DEUTSCHES ELEKTRONEN-SYNCHROTRON

DESY-THESIS-1999-018
June 1999

Measurements of the Hadronic Final State in Deep Inelastic Scattering at HERA

by

J. R. Okrasiński

ISSN 1435-8085

NOTKESTRASSE 85 - 22607 HAMBURG

To be sure that your reports and preprints are promptly included in the HEP literature database
send them to (if possible by air mail):

DESY Zentralbibliothek Notkestraße 85 22603 Hamburg Germany	DESY Bibliothek Platanenallee 6 15738 Zeuthen Germany

The Pennsylvania State University

The Graduate School

Eberly College of Science

MEASUREMENTS OF THE HADRONIC FINAL STATE IN DEEP INELASTIC SCATTERING AT HERA

A Thesis in

Physics

by

Jan R. Okrasiński

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 1999

We approve the thesis of Jan R. Okrasiński

Date of Signature

Benedict Y. Oh Professor of Physics Chair of Committee	
James J. Whitmore Professor of Physics Thesis Adviser	
Jayanth R. Banavar Professor of Physics Head of the Department of Physics	
Gerald A. Smith Professor of Physics	
Peter D. Usher Professor of Astronomy and Astrophysics	

Abstract

Charged hadrons produced in Deep Inelastic e^+p Scattering (DIS) at a center-of-mass energy of $\sqrt{s} = 300$ GeV are studied with the ZEUS detector at HERA. Single-particle distributions as a function of the scaled momentum, x_P, as well as correlations in the two-particle relative angle are measured in the current region of the Breit frame-of-reference for the four-momentum transfer squared, $10 < Q^2 < 10240$ GeV^2, and the Bjorken scaling variable, $0.0006 < x_{Bj} < 0.40$. The results are compared to analytic QCD calculations, leading order Monte Carlo models and existing experimental data. Support for fragmentation universality is found in the scaled momentum distributions; discrepancies between e^+e^- and DIS exist for angular correlations. QCD calculations in the analytic perturbative approach typically provide a qualitative, but not always quantitative, description of the data indicative of possible limitations of the Local Parton-Hadron Duality hypothesis, or importance of subleading corrections in the analytic part of the calculation.

Contents

List of Figures ix

List of Tables xiii

Acknowledgements xiv

1 Introduction 1

2 Hadronic Final States in DIS 3

2.1 Deep Inelastic Scattering 4

2.1.1 Kinematics 4

2.1.2 Historical Background 6

2.1.3 DIS at HERA 8

2.2 Charged Particles in Hadronic Final State 12

2.2.1 The Breit Frame-of-Reference 13

2.2.2 Parton Cascades in QCD 15

3 The ZEUS Experiment at HERA 19

3.1 The HERA Storage Ring 19

3.2 The ZEUS Experiment Environment 22

3.2.1 The Uranium Calorimeter 25

3.2.2 The Central Tracking Detector . . . 26
3.2.3 The Forward and Rear Tracking Detectors . . . 27
3.2.4 Peripheral Subdetectors . . . 28
3.2.5 The Trigger . . . 30

4 Reconstruction **31**

4.1 Reconstruction at ZEUS . . . 32
4.2 DIS Electron . . . 33
4.2.1 Electron Identification Algorithms . . . 33
4.2.2 Sinistra: a Neural Network Electron Finder . . . 35
4.2.3 Position Correction . . . 36
4.2.4 Energy Corrections . . . 36
4.3 Charged Particles . . . 38
4.3.1 The VCTRAK package . . . 38
4.3.2 Track Reconstruction Algorithm . . . 39
4.3.3 Primary Vertex . . . 40
4.4 Kinematic Variables . . . 41
4.4.1 Electron Method . . . 42
4.4.2 Jacquet-Blondel Method . . . 45
4.4.3 Double Angle Method . . . 46
4.4.4 Choice of the Method . . . 51
4.5 Breit Frame . . . 51
4.5.1 Virtual Photon . . . 51
4.5.2 Lorentz Boost and Rotation . . . 52
4.5.3 Current Fragmentation Region . . . 55

5 Monte Carlo and Detector Simulation **58**

5.1 Event Generators . . . 59
5.1.1 Parameterizations . . . 60
5.1.2 Hard Subprocess . . . 61
5.1.3 Parton Showers . . . 61
5.1.4 Hadronization . . . 62
5.1.5 QED Radiative Corrections . . . 63
5.2 Detector Simulation . . . 63
5.3 Monte Carlo Samples . . . 64
5.3.1 Description of the Kinematic Variables . . . 64
5.3.2 Description of the Tracking Variables . . . 69
5.4 Detector Effects . . . 71
5.4.1 Resolution . . . 74
5.4.2 Efficiency and Purity . . . 74
5.4.3 Correction Methods . . . 75
5.5 Monte Carlo Predictions . . . 75

6 Data Selection **77**

6.1 Trigger . . . 77
6.1.1 First Level Trigger . . . 78
6.1.2 Second Level Trigger . . . 78
6.1.3 Third Level Trigger . . . 79
6.2 Off-line Preselection . . . 80
6.2.1 Low Q^2 Data Sample . . . 81
6.2.2 High Q^2 Data Sample . . . 81
6.3 Track Selection . . . 82
6.4 Cleaning Cuts, Backgrounds . . . 83

6.5 Kinematic Selection 85
6.5.1 Phase Space Cuts 86
6.5.2 Binning 86
6.6 Final Data Sample 89

7 Fragmentation Functions **91**
7.1 Resolution and Binning 92
7.2 Uncorrected Data 95
7.3 Correction Factors 100
7.4 Corrected Distributions and Systematic Uncertainties 103
7.5 Summary 109

8 Angular Correlations **112**
8.1 Measurement Techniques 113
8.1.1 Two-particle Production Density 113
8.1.2 Scaling Variable 114
8.1.3 Resolution and Binning 115
8.1.4 Normalization 115
8.2 Uncorrected Data 118
8.3 Correction Procedure 121
8.4 Corrected Data 125
8.5 Summary 132

9 Discussion **133**
9.1 Single-Particle Distributions 133
9.1.1 Charged Multiplicity 133
9.1.2 Scaled Momentum Distributions 134

9.1.3 $\ln(1/x_p)$ Distributions vs MLLA Predictions 140
9.1.4 Summary 153
9.2 Two-Particle Angular Correlations 153
9.2.1 Relative Angular Distribution $\hat{r}_2$ 153
9.2.2 Angular Correlation Function r_2 158
9.2.3 Summary 161

