies of $R(\text{charm})$ at low x_P . The first crucial feature of these plots to note is the effect of the charm mass. The hadronic system X will contain two charm quarks, setting a lower limit on M_X of $2m_c$. This feeds through to β , setting an upper limit, β_{max} , given by

$$\beta_{max} = \frac{Q^2}{Q^2 + 4m_c^2} \quad (2.21)$$

This feature is clearly demonstrated in figure 2.9.

The first model shown in figure 2.9 is that proposed by Nikolaev and Zakharov (labelled “NZ” in the figure). In their original publication [62], a 1 % charm contribution was predicted. However, more recently, they have proposed a larger charm contribution [77], but only at very low x_P . This is because very low x_P implies very low Bjorken x , where the gluon density becomes large, cancelling out the effects of colour transparency and hence the suppression of charm production. A charm contribution of roughly 10 % at $x_P = 0.005$ is predicted, as indicated in the figure. The β and Q^2 dependencies are predicted to be flat in this model, the only departure being caused by the β cut-off.

The next models listed on the plots are those proposed by Gehrmann and Stirling based on the Ingelman-Schlein model. The first, labelled GS(I), is the model in which the pomeron is treated like the proton with the equivalent of the momentum sum rule given in equation 1.8 imposed. The Q^2 and β dependencies are those expected from the DGLAP evolution equations. The second model, labelled GS(II), includes a contribution from direct coupling to quarks, leading to an enhancement in the cross section at high β . This is clear when the predictions from the two models are compared.

The next prediction shown in the figure is that made by Diehl, based on the Landshoff-Nachtmann model [78]. As was discussed in the previous section, the shapes of the spectra from this model are very similar to those from the Nikolaev-Zakharov model, although due to the difference in normalisation $R(\text{charm})$ is predicted to be smaller for non-perturbative

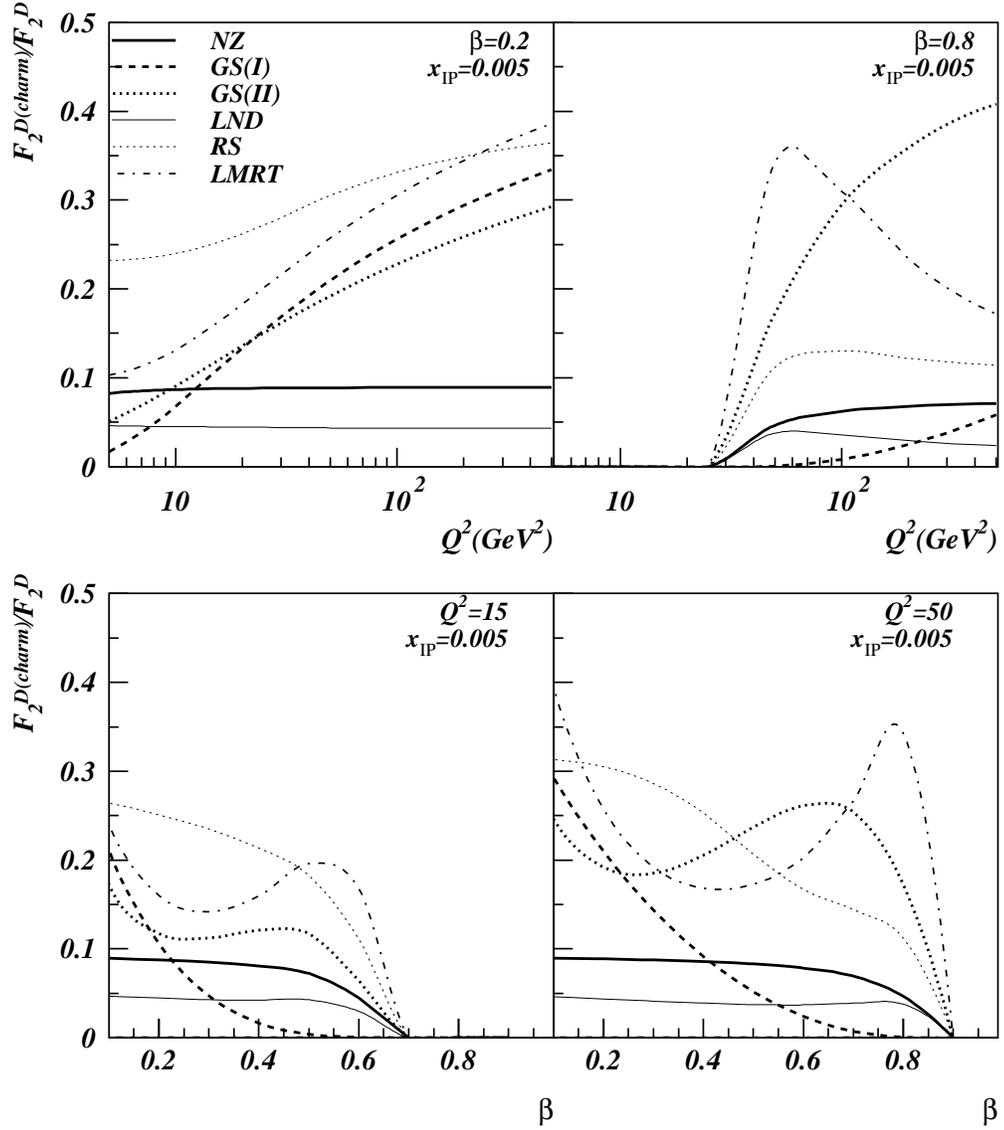


Figure 2.9: The ratio of diffractively produced charm at $x_P = 0.005$ as a function of Q^2 at fixed β and β at fixed Q^2 . Various different models are compared, the details of which are given in the text.

gluon exchange. Figure 2.9 clearly shows that this model (labelled “LND”) predicts a small charm contribution compared to the perturbative gluon models.

The last two models in figure 2.9 are the Monte Carlo simulation by Ryskin and Solano (labelled “RS”) and the LMRT model, both of which are based on the Ryskin model [65].

The leading order cross sections predicted by LMRT applicable in all but the β cut-off region, are given by

$$\begin{aligned}\sigma_L &\sim \frac{2 \langle m_T^2 \rangle Q^2}{(Q^2 + M_X^2)^3} \\ \sigma_T &\sim \frac{M_X^4}{(Q^2 + M_X^2)^3}\end{aligned}\tag{2.22}$$

where $\langle m_T^2 \rangle = m_c^2 + \langle k_T^2 \rangle$. $\langle k_T^2 \rangle$ is the average quark transverse momentum and is of the order of m_c^2 . These cross sections, when combined, give the steep rise of $R(\text{charm})$ with increasing β . The sharp drop at high β is due to the β cut-off. The rise at low β is caused by the contribution from the higher order real gluon corrections. A comparison of the Nikolaev and Zakharov and LMRT predictions clearly shows the enhancement which is the result of including the virtual gluon corrections.

The differences between LMRT and RS are caused by the different implementation of the higher order corrections and the choice of gluon density parametrisation.

Although they are not shown in figure 2.9, other predictions have been made for charm production. For example, Buchmüller, Hebecker and McDermott [79] find that the leading order process $\gamma^* \rightarrow c\bar{c}$ is suppressed, resulting in the higher order $\gamma^* \rightarrow c\bar{c}g$ process becoming the dominant charm production mechanism. It is also found that this higher order process dominates in the diffractive production of light quarks, implying that the ratio of charm in diffraction should be similar to the ratio of charm production in non-diffractive DIS. At $x = 10^{-3}$ and $Q^2 = 36 \text{ GeV}^2$, they predict $R(\text{charm}) \approx 0.2$. This is similar to the non-diffractive measurement of $F_2^{c\bar{c}}/F_2$ made by both ZEUS [16] and H1 [80].

Predictions for charm production have also been made by Alvero, Collins and Whitmore [81] based on the fits made in [55]. They do not make predictions for $R(\text{charm})$, but do find that the factorisable pomeron model describes well the early diffractive open charm measurements from ZEUS [82] and H1 [83], providing a substantial hard gluon contribution is included. Models without a gluon contribution significantly undershoot the data.

More recent results on diffractive open charm production from ZEUS [59], have been compared to the Alvero, Collins and Whitmore results as well as a simpler hard gluonic factorisable pomeron. The predictions from Buchmüller, Hebecker and McDermott are also included in the comparison. All three predictions are compatible with the data. However, it is not possible to distinguish between the three due to the limited statistics of the data. No comparison was made to any two-gluon exchange models. However, the level of diffractive open charm production observed so far is clearly too large to be accounted for by the original prediction by Nikolaev and Zakharov. Further comparisons to data, as well as higher statistics, are definitely required.

2.5 Summary

In this chapter the background to modern diffractive theory has been described and the three main approaches to the theoretical modelling of diffraction within the context of QCD have been discussed in detail. The main expectations of these different approaches have been illustrated with examples from a range of the more popular models.

The models currently available to describe diffractive interactions can be divided broadly into three categories: The first is the factorisable non-perturbative pomeron model, in which the pomeron is treated as an extended object with a partonic structure whose evolution with β and Q^2 can be described by the DGLAP equations. The process can be factorised into the non-perturbative emission of the pomeron by the proton and the hard scatter of the virtual boson with one of the partons in the pomeron. Charm production proceeds via boson gluon fusion with one of the gluons in the pomeron, resulting in an unsuppressed charm production rate. The second category are the perturbative QCD pomeron models, in which the diffractive exchange is described by the t -channel exchange of two gluons or a gluon ladder in a colour singlet state, resulting in a cross section which is proportional to the square of the proton gluon density. All calculations are performed in the proton rest frame in which the critical parameter in the calculation of the interaction cross section is the transverse separation, ρ , of the $q\bar{q}$ pair into which the photon fluctuates in the leading order case. It is found that $\rho \sim 1/m_q$, where m_q is the mass of the quark produced, indicating that heavy flavours are produced with a small transverse separation. Such configurations are suppressed as a result of colour transparency. However, it has been found that higher order real and virtual gluon corrections play an important role in

cancelling out the suppression.

The third and final category is the soft colour interaction model, in which the hadronic system produced by photon fluctuation undergoes soft interactions in the proton colour field. In approximately 1/9 of all events, the colour structure of the hadronic system is altered to produce a colour singlet state, causing a rapidity gap. Once again heavy flavours are suppressed at leading order, but that the suppression is cancelled by higher order corrections.

The importance of charm production in diffraction has also been discussed. Its importance from an experimental point of view arises from its sensitivity to the role of gluons in the diffractive exchange. Charm production is clearly also an effective discriminator between the different theoretical models. From the theoretical point of view, it is a valuable process given the hard scale provided by the large charm mass, which indicates the validity of using perturbative QCD.

In order to study diffractive charm production effectively, it is necessary to measure the dependence of the charm cross section on a variety of different kinematic variables. The remaining chapters of this thesis describe the method through which the necessary data is recorded, reconstructed and analysed in order to achieve this aim.

Chapter 3

HERA and the ZEUS Detector

This chapter gives an overview of the HERA accelerator and its performance during the 1995 — 1997 data-taking periods, as well as a description of the components of the ZEUS detector which are used in this analysis.

3.1 The HERA Collider

HERA (Hadronen Elektronen Ring Anlage), situated at the DESY laboratory, Hamburg, was the first electron¹-proton collider ever to be built. The electron and proton rings are located 10 — 25 m underground in a circular tunnel a little over 6 km in circumference. Figure 3.1 shows the layout of the HERA ring, along with the four experimental halls in which the 820 GeV proton beam and the 27.5 GeV electron beam are brought into collision. ZEUS is housed in the South Hall with HERMES to the east and HERA-B to the west. The other general purpose detector, H1, is located in the North Hall.

Figure 3.1 also shows the pre-accelerator chain required to accelerate the electron and proton beams prior to injection into the main HERA ring. The electrons are radiated from a heated Tungsten wire and accelerated to 500 MeV in a linear accelerator (Linac) prior to being collected into bunches. They are then passed on to a small synchrotron ring (DESY II) and accelerated to 7 GeV. The bunches are then transferred to PETRA, in

¹Although HERA was originally designed as an e^-p collider, it changed to e^+p collisions during 1994 due to beam lifetime problems. In this chapter *electron* is used as a generic name for the lepton beam. In subsequent chapters, it will be referred to as the positron beam.

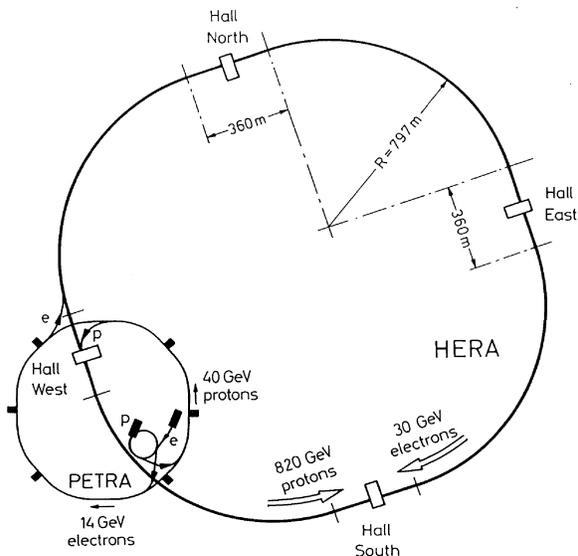


Figure 3.1: A Schematic of the HERA ring, including the pre-accelerator chain.

which up to 70 bunches may be stored, and accelerated to 14 GeV before they are injected into the main HERA ring.

The protons are produced by the ionisation of hydrogen, and are again accelerated in a Linac to 50 MeV before being collected into bunches in another synchrotron ring (DESY III). They are then also transferred to PETRA, where up to 70 proton bunches can be accelerated to 40 GeV. They are then also injected into the main HERA ring.

The main proton ring at HERA uses superconducting magnets to generate the 4.68 T field which is required to keep the protons circulating at 820 GeV. The protons are accelerated to their colliding energy in the HERA ring through the use of radio frequency cavities and a 208 MHz feedback system. Conventional magnets are used to keep the electrons circulating at 27.5 GeV. Their acceleration to colliding energy is achieved through the use of standard and superconducting RF cavities. The two beams are separated into 174 equidistant buckets which contain the particle bunches and are separated in time by 96 ns. Colliding bunches are those in which a “full” proton bucket coincides with a “full” electron bucket. The bunch structure can also be arranged so that a electron bunch can coincide with an empty bunch in the proton beam (or vice versa). These are known as pilot bunches. It is equally possible to have two empty bunches coinciding. These pilot bunches and empty bunches are used to study background contributions from interactions of the beam with gas in the beam pipes and also from cosmic rays.

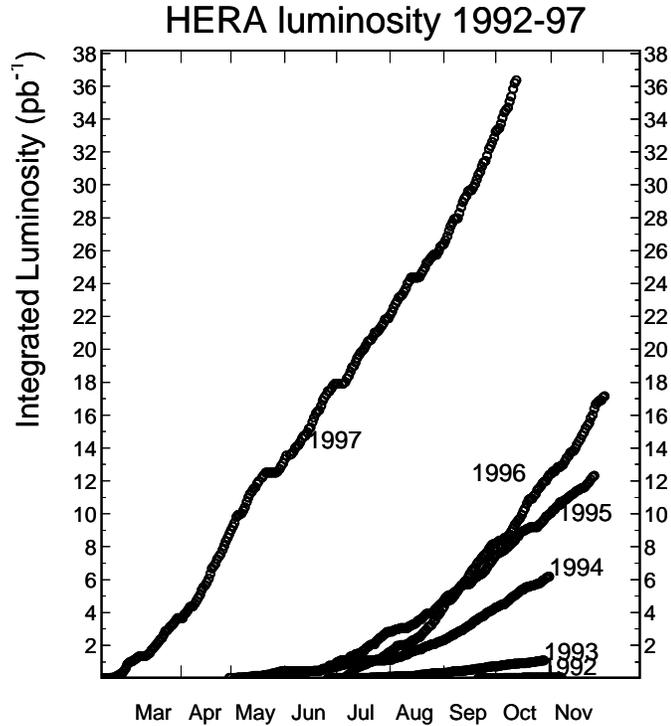


Figure 3.2: The HERA luminosity delivered to the ZEUS experiment since the start of data-taking in 1992.

3.2 HERA Performance 1995 - 1997

Since the start of data-taking in 1992, the HERA machine has gradually increased the integrated luminosity it is able to provide each year. This can be clearly seen in figure 3.2. The sharp increase in luminosity from 1994 onwards was caused by HERA switching from electrons to positrons, which resulted in a substantially improved lepton lifetime.

The data on which this analysis is based comes from the 1995, 1996 and 1997 datasets. The performance of the HERA machine during these three periods is summarised in table 3.1 and compared to the original design values.

3.3 The ZEUS Detector

Figures 3.3 and 3.4 show the overall layout of the central section of the ZEUS detector. A detailed description of the ZEUS detector can be found in [84, 85], but a description of the main components used in this analysis will be given here. ZEUS is multipurpose

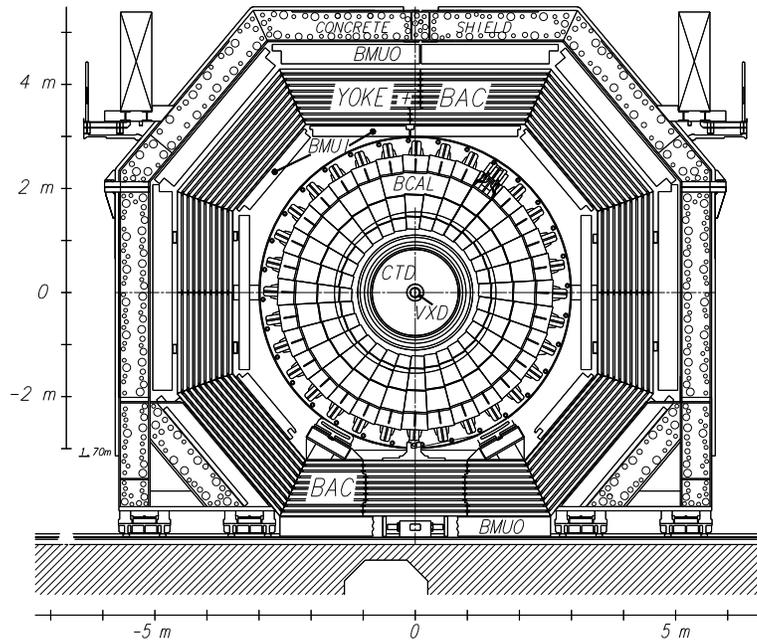


Figure 3.3: 2-d cross section of ZEUS in the $r - \phi$ plane.

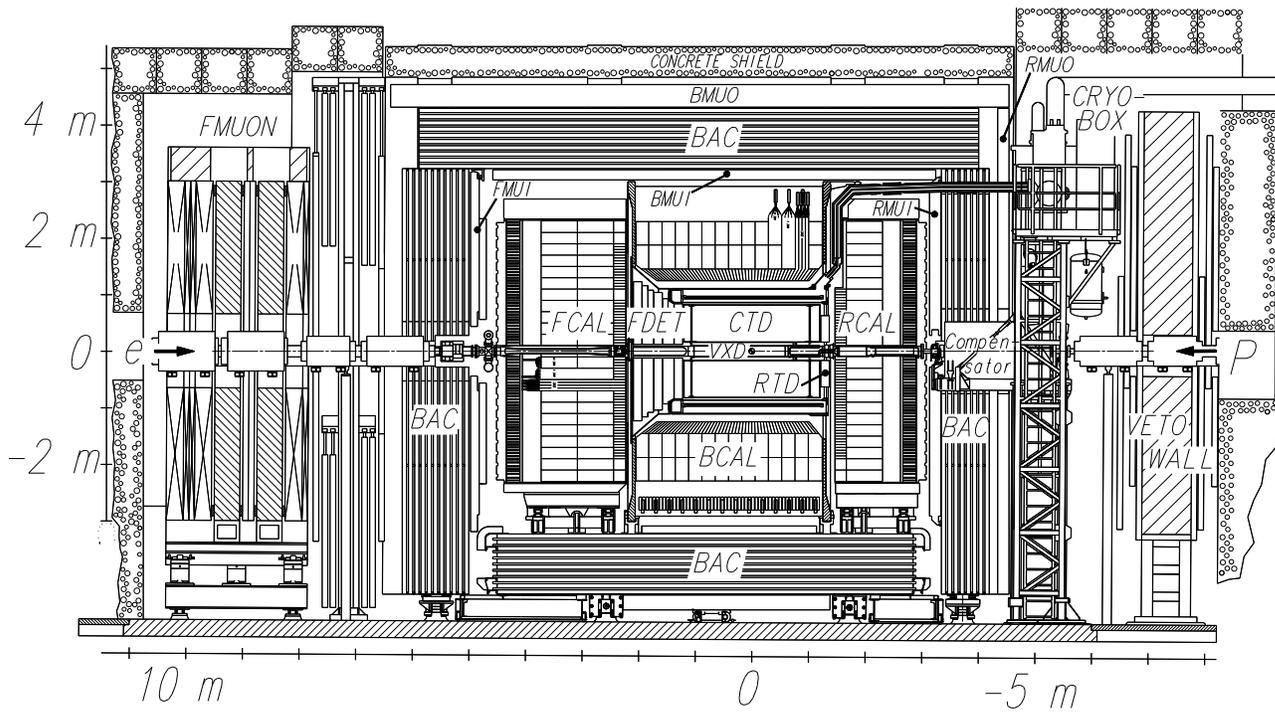


Figure 3.4: 2-d cross-section parallel to the beam pipe of the ZEUS detector.

Parameter	1995		1996		1997	
	e^+	p	e^+	p	e^+	p
Beam Energies (GeV)	27.5	820	27.5	820	27.5	820
Beam Currents (mA)	30-41	55-73	33	65	36-43	77-105
Instantaneous Luminosity \mathcal{L} ($10^{30}\text{cm}^{-2}\text{s}^{-1}$)	3.9-7.2		6.1-10.3		8.44-14.0	
Integrated Luminosity (pb^{-1})	12.15		17		36.5	
Centre of Mass Energy (GeV)	300		300		300	
Number of Colliding Bunches	174		174		174	
Number of Pilot Bunches	6	15	6	15	6	15
Parameter	Design					
	e^\pm			p		
Beam Energies (GeV)	30			820		
Beam Currents (mA)	58			163		
Instantaneous Luminosity \mathcal{L} ($\text{cm}^{-2}\text{s}^{-1}$)	1.7×10^{31}					
Annual Integrated Luminosity (pb^{-1})	100					
Centre of Mass Energy (GeV)	314					
Number of Colliding Bunches	210					
Number of Pilot Bunches	15			17		

Table 3.1: Performance of the HERA Machine for the period 1995 — 1997

detector with almost complete solid angle coverage (the exceptions being the beam pipe holes), which has been designed to give the optimum detection of the heavily boosted event topologies seen at HERA. In the layout shown in figure 3.4, the electron beam comes in from the left and the proton beam from the right. The coordinate system is defined such that the incident beams travel along the Z axis with $Z = 0$ as the nominal interaction point. The direction of the proton beam defines the positive Z -direction and $\theta = 0$, where θ is the polar angle. The X - Y plane is perpendicular to the beam axis with X horizontal. The azimuthal angle ϕ is measured with respect to the X axis. Thus the region of positive Z (to the left of the interaction point) is referred to as the forward region, while the negative Z region is known as the rear region. Due to the large boost that the interaction products receive due to the asymmetry in the beam energies, the detectors in the forward region are of much greater depth compared to those in the rear region.

The tracking system consists of several chambers, starting with the vertex detector (VXD) directly surrounding the central beam pipe. This was present until the end of the 1995 data-taking period, but was only active in a small fraction of the data taken that year. The VXD is surrounded by the Central Tracking Detector (CTD), which is a large drift chamber used in the identification and measurement of charged particles. The tracking system is completed by the forward and rear tracking detectors (FTD and RTD), which provide added coverage in the forward and rear beam pipe regions. The CTD is surrounded by a superconducting magnet which provides a field of 1.43 T.

The tracking system is surrounded by the high-resolution Uranium Scintillator Calorimeter (UCAL). The UCAL is subdivided into three sections: The FCAL which covers the forward end of the tracking system, the RCAL which covers the rear end and the BCAL, which provides coverage in the barrel section between the FCAL and RCAL. At a depth of $3X_0$ (one radiation length (X_0) is roughly 2.6 mm of scintillator and 3.3 mm of Uranium) in the RCAL, the Hadron-Electron Separator (HES) can be found. This is used to discriminate between electrons and hadrons through the measurement of longitudinal shower profiles. There is also a small calorimeter at the back of the RCAL directly adjacent to the beam pipe (BPC). This is designed to detect scattered electrons which would otherwise pass undetected through the rear beam pipe hole.

The Rear Presampler (RPRES) is located on the inner face of the RCAL and provides information about the showering caused by dead material between the interaction point and the calorimeter. The Small Angle Rear Tracking Detector (SRTD) is also mounted on the inner face of the RCAL in the beam pipe region and provides further information about particle showering as well as excellent reconstruction of the calorimeter impact position of electrons scattered through small angles. The SRTD will be discussed in detail in chapter 5.

The remainder of the central detector consists of the backing calorimeter (BAC) and the muon detection system. The BAC is sandwiched between the two sets of muon detectors and is designed to pick up particles which do not shower until after passing through the main calorimeter. It also serves as the return yoke for the magnetic field.

The vetowall and C5 counter are situated outside the main detector, behind the rear muon chambers. The vetowall protects the main detector from the effects of the proton beam halo and is used to reject beam-related backgrounds. The C5 counter surrounds the beam

pipe at $Z = -3.15$ m and is used to monitor the beam profiles using timing measurements.

There are also several other detector components, which lie along the beam pipes outside the main detector. In the rear direction, there are several electron taggers, known as the 8 m tagger, the 35 m tagger and the 44 m tagger. The 35 m tagger is one half of the Luminosity Monitor (LUMI), which is designed to detect the final state electron and photon from the Bethe-Heitler process, $ep \rightarrow ep\gamma$, through which the luminosity can be determined.

There are also several components in the forward direction, such as the Leading Proton Spectrometer (LPS) which consists of a series of Roman pots distributed along the proton beam line. This is used to measure the momentum of the scattered proton in the small fraction of events in which the proton does not fragment. The Proton Remnant Tagger (PRT) and Forward Neutron Calorimeter (FNC) are also located in the forward direction.

3.3.1 The Central Tracking Detector

The Central Tracking Detector (CTD) is a cylindrical drift chamber which is 205 cm long and which has an inner(outer) radius of 18.2 cm(79.4 cm), covering an angular range of $15^\circ < \theta < 64^\circ$. It consists of 24192 wires, 4608 of which are sense wires. The chamber is separated into nine superlayers which are further divided into octants. Each superlayer contains eight sense wire layers which can provide up to 72 $r\phi$ measurements or “hits” from which tracks can be reconstructed. The wires in the odd-numbered superlayers are parallel to the beam direction (axial layers), while those in the even-numbered layers are tilted by $\pm 5^\circ$ with respect to the beam (stereo layers). This system allows the reconstruction of tracks in all three dimensions. The angle for the stereo layers is chosen so that the polar and azimuthal angular resolutions are roughly equal. The layout of a CTD octant is shown in figure 3.5, where the sense wires are indicated by the larger dots. Table 3.2 gives more information about the properties of the different superlayers. Each superlayer is divided up into cells, each of which contains eight sense wires. The cells are arranged such that the sense wire plane is at 45° to the radius vector. This ensures that there is no left-right ambiguity. The magnetic and electric fields and the gas mix are all chosen to give a Lorentz angle of approximately 45° , ensuring that the ionisation drift direction is perpendicular to the track direction. This makes the radial position of each hit roughly equal to that of the sense wire which detected it.

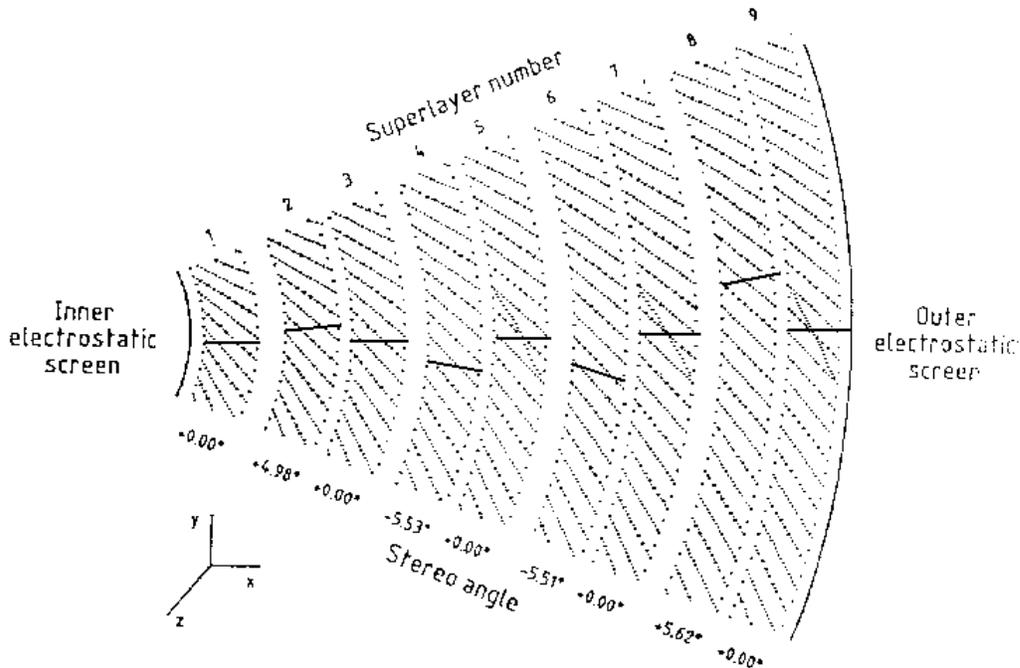


Figure 3.5: An octant of the CTD. The larger dots indicate the sense wires

The presence of the large magnetic field allows precise momentum measurements. The sense wires are read out using 100 MHz flash analogue to digital converters (FADCs) in $r\phi$ (and z for the stereo layers). In addition, the first three axial superlayers are equipped with *Z-by-Timing* electronics, which determine Z through the measurement of the arrival time of the pulses at the two ends of the wire. The *Z-by-Timing* method has a resolution of ≈ 4 cm, while the resolution using the stereo layers is ≈ 1 mm. The two track resolution is about 2.5 mm. Providing a track has hits in all nine superlayers, it will be reconstructed with a resolution of $\frac{\sigma(p_T)}{p_T} = 0.005p_T \oplus 0.0016$ GeV.

A more detailed description of the CTD can be found in [86].

3.3.2 The Calorimeter

At HERA it is important to be able to distinguish between electromagnetic and hadronic objects and to be able to reconstruct their energy and position as accurately as possible, to ensure that the kinematic variables x , y and Q^2 are well-reconstructed. This is best done using a sampling calorimeter, which consists of alternating layers of active and inactive material. Lead or uranium are common choices for the inactive material, while scintillator

Superlayer	No. of Cells	Mean Stereo Angle	Centre Radius (cm)	θ range
1	32	0°	20.97	11.3 — 168.2°
2	40	4.98°	27.23	14.5 — 164.8°
3	48	0°	35.00	18.4 — 160.7°
4	56	-5.31°	41.30	21.5 — 157.3°
5	64	0°	48.73	24.9 — 154.0°
6	72	-5.51°	55.52	27.9 — 151.0°
7	80	0°	62.74	30.9 — 147.9°
8	88	5.62°	69.46	33.5 — 145.2°
9	96	0°	76.54	36.1 — 142.6°

Table 3.2: The dimensions of the CTD Superlayers.

or gas provide the active material layers.

When the UCAL was originally designed, the requirements made were as follows

- Energy resolutions of $35 \pm 0.3\%/\sqrt{E}$ for hadrons and jets and $18 \pm 0.1\%/\sqrt{E}$ for electrons, where E is energy measured in GeV,
- Jet angular resolution better than 10 mrad,
- Timing resolutions of less than 1 ns, to allow the rejection of cosmic rays and beam-gas interactions.

In order to achieve these requirements, 3.3 mm thick plates of depleted uranium (DU) were chosen along with 2.6 mm thick tiles of plastic scintillator. The plate thicknesses were optimised to achieve equal response to electromagnetic and hadronic showers.

The calorimeter is divided up into three sections. The forward calorimeter (FCAL) covers the angular range 2.2° — 36.7° , while the rear calorimeter (RCAL) covers the range 129.1° — 176.5° . The barrel calorimeter (BCAL) covers the intermediate angular range.

Each calorimeter is divided longitudinally into electromagnetic (EMC) and hadronic (HAC) sections, which are divided transversely into cells. Each cell is made up of a scintillator tile and a depleted uranium plate. In the FCAL and BCAL, the EMC cells are 5×20 cm,

while in the RCAL they are 10×20 cm. The HAC cells are 20×20 cm in all sections. The layout of the FCAL, as seen from the interaction point, is shown in figure 3.6. It is divided into 24 modules along the Y axis, each of which is made up of towers 20×20 cm, going away from the interaction point. The smaller divisions indicate the dimensions of the EMC cells.

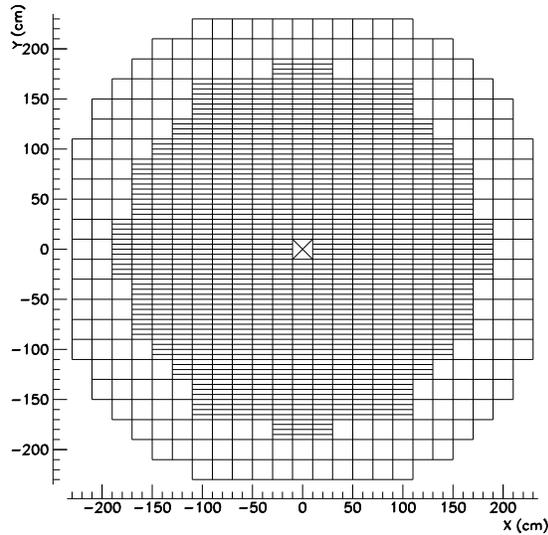


Figure 3.6: The FCAL as seen from the interaction point

Figure 3.7 depicts an FCAL module, and shows how it is separated into EMC and HAC sections. It also indicates the wavelength shifters (WLS) down each side of each cell, through which the energy deposited in the scintillator is passed to the photomultiplier tubes (PMTs) at the rear of the module. Each cell is readout by two PMTs providing redundancy and position reconstruction. The natural radioactivity of the uranium provides a energy calibration method on a channel-by-channel basis. The calibration using this method is good to 1 %.

3.3.3 Preshower Detectors

Although ZEUS was designed to minimise the amount of material through which particles have to pass before detection, it is inevitable that there will be some dead material (cables, supports etc.), which will cause particles to lose energy prior to detection. It is therefore necessary to determine the exact effect of the dead material, particularly in the case of the

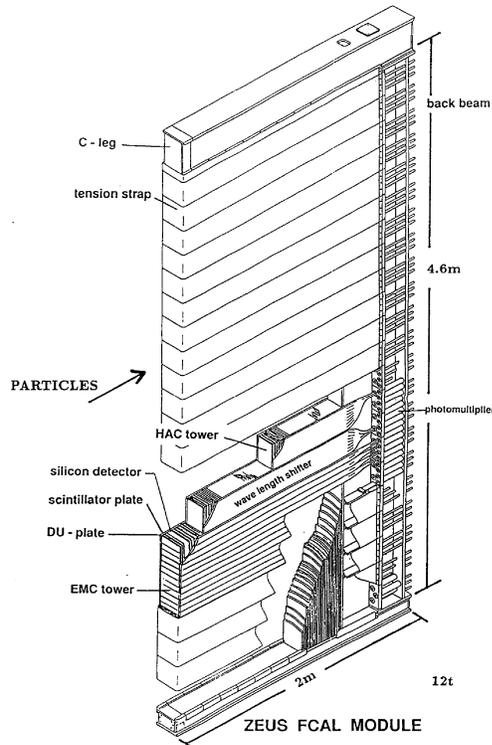


Figure 3.7: One of the FCAL modules in the ZEUS detector.

scattered electron, as this can effect the measurement of the kinematic variables. There are several methods of correcting for energy loss from the scattered electron, involving the tracking detectors and more importantly the series of preshower detectors installed within ZEUS for this purpose. For electrons which hit the calorimeter in the rear beam pipe region, there is the Small Angle Rear Tracking Detector (SRTD) which will be discussed in detail in chapter 5. The alternatives are offered by the Forward and Rear Presamplers. A Barrel Presampler is also being installed in time for 1998 running. However, the majority of the scattered electrons in the sample on which this analysis is based are detected in the RCAL and hence energy loss corrections based on presampler information are only required for this region. Thus the forward presampler will not be discussed here.

The Rear Presampler (RPRES) consists of a layer of scintillator tiles of the same dimensions as the HAC cells of the calorimeter. These are read out by wavelength shifters which guide the scintillator signal to a series of PMTs. The RPRES is calibrated using Minimum Ionising Particles (MIPs), which produce at least 5 photoelectrons per MIP at the photocathode of the PMT. These tiles are mounted on the inner face of the RCAL and cover all but the outer two rings of cells.

The RPRES does not provide accurate position reconstruction, but the energy deposited in the RPRES tile associated with the scattered electron can be used to correct for any energy loss that the electron may suffer.

3.3.4 Luminosity Measurement

Luminosity is determined at ZEUS by measuring the rate at which bremsstrahlung events from the Bethe-Heitler process $ep \rightarrow e'p\gamma$ [87] are produced. The final state electron and photon are detected by a pair of lead scintillator calorimeters behind the main detector, as shown in figure 3.8. The Bethe-Heitler process is used for luminosity determination because its cross section is large and very accurately predicted by QED. Such interactions also have a very characteristic topology. As shown in figure 3.8, the bremsstrahlung

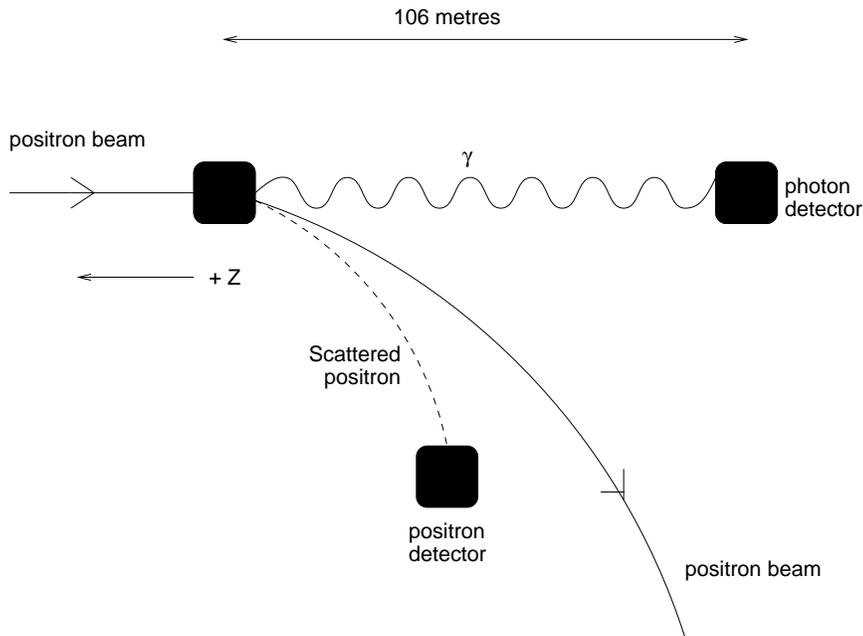


Figure 3.8: The Layout of the ZEUS Luminosity Monitor and paths of the different particles from the Interaction Point (IP).

photons leave the electron beam pipe via an exit window and hit the photon detector (referred to as the LUMI- γ). The LUMI- γ has an energy resolution of $\sigma(E_\gamma) = 0.18\sqrt{E_\gamma}$, where E_γ is the photon energy measured in GeV.

After emitting the bremsstrahlung photon, the final state electron is bent away from the beam direction by the bending magnets of the electron ring. It is then detected in the LUMI- e , providing it is scattered through an angle of less than 6 mrad and its energy lies

in the range $0.2E_e < E'_e < 0.9E_e$, where E_e is the nominal electron beam energy and E'_e is the scattered electron energy. Both the LUMI- e and the LUMI- γ are known to have high detection efficiencies.

The LUMI- e is also used to tag electrons in the range $10^{-7} < Q^2 < 2 \times 10^{-2} \text{ GeV}^2$, while the LUMI- γ can be used to tag initial state radiative photons from deep inelastic scattering events.

Chapter 4

Data Preselection

The short time interval between bunch crossings at HERA (96 ns) results in a high beam crossing rate (~ 10 MHz) which poses a challenge for the readout and storage of data. In addition the interaction rate varies between 10 and 100 kHz, with useful ep physics accounting for only a small fraction of this rate (\sim few Hz). The dominant contribution to the interaction rate comes from non- ep background which must be identified and rejected. In fact, it is necessary to reduce the rate to the level of a few Hz so that the data can be written to tape. This is the purpose of the ZEUS online three-level trigger.

The aim of this chapter is to describe the techniques used to convert the raw information from the detector into a high quality, fully reconstructed dataset for use in the studies described in subsequent chapters. First, the ZEUS online trigger and the backgrounds it must reject will be discussed, with particular emphasis on the requirements for triggering on DIS interactions. Then the offline reconstruction software will be discussed, along with the post-reconstruction analysis tools used to prepare the data for full physics analysis. The parallel processes used to simulate this chain for Monte Carlo will also be discussed.

4.1 The Data Acquisition System

As has already been explained, the interaction rate is dominated by backgrounds which do not arise from ep collisions. In addition, background in a DIS data sample can come from other physics processes, such as photoproduction. In this section, the principal sources of background and their rejection will be discussed, along with how this rejection

is implemented in the online trigger system.

4.1.1 Backgrounds

The main sources of background which are not the result of an ep collision are

Beam Gas. These events are caused by collisions between the proton beam and any residual gas in the beampipe and are produced at a rate of approximately 50 kHz. Collisions between the positron beam and residual gas can also occur, but are much less frequent as the interaction cross section for positrons is much lower than that for protons.

Synchrotron Radiation. This is the main source of background from the positron beam and is caused by the bending of the beam in the HERA magnetic field. This bending does not occur near the ZEUS detector, limiting this contribution to the background rate. There are also collimators and shielding in place to further reduce such backgrounds.

Proton Beam Halo. These events are caused by interactions between protons in the beam halo and the beampipe wall or other beamline components. These interactions tend to produce muons which leave a characteristic signature in the detector, allowing them to be identified and rejected.

Cosmic Muons. This is one of the more significant backgrounds which occur within the main detector volume, but once again, they have a characteristic signature. The rate of cosmic muons can be as high as 5 kHz.

The majority of these backgrounds are rejected by the trigger system using two main techniques. The first method is based on timing information from the calorimeter and the SRTD. For a standard ep collision occurring at the nominal interaction point, $t = 0$ is defined as the time at which the interaction products hit a given detector section, as depicted in figure 4.1(a). Proton beam-gas and beam halo interactions occurring prior to the entry of the proton beam into the main detector volume (as indicated in figure 4.1b) would have a large negative RCAL/SRTD time (~ -10 ns), as well as a large positive difference between the time measured in the FCAL and the RCAL. These backgrounds would also produce a signal in the vetowall and the C5 collimator.

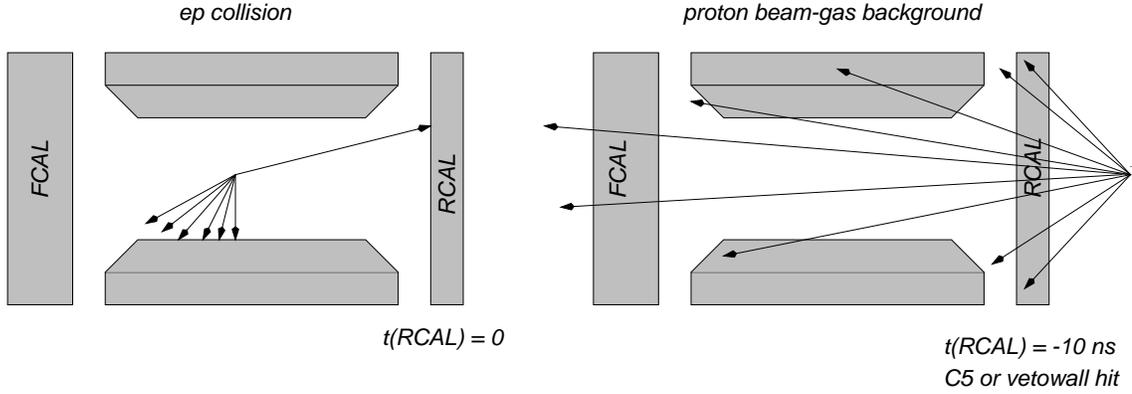


Figure 4.1: The topology expected from (a) a standard ep event and (b) a proton beam-gas interaction occurring prior to the entry of the proton beam into the main detector volume.

The other main background rejection method is based on momentum conservation, using the quantity

$$\delta = E - p_z (+ 2E_\gamma) = \sum_i E_i(1 - \cos \theta_i) (+ 2E_\gamma) \quad (4.1)$$

where i indicates the sum over the calorimeter cells and E_γ is the energy measured in the photon calorimeter of the luminosity monitor, indicating the presence of an initial state photon. A proton beam-gas interaction of the type shown in figure 4.1(b) would give a very high value of δ , while a well-contained NC DIS event such as that shown in figure 4.1(a), would have $\delta \simeq 2E_e$, where E_e is the positron beam energy. A proton beam-gas interaction occurring within the main detector volume would send a shower of particles in the proton beam direction, resulting in a very small value of δ . Appropriately placed cuts on the value of δ provide a powerful background rejection tool.

The dominant physics background in a DIS data sample arises from photoproduction processes. In such interactions, the scattered positron disappears undetected down the rear beampipe, resulting in a low value of δ . Thus, δ also plays an important part in offline data selection, as well as in the online triggering process.

4.1.2 The ZEUS Trigger System

The purpose of the online three-level trigger system is to reduce the interaction rate to the level of a few Hz by rejecting background and specifically selecting interesting physics processes. Each level has a system of accepting or rejecting events using increasingly strict requirements and sophisticated selection algorithms, known as trigger “slots”. In order to

decrease the rate further, slots can also be prescaled, which causes them to accept only a fraction of the events which actually pass the selection algorithm. This will be discussed in more detail in the context of the physics studies described in later chapters.

The First Level Trigger (FLT)

The FLT is designed to reduce the input interaction rate of 10-100 kHz (depending on beam conditions) to about 1 kHz. The global first level trigger (GFLT) is a pipeline system which receives information from the first level triggers of the major detector components (referred to as component FLTs). The use of a pipeline increases the time available for data processing and decision-making to 4.4 μ s. If the event is accepted by the GFLT, the data are transferred to the second level trigger, otherwise it is overwritten.

The different FLT slots are based predominantly on information from the calorimeter FLT (CFLT) and the CTD (CTD-FLT). In order to calculate the required quantities in the CFLT, the calorimeter is divided up into 448 “trigger towers” in the REMC, BEMC, FEMC, RHAC, BHAC and FHAC sections. The trigger towers in a particular component, eg. REMC, are chosen such that they do not overlap with each other and usually consist of two adjacent cells.

The key characteristic of a DIS event is the presence of a (usually isolated) scattered positron. Thus the FLT selection requirements for DIS events are based predominantly on the identification of this positron using CFLT information. The quantities used in the identification are

RcalIsoe. This is an algorithm which finds isolated positrons in the rear calorimeter. It simply finds groups of up to four trigger towers with electromagnetic energy deposits, surrounded by trigger towers with no energy deposits.

REMCth. This is the total electromagnetic energy deposited in the RCAL.

REMC. This is the electromagnetic energy deposited in the RCAL excluding the trigger towers immediately surrounding the rear beampipe.

BEMC. This is simply the total electromagnetic energy deposited in the BCAL.

CAL_E. This is the energy summed over all calorimeter towers, except those in the inner three rings in the FCAL and the innermost ring of the RCAL.

Background rejection is performed using the C5, vetowall and SRTD. A signal in the C5 or vetowall causes an event to be rejected, while if any of the SRTD quadrants have bad timing measurements, the event will also be rejected. If the SRTD cannot provide any timing information, the event will not be rejected. These SRTD requirements are included in some of the triggers using the collective name **SRTDgood**. The two FLT slots used for DIS selection in 1995 combine the above CFLT quantities in the following way

FLT30 $R_{callSoe} * REMC_{th} > 3.75 * SRTD_{good}$, where * indicates a logical **.and.**

FLT44 $REMC > 3.4$ **.or.** $BEMC > 4.8$.

In 1996/7 (which accounts for approximately 84 % of the total data sample considered in this thesis), the requirements made on CFLT information changed and information from the CTD-FLT was included in order to decrease further the rate. The quantities from the CTD-FLT which are used are

anyTRK. This is the requirement that at least one FLT track should be found.

TRKveto. This is the requirement that a high proportion of the FLT tracks found in an event must point to the nominal vertex region. If the CTD-FLT finds more than five tracks in an event which do not point to the nominal vertex, the event is rejected.

The principal FLT DIS slots in 1996/7 then became

FLT30 $R_{callSoe} * (REMC \geq 2.03$ **.or.** $REMC_{th} \geq 3.75$ **.or.** $(CAL_E \geq 0.5 * SRTD_{good}))$

FLT44 $(BEMC > 4.8 * anyTRK)$ **.or.** $REMC > 3.4$

FLT46 $FLT30 * TRK_{veto}$

The Second Level Trigger (SLT)

The SLT is designed to reduce the FLT output rate to about 100 Hz through the use of more sophisticated algorithms. Unlike the FLT, the SLT is software-based, but also consists of a global second level trigger (GSLT) which makes its decisions based on information from several component SLTs. If an event is accepted by the GSLT, it is passed on to the “Event Builder” which prepares it for the third level trigger.

The SLT background rejection includes both timing cuts (based on the calorimeter) and the use of δ . More sophisticated algorithms are employed to select specific physics processes. No SLT triggers are explicitly used in the physics studies described in this thesis — the events will therefore have been accepted by any of the wide range of SLT slots available.

The Third Level Trigger (TLT)

The TLT is also software-based and is designed to make the final reduction of the trigger rate to a few Hz. The data from all detector components are fed from the Event Builder into the TLT, which runs a reduced version of the offline reconstruction software. These reconstructed data are used to define physics-motivated filters.

DIS event selection at the TLT requires the use of several positron finders. During the 1995 data-taking period, the ELEC5 [88] and LOCAL [89] finders were in use, while in 1996 and 1997, SINISTRA [90] and EMILLE [91, 111, 112] were added.

SINISTRA is the positron finder which is used in the final data selection and is described in section 4.3.2. The other positron finders in the TLT are based on either cones or islands. Cone-based positron finders simply use calorimeter cells contained within a fixed radius cone around the highest energy cell to define the positron. Only these cells contribute to the measurement of the energy and impact position of the positron. Islands are described in more detail in section 4.2. These two different clustering methods are illustrated in figure 4.2.

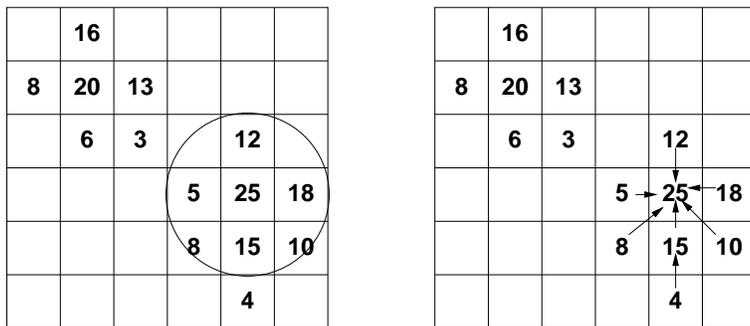


Figure 4.2: The different methods of clustering used in the positron finders implemented in the Third Level Trigger. (a) shows the cone method of clustering, while (b) shows the island method.

The $x - y$ impact position of the scattered positron when it is detected in the RCAL can be reconstructed using either the calorimeter or a reduced version of the SRTD reconstruction

software (SRTLTP0 [92]), which is described in section 4.3.3. If SRTD information is available it is used in preference to the calorimeter information due to its superior spatial resolution.

There are three main TLT slots for selecting NC DIS events. Each of these requires that one of the DIS-related FLT slots should have fired, that $\delta > 30$ GeV and that $E - p_z < 100$ GeV. A scattered positron with energy greater than 4 GeV should also have been found by one or more of the positron finders. If the positron is found in the RCAL, an $x - y$ impact position requirement (known as the “box cut”) is applied in each of the three TLT slots with increasing severity, equivalent to applying a cut on Q^2 . For this reason the slots are known as the “low”, “medium” and “high” Q^2 slots. In the low Q^2 slot, the cut is made at $|x| > 12$.or. $|y| > 6$ cm, while in the medium Q^2 slot the cut is made at $|x| > 14$.or. $|y| > 14$ cm. In 1996/7 an alternative medium Q^2 slot was also used in which a cut was placed on the radius of the impact position from the centre of the RCAL beampipe at 25 cm. The high Q^2 slot was not used in any of the physics studies presented here.

Events which are accepted by the TLT are then written to tape and made available for offline reconstruction — the subject of the next section.

4.1.3 Simulation of the Detector and DAQ System

In order to extract physics results from the reconstructed data, it is necessary to correct for effects caused by the detector and trigger system. The corrections are obtained by simulating the underlying ep collisions using Monte Carlo techniques and then processing these data using packages which simulate the effects of various parts of the detection system.

The different Monte Carlo generators used in the physics studies described in this thesis are discussed in later chapters. They are all subjected to a common simulation chain which starts with ZDIS — a package which outputs the four-vectors of the generated particles in the format required for the standard ZEUS analysis package, EAZE. The Monte Carlo is then processed using MOZART [84], the ZEUS detector simulation package. MOZART uses GEANT [93] to describe the detector geometry and includes a full simulation of the magnetic field. It also simulates energy loss, multiple scattering and shower development. The output from MOZART is very similar to the raw data from the detector, except that

the ZDIS four-vectors are retained. The ZGANA [94] package is then used to simulate the ZEUS trigger system. Once the Monte Carlo data has passed through this simulation chain, it undergoes the same reconstruction processes as the real data. Thus, from this point onwards no distinction is made between data and Monte Carlo.

4.2 Offline Reconstruction

Once the raw data has been written to tape it is processed by the reconstruction program ZEPHYR. First, the data from each detector component is reconstructed individually, then information from the different components are matched up (where applicable). These reconstructed data can then be used to select specific event types. These events are recorded in the form of a Data Summary Tape (DST). The data from the CTD and calorimeter are the most important to this analysis and so their reconstruction methods are described in some detail here.

4.2.1 Calorimeter Clustering

Each particle which enters the calorimeter deposits energy in a manner characteristic of its identity. For example, electrons and photons shower electromagnetically over a relatively small volume, while hadron-induced showers cover a much larger area, typically several cells. Thus it is important to have a reliable method of grouping calorimeter cells in order to form objects whose characteristics can be used to make deductions about the particle(s) which made them.

Before any local clustering algorithms are used, basic noise suppression is performed. Calorimeter noise comes from the natural radioactivity of the uranium in the calorimeter or noise in the readout electronics and can be suppressed by requiring that the energy in any EMC cell must be greater than 60 MeV, while that in any HAC cell must be greater than 100 MeV.

Once this basic noise suppression has been applied, several clustering algorithms are used. However, only those which are used in the analyses presented in later chapters are described here.

Condensates

Condensates are formed by making a list of all cells ordered by energy, taking the highest energy cell and checking the energy deposited in neighbouring¹ cells within a specific calorimeter section (FCAL, BCAL, RCAL). No attempt is made to cross the boundaries between calorimeter sections during the formation of a condensate. If the energy deposited in the neighbouring cell is greater than some predefined limit (20 MeV for an EMC cell and 50 MeV for a HAC cell), it is combined with the first cell. This process continues until no more adjacent cells with sufficient energy can be found. Each time a cell is used in the formation of a condensate, it is deleted from the original list. Once the list has been completely emptied, the formation process is complete. The condensates are kept for analysis if they pass a further energy cut (100 MeV for purely electromagnetic condensates and 200 MeV for all other condensates), otherwise they are ignored. The energy of the condensate is simply the sum of the constituent cell energies, while the position is the energy weighted average of the cell positions.

Islands

Islands are formed using calorimeter towers and may extend across section boundaries. The energy of each tower is compared with those of its immediate and diagonally adjacent neighbours. The tower will then be associated to its highest energy neighbour, or itself if none of its neighbours has a higher energy. Clusters of towers will be formed as a result, which link back to the nearest high energy tower, as shown in figure 4.2. These clusters are known as Islands. This type of clustering algorithm can also be performed on cells rather than towers — these are referred to as Cell-Islands.

4.2.2 Track and Vertex Reconstruction

The full ZEUS offline track reconstruction is described in detail in [95]; a brief overview is given here. Two different methods of reconstruction are available: regular or CTD-only. The former uses information from the forward and rear tracking detectors (FTD and RTD — see chapter 3) in addition to the CTD in order to increase the angular

¹A neighbour is defined as a cell which shares at least part of one side with the cell it is neighbour to. This process excludes diagonally adjacent cells.

acceptance for track reconstruction. However, in the analysis presented here, the CTD-only offline reconstruction method is used, as it reduces uncertainties in the determination of efficiencies and acceptances due to detector alignment and the presence of dead material.

Track Finding

A track candidate starts out as a track segment, which is defined by three adjacent CTD hits from an axial superlayer in an outer part of the CTD. This track segment is then extrapolated inwards, guided by an additional “virtual hit” at the beamline. As the track is extrapolated inward, it picks up additional axial hits, improving the accuracy with which the trajectory parameters are determined. This series of axial hits will form an arc in the xy plane. If any of the axial hits provides *Z-by-Timing* information this can be used to make a first estimate of the z parameter of the trajectory. The stereo superlayers crossed by this trajectory are then searched for hits, which are added to the trajectory providing more accurate z -component information. All tracks must ultimately have 3D information. Short tracks which only have axial hits in superlayer (SL) 1 are extrapolated to SL2 in order to pick up stereo information.

This method of pattern recognition is performed twice. The first time only track candidates which continue all the way in to SL1 are accepted, while in the second pass, those which have hits in at least two axial superlayers and the intermediate stereo layer, but fail to reach SL1 are also kept. The candidates which fall into the latter category are kept as they may be due to long-lived particle decays. This procedure provides a series of track candidates with 3D information which are then fitted using a five-parameter helix fit [95].

It is also useful, as will be seen in section 4.3.6, to match tracks to calorimeter energy deposits. The matching is performed by “swimming” the track out towards the calorimeter using a GEANT Runge Kutta [93] extrapolation which uses a comprehensive map of the magnetic field. If a low momentum track heading for the FCAL or RCAL reverses direction in z or, when heading towards the BCAL, begins to spiral back towards the centre, the extrapolation to the calorimeter face is abandoned.

Studies have been performed using Monte Carlo in order to establish the efficiency achieved using this package [96]. The studies were carried out using hard photoproduction events which tend to have a higher track multiplicity and used only generated charged particles

which satisfy $p_T > 0.1$ GeV and $|\eta| < 1.75$, ie. those which would be expected to enter the CTD and create a track. It was found that approximately 95 % of all these “true” tracks were reconstructed in the CTD. The remaining 5 % are lost in the beampipe or CTD inner wall. 97 % of tracks which make it into the CTD are reconstructed, 96 % of which will be associated to the primary vertex. These efficiencies increase further if requirements are placed on the true track such that it would be expected to reach the third superlayer. It is also known that for “typical” $D^{*\pm}(2010)$ decay tracks ($p_T \sim 2 - 3$ GeV), the probability of assigning the wrong charge to a track is less than 0.1 % [96].

As was already mentioned in the previous chapter, the transverse momentum resolution achieved using the CTD is $\sigma(p_T)/p_T = 0.005p_T \oplus 0.0016$ (with p_T is GeV) and the two-hit resolution, that is, how far apart two hits must be in order to be distinguished, is of the order to 2 - 3 mm [84].

Primary Vertex Fitting

The reconstruction package also contains a vertex finder which uses standard vertex fitting techniques [95]. Before the vertex fit is performed, track candidates which do not appear to originate from the beamline are removed. Then a simple vertex algorithm is applied which determines the vertex centre of gravity of the remaining tracks. Any tracks which make a large contribution to the χ^2 of the vertex are gradually discarded until the fit achieves an acceptable quality. Once this preliminary vertex and list of tracks has been determined, a full vertex fit is carried out which constrains the remaining track trajectories to come from the vertex and refits their direction and curvature.

4.3 Post-Reconstruction Data Analysis

Once the data has passed through ZEPHYR, it is ready for the final stage of the chain. The remaining sections of this chapter describe those techniques and tools required for the selection and analysis of a DIS event sample. The datasets used in the remaining chapters of this thesis come from the 1995, 1996 and 1997 data-taking periods. The integrated luminosity versus days of running, after the data quality monitoring package EVTAKA [98] has been applied, is shown in figure 4.3.

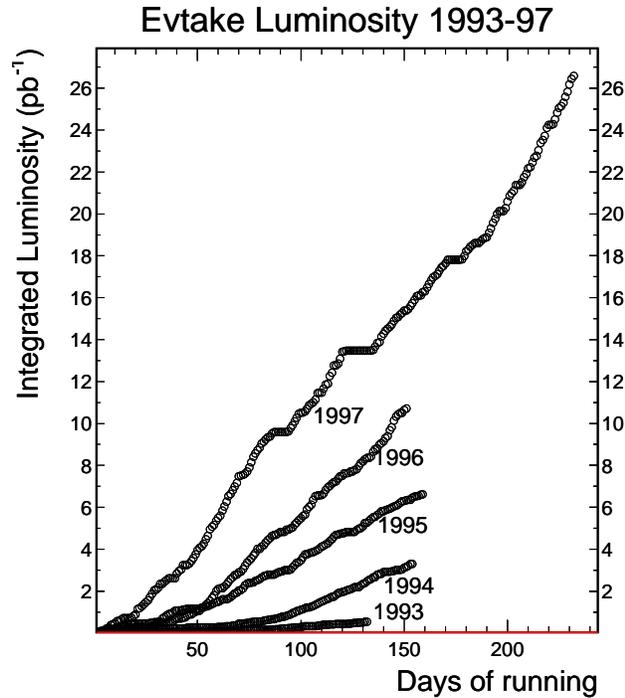


Figure 4.3: Integrated luminosity for high quality data versus day of running for all the positron-proton datasets. EVTAKЕ is part of the data quality monitoring procedure.

EVTAKЕ is a procedure which rejects data in which one or more of the major detector components were not operating normally. It makes the following specific requirements

- The luminosity monitor must be fully operational.
- The magnetic field must be on.
- The CTD must be at full high voltage, with no large dead regions.
- The calorimeter must be fully operational.

Table 4.1 gives the total luminosities available for each year after the application of EV-TAKЕ. The systematic errors on these luminosity measurements range from 1.1 % for 1995 data to 1.7 % for 1997 data [99]. During 1995 a small amount of data was taken with a shifted interaction point. 236.6 nb^{-1} were taken with the interaction point shifted towards the FCAL and 63.3 nb^{-1} were taken with the interaction point shifted towards the RCAL. These two datasets were excluded from the final data sample.

Year	1995	1996	1997
Luminosity after EVTAK (pb ⁻¹)	6.6	10.4	26.6
TOTAL after EVTAK	43.6 pb ⁻¹		

Table 4.1: A Summary of the good data available for analysis for 1995 —1997.

4.3.1 Noise Suppression

The majority of the kinematic variable reconstruction methods are heavily reliant on calorimeter information, indicating that the suppression of calorimeter noise is an important issue. In particular, events which contain a low level of hadronic activity can produce very distorted measurements of the hadronic quantities if the noise is left untreated.

As was discussed previously, noise can come from the natural radioactivity of the calorimeter or noise in the readout electronics. It can also arise from mini-sparks, where one of the pair of PMTs associated to a particular cell fires, but the other does not. Detailed studies have been made of these sources of noise [100, 101, 102] using a data sample made up of random trigger events which would not have passed any physics trigger and a Monte Carlo sample produced by generating events in which all particles pass down the beampipe.

Using these samples, it was found that the data contained a significant number of cells whose energy imbalance, $E_{L,R} = E_L - E_R$ (where E_L and E_R are the energies measured by the two PMTs associated with a given cell) is large. Such cells were not present in the Monte Carlo sample, as can be seen in figure 4.4 and were attributed to mini-sparks and can be removed by requiring $E_{L,R} < 0.49 \cdot E_{cell} + 0.03$ GeV, where $E_{cell} = E_L + E_R$.

Once the mini-sparks had been removed, it was found that the basic noise suppression energy cuts applied during reconstruction could be tightened to $E_{cell}^{EMC} > 80$ MeV and $E_{cell}^{HAC} > 140$ MeV if the cell was isolated. The remaining discrepancies then observed between data and Monte Carlo were attributed to the presence of “hot cells” in the data.

A “hot” cell is one which fires more frequently and measures a higher energy than the average for the calorimeter section in which it is situated. If the hot cell is known to have a malfunctioning PMT, then E_{cell} is set to twice the energy deposited in the working PMT,

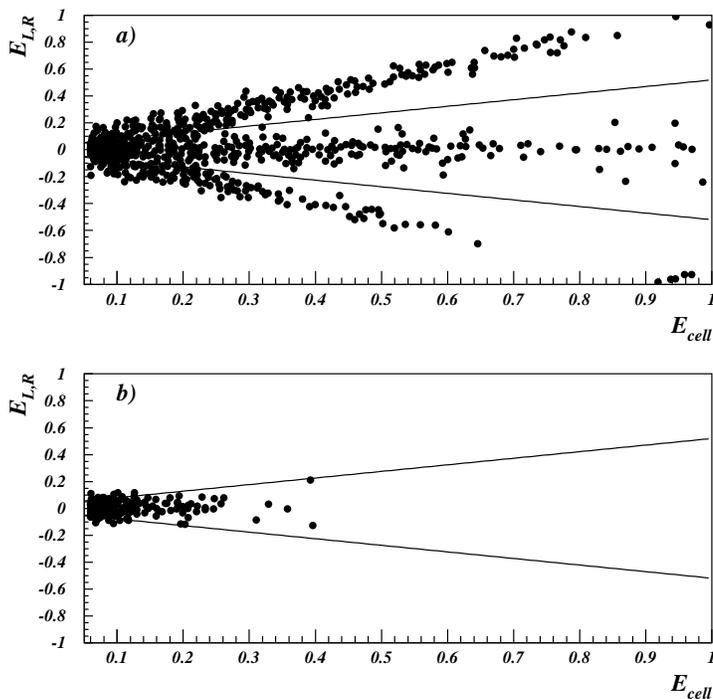


Figure 4.4: The Cell energy imbalance as a function of cell energy in (a) data, and (b) Monte Carlo.

providing $E_{L,R}$ is not zero. If, however, $E_{L,R} = 0$, then the cell is ignored if the energy it measures in a normal ep event is less than the maximum noise energy for that cell.

After all these measures have been applied, good agreement is observed between data and Monte Carlo.

4.3.2 Positron Identification

After the noise suppression algorithm described in the previous section has been applied, the data are searched for scattered positron candidates. This is clearly one of the most important aspects of any DIS analysis and hence will be described in some detail here. It is also important that both the positron's energy and scattering angle are reconstructed as accurately as possible, given the important role these quantities play in kinematic variable reconstruction.

Over a large area of the kinematic plane accessible at HERA the positron in the event will be well separated from the hadronic activity making it relatively easy to identify. However,

there are several problems which can make identification more difficult, such as showering caused by the positron passing through regions of high dead material prior to reaching the calorimeter, making the positron appear more like a hadron in terms of its energy deposit. In the high y region it is also found that the low energy positron can become mixed up in the hadronic activity. A further complicating factor comes from low energy hadrons and $\pi^0 \rightarrow \gamma\gamma$ decays in photoproduction events as they can leave a very similar energy deposit to that created by the true scattered positron. A sophisticated positron-finding algorithm is clearly required in order to overcome these problems.

The Positron Finding Algorithm

In both the physics studies described in subsequent chapters the neural net positron finder SINISTRA [90] is used to identify the scattered positron. First, electromagnetic islands are selected according to their shower characteristics. The 54 PMT energies of the 3×3 array of calorimeter towers centred on the highest energy tower of each selected island then form the input for the neural net. The neural net program returns a probability which will be close to 1 for the scattered positron and close to 0 for an island created by hadrons. The neural net was trained using Monte Carlo to provide a clean distinction between the scattered positron and hadronic energy deposits.

SINISTRA Efficiency and Purity

Studies have been performed in order to establish the efficiency of SINISTRA as a positron finder using elastic QED Compton event samples from both data and Monte Carlo [8]. Inefficiencies in SINISTRA can arise from the presence of dead material, which can cause the positron to preshower, lowering its energy and broadening its shower in the calorimeter. Other studies [92] have indicated that there is some discrepancy in the description of the dead material distribution in Monte Carlo. It is therefore important to study any effects this may have on the efficiency. Figure 4.5 shows the results of such a study performed using 1994 data, taken from [8]. The efficiency is acceptable above about 6 GeV and reasonable agreement is achieved between data and Monte Carlo.

More recent studies [103] performed using the higher statistics from 1996-97 data, in which

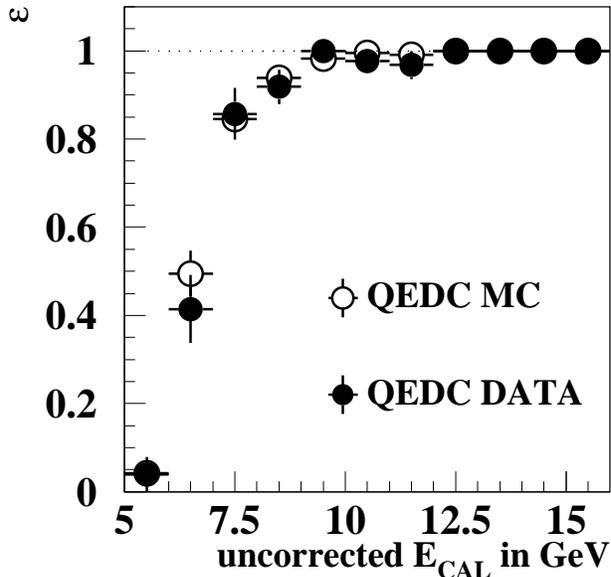


Figure 4.5: SINISTRA efficiency estimated from QED Compton data and Monte Carlo taken from [8].

the dead material was reduced compared to previous years², have shown similarly high efficiencies. The data and Monte Carlo are in good agreement above 10 GeV, while at lower energies a maximum of 5 % difference between data and Monte Carlo is observed.

Studies have also been performed using Monte Carlo to establish the purity of SINISTRA [5]. It has been shown that the purity is above 80 % for energies greater than 5 GeV.

4.3.3 Positron Position Reconstruction

The impact position of the positron on the calorimeter can be reconstructed either with the calorimeter itself or, if it lies within its geometrical acceptance, the position can be reconstructed using the Small Angle Rear Tracking Detector (SRTD). The SRTD is used in preference to the calorimeter as, on average, its spatial resolution is significantly better than that of the calorimeter (approximately 3 mm, as compared to ≈ 1 cm for the calorimeter). The positron impact position is initially calculated using the energies and positions of the cells assigned to the positron by the positron finder. Using this, the position is then tuned using either the SRTD or a more sophisticated calorimeter position

²Between the 1995 and 1996 data-taking periods, the cables from the old vertex detector were removed, substantially reducing the level of dead material in the RCAL beampipe region. This will be discussed further in the next chapter.

reconstruction algorithm when the SRTD is not available.

SRTD Position

The SRTD is described in chapter 5 and in [104], but a brief overview of the hit reconstruction methods and SRTD-CAL matching is given here.

First, clusters are formed in each SRTD plane separately, where strips are assigned to a cluster if there is a gap of no more than two “empty”³ strips between them. The clusters in the two planes are then matched by using the fact that the energy deposit in each plane should be roughly the same if caused by the same particle. Having formed the clusters, the energy is simply the sum of the energies of the all the associated strips. The position is determined by finding the strip for which $0.5E(X - 1) + E(X) + 0.5E(X + 1)$ is maximal, where X is the co-ordinate to be determined and E is the cluster energy in MIPS⁴. The 3 strips central to this shower maximum are then used to calculate the energy-weighted position of the cluster. A small correction is then applied to account for the natural bias towards the central strip caused by this type of procedure.

The cluster is then matched to the calorimeter energy deposit identified as the positron in the following manner: a linear extrapolation is made between the vertex position and the calorimeter position and if this lies within 1 cm of the SRTD a match is attempted. In the first attempt, only hits with more than 5 MIPS/plane are used. The closest hit (provided it lies within 15 cm of the calorimeter position) will be returned as the associated SRTD cluster. If nothing is found, then a second attempt is made with all hits, regardless of energy. If still nothing is found, then the calorimeter position must be used instead.

Calorimeter Position

In the RCAL, the x and y coordinates of the positron impact position are calculated separately by different methods, using the reconstruction algorithm ELECPO [105]. ELECPO uses the same 3×3 array of calorimeter towers as used by SINISTRA to determine x and y , as illustrated in figure 4.6. The y position is determined from strips s1, s2 and

³an empty strip is defined as one whose energy lies below the noise suppression threshold, which is currently set to 0.2 MIPS.

⁴Minimum ionising particle

s3, using a logarithmic energy-weighted average. The x position is calculated using the three cells in the central strip s2. Information from s1 and s3 is only used if the strips contain more than 25 % of the positron energy. The position within each individual cell is determined from the cell imbalance, and are then weighted by the cell energy and averaged to determine x . Once this x - y position has been determined, the RHES is checked for a similarly-positioned energy cluster. If one is found, then the HES position will be used in preference, due to its superior resolution. The z position is then calculated, taking into account the depth in the RCAL of the shower centre.

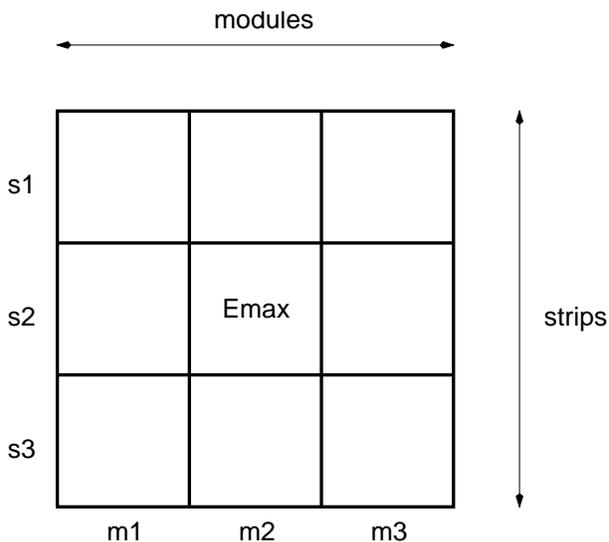


Figure 4.6: The layout of the 3×3 array of towers used by ELECPO to determine the position of the scattered positron. “Emax” indicates the highest energy cell and the strips and module labels are indicated.

If the positron is found in the BCAL, where no HES information is available, the impact position is determined using parametrisations from test beam results [105].

4.3.4 Positron Energy Loss Corrections

The effects of energy loss from the scattered positron when it is detected in the RCAL can be corrected for using either the SRTD or the rear presampler. The corrections based on SRTD information and their extraction are the subject of the next chapter. When it is not possible to use the SRTD correction, an alternative is offered by the presampler. This correction has been extracted from test beam data and is described in [106]. The

correction is simple and is the same for both data and Monte Carlo. It is given by

$$E_{cor} = E_{CAL} + 0.072 \cdot E_{pre} \quad (4.2)$$

where E_{cor} is the corrected energy, E_{cal} is the energy measured by the calorimeter and returned by the positron finder and E_{pre} is the energy of the associated cluster measured in the presampler. All the energies are in GeV, except E_{pre} which is measured in MIPs.

This correction only applies in the RCAL — no correction is available for the FCAL (there is no presampler for the BCAL). However, in the $D^{*\pm}$ analysis described in chapters 6 and 7, none of the events have a sufficiently high Q^2 for the positron to be identified anywhere but in the RCAL.

4.3.5 Calorimeter Energy Scale

Studies have been made of the energy scale in different calorimeter sections by matching tracks in the CTD to isolated energy clusters in the calorimeter. Both positrons and isolated hadronic deposits have been studied. In the studies performed on 1994 data [107, 108], it was found that there was a mismatch in the energy scale between data and Monte Carlo. More specifically, in the RCAL, the data energy scale was low by 2.5 ± 0.5 %, while in the BCAL it was found to be low by 6 ± 2 %. After these initial studies were made, an error was discovered in the Monte Carlo simulation package which resulted in the Monte Carlo energy scale being 1 % too high in the BCAL [109]. Thus a correction formula was implemented which scales BCAL energies up by 5 % and RCAL energies up by 2.5 % in the data [110]. The FCAL energy scale was found to agree well between data and Monte Carlo.

Since then, more detailed studies have been performed on the BCAL using the larger statistics available from 1996-7 data. These studies [111, 112] indicate that 2 % of the 1994 BCAL correction can be attributed to a variety of different sources. There is, however, still an unexplained mismatch of 3 ± 1 %. Similar studies made on RCAL data have attributed some of the mismatch to non-uniform energy response between cells, reducing the unexplained mismatch to 1 %. A cell-by-cell correction procedure for the RCAL has been determined from these studies.

For 1995 data, both in the SRTD corrections described in the next chapter and in the $D^{*\pm}$ analysis described in later chapters, the global factors, ie. 2.5 % for the RCAL and 5 %

for the BCAL, are applied to all cells except those associated with the scattered positron. The positron energy is then corrected using either SRTD or presampler information. If this information is not available, then the energy is scaled up by the appropriate factor.

For 1996 and 1997 data, the cell-by-cell energy scale corrections for the RCAL or the 5 % correction factor for the BCAL are applied to all cells, regardless of whether the cell is associated to the positron or not. No further correction is applied.

4.3.6 Hadronic Energy Flow Reconstruction

Historically within ZEUS event variables associated with hadronic activity have been reconstructed using only calorimeter information. More recently, however, a more comprehensive package has been developed [113] which combines calorimeter information with tracks from the CTD. As can be seen from figure 4.7, this ensures that all different “types” of particle are included in the reconstruction process with the highest possible accuracy. To ensure that the best possible quality data is used in this reconstruction method both noise suppression and calorimeter energy scale corrections must be used before the algorithm is applied. Full details of this package can be found in [113], but an overview is given here and is referred to as the EFO (energy flow object) method.

The first stage of the reconstruction process involves the formation of cell islands (see section 4.2.1) separately in the EMC, HAC1 and HAC2 sections of the calorimeter. Every calorimeter cell is assigned to a cell island. The position of the cell island is then determined from the logarithmically energy-weighted average of the positions of all the constituent cells. Having established a set of cell islands, these are then joined to form cone islands. This is achieved by matching cell islands in the hadronic and electromagnetic sections of the calorimeter together, based on their angular separation, starting in the outer hadronic sections and working inwards. EMC cell islands are then matched with each other using a similar process.

Once the cone islands have been formed, a set of high quality tracks must then also be selected for matching purposes. A track is considered to be of high enough quality if it has traversed at least three CTD superlayers and has $0.1 < p_T < 15$ GeV. These tracks are then extrapolated to the face of the calorimeter and matched to a cone island, where possible. The energy and momentum of these matched objects are then calculated using

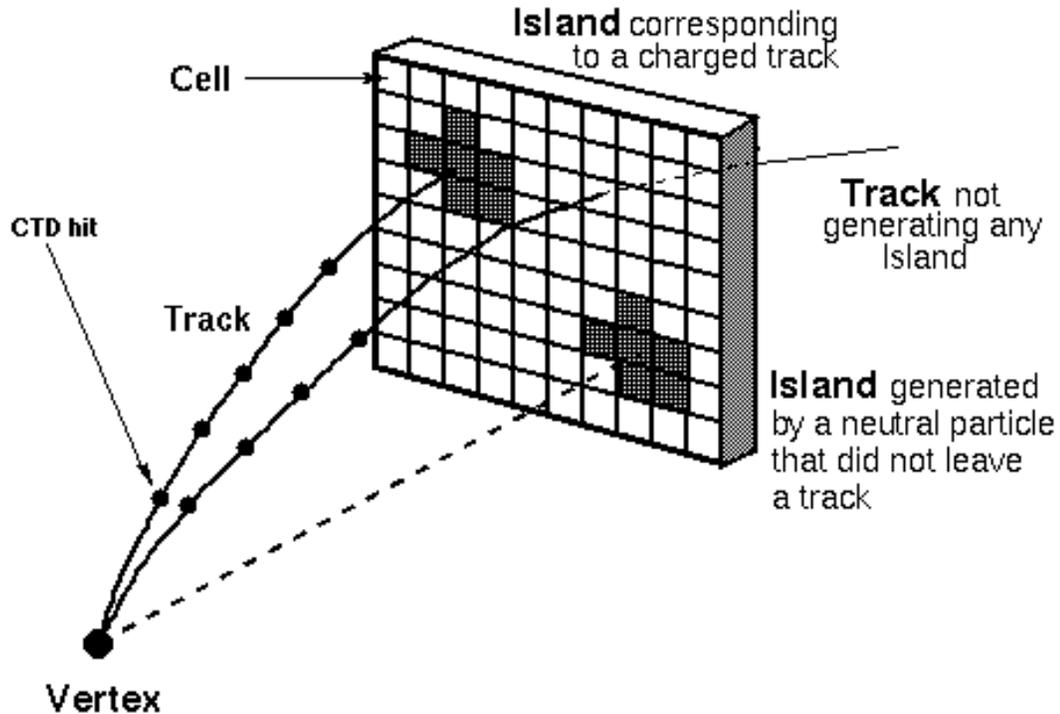


Figure 4.7: The different types of energy flow objects which can be formed using a combination of calorimeter and CTD information.

the following rules

1. If a track has not been matched to a cone island, then its energy is calculated using the momentum determined from the track under the assumption that the particle which produced the track was a pion.
2. If a cone island has not been matched to a track, then its energy and momentum are calculated by assigning a momentum vector \mathbf{p} to it such that $E^2 = \mathbf{p}^2$, where E is the energy measured by the calorimeter and \mathbf{p} has the direction from the reconstructed vertex to the cone island position.
3. If a cone island has more than three tracks associated to it, then its energy and momentum are calculated using the calorimeter information alone as in 2. above.

All other objects are then assessed to decide whether to use the calorimeter or CTD information associated to the object. The tracking information will be used in preference to the calorimeter information if

- $E/p < 0.9 + \Delta(E/p)$, where E is the calorimeter energy in GeV and p is the momentum measured by the CTD. $\Delta(E/p)$ is the 1σ error on E/p .
- and the resolution on the momentum measurement from the CTD is better than the resolution on the energy measurement from the calorimeter.

The first requirement ensures that the calorimeter energy is due only to the associated track ie. there are no extra neutral particles involved. The second requirement ensures that the highest possible accuracy is achieved. The objects which have been reconstructed are referred to as ZEUS Unidentified Flow Objects, or ZUFOS.

4.4 Kinematic Variable Reconstruction

The reconstruction of the kinematic variables x , y and Q^2 is one of the most important aspects of any analysis of DIS data. There are three different methods in use within ZEUS, all of which are described here.

Figure 4.8 depicts a standard ep collision and indicates which quantities can be measured, any two of which can be used to reconstruct the kinematic variables. Naturally F and γ_h , the energy and scattering angle of the struck quark, cannot be directly measured, but can be well approximated by reconstructing the energy and angle of the resulting hadronic system. This will be discussed further in section 4.4.2.

Figure 4.9 shows isolines of the four quantities, E'_e , θ_e , F and γ_h , over the $x - Q^2$ plane. Where the isolines are close together, the quantity which the isolines represent will be relatively insensitive to any mismeasurement. An example of this is shown in figure 4.9(a) for low scattered positron energies. Conversely, at high positron energies, where the isolines are much further apart (low y), a small mismeasurement of E'_e can cause serious distortion in the kinematic variables. This will be discussed further in the next section.

4.4.1 The Electron Method

The simplest reconstruction method available is based purely on E'_e and θ , the energy and polar angle of the scattered positron, and is referred to as the electron method. It is also the reconstruction method normally used by fixed target experiments. The formulae for

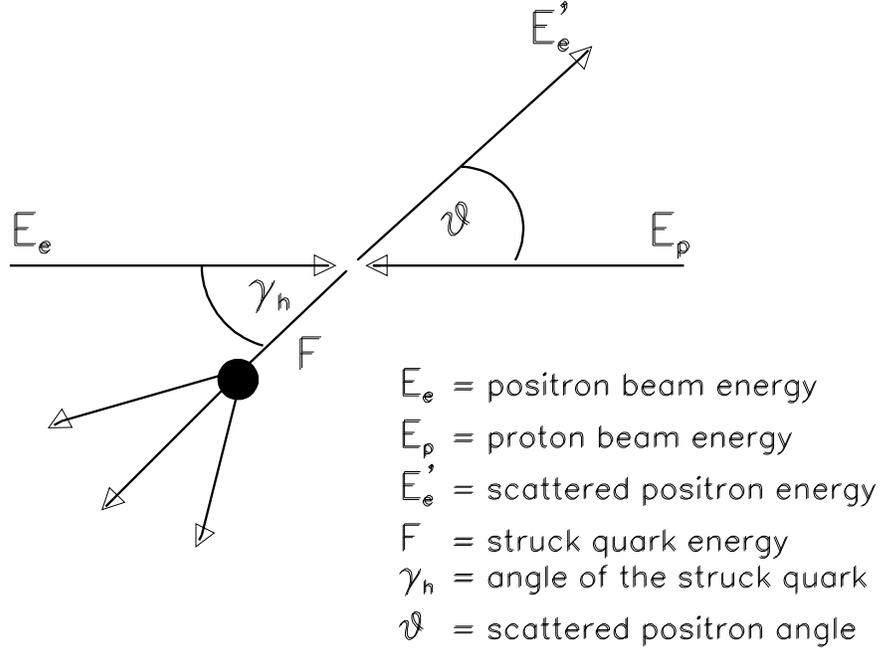


Figure 4.8: Schematic representing a standard ep collision, indicating the main quantities which can be used to reconstruct the event kinematic variables x , y and Q^2 .

the kinematic variables, x_e , y_e and Q_e^2 can be derived from equation 1.1 and are found to be

$$\begin{aligned}
 Q_e^2 &= 2E_e E'_e (1 + \cos \theta) \\
 y_e &= 1 - \frac{E'_e}{2E_e} (1 - \cos \theta) \\
 x_e &= \frac{Q_e^2}{s y_e}
 \end{aligned} \tag{4.3}$$

where E_e , E'_e and θ are as defined in figure 4.8 and s is the square of the centre of mass energy.

As can be seen from figure 4.9, the electron method will perform well at high y , indicated by the closeness of the isolines in E'_e and θ is this region. However, at low y the isolines are far apart, indicating that the electron method will not perform so well. This is reflected in figure 4.10 which shows the difference between the reconstructed and true kinematic variables obtained from a DIS Monte Carlo sample, using the electron method. Events in which the positron radiates a photon in the initial or final state require special treatment. This will be discussed in section 4.4.4. It is clear that x_{el} is systematically shifted from the true value and this is due to the poor performance of the electron method at low y and Q^2 — the kinematic region which dominates in the $D^{*\pm}$ analysis presented in chapters 6

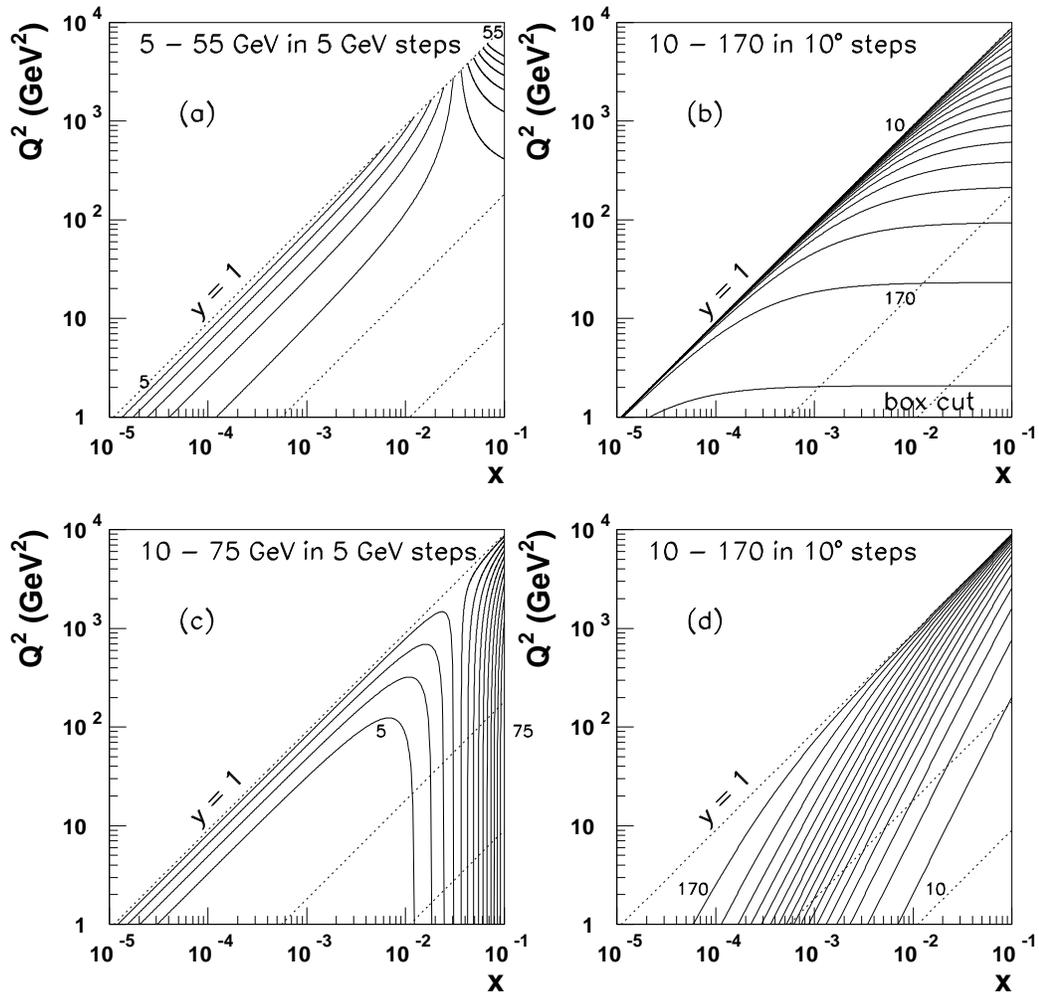


Figure 4.9: Isolines in the x - Q^2 plane of (a) positron energy, (b) positron angle, (c) struck quark energy and (d) struck quark angle.

and 7.

The electron method is also hampered by the difficulties in measuring the positron energy accurately. These arise from energy loss in dead material and also from calorimeter energy scale uncertainties. For these reasons, the electron method is not used in the $D^{*\pm}$ analysis presented in this thesis.