10 Summary **164**

Bibliography **167**

List of Figures

2.1 Leading order diagram for DIS. 5

2.2 Phase space for various DIS experiments in the (x, Q^2) plane. 9

2.3 A high Q^2 DIS event. 11

3.1 The HERA accelerator layout. 20

3.2 Overview of the ZEUS detector 24

4.1 Resolution of x_e in kinematic bins. 43

4.2 Resolution of Q_e^2 in kinematic bins. 44

4.3 Resolution of y_{JB} in kinematic bins. 47

4.4 Resolution of x_{DA} in kinematic bins. 49

4.5 Resolution of Q_{DA}^2 in kinematic bins. 50

4.6 DIS in the QPM approximation as seen in the Breit frame. 52

4.7 Resolution of E_e^γ from the electron method plotted in kinematic bins. 53

4.8 Resolution of E_{DA}^γ from the Double Angle method plotted in kinematic bins. 54

5.1 Feynman diagrams for the $\mathcal{O}(\alpha_s)$ processes in DIS. 61

5.2 Electron energy distribution in uncorrected data, ARIADNE and HERWIG. 66

5.3 $E - p_Z$ distributions for uncorrected data, ARIADNE and HERWIG 67

5.4 Normalized distributions of kinematic variables for uncorrected data, ARIADNE and HERWIG. 68

5.5 η_{max} distribution for uncorrected data, ARIADNE and HERWIG 69

5.6 z_{VTX} distribution in uncorrected data and MC. 70

5.7 Track multiplicity distributions in uncorrected data and MC. 72

5.8 Track pseudorapidity distributions in kinematic bins for uncorrected data and ARIADNE. 73

6.1 Analysis bins in the x–Q^2 plane. 88

7.1 Resolution of x_P in kinematic bins. 93

7.2 Resolution of $\ln(1/x_P)$ in kinematic bins. 94

7.3 Uncorrected x_P distributions from the 1995 data sample 96

7.4 Uncorrected $\ln(1/x_P)$ distributions from the 1995 data sample 97

7.5 Ratio of the uncorrected x_P distributions from the 1996 and 1997 data samples. . . . 98

7.6 Ratio of the uncorrected $\ln(1/x_P)$ distributions from the 1996 and 1997 data samples. 99

7.7 Correction factors to the normalized x_P distributions. 101

7.8 Correction factors to the normalized $\ln(1/x_P)$ distributions. 102

7.9 Dominant systematic shifts to the x_P distribution. 105

7.10 Corrected x_P distribution from the 1995 data sample 107

7.11 Corrected $\ln(1/x_P)$ distributions from the 1995 data sample 108

7.12 Comaprison of the corrected x_P distributions in the 1995 and 1996/97 data samples. . 110

7.13 Comaprison of the corrected $\ln(1/x_P)$ distributions in the 1995 and 1996/97 data samples. 111

8.1 The principle of a measurement of relative angle. 113

8.2 Resolution of the scaling variable ϵ in kinematic bins 116

8.3 Uncorrected relative angular distributions from the 1995 data sample, ARIADNE and HERWIG. 119

8.4 Uncorrected angular correlation function from the 1995 data sample, ARIADNE and HERWIG. 120

8.5 Ratio of the uncorrected relative angular distributions from the 1996 and 1997 data samples. . . . 122

8.6 Diffrence of the uncorrected correlation functions from the 1996 and 1997 data samples. 123

8.7 Correction factors to the relative angular distribution. . . . 124

8.8 Correction factors to the angular correlation function. . . . 126

8.9 Dominant systematic shifts to the relative angular distribution. . . . 127

8.10 Dominant systematic shifts to the angular correlation function. . . . 128

8.11 Corrected relative angular distribution from the 1995 data sample. . . . 130

8.12 Corrected angular correlation function from the 1995 data sample. . . . 131

9.1 The mean charged multiplicity evolution with energy. . . . 135

9.2 The mean charged multiplicity compared to MC calculations. . . . 136

9.3 The measured scaled momentum cross-section, $1/\sigma\ d\sigma/dx_P$, as a function of Q compared to DIS and e^+e^- data. . . . 139

9.4 The measured scaled momentum cross-section, $1/\sigma\ d\sigma/dx_P$, as a function of Q compared to QCD calculations. . . . 141

9.5 The measured scaled momentum cross-section, $1/\sigma\ d\sigma/dx_P$, as a function of Q compared to MC models. . . . 142

9.6 Limiting MLLA spectrum fit to the $\ln(1/x_P)$ distribution. . . . 143

9.7 Evolution with energy of Λ_{eff} from MLLA fit to $\ln(1/x_P)$ distribution. . . . 145

9.8 The Distorted Gaussian spectrum fit to the $\ln(1/x_P)$ distribution. . . . 146

9.9 Energy evolution of the peak of the $\ln(1/x_P)$ distribution. A single-parameter MLLA fit is overlaid. . . . 147

9.10 Energy evolution of the width of the $\ln(1/x_P)$ distribution. A single-parameter MLLA fit is overlaid. . . . 148

9.11 Evolution with energy of the peak of the $\ln(1/x_P)$ distribution from a distorted Gaussian fit. . . . 150

9.12 Evolution with energy of the width of the $\ln(1/x_P)$ distribution from a distorted Gaussian fit. . . . 151

9.13 Evolution with energy of the skewness of the $\ln(1/x_P)$ distribution from a distorted Gaussian fit. . . . 152

9.14 The relative angular distribution dependence on the scale parameter Λ. . . . 154

9.15 The relative angular distribution dependence on the cone width and comparison to Monte Carlo models. . . . 156

9.16 The relative angular distribution compared to e^+e^- data and DLA calculations. . . . 157

9.17 The angular correlation function dependence on the scale parameter Λ. . . . 159

9.18 The angular correlation function dependence on the cone width and comparison to Monte Carlo models. . . . 160

9.19 The angular correlation function compared to e^+e^- data and DLA calculations. . . . 162

List of Tables

4.1 Track migration rates in the Breit frame. 57

6.1 Kinematic bin definition, efficiency and purity (ARIADNE). 87

6.2 Statistics of the data in kinematic bins. 90

9.1 Mean charged multiplicity in kinematic bins. 137

9.2 Peak position, width and skewness of the $\ln(1/x_P)$ distribution from distorted Gaussian fit. 150

Acknowledgements

To Elżbieta, Alicja and Adam

I would like to thank all the patient people around me who helped me to survive the process of the preparation of this thesis.