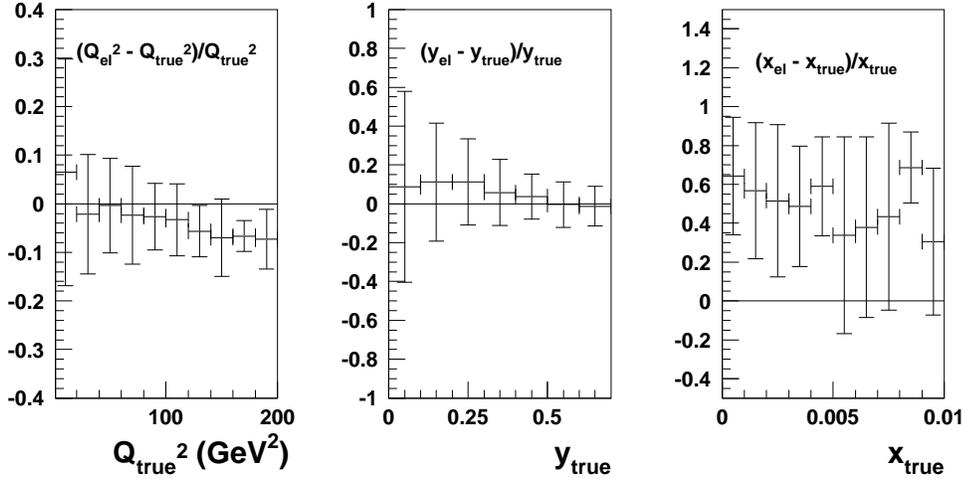


Figure 4.10: The fractional difference between the reconstructed and true values of Q^2 , y and x , using the electron method, from a sample of DIS Monte Carlo as a function of the true variable. The error bars are the RMS widths of the fractional differences.

4.4.2 The Jacquet-Blondel Method

Due to the more comprehensive coverage of the ZEUS and H1 detectors (compared to fixed target experiments), it is possible to fully reconstruct the struck quark energy and angle, F and γ_h . Thus these quantities can be used to reconstruct the kinematic variables. This method was suggested by Jacquet and Blondel [114] with charged current events in mind. They derived the formulae:

$$\begin{aligned}
 y_{JB} &= \frac{\sum_h (E^h - p_z^h)}{2E_e} \\
 Q_{JB}^2 &= \frac{(\sum_h p_x^h)^2 + (\sum_h p_y^h)^2}{1 - y_{JB}} \\
 x_{JB} &= \frac{Q_{JB}^2}{s y_{JB}}
 \end{aligned} \tag{4.4}$$

where p_x^h , p_y^h and p_z^h are the momentum components in the x , y and z directions of the hadrons and where \sum_h indicates the sum over all hadrons in the event, except those associated with the proton remnant. The hadronic quantities required in this method can be calculated using either calorimeter cells or ZUFOS. The latter are used for this purpose in the $D^{*\pm}$ analysis. This method is also sensitive to the effects of dead material and calorimeter energy scale uncertainties. It is also particularly susceptible to calorimeter noise when the level of hadronic activity is low (low y events). This results in poor resolution as can be seen in figure 4.11. Thus, this reconstruction method is not used in the $D^{*\pm}$ analysis, although y_{JB} can be used to exclude events with a low level of hadronic activity, in which γ_h would be most susceptible to distortion caused by calorimeter noise.

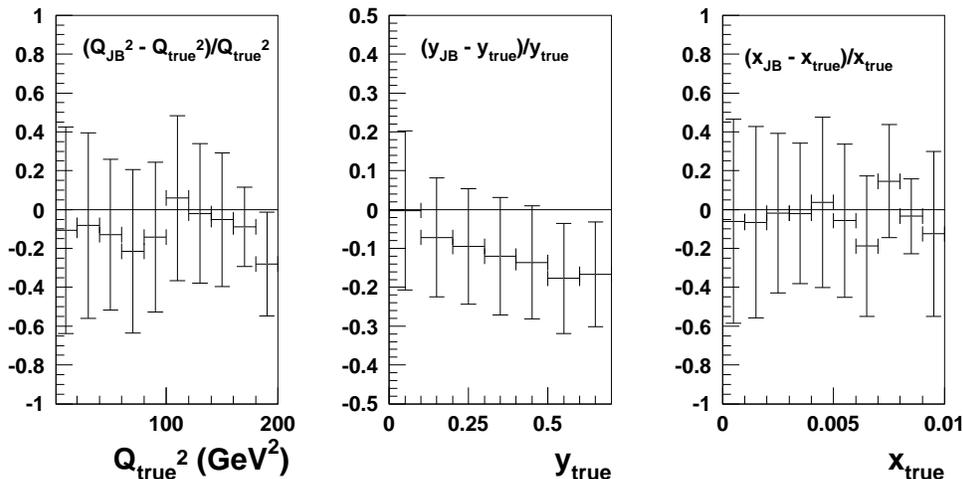


Figure 4.11: The fractional difference between the reconstructed and true values of Q^2 , y and x using the Jacquet-Blondel method from a sample of DIS Monte Carlo as a function of the true variables. The error bars are the RMS widths of the fractional difference distributions.

4.4.3 The Double Angle Method

This final method [115] relies, as the name implies, on the two scattering angles, θ and γ_h , where γ_h is determined using

$$\cos \gamma_h = \frac{Q_{JB}^2(1 - y_{JB}) - 4E_e^2 y_{JB}^2}{Q_{JB}^2(1 - y_{JB}) + 4E_e^2 y_{JB}^2} \quad (4.5)$$

The scattering angles tend to be much more accurately reconstructed than the energies and are much less sensitive to calorimeter energy scale considerations. It is clear from figure 4.9, that the only region where the Double Angle method will not perform well is at low y — a region which can be removed using a cut on y_{JB} , as described in the previous section.

The formulae for the Double Angle method kinematic variables can be derived [115] using the naive quark parton model as

$$\begin{aligned}
 Q_{DA}^2 &= 4E_e^2 \frac{\sin \gamma_h (1 + \cos \theta)}{\sin \theta + \sin \gamma_h - \sin(\theta + \gamma_h)} \\
 y_{DA} &= \frac{\sin \theta (1 - \cos \gamma_h)}{\sin \theta + \sin \gamma_h - \sin(\theta + \gamma_h)} \\
 x_{DA} &= \frac{E_e \sin \theta + \sin \gamma_h + \sin(\theta + \gamma_h)}{E_p \sin \theta + \sin \gamma_h - \sin(\theta + \gamma_h)}
 \end{aligned} \tag{4.6}$$

From figure 4.12 it is clear that the Double Angle method does not suffer from large systematic shifts and also has a better resolution (RMS width) than the Jacquet-Blondel method and is consequently used as the reconstruction method for the $D^{*\pm}$ analysis described in chapters 6 and 7.

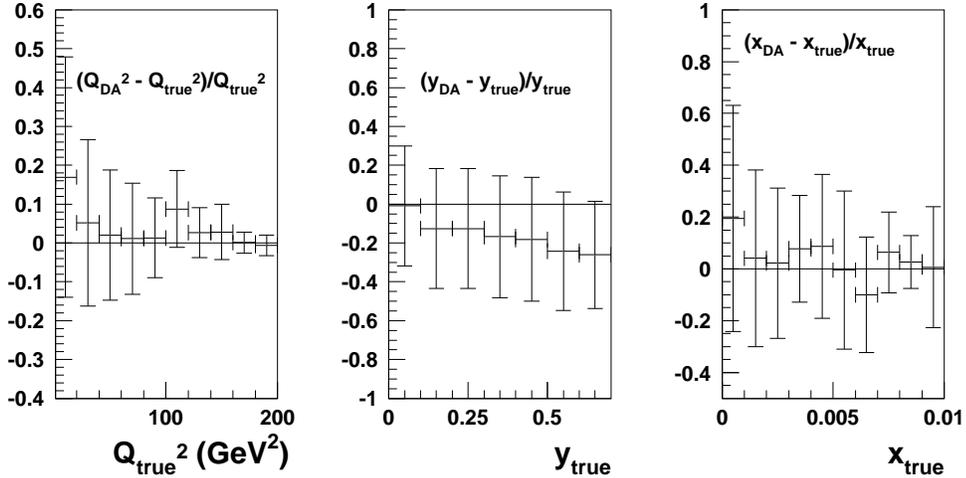


Figure 4.12: The fractional difference between the reconstructed and true values of Q^2 , y and x using the Double Angle method, from a sample of DIS Monte Carlo as a function of the true variables. The error bars are the RMS widths of the fractional difference distributions.

4.4.4 Radiative Events

The kinematic variable reconstruction methods which have been described in the previous sections do not account for the possibility that the incoming or outgoing positron may radiate a photon. This will cause the kinematics of the event to be incorrectly reconstructed. As the photons tend to be radiated close to the direction of the positron, the effect of final state radiation (FSR) is not serious. This is because the calorimeter deposits of the positron and the photon are sufficiently close together that they are identified as a single cluster.

Initial state radiation (ISR), however, poses a more serious problem as the photon usually escapes undetected down the beampipe. This means that the positron beam energy, E_e has been reduced to $E_e - E_\gamma$. This effect must be accounted for during the unfolding procedure used in the extraction of cross sections from measured data, as will be discussed in chapter 7.

4.5 Summary

In this chapter we have discussed the procedure through which data from the detector are filtered online to reject background and the selection requirements for DIS interactions. The main features of the reconstruction package have also been discussed, with emphasis placed on the detector components most relevant to the physics studies presented in this thesis. The post-reconstruction analysis tools required to produce the final high quality data sample have been described and the available kinematic variable reconstruction methods have been assessed. The Monte Carlo simulation of the detector and trigger systems have been described.

Chapter 5

SRTD Energy Corrections

5.1 Introduction

At the end of the previous chapter the different methods available for kinematic variable reconstruction were discussed. It was explained that the simplest and potentially most accurate method is the so-called “electron” method, in which x , y and Q^2 can be reconstructed from the angle and energy of the scattered positron. This method, however, can only reach its full potential if both the angle and the energy are accurately measured. Of these two quantities, the energy is the most difficult to measure well because it is the most sensitive to the different degrading factors (showering caused by dead material and calorimeter energy scale uncertainties) discussed in the previous chapter. Figure 5.1 shows the energy spectra of positrons detected in the small angle rear tracking detector (SRTD) region for 1995 data and Monte Carlo. The peak of the distribution is expected to be at approximately the positron beam energy (27.5 GeV). However, it is clear that this is not the case, both for data and Monte Carlo, and indicates that some energy loss has occurred. The noticeable discrepancy between data and Monte Carlo indicates that the scattered positron energy loss prior to detection in the calorimeter is not well simulated in the Monte Carlo.

The mismatch between data and Monte Carlo observed in the scattered positron energy distribution could be due to missing dead material between the interaction point and the calorimeter in Monte Carlo. It could also be due to differences in the energy scale between data and Monte Carlo (as discussed briefly in the previous chapter), or some combination

of both effects. Previous analyses of low Q^2 data [116] have shown that information from

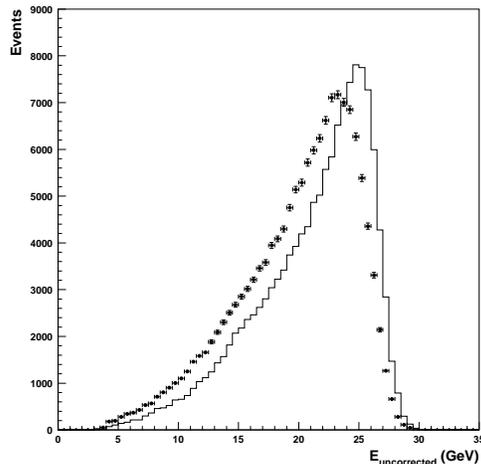


Figure 5.1: Scattered positron energy distributions for 1995 data (black dots) and Monte Carlo (histogram) where the positron is detected in the SRTD region. Standard DIS cuts have been applied to obtain the distributions. Energy loss by the scattered positron is indicated by the shift of the peak to a value which is lower than the nominal positron beam energy (see text). There is also a noticeable discrepancy between data and Monte Carlo, indicating that the energy loss is inaccurately simulated in Monte Carlo.

the SRTD can be used to correct for energy loss from the scattered positron, improving the resolution on the measurement of the positron energy and ensuring good agreement between data and Monte Carlo.

After the end of data-taking in 1994, the RCAL module which contains the beampipe was modified so that the beam hole was no longer 20×20 cm, but $20(x) \times 8(y)$ cm, allowing the detection of lower Q^2 positrons than was previously possible. This, however, also changed the dead material configuration in the rear beampipe region, requiring the determination of a new set of energy loss corrections for data and Monte Carlo using the SRTD. There was also a sharp increase in the amount of data taken between 1994 and 1995, making it possible for the first time to extract a detailed functional form for these corrections.

This chapter describes the extraction of these energy loss corrections using the SRTD. First an overview of the SRTD itself will be given, followed by a description of the data and Monte Carlo samples which are used for this study and how they are selected. The methods available for extracting the corrections will be discussed and their results compared to find

the best set of energy loss corrections for 1995 data. Finally, comparisons will be made between data and Monte Carlo after correction to ensure that the objectives of this study have been achieved.

5.2 The Small Angle Rear Tracking Detector

The SRTD is situated on the inner face of the RCAL (in front of the RPRES — see chapter 3) and covers the whole of the inner ring (plus some of the second ring) of calorimeter towers around the RCAL beampipe, as shown in figure 5.2. It's total angular coverage is $162^\circ < \theta < 176^\circ$.

The SRTD consists of two layers made up of strips of scintillator; the first layer consists of horizontal strips, while the second layer consists of vertical strips. The strips are 10 mm wide, 5 mm thick and either 240 mm or 440 mm long, depending on which layer they are in. Each layer is divided up into four quadrants which, as with the RPRES, are read out using wavelength shifters which pass the scintillation light to the PMTs. The SRTD is also calibrated strip-by-strip using MIP-like particles.

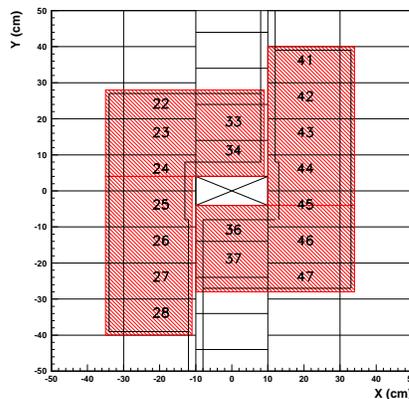


Figure 5.2: Layout for the SRTD as in the 1995 Data Taking Period. The grey hatched areas indicate the four SRTD quadrants, which are numbered 1 to 4 starting in the top right-hand corner and going round anti-clockwise. The RCAL cells are also indicated and are numbered for use in the comparison of data and Monte Carlo kinematic peak events, as performed in section 5.6.3.

5.3 Selection of Data Samples

In order to study scattered positron energy loss, event types are required in which the energy of the scattered positron can be determined independently of the calorimeter. Three such event types have been used in this study: Kinematic peak events, DIS ρ^0 events and elastic QED Compton events. Separate samples of each of these event types are required for both data and Monte Carlo. The selection criteria for data and Monte Carlo and the method of “predicting” the energy of the scattered positron are described in this section.

All the analysis tools described in the previous chapter are used in the selection of all the different event types. However, at this stage, ZUFOS had not been developed and consequently, calorimeter cell positions and energies were used in their place. There are two specific criteria which are not previously discussed, but which are common to all the data and Monte Carlo samples in this study. These are

SRTD hit All the scattered positrons in these event samples were required to be associated to a good hit in the SRTD (in the case of the QED Comptons, at least one of the clusters was required to hit the SRTD). A good SRTD hit is normally defined as one found by the reconstruction routine, SRTDELEC, for which the error flag is returned as zero. When error = 0, it means that the SRTD cluster to which the scattered positron is associated is well-contained within the SRTD and has sufficiently high energy to ensure that the cluster was reconstructed properly. When error > 0, it means that the highest energy strip within the cluster was on the edge of the SRTD or on the calorimeter crack (the line along which the two halves of the calorimeter meet). It can also mean that the energy deposit in each of the two SRTD planes was too small to ensure a good match when the cluster was made. Details of the construction of SRTD clusters and the matching procedure to the scattered positron are given in the previous chapter. The “ambiguous” hits make up about 20 % of the data sample. Both of these hit types are kept in this sample as it is not known how well simulated the ambiguous hits are in Monte Carlo. To ensure that the hits are well-contained, fiducial cuts are applied such that the hit position is at least 1 cm from the SRTD edge.

Calorimeter Crack Any scattered positrons which are found in the cracks where the two

halves of the RCAL/SRTD meet are excluded from all samples.

The remainder of the cuts are specific to the event type and are applied to both data and Monte Carlo, unless otherwise stated. They are described in detail below.

5.3.1 Kinematic Peak Events

From the $x - Q^2$ plane shown in figure 5.3, it can be seen that for low y and low Q^2 ($Q^2 < 100 \text{ GeV}^2$) the energy of the scattered positron becomes approximately independent of the kinematics of the event and is roughly equal to the beam energy (in 1995 this was 27.52 GeV). This region of phase space is known as the kinematic peak region.

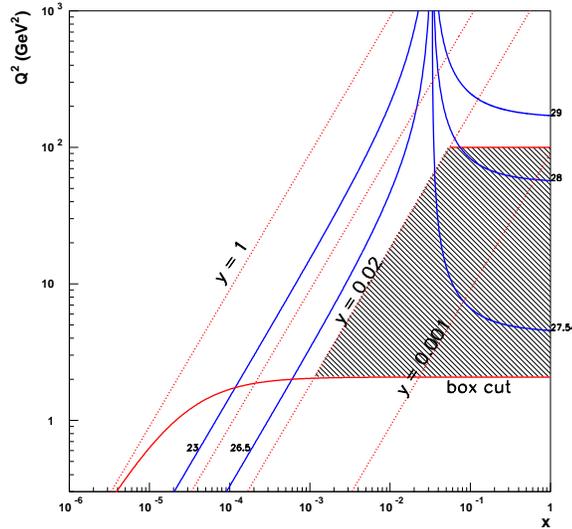


Figure 5.3: Isolines of Scattered Positron Energy (GeV) as seen in the $x - Q^2$ plane.

The data which form the kinematic peak sample correspond to roughly half the data taken in 1995 in which the low Q^2 TLT trigger was switched on and unrescaled. The triggers are discussed in more detail in chapters 4 and 6. The corresponding Monte Carlo sample was generated using DJANGO6V2.4, which interfaces LEPTO [117] and HERACLES [118] in order to produce non-diffractive DIS events with radiative corrections. Parton shower evolution was performed using MEPS [119], while the hadronisation was performed using the Lund string model [120] as implemented in JETSET [121]. The input proton structure function was GRV94(HO), with F_L set to zero. This latter requirement only applies to this particular Monte Carlo sample and will not significantly alter the results of this study.

The samples of kinematic peak data and Monte Carlo are selected using a combination of requirements which are as follows

- One of the FLT DIS trigger slots has fired.
- TLT low Q^2 bit set. This ensures good coverage of the SRTD.
- A positron with energy $E'_e > 8$ GeV, found using the SINISTRA positron finder (see chapter 4 for more details).
- A fiducial cut around the RCAL beam hole was applied by requiring the positron to hit the calorimeter outside a $26(x) \times 16(y)$ cm box, centred on the middle of the RCAL beam hole. Due to the choice of $x = 0$, $y = 0$, the box actually runs from -13 cm to $+13$ cm in x and from -8 cm to $+8$ cm in y . Thus it is normally written $13(x) \times 8(y)$ cm and is referred to as the “box cut”.
- A good hit in the SRTD.
- $35 < E - P_z < 65$ GeV.
- $-40 < Z_{vertex} < +40$ cm.
- $Q_{DA}^2 > 3.5$ GeV². This lower Q^2 limit was chosen because the Monte Carlo sample used was generated with $Q^2 > 3$ GeV².

The kinematic variables were reconstructed using the Jacquet-Blondel and Double Angle methods in order to minimise the input from the scattered positron energy. The significance of all these cuts is discussed fully in the context of the $D^{*\pm}$ analysis presented in subsequent chapters.

Having established a sample of “good” DIS events, the kinematic peak sample was selected by requiring $y_{JB} < 0.02$ and $Q_{DA}^2 < 100$ GeV². This leaves 32368 data events and 18484 Monte Carlo events for the kinematic peak samples.

5.3.2 DIS ρ^0 Events

In elastic ρ^0 events, the proton remains intact and scatters through a small angle such that it escapes through the forward beampipe. The ρ^0 itself decays almost 100 % of the time to a pair of charged pions. Thus in exclusive DIS ρ^0 events, the energy of the scattered

positron can be calculated from the conservation of energy and longitudinal momentum, under the assumption that the transverse momentum of the scattered proton is small compared to the longitudinal component. The energy is given by the following formula:

$$E'_e = \frac{2E_e - (E_\rho - |p_\rho| \cos \theta_\rho)}{1 - \cos \theta_e} \quad (5.1)$$

where E_e is the positron beam energy, E_ρ , p_ρ and θ_ρ are the energy, momentum and angle of the ρ^0 (as determined from the tracks of the two pions) and θ_e is the angle of the scattered positron. In the case of initial state radiation, E_e becomes $E_e - E_{ISR}$.

The data which form the DIS ρ^0 sample were taken from the full 1995 data-taking period. The Monte Carlo DIS ρ^0 sample was generated using HERACLES [118], with the requirement that $Q^2 > 1 \text{ GeV}^2$. After noise suppression and hadronic energy scaling the following further cuts were then applied to both data and Monte Carlo:

- A scattered positron with energy greater than 5 GeV.
- $E - P_z > 35 \text{ GeV}$.
- $-50 < Z_{vertex} < +40 \text{ cm}$.
- box cut $13(x) \times 8(y) \text{ cm}$
- Two oppositely-charged tracks were required which were associated to the primary vertex and which were also required to have at least one hit in superlayer 3. The two tracks were only accepted as pion candidates if they also each had $p_T > 0.16 \text{ GeV}$. In the area of angular coverage of the SRTD, a track associated to the scattered positron is not expected.
- $E_{CAL}^\rho / P_\rho < 1.5$, where E_{CAL}^ρ is the calorimeter energy excluding that deposited by the scattered positron and P_ρ is the momentum of the ρ^0 candidate as measured in the CTD.
- $0.6 < M_{\pi^+\pi^-} < 1.0 \text{ GeV}^2$.

The last three cuts on this list are those needed to select $\rho^0 \rightarrow \pi^+\pi^-$ decays in elastic events. The nominal ρ^0 mass is $770.0 \pm 0.8 \text{ MeV}$ [18], hence the cut on the $\pi^+\pi^-$ invariant mass. The width of the mass window reflects the resolution with which the mass can

be reconstructed. The E_{CAL}^{ρ}/P_{ρ} requirement ensures that the only hadronic energy deposited in the calorimeter comes from the ρ^0 . This excludes ρ^0 events in which the proton dissociates.

This is a minor adaption of the cuts used in previous analyses [122]. After applying all these cuts, 2200 data events and 2806 Monte Carlo events were left.

5.3.3 Elastic QED Compton Events

The Feynman diagrams for the elastic QED Compton process are shown in figure 5.4. In such events, the incoming positron scatters off a quasi-real photon ($q_2^2 \rightarrow 0$), producing a positron and a photon in the final state, each with a finite scattering angle. Hence, both the positron and the photon can be measured in the main detector. The energy transfer to the proton is small, allowing the proton to remain intact and disappear undetected down the beampipe.

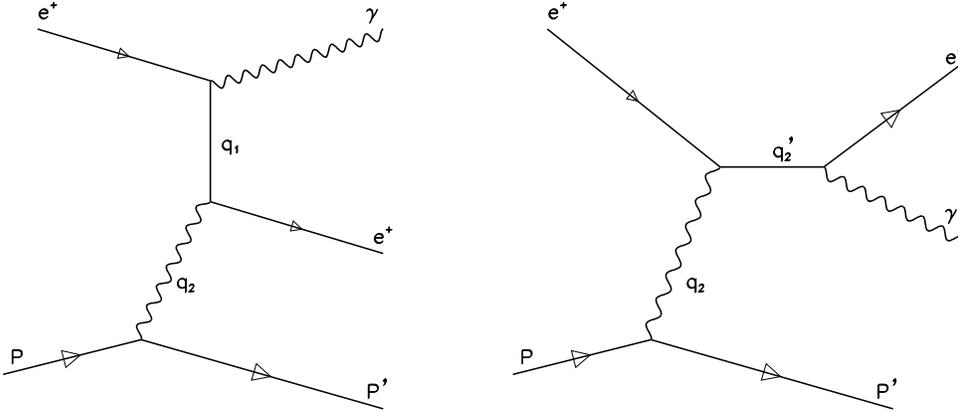


Figure 5.4: The Feynman diagrams for the elastic QED Compton process.

Given the kinematics of these events, it is possible to determine the positron and final state photon energies independently of the calorimeter using conservation of transverse energy and momentum, under the assumption that the transverse momentum of the scattered proton is negligible. Using these requirements the positron and photon energies are given by

$$E_e' = 2E_e \frac{\sin\theta_\gamma}{\sin\theta_e + \sin\theta_\gamma - \sin(\theta_e + \theta_\gamma)} \quad (5.2)$$

$$E_\gamma = 2E_e \frac{\sin\theta_e}{\sin\theta_e + \sin\theta_\gamma - \sin(\theta_e + \theta_\gamma)} \quad (5.3)$$

where E_e is the positron beam energy, θ_e is the polar angle of the scattered positron and θ_γ is the polar angle of the final state photon. Again, in the case of ISR events, E_e should be replaced with $E_e - E_{ISR}$.

The QED Compton data sample was selected from the whole 1995 data-taking period using the following combination of DST requirements

- Any one of the FLT DIS triggers or the FLT QED Compton trigger should have fired. The former are discussed in the previous chapter, while the latter looks specifically for at least two isolated electromagnetic deposits in the calorimeter, with energy threshold cuts set according to where the deposits are located [123].
- $E - P_z + 2E_\gamma > 30$ GeV.
- Exactly two electromagnetic islands ($E_{EMC}/E_{total} > 0.9$) must be found with $\eta_{island} < 2.5$.
- Each island should have $E_{island} > 2$ GeV and at least one island should have $E_{island} > 4$ GeV.
- $E_{FCALBP} < 50$ GeV, where E_{FCALBP} is the energy deposited in the inner rings of the FCAL.
- No hadronic islands should be found. An island is considered hadronic if it satisfies $E_{EMC}/E_{total} < 0.9$.
- The islands should be back-to-back in ϕ , ie. $|\pi - |\phi_1 - \phi_2|| < \pi/4$.

The sample of elastic QED Compton Monte Carlo was generated using Compton2.0 [124]. After noise suppression and energy scaling, the following further cuts were then applied to both data and Monte Carlo:

- Two electromagnetic clusters were required, each with energy greater than 5 GeV.
- Both clusters were required to lie outside a box cut of $13(x) \times 8(y)$ cm.

- $35 < E - P_z < 60$ GeV.
- $-50 < Z_{vertex} < +50$ cm.
- $|\theta_e - \theta_\gamma| < 85^\circ$. This removes events in which hadrons were identified as electromagnetic clusters.
- $E_{had} = E_{tot} - E'_e - E_\gamma$ must be less than 2 GeV.
- $|\psi| < 3.4^\circ$, where $\psi = \pi - |\phi_e - \phi_\gamma|$. ϕ_e and ϕ_γ are the azimuthal angles of the scattered positron and final state photon respectively (calculated from the SRTD co-ordinates where possible).

The last two cuts are those required to select elastic (as opposed to inelastic) QED Compton events. The vertex requirement should be viewed in the knowledge that the vertex is reconstructed successfully in only a small number of the events. Where the vertex is not found, it is assumed to be zero. These cuts are a minor adaptation of those used previously [125, 126].

After applying all these cuts, 5134 data events and 19472 Monte Carlo events remain for use in this study.

5.4 Methods of Energy Prediction

As described in the previous section, each of the event types provides an alternative method for predicting the energy of the scattered positron to direct measurement using the calorimeter. In order to confirm that these methods were reliable, studies were carried out on each of the Monte Carlo samples. The results of these are summarised below.

5.4.1 Kinematic Peak Sample

The y and Q^2 cuts on this sample select out a small spread in scattered positron energies around the beam energy (see figure 5.5). In order to determine the peak value, the Monte Carlo sample must be used. After cutting on y_{JB} and Q_{DA}^2 , the “true” scattered positron energy distribution from the Monte Carlo generator level can be extracted. To find the peak value of this “true” kinematic peak sample, the energy must be smeared by $24\%\sqrt{E}$

in order to simulate the measured calorimeter resolution [127]. Then the peak of this smeared distribution can be fitted with a Gaussian to find the “predicted” energy. By fitting the peak shown in figure 5.5, the predicted energy for the kinematic peak samples is found to be 27.18 ± 0.02 GeV. The resolution of this method can also be obtained from this Gaussian fit, and is found to be 1.30 ± 0.04 GeV. This resolution is then also assumed to be correct for the data kinematic peak sample.

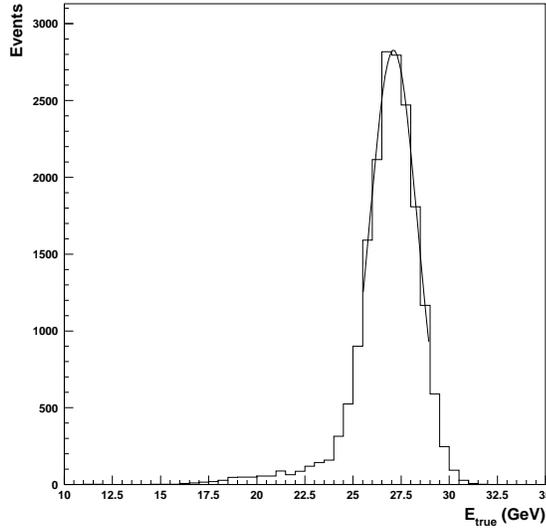


Figure 5.5: The “True” Kinematic Peak energy distribution, taken from Monte Carlo. The line indicates a fit to the peak to obtain the predicted energy for this sample.

5.4.2 DIS ρ^0 Sample

As described in section 5.3.2, the energy of the scattered positron can be determined from the kinematics of the ρ^0 . When this energy was calculated from the reconstructed ρ^0 in the Monte Carlo and the value compared to the generator-level scattered positron energy (“true” positron energy), the agreement was not as good as expected. In order to investigate whether the prediction method itself was at fault, the energy was calculated using equation 5.1 and the four-vector of the generator level ρ^0 and compared to the true positron energy. The two main factors which were investigated were the assumption of negligible proton p_T and the consequences of neglecting to subtract any initial state photon energy from the beam energy. The results of these investigations are shown in figure 5.6. All the quantities used to make these plots come from the generator-level Monte Carlo,

ensuring that there are no detector effects contributing to any discrepancies. It is clear that the assumption of negligible proton p_T is perfectly reasonable, but that failing to deal with ISR events can have a significant effect. This is to be expected as the energy predicted from the conservation of longitudinal energy and momentum, which implies a small contribution from the proton, but a large one from any initial state radiation.

Going then to the reconstructed predicted energy, a method of dealing with ISR events was sought. Using the LUMI photon calorimeter (LUMI- γ) to reject possible ISR events was tested, but due to the limited geometrical acceptance of the LUMI- γ , too few photons were tagged to remove the problem. The method which was finally used was based on the assumption that the ρ^0 and the scattered positron should balance in transverse momentum, and that any imbalance was due to the effect of energy loss from the scattered positron. Thus the ratio of the transverse momenta could be used as a rough correction factor for the positron $E - p_z$. This then implies that any tail in this corrected $E - p_z$ distribution is due to ISR rather than energy loss from the scattered positron. Thus a relatively harsh cut could then be placed on this corrected $E - p_z$ (referred to as $(E - p_z)_{corr}$) to reduce the ISR tail. A cut of $(E - p_z)_{corr} > 50$ GeV was chosen after examining figures 5.7b and c. Figure 5.7b shows the number of ISR (non-ISR) events which survive the $(E - p_z)_{corr}$ cut as a fraction of the total number of events which survive the cut (referred to as fraction 1), versus cut value, while figure 5.7c shows the number of ISR (non-ISR) events which survive the $(E - p_z)_{corr}$ cut as a fraction of the total number of events in the starting sample (referred to as fraction 2), versus cut value. It is clear from this latter figure that $(E - p_z)_{corr} > 50$ GeV is a reasonable cut, as it removes ~ 50 % of the ISR events, but only removes ~ 20 % of the non-radiative events. After applying this cut 1630 data events and 1849 Monte Carlo events remain for this study.

Having obtained these final samples, it was necessary to determine the resolution of the prediction method. To do this, $E_{predicted} - E_{true}$, where E_{true} is the true positron energy from the Monte Carlo generator level, was plotted in bins of predicted energy. As can be seen from figure 5.8a, the distributions of the absolute difference are not Gaussian, and so the RMS of each distribution was taken as the resolution. This RMS was then plotted as a function of the mean predicted energy and fitted with a polynomial to obtain a parametrisation of the resolution. However, when this parametrised form of the resolution was used in the extraction of the corrections for both data and Monte Carlo (the extraction is described in section 5.6), it was found that the contribution from the DIS ρ^0 events to

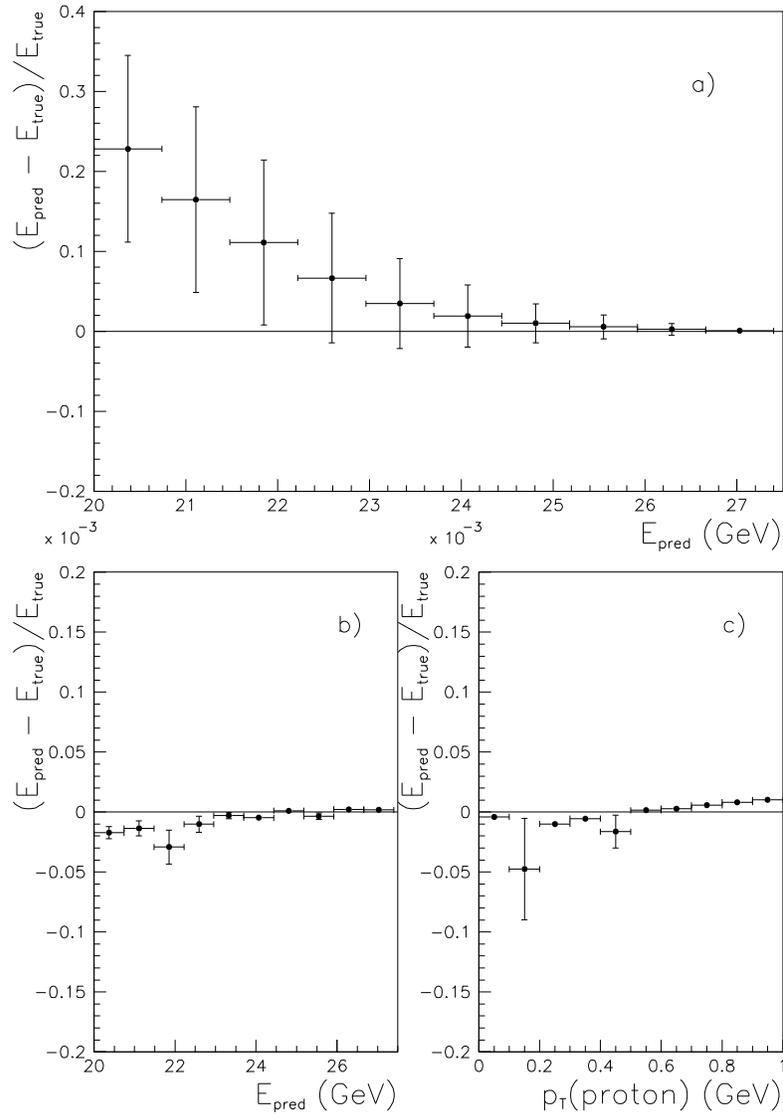


Figure 5.6: a) Comparison of the predicted energy and the “true” energy from the Monte Carlo generator level, without taking into account any initial state radiation in the prediction method, b) the same comparison, but taking into account initial state radiation and c) comparison of the predicted and true energies as a function of proton p_T .

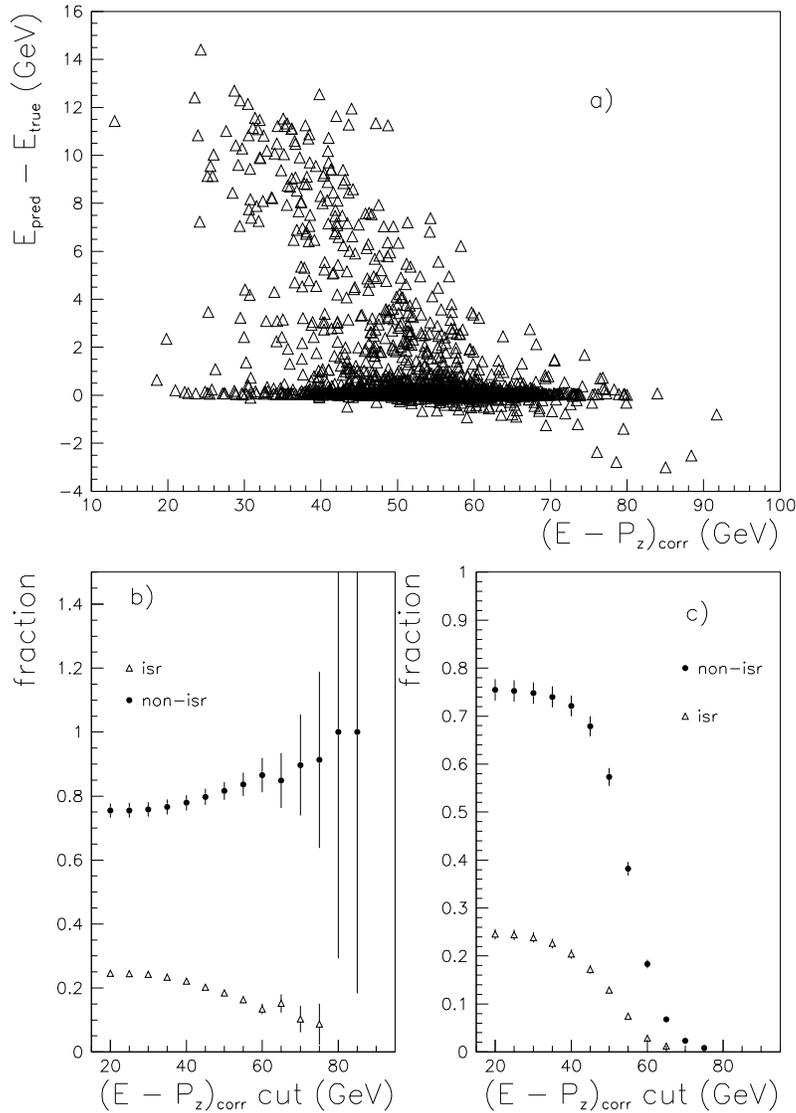


Figure 5.7: a) Comparison of $E_{\text{pred}} - E_{\text{true}}$ as a function of the p_T -corrected $E - P_z$. The tail to low values of $E - P_z$ is predominantly due to ISR events. b) Fraction 1 (as described in the main text) as a function of $(E - P_z)_{\text{corr}}$ cut. c) Fraction 2 (also described in the main text) as a function of $(E - P_z)_{\text{corr}}$ cut.

the χ^2 in the second method was very high. This was traced to the resolution being extremely small at the high end of the DIS ρ^0 predicted energy spectrum. These events in which the scattered positron is almost at beam energy correspond to the lowest y events. This also means that the tracks from the ρ^0 will be at the forward limit of the CTD acceptance. This in turn will cause systematic effects on the track measurements, which will deteriorate the resolution for these events. For the purposes of this study, it was assumed that the highest resolution which could be attained in this sample was ~ 200 MeV. The lowest resolution obtained from this event type is ~ 400 MeV, as can be seen at low energies in figure 5.8. Thus a linear interpolation between 400 MeV at 19 GeV and 200 MeV at 27.3 GeV was made, which gives the resolution of the DIS ρ^0 events as $\sigma = 0.86 - 0.024 * E_{predicted}$. This value was then, once more, assumed to also be the case of the data DIS ρ^0 sample.

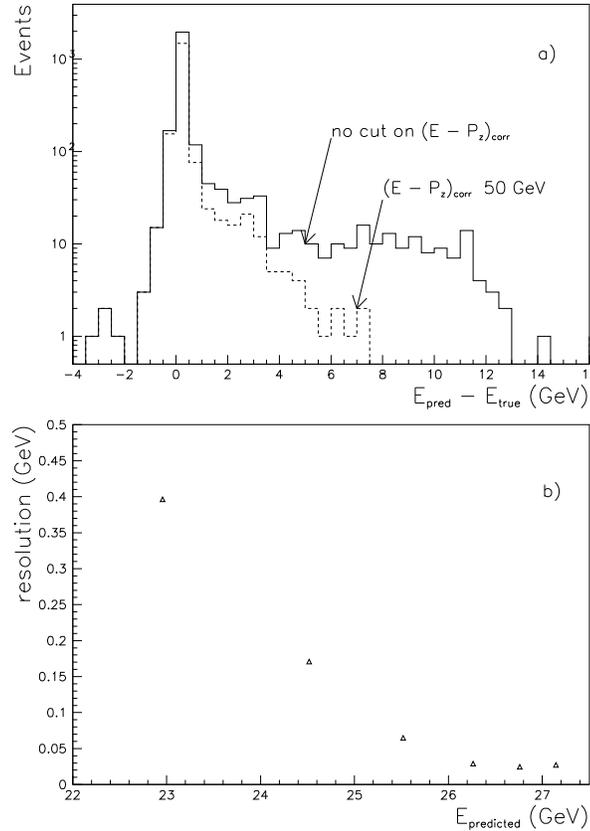


Figure 5.8: a) The effect of requiring $(E - P_z)_{corr} > 50$ GeV on the absolute difference between the predicted and true energies. b) The resolution of the ρ^0 energy prediction method as a function of the predicted energy.

5.4.3 Elastic QED Compton Sample

As described in section 5.3.3, the energy of both the positron and the photon can be predicted from transverse momentum balance, assuming the proton p_T to be small compared to its longitudinal component. As a consequence, it is reasonable to expect that the effect of initial state radiation on this prediction method should be less significant than for the DIS ρ^0 sample as ISR photons have negligible p_T . However, it is equally reasonable to expect that non-zero proton p_T will have a more significant effect on this sample than on the DIS ρ^0 sample. In order to test these expectations, the generator-level four-vectors of the positron and photon were used to predict the energies, which were then compared to the actual generator-level values. The results of these investigations are shown in figure 5.9. There is clearly some contribution from both ISR and proton p_T , but by applying the cuts described in the section 5.3.3 and increasing the $E - P_z$ cut from 35 to 38 GeV, the agreement at reconstruction level (as shown in figure 5.10a) is good between $6 < E_{predicted} < 23$ GeV. The candidates outside this range are the highly asymmetric QED Comptons, which are clearly much more sensitive to the effects of proton p_T and ISR and so have been excluded from this study.

The resolution for this prediction method was determined in the same manner as that for the DIS ρ^0 sample. The resolution as a function of predicted energy can be seen in figure 5.10b and can be parametrised as $\sigma = -0.3 + 0.145 * E_{pred} - 0.005 * E_{pred}^2 + (3.4 * 10^{-5}) * E_{pred}^3$.

5.5 Dead Material Distributions

Information from the SRTD can be used to look at the distribution of dead material between the interaction point and the face of the RCAL. Positrons tend to shower when they pass through a layer of dead material, and the larger the amount of dead material, the greater the amount of energy the positron loses prior to hitting the calorimeter. The distribution of SRTD hits for different ranges in SRTD energy deposit can be examined in order to build up a picture of the dead material distribution in data and Monte Carlo. The density of hits in any region for a particular energy range indicate the level of dead material. For example, the x - y distribution of hits with low SRTD energies has the highest density of hits in the regions of little dead material. Conversely, the same distribution, but for hits with high SRTD energies will have the highest density of hits in regions of

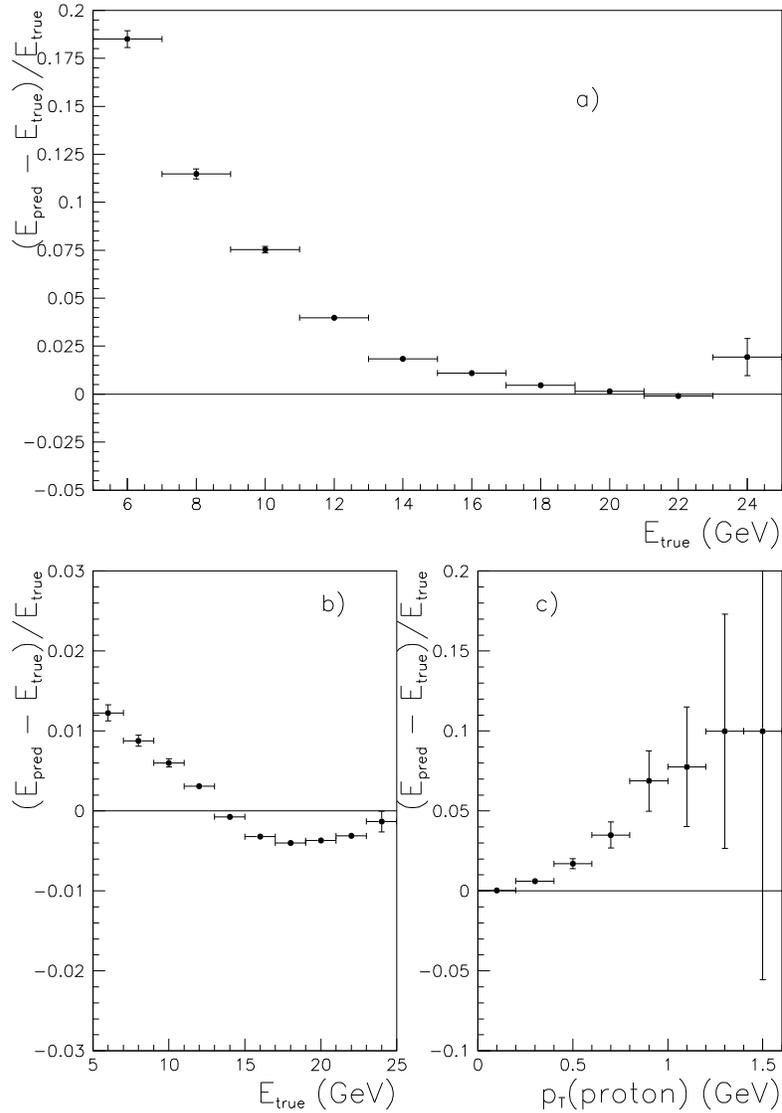


Figure 5.9: a) Comparison of the predicted energy and the “true” energy from the Monte Carlo generator level, without taking into account any initial state radiation and under the assumption that proton p_T is small compared to p_L . b) The same comparison, taking into account ISR. c) Comparison of the predicted and “true” energies as a function of proton p_T .