Although it is expected of faculty in general and thesis advisers in particular to positively interfere with PhD candidates, I was lucky indeed to be supervised by professors of not only deep and broad knowledge of the subject but also of nearly infinitely patient attitude towards the choking process of my analysis and later writing up. For that reason I am indebted to Prof. James Whitmore, my thesis adviser, and Prof. Benedict Oh, the chair of my doctoral committee. Prof. William Toothacker, during his stay in Hamburg, allowed me to ask all kinds of silly questions providing correct yet often funny answers to most of them.

The first two years of my research assistantship were supported by Division of High Energy Physics, Argonne National Lab. At the same time I was given an opportunity to be guided by Malcolm Derrick, Jose Repond and Steve Magill. I hope they consider their effort to have been allocated well even though my first steps in ZEUS were rather difficult making my progress not too fast.

My studies would have been impossible without kind assistance from my family: they allowed me to live away from home for a couple of long years barely complaining of my prolonged absence most of the time limiting themselves to words of love and encouragement. I couldn't find an appropriate expression of gratitude in my native Polish language let alone in English but I must at least say

"Thank you, Elżbieta, Alicja and Adam" here. I'm coming home, at last.

My Mother-in-law and Mother did their best to substitute for me in the home routine so that I was able to work in Hamburg. Their efforts *are* appreciated.

There are four people who contributed a lot to the final shape of my analysis. They introduced me to the subject and then shared their knowledge with me in long discussions. They are Nick Brook, Sergei Chekanov, Tony Doyle and Leszek Zawiejski. I would have probably abandoned this subject without all of you!

Since my memory is not that good and I've been studying for a long time indeed, I may not be able to recall all the people deserving a credit here. Certainly they are Jeff Bulmahn, Greg Feild, and Jon Butterworth, former Penn Staters. Working on the SRTD-FLT project and the detector maintenance I had a pleasure to meet Richard Talaga, John Dawson, Henk Boterenbrood, Vladimir Hain, Jon Labs, Johnny Ng, Katsuo Tokushuku, Roberto Carlin, Antonio Pellegrino, and Pat Saull. There was a lot to learn from them and I tried to do that.

The strong and nice Polish society at DESY was there when I wanted fun or help. Let me name only a few of them: Radek Pawlak, Kaśka Klimek, Maciek Przybycień, Jurek Andruszków, Karol Bartkiewicz, and Janusz Chwastowski. See you in Poland!

Teresa and Andrzej Badzian very kindly greeted me as a newcomer to State College, hosted me and my wife and were the people I could count on.

Jerie Ann Zitek's remote help with formalities was always so transparent and efficient that I could very easily overlook it.

Laura Iannotti joined the Penn State team during my final struggle with the thesis and she kept me from wasting my time on anything else but editing: this is why I was able to meet the deadlines.

Chapter 1

Introduction

Deep inelastic scattering is a way of probing the properties of hadronic matter by means of a leptonic probe. At HERA, insight is possible to a depth not accessible before: the structure of the proton at scales down to 10^{-18} m can be studied. Another approach is to treat the proton as a source of nearly free partons and to study the nature of their interactions with a lepton to obtain, in a perturbative regime, the strong coupling constant α_S. The third possibility, adopted in this thesis, focuses on the properties of the hadronic final state and the process of parton fragmentation. It is in many aspects complementary to the first one, as non-perturbative phenomena are encountered in both approaches so that a reference to phenomenological description is needed.

The organization of this thesis is as follows. Chapter 2 contains a general overview of physics related to deep inelastic scattering with an emphasis on the hadronic final states to which this report is devoted. The current standing of applicable theoretical models is summarized with a connection to experimental status. In chapter 3 the ZEUS detector and HERA accelerator are described with a limited amount of details relevant for the analysis. Reconstruction of deep inelastic scattering events is the topic of chapter 4. Limitations of the reconstruction process and effects introduced by the experimental apparatus are studied in chapter 5 using Monte Carlo techniques and detector simulation. The understanding gained is employed in chapter 6 to justify and describe the selection of the data sample. A measurement of the fragmentation function, a single particle inclusive observable, is presented in chapter 7 together with correction procedures and an uncertainty estimation. Similarly,

results on collective observables, the relative angular correlations of charged particles, are reported in chapter 8. A discussion of the results in the context of Quantum Chromodynamics follows in chapter 9. In particular the universality of hadronization is studied by making a comparison to the data from electron-positron annihilation. Finally, chapter 10 summarizes the physics message of this thesis.

Chapter 2

Hadronic Final States in DIS

Having both a lepton and a hadron in the inital state, deep inelastic scattering (DIS) is situated in the middle between e^+e^-annihilation and hadron-hadron scattering. On one hand, due to the presence of a strongly interacting hadron DIS cannot benefit from the simplicity and clarity of the leptonic annihilation experiments; on the other hand, the cross-sections involved are orders of magnitude lower compared to hadron colliders. Had there been no advantages to outweigh these drawbacks DIS would never have been paid attention to but it has been pursued for more than three decades now using apparatus worth billions and making dozens of physicists busy. Though the initial interest in DIS of the possibility to probe hadronic inner structure at yet smaller and smaller scales has not been abandoned, new areas have opened with the advent of HERA: high enough energies allow a deeper insight into the nature of the strong interactions as well as lead to an enriched structure of the hadronic final state making it a valuable field of study. In this thesis particles building up the final state are investigated in terms of inclusive distributions as well as their angular relationships.

This chapter tells the story of the DIS with its motivation, progress made over the years and, finally, gives an overview of the specifics of the world's highest energy DIS environment provided by HERA. An introduction follows to the main topics of the hadronic final state in DIS in general and to the measurements reported in this thesis in particular. Some theoretical background is also outlined.

2.1 Deep Inelastic Scattering

It is evident that the role of deep inelastic scattering in the field of elementary particle physics can be hardly overestimated due to its influence on the birth and early period of Quantum Chromodynamics (QCD), the theory of strong interactions that bind together quarks into hadrons. Today, a few decades after the experimental and theoretical milestones were made, the theory is unquestioned but its ultimate goal, the predictive power, is still impeded by serious mathematical difficulties. As a result the experimental data are used not only to constrain the only free parameter of QCD, the strong coupling constant, α_S, but also to guide the theory through the unavoidable calculational approximations. DIS remains an important test field of QCD and phenomenological concepts devised to help it in making a connection between its natural domain, the partons, and the world accessible to the experiment, the hadrons.