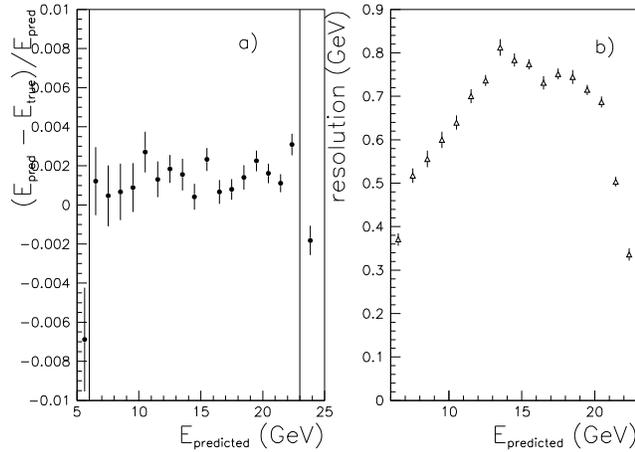


Figure 5.10: a) Comparison of the reconstructed predicted energy and the generator level energy as a function of the predicted energy. The agreement is good within the range $6 < E_{\text{pred}} < 23$ GeV (as shown by the vertical lines). b) The resolution of the elastic QED Compton energy prediction method as a function of predicted energy. This is fitted to determine the parametrisation of the resolution used in the correction determination.

large dead material. Figures 5.11(a)-(d) shows such hit distributions for data and Monte Carlo.

These distributions clearly show the ring structures which were seen in previous studies of energy loss by the scattered positron using the SRTD [92]. These are attributed to the VXD cables and the CTD inner wall. There are also areas of high hit density at the corners of the box cut. These are caused by the presence of water cooling pipes. A comparison between data and Monte Carlo of the mean SRTD energy deposit as a function of radius is also shown in figure 5.11e). This is essentially an indication of the amount of dead material as a function of radius. Through examination of the plots in figure 5.11, it can be seen that there is clearly insufficient dead material in Monte Carlo, but that the dead material that is already implemented in the Monte Carlo is distributed in a similar manner to that in the data.

5.6 Positron Energy Correction Results

Two different methods have been used to determine the functional forms and parameter values of these corrections, which are described below.

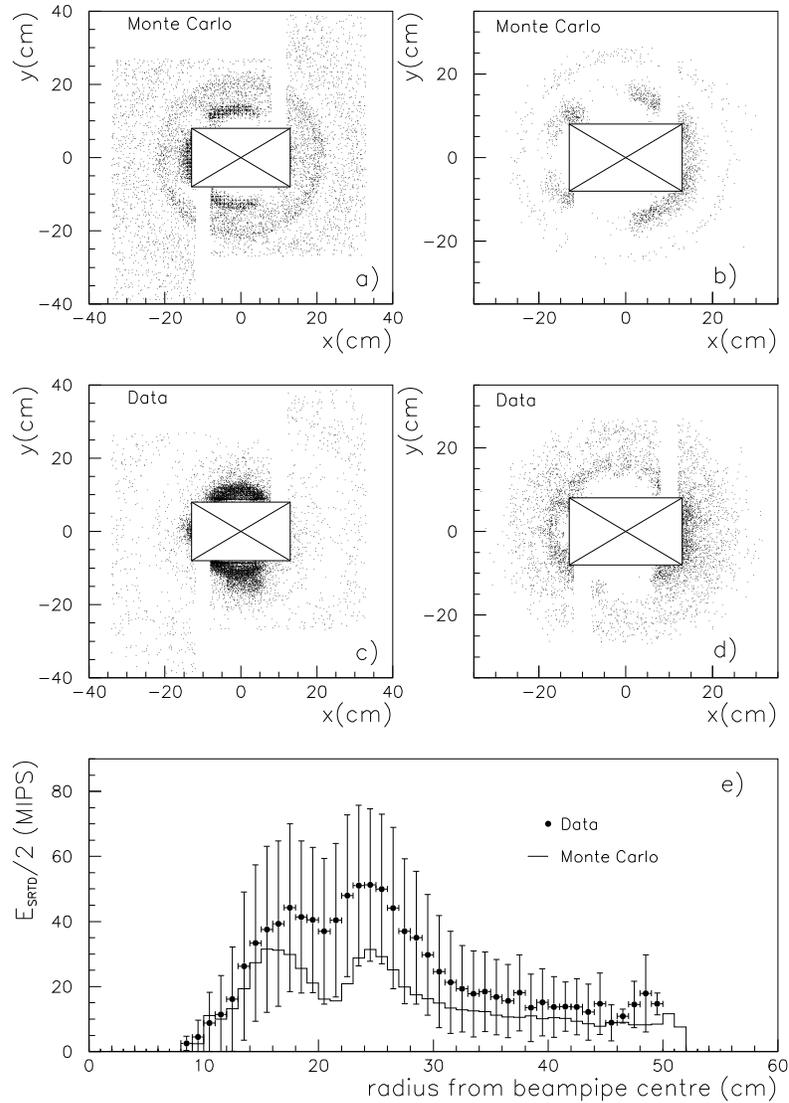


Figure 5.11: a) The SRTD hit distribution in the Monte Carlo for hits with $E_{SRTD}/2 < 10$ MIPS. b) The hit distribution for the same sample, but for hits with $E_{SRTD}/2 > 60$ MIPS. c) The SRTD hit distribution in the data kinematic peak sample for hits with $E_{SRTD}/2 < 10$ MIPS. d) The hit distribution for the same sample, but for hits with $E_{SRTD}/2 > 60$ MIPS. e) A Comparison of $E_{SRTD}/2$ versus radius for data and Monte Carlo.

Method I

In this method, for each data and Monte Carlo sample the calorimeter energy is binned in SRTD energy and the peak of the calorimeter energy is then fitted with a Gaussian. This peak value is plotted as a function of the mean SRTD energy in each bin. This distribution is then fitted to extract both the functional form and the parameter values. These values can then be examined as a function of corrected scattered positron calorimeter energy to determine any energy dependence.

Method II

This method is an event-by-event χ^2 minimisation technique, where the χ^2 is defined as

$$\chi^2 = \frac{(E_{predicted} - E_{corrected}(E_{CAL}, E_{SRTD}))^2}{\sigma^2} \quad (5.4)$$

σ is defined as the resolution of the prediction method (as described in section 5.4) and the functional form of the correction is assumed from Method I. The χ^2 is then minimised to find the best set of correction parameters. Although the functional form from Method I is assumed, this is an independent way of determining the parameter values required.

Both these methods have their advantages: by fitting only the peak of each calorimeter energy distribution in Method I, the problematic ISR events are excluded from the study. However, the results of the method will be dominated by the sample with the largest statistics, which in this case is the kinematic peak sample. Unfortunately, this also has the poorest resolution. Conversely, the DIS ρ^0 sample has the best resolution, but the lowest statistics. This method is also very subjective, as it depends on the user choice of what the “best fit” is.

Method II has the advantage that the sample with the best resolution will dominate the result, but ISR events have to be treated correctly to make the method work. Results from both these methods are given below and compared.

5.6.1 Results from Method I

The functional form of the correction is obtained from the kinematic peak samples, because these samples have the largest statistics. The kinematic peak can only be used to give one set of correction parameters at a single predicted energy. Although the DIS ρ^0 sample covers a range of predicted energies (20 GeV to 27.4 GeV), the statistics are only large

enough to provide one set of correction parameters at $\langle E_{predicted} \rangle \sim 25.6$ GeV. The elastic QED Compton samples cover a range in predicted energy from 6 to 23 GeV (as explained in section 5.4), but these samples contain sufficient statistics to provide four sets of correction parameters. The bins in which each set of correction parameters have been determined (in both data and Monte Carlo) are given in table 5.1.

The functional form of the correction in data is found to be linear, while the form of the correction for the Monte Carlo was found to be well described, at least for the kinematic peak and DIS ρ^0 samples, by a second order polynomial. The results of these fits can be seen in figures 5.12 and 5.13. The slope parameter for each of the data fits are given in figure 5.12. For the Monte Carlo, the samples and their linear and quadratic parameter values are given in table 5.1.

Sample	Linear term	Second order term
QED Compton $6 < E_{pred} < 12$ GeV	0.045 ± 0.008	0.00010 ± 0.00016
QED Compton $11 < E_{pred} < 14$ GeV	0.053 ± 0.007	0.00016 ± 0.00015
QED Compton $14 < E_{pred} < 16$ GeV	0.048 ± 0.006	0.00020 ± 0.00011
QED Compton $16 < E_{pred} < 23$ GeV	0.051 ± 0.007	0.00012 ± 0.00012
DIS ρ	0.031 ± 0.007	0.0004 ± 0.0001
Kinematic Peak	0.039 ± 0.002	0.00031 ± 0.00003

Table 5.1: The parameters of the fit to each of the Monte Carlo Samples.

In the Monte Carlo samples, it is clear that fitting a quadratic function gives a good description of the E_{CAL} versus $E_{SRTD}/2$ distributions for the kinematic peak and DIS ρ^0 samples. However, for the elastic QED Comptons, it is not so clear whether the fit should be linear or quadratic. This ambiguity can be explained by the fact that the statistics are limited in each fit, but also by the fact that the QED Compton predicted energies

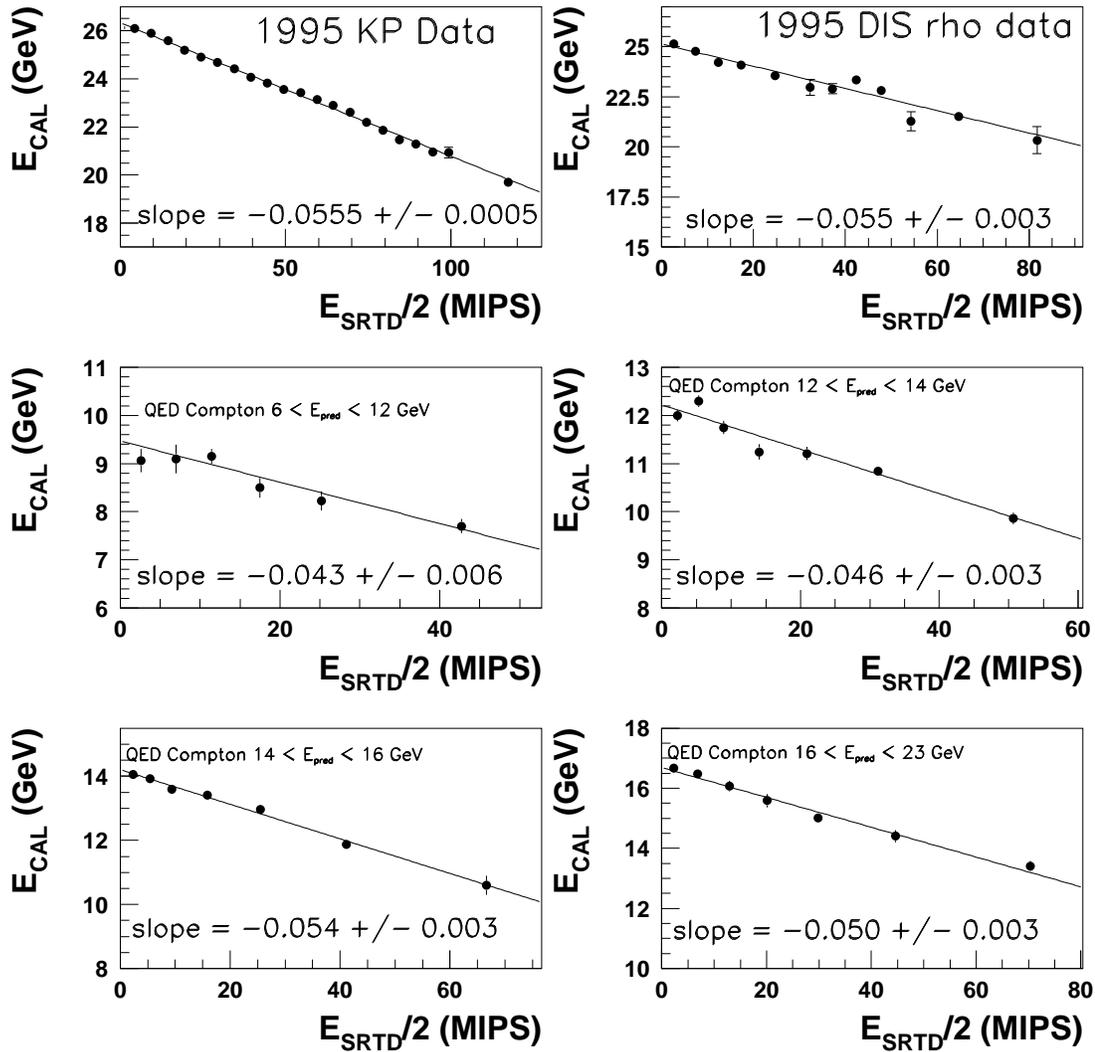


Figure 5.12: Uncorrected positron energy, E_{CAL} , versus SRTD energy (in MIPS) for each of the three data samples. The elastic QED Compton sample has been divided up into four bins in predicted positron energy to gain wider positron energy coverage. The behaviour is linear in all bins and the number given on each plot is the slope from a linear fit.

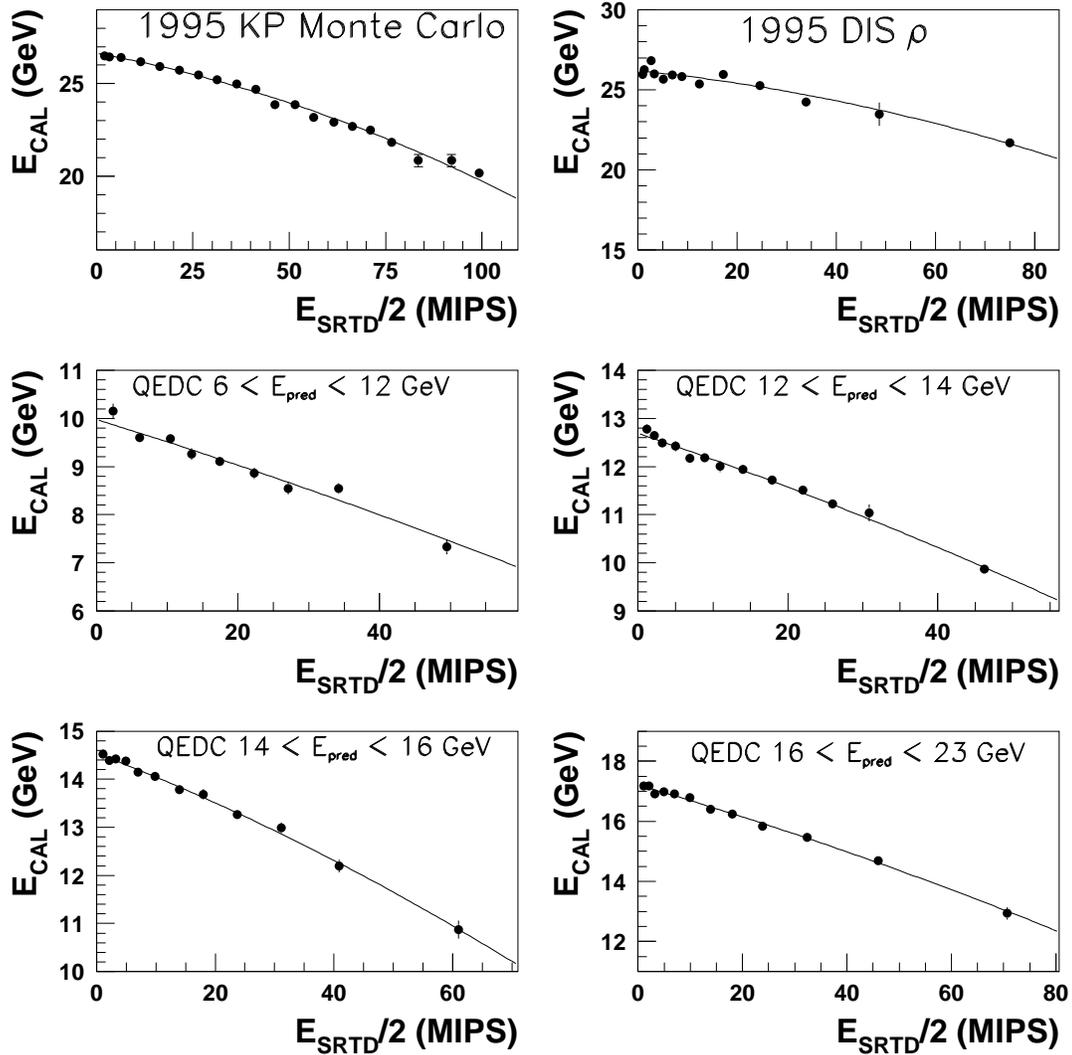


Figure 5.13: Uncorrected positron energy, E_{CAL} , versus SRTD energy for each of the three Monte Carlo samples. The elastic QED Compton sample has been divided into four bins in predicted positron energy. The curves on the plots are the results of fitting a second order polynomial and the fit parameters are summarised in table 5.1.

are substantially lower than for the DIS ρ^0 or kinematic peak samples. This will mean that the SRTD energy deposits for the QED Comptons will be naturally lower. Whether the better fit is linear or quadratic is due to the behaviour at the highest SRTD energy deposits, which will make up a smaller fraction of the QED Compton sample relative to the kinematic peak and DIS ρ^0 samples. Another possible solution to this ambiguity could be due to the correction parameters themselves being dependent on the calorimeter energy of the scattered positron. When the fits shown in figures 5.13 and 5.12 are performed, the measured calorimeter energy at $E_{SRTD}/2 = 0$ MIPS (ie. the y-intercept of the fit) is allowed to vary as one of the free parameters of the fit function. This value can be interpreted as the mean energy loss-corrected scattered positron energy for each sample. Thus by plotting, for example, the slope parameter for each sample versus this mean energy, this dependence can also be investigated for both data and Monte Carlo. The results of these investigations are shown in figure 5.14. It is clear from examining figure 5.14 that statistics are too limited to give a clear indication of any energy dependence. In order to use information from all the samples available, an error-weighted mean is taken for each parameter, indicated by the flat lines in figure 5.14. Thus the functional forms of the final corrections using Method I and the corresponding parameter values are given as

$$E_{corr} = E_{CAL} + m \cdot E_{SRTD}/2 \quad (5.5)$$

where $m = 0.0550 \pm 0.0005$ GeV/MIP for data and

$$E_{corr} = E_{CAL} + a \cdot E_{SRTD}/2 + b \cdot (E_{SRTD}/2)^2 \quad (5.6)$$

where $a = 0.0404 \pm 0.0016$ GeV/MIP and $b = (2.9 \pm 0.2) \times 10^{-4}$ GeV/MIP².

However, if these corrections are applied to standard samples of low Q^2 NC data and Monte Carlo and the corrected scattered positron energy distributions compared, it is clear that the Monte Carlo distribution is shifted to higher values than the data distribution (see figure 5.15). This is principally due to the technical method by which the samples are selected. The calorimeter energy scale corrections described in the previous chapter are only applied to cells not associated to the scattered positron. Thus a scaling factor must be determined to account for this after the energy loss correction has been applied. In the results from both Method I and II, the scaling factors for data and Monte Carlo show a relative difference which is approximately the 2.5 % expected for RCAL cells.

This scaling factor is obtained independently for data and Monte Carlo by comparing the dead-material corrected scattered positron energy with its predicted energy. Thus the

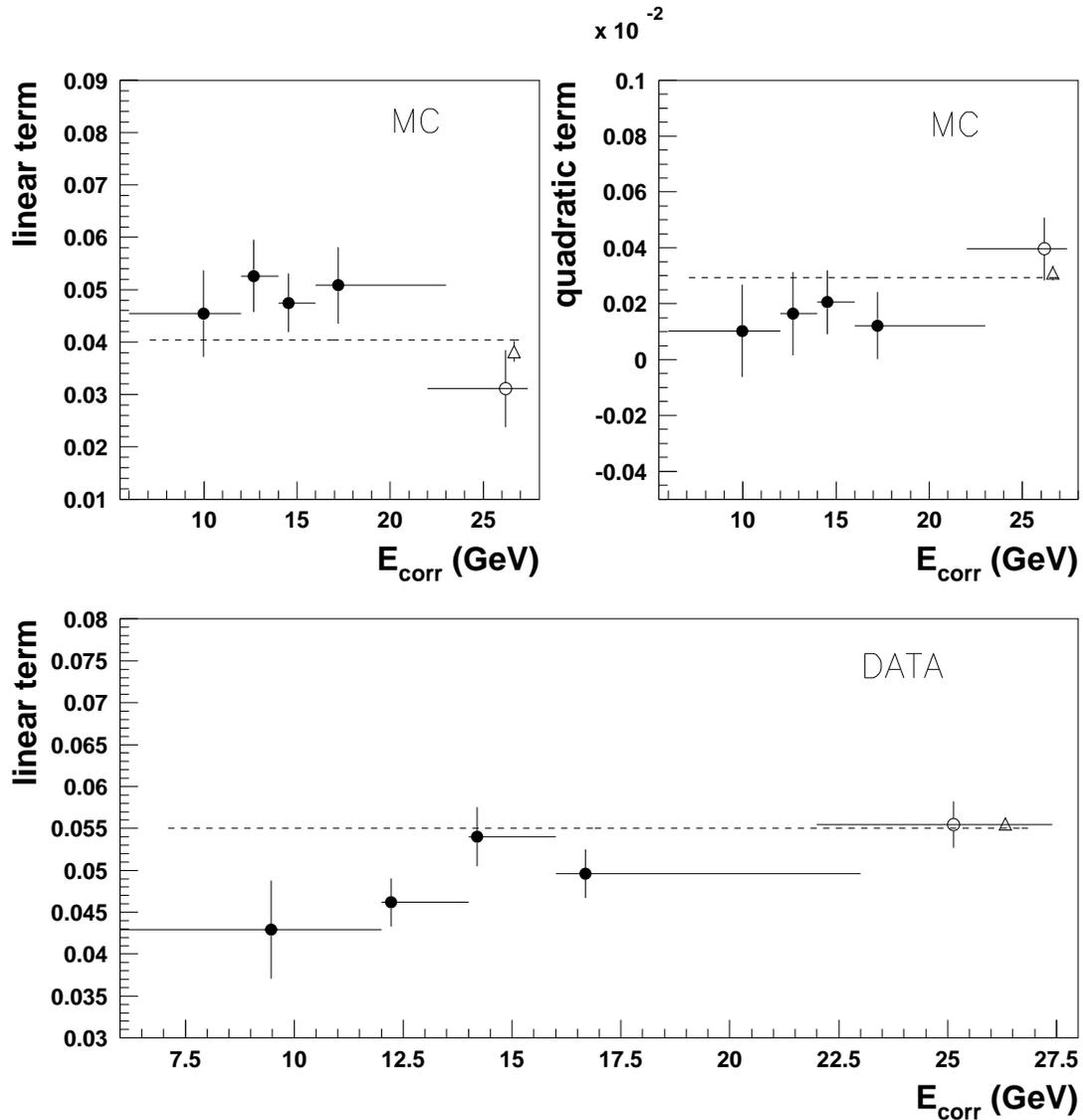


Figure 5.14: The energy dependence of the correction parameters for both Data and Monte Carlo. The linear and quadratic fit parameters are plotted versus the mean corrected energy of the sample. The filled circles indicate the elastic QED Compton sample, the open circles the DIS ρ sample and the open triangles the kinematic peak sample. The dotted lines indicate the weighted mean values of the correction parameter in each plot.

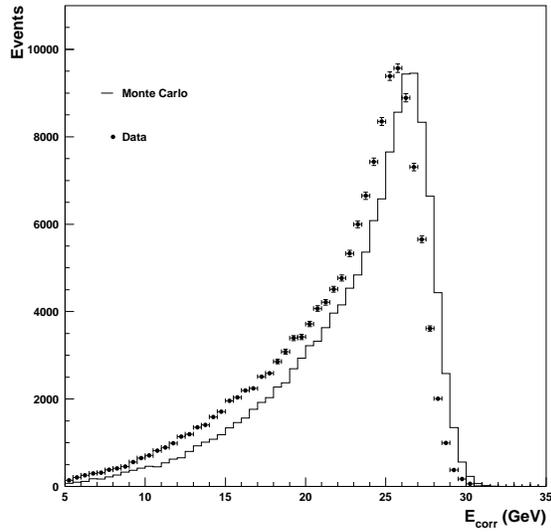


Figure 5.15: Comparison between data and Monte Carlo scattered positron energy spectra, E_{corr} , after the dead material correction has been applied. There is a clear shift between the two distributions even after the SRTD corrections have been applied.

kinematic peak and DIS ρ^0 samples in both data and Monte Carlo each yield one scaling factor. Using the bins in predicted energy which were used for the determination of the dead material correction from the elastic QED Compton samples, four scaling factors of both data and Monte Carlo can be determined. The value of the scaling factor is plotted versus average dead-material corrected positron energy in figure 5.16 for both data and Monte Carlo. Two things can be surmised from this distribution: Firstly, the difference between data and Monte Carlo remains constant over the whole energy range and secondly, as with the other correction parameters, there appears to be some systematic variation in the value of the scaling factor with scattered positron energy. However, as before, statistics are too limited to be able to parametrise this dependence. As a result, the scaling factor obtained from the data and Monte Carlo kinematic peak samples are used. But as can be seen in figure 5.16, the kinematic peak scaling factor is about 1 % lower than the scaling factor for the low energy QED Comptons, so using this value could cause a undercorrection at lower energies. The scaling factor data using Method I is 1.0315 ± 0.0005 and for Monte Carlo is 1.0181 ± 0.0008 .

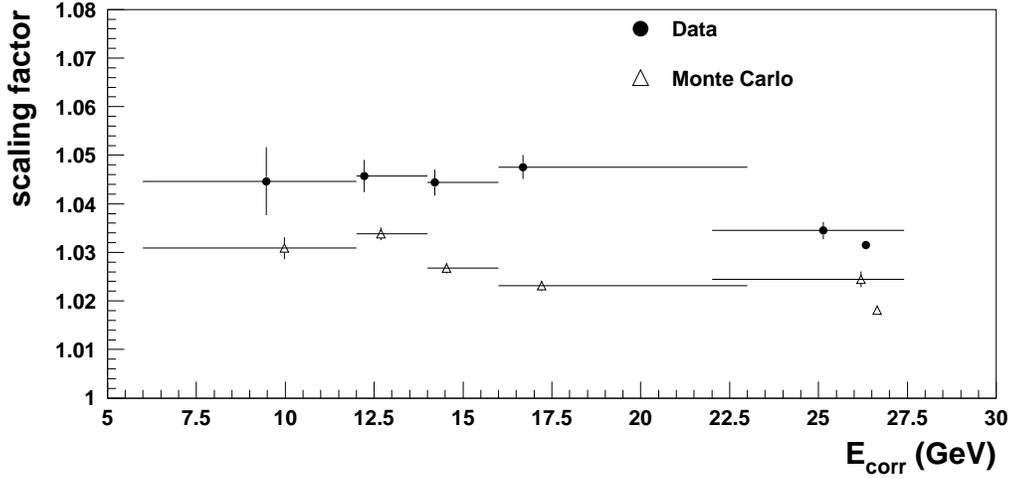


Figure 5.16: The distribution of the shifts to the predicted energy in data and Monte Carlo as a function of corrected energy (before the shift is applied).

5.6.2 Results from Method II

Using the functional form for the SRTD energy correction indicated by Method I and the resolutions described in section 5.4, the χ^2 given in equation 5.4 was minimised to obtain the optimum values of the correction parameters for data and Monte Carlo. These fits were based on a total of 41,820 data events and 38,573 Monte Carlo events (combining all three samples) and gave the following results:

$$E_{corr} = s \cdot (E_{CAL} + m \cdot E_{SRTD}/2) \quad (5.7)$$

where $s = 1.0372 \pm 0.0003$ and $m = 0.05619 \pm 0.0002$ GeV/MIP for the data, resulting in a $\chi^2/ndof$ of 3.0. For the Monte Carlo the result was given by

$$E_{corr} = s \cdot (E_{CAL} + a \cdot E_{SRTD}/2 + b \cdot (E_{SRTD}/2)^2) \quad (5.8)$$

where $s = 1.0099 \pm 0.0003$, $a = 0.0791 \pm 0.0005$ GeV/MIP and $b = (-1.41 \pm 0.06) \times 10^{-4}$ GeV/MIP². This fit was found to have a $\chi^2/ndof$ of 3.3.

5.6.3 Comparison of Method Results

The fit parameters obtained from both methods are summarised in table 5.2. The parameters obtained from each method for data appear to be reasonably consistent for the SRTD

energy dependent slope parameter. There is a small difference between the scaling factor obtained from methods I and II, but this is to be expected as the result from method I is from the kinematic peak sample only, while the result from method II is dominated by the contribution from the DIS ρ^0 sample. Examination of figure 5.16 shows that the scaling factor from the DIS ρ^0 sample is a little higher than that from the kinematic peak sample. However, when a comparison between the two methods is made for the Monte Carlo

DATA			
METHOD I		METHOD II	
Scaling factor	Linear term	Scaling Factor	Linear term
1.0315 ± 0.0005	0.0549 ± 0.0005	1.0372 ± 0.0003	0.0562 ± 0.0002
MONTE CARLO			
METHOD I		METHOD II	
Scaling factor	Linear term	Scaling Factor	Linear term
1.0181 ± 0.0008	0.0404 ± 0.0016	1.0099 ± 0.0003	0.0791 ± 0.0005
Quadratic term		Quadratic term	
$(2.9 \pm 0.2) \times 10^{-4}$		$(-1.41 \pm 0.06) \times 10^{-4}$	

Table 5.2: Comparison of the results obtained from correction methods I and II for data and Monte Carlo

samples, there is a clear discrepancy in all the fitted parameters. In order to determine the source of this discrepancy, a comparison is made between the predicted and corrected positron energies as a function of SRTD energy deposit, for the two Monte Carlo results. The distributions can be seen in figure 5.17. While the agreement is good up to SRTD energies of about 100 MIPS, the method I correction clearly fails above this value (this is only reached in the kinematic peak sample). This is due to the low statistics at these energies, which will have a more significant effect on method I than on method II. This indicates that method II is more reliable. In addition a comparison of the parameter error sizes between the two methods indicates that method II has a higher precision. Thus the final 1995 SRTD corrections are those obtained from method II.

Having obtained the final correction values, a comparison was made between data and

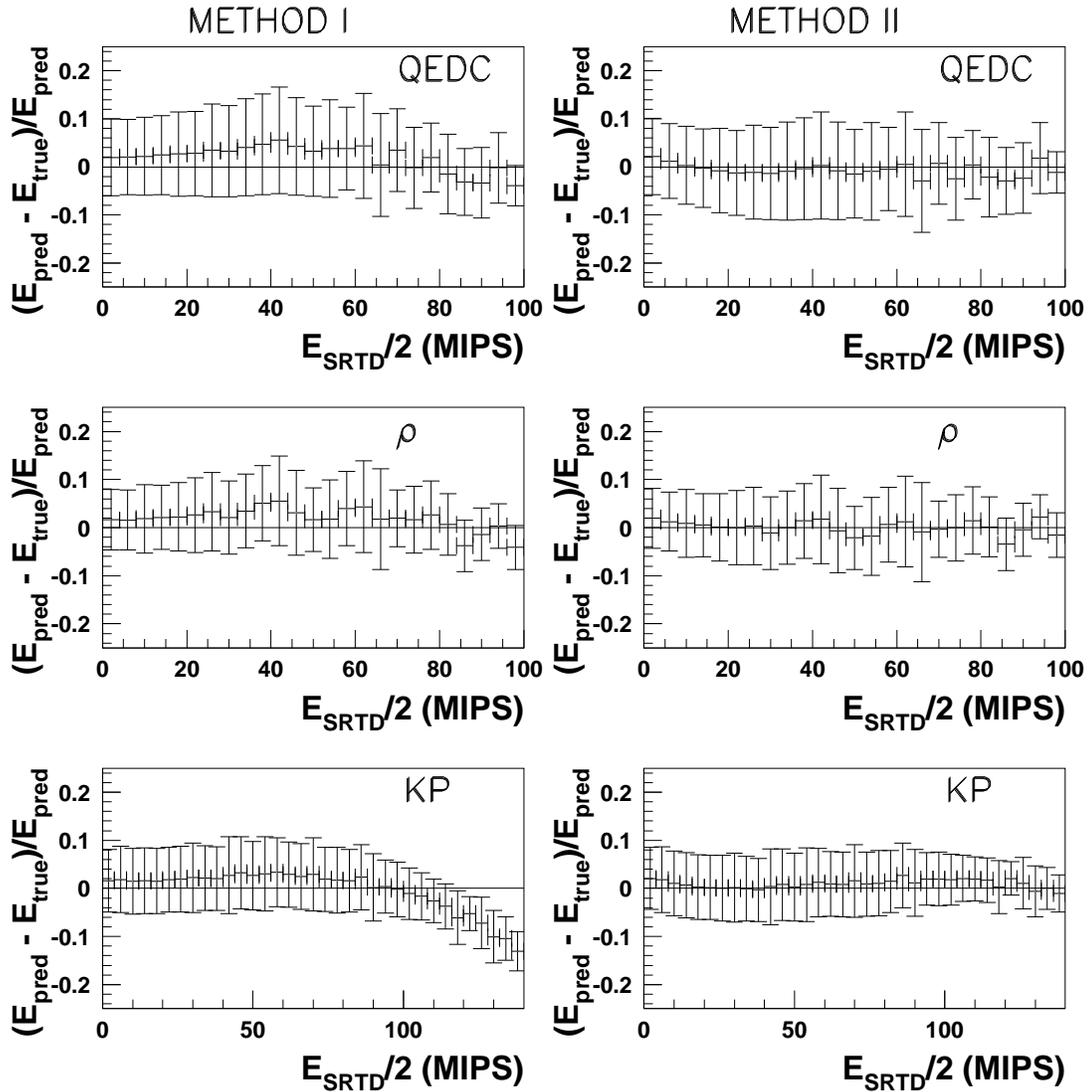


Figure 5.17: Comparison of the predicted energy to the corrected energy in Monte Carlo for the three samples as a function of SRTD energy, using the two different methods (see text). The error bars indicate the RMS in each bin of SRTD energy. A clear discrepancy can be seen in the kinematic peak sample corrected using method I.

Monte Carlo before and after correction, see figure 5.18. These plots show that after correction, the agreement between data and Monte Carlo is generally quite good, but that there are still small discrepancies. For example, in the distribution of the corrected scattered positron energy, the data are consistently slightly higher than the Monte Carlo in the low energy region and there is a slight difference between the two distributions on the high side of the peak. The Monte Carlo sample contains no photoproduction, diffraction or any F_L contribution. The low energy tail is the region which is most sensitive to background or differences in structure function, thus any or all of these could contribute to the discrepancy observed. However, none of these contributions could be the cause of

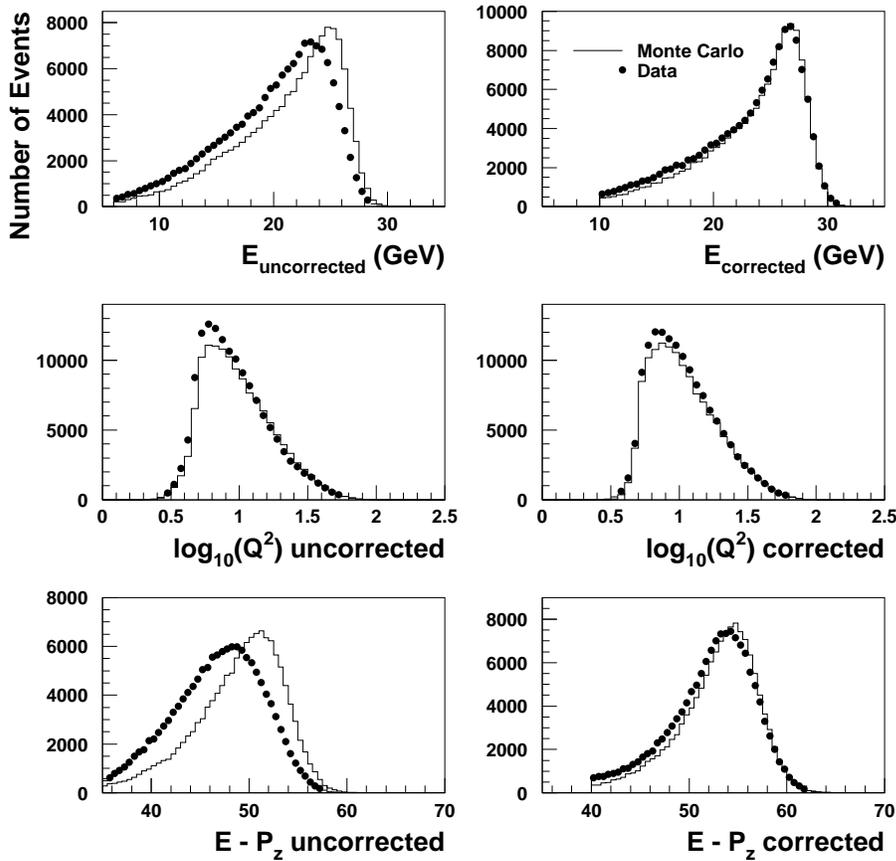


Figure 5.18: Comparison of scattered positron variables, before and after correction, between Data and Monte Carlo. The Monte Carlo is weighted by luminosity and is reweighted to the structure function CTEQ4D.

the slight discrepancy in the peak of the distribution. One possibility is a slight energy dependence in the correction parameter values, as has been previously discussed. However,

these parameter values are dominated by the contribution from the DIS ρ^0 sample, which lies just below the kinematic peak in terms of predicted energy, so the energy dependence argument should be ruled out for the peak region, but could be the explanation for the discrepancy in the low energy tail region. There is also the possibility that this discrepancy is due to position in the calorimeter. This can be investigated by comparing data and Monte Carlo kinematic peak events, after correction, in different calorimeter cells. In order to perform this investigation, a numbering scheme was devised for cells covered by the SRTD and is illustrated in figure 5.2. The hatched area is the SRTD and the thick black lines indicate the region excluded by the box cut, the removal of the crack regions and the fiducial cuts placed on the SRTD in order to keep ambiguous hits. Ten cells were then compared as shown in figure 5.19. The highest populated cells (34 and 36) are those directly above and below the beampipe and will therefore have made up the bulk of the events used to determine the SRTD corrections. Thus, it is not surprising that the agreement is good. However, there are noticeable differences in other cells which could be accounted for by differences in calorimeter cell calibration. In order to test this possibility, there are two important factors to be dealt with. Firstly, the position-dependence of the true kinematic peak value must be removed and secondly, the effects of dead material must be accounted for. This can be done by comparing cells radially equidistant from the beampipe and only using positrons which are associated with a small SRTD energy deposit. Unfortunately, the statistics are too limited in 1995 data to be able to check this possibility, but the 1996 dataset should be large enough to make this investigation.

5.7 Conclusions

Dead material effects on the scattered positron using the SRTD in 1995 data have been studied. It has been found that although there has been a great improvement in the simulation of dead material in Monte Carlo between the 1995 version of MOZART and previous versions, there is still some discrepancy between Monte Carlo and data. Three data types have been used to extract a correction based on SRTD information, these are kinematic peak events, DIS ρ^0 events and elastic QED Compton events. The methods of predicting the energy of the scattered positron in each of these samples have been thoroughly investigated to ensure that they are valid. In the case of the kinematic peak sample, it was found that the “true” peak value in the Monte Carlo was actually 27.18 ± 0.02

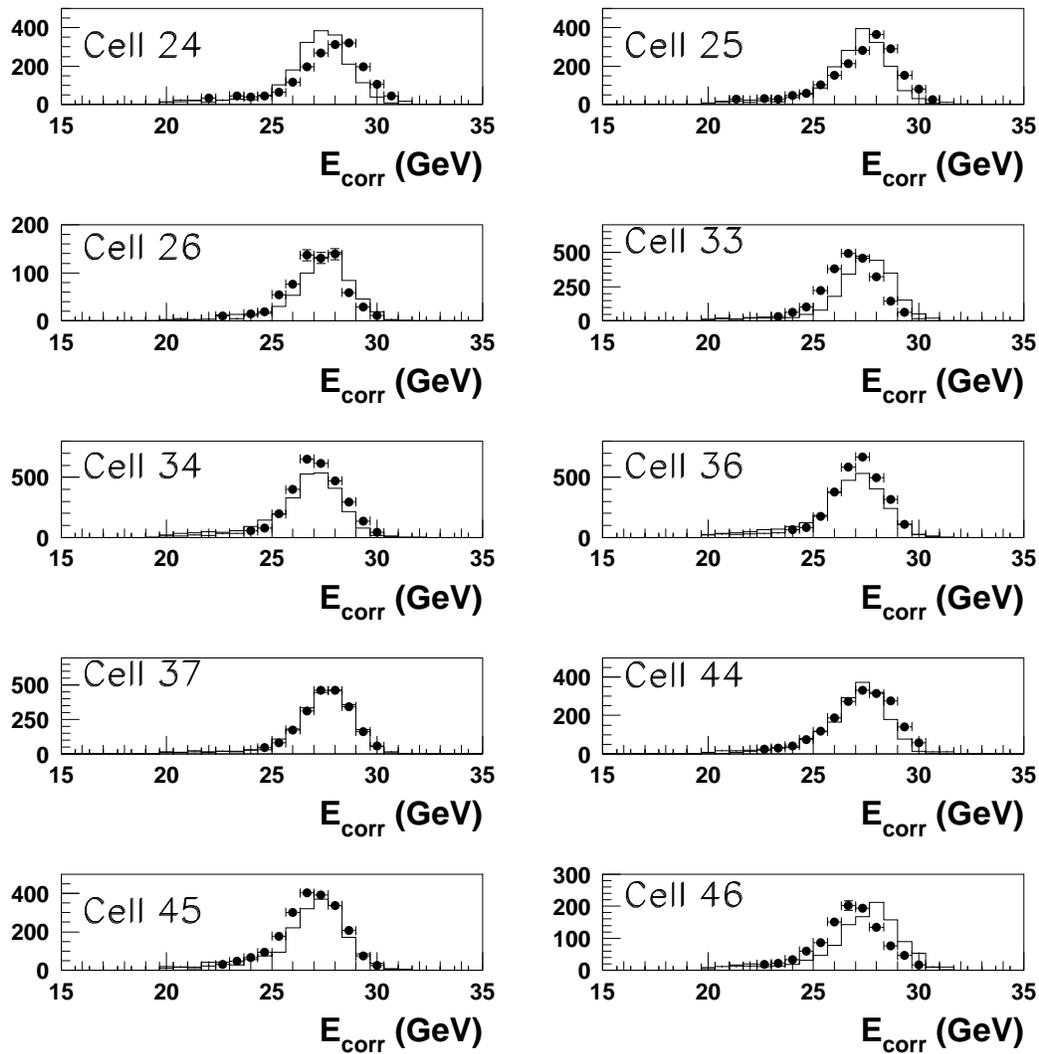


Figure 5.19: Cell-by-cell comparison of the data and Monte Carlo kinematic peak samples. There are clear cell-to-cell variations, which could be accounted for by differences in cell-by-cell calibration between data and Monte Carlo.

GeV, ie. slightly lower than the nominal beam energy. It was also found that the method used for the DIS ρ^0 sample was insensitive to the effects of non-zero proton p_T , but that the presence of initial state radiation has a significant effect on the prediction method. p_T balance was used to reduce the effect of ISR to a manageable level.

In the elastic QED Compton sample, it was found that the effect of ISR was noticeably smaller than in the DIS ρ^0 sample, but that the effect of proton p_T was much more significant. However, by using appropriate cuts these effects were minimised in the range $6 < E_{predicted} < 23$ GeV. The candidates outside this range are the highly-asymmetric QED Comptons, which were found to be much more sensitive to proton p_T and thus not usable for the extraction of corrections.

Having established “safe” samples of events, the corrections were extracted for both data and Monte Carlo. It was found that the form of the correction for the data was linear, but that the kinematic peak and DIS ρ^0 Monte Carlo samples were well described by a second order polynomial. The situation with regard to Monte Carlo elastic QED Comptons is not so clear due to limited statistics and low SRTD energy deposits. In the absence of concrete evidence of the correction parameters varying systematically with calorimeter positron energy, the same function and parameter values were assumed for the full scattered positron energy range. Two methods were used to extract the corrections: the bin fitting method and the event-by-event χ^2 minimisation method. The latter method gave the best results.

The full scattered positron energy spectra taken from neutral current data and Monte Carlo samples were compared after correction above $E'_e = 10$ GeV and the overall agreement was found to be reasonably good. However, there were small discrepancies; the data were consistently slightly higher than the Monte Carlo in the low energy region and there was a slight mismatch on the high side of the peak. The former could be due to several things, including the absence of diffractive and F_L contributions in the Monte Carlo. The latter, however, cannot be explained by any of these factors. There are indications that this small discrepancy could be due to the simulation of calorimeter calibration in Monte Carlo.

Chapter 6

Selection of Candidates

In chapter 4, the establishment of a high quality, fully-reconstructed data sample was described in detail. This sample then forms the starting point for the selection of a DIS sample from which diffractive events containing a $D^{*\pm}(2010)$ candidate in the decay channel

$$D^{*+} \rightarrow D^0 \pi_s^+ \rightarrow (K^- \pi^+) \pi_s^+ \quad + \text{c.c.}$$

can be selected.

In this chapter the criteria for the selection of such events is discussed, along with the remaining backgrounds which must be removed. The different Monte Carlo generators used will also be described. After a full discussion of the trigger strategy used in this analysis, the offline DIS and $D^{*\pm}$ selection methods are covered. After the DIS $D^{*\pm}$ signal has been studied, the diffractive requirements are discussed and the signal investigated. Finally, the characteristics of the data sample are presented.

6.1 Monte Carlo Simulation

Several different Monte Carlo generators were used in the tuning of selection requirements and subsequently for cross section unfolding, all of which are described in this section. In the case of the $D^{*\pm}$ Monte Carlos (RAPGAP [128] and RIDI [65, 67]), two samples of each type were generated. One of these was processed with the 1995 versions of MOZART and ZGAN and the other with the 1996 version of MOZART and the 1997 version of

ZGANA. The former is used to describe 1995 data and the latter is used to describe both 1996 and 1997 data. This strategy is used because the detector changed significantly between the 1995 and 1996 data-taking periods, but not between the 1996 and 1997 data-taking periods. The same is true of the DIS trigger slots, and hence of ZGANA. All $D^{*\pm}$ Monte Carlo samples contain at least one $D^{*\pm}$ decaying in the $D^* \rightarrow (K\pi)\pi_s$ channel in each event.

RAPGAP

The principal diffractive $D^{*\pm}$ sample was generated using RAPGAP [128]. The Ingelman-Schlein model is used to simulate the diffractive exchange, assuming the Donnachie-Landshoff pomeron flux with $\alpha_P(0)$ set to 1.157 (as measured by ZEUS [129]). The pomeron is assumed to consist predominantly of hard gluons, that is the pomeron structure function is defined by $\beta f_{\mathbb{P}/g}(\beta, Q_0^2) = 6\beta(1 - \beta)$ for the initial gluon distribution and $\beta f_{\mathbb{P}/q}(\beta, Q_0^2) = \frac{1}{4}f_{\mathbb{P}/g}(\beta, Q_0^2)$ for the initial quark distributions, where it is assumed that only u and d quarks contribute. Q_0^2 is set to 4 GeV² and the parton distributions are evolved using the next-to-leading order DGLAP equations. No momentum sum rule is applied to the pomeron.

RAPGAP simulates charm production at the parton level using the leading order massive charm scheme. Hence, charm is produced only via boson gluon fusion, where the gluon originates from the pomeron. Charm fragmentation is performed according to the Peterson fragmentation model with $\varepsilon = 0.035$ (see chapter 1). The colour dipole model implemented in ARIADNE [130] is used to simulate QCD parton shower evolution, while the LUND string model [120] as implemented in JETSET [121] is used for hadronisation.

RAPGAP is also used to produce a non-diffractive $D^{*\pm}$ sample via standard boson gluon fusion, with GRV94(HO) as the input proton structure function. However, in this case QED radiative corrections were also included using an interface to the HERACLES [118] generator. In both the diffractive and non-diffractive samples the charm mass is set to 1.25 GeV. In order to reduce the sample sizes, only $D^{*\pm}$ mesons with $p_T(D^{*\pm}) > 0.8$ GeV are retained. In both samples events were generated with $Q^2 > 0.6$ GeV².

RIDI

As described in chapter 2, RIDI [67] simulates diffractive interactions via the t -channel exchange of a perturbative gluon ladder using the Ryskin model [65] and includes $q\bar{q}$ and $q\bar{q}g$ final states, as well as virtual gluon corrections. Fragmentation, parton shower evolution and hadronisation are all performed using JETSET. This sample includes QED radiative corrections and has the charm mass set to 1.25 GeV. Given that in this case the diffractive cross section is proportional to the square of the proton gluon density, it is important to note that the input proton structure function is GRV94(HO). The other parameter to which the cross section is sensitive is the choice of “ K -factor” or the enhancement caused by the virtual gluon corrections. In this sample it is set to $\exp(\alpha_s(M_X^2)\pi C_F)$ where C_F is the colour factor described in chapter 2. In this sample, events were generated with $Q^2 > 0.6 \text{ GeV}^2$ but also with the requirement that $Q^2 < 300 \text{ GeV}^2$.

DJANGO

An inclusive non-diffractive DIS sample is also used for background studies and has been generated using DJANGO6V2.4 [117, 118] in the same way as the sample described in the previous chapter. It is important to note that this sample does not contain any diffractive interactions.

6.2 Trigger Strategy

The FLT and TLT trigger requirements used in this analysis are crucial, as not only are they required to provide an appropriate starting sample, but they should also be chosen to maximise the sample size. Clearly a DIS $D^{*\pm}$ sample will be a relatively small subset of an inclusive DIS sample, hence a trigger strategy based on inclusive DIS triggers provides the appropriate starting point for this analysis. These trigger slots have already been described in chapter 4. The size of the starting sample can be maximised by requiring either FLT30 or FLT44 to have fired and by requiring the low Q^2 TLT slot (referred to as DIS01) to have fired. However, due to trigger rate considerations, FLT30 and DIS01 (the former being the FLT low Q^2 slot) were not available for the total data-taking period

and hence alternatives had to be sought. This section describes the FLT and TLT trigger strategy in full detail.

6.2.1 FLT Strategy

During the 1995 and 1996 data-taking periods, DIS selection at the FLT was performed using a logical *.or.* of FLT30 and FLT44 (see section 4.1.2 for details). However, in 1997, FLT30 was intermittently prescaled by a factor of 8 and hence FLT46 was used instead of FLT30.

The efficiency for both FLT30 and FLT44 has been studied in detail previously [5] and has been found to be above 99 % for events which pass offline DIS selection cuts. FLT46 is simply FLT30 with the additional requirement that most FLT tracks found in an event come from the nominal interaction region. This is suitable for $D^{*\pm}$ event selection as $D^{*\pm}$ candidates are selected offline in such a way that every event contains a minimum of three well-reconstructed tracks associated to the primary event vertex. For this reason, FLT46 is expected to have a high efficiency in the selection of DIS $D^{*\pm}$ candidates. This can be checked using the final DIS $D^{*\pm}$ sample from data and Monte Carlo for 1997, defining the efficiency as

$$\varepsilon = \frac{\text{Final DIS } D^{*\pm} \text{ selection} * \text{FLT30} * \text{FLT46}}{\text{Final DIS } D^{*\pm} \text{ selection} * \text{FLT30}} \quad (6.1)$$

where $*$ is used to indicate a logical *.and.*. The efficiency is plotted in bins of Q^2 , W , $p_T(D^{*\pm})$ and $\eta(D^{*\pm})$ for both data and Monte Carlo in figure 6.1. Figure 6.1 shows that the efficiency of FLT46, relative to FLT30, for $D^{*\pm}$ selection is very high (greater than 97 %) and that the agreement between data and Monte Carlo is good to within 2 %. This systematic difference between data and Monte Carlo is much smaller than the statistical errors associated with the final results discussed in chapter 7 and therefore can be neglected.

6.2.2 TLT Strategy

In all three years, 1995 — 1997, DIS01 was only switched on and unprescaled for limited periods of time due to its large contribution to the TLT output rate. When DIS01 is unavailable, the medium Q^2 TLT trigger (DIS03) is used in its place. However, this

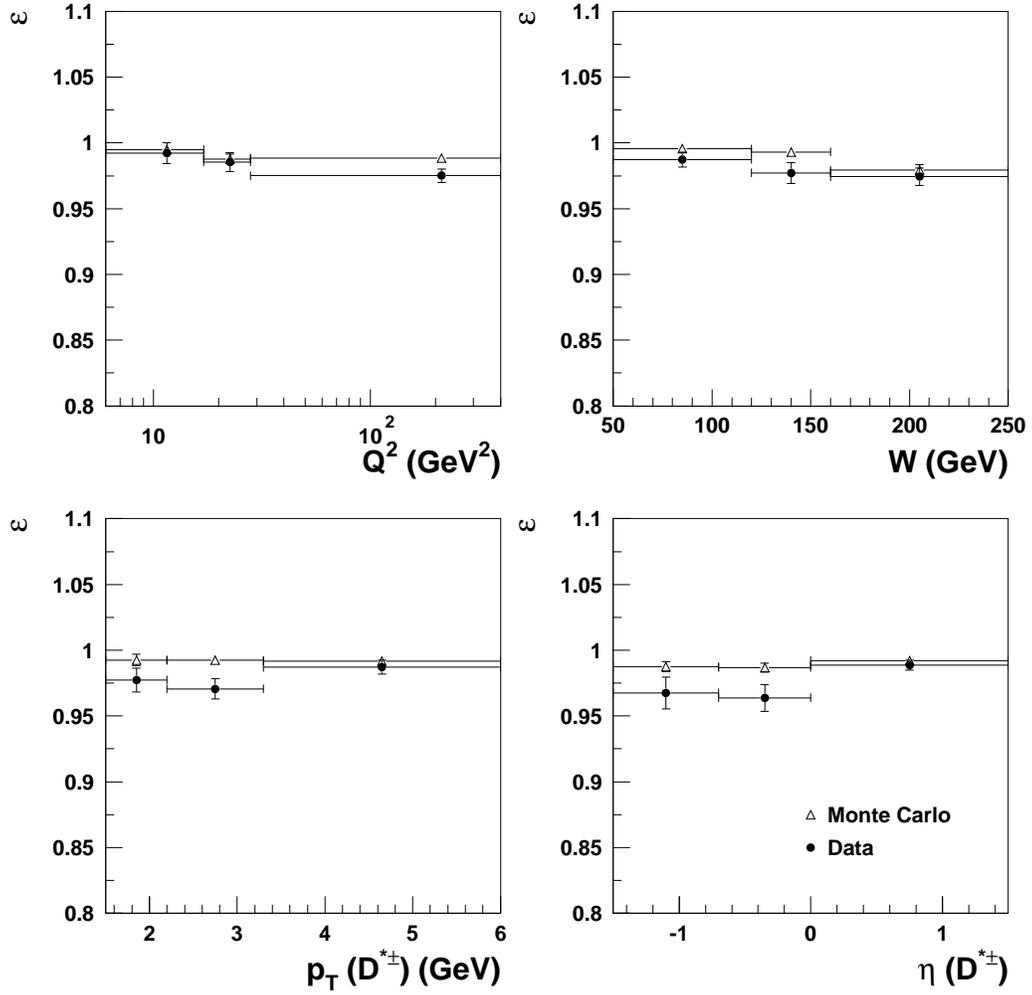


Figure 6.1: The efficiencies for FLT46 (see text) after the final DIS $D^{*\pm}$ cuts in both data (closed circles) and Monte Carlo (open triangles) as a function of Q^2 , W , $D^{*\pm}$ transverse momentum and $D^{*\pm}$ pseudorapidity.

trigger was sometimes implemented with a 14×14 cm box cut and at other times with a 25 cm radius cut. In addition, a special TLT trigger was available during 1996 and 1997 which combined DIS01 with a $D^{*\pm}$ filter. The $D^{*\pm}$ filter was designed to select candidates via the reconstruction method used for offline $D^{*\pm}$ selection (described later in this chapter), using the following cuts:

- $1.4 < M(D^0) < 2.3$ GeV.
- $(M(D^{*\pm}) - M(D^0)) < 0.17$ GeV.
- $p_T(\pi_s) > 0.05$ GeV.
- $p_T(K) > 0.15$ GeV.
- $p_T(\pi) > 0.2$ GeV.
- $p(D^0)/p(\pi_s) > 8$.
- Combinations with the correct charge only.

This trigger, known as DIS13, is used in addition to DIS01 and DIS03, and is particularly important when only DIS03 is available. The inclusive DIS TLT trigger strategy is summarised in table 6.1. DIS13 was available for runs greater than 21504 in 1996, ie. about 92 % of the 1996 dataset, and for the whole of 1997.

The efficiency of DIS13, relative to the inclusive DIS triggers, can be checked using the final DIS $D^{*\pm}$ sample from data and Monte Carlo from the 1996/7 data-taking periods. The efficiency is defined as

$$\varepsilon = \frac{\text{Final DIS } D^{*\pm} \text{ selection} * \text{DIS01} * \text{DIS13}}{\text{Final DIS } D^{*\pm} \text{ selection} * \text{DIS01}} \quad (6.2)$$

The efficiency of DIS13 is measured using events triggered by DIS01 in order to make a proper assessment of the efficiency of the $D^{*\pm}$ filter. This means that only those parts of the data in which DIS01 was available may be used for this study (see table 6.1). The efficiency is plotted in bins of Q^2 , W , $p_T(D^{*\pm})$ and $\eta(D^{*\pm})$ in figure 6.2. The efficiency is found to be approximately 80 % and the agreement between data and Monte Carlo is good to within approximately 5 %. This overall difference in efficiency between data and Monte Carlo as well as the small differences observed in the shapes of the Q^2 and $p_T(D^{*\pm})$ distributions will be taken into account when the systematic errors are evaluated in the next chapter.

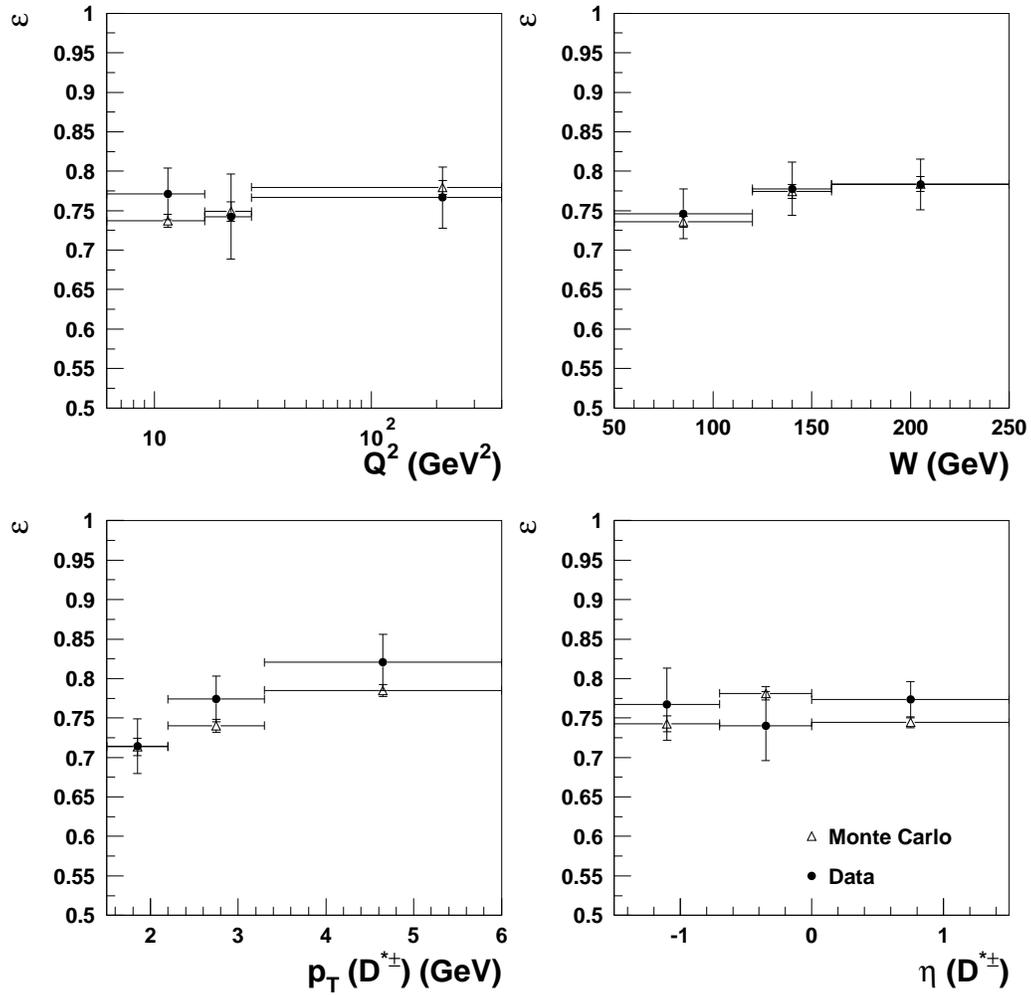


Figure 6.2: The efficiencies for DIS13 (see text) after the final DIS $D^{*\pm}$ cuts in both data (closed circles) and Monte Carlo (open triangles) as a function of Q^2 , W , $D^{*\pm}$ transverse momentum and $D^{*\pm}$ pseudorapidity.

Year	Run Range	TLT trigger	RCAL fiducial cut (cm)	Integrated Luminosity (pb ⁻¹)
1995	11539 - 12780	DIS01	12 × 6	2.41
	12781 - 14397	DIS03	14 × 14	3.91
1996	21186 - 21504	DIS01	12 × 6	0.80
	21505 - 21631	DIS01	12 × 6	0.50
	21634 - 22448	DIS03	14 × 14	5.73
	22449 - 22462	DIS01	12 × 6	0.15
	22466 - 22660	DIS03	radius > 25	0.80
	22662 - 22954	DIS01	12 × 6	2.42
1997	25190 - 25337	DIS01	12 × 6	0.71
	25344 - 27889	DIS03	radius > 25	25.89

Table 6.1: A Summary of availability of the different TLT triggers in the three different years of data-taking.

6.3 DIS Event Selection

As was discussed in chapter 4, the key feature of a NC DIS event is the identification of the scattered positron in the main detector. The various complicating factors, such as dead material and low energy forward-going hadrons, have also been discussed and in this section the steps which can be taken to minimise their effects offline will be discussed. The principal source of physics background, photoproduction, in which the scattered positron disappears undetected down the rear beampipe but electromagnetic energy deposits from, for example, $\pi^0 \rightarrow \gamma\gamma$ decays, fake a positron, is also discussed.

The cuts required to extract a clean DIS sample with high efficiency are described here and are based on the selection methods used in the extraction of the proton structure functions F_2 and $F_2^{c\bar{c}}$ [5, 16].

Positron Identification

As was discussed in chapter 4, the scattered positron is identified using the neural net positron finder SINISTRA [90]. In each event, the candidate with the highest probability

output by the neural net is chosen as the scattered positron candidate. If its probability is greater than 0.9, then the event is kept in the final sample. Figure 6.3 shows the probability distribution from the non-diffractive RAPGAP sample, indicating that the majority of true positron candidates have a very high probability and would therefore pass this selection cut. The candidate is also required to have a corrected energy greater than 8 GeV. This ensures that SINISTRA operates with high efficiency and purity and is equivalent to requiring $y < 0.7$. This, in turn, excludes the region of phase space in which photoproduction events containing a low energy “fake” positron in the FCAL are found. This effect is illustrated in figure 6.4 which shows the y_e distribution observed in the data. The shaded area indicates those events excluded by requiring $y_{DA} < 0.7$. The tail in the shaded distribution is caused by the migrations associated with y_{DA} being different from those associated with y_e .

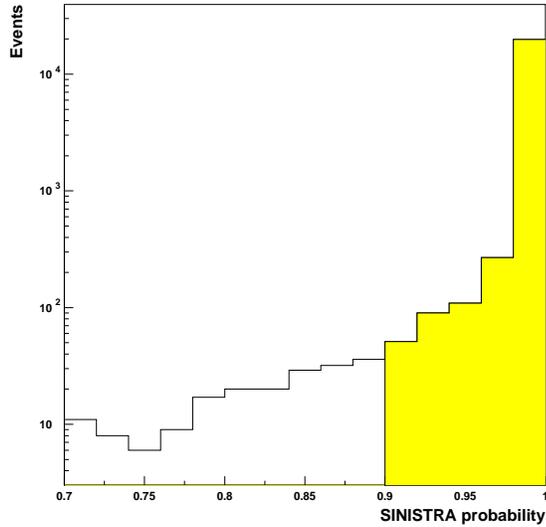


Figure 6.3: The SINISTRA probability distribution from non-diffractive RAPGAP Monte Carlo. The shaded area indicates those candidates which are kept in the final $D^{*\pm}$ sample. Note that the peak near 1 is several orders of magnitude larger than the tail.

Energy-Momentum Conservation

The variable δ was discussed in chapter 4 as a quantity which can be used to reject both non- ep and photoproduction background. The variable which is used in offline selection

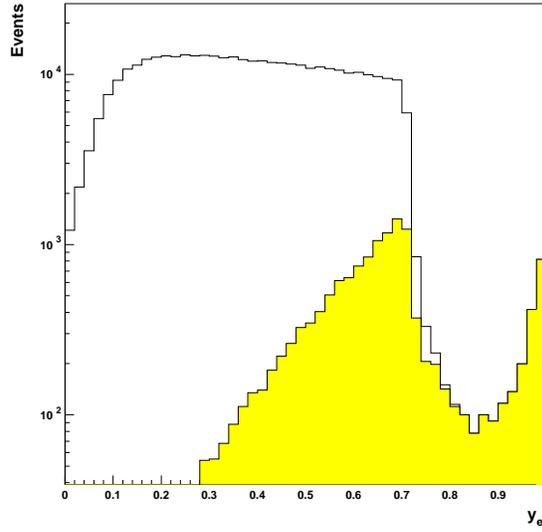


Figure 6.4: The y_e distribution from the data, after all DIS selection requirements have been applied. The peak at very high y comes from “fake” positron events. The shaded region indicates the events which are excluded by the $y_{DA} < 0.7$ requirement. This clearly excludes the “fake” positron background.

is actually

$$\delta = E - p_z = \sum_i E_i(1 - \cos \theta_i) \quad (6.3)$$

where the sum over i indicates the sum over all ZUFOs found in an event plus the scattered positron. Once again the positron energy is corrected before it is used in the determination of $E - p_z$.

Due to the definition of the proton beam direction as the positive z or forward direction, particles which travel forward will make a small contribution to the total $E - p_z$ of the event, while particles which travel backwards will make a large contribution to the total $E - p_z$. Thus, based on energy-momentum conservation, one expects a well-contained NC DIS event to have $\delta \simeq 2E_e$. However, in photoproduction, where the scattered positron disappears undetected down the rear beampipe, or in a DIS event in which an initial state photon is radiated, again escaping down the rear beampipe, a low value of $E - p_z$ is expected. Thus, requiring $E - p_z > 40$ GeV, will help to remove any remaining photoproduction background and will limit the contribution from ISR events. This latter point is important as it is not currently possible to generate diffractive charm events with radiative corrections using RAPGAP.

At the other end of the $E - p_z$ scale, some proton beam gas interactions can result in unphysically high values of $E - p_z$ and so requiring $E - p_z < 65$ GeV will help to remove any remaining non- ep background.

Event Vertex Requirements

It is expected that $D^{*\pm}$ events will tend to have a well-defined primary interaction vertex due to the nature of the decay channel being studied (the resolution of the CTD is not sufficient to distinguish the displaced secondary vertex which could arise from the D^0 decay). It is therefore important only to accept events which have a well-reconstructed vertex which lies within the main detector volume, that is, $-100 < z_{vtx} < +100$ cm. This also helps to exclude any remaining non- ep interactions which cannot usually satisfy such a requirement.

Box Cut

The majority of the events within the DIS sample have a relatively low value of Q^2 so that the scattered positron is detected in the RCAL. At the lowest Q^2 values, the scattered positron is detected near the beamhole. When the scattered positron is detected close to the beamhole, the cluster may not be fully-contained within the calorimeter, causing the energy and impact position to be incorrectly measured. To ensure that only well-reconstructed positrons are kept in the final DIS sample, a box cut is applied (see chapter 5 for the definition of this fiducial cut). As was discussed in chapter 4, each of the DIS TLT trigger slots contains some sort of box cut. However, the box cuts are based on simplified positron position reconstruction algorithms compared to those used in offline reconstruction. Thus, to ensure that the offline position is accurately reconstructed, the box cut is made 1 cm tighter than that implemented in the trigger. This choice of offline box cut is consistent with the position resolution of the calorimeter.

For example, when DIS01 or DIS13 are in use, both of which require $|x| > 12$ cm or $|y| > 6$ cm (written $12(x) \times 6(y)$ cm), the offline cut is made at $13(x) \times 8(y)$ cm. The extra centimetre on y is to remove positrons which have hit the very high dead material regions at the corners of the rear beampipe (see the previous chapter for more details). When DIS03 is in use, the box cut is made at $15(x) \times 15(y)$ cm or radius > 26 cm, depending

on which was implemented in the trigger, as described in section 6.2.

It is useful to note that the loosest box cut which is applied in this analysis is equivalent to setting a minimum value of $Q^2 \simeq 1 \text{ GeV}^2$, as can be seen in figure 6.5. The loosest box cut corresponds to a scattering angle of approximately 177° for the positron. In figure 6.5 there are a small number of events which lie in the region of the kinematic plane which is excluded by the box cut. These events all pass the box cut, but have an interaction vertex which is significantly shifted compared to the nominal values. This results in a scattering angle for the positron which is smaller than 177° , reducing the Q^2 of the event.

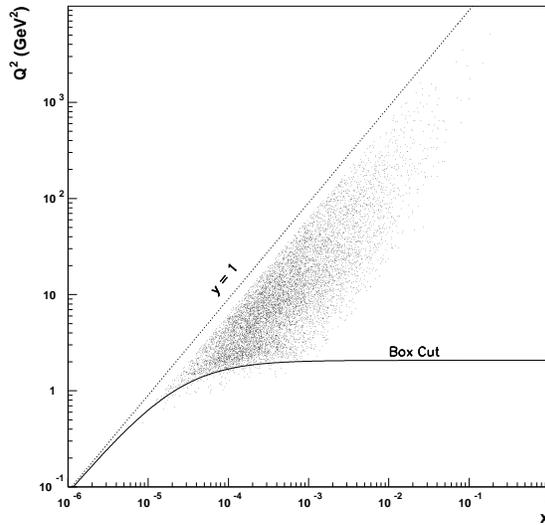


Figure 6.5: The $Q^2 - x$ distribution for data which satisfy all DIS requirements and $D^{*\pm}$ preselection. The data clearly lie predominantly above $Q^2 = 1 \text{ GeV}^2$.

Hadronic Energy Reconstruction

The final requirement used in the selection of DIS events, is $y_{JB} > 0.02$. y_{JB} is defined in section 4.4 and is calculated using ZUFOS. This cut ensures that there is a significant level of hadronic activity in the event. This is important as noise can distort the reconstruction of event kinematics using the Double Angle method, and is most prominent in events which only contain a low level of hadronic activity.

6.4 $D^{*\pm}(2010)$ Candidate Selection

$D^{*\pm}(2010)$ mesons which decay in the $D^* \rightarrow K\pi\pi_s$ channel provide a powerful method of tagging charm. The kinematic constraints on the $D^* \rightarrow D^0\pi$ stage of the decay cause a very clean signal to be produced in the $\Delta M = M(D^*) - M(D^0)$ distribution in a region where the combinatorial background is suppressed [24]. In addition, this decay channel results in three charged particles which can be detected in the tracking chamber.

The $D^{*\pm}$ reconstruction procedure is relatively straightforward. First, pairs of oppositely-charged tracks which come from the primary event vertex are taken, assuming one to be the kaon and the other to be the pion from the decay of the D^0 . Their combined invariant mass is calculated, and if it is found to lie in a restricted “window” around the nominal D^0 mass, then a further track of opposite sign to the kaon candidate is added to the pair under the assumption that it is the slow pion. After the application of suitable cuts to suppress the combinatorial background, a prominent signal is seen in the $\Delta M = M(K\pi\pi_s) - M(K\pi)$ distribution, from which the number of $D^{*\pm}$ candidates can be determined.

In this selection method, signal “windows” for both ΔM and $M(D^0)$ must be chosen in order to isolate clean peaks in the invariant mass distributions. For example, in order to achieve a clean signal in the ΔM distribution, the $K\pi$ invariant mass must be restricted to a region around the nominal D^0 mass, $M(D^0) = 1864.6 \pm 0.5$ MeV [18]. The region chosen is $1.79 < M(K\pi) < 1.93$ GeV. Conversely, when a clean signal is required in the $M(D^0)$ distribution, ΔM must be constrained to the region around the nominal value, $\Delta M = 145.397 \pm 0.030$ MeV [18]. The window chosen is $0.142 < \Delta M < 0.149$ GeV. The regions chosen are consistent with the resolution achieved in the measurement of the invariant masses.

The DIS sample, selected as described in the previous section, is searched in this manner for $D^{*\pm}$ candidates, and in this section the cuts used to suppress the combinatorial background, as well as the methods of extracting the number of signal candidates from the ΔM distribution, are discussed.

6.4.1 Track Momentum Requirements

The transverse momentum of the produced charm (or anticharm) quark is typically of the order of its mass and, given that the $D^{*\pm}$ takes, on average, 70 % of that transverse

momentum [15], it can be expected that the $D^{*\pm}$ decay products will also carry a significant p_T . It is therefore reasonable to expect that cutting on the p_T of these decay products should reduce the combinatorial background.

As was mentioned in the previous section, only tracks associated to the primary vertex are used in $D^{*\pm}$ reconstruction. In addition, only tracks reconstructed in the CTD which reach the third superlayer (SL3) are considered, ensuring that the z -coordinate of the track is reconstructed with stereo information from SL2. The SL3 requirement is also an effective p_T cut of approximately 100 MeV — a fact which is of particular importance when choosing a p_T cut for the slow pion candidate (see below).

After this basic track selection has been made, cuts which depend on the candidate particle type are applied. These are summarised below.

$p_T(\pi_s) > 0.12$ GeV. As can be seen in figure 6.6, the transverse momentum of the slow pion peaks at very low values, as expected from the kinematics of the $D^{*\pm}$ decay. For the same reason, the p_T of the slow pion and the parent $D^{*\pm}$ are highly correlated — thus it is clearly beneficial to keep this cut as low as possible. The p_T cut implied by the SL3 requirement is only an approximate requirement of 100 MeV. Thus, an explicit cut must also be applied, for which the value of 120 MeV is chosen.

$p_T(K, \pi) > 0.5$ GeV. The requirement placed on the transverse momenta of the D^0 decay products is motivated by the fact that they are expected to have a larger p_T than the light quark combinatorial background. However, these cuts must be tuned carefully. Figure 6.7 shows the efficiency and purity versus p_T cut, where the efficiency is defined as

$$\frac{\text{No. signal candidates after cut}}{\text{No. of signal candidates } (p_T(K, \pi) > 0.3)} \quad (6.4)$$

and the purity is

$$\frac{\text{No. signal candidates after cut}}{\text{No. of signal + background candidates after cut}} \quad (6.5)$$

The efficiency is normalised to the signal for $p_T(K, \pi) > 0.3$ GeV as this is the value used during the preselection of candidates.

For $p_T(K, \pi) > 0.5$ GeV, the loss in efficiency is approximately 15 %, while the gain in purity is approximately 30 %. Figure 6.7 also shows the signal p_T distributions from Monte Carlo and estimates of the background p_T distributions from data, which clearly show that the background is shifted to lower p_T than the signal.

$|p_T(K)| + |p_T(\pi)| > 1.9$ GeV. This additional scalar requirement is motivated by the distributions in figure 6.8, which show the efficiency and purity versus cut value (defined as above, but the efficiency is normalised with respect to $|p_T(K)| + |p_T(\pi)| > 1.7$ GeV), as well as $p_T(K)$ versus $p_T(\pi)$ for signal Monte Carlo and an estimate of the background in data. The efficiency and purity plots show that there is no drop in efficiency until the cut value reaches 2 GeV, while the purity has increased by approximately 15 %. Figures 6.8(c) and (d) clearly show that the signal lies predominantly above 1.9 GeV, while the background makes a significant contribution below this value.

$p(D^0)/p(\pi_s) > 8$. This final cut is slightly different to the previous requirements. This is a background “shaping” cut rather than a background rejection cut and is used in the TLT DIS $D^{*\pm}$ trigger to control the number of events selected in the high ΔM shoulder, ie. to reduce the overall rate of this trigger. This is illustrated in figure 6.9 where it can be seen that the cut removes no candidates from the signal region.

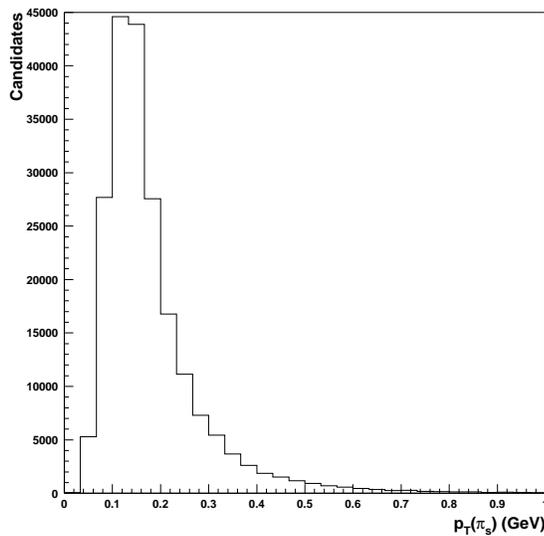


Figure 6.6: The generated transverse momentum distribution of the slow pion taken from the non-diffractive RAPGAP sample.

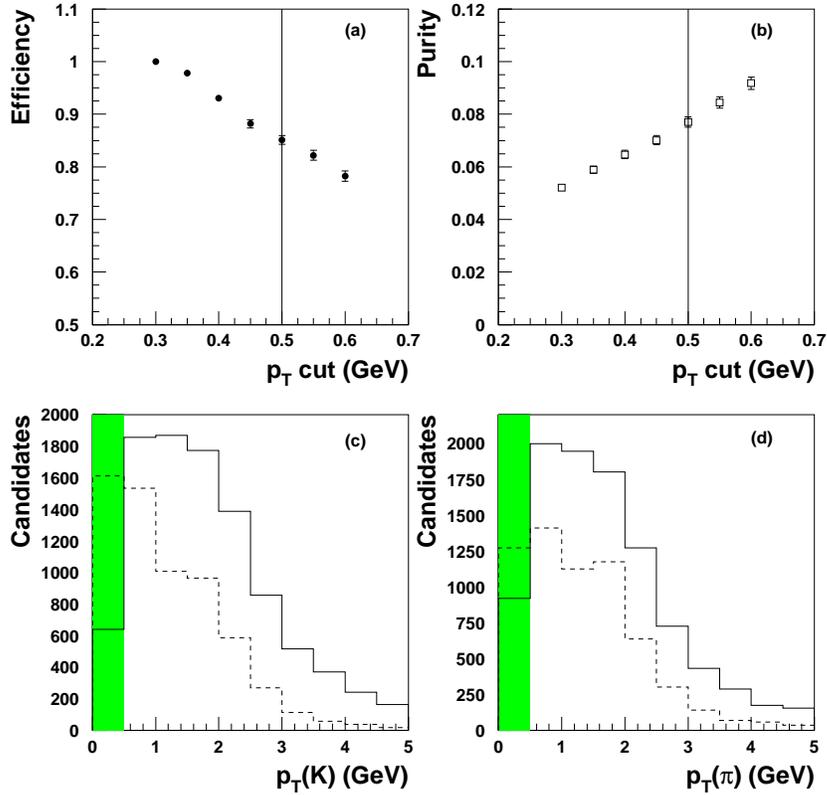


Figure 6.7: The relevant distributions for the selection of the $p_T(K, \pi)$ cut. Figure (a) shows the efficiency of $D^{*\pm}$ reconstruction versus cut value, while figure (b) shows the purity versus cut value. The definitions of efficiency and purity are given in the text. Figures (c) and (d) show the p_T distributions for K and π candidates from Monte Carlo (solid line) and from the wrong charge combinations in the data, as a background example (dotted line).

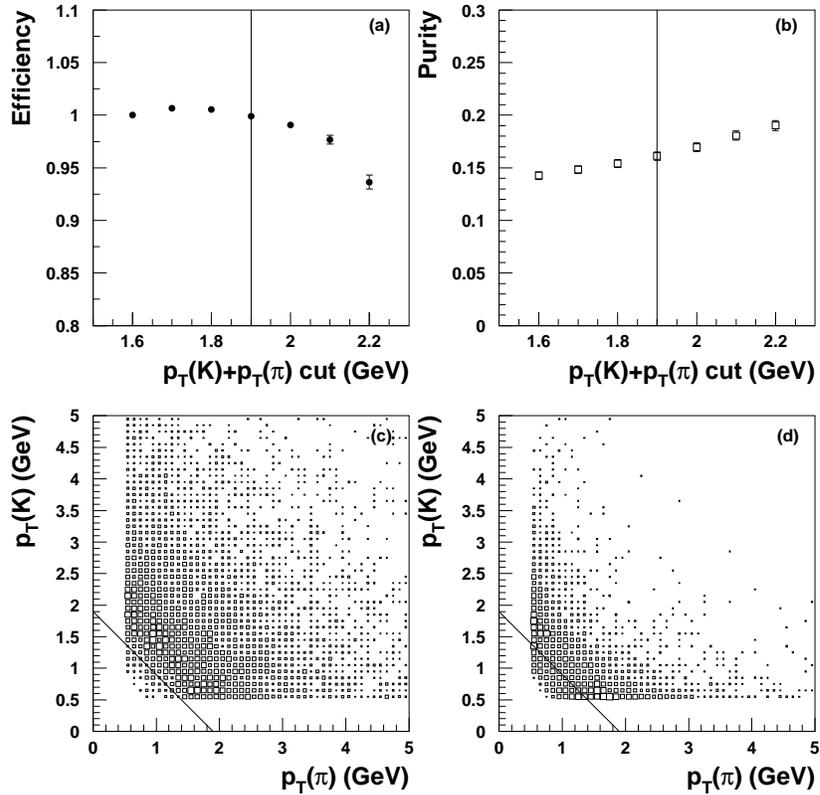


Figure 6.8: The relevant distributions for the selection of the $|p_T(K)| + |p_T(\pi)|$ cut. Figures (a) and (b) show the efficiency and purity respectively versus cut value. The definitions of efficiency and purity are similar to those used in the $p_T(K, \pi)$ cuts. Figures (c) and (d) show $p_T(K)$ versus $p_T(\pi)$ in the Monte Carlo and in the background from the data respectively. The diagonal line indicates the cut which is used in this analysis. The size of each box is proportional to the density of the points on the original scatter plot.

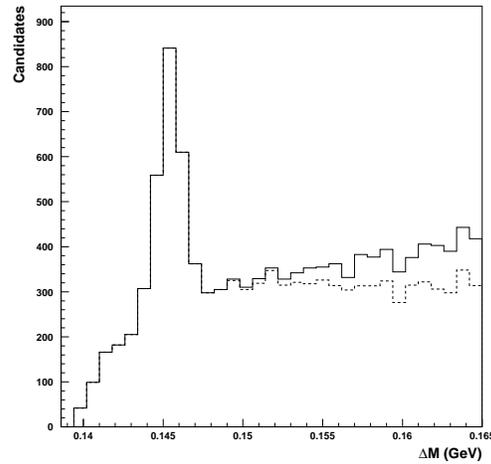


Figure 6.9: The ΔM distribution in data after the application of all cuts except the $p(D^0)/p(\pi_s)$ cut (solid line). The dashed line shows the effect of applying the $p(D^0)/p(\pi_s)$ cut.

Selection Efficiencies

The efficiencies of the different selection requirements are summarised in table 6.2 and have been calculated using the non-diffractive RAPGAP samples. The combined efficiencies are calculated by weighting each of the two efficiencies by the fraction of the total data sample they represent and then summing them.

The efficiencies are all calculated under the assumption that DIS01 is always available, so that the effect of the cuts alone can be studied, without folding in the efficiency of the trigger strategy. Line 3 of the table shows the drop in efficiency caused by using DIS03 instead. The large difference between the two samples is due to DIS03 in 1995 having the 14×14 cm box cut and in 1996/7 having the radius > 26 cm cut. The method of simulating the trigger strategy in Monte Carlo is discussed in the next chapter.

It is clear that the trigger efficiencies for this sample are fairly high. The small difference between the two Monte Carlo samples at this stage is due to the reduced dead material within the main detector volume from the start of 1996 data-taking and will be most noticeable in calorimeter-based quantities.

The most significant selection requirement after these is the identification of $D^{*\pm}$ candidates (with preselection cuts) which lie in both the signal windows. After the application

Selection cut	1995 Monte Carlo %	1996/7 Monte Carlo %	Combined %
FLT selection	94.77 ± 0.21	93.77 ± 0.15	93.92 ± 0.13
TLT selection			
DIS01	92.27 ± 0.24	92.52 ± 0.17	92.48 ± 0.15
(DIS03)	(83.33 ± 0.35)	(45.74 ± 0.32)	(51.23 ± 0.27)
All DIS selection requirements	89.12 ± 0.30	87.48 ± 0.21	87.71 ± 0.18
Candidate in the signal windows	51.37 ± 0.60	51.70 ± 0.40	51.65 ± 0.35
$p_T(\pi_s) > 0.12$ GeV	47.74 ± 0.60	48.24 ± 0.40	48.33 ± 0.35
$p_T(K, \pi) > 0.5$ GeV	40.12 ± 0.59	40.88 ± 0.40	40.77 ± 0.35
$ p_T(K) + p_T(\pi) > 1.9$ GeV	39.37 ± 0.59	40.15 ± 0.39	40.04 ± 0.35
$p(D^0)/p(\pi_s) > 8$	39.37 ± 0.59	40.15 ± 0.39	40.04 ± 0.35

Table 6.2: A Summary of the cumulative selection efficiency from each Monte Carlo sample separately and combined. The errors given are statistical.

of all the remaining cuts, the combined efficiency is approximately 40 %.

6.4.2 Invariant Mass Distributions

Figure 6.10 shows the ΔM and $M(K\pi)$ distributions after the application of all the trigger requirements and selection cuts which have been described so far in this chapter. Additional requirements have also been made on the candidates in these distributions to ensure that they lie within the kinematic region used for the cross section measurements described in the next chapter.

The number of $D^{*\pm}$ candidates can be extracted from the ΔM distribution using a variety of different methods. The method used in this analysis is to make a maximum likelihood fit to the ΔM distribution, using a Gaussian to fit the peak and a function of the form

$$f(\Delta M) = a \cdot (\Delta M - m_\pi)^b \cdot \exp(c \cdot (\Delta M - m_\pi)) \quad (6.6)$$

to describe the background. a , b and c are free parameters of the fit, as are the mean, width and number of events from the Gaussian. The exponential factor was introduced to

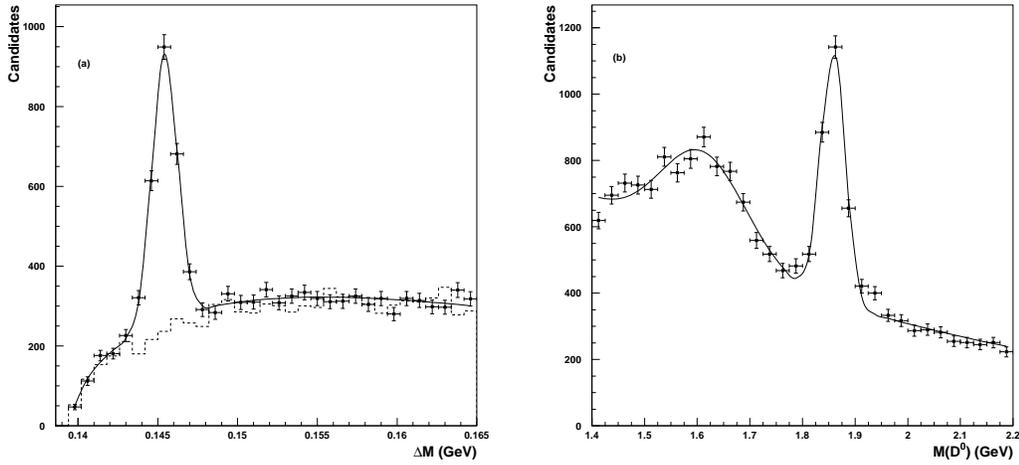


Figure 6.10: (a) The ΔM distribution after the application of all DIS and $D^{*\pm}$ selection cuts. The data exhibit a clear peak (black dots). The solid line is a fit performed to the distribution (see text) and the dashed line is an estimate of the combinatorial background using wrong charge combinations from the data. (b) The $M(D^0)$ distribution after the application of all DIS and $D^{*\pm}$ selection cuts. The distribution has been fitted (see text) and the results are indicated by the solid line. The expected two-peak structure is observed (see text for details).

describe the suppression of the high ΔM shoulder caused by the $p(D^0)/p(\pi_s)$ requirement. A fit to the ΔM distribution in figure 6.10(a) gives the number of $D^{*\pm}$ candidates as 1650 ± 64 . The mean of the ΔM peak is found to be 145.44 ± 0.03 MeV — in good agreement with the Particle Data Group (PDG) value. The width is found to be 0.79 ± 0.03 MeV, in agreement with that observed in the Monte Carlo.

An alternative method of extracting the number of $D^{*\pm}$ candidates in the ΔM distribution is through background subtraction. The background can be estimated using so-called “wrong-charge” combinations, which are three-track combinations in which the D^0 decay product candidates are chosen to have the same charge. Such combinations will have an overall charge which lies between -3 and $+3$. This background distribution is then normalised to the signal ΔM distribution away from the peak region ($0.155 < \Delta M < 0.165$ GeV). The result of this background estimation method is shown in the ΔM distribution in figure 6.10.