2.1.1 Kinematics

It would be instructive to introduce the kinematics of DIS in the very beginning. The leading order Feynman diagram of the process is shown in Fig. 2.1. Apart from the center-of-mass energy squared, $s = (p_l + p_h)^2 \approx 2p_l \cdot p_h$, fixed for any given experimental conditions by the four-momenta of the colliding lepton, p_l, and that of the hadron, p_h, only two independent Lorentz invariants are needed to fully define the kinematics.[1] These are customarily drawn from the set of the following three:

- the negative four-momentum transfer to the hadron squared, Q^2;
- the lepton fractional energy transfer *in the hadron rest frame*, y;
- the Bjorken scaling variable, x.

[1] Transverse polarization of the particles in collision would break the axial symmetry but no such data have been collected to date.

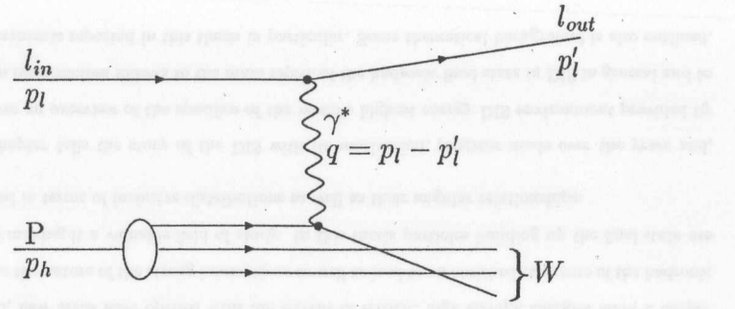

Figure 2.1: Leading order diagram for DIS.

It is an advantage of DIS that the kinematics can be deduced independent of the hadronic final state from the scattered lepton alone.[2] The relevant formulae are:

$$Q^2 = -(p_l - p'_l)^2 = 2p_l \cdot p'_l, \tag{2.1}$$

$$y = \frac{p_h \cdot (p_l - p'_l)}{p_h \cdot p_l}, \tag{2.2}$$

and

$$x = \frac{-(p_l - p'_l)^2}{2p_h \cdot p_l} \tag{2.3}$$

where p'_l is the four-momentum of the outgoing lepton. The transverse resolution at which the hadron is probed is inversely proportional to Q; at the same time, Q sets the energy scale of the hard scattering. In an approximation in which a rapidly moving hadron is formed of a bunch of non-interacting partons, as stated by the Quark Parton Model introduced in the next subsection, x can be interpreted as the hadron's fractional longitudinal momentum carried by the scattered parton. y is of particular interest in fixed-target DIS where the hadron rest frame coincides with the laboratory frame. Both dimensionless variables can vary in the range between 0 and 1, whereas Q^2 is bound to

[2] Except for the case of the undetectable outgoing lepton, the neutrino. At HERA these (charged current) events are rare and are beyond the scope of this thesis.

remain below s but practically the rapid fall-off of the cross-section imposes a more restrictive limit dependent on the statistics of the data. There exists a simple relation between the three quantities

$$Q^2 = sxy \tag{2.4}$$

given here in an approximation of negligible masses of the colliding particles.

In addition to the above variables, the invariant mass squared of the hadronic system,

$$W^2 = (p_h + p_l - p_l')^2 \approx sy - Q^2, \tag{2.5}$$

is used mainly by analyses performed in the hadron CMS system to define the phase space available for particle production.

2.1.2 Historical Background

As pointed out in the previous subsection, DIS provides a relatively clean way of probing hadronic structure with a resolution which improves with energy. In the late 1960's, in addition to hints based on a neat classification of hadrons founded by Gell-Mann in 1961 and subsequently developed by Gell-Mann and Zweig, evidence from DIS experiments at SLAC for point-like constituents of the nucleon began to pour in. The evidence relied on the interpretation of the observed energy independence of the nucleon structure functions, F_i, entering a convenient parameterization of the double-differential DIS cross-section

$$\frac{d^2\sigma}{dQ^2dx} = \frac{4\pi\alpha^2}{xQ^4}[y^2xF_1(x,Q^2)] + (1-y)F_2(x,Q^2)]. \tag{2.6}$$

The independence, usually referred to as scaling behavior, had been predicted by Bjorken as a consequence of quasi-elastic scattering off pointlike partons to be found inside a nucleon. Furthermore, the Callan-Gross relation between the two scaling structure functions,

$$2xF_1(x) = F_2(x), \tag{2.7}$$

following from the fermionic nature (spin 1/2) of the partons as postulated in the constituent quark model by Gell-Mann and Zweig was in fact confirmed to hold. Yet another characteristic of the quarks á la G-M–Z, their fractional electric charges, was positively verified with the help of bubble-chamber

neutrino DIS experiments: the ratio of the respective structure functions measured in νN and eN was consistent with the expected value of 18/5.[3]

Taking into account the facts listed above, the Quark Parton Model (QPM) was formulated. It claims a hadron to consist of non-interacting partons identified with quarks. In addition to the valence partons, three quarks in the case of baryons and a quark-antiquark pair for mesons, a multitude of additional sea-quark pairs was proposed to exist in a hadron. The QPM explanation of the DIS data as scattering on quasi-elastic partons was done in the frame of a fast moving hadron; it relied on the relativistic time-dilation effectively freezing the residual interactions between quarks bound inside a hadron over the short interval of the actual hard scattering process. In this frame, the physical content of the structure function is in its relation to the probability density, $f_q(x)$, for a quark to carry a longitudinal fractional momentum x:

$$F_2(x) = \sum_q e_q^2 x f_q(x), \tag{2.8}$$

where the sum runs over all quark and antiquark species.

The QPM was never expected to be a complete approach to the hadronic substructure for the obvious reason of neglecting the force gluing quarks together. Soon an experimental proof of the existence of partons not taking part in DIS was given through a measurement of the total quark contribution to the longitudinal momentum of a hadron, proportional to the integral of the structure function, $\int_0^1 dx F_2(x)$: the number turned out to be about a half, indicating the presence of a significant admixture of matter transparent to the DIS probes. At this point in time it became necessary to formulate a quantum field theory of the nucleon binding forces. For this purpose the long-lived color degree of freedom was revisited: it was originally postulated as an *ad hoc* fix for the apparent violation of the Pauli exclusion principle in the baryon decuplet.[4] A gauge theory of three-colored quarks and spin-one gluons mediating the strong interactions was gradually developed and called Quantum Chromodynamics.

QCD is much like the successful Quantum Electrodynamics (QED), but several important differ-

[3] To be accurate, for the isosinglet target in which there are equal numbers of u and d quarks.