The $K\pi$ invariant mass distribution is also shown in figure 6.10 and contains a clear signal peak. The second broader peak is a result of $D^0 \rightarrow K^\pm \pi^\mp \pi^0$ decays from which only

the charged kaon and pion are selected, resulting in a low value of $M(K\pi)$. The $M(K\pi)$ distribution can be fitted with a simple polynomial to describe the background, plus two Gaussians. This fit gives a D^0 mass of 1857.8 ± 0.9 MeV with a width of 23 ± 1 MeV. The measured D^0 mass is lower than the particle data group value quoted earlier in this section. This effect has been observed in the $F_2^{c\bar{c}}$ analysis [16], as well as in the reconstruction of the J/ψ in the $\mu^+\mu^-$ decay channel [43], and has been attributed to a small calibration shift (-0.3%) in the magnetic field.

6.5 Diffractive Event Selection

The key characteristic of a diffractive event is the presence of a large rapidity gap between the outgoing proton, which disappears undetected down the forward beampipe, and the hadronic system X (see chapter 2) from the dissociated photon. This means that the gap will be visible in the region surrounding the forward beamhole.

Traditionally within ZEUS, such events are selected using the η_{MAX} method [132]. η_{MAX} is defined to be the pseudorapidity of the ZUFO closest to the direction of the outgoing proton. The scattered positron is not considered in the determination of η_{MAX} and only ZUFOs with energy greater than 400 MeV are considered. The 400 MeV requirement is made to avoid distortion from calorimeter noise.

The η_{MAX} distribution from the final DIS $D^{*\pm}$ sample is shown in figure 6.11. The peak at $\eta_{MAX} \sim 3.5$ is that from non-diffractive events, ie. those without a rapidity gap. The shaded area in figure 6.11 indicates those candidates which pass an $\eta_{MAX} < 2$ requirement and which lie in the kinematic region chosen for cross section measurements in the next chapter. In the recent ZEUS diffractive $D^{*\pm}$ results [59], a cut of $\eta_{MAX} < 1.5$ was used which selects a very clean sample of events containing a large rapidity gap.

Although an η_{MAX} cut of 1.5 is a good method of selecting a high purity diffractive sample, it is a very severe selection requirement which, given the limited statistics available for this analysis, is clearly not desirable. An alternative method was used in the recent ZEUS results on the measurement of the inclusive diffractive cross section [129]. Here, the $\ln M_X^2$ distribution is fitted to extract the number of diffractive events. While this method would be suitable for the measurement of the total diffractive $D^{*\pm}$ cross section, the statistics are too limited for this method to be beneficial in the measurement of differential cross

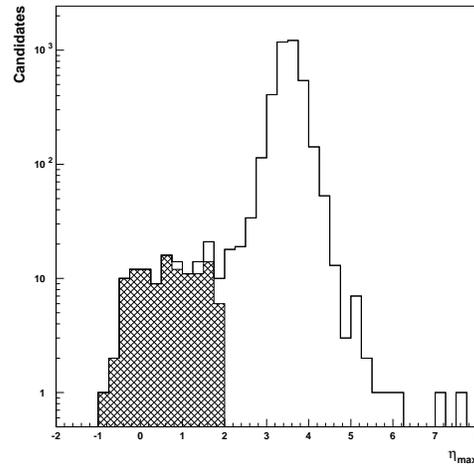


Figure 6.11: The η_{MAX} distribution for data $D^{*\pm}$ candidates which lie in both the $M(D^0)$ and ΔM signal windows. The shaded area indicates the candidates which are kept for the measurement of the cross section (see text).

sections. Thus, this method is not considered here.

The final method available is to combine the η_{MAX} selection requirement with a requirement on θ_H , where θ_H is defined by

$$\cos \theta_H = \frac{\sum_h p_z^h}{|\sum_h p^h|} \quad (6.7)$$

where h indicates the sum over all ZUFOs in the event. For non-diffractive events $\cos \theta_H$ will be peaked near 1, while for diffractive events, a large fraction of the events will have $\cos \theta_H < 1$. The $\eta_{MAX} - \theta_H$ distribution for the final DIS $D^{*\pm}$ sample is shown in figure 6.12, along with those from various different Monte Carlo samples. The conclusions which can be drawn from these plots are that, when considering an inclusive DIS sample, it is clearly beneficial to use a combination of an η_{MAX} requirement and a θ_H requirement as used, for example, in the 1993 diffractive results published by ZEUS [132]. However, as soon as a $D^{*\pm}$ candidate is also required (whether it is a genuine one or not), the η_{MAX} requirement can be loosened without any additional requirement on θ_H being necessary. Therefore for the purposes of this analysis, a cut of $\eta_{MAX} < 2$ is used without an additional θ_H requirement.

The ΔM and $M(D^0)$ distributions after the application of the η_{MAX} requirement are shown in figure 6.13. A clear signal still remains in both distributions. Using the mass fits described in the previous section, the number of events from the ΔM distribution is

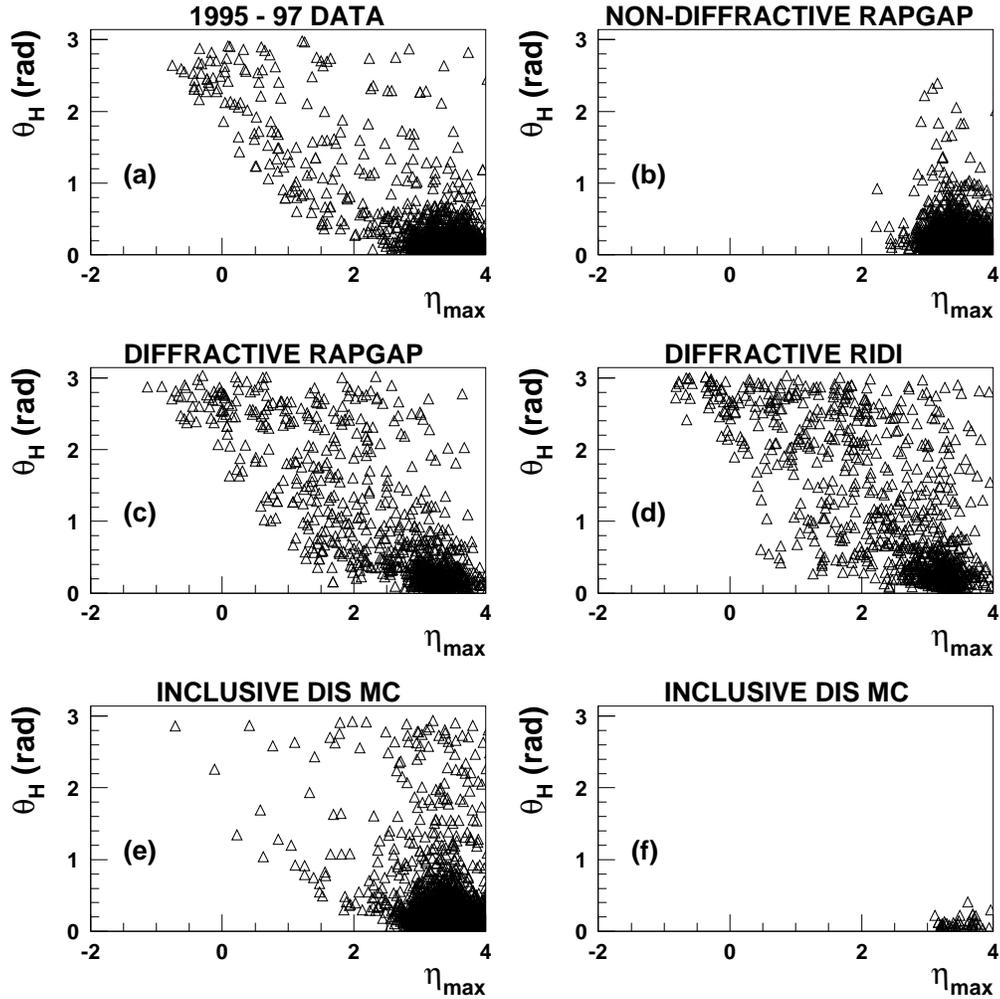


Figure 6.12: The θ_H versus η_{MAX} distribution for (a) data and (b) - (f) Monte Carlo. Figure (b) shows the distribution for non-diffractive $D^{*\pm}$ Monte Carlo, while (c) and (d) are from the two different diffractive samples. Figures (e) and (f) both come from the inclusive DIS Monte Carlo sample, but the latter has $D^{*\pm}$ finding performed, while the former does not.

found to be 82 ± 13 , the mean value of ΔM is 145.69 ± 0.17 MeV and the width of the peak is 1.11 ± 0.15 MeV. Both the width and the mean value of ΔM are found to be reasonably consistent with those observed in the DIS $D^{*\pm}$ signal. The D^0 mass from the diffractive $M(D^0)$ distribution yields a mean of 1860.4 ± 2.9 MeV, with a peak width of 18 ± 3 MeV, both of which are in good agreement with the values obtained from the DIS $D^{*\pm}$ signal. The results of the fits to both the diffractive and non-diffractive ΔM and $M(D^0)$ distributions, along with the PDG values are summarised in table 6.3.

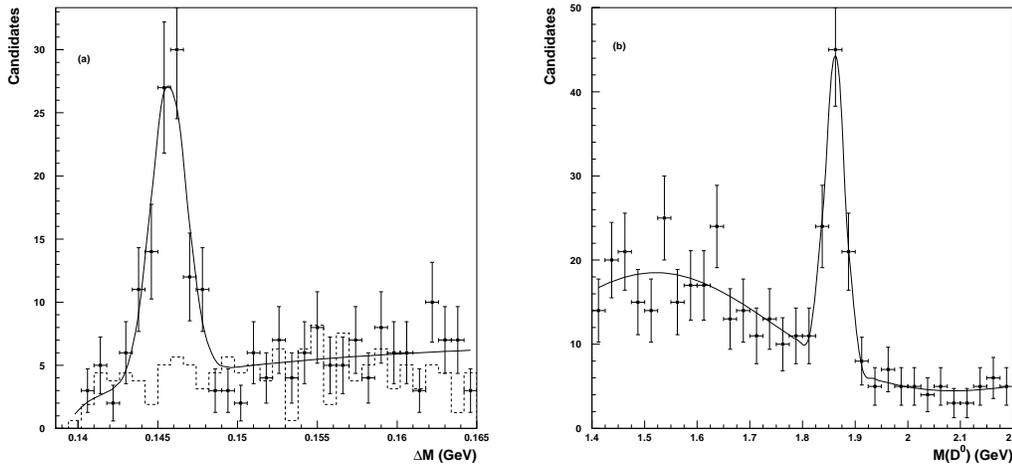


Figure 6.13: (a) The ΔM distribution for events which satisfy all diffractive selection requirements, as well as the DIS and $D^{*\pm}$ requirements. Once again, the solid line is a fit to the distribution from which the number of events has been extracted and the dashed line is an estimate of the combinatorial background made using wrong-charge combinations. (b) The $M(D^0)$ distribution for events which satisfy all selection requirements fitted in the same manner as the distribution in figure 6.10. The expected two-peak structure is not as clear due to the limited statistics available.

The characteristics of this diffractive $D^{*\pm}$ sample are discussed in the next section.

6.6 Characteristics of the Data Sample

Having established a clear signal for $D^{*\pm}$ production in events containing a large rapidity gap, it is important to ensure that the data exhibit the other characteristics expected of diffractive interactions. For example, diffractive events usually have a significantly lower

Quantity	DIS data	Diffractive DIS data	PDG/Monte Carlo
N_D	1650 ± 64	82 ± 13	N/A
ΔM	145.44 ± 0.03 MeV	145.69 ± 0.17 MeV	145.397 ± 0.030 MeV
ΔM resolution	0.79 ± 0.03 MeV	1.11 ± 0.15 MeV	~ 0.8 MeV
$M(D^0)$	1857.8 ± 0.9 MeV	1860.4 ± 2.9 MeV	1864.6 ± 0.5 MeV
$M(D^0)$ resolution	23 ± 1 MeV	18 ± 3 MeV	~ 22 MeV

Table 6.3: A Summary of fit parameters determined from the DIS and diffractive DIS ΔM and $M(D^0)$ distributions. The values of ΔM and $M(D^0)$ are those given by the Particle Data Group [18], while the resolutions are those estimated from the different Monte Carlo samples.

multiplicity than non-diffractive events, indicating that the combinatorial background in the diffractive ΔM distribution should be lower than that in the DIS ΔM distribution. Inspection of figures 6.10 and 6.13 shows this to be the case. This can be quantified by looking at the ratio of signal-to-background in the ΔM window. For the DIS ΔM signal this is $1650/2239 = 0.74$, while for the diffractive case it is $82/34 = 2.41$. It is also demonstrated by the ZUFO multiplicities, as shown in figure 6.14.

The M_X distribution for the total $D^{*\pm}$ sample should also show a low mass tail indicating the presence of diffractive events. M_X is calculated using ZUFOs and is defined by

$$M_X^2 = \left(\sum_h E^h \right)^2 - \left(\sum_h p_x^h \right)^2 - \left(\sum_h p_y^h \right)^2 - \left(\sum_h p_z^h \right)^2 \quad (6.8)$$

where h indicates the sum over all ZUFOs in the event and excludes the scattered positron. For a diffractive event, this is simply the mass of the hadronic system produced when the photon dissociates (see chapter 2). For non-diffractive events, this is the hadronic final state arising from the hard interaction, plus any of the proton remnant which does not escape down the forward beampipe. Thus, the reconstructed M_X must be corrected to account for the partial loss of any of the diffractive system X which escapes down the rear beampipe, as well as for dead material effects. This correction is determined from Monte

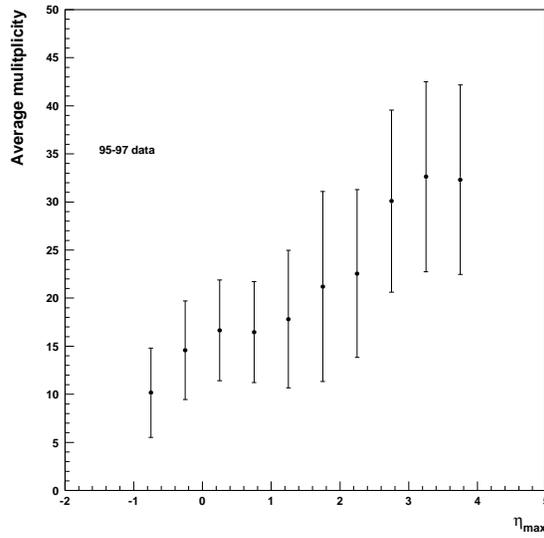


Figure 6.14: The average multiplicity versus η_{MAX} in data, calculated using ZUFOs. The expected increase in multiplicity with increasing η_{MAX} is observed.

Carlo by studying M_X^{rec}/M_X^{true} and yields a correction factor of 0.86, regardless of which Monte Carlo and which year is used. It is also found to be approximately independent of M_X , W and Q^2 . If cells are used to calculate M_X instead of ZUFOs, the correction factor is 0.70. The corrected $\ln M_X^2$ distribution in data using ZUFOs is shown in figure 6.15; the diffractive candidates all populate the low M_X region as expected. Figure 6.15 also shows the M_X versus W distribution which indicates that the diffractive candidates occupy the low M_X region over the full W range, unlike the non-diffractive candidates.

Finally figure 6.16 shows the $x - Q^2$ distribution from both diffractive and non-diffractive $D^{*\pm}$ candidates. As expected, the two sets of candidates exhibit very similar characteristics. It is, however, interesting to note that the x_p and β requirements made on the diffractive sample place an additional limit on Bjorken x of $x < 0.0128$, as discussed in the next chapter.

Figure 6.17 shows one of the diffractive $D^{*\pm}$ candidates from 1995 data. It clearly contains a substantial rapidity gap, with the forward half of the calorimeter completely devoid of hadronic activity. The event also has a very low multiplicity and a well-defined scattered positron.

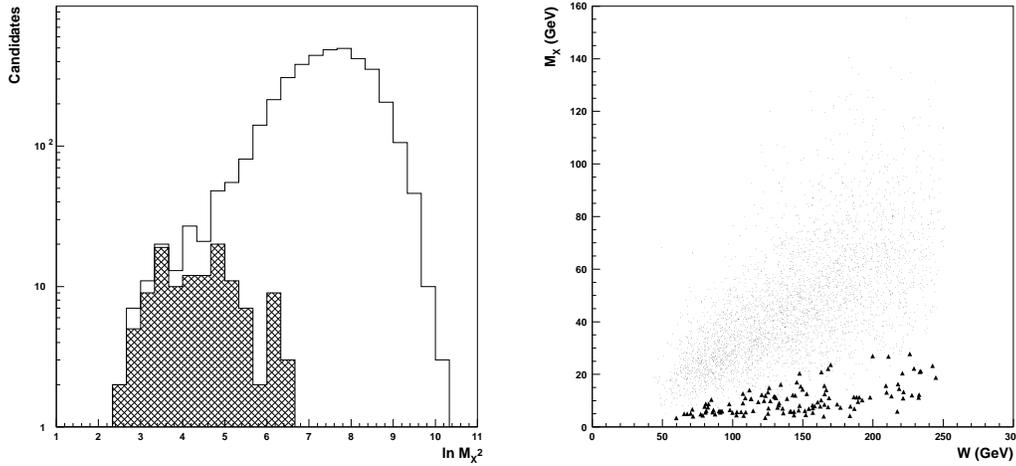


Figure 6.15: The $\ln M_X^2$ distribution and the M_X versus W distribution from the final DIS $D^{*\pm}$ data sample. The candidates which satisfy the $\eta_{MAX} < 2$ requirement and which lie in the kinematic region chosen for cross section measurements in the next chapter are marked by the hatched area in the left-hand plot and the larger triangles in the right-hand plot.

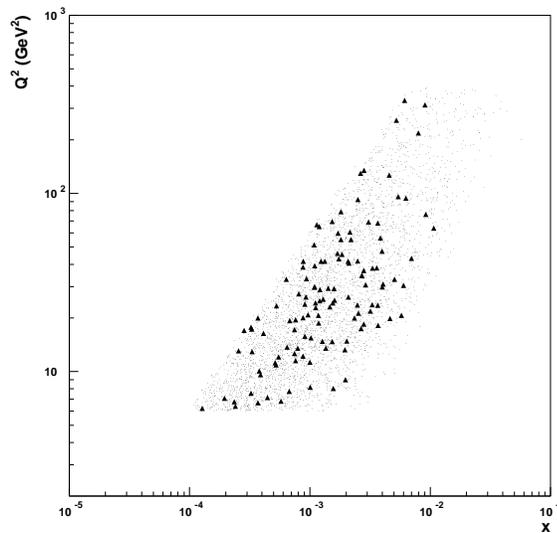


Figure 6.16: The $Q^2 - x$ distribution from the final DIS $D^{*\pm}$ sample. Those marked by triangles are the candidates which satisfy the diffractive requirements.

6.7 Summary

In this chapter we have discussed the methods by which a diffractive DIS $D^{*\pm}$ sample can be extracted from the full ZEUS data set with optimum efficiency and purity. First, the Monte Carlo samples used in the choice of the selection requirements were discussed — these will be used in the next chapter to perform the cross section unfolding. Then the trigger strategy used in this analysis was discussed in some detail, as this plays a crucial part in ensuring that the maximum possible statistics are available. This was followed by a discussion of the DIS and $D^{*\pm}$ selection requirements, along with some motivation for the cuts used.

Having reached this stage, the DIS $D^{*\pm}$ signal was studied and the number of candidates extracted using a fit to the ΔM distribution. Other methods of extracting the number of events were also reviewed. This signal will be used again in the calculation of the ratios performed in the next chapter.

Finally, the diffractive event selection requirements were discussed and used to extract the diffractive $D^{*\pm}$ signal in a similar manner to the total DIS $D^{*\pm}$ signal. The ΔM distribution was found to contain 82 ± 13 candidates which will be used in the cross section calculations performed in the next chapter.

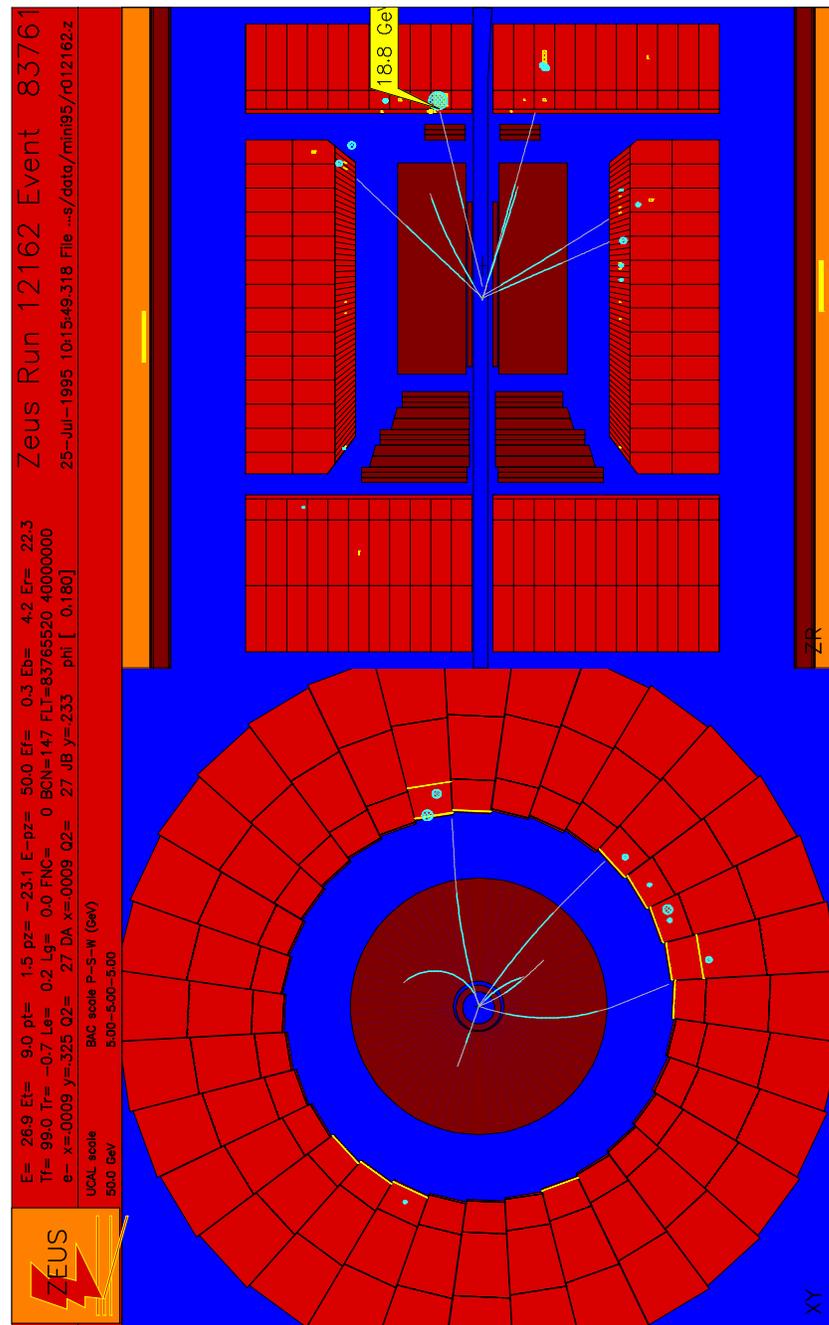


Figure 6.17: One of the diffractive $D^{*\pm}$ candidates selected from 1995 data. The r - ϕ view (left) shows the small number of tracks identified in the event, three of which make up the $D^{*\pm}$ candidate. The r - z view (right) shows that the event clearly contains a rapidity gap in the forward region of the detector. The calorimeter cluster identified as the scattered positron candidate is also indicated.

Chapter 7

Extraction of Cross Sections

In the previous chapter a $D^{*\pm}$ signal was established in the combined 1995 — 1997 data sample, corresponding to an integrated luminosity of 43.3 pb^{-1} . Through fits to the ΔM distribution the total DIS $D^* \rightarrow (K\pi)\pi_s$ signal was found to contain 1650 ± 64 candidates, which was reduced to 82 ± 13 candidates after the application of the diffractive requirements.

In this chapter, these signals will be used to extract the total diffractive $D^{*\pm}$ cross section, along with differential cross sections and the ratio of diffractive $D^{*\pm}$ production to DIS $D^{*\pm}$ production. First, the kinematic region accessible in this analysis will be discussed, then various quantities in the data will be compared to the two different diffractive Monte Carlo samples (described in the previous chapter). These Monte Carlo samples, along with the data distributions, will then be used to choose the best bins for differential cross section extraction. The remaining backgrounds will then be discussed along with how they are treated. Once all these different issues have been resolved, the cross sections and the ratios can be extracted. After a discussion and calculation of the systematic errors, the differential cross sections are compared to various different theoretical models.

7.1 The Kinematic Region

The kinematic range accessible within any analysis is constrained by several different factors, such as detector acceptance or event selection method. It is also important to ensure that Monte Carlo input is minimised in order to avoid any bias from the theoretical

model on which the Monte Carlo is based. This last point is particularly important in an analysis such as this one, where one of the principal goals is to distinguish between the different theoretical models on offer.

The kinematic region in which the cross sections are extracted is

$$\begin{aligned}
 6 < Q^2 < 400 \text{ GeV}^2 \\
 0.02 < y < 0.7 \\
 p_T(D^{*\pm}) > 1.5 \text{ GeV} \\
 -1.5 < \eta(D^{*\pm}) < 1.5 \\
 0.001 < x_{\mathcal{P}} < 0.016 \\
 0 < \beta < 0.8
 \end{aligned}$$

The Q^2 limits are dictated by the Q^2 region in which diffractive $D^{*\pm}$ candidates are observed. In the case of the lower limit, although the geometrical acceptance of the calorimeter would allow a lower Q^2 limit of 1 GeV^2 , a decrease of only one signal candidate is observed when the Q^2 limit is moved from 1 to 6, indicating that very few diffractive $D^{*\pm}$ candidates can be measured in this region. Equally the majority of the data lie below $Q^2 \simeq 400 \text{ GeV}^2$, indicating this to be an appropriate choice for the upper limit.

The accessible y region for this analysis is defined by various aspects of the DIS event selection. In the case of the lower limit, it is set by the $y_{JB} > 0.02$ requirement, while the upper limit is set by the $E'_e > 8 \text{ GeV}$ requirement, which is equivalent to setting $y < 0.7$.

The constraints on $p_T(D^{*\pm})$ and $\eta(D^{*\pm})$ arise from the selection method — in particular, the cuts applied to the slow pion candidate limit the accessible p_T and η range due to the high correlation between the $D^{*\pm}$ and the slow pion. The overall effect is illustrated in figure 7.1, which shows the acceptance versus $p_T(D^{*\pm})$ and $\eta(D^{*\pm})$ and clearly indicates the reason for the choice of kinematic range.

The upper limit on β is motivated partly by the data themselves and partly by the effect of the M_X cut-off. Due to the requirement that each event must contain a $c\bar{c}$ pair, a minimum is set on M_X of approximately twice the $D^{*\pm}$ mass, ie. $M_X^{\text{min}} \simeq 4 \text{ GeV}$. When this is combined with the upper Q^2 limit ($Q^2 = 400 \text{ GeV}^2$) this places a maximum on β

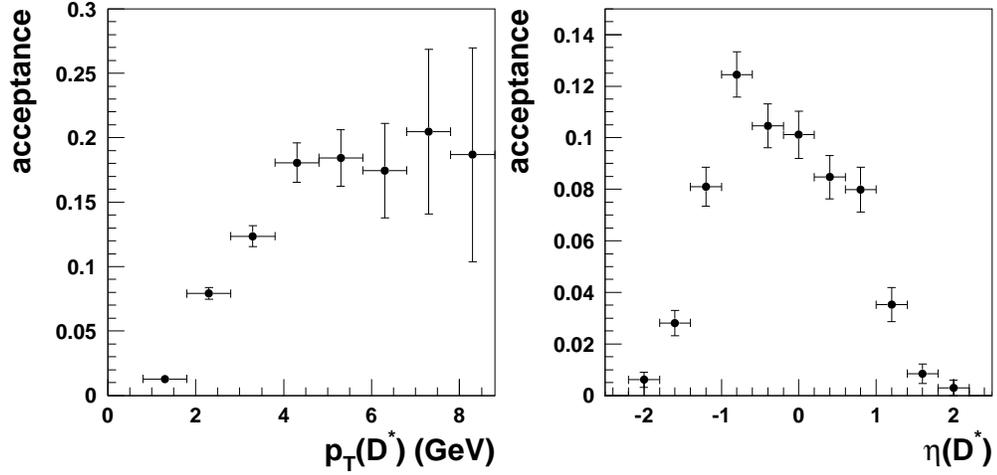


Figure 7.1: The acceptances versus $p_T(D^{*\pm})$ and $\eta(D^{*\pm})$ from the diffractive RAPGAP sample. The plots indicate that $p_T(D^{*\pm}) > 1.5$ GeV and $|\eta(D^{*\pm})| < 1.5$ is an acceptable kinematic region for the cross section measurements.

of

$$\beta_{max} = \frac{Q_{max}^2}{M_X^2(min) + Q_{max}^2} = 0.96$$

However, when the $x_P - \beta$ distributions for both data and Monte Carlo (figures 7.3(a) and 7.2 respectively) are inspected it becomes apparent that no candidates populate the $\beta > 0.8$ region — hence the choice of upper limit.

The choice of upper x_P limit is heavily influenced by the choice of η_{MAX} cut, as can be seen in figure 7.2. In order to study this effect more closely the efficiency as a function of x_P cut can be used, where the efficiency is defined as

$$\frac{\text{no. reconstructed } D^{*\pm} \text{ candidates with } \eta_{MAX} < 2 \text{ and } x_P < x_P^{cut}}{\text{no. generated } D^{*\pm} \text{ candidates with } x_P < 0.1}$$

Figure 7.3(b) shows this efficiency as a function of x_P^{cut} and indicates that the efficiency saturates at the point where every candidate which passes the η_{MAX} requirement also lies within the chosen x_P region.

It would appear from figure 7.3(b) that a suitable choice for the upper limit on x_P would be $x_P < 0.025$. However, it is found that a decrease of only four signal candidates is observed in the data if the cut on x_P is moved from 0.025 to 0.016. This is a result of the low acceptance in this region which, in turn, is a result of the η_{MAX} requirement. In

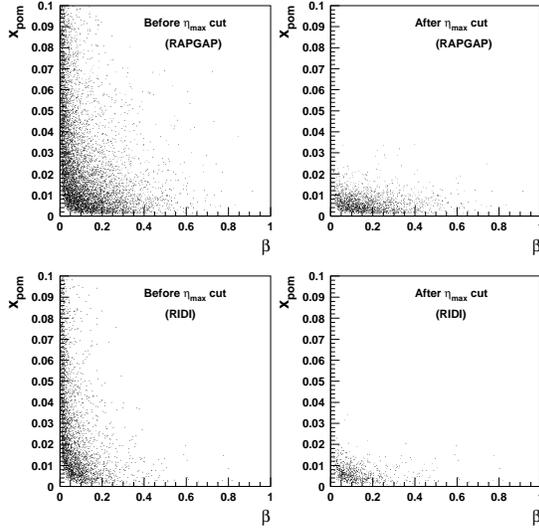


Figure 7.2: The $x_{\mathcal{P}} — \beta$ distributions from the diffractive RAPGAP and RIDI Monte Carlo samples, before and after the application of the η_{MAX} cut. There is a substantial effect on the accessible $x_{\mathcal{P}}$ range. It should also be noted that x_{pom} is equivalent to $x_{\mathcal{P}}$ in all subsequent figures.

order to avoid entering a region of phase space in which the acceptance is low, the upper limit on $x_{\mathcal{P}}$ is chosen to be 0.016.

The lower limit on $x_{\mathcal{P}}$ is kinematically given by

$$x_{\mathcal{P}}^{min} = \frac{Q_{min}^2 + M_x^2(min)}{Q_{min}^2 + W_{max}^2} = \frac{6 + 16}{6 + 62500} \simeq 0.00035$$

However, all the data are found to lie above $x_{\mathcal{P}} \simeq 0.001$. This is caused by two factors: the choice of β_{max} and the Q^2 distribution in the data. From figure 6.16, it can be seen that all but a small number of the diffractive candidates lie below $Q^2 \simeq 250 \text{ GeV}^2$ which, when combined with $\beta_{max} = 0.8$, implies an effective minimum M_X^2 of

$$M_X^2(min) = \frac{Q_{max}^2(1 - \beta_{max})}{\beta_{max}} \simeq 50 \text{ GeV}^2$$

If this value of $M_X^2(min)$ is then used to determine $x_{\mathcal{P}}^{min}$, it is found to be

$$x_{\mathcal{P}}^{min} \simeq \frac{6 + 50}{6 + 62500} = 0.0009$$

It is also observed that the candidates which have $Q^2 > 250 \text{ GeV}^2$ also have $M_X \gg M_X(min)$, resulting in values of $x_{\mathcal{P}}$ which are significantly larger than $x_{\mathcal{P}}^{min}$. Thus a minimum requirement of $x_{\mathcal{P}} = 0.001$ is chosen. These effects are also observed in the diffractive Monte Carlo samples.

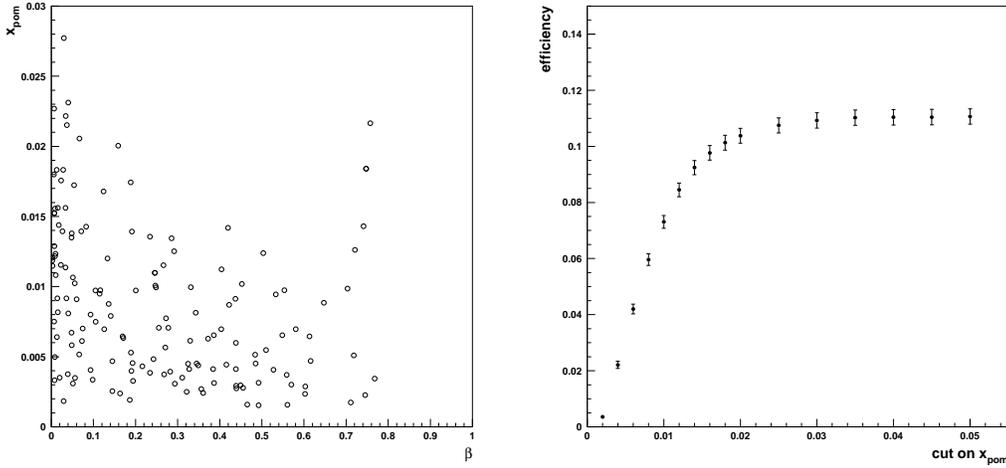


Figure 7.3: (a) The $x_P - \beta$ distribution from the data, indicating that the majority of candidates lie within the chosen kinematic region, (b) the efficiency versus x_P cut determined using the diffractive RAPGAP sample (see text for details).

7.2 Data — Monte Carlo Comparisons

Before either the RAPGAP or RIDI samples can be used with confidence to unfold the cross section, it must be ensured that they describe a wide variety of different quantities in the data — in particular, those quantities on which selection cuts are placed. In order to compare data and Monte Carlo distributions, only those candidates which satisfy both mass windows are considered. In addition, the data must have the combinatorial background subtracted. This is carried out using appropriately normalised wrong charge combinations (see the previous chapter for details). The Monte Carlo must also be normalised to the data to allow comparisons to be made. Usually each Monte Carlo sample is normalised to the data based on the luminosities which the data and Monte Carlo samples represent. This method is used because it provides information relating to both distribution shapes and the production rate. The luminosity of each Monte Carlo sample is calculated as follows: each sample contains N_D events, each of which contain at least one $D^{*\pm} \rightarrow K\pi\pi_s$ with $p_T(D^{*\pm}) > p_T^{min}$, and has a charm production cross section, $\sigma_{c\bar{c}}$, calculated by the Monte Carlo generator. p_T^{min} is the minimum p_T the $D^{*\pm}$ must have in order to be kept during the generation of the Monte Carlo samples. For the RAPGAP samples $p_T^{min} = 0.8$ GeV and for the RIDI samples $p_T^{min} = 0$. The luminosity, \mathcal{L} , is then

given by

$$\mathcal{L} = \frac{N_D}{2 \cdot f(c \rightarrow D^{*\pm}) \cdot BR(D^{*\pm} \rightarrow K\pi\pi_s) \cdot \sigma_{c\bar{c}} \cdot \alpha(p_T > p_T^{min})} \quad (7.1)$$

where $f(c \rightarrow D^{*\pm})$ is the probability for a charm quark to fragment to a $D^{*\pm}$ which is set to 0.26 in the Monte Carlo generators, and $BR(D^{*\pm} \rightarrow K\pi\pi_s)$ is the appropriate branching ratio, set to 0.037 in the generators. This value for the branching ratio has since been superseded by a more accurate measurement, which gives $BR(D^{*\pm} \rightarrow K\pi\pi_s) = 0.026 \pm 0.001$ [18]. This value will therefore be used for the rest of the results presented in this chapter. This will not make a significant difference in the Monte Carlo, as the way the $D^{*\pm}$ samples are selected effectively sets the branching ratio to 1 (to a good approximation). $\alpha(p_T > p_T^{min})$ is the fraction of events containing a $D^{*\pm}$ with $p_T(D^{*\pm}) > p_T^{min}$ in a sample generated with $p_T^{min} = 0$. For the RIDI samples $\alpha = 1$, as p_T^{min} was set to zero during the generation of the original sample. In order to determine α for the RAPGAP samples, a sample of 10,000 events was generated under the same conditions as the RAPGAP samples used for cross section unfolding, except that p_T^{min} was set to zero. It was found that 7183 of these events contained a $D^{*\pm}$ with $p_T(D^{*\pm}) > 0.8$ GeV, implying that $\alpha(p_T > 0.8) \simeq 0.72$ for the RAPGAP samples.

Given that the RAPGAP sample contains 140,000 events and has a cross section of $\sigma_{c\bar{c}} = 46.13$ nb, this corresponds to an integrated luminosity of 308.87 pb^{-1} . The RIDI sample contains 60,000 events and has $\sigma_{c\bar{c}} = 27.37$ nb, giving an integrated luminosity of 161.10 pb^{-1} . When the data and the two Monte Carlo samples are compared, after luminosity normalisation has been performed, it is found that both Monte Carlo samples contain approximately double the number of $D^{*\pm}$ candidates observed in the data. This is due to the Monte Carlo cross sections being too large in comparison to the data rather than from differences in the shapes of the distributions. This conclusion is based on figures 7.4 and 7.5 in which comparisons are made between the data and the two Monte Carlo samples after each Monte Carlo sample has been normalised to the number of data $D^{*\pm}$ candidates in each distribution. All distributions are shown for $Q^2 < 300 \text{ GeV}^2$ due to the upper Q^2 limit on the RIDI sample. There is reasonably good agreement in the shapes of the distributions in figure 7.4, indicating that the discrepancies observed when the Monte Carlo is normalised to the data luminosity must arise from an overestimation of the cross sections in the Monte Carlo generators. It also indicates that the samples are acceptable for use in cross section unfolding. The agreement appears to be slightly better between the data and the RAPGAP Monte Carlo samples than between the data and the RIDI

Monte Carlo samples, so the RAPGAP samples are used to determine the nominal cross section, while RIDI is used as a systematic check.

When figure 7.5 is considered however, there is a clear discrepancy in the shape of the β distribution with the data favouring a harder distribution than is present in either Monte Carlo sample. For this reason differential cross sections with respect to β will not be determined. Possible explanations for the data β distribution will be discussed in section 7.8.

7.3 Bin Selection

In order to extract differential cross sections, it is necessary to select appropriate bins. The choice of bins is dictated by two factors: the resolution of the different variables and the number of candidates in each bin. Due to the limited statistics involved in this analysis, the latter factor is the dominant one. The aim is to select bins which contain equal numbers of $D^{*\pm}$ candidates. Thus three bins in each variable, containing approximately 25 signal candidates, are chosen which have widths which are larger than the resolution on each variable. This latter requirement ensures that the cross sections are not distorted by migrations.

Suitable bin choices for the single differential cross sections which give ~ 25 signal candidates in each bin are

- Q^2 : 6 — 17 — 28 — 400
- W : 50 — 120 — 160 — 250
- x_F : 0.001 — 0.0045 — 0.009 — 0.016
- $p_T(D^{*\pm})$: 1.5 — 2.2 — 3.3 — 6.0
- $\eta(D^{*\pm})$: -1.5 — -0.7 — 0 — 1.5

The ΔM distributions in each of these bins, together with fits to signal plus background, are shown in figure 7.6. The fits are of the same functional form as those described in section 6.4.2.

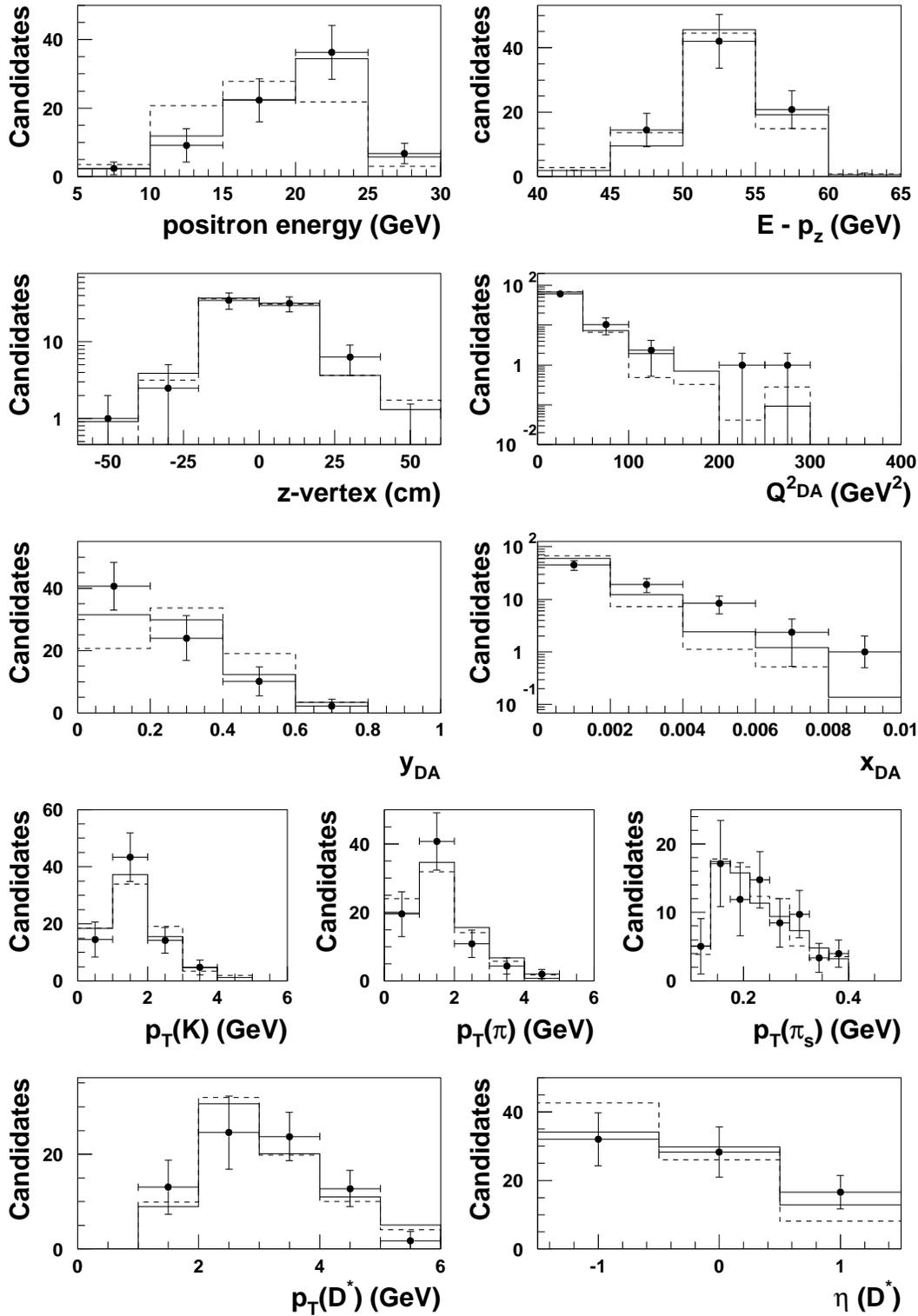


Figure 7.4: Comparisons between data (black dots), RAPGAP (solid lines) and RIDI (dotted lines) Monte Carlo for a wide variety of DIS and $D^{*\pm}$ quantities.

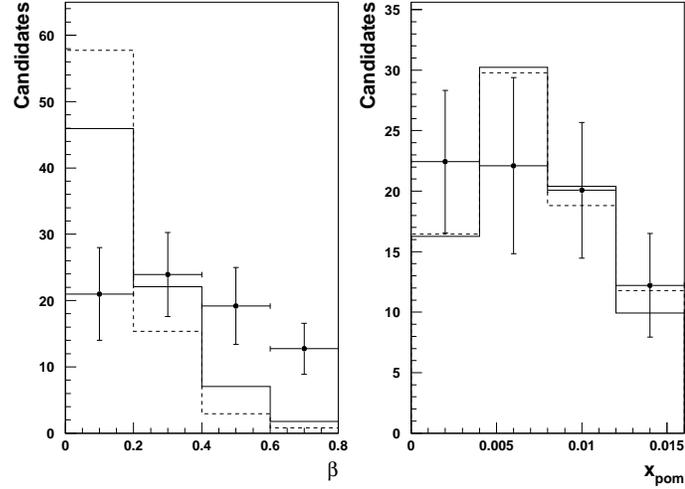


Figure 7.5: A comparison of the β and x_P distributions in data and Monte Carlo. The data is indicated by black dots, while RAPGAP is represented by the solid lines and RIDI by the dotted lines. See text for discussion of the distribution shapes.