[4] Among the lightest spin 3/2 baryons there are three composed of *three* like quarks each with a completely symmetric wavefunction excluding the proposed color degree of freedom.

ences need to be pointed out. The gluons are themselves strongly interacting particles due to their non-vanishing color charge. A non-trivial implication of this fact is the property of asymptotic freedom possessed by QCD: the strength of the coupling, *decreases* with energy scale as opposed to the growth observed in QED; this is because at high energies only the true color charge of a particle is probed, whereas at low energies (large distances) the entire cloud of virtual particles carrying their own color charges is sensed. It also follows from this anti-screening that at some large distance the growing coupling makes the perturbative approach inappropriate for describing bound states of partons; color charges are thus *confined* into singlet combinations with the total color charge exactly balanced. Even though this is a challenge for those who want to use QCD, it also helps to explain why only so late the color degree of freedom was discovered.

Introduction of an additional class of partons, the gluon, is not the only implication of QCD for DIS. It was soon predicted that the scaling of structure functions is only approximate as a logarithmic dependence on Q^2, the scaling violation, is expected. Newer more accurate measurements at SLAC readily confirmed that, manifestly supporting the just founded theory. Since then, QCD has matured into an unquestioned model of strong interactions, thus becoming a part of the Standard Model of elementary particles. Implications of QCD relevant for the analysis presented in this thesis are gathered in section 2.2.

2.1.3 DIS at HERA

With the commissioning of HERA a completely new kinematic range has been opened for the deep inelastic scattering to explore. This is illustrated in Fig. 2.2 in the kinematic plane (x, Q^2). The fixed target experiments, including HERMES operating in HERA's lepton beam, are typically bound below $y < 0.005$, whereas the HERA collider data can reach almost up to the physical limit of $y = 1$; by the same means both the x and Q^2 ranges have been extended by more than two orders of magnitude. With beam energies at $E_l = 27.5$ GeV and $E_p = 820$ GeV the center-of-mass energy of $\sqrt{s} = 300$ GeV is reached. Since the commissioning, the collider has been operated with either electrons or positrons brought into head-on scattering with protons; running with heavier nuclei is under consideration for the more distant future. In this thesis a high luminosity e^+p sample is analysed; the term "electron"

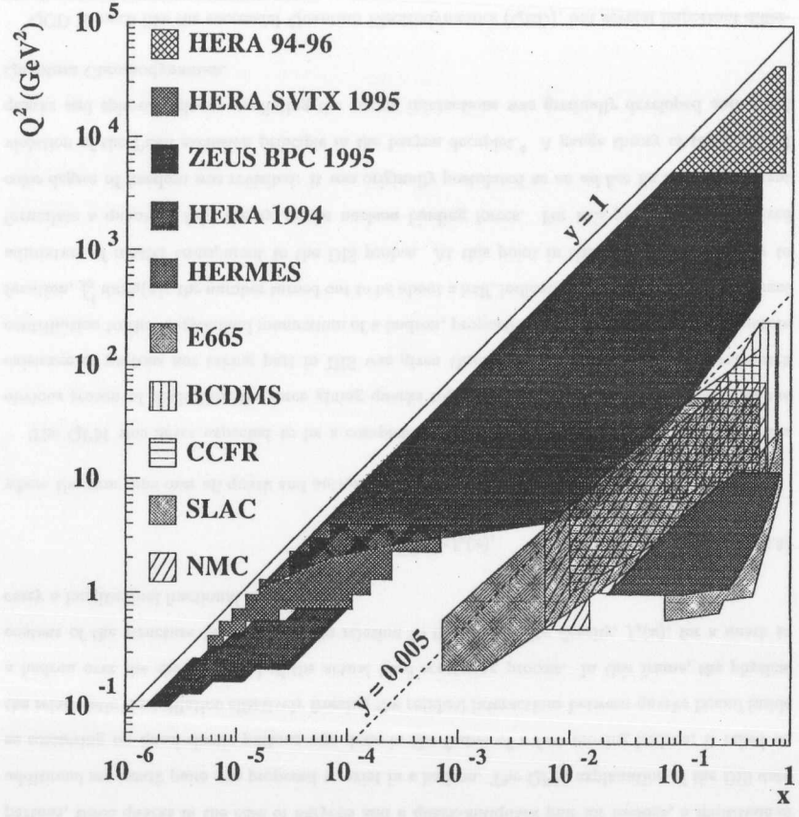

Figure 2.2: Phase space for various DIS experiments in the (x, Q^2) plane. Fixed-target experiments typically are bound below $y = 0.005$ (dashed line); HERA reaches almost up to the kinematic limit $y = 1$ (solid).

is used most of the time in reference to the positron unless clearly stated otherwise.

An example of a high Q^2 event registered by the ZEUS detector is depicted in Fig. 2.3: an energetic DIS positron is balanced by a well collimated hadronic jet as seen in both the tracker and the calorimeter; calorimetric activity in the vicinity of the beam pipe is due to the proton remnant. It is a characteristic feature of HERA, that the laboratory frame is neither the center-of-mass system nor the rest frame of any of the colliding particles. Moreover, a large imbalance exists in favor of the proton for the longitudinal momentum, typically leading to a significant amount of energy originating from the proton debris being measured in the forward direction. More details on the HERA machine and the ZEUS detector are provided in chapter 3.

Since the first physics results from HERA saw the light in 1992, more than 150 publications have been published or submitted to major journals. In the competition between photoproduction[5] and DIS, the latter claims a significant fraction of the experimental attention. The main areas of interest, only briefly enumerated here, include:

Proton structure The proton structure functions serve as a constraint for the parton distribution functions (PDFs), an unavoidable input for many QCD-based calculations. In particular, the gluon content of the proton has been extracted at low x using the logarithmic scaling violations in $F_2(x, Q^2)$ [1]. The F_2 itself has been measured in formerly unexplored kinematic regions [2, 3]. Observation of particles composed of heavy quarks shed light on the flavor composition of the quark sea in the proton, including the charm contribution to F_2 [4]. The hadronic structure of the virtual photon is accessible as well [5]. An investigation of alternative scenarios for the PDFs evolution with energy has been performed based on the low x events with a forward jet [6]. Cross sections for the charged current DIS have been reported [7].

Hard scattering Jets, the hadron level mirror of the hard scattering at parton level, have been studied to reveal details of the QCD dynamics. The dijet production in DIS is a direct consequence of the order $\mathcal{O}(\alpha_S)$ corrections to the naive QPM picture; the way back from the dijet rate measurement to the coupling constant has been accomplished [8] and extended towards analysis

[5] Scattering of quasi-real photons, $Q^2 \approx 0$.

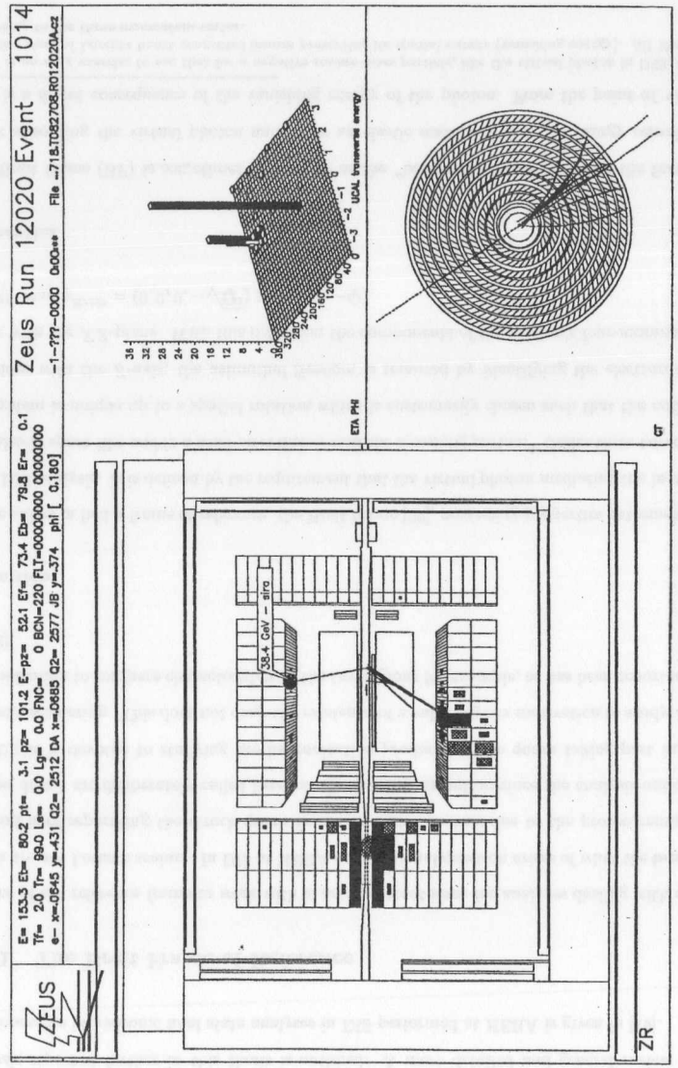

Figure 2.3: A high Q^2 DIS event recorded by the ZEUS detector.

of multi-jet events [9].

Fragmentation The structure of the hadronic final state was initially studied in DIS in terms of the energy flows but with the gradually increasing luminosity more detailed analyses have become possible. Currently, the published results range from the jet shapes [10] through a QCD-motivated analysis of event shape variables [11] to numerous studies of charged hadrons including single-particle spectra [12, 13] and Bose-Einstein correlations [14]. The subject of this thesis, single- and two-particle distributions, falls into this class of problems.

Searches Though HERA cannot be equal in CMS energy with hadron colliders, different exotic phenomena beyond the Standard Model can be sought for. The famous "high Q^{2}" excesses reported simultaneously by the H1 and ZEUS collaborations early in 1997 [15] seem to have been diluted by the newer data. However, investigations continue towards possible discovery of new particles like leptoquarks, excited fermions or supersymmetric counterparts of the known particles as well as new types of interactions [16]. Some of the searches are not necessarily limited to the DIS sample but there is typically a significant overlap, especially at the highest Q^{2}.

A separate branch of experimental analyses has been established with the observation of a significant rate of the so-called rapidity gap events [17]. These diffractive results include, among others, the respective proton structure function, $F_{2}^{D(4)}$ [18]. Virtually any measurement done for inclusive DIS can be fruitfully restricted to the diffractive subsample. The interplay of the soft (the nearly intact proton) and hard (DIS) physics to be found in diffraction keeps the field very active and interesting.

2.2 Charged Particles in Hadronic Final State

The measurement of individual particles is the most detailed approach one can adopt in a study of final states of any high energy experiment. However, only charged particles can be easily detected and tracked allowing unambiguous determination of their kinematics. In this section, after introducing a special frame known as the Breit frame-of-reference, the theoretical background to the experimental

analysis reported further in this thesis is outlined. A more detailed and general review of recent developments in hadronic final state analyses in DIS performed at HERA is given in [19].

2.2.1 The Breit Frame-of-Reference

Choice of the reference frame to work with is an important issue for analyses dealing with quantities which are not Lorentz scalars. In DIS at HERA an additional question arises of what the best possible method is of separating the struck quark from the contamination due to the proton remnant. The proton debris are deliberately called here an unwanted contribution since the analysis undertaken in this thesis is devoted to studying the hadronization products of the quark taking part in the deep inelastic scattering. This does not deny the existence of a valid physics motivation to study the target hadronization to compare characteristics of the two regions for example, as has been reported recently in [13].

Definition

There exists in fact a frame of reference, the Breit frame [20], possessing properties extremely suitable for a DIS analysis. It is defined by the requirement that the virtual photon mediating the interaction is completely space-like *and* in a head-on collision with the incoming proton.[6] Under these two conditions the system is unique up to a spatial rotation which is customarily chosen such that the collision axis coincides with the Z-axis; the azimuthal freedom is removed by identifying the electron scattering plane with the XZ-plane. With this definition the components of the photon's four-momentum, q, in the BF read $q^{Breit} = (0, 0, 0, -\sqrt{Q^{2}}) \equiv (0, 0, 0, -Q)$.

Properties

The Breit frame (BF) is sometimes referred to as the "brickwall" frame reflecting the fact that the quark absorbing the virtual photon undergoes an elastic scattering *i. e.* its energy remains intact. This is a direct consequence of the vanishing energy of the photon. From the point of view of the

[6] It is an easy exercise to see that for a negative square-mass particle, like the virtual photon in DIS, there exists an entire class of Lorentz boost-connected frames preserving its spatial nature (vanishing energy). All the boosts are *transverse* to the three-momentum vector.

QPM, where partons carry no transverse momentum with respect to the fast proton direction, it is true that the quark itself is aligned with the Z-axis and is therefore scattered backwards, exactly. It is then easily seen that the quark's four-momentum is $(Q/2, 0, 0, \pm Q/2)$, where the plus and minus signs refer to the state before and after the collision, respectively. A more realistic picture emerges from considerations including QCD corrections: the quark no longer needs to be on-mass-shell nor collinear with the photon (proton); these conditions are relaxed due to the possibility of an additional interaction with a gluon. Irrespective of the approximation adopted the proton remnant proceeds in the positive Z direction.