Having chosen these bins, the resolution must be checked for each one. Figure 7.7 shows an estimation of the resolution for each bin and indicates that all bins are large enough to ensure that bin-to-bin migrations are small. This can be further investigated by determining the purity in each bin where the purity, p , is defined as

$$p = \frac{\text{no. of } D^{*\pm} \text{ candidates rec. and gen. in bin}}{\text{no. of } D^{*\pm} \text{ reconstructed in bin}}$$

This has been calculated for each cross section bin and is also illustrated in figure 7.7. The purities are very high in all except the highest x_P bin, but even this has an acceptable level of purity, indicating that reasonable bins have been chosen.

7.4 Backgrounds

It is important to check the level of different background contributions after all selection requirements have been applied. The three key backgrounds which must be considered are from photoproduction processes, non-diffractive $D^{*\pm}$ production and proton dissociative events.

Studies have been performed of photoproduction background in the context of the extraction of $F_2^{c\bar{c}}$ from a DIS $D^{*\pm}$ sample [16]. The studies indicate that photoproduction

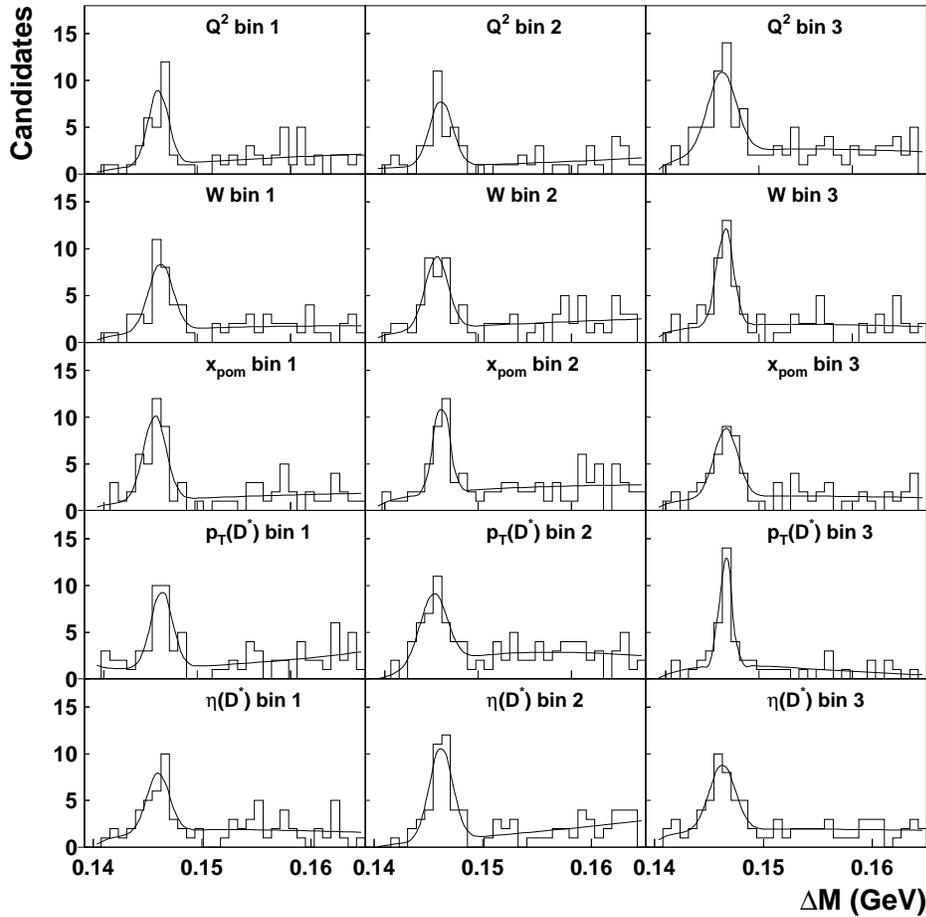


Figure 7.6: The ΔM distributions in each of the cross sections bins. The bin numbers increase with the bin quantity. The curves indicate maximum likelihood fits to each distribution. The numbers of $D^{*\pm}$ candidates are given in table 7.1.

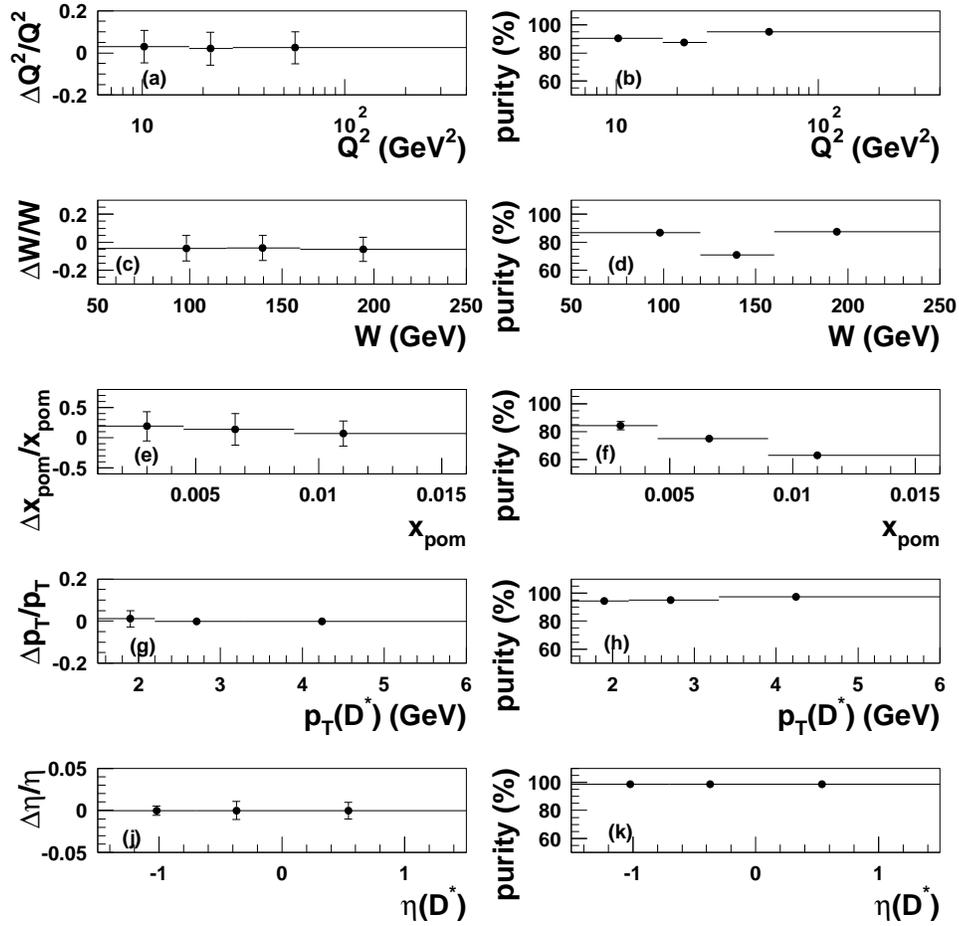


Figure 7.7: The resolutions and purities in each cross section bin. The left-hand column of plots shows $(X_{rec} - X_{true})/X_{true}$ in each of the cross section bins. The error bars are the width of the distribution in each bin, giving an estimate of the resolution for each bin. The central point gives an estimate of the accuracy with which the quantity is reconstructed. The right-hand column of plots shows the purity of each cross section bin.

contamination is less than 1 %, much lower than the level of the statistical errors on the cross section measurement. The photoproduction background in a diffractive sample is expected to be no worse than in the full DIS $D^{*\pm}$ sample and so this source of background is neglected.

Figure 7.8 shows the η_{MAX} distribution from the non-diffractive RAPGAP Monte Carlo sample. This corresponds to an integrated luminosity of 106.30 pb^{-1} and indicates that no events from this sample would pass all the selection requirements. Thus this source of background may also be safely ignored.

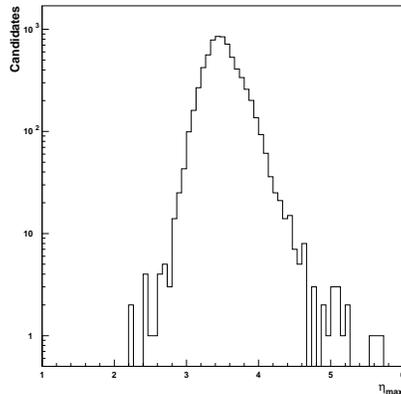


Figure 7.8: The reconstructed η_{MAX} distribution from the non-diffractive RAPGAP Monte Carlo sample.

The final source of background to be considered is from diffractive proton-dissociative events. These are events which contain a colourless exchange (and hence a rapidity gap), but in which the proton breaks up. When the proton dissociates into a sufficiently high mass state it will deposit energy in the forward calorimeter and be removed by the η_{MAX} cut. However, if the proton dissociates into a low mass state, it may still pass undetected down the forward beampipe. The size of this background contribution has been estimated by comparing diffractive results in a separate experiment in which the proton was detected in the leading proton spectrometer (LPS) with results from the most recent diffractive cross section measurements using the $\ln M_X^2$ method, which contain both single and double dissociative diffractive events. The proton dissociation contribution was determined to be $31 \pm 15 \%$ [129]. This is clearly the most significant background for this analysis and must therefore be subtracted from all cross sections.

7.5 Cross Section Extraction

In order to turn the measured number of $D^{*\pm}$ candidates into a cross section, the effects of finite detector resolutions and the inefficiencies of the trigger system, offline reconstruction and selection must be corrected for. This is done by using Monte Carlo to calculate the acceptance, which is given by

$$\varepsilon = \frac{\text{no. of reconstructed } D^{*\pm} \text{ candidates}}{\text{no. of generated } D^{*\pm} \text{ candidates}} \quad (7.2)$$

for each bin as well as for the full kinematic region. The true number of $D^{*\pm}$ candidates is then given by $N_D^{true} = N_D^{meas} / \varepsilon$. This quantity is calculated using the diffractive RAPGAP Monte Carlo samples. In order to determine the acceptance, a method of simulating the TLT trigger strategy must be sought. In order to simulate the inclusive DIS TLT trigger slots, it is sufficient to use the appropriate box cuts, as the other offline cuts are more strict than those implemented in the trigger algorithms. The DIS TLT slots have to be treated in this way, as in the 1997 version of ZGANA the medium Q^2 slot is included with the radius cut and therefore cannot be used to represent the medium Q^2 slot with the box cut which was used during 1996.

Using the two diffractive Monte Carlo samples, the trigger strategy is simulated as described below.

1995 Monte Carlo

FLT Require FLT30 or FLT44 to have fired.

TLT Positron detected in (see figure 7.9 for regions):

region A: weight event by $\mathcal{L}_{DIS01} / \mathcal{L}_{95} = 2.4/6.3$

region B or C: event weight is set to 1.

The weights for the TLT are chosen in this manner because a positron can be detected in region A only when the low Q^2 TLT slot (DIS01) is available. A positron detected in region B or C will pass either DIS01 or DIS03 and therefore the event weight is 1.

1996/7 Monte Carlo

FLT Require FLT44 or FLT46 to have fired. The relative efficiency of FLT46 is suffi-

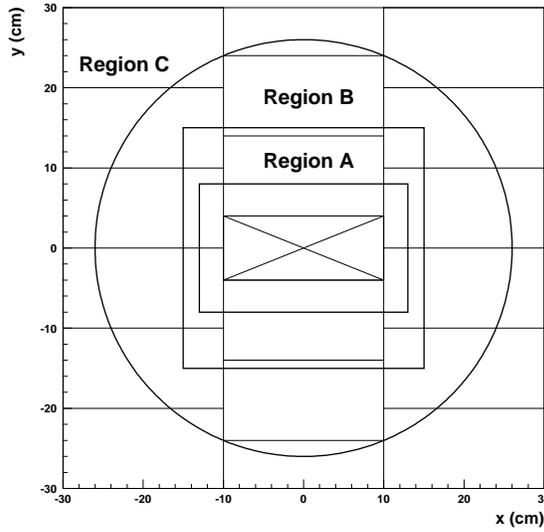


Figure 7.9: The three regions into which the face of the RCAL is divided for the simulation of the TLT trigger strategy in Monte Carlo.

ciently high that it can be assumed that this is no different to requiring FLT30 to have fired instead (see chapter 4 for more details).

TLT Positron detected in:

region C: accept event with weight = 1.

region B: generate a random number from a flat distribution between 0 and 1. If this number is less than $\mathcal{L}_{DIS01+DIS03}/\mathcal{L}_{1996+7} = 10.3/37.0$ then accept the event. If the number is greater than $\mathcal{L}_{DIS03}/\mathcal{L}_{1996+7}$ then only accept if DIS13 has fired.

region A: generate a random number as before. If the number is less than $\mathcal{L}_{DIS01}/\mathcal{L}_{1996+7} = 4.6/37.0$ then accept the event. If the number is greater than $\mathcal{L}_{DIS01}/\mathcal{L}_{1996+7}$, then only accept the event if DIS13 has fired.

The total acceptance, ε_{tot} , is then defined as

$$\varepsilon_{tot} = \frac{\mathcal{L}_{95}}{\mathcal{L}_{95+96+97}} \cdot \varepsilon_{95} + \frac{\mathcal{L}_{96+97}}{\mathcal{L}_{95+96+97}} \cdot \varepsilon_{96+97} \quad (7.3)$$

Figure 7.10 shows the acceptance in the cross section bins, while the total acceptance for the defined kinematic region is 20.36 ± 0.57 %. The cross sections can then be calculated using

$$\sigma = \frac{N_D^{meas}}{\mathcal{L} \cdot BR \cdot \varepsilon}$$

where N_D^{meas} is the number of signal candidates determined by fitting the appropriate ΔM distribution, $\mathcal{L} = 43.3 \pm 1.0 \text{ pb}^{-1}$, the total 1995-97 integrated luminosity and $BR = 2.6 \pm 0.1 \%$ [18] is the combined branching ratio for the $D^{*\pm} \rightarrow K\pi\pi_s$ decay channel.

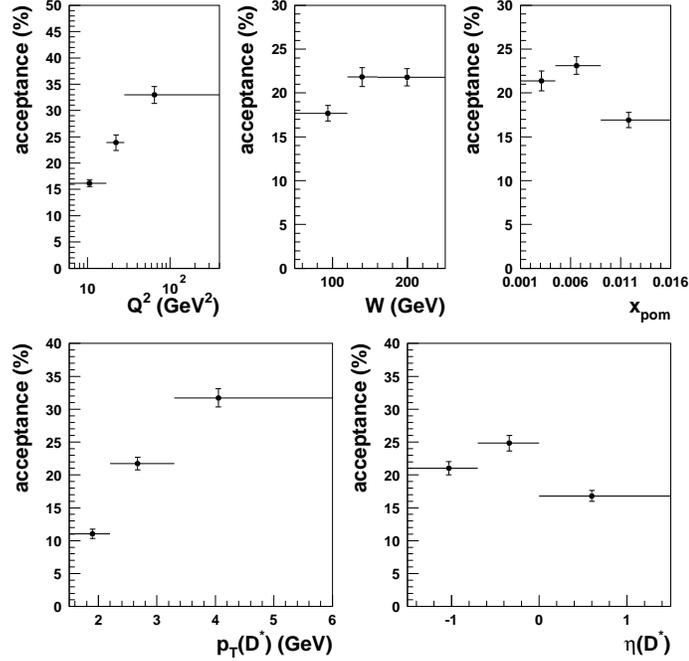


Figure 7.10: The acceptances in the different cross section bins calculated using the diffractive RAPGAP Monte Carlo sample.

Using all these numbers, the cross section in the kinematic region defined in section 7.1 is

$$\sigma(ep \rightarrow eD^{*\pm}XY) = 355 \pm 57 \text{ pb} \quad (7.4)$$

where X indicates the final state hadronic system produced at the photon-pomeron vertex and Y indicates the final state hadronic system produced at the pomeron-proton vertex (see chapter 2). The latter system can either be an elastically scattered proton or a hadronic system with low mass ($M_Y < 4 \text{ GeV}^2$ [132]). The error on the cross section is determined from the errors on N_D^{meas} , ε and the branching ratio. The cross section is dominated by the error on the number of signal $D^{*\pm}$ candidates observed in the data, N_D^{meas} . The error on the luminosity is small ($\sim 2.3 \%$) compared to all other sources of systematic error and is therefore neglected.

After subtracting the proton dissociative background (see previous section), the final cross

section is

$$\sigma(ep \rightarrow eD^{*\pm}Xp) = 245 \pm 57 \text{ pb} \quad (7.5)$$

The error on the proton dissociation background fraction has not been included in the evaluation of the statistical error on the cross section. This is because the error arises from the combined statistical and systematic errors on two separate sets of results, as described in section 7.4. The error on the cross section arising from the uncertainty on the fraction of proton dissociation background will be treated as a systematic error and will be discussed in this context in the next section.

The differential cross sections and the numbers of measured $D^{*\pm}$ candidates and acceptances which go in to the determination of the cross sections are summarised in table 7.1. The errors on the number of $D^{*\pm}$ candidates shown in the table are determined from the fit to the ΔM distribution in each bin, while the error on the acceptance is simply a reflection of the statistics of the RAPGAP Monte Carlo samples. The errors on the cross sections are determined from the errors on the number of $D^{*\pm}$ candidates in the data and the acceptance in each bin, along with the error on the branching ratio. As before the error on the luminosity has been ignored. The proton dissociation background has been subtracted from all of the cross sections quoted in table 7.1.

The column labelled “ $\langle X \rangle$ ” in table 7.1 contains the mean value of X for each bin, where X is Q^2 , W , x_F , $p_T(D^{*\pm})$ or $\eta(D^{*\pm})$. The values quoted in table 7.1 are taken from the data before any acceptance correction has been applied and are used in all subsequent cross section figures. This is a reasonable choice for the mean values of each quantity, as can be seen by inspecting the numbers in table 7.2. This shows the mean values for the Q^2 and W bins determined from the data and the RAPGAP Monte Carlo before any acceptance corrections have been applied as well as from the Monte Carlo generator level. The final column gives the “corrected” data values which are calculated by assuming

$$\frac{\text{data uncorrected}}{\text{Monte Carlo uncorrected}} = \frac{\text{data corrected}}{\text{Monte Carlo corrected}}$$

which should give a good approximation of the data value after acceptance correction. All the values quoted in table 7.2 are reasonably close, with the possible exception of the highest Q^2 bin. This is however, a very wide bin with low statistics and therefore a large uncertainty on the mean value is to be expected. When comparisons are made between the cross sections and the predictions from theoretical models, $d\sigma/dQ^2$ is displayed on a log scale, minimising any uncertainty on the mean value of Q^2 .

Bin (X)	No. of $D^{*\pm}$ Candidates	Acceptance (%)	Cross Section (pb)	$\langle X \rangle$
Q^2 : 6 — 17 GeV ²	23.6 ± 6.1	16.18 ± 0.64	89 ± 23	10.6 GeV ²
Q^2 : 17 — 28 GeV ²	23.3 ± 5.9	23.89 ± 1.49	59 ± 16	22.2 GeV ²
Q^2 : 28 — 400 GeV ²	35.6 ± 10.8	32.95 ± 1.60	66 ± 20	66.9 GeV ²
W : 50 — 120 GeV	25.3 ± 8.1	17.69 ± 0.90	87 ± 28	94.2 GeV
W : 120 — 160 GeV	25.6 ± 7.3	21.82 ± 1.07	71 ± 21	139.5 GeV
W : 160 — 250 GeV	25.5 ± 7.0	21.79 ± 0.99	71 ± 20	199.0 GeV
$x_{\mathcal{P}}$: 0.001 — 0.0045	29.8 ± 7.5	21.38 ± 1.12	85 ± 22	0.0031
$x_{\mathcal{P}}$: 0.0045 — 0.009	22.3 ± 6.5	23.13 ± 0.99	59 ± 17	0.0066
$x_{\mathcal{P}}$: 0.009 — 0.016	25.5 ± 7.2	16.93 ± 0.87	92 ± 27	0.0118
$p_T(D^{*\pm})$: 1.5 — 2.2 GeV	25.9 ± 6.5	11.04 ± 0.71	143 ± 37	1.90 GeV
$p_T(D^{*\pm})$: 2.2 — 3.3 GeV	29.5 ± 9.3	21.71 ± 0.96	83 ± 27	2.67 GeV
$p_T(D^{*\pm})$: 3.3 — 6.0 GeV	20.8 ± 6.0	31.73 ± 1.37	40 ± 12	4.05 GeV
$\eta(D^{*\pm})$: -1.5 — -0.7	20.6 ± 6.8	21.03 ± 0.99	60 ± 20	-1.03
$\eta(D^{*\pm})$: -0.7 — 0.	33.9 ± 8.6	24.83 ± 1.18	83 ± 22	-0.34
$\eta(D^{*\pm})$: 0. — 1.5	26.6 ± 7.5	16.80 ± 0.84	96 ± 28	0.60

Table 7.1: The cross sections in the different bins chosen for the differential distributions.

It is important to note that the RAPGAP sample does not contain radiative corrections. This is expected to be, at most, a $\sim 10\%$ effect [132]. This is smaller than the statistical error on the cross sections, but should be kept in mind.

7.6 Systematic Error Evaluation

Several different sources of systematic error have been considered, principally those arising from the variation of the selection cuts. For each check the analysis was repeated and the resulting cross section compared with the nominal values calculated in the previous section. The 24 checks which were made are listed below.

1. Increase E'_e cut to 9 GeV.
2. Decrease E'_e cut to 7 GeV.

Bin (X)	Mean Data (uncorrected)	Mean MC (uncorrected)	Mean MC (corrected)	Mean Data (corrected)
Q^2 : 6 — 17 GeV ²	10.6 GeV ²	10.3 GeV ²	10.3 GeV ²	10.6 GeV ²
Q^2 : 17 — 28 GeV ²	22.2 GeV ²	22.1 GeV ²	21.5 GeV ²	21.7 GeV ²
Q^2 : 28 — 400 GeV ²	66.9 GeV ²	58.7 GeV ²	56.7 GeV ²	64.6 GeV ²
W : 50 — 120 GeV	94.2 GeV	95.8 GeV	95.6 GeV	93.7 GeV
W : 120 — 160 GeV	139.5 GeV	139.8 GeV	139.3 GeV	139.3 GeV
W : 160 — 250 GeV	199.0 GeV	193.9 GeV	193.3 GeV	199.1 GeV

Table 7.2: The mean values of Q^2 and W in the cross section bins from data and Monte Carlo before and after acceptance correction.

3. Decrease SINISTRA probability cut to 0.7.
4. Increase $E - p_z$ cut to 42 GeV.
5. Decrease $E - p_z$ cut to 38 GeV.
6. Tighten all box cuts by 1 cm.
7. Use cells for hadronic quantities.
8. Tighten $p_T(\pi_s)$ cut by 15 %.
9. Loosen $p_T(\pi_s)$ cut by 15 %.
10. Tighten $p_T(K, \pi)$ cut by 15 %.
11. Loosen $p_T(K, \pi)$ cut by 15 %.
12. Tighten $|p_T(K)| + |p_T(\pi)|$ cut by 15 %.
13. Loosen $|p_T(K)| + |p_T(\pi)|$ cut by 15 %.
14. Loosen $M(D^0)$ window to 1.78 \rightarrow 1.94 GeV.
15. Tighten $M(D^0)$ window to 1.80 \rightarrow 1.92 GeV.
16. Determine number of signal candidates using wrong charge combinations to subtract the combinatorial background.
17. Exclude DIS13 as a check on the trigger efficiency

18. Use condensates to determine η_{MAX} and not ZUFOS.
19. Move η_{MAX} noise threshold to 300 MeV.
20. Move η_{MAX} noise threshold to 500 MeV.
21. Tighten η_{MAX} cut to 1.5.
22. Increase proton-dissociation background contribution to 46 %.
23. Decrease proton-dissociation background contribution to 16 %.
24. Use RIDI to calculate the acceptance instead of RAPGAP. It is important to remember the upper Q^2 limit of 300 GeV² applied to this sample during generation. However, it is not expected that this should have a significant effect on the acceptances calculated with the sample.

The cross sections resulting from each of the 24 checks for each cross section bin, as well as the total cross section are illustrated in figures 7.11 and 7.12.

When considering the systematic errors on the total cross section, it is clear that the largest systematic error contributions are those from checks 21 — 24, which all relate specifically to requirements which are sensitive to model dependencies or are beyond the control of this analysis. When the binned cross sections are considered more fluctuation is seen, with check 7 becoming particularly bad. This is due to the fact that the bin choices have been optimised using variables based on ZUFOS which will have different migrations to those calculated using cells. All the cross section bins exhibit greater fluctuations than the total cross section, due to the limited statistics involved in this analysis. This can cause overestimation of the systematic errors, making them more a reflection of the statistical error than a true systematic error.

On inspection of figures 7.11 and 7.12 the variation of the fraction of proton dissociation background is consistently one of the largest systematic errors. This is supported by equation 7.6 which quotes the error on the total cross section which arises from the uncertainty on the level of proton dissociation background separately from the rest of the systematic errors.

The total systematic error is calculated by converting the cross section from each systematic check into a percentage difference from the nominal cross section. Those systematic

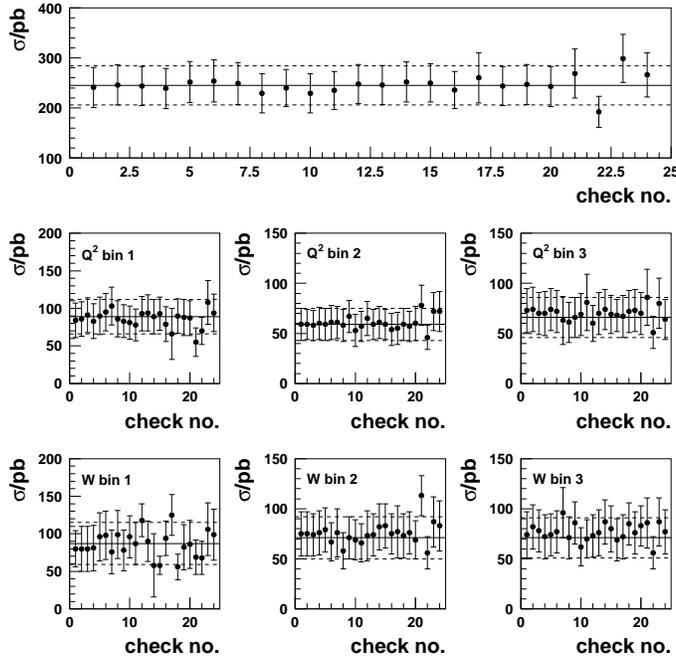


Figure 7.11: The cross sections resulting from each systematic error check versus check number for the total cross section (top) and in the Q^2 and W bins.

checks which increase the cross section are then summed in quadrature separately from those which decrease the cross section to give the upper and lower systematic errors. The final cross section can therefore be written as

$$\sigma(ep \rightarrow eD^{*\pm}Xp) = 245 \pm 39(stat.)_{-28}^{+38}(syst.) \pm 53(bkgd) \text{ pb} \quad (7.6)$$

where “*bkgd*” indicates the error arising from the uncertainty on the level of proton dissociation background.

The correlation between the different systematic checks has been tested by repeating the analysis first with all the systematic checks which increase the cross section applied simultaneously and then again with all the systematic checks which decrease the cross section applied simultaneously. The cross sections determined using this method give consistent results for the systematic errors to those given in equation 7.6, indicating that there is little correlation between the different checks.

Figure 7.13 shows the final differential cross sections, along with the preliminary ZEUS results shown at ICHEP98 [59] which were measured in the kinematic region defined by

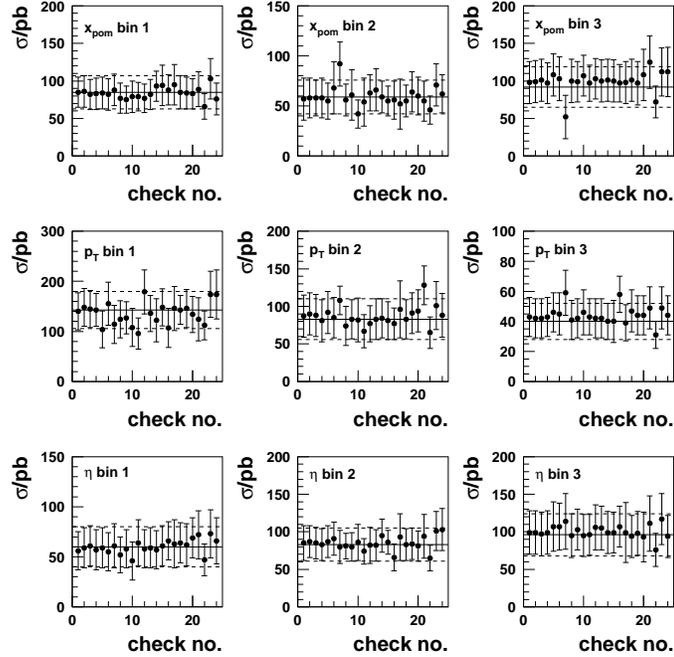


Figure 7.12: The cross sections resulting from each systematic error check versus check number for the x_P , $p_T(D^{*\pm})$ and $\eta(D^{*\pm})$ bins.

$$\begin{aligned}
 3 < Q^2 < 150 \text{ GeV}^2 \\
 0.02 < y < 0.7 \\
 p_T(D^{*\pm}) > 1.5 \text{ GeV} \\
 -1.5 < \eta(D^{*\pm}) < 1.5 \\
 0.002 < x_P < 0.012 \\
 0 < \beta < 0.8
 \end{aligned}$$

Good agreement is observed between the two sets of results, despite the difference in choice of kinematic region.

The differential distributions exhibit the expected behaviour, such as the sharp fall-off of $d\sigma/dQ^2$ and $d\sigma/dp_T$ with increasing Q^2 and $p_T(D^{*\pm})$. The W dependence is, as with the ICHEP98 results, reminiscent of that observed in the inclusive DIS $D^{*\pm}$ analysis [16] and is mainly determined by the $p_T(D^{*\pm})$ and $\eta(D^{*\pm})$ restrictions. Within errors, the $d\sigma/dx_P$ distribution is consistent with being flat, although there is a hint of an increase in the

cross section at very low x_P , while $d\sigma/d\eta$ also exhibits similar behaviour to that expected in the DIS $D^{*\pm}$ analysis [16], apart from in the forward region which will be reduced due to the presence of the rapidity gap.

Comparisons can also be made with the most recent H1 result [60] for the diffractive $D^{*\pm}$ cross section, $\sigma(ep \rightarrow eD^{*\pm}XY) = 154 \pm 40(stat.) \pm 35(syst.)$ pb in the kinematic region

$$\begin{aligned}
 2 < Q^2 < 100 \text{ GeV}^2 \\
 0.05 < y < 0.7 \\
 p_T(D^{*\pm}) > 2 \text{ GeV} \\
 -1.5 < \eta(D^{*\pm}) < 1.5 \\
 x_P < 0.04 \\
 M_Y < 1.6 \text{ GeV} \\
 |t| < 1 \text{ GeV}^2
 \end{aligned}$$

where X and Y are defined in the same way as in equation 7.4, M_Y is the mass of the hadronic system Y and $|t|$ is the square of the four-momentum transfer to the hadronic system Y . The majority of these requirements can be equivalently applied to ZEUS data, with the exception of the M_Y and $|t|$ restrictions. The last two requirements arise as a result of the extended coverage of the H1 detector in the forward region compared to ZEUS [19] and cannot therefore be applied when the cross section is measured from ZEUS data for comparison. The $|t| < 1 \text{ GeV}^2$ requirement should not be important as the cross section decreases sharply with increasing $|t|$. The H1 Collaboration find that by applying these cuts on $|t|$ and M_Y , they restrict the proton dissociation background in their diffractive data sample to the level of 5 % [83] compared to the 31 % estimated from ZEUS data. Thus subtracting the proton dissociation background appears to have the same effect as restricting M_Y when the cross section in the H1 kinematic region is measured from ZEUS data.

The cross section measured from ZEUS data using the H1 kinematic region, but without the $|t|$ and M_Y restrictions applied, is found to be $\sigma(ep \rightarrow eD^{*\pm}XY) = 426 \pm 80$ pb, where only the statistical error has been given. The proton dissociation background can then be subtracted giving a final cross section of $\sigma(ep \rightarrow eD^{*\pm}Xp) = 294 \pm 55(stat) \pm$

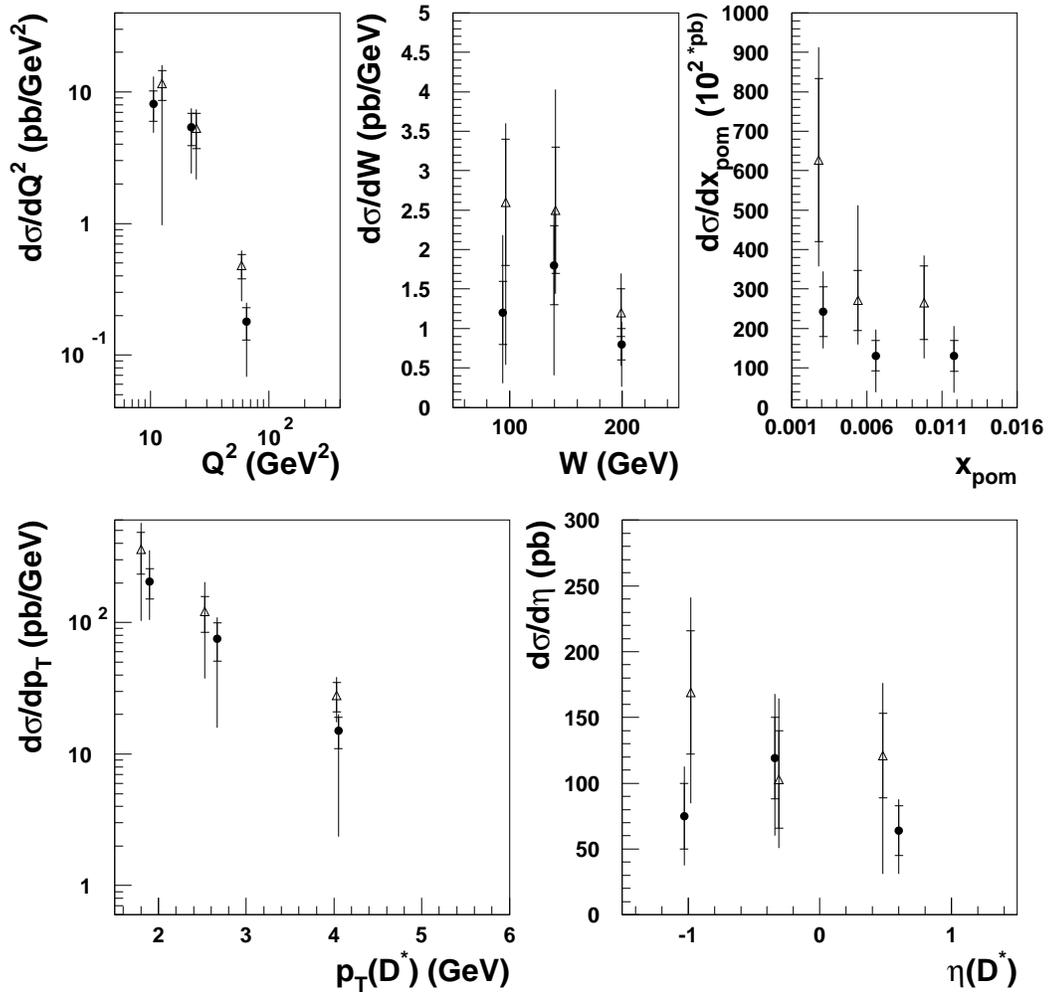


Figure 7.13: The differential cross sections calculated from the diffractive $D^{*\pm}$ data sample, using diffractive RAPGAP Monte Carlo to unfold from the measured number of candidates. The black dots are the points from this analysis, while the open triangles are the preliminary ZEUS results shown at ICHEP98. The inner error bars indicate the statistical error, while the outer bars are the statistical and systematic errors added in quadrature. The systematic shift between the two sets of results is caused by the kinematic region in which the cross sections are measured being different.

64(*bkgd*) pb, where “*bkgd*” indicates the error arising from the uncertainty on the level of proton dissociation background in diffractive ZEUS data. Reasonable agreement is observed when this value is compared with the H1 preliminary cross section for diffractive $D^{*\pm}$ production in this kinematic region.

7.7 Ratio Extraction

It is also interesting to look at the ratio of diffractive $D^{*\pm}$ production to all $D^{*\pm}$ production in the kinematic region

$$\begin{aligned}
 6 < Q^2 < 400 \text{ GeV}^2 \\
 0.02 < y < 0.7 \\
 p_T(D^{*\pm}) > 1.5 \text{ GeV} \\
 -1.5 < \eta(D^{*\pm}) < 1.5 \\
 0.001 < x_{\mathcal{P}} < 0.016 \\
 \beta < 0.8
 \end{aligned}$$

It is not possible, however, to apply the β and $x_{\mathcal{P}}$ requirements directly to the inclusive $D^{*\pm}$ sample, but they can be replaced by requirements on Bjorken x using the relation $x_{\mathcal{P}}\beta = x_{Bj}$. The upper limit is therefore $x_{\mathcal{P}}^{max}\beta_{max} = 0.0128$. For the lower limit, the minimum value of β accessible in this analysis, β_{min} , is required and is given by

$$\beta_{min} = \frac{Q_{min}^2}{M_X^2(max) + Q_{min}^2}$$

Figure 6.15 indicates that the maximum value of M_X accessible in this analysis is approximately 30 GeV, making $\beta_{min} \simeq 0.007$. This then gives $x_{Bj}^{min} \simeq 7 \times 10^{-6}$. However, the lowest value of x_{Bj} accessible for the given Q^2 and y limits is

$$x_{Bj}^{min} = \frac{Q_{min}^2}{y_{max}^8} = 9.5 \times 10^{-5}$$

indicating that both the diffractive and the inclusive $D^{*\pm}$ samples have the same implicit low x_{Bj} limit and that no explicit low x_{Bj} cut is required.

The DIS $D^{*\pm}$ ΔM distributions for the above kinematic region and in the Q^2 and W bins used for the diffractive differential cross sections are shown in figure 7.14. The numbers of

$D^{*\pm}$ candidates and acceptances calculated using the non-diffractive RAPGAP samples, along with the cross sections are summarised in table 7.3. The mean values of Q^2 and W for each of the bins are also given in table 7.3 and are taken from the acceptance-corrected data inclusive $D^{*\pm}$ sample. Good agreement is observed in the mean values from both data samples, with the exception of the highest Q^2 bin. The inclusive $D^{*\pm}$ data and Monte Carlo samples are better understood theoretically than the diffractive samples, hence the choice of mean values.

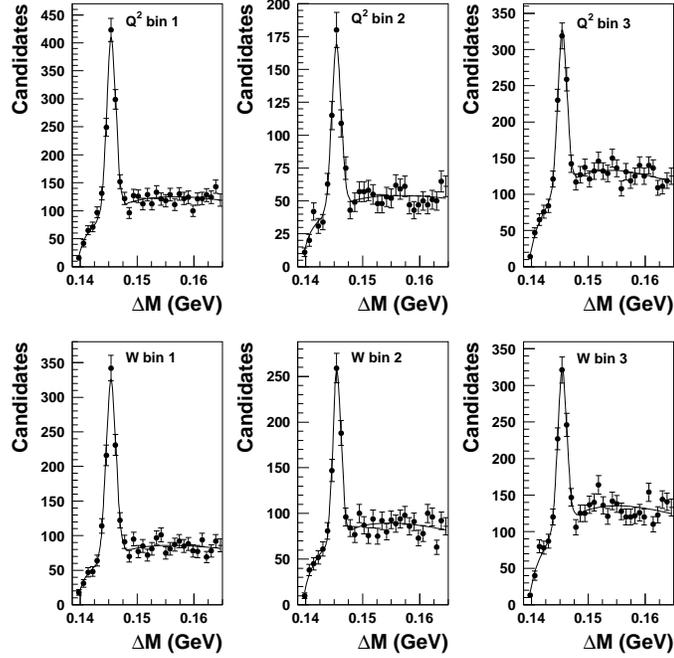


Figure 7.14: The DIS ΔM distributions in the Q^2 bins (top row) and W bins (bottom row) used in the calculation of the $D^{*\pm}$ ratios.

The total ratio in the above kinematic region is

$$\frac{\sigma(ep \rightarrow eD^{*\pm}Xp)_{diff}}{\sigma(ep \rightarrow eD^{*\pm}XY)} = 6.3 \pm 1.1(stat)_{-1.4}^{+1.6}(syst) \% \quad (7.7)$$

where in the denominator X indicates the hadronic system arising from the hard scatter and Y indicates the proton remnant. The systematic errors have been determined from systematic checks 18 – 24 only, under the assumption that all other checks will cancel out when the ratio is taken. The total systematic error is then calculated by converting the results of each check into a percentage difference from the nominal diffractive cross section and summing in quadrature those checks which increase the cross section and those checks which decrease the cross section separately. The resulting upper and lower

systematic errors are then divided by the nominal inclusive $D^{*\pm}$ cross section to obtain the upper and lower systematic errors on the ratio.

The ratio measured in this analysis is consistent with the preliminary ZEUS results presented at ICHEP98, which is

$$\frac{\sigma(ep \rightarrow eD^{*\pm}Xp)_{diff}}{\sigma(ep \rightarrow eD^{*\pm}XY)} = 7.0 \pm 1.3(stat)_{-1.8}^{+1.7}(syst) \%$$

It is also consistent with the fraction of diffractively produced events in inclusive DIS, as measured by ZEUS [129], where it is found that the ratio is independent of W , but varies with both Q^2 and M_X for low values of M_X ($M_X < 7.5$ GeV). However, in the interval $7.5 < M_X < 15$ GeV (the most appropriate for comparison with this analysis), no strong Q^2 dependence is observed and the ratio is approximately 4 %. The Q^2 and W dependencies of the ratio measured in this analysis are shown in figure 7.15. The ratio appears to be independent of both Q^2 and W within errors, as observed in the inclusive DIS case.

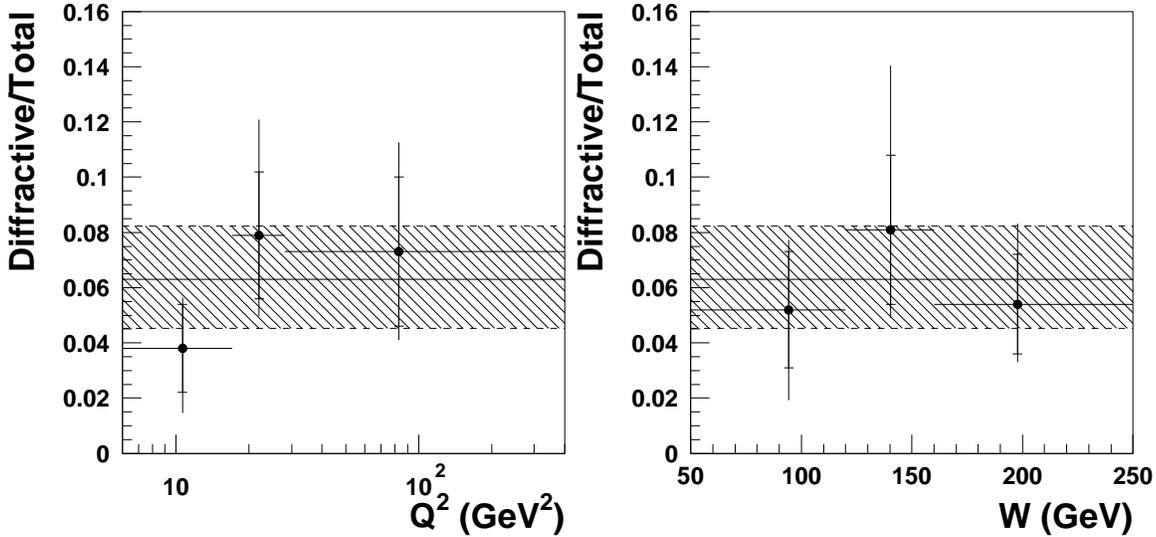


Figure 7.15: The differential ratios of diffractive $D^{*\pm}$ production to DIS $D^{*\pm}$ production as a function of Q^2 and W . The inner error bars indicate the statistical errors, while the outer bars indicate the statistical and systematic errors added in quadrature. The solid horizontal line indicates the total ratio, while the shaded band indicates its statistical and systematic errors added in quadrature.

Bin (X)	No. of $D^{*\pm}$ Candidates	Acceptance (%)	Cross Section (nb)	$\langle X \rangle$
Q^2 : 6 — 17 GeV ²	761.8 ± 41.6	28.92 ± 0.44	2.32 ± 0.16	10.6 GeV ²
Q^2 : 17 — 28 GeV ²	313.4 ± 26.9	36.86 ± 0.78	0.75 ± 0.07	22.0 GeV ²
Q^2 : 28 — 400 GeV ²	534.9 ± 37.6	52.45 ± 0.68	0.90 ± 0.07	83.0 GeV ²
W : 50 — 120 GeV	683.5 ± 38.0	36.24 ± 0.51	1.66 ± 0.11	94.3 GeV
W : 120 — 160 GeV	403.1 ± 30.3	40.31 ± 0.68	0.88 ± 0.08	140.4 GeV
W : 160 — 250 GeV	522.3 ± 37.3	34.90 ± 0.64	1.32 ± 0.11	197.7 GeV

Table 7.3: A summary of the information relating to the inclusive $D^{*\pm}$ cross sections.

7.8 Comparison with Theoretical Predictions

Having extracted the cross sections shown in figure 7.13, it is interesting to compare them to the predictions of different theoretical models. To start with it is useful to compare the data cross sections with those predicted by the model on which the diffractive RAPGAP Monte Carlo samples are based. However, having observed discrepancies in the numbers of $D^{*\pm}$ candidates predicted by this model when data and Monte Carlo were compared (see section 7.3), it is important to first understand where these discrepancies come from. The solution comes from the form of the pomeron flux which contains an arbitrary normalisation factor, which can only be determined by performing fits to the data. Given that the pomeron flux factorises from the pomeron structure function, this will affect the production rate and not the shapes of the distributions predicted by this model. Thus comparisons can still be made between the data and the predictions of this model, but the predictions must first be normalised to the data production rate. The appropriate normalisation factor required is given by $\sigma_{data}/\sigma_{MC}$, where σ_{MC} is the diffractive $D^{*\pm}$ cross section predicted by the model for the kinematic region used in this analysis and σ_{data} is the equivalent cross section measured from ZEUS data in this analysis.