A clear motivation for the choice of the BF is seen from the above considerations: the frame provides a good geometric separation between the outgoing struck quark (and products of its later fragmentation) and the proton remnant since they are back-to-back. The two regions, hemispheres centered on the scattering axis, are called the current and target fragmentation regions, respectively. Following that, hadrons observed in the final state are assigned to the current region based on the sign of the Z component of momentum after transformation to the Breit frame if $p_Z^{Breit} < 0$, otherwise they fall into the target region.

Analogy to the e^+e^- Annihilation

Actually, one can consider other systems leading to a similar geometric separation as for instance the hadronic center-of-mass system, related to the Breit frame by a boost along the photon axis. The advantage of the BF is twofold. Firstly, the phase space for the struck quark fragmentation is governed by the same energy scale, Q, as are the dynamics of the hard scattering; this is opposite to the situation in the $\gamma^* p$ center-of-mass system, where the quark and the remnant balance each other with momenta $W/2$; in the BF there is no such symmetry.[7] Secondly, the current region of the BF can be considered analoguous to a single hemisphere of an e^+e^- annihilation event with a pair of quarks of energies $\sqrt{s}/2$ produced back to back, each of them equivalent to the struck quark of energy $Q/2$ in DIS. It should be stressed that there are important *dynamical* differences between the two types of hard scattering: for lepton pair annihilation, unlike DIS, there is no strongly interacting particle in

[7] In BF, proton remnant carries a higher momentum, $Q(1-x)/2x$, than the struck quark, $Q/2$.

the initial state. As far as the final state is concerned the proton remnant in DIS is not a pointlike color charge but rather an object of finite size and internal structure, again a more complex situation than in e^+e^- annihilation. Yet another difference, experimental this time, lies in the determination of the axis of an event: its choice is dependent on the event topology in e^+e^-, whereas the definition in DIS relies on detection of the scattered lepton only to reconstruct the virtual photon direction. Hence, there is no reason to expect exact equivalence between the two processes.

2.2.2 Parton Cascades in QCD

Parton Showering

In order to model distributions of individual particles in DIS, QCD needs to make a connection between a single energetic quark kicked out of the proton and a collection of hadrons observed in the final state. Somewhere on the way, a fundamental change of the treatment is necessary: there is no smooth calculable transition from partons to hadrons. The qualitative description of the process provided by QCD is the following. The hard parton must be accompanied by another object balancing its color charge: this is a consequence of the confinement and it holds true for both DIS and e^+e^- as well as any other process. Existence of the counterpart causes the parton to be decelerated and, in analogy to QED, the radiation to be emitted in form of color field quanta, the gluons. This process intensifies as the quark loses its energy and separates from its counterpart due to the increase of the coupling constant with the increasing distance as well as due to the presence of the produced avalanche of gluons themselves taking part in the color interaction. Even though at some point the perturbative description breaks down, it is assumed that the parton multiplication process continues until the relative motion of the partons in the cascade becomes small enough for them to form colorless hadrons through appropriate recombination. An important feature of the cascade evolution should be noted: as a consequence of quantum-mechanical coherence a constraint is imposed on the multiple gluon radiation off the primary quark in such a way that, effectively, consecutive emissions are suppressed unless ordered in the polar angles to yet narrower cones; this is happening in an analogy to coherence observed in QED bremsstrahlung: large angle emissions cannot resolve individual partons in the cascade and thus are governed by the total color charge which is equivalent, due to the charge conservation in the

process of splitting, to radiation off the initial parton as claimed above.

Analytic Description of Parton Cascade

An analytic description of the parton cascade formation is done in QCD in terms of generating functional formalism [21]. In this method, the generating functional is constructed for parton distributions inside a cascade of given energy and contained in a cone of a given width. It adopts a probabilistic approach where, in the simplest Double Logarithmic Approximation (DLA), an expression for the rate of gluon emission off a quark can be written:

$$d\omega^{q\to q+g} \propto \alpha_S(E\theta)\frac{dE}{E}\frac{d\theta}{\theta}, \tag{2.9}$$

where E and θ are the energy and the polar angle of the gluon. The above formula exhibits the soft, $E \to 0$, and collinear, $\theta \to 0$, singularities of the underlying QCD matrix element.[8] The coherence effects are put in by hand by forcing a strong ordering in polar angle, $\theta_1 \gg \theta_2 \gg \theta_3 \ldots$. The recoil effects, however, are neglected in the integral Master Equation governing the evolution of the generating functional with the energy and cone width. This simplification causes the resulting solutions to the Master Equation to be valid in the asymptotic energy limit.

In a refined approach, the Modified Leading Logarithmic Approximation (MLLA), both the coherence and the recoil effects are taken care of more accurately. The angular ordering is incorporated via an azimuth-averaged exact QCD matrix element for the splitting which roughly translates into *ordinary* angular ordering: $\theta_1 > \theta_2 > \theta_3 \ldots$. At the same time, subleading corrections due to gluon conversion into quark-antiquark pairs, $g \to q\bar{q}$, is accounted for. These improvements over the DLA approach, however, are achieved at the cost of significant numerical complications. As a result, not all predictions available in DLA have been successfully upgraded to the MLLA level.

Perhaps the simplest possible method is used to translate the *parton-level* QCD calculations to the directly measurable *hadron-level*. The Local Parton-Hadron Duality (LPHD) hypothesis [22], formulated in connection with the pre-confinement properties of QCD, states that on the average there is a direct proportionality between hadronic and partonic distributions. The pre-confinement in

[8] Each singularity is of logarithmic type ($\int dx/x$), hence the name of the approximation.

turn ensures local compensation of color charge in parton cascades, a prerequisite for a smooth *local* recombination of partons into hadrons. The key asumption of LPHD is that the perturbative evolution of the cascade can be extended down to very low energy scales, comparable to typical hadron masses, just about where the strong coupling constant diverges. Thus, perturbative QCD aided with assumed LPHD validity is capable of making quantitative predictions concerning hadronic distributions in any system produced from an energetic parent parton like DIS or e^+e^- annihilation.

For a recent extensive discussion of the application of perturbative QCD calculations to multiparton production and respective experimental data, the review [23] can be referred to.