Figure 7.16 shows the comparisons made between the measured differential cross sections and the predictions made using the Ingelman-Schlein model, as implemented in the RAPGAP Monte Carlo generator. Two different forms for the initial pomeron parton distributions are shown: the first assumes hard parton distributions of the form

$$\beta f_{g/\mathbb{P}}(\beta, Q_0^2) = 6\beta(1 - \beta)$$

$$\beta f_{q/\mathbb{P}}(\beta, Q_0^2) = \frac{1}{4} f_{g/\mathbb{P}}(\beta, Q_0^2) \quad (7.8)$$

These are the same as those implemented in the diffractive RAPGAP Monte Carlo samples used for cross section unfolding. σ_{MC} is found to be 944 pb, giving a normalisation factor of $245/944 = 0.260$. The second model shown in figure 7.16 assumes soft parton distributions of the form

$$\begin{aligned} \beta f_{g/\mathbb{P}}(\beta, Q_0^2) &= 6(1 - \beta)^5 \\ \beta f_{q/\mathbb{P}}(\beta, Q_0^2) &= \frac{1}{4} f_{g/\mathbb{P}}(\beta, Q_0^2) \end{aligned} \quad (7.9)$$

and gives a value of 580 pb for σ_{MC} . This implies a normalisation factor of $245/580 = 0.422$. The data shapes are reasonably well reproduced by both sets of parton distributions, although there is perhaps a slight preference for the $d\sigma/dx_{\mathbb{P}}$ prediction based on the hard parton distributions. However, given the limited statistics involved in this analysis, it is not possible to make any definitive statements.

HERA data have been fitted by Alvero et al. [55] using the Ingelman-Schlein model and assuming five different initial configurations for the pomeron parton distributions. In four of the fits (**A** — **D**), the initial forms of the different parton distributions are given by

$$\begin{aligned} \beta f_{q/\mathbb{P}}(\beta, Q_0^2) &= a_q [\beta(1 - \beta) + \tilde{a}_q(1 - \beta)^2] \\ \beta f_{g/\mathbb{P}}(\beta, Q_0^2) &= a_g \beta(1 - \beta) \end{aligned} \quad (7.10)$$

For the fifth fit (**SG**), “super-hard” distributions are assumed which have the form

$$\begin{aligned} \beta f_{q/\mathbb{P}}(\beta, Q_0^2) &= a_q \beta(1 - \beta) \\ \beta f_{g/\mathbb{P}}(\beta, Q_0^2) &= a_g \beta^8(1 - \beta)^{0.3} \end{aligned} \quad (7.11)$$

and are based on results determined by the H1 Collaboration from QCD fits of their measurements of the diffractive structure function $F_2^{D(3)}$ [54]. The values of the constants a_q , \tilde{a}_q and a_g which are free parameters of the different fits are summarised in table 7.4. The fits were performed three times, each time with $\alpha_{\mathbb{P}}(0)$ set to a different value (1.08, 1.14 or 1.19). For fits **A** — **C** and **SG**, the fits with the lowest χ^2 were achieved with $\alpha_{\mathbb{P}}(0) = 1.14$. For fit **D** the lowest χ^2 was achieved with $\alpha_{\mathbb{P}}(0)$ set to 1.19. The parton distributions were evolved from the starting scale of $Q_0^2 = 4 \text{ GeV}^2$ using the standard NLO DGLAP equations.

Using the results of these fits, Alvero et al. [81] then calculated the diffractive $D^{*\pm}$ differential cross sections in the bins used for the preliminary ZEUS results shown at ICHEP98

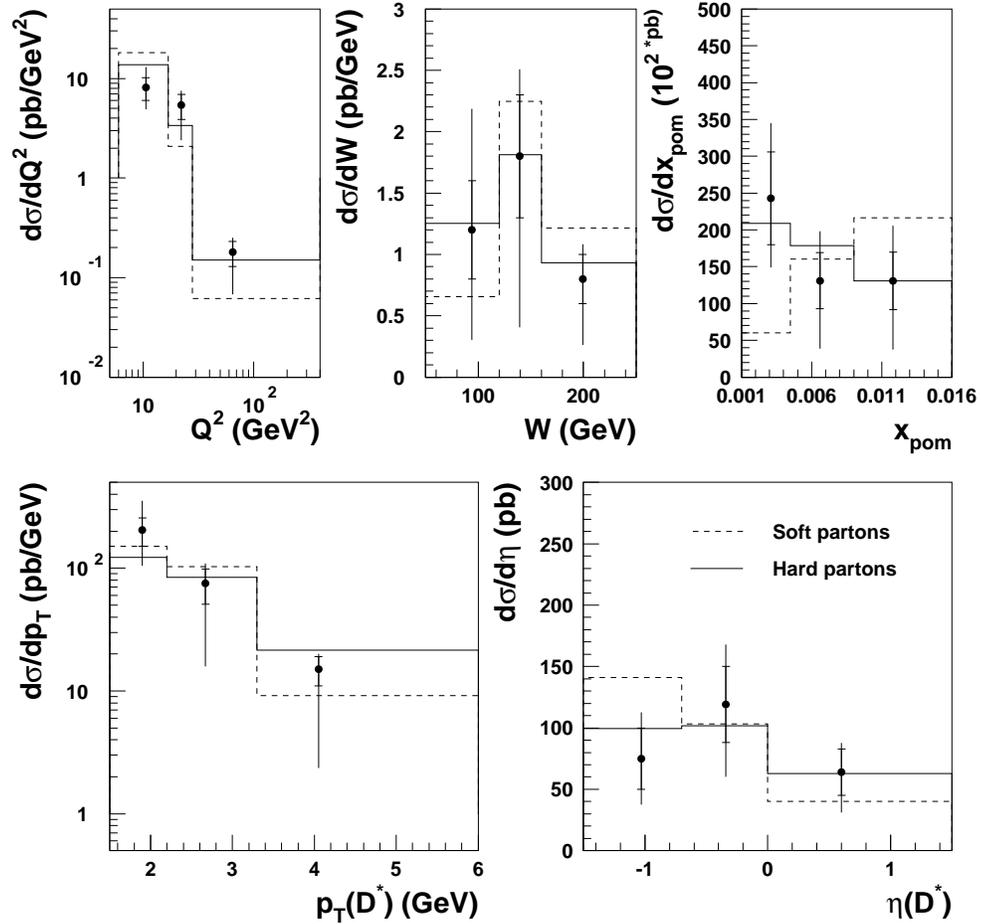


Figure 7.16: The final differential cross sections compared to the RAPGAP implementation of the Ingelman-Schlein model using hard (solid line) and soft (dashed line) parton distributions, appropriately normalised (see text).

Fit	a_q	a_g	\tilde{a}_q
A	0.0240 ± 0.006	0	0
B	0.0239 ± 0.006	4.5 ± 0.5	0
C	0.0249 ± 0.011	0	-0.031 ± 0.029
D	0.1750 ± 0.008	6.7 ± 1.0	-0.191 ± 0.026
SG	0.2250 ± 0.008	7.4 ± 2.2	0

Table 7.4: The fit parameters from the Alvero et al. calculations.

[59]. The calculations were made using the Peterson model for charm fragmentation [26] and assuming a charm mass of 1.5 GeV.

Figure 7.17 shows comparisons between these calculations and the cross sections measured in this analysis, which have been recalculated in the ICHEP98 bins to allow direct comparisons to be made. For this reason no systematic errors are available for this comparison. From Figure 7.17 it can be concluded that only fits which contain a significant hard gluon component (fits **B** and **D**) are able to reproduce the cross sections measured in this analysis. This is consistent with the fit qualities observed by Alvero et al. and with the results determined from diffractive dijet production at HERA [53]. It is also observed that the “super-hard” gluon distribution suggested by the H1 Collaboration [54] does not describe the data as well as the standard $\beta(1 - \beta)$ distribution.

Since the predictions were made by Alvero et al. for the ICHEP98 results, the range of favoured charm masses has decreased to 1.1 – 1.4 GeV. It might be expected that the choice of charm mass could affect the normalisation of the cross sections predicted using the Ingelman-Schlein model — a source of theoretical uncertainty which must be investigated. Other sources of theoretical uncertainty which must also be studied are the choice of $\alpha_P(0)$, the exact value of $f(c \rightarrow D^{*\pm})$ and the choice of hadronisation model.

In order to investigate the different theoretical uncertainties, small samples of events were generated with one of the parameters being studied changed from its nominal value. The “nominal” values are those implemented in the RAPGAP Monte Carlo samples used for cross section unfolding, ie. $m_c = 1.25$ GeV, $\alpha_P(0) = 1.157$, $f(c \rightarrow D^{*\pm}) = 0.26$ and the ARIADNE hadronisation scheme [130]. From these samples, $d\sigma/dW$ is then determined as an example distribution to illustrate the magnitude of the different sources of uncertainty.

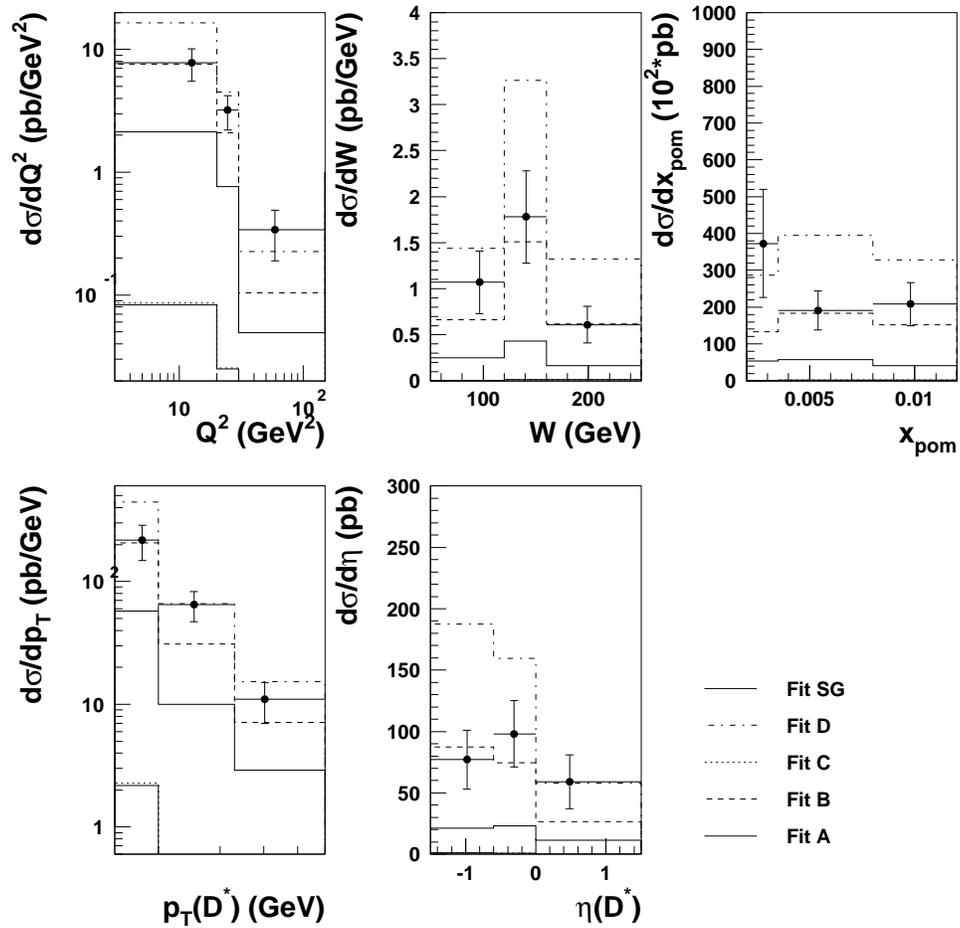


Figure 7.17: A comparison of the final differential cross sections with the five different fits made by Alvero et al., the details of which are in the text. The data appear to favour fits **B** and **D**. It should be noted that the upper solid line is fit **SG** and is in some cases the only solid line visible.

Figure 7.18 shows the effect of varying the different parameters on the $d\sigma/dW$ distribution. The variation caused by changing from ARIADNE to the MEPS hadronisation scheme [119] is even smaller than that caused by the uncertainty on $f(c \rightarrow D^{*\pm})$ and is therefore not shown. The two values of $f(c \rightarrow D^{*\pm})$ shown are the old [25] and new [27] values determined by the OPAL Collaboration. The old value is implemented in the RAPGAP generator. Figure 7.18 shows that the largest sources of uncertainty arise from the charm mass and from varying $\alpha_P(0)$. These two effects are then combined for the final plot in figure 7.18 to give an estimate of the overall uncertainty. There is no obvious evidence for any correlation between the effects of the charm mass and the pomeron intercept indicating that, in future, indirect measurements of both these quantities may be possible.

It is also interesting to compare the data with predictions from the other types of diffractive models, such as the Ryskin model implemented in the RIDI Monte Carlo generator [65, 67] and the soft colour interaction (SCI) model of Buchmüller et al. [75, 79], as implemented in the RAPGAP Monte Carlo generator. These comparisons are shown in figure 7.19. It is important to remember that the RIDI Monte Carlo sample was generated with $Q^2 < 300 \text{ GeV}^2$. This, however, is not expected to have a noticeable effect on the predicted cross sections. The predictions of Buchmüller et al. in figure 7.19 are shown for a charm mass of 1.4 GeV, while the RIDI predictions are shown for a charm mass of 1.5 GeV. The former describes both the shape and the normalisation of the data quite well. The latter, however, clearly overestimates the absolute values of the cross sections, although the agreement in shape is quite good. Unlike the Ingelman-Schlein model, the Ryskin model predicts the absolute values of the cross sections as well as the shapes of the distributions. This means that the discrepancies observed in figure 7.19 cannot be attributed to the presence of an arbitrary normalisation factor. There are, however, other theoretical uncertainties associated with this model which may account for the observed discrepancies. For example, there are large uncertainties associated with the precise form of the K -factor (the factor by which the cross sections are enhanced by the inclusion of virtual gluon corrections [64]), which could decrease the absolute values of the cross sections. The Ryskin model is also sensitive to the choice of input proton gluon density — better agreement may be achieved through the use of a different parametrisation. There are also several other two-gluon exchange models available which may better describe the data, but comparisons with these models are not currently possible. It is also interesting to note that the SCI model of Buchmüller et al. only contains real gluon corrections and not the virtual corrections,

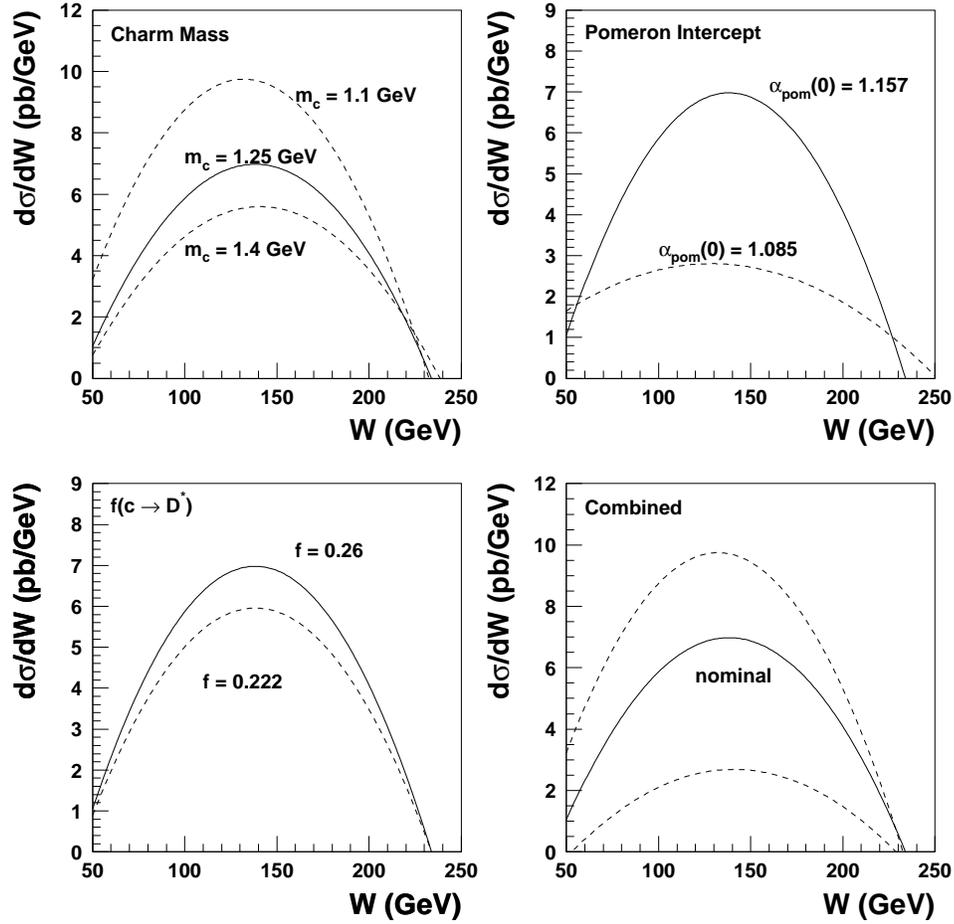


Figure 7.18: The theoretical uncertainties associated with the Ingelman-Schlein model. The solid line shows the “nominal value”, ie. that which is implemented in the RAPGAP Monte Carlo samples, while the dotted lines indicate the effect of varying each parameter. In the final plot the upper dotted line shows the combined effect of $m_c = 1.1$ GeV and $\alpha_{\mathcal{P}}(0) = 1.157$, while the lower dotted line shows the combined effect of $m_c = 1.4$ GeV and $\alpha_{\mathcal{P}}(0) = 1.085$.

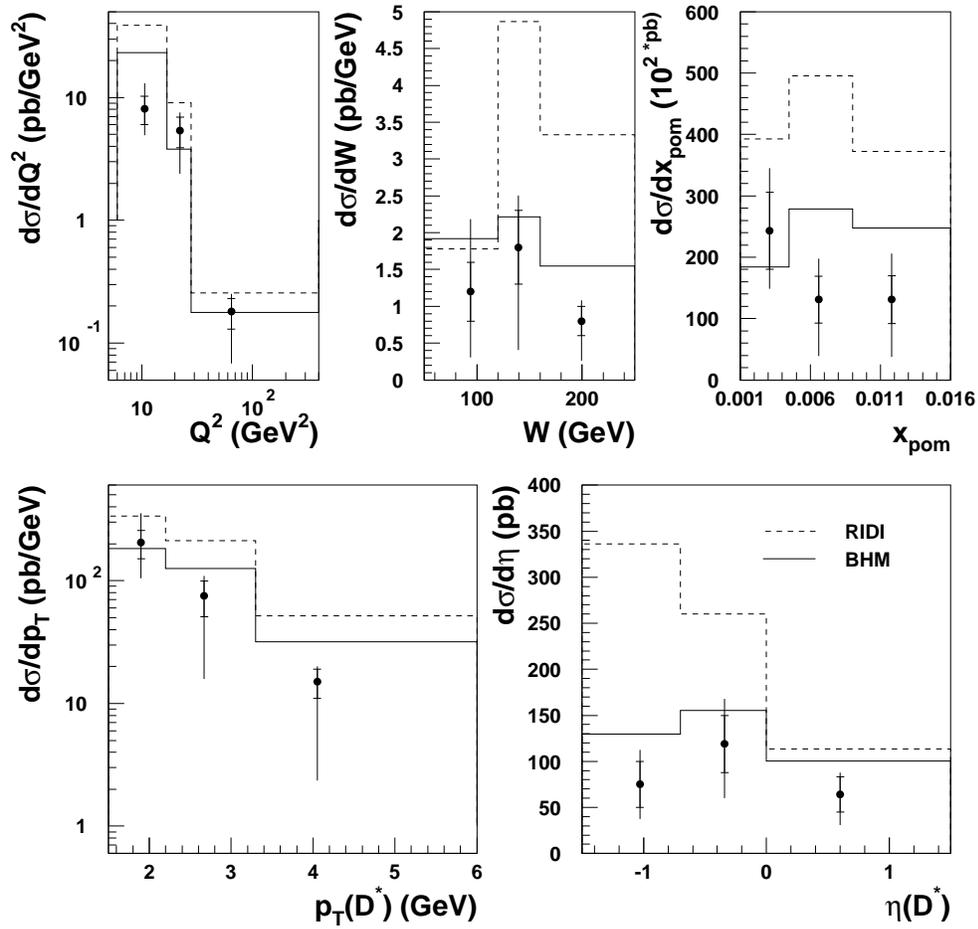


Figure 7.19: A comparison of the final differential cross sections with the predictions from the model of Buchmüller et al. (“BHM”) and from the RIDI Monte Carlo samples (“RID”). Further details of the models are given in the text.

which may account for the overestimation of the cross sections by RIDI. This absence of K -factor enhancement could also explain the harder than expected β distribution observed in the data, providing the level of enhancement is different for $c\bar{c}$ and $c\bar{c}g$ final states. Another explanation is that more charm is produced through two-jet processes than through three-jet processes, regardless of whether the K -factor enhancement is included or not. In terms of the Ingelman-Schlein model, the observed β distribution could be explained by the presence of an additional direct coupling component, as this would enhance the high β region (see chapter 2).

In conclusion, the data can be reasonably well described by both the Ingelman-Schlein model and the model of Buchmüller et al., bearing in mind the effect which the uncertainty on the charm mass may cause. The two-gluon exchange model of Ryskin [65], as implemented in the RIDI Monte Carlo generator [67], does not describe the data as well, but it is known that this model is sensitive to the choice of K -factor and the input proton gluon density which, if carefully chosen, may improve the agreement of this model with the data.

Despite all these uncertainties, it is clear that within the context of the Ingelman-Schlein model, the pomeron must contain a significant gluon component in order to describe the data well. It is also noted that the “super-hard” gluon distribution suggested by the H1 Collaboration [54] (fit **SG** in figure 7.17) does not appear to describe the data as well as the standard hard gluon density given in equation 7.10 and used in fits **B** and **D**. It is also clear that charm accounts for significantly more than the 1 % of diffractive interactions predicted by the early calculations of Nikolaev and Zakharov [62]. This may indicate the need for higher order corrections to be included in the two-gluon exchange models. Alternatively, this higher level of charm production in diffraction could be attributed to the cancellation of heavy flavour suppression at low x_P (x_{Bj}) by the fast increase of the proton gluon density, as suggested in later work by Nikolaev and Zakharov [77].

Chapter 8

Conclusions and Outlook

An analysis of diffractive open charm production in deep inelastic scattering (DIS) data has been presented, using the full reconstruction of $D^{*\pm}$ (2010) mesons in the $D^* \rightarrow D^0 \pi_s \rightarrow (K\pi)\pi_s$ decay channel as the method of tagging charm. The analysis was performed on the combined 1995-97 ZEUS data sample, which corresponds to an integrated luminosity of 43.3 pb^{-1} . Diffractive events are selected from this sample by requiring the presence of a large rapidity gap in the forward region of the detector.

Using this data sample, the cross section for diffractive $D^{*\pm}$ production has been measured in the kinematic region

$$\begin{aligned} 6 < Q^2 < 400 \text{ GeV}^2 \\ 0.02 < y < 0.7 \\ p_T(D^{*\pm}) > 1.5 \text{ GeV} \\ -1.5 < \eta(D^{*\pm}) < 1.5 \\ 0.001 < x_F < 0.016 \\ 0 < \beta < 0.8 \end{aligned}$$

and is found to be $245 \pm 39(\text{stat})_{-28}^{+38}(\text{syst}) \pm 53(\text{bkgd})$ pb. The ratio of diffractive $D^{*\pm}$ production to total $D^{*\pm}$ production in this region is found to be

$$\frac{\sigma(ep \rightarrow eD^{*\pm}Xp)}{\sigma(ep \rightarrow eD^{*\pm}XY)} = 6.3 \pm 1.1(\text{stat})_{-1.4}^{+1.6}(\text{syst})\%$$

which is consistent with the fraction of diffractively produced events in inclusive DIS [129]. The extension of the kinematic region with respect to previous ZEUS measurements

[59, 82] has been made possible by the increase in size of the available dataset as well as the inclusion for the first time of data taken with a specially-tailored DIS D^* trigger. These latest results remain consistent both with previous ZEUS results [59] and with recent results produced by the H1 Collaboration [60]. The level of diffractive open charm production observed through $D^{*\pm}(2010)$ reconstruction already shows that heavy flavour production is not suppressed, contrary to the original predictions of Nikolaev and Zakharov [62]. This lack of suppression may indicate the need for the NLO corrections proposed by Levin et al. [64] and Buchmüller et al. [79] to be included in calculations based on two-gluon exchange models. Equally this effect could be attributed to the cancellation of heavy flavour suppression by the fast increase of the proton gluon density at low x , as suggested in the more recent work of Nikolaev and Zakharov [77].

Differential cross sections with respect to Q^2 , W , x_P , $p_T(D^{*\pm})$ and $\eta(D^{*\pm})$ have also been measured and compared to several different theoretical predictions. Comparisons with the predictions of Alvero et al. [81] which are based on the Ingelman-Schlein model indicate that the data prefer a pomeron which contains a substantial gluon component, as was observed in diffractive dijet production at ZEUS [53]. A slight preference for hard pomeron parton distributions over soft parton distributions in the data is also observed. Comparisons are also made with predictions from the soft colour interaction (SCI) model of Buchmüller et al. [75, 79] and the perturbative two-gluon exchange model of Ryskin [65]. It is observed that while the SCI model appears to describe the data quite well, the Ryskin model, as implemented in the RIDI Monte Carlo generator, does not perform as well. There are, however, several different theoretical uncertainties associated with this model which may account for the discrepancies observed.

The final results presented are the Q^2 and W dependencies of the ratio of diffractive $D^{*\pm}$ production to all $D^{*\pm}$ production in the previously defined kinematic region. It is found that within the errors, the measured ratios are approximately independent of both Q^2 and W . Similar behaviour is observed in the inclusive DIS case [129].

It will not be possible to improve the accuracy of these measurements until after the HERA luminosity upgrade in 2000. Although the current results have yielded important information — the absence of heavy flavour suppression and the need for a substantial hard gluon contribution in the context of the Ingelman-Schlein model — there are still more important discoveries to be made in this area of diffractive physics. In the medium term,

it will be possible to extract the diffractive open charm cross section and consequently to determine the ratios of diffractive charm to all charm and all diffraction. This is not currently possible, as the $D^{*\pm}$ cross sections from which the charm cross sections will be extracted are measured in a kinematic region restricted with respect to $p_T(D^{*\pm})$ and $\eta(D^{*\pm})$. In order to determine the charm cross sections, the $D^{*\pm}$ cross sections must be extrapolated to the full $p_T(D^{*\pm})$, $\eta(D^{*\pm})$ region using some sort of theoretical calculation. Such an extrapolation could be performed using Monte Carlo, but it is not certain that the Monte Carlo can be trusted for such a procedure. Once the $D^{*\pm}$ calculations of Alvero et al. are publicly available as a Monte Carlo-like program, this could be used to perform the extrapolation. This, however, is not expected to be available for several months.

In the long term, post-luminosity upgrade data will provide substantially larger samples for analysis which will also benefit from the addition of the new MicroVertex Detector. Only then can the errors on the cross sections be reduced sufficiently to allow discrimination between the different theoretical models and hence to answer the fundamental question — what is the Pomeron ?

Bibliography

- [1] E. D. Bloom et al., Phys. Rev. Lett. **23** (1969) 930.
A. Bodek et al., Phys. Rev. Lett. **30** (1973) 1087.
- [2] J. D. Bjorken, Phys. Rev. **179** (1969) 1547.
- [3] H. E. Fisk & F. Sciulli, Ann. Rev. Nucl. Part. Sci. **32** (1982) 499.
- [4] R. P. Feynman, *Photon-Hadron Interactions*, W.A. Benjamin inc., 1972.
- [5] M. Derrick et al., Z. Phys. **C 72** (1996) 399.
- [6] G. Altarelli & G. Parisi, Nucl. Phys. **B 81** (1977) 1298.
G. Altarelli, Nucl. Phys. **B 126** (1981) 1.
- [7] V. N. Gribov & L. N. Lipatov, Sov. J. Nucl. Phys. **15** (1972) 438.
Y. L. Dokshitzer, Sov. Phys. JETP **46** (1977) 641.
- [8] A. Quadt, *Measurement and QCD Analysis of the Proton Structure Function F_2 from the 1994 HERA Data using the ZEUS Detector*, PhD. thesis, University of Oxford, 1996.
- [9] Y. Balitsky & L. N. Lipatov, Phys. Rev. Lett. **28** (1978) 822.
E. A. Kuraev, L. N. Lipatov & V. S. Fadin, Phys. Rev. Lett. **44** (1976) 443; Phys. Rev. Lett. **45** (1977) 199.
- [10] G. t'Hooft, Nucl. Phys. **B72** (1974) 461.
- [11] J. Breitweg et al., Phys. Lett. **B 407** (1997) 402.
- [12] G. Ingelman, L. Jönsson & M. Nyberg, Phys. Rev. **D 47** (1993) 4872.
- [13] J. J. Aubert et al., Phys. Lett. **B 110** (1982) 73.
J.J. Aubert et al., Nucl. Phys. **B 213** (1983) 31.

- [14] H1 Coll., *Z. Phys.* **C 72** (1996) 593.
- [15] Proceedings of the Future Physics at HERA Workshop (1991), Vol. 2, edited by W. Buchmüller and G. Ingelman, p. 739.
- [16] ZEUS Coll., Abstract 768, *Measurement of D^* Cross Sections and the Charm Structure Function of the Proton in Deep Inelastic Scattering at HERA*, ICHEP98, Vancouver, July 1998.
- [17] H1 Coll., Abstract 538, *Determination of the gluon density in the proton from charm electro- and photo-production at HERA using NLO QCD*, ICHEP98, Vancouver, July 1998.
- [18] C. Caso et al., *Eur. Phys. J.* **C 3** (1998) 1.
- [19] A.M. Cooper-Sakar et al., *Structure Functions of the Nucleon and their Interpretation*, DESY 97-226.
- [20] J.C. Collins et al., *Charm in Deep Inelastic Scattering*, hep-ph/9709440.
- [21] A.D. Martin et al., *Consistent Treatment of Charm Evolution in Deep Inelastic Scattering*, RAL-TR-96-103 (revised January 1997).
- [22] R.S. Thorne & R.G. Roberts, *A Practical Procedure for Evolving Heavy Flavour Structure Functions*, hep-ph/9711223.
- [23] H.L. Lai & W.K. Tung, *Z. Phys.* **C 74** (1997) 463.
- [24] G. Feldman et al., *Phys. Rev. Lett.* **38** (1977) 1313.
- [25] R. Akers et al., *Z. Phys.* **C 67** (1995) 27 and references therein.
- [26] C. Peterson et al., *Phys. Rev.* **D 27** (1983) 105.
- [27] K. Ackerstaff et al., *Eur. Phys. J.* **C 1** (1998) 439.
- [28] A.D. Martin et al., *Phys. Lett.* **B 387** (1996) 419.
- [29] A.D. Martin et al., *Parton Distributions: A New Global Analysis*, RAL-TR-1998-029.
- [30] M. Glück et al., *Z. Phys.* **C 67** (1995) 433.
- [31] M. Glück et al., *Nucl. Phys.* **B 422** (1994) 37.

- [32] J. R. Forshaw & D. A. Ross, *Quantum Chromodynamics and the Pomeron*, CUP 1997.
- [33] T. Regge, *Nuovo Cimento* **14** (1959) 951.
- [34] G. F. P. Chew & S. C. Frautschi, *Phys. Rev. Lett.* **7** (1961) 399,
Phys. Rev. Lett. **8** (1962) 41.
- [35] I. Pomeranchuk, *Sov. Phys. JETP* **7** (1958) 499.
- [36] L. F. Foldy & R. F. Peierls, *Phys. Rev.* **130** (1963) 1585.
- [37] V. N. Gribov, *JETP Lett.* **41** (1961) 667.
- [38] A. Donnachie & P. V. Landshoff, *Phys. Lett.* **B 296** (1992) 227,
Nucl. Phys. **B 267** (1986) 690.
- [39] S. Abatzis et al., *Phys. Lett.* **B 324** (1994) 509.
- [40] J. Bjorken, *Rapidity Gaps in DIS*, talk at ITEP, Moscow 1995, SLAC-PUB-7096.
- [41] M. Derrick et al., *Phys. Lett.* **B 293** (1992) 465.
- [42] R. G. Roberts, *The Structure of the Proton*, CUP 1990.
- [43] J. Breitweg et al., *Exclusive electroproduction of ρ^0 and J/ψ Mesons at HERA*, DESY 98-107.
- [44] J. Breitweg et al., *Z. Phys.* **C 75** (1997) 215.
- [45] G. Ingelman & P. E. Schlein, *Phys. Lett.* **B 152** (1985) 256.
- [46] A. Brandt et al., *Phys. Lett.* **B 297** (1992) 417,
R. Bonino et al., *Phys. Lett.* **B 211** (1988) 239.
- [47] A. Donnachie & P. V. Landshoff, *Phys. Lett.* **B 191** (1987) 309, *Phys. Lett.* **B 198** (1987) 590(E).
- [48] R. W. McAllister & R. Hofstadter, *Phys. Rev.* **102** (1956) 851.
R. Hofstadter, *Rev. Mod. Phys.* **28** (1956) 214.
- [49] E. Berger, J. Collins, D. Soper & G. Sterman, *Nucl. Phys.* **B 286** (1987) 704.
- [50] P. Bruni & G. Ingelman, *Phys. Lett.* **B 311** (1993) 317.

- [51] T. Gehrmann & W. J. Stirling, *Z. Phys.* **C 70** (1996) 89.
- [52] Kniehl, Kohrs and Kramer, *Z. Phys.* **C 65** (1995) 657.
- [53] M. Derrick et al., *Phys. Lett.* **B 356** (1995) 129.
- [54] C. Adloff et al., *Z. Phys.* **C 76** (1997) 613.
- [55] L. Alvero et al., *Diffraction Production of Jets and Weak Bosons and Tests of Hard-Scattering Factorization*, hep-ph/9805268.
- [56] F. Abe et al., *Phys. Rev. Lett.* **79** (1997) 2636.
- [57] F. Abe et al., *Phys. Rev. Lett.* **78** (1997) 2698.
- [58] J. C. Collins, *Phys. Rev* **D 57** (1998) 3051.
- [59] ZEUS Collaboration, *Open Charm Production in Deep Inelastic Diffractive Scattering at HERA*, abstract 785, ICHEP98, Vancouver 1998.
- [60] H1 Collaboration, *Production of $D^{*\pm}$ Mesons in Diffractive Interactions at HERA*, abstract 564, ICHEP98, Vancouver 1998.
- [61] F. E. Low, *Phys. Rev.* **D 12** (1975) 163,
S. Nussinov, *Phys. Rev. Lett.* **34** (1975) 1886,
Phys. Rev. **D 14** (1976) 246.
- [62] N. N. Nikolaev & B. G. Zakharov, *Z. Phys.* **C 53** (1992) 331.
- [63] H. Abramowicz et al., *Future Physics at HERA Workshop 1995/96*, Vol. 2, p.635.
- [64] E. M. Levin et al., hep-ph/9606443 (1996).
- [65] M. G. Ryskin, *Sov. Journ. Nucl. Phys.* **52** (1990) 529.
- [66] J. Kubar-André & F. E. Paige, *Phys. Rev.* **D 19** (1979) 221.
G. Altarelli, R. K. Ellis & G. Martinelli, *Nucl. Phys.* **B 143** (1978) 521.
- [67] M. G. Ryskin, S. Y. Sivoklokov & A. Solano, *Proc. of the Int. Conf. on Elastic and Diffractive Scattering*, Ed. H. M. Fried et al., (1993) World Scientific.
- [68] J. Bartels & M. Wüsthoff, *J. Phys. G: Nucl. Part. Phys.* **22** (1996) 929.
- [69] A. Bialas & R. Peschanski, hep-ph/9605298 (1996).

- [70] E. Gotsman, E. Levin & U. Maor, hep-ph/9606280 (1996).
- [71] M. Diehl, Z. Phys. **C 66** (1995) 181.
- [72] P.V. Landshoff & O. Nachtmann, Z. Phys. **C 35** (1987) 405.
- [73] A. Edin, G. Ingelman & J. Rathsman, Phys. Lett. **B 366** (1996) 371.
- [74] W. Buchmüller & A. Hebecker, Phys. Lett. **B 355** (1995) 573.
- [75] W. Buchmüller, A. Hebecker & M. F. McDermott, *Gluon radiation in Diffractive Electroproduction*, DESY 96-126.
- [76] M. F. McDermott & G. Briskin, *Future Physics at HERA Workshop 1995/96*, Vol. 2, p.691.
- [77] M. Genovese, N. N. Nikolaev & B. G. Zakharov, Phys. Lett. **B 378** (1996) 347.
- [78] M. Diehl, *Open Charm Production in Diffractive ep Collisions*, CPTH-S492-0197.
- [79] W. Buchmüller, A. Hebecker & M. F. McDermott, *Charm as a Key to Diffractive Processes*, DESY 97-035.
- [80] C. Adloff et al., Z. Phys. **C 72** (1996) 593.
- [81] L. Alvero, J. C. Collins & J. J. Whitmore, *Tests of Factorisation in Diffractive Charm Production and Double Pomeron Exchange*, hep-ph/9806340.
- [82] ZEUS Collaboration, *$D^{*\pm}$ Meson Production in Deep Inelastic Scattering at HERA with the ZEUS Detector*, contribution N-643 to EPS97, Jerusalem 1997.
- [83] H1 Collaboration, *A Measurement of the Production of $D^{*\pm}$ Mesons in Deep-Inelastic Diffractive Interactions at HERA*, contribution pa02-060 to ICHEP96, Warsaw 1996.
- [84] *The ZEUS detector*, Status Report, DESY, 1989.
- [85] *The ZEUS detector*, Status Report, DESY, 1993.
- [86] N. Harnew et al., Nucl. Inst. Meth. **A279** (1989) 290.
- [87] H. Bethe and W. Heitler, Proc. Roy. Soc. **A146** (1934) 83.
- [88] ELEC5 Electron Finder, Phantom User Library.
- [89] LOCAL Electron Finder, Phantom User Library.

- [90] A. Abramowicz, A. Caldwell & R. Sinkus, Nucl. Instr. Meth. **A365** (1995) 508.
R. Sinkus, Nucl. Inst Methods **A 361** 290 (1995).
- [91] EMILLE Electron Finder, Phantom User Library.
- [92] J. Ng and W. Verkerke *An Overview of SRTD Analysis*, ZEUS-note 95-037.
- [93] GEANT 3.13: R. Brun et al., CERN DD/EE/84-1 (1987).
- [94] ZGANA, ZEUS trigger simulation library, Els de Wolf (Editor) et al.
- [95] G.F. Hartner et al. *VCTRAK (3.07/04): Offline OutPut Information*, ZEUS-note 97-064.
- [96] L. Gladilin, Private Communication;
N. McCubbin, Private Communication.
- [97] P. Billoir and S. Qian, Nucl. Inst. Meth. **A311** (1992) 139.
- [98] S. Schlenstedt, EVTAK routine, Phantom Library.
- [99] K. Piotrkowski and M. Zachara, *Determination of the ZEUS Luminosity in 1994*, ZEUS-note 95-138.
- [100] Song Ming Wang *Study of Calorimeter Noise in the '95 data*, ZEUS-note 96-121.
- [101] S. Kananov *The Noise Study for the 1995 Data in the ZEUS Calorimeter*, ZEUS-note 97-018.
- [102] A. Savin *Study of Calorimeter Noise in the 1996 data*, ZEUS-Note 98-07.
- [103] C. Cormack, Private Communication;
ZEUS Coll., Abstract 769, *Measurement of the Proton Structure Function F_2 in e^+p Collisions at HERA*, ICHEP98, Vancouver, July 1998.
- [104] A. Bamberger et al. NIM **A401** (1997) 63-80.
- [105] T. Doeker, A. Frey & M. Nakao, *Electron Position Reconstruction — Update of the ELECPO routines — (May 1994)*, ZEUS-note 94-123.
- [106] H.-J. Grabosch, A. Meyer & S. Schlenstedt, *X_0 determination and electron energy correction with the presampler*, ZEUS-note 98-045.

- [107] Bin Lu and Luke Mu, *Calorimeter Energy Corrections for the 1994 Data*, ZEUS-note 95-098.
- [108] G. Briskin, A. Caldwell *Comparison of CAL Energy for Electrons with CTD momentum*, ZEUS-note 95-035.
- [109] Y. Iga, S. Schlenstedt *Status and Improvements of the ZEUS Calorimeter Simulation*, ZEUS-note 97-019.
- [110] J. Labs, CALCORR ROUTINE, 22nd May 1996.
- [111] A. Caldwell, W. Liu, B. Straub, *BCAL Electron Studies — Part I: Monte Carlo and Test Beam Energy Studies*, ZEUS-note 98-002.
- [112] A. Caldwell, W. Liu, B. Straub, *BCAL Electron Studies — Part II: Energy Scale Calibration*, ZEUS-note 98-018.
- [113] G. Briskin, *Diffractional Dissociation in ep Deep Inelastic Scattering*, PhD. Thesis, School of Physics and Astronomy, Tel Aviv University (1997).
- [114] F. Jacquet & A. Blondel, *An ep Facility for Europe*, 79/48 (1979) 391, ed. U. Amaldi.
- [115] Proceedings of the Future Physics at HERA Workshop (1991), Vol. 1, edited by W. Buchmüller and G. Ingelman, p. 23.
- [116] M. Derrick et al. *Zeitschrift f. Phys.* **C69** (1996) 607-620.
- [117] G. Ingelman, *Physics at HERA*, Vol. 3 (1992) 1366, Eds W. Buchmüller & G. Ingelman.
- [118] A. Kwiatkowski, H. Spiesberger & H.-J. Möhring, Proc. of the HERA workshop, DESY, Hamburg, Vol. 3 (1992) 1294.
- [119] B. Andersson et al., *Z. Phys.* **C43** (1989) 621.
- [120] B. Andersson et al., *Phys. Rep.* **97** (1983) 31.
- [121] T. Sjöstrand, *Comp. Phys. Comm.* **39** (1986) 347;
T. Sjöstrand & M. Bengtsson, *Comp. Phys. Comm.* **43** (1987) 367.
- [122] ZEUS Coll. *Exclusive ρ^0 Production in Deep Inelastic Scattering at HERA*, DESY 95-133.

- [123] E. Gallo *The TLT and DST Filter for the DIS group in 1995*, ZEUS-note 96-001.
- [124] T. Carli et al., *Physics at HERA*, Vol. 3 (1991) 1468.
- [125] A. Ebel and F. Zetsche *Analysis of Elastic QED Compton Scattering Events using the 1994 ZEUS Data*, ZEUS-note 96-020.
- [126] D. Roff *Luminosity Measurements from Virtual Compton Scattering*, ZEUS-note 97-026.
- [127] M. Derrick et al., *Zeitschrift f. Phys.* **C72** (1996) 399-424.
- [128] H. Jung, *Comp. Phys. Comm.* **86** (1995) 147.
- [129] J. Breitweg et al., *Measurement of the Diffractive Cross Section in Deep Inelastic Scattering using ZEUS 1994 data*, DESY 98-084.
- [130] L. Lönnblad, *Comp. Phys. Comm.* **71** (1992) 15.
- [131] G. Marchesini et al., *Comp. Phys. Comm.* **67** (1992) 465.
- [132] M. Derrick et al., *Zeit. f. Phys.* **C 68** (1995) 569.

Acknowledgments

There are many, many people who have helped and supported me during the last three years to whom I would like to say thank you — if I've forgotten anyone, I'm sorry !

First of all, my thanks go to my supervisor, Dr Julia Sedgbeer and her “unofficial deputies”, Drs Alex Prinias and Luciano Lamberti, for all their advice and assistance right from the start of my time on ZEUS. None of this thesis would have been possible without their support and calming influence over the last three years.

I also owe a lot to my family for being there when I needed them (even if it does mean getting up at 6 am every day !) and for providing me with regular doses of “de-stress”.

There are many people both at DESY and at IC who have helped me out on many different occasions: Halina Abramowicz and Kryzstof Piotrkowski, who provided valuable assistance and generally great physics coordinating and Arnulf Quadt for sound advice and proving the old adage, “ If a job is worth doing, it is worth doing properly”. I also want to thank Ken Long for his enthusiasm, bright ideas and sound career advice.

My thanks also go to the diffractive D^* gang — Olivier and especially Jason for emergency analysis and assistance. The thanks in this department would not, however, be complete without mentioning Richard — all-round hero, friend and partner in crime — the last three years would have been an awful lot less fun without you !

I also send my general thanks to the Zarah team, sorry for all those inadvertent jobpurges !

Then, of course, there are all the guys in the ZEUS UK group who made my time in Hamburg an entertaining experience. Thanks to: Helen, Neil, Jo, Rob, Matt, Dave B. Rod, Alex T., Mark, Olaf, Chris, Laurel and Pat, Dimitri and Gareth. Special thanks to Val for lots of sound advice and a good gossip ! Thanks also go to the gang at IC for all those coffees and beer trips to the Union: Sue, Phil, Dave C., Richard, Katherine, Malcolm, Roly, Alex, Matt and Anna.

More special thanks to Martin for answering my daft questions and also to Ash and Don for long-distance HV advice. Finally, my thanks to PPARC for their funding of me and the research in this thesis.