Single Particle-Inclusive Distributions

The key prediction of the MLLA calculation is the nearly Gaussian shape, often referred to as "humpbacked plateau," of the logarithmic scaled momentum, $\ln(1/x_P) = \ln(p_{max}/p_h) = \ln(Q/2p_h)$, distribution. It is a combined coherence and recombination effect that, despite the rapid growth of the strong coupling constant governing the soft gluon radiation, softest particles are not the most abundant in the final state: such gluons need to be radiated at large angles in order to have a chance to give rise to a separate hadron but coherence suppresses that kind of emission. At the quantitative level the evolution of the distribution around its maximum has been calculated and the numerical results are presented in chapter 9 in connection with the measurement reported in this thesis.

A variant of the prediction for the hard part of the distribution exists up to about the peak region: the parton cascade is evolved down to the very lowest, analytically allowed energy scales to yield the so-called limiting spectrum. Its shape depends on a single parameter, an effective energy scale Λ_{eff}, expected by the theory and measured to be of the order 0.2–0.3 GeV. This prediction, though not applicable to very soft particles still exhibits a signature of coherence effects: a peak is seen in the calculated distribution in the not too soft particle region of the spectrum.

An alternative approach to the hard end of the particle spectrum is the fixed-order QCD calculation of the scaled momentum distribution. This formalism, in analogy to the structure functions, defines phenomenological probabilities, the fragmentation functions, for finding a hadron of a given

momentum among daughter particles of a given hard parton. These functions can be constrained using a set of relevant data and then used in conjunction with appropriate hard scattering matrix elements and parton distribution functions to predict alternative experimental results. In particular, with a belief in the fragmentation universality, e^+e^- annihilation data can be confronted in such a way with DIS, as will be done in chapter 9.

Angular Correlations

Due to the more complicated analytic structure of multiparticle distributions predictions are only available in the DLA framework. It is sufficient, however, to find the asymptotic behavior of distributions in the relative angle between two hadrons observed in the final state. In particular, a scaling variable which is more natural than the directly measurable relative angle is advocated: the distributions should become independent of the parent-parton energy if plotted as a function of this variable. Not suprisingly, the correlations are predicted to be most pronounced for nearby particles; universality between DIS and e^+e^- annihilation is expected to hold, in a first approximation. Being related to multiparticle distributions the correlations may constitute a test of the possibility to generalize the LPHD beyond single particle spectra. Relevant comparisons and discussion of the results of the current analysis can be found in chapter 9.

Chapter 3

The ZEUS Experiment at HERA

3.1 The HERA Storage Ring

HERA (*Hadron-Elektron-Ring-Anlage*) is an electron-proton collider operated by DESY (Deutsches Elektronen-Synchrotron), Germany's center for high energy physics, on its main site in Hamburg. The cost in excess of one billion DM incurred during construction in 1984-1990 was covered mainly by the German federal government; the remaining part was split between an international collaboration of institutes from a dozen countries. Since its commissioning in 1991 the accelerator has been a unique facility in the world providing colliding beams of electrons and protons at a center-of-mass energy of $\sqrt{s} \approx 300$ GeV. With a diameter of almost 2 km the machine by far exceeds the area of the laboratory so that in large part the tunnel is buried up to 25 m under the city's recreational, residential and industrial neighborhoods surrounding DESY. The circumference of the ring is 6.3 km and consists of four bending areas of radius 0.73 km and four straight sections accommodating experimental halls. A sketch of the HERA machine is given in Figure 3.1. The tunnel houses two separate vacuum systems (pipes), one per beam. They are only connected near the interaction points where the particles traverse a single beampipe to collide at a relative angle of 0 mrad every 96 ns.

A clear imbalance exists between the beam energies: 820 GeV for protons and 27.5 for electrons. It results in an interesting situation where the HERA frame of reference coincides with neither the center-of-mass system nor the rest frame of one of the colliding particles. There are different limiting

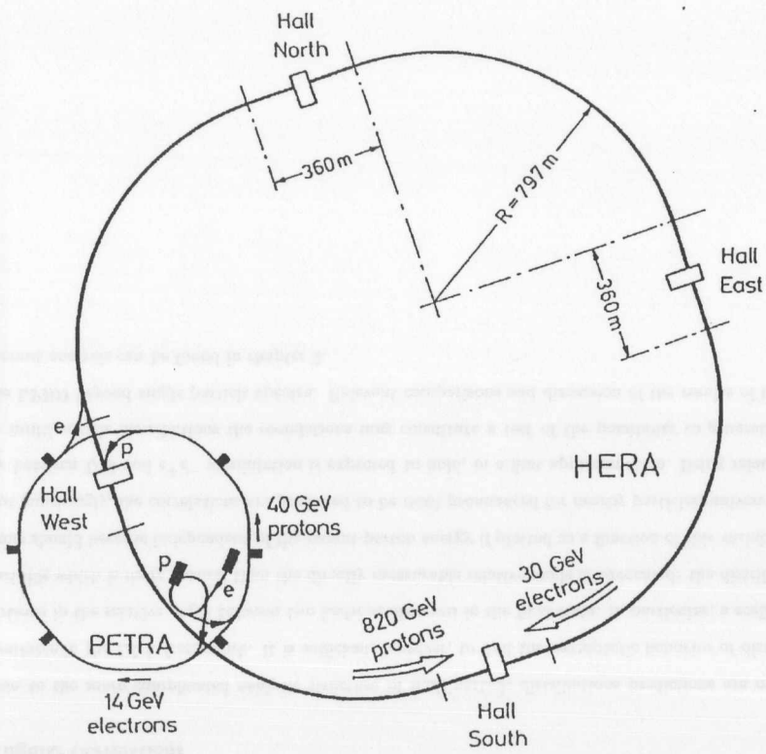

Figure 3.1: The HERA accelerator layout.

factors for both accelerators. In the lepton case it is a high energy dissipation rate due to synchrotron radiation from electrons. This has to be compensated by a highly efficient system of RF accelerating cavities. At the nominal energy an electron beam current of up to 60 mA can be sustained. For protons, due to the much higher mass, the radiative losses are negligible even at the higher momenta. However, the main dipole magnets must be able to provide an appropriate bending force to keep the particles in their orbit. Superconducting coils are used and a nominal magnetic field of 4.7 T is obtained.

The two beams follow different patterns of injection, both making use of the pre-HERA facilities at DESY. The machine is first filled up with protons. They originate from a source of negative hydrogen ions (H^-). The pre-acceleration phase is done in the 50 MeV Proton Linac. At this stage the electrons get stripped off and the protons are fed into the synchrotron DESY III to reach an energy of 7.5 GeV. The PETRA storage ring takes them over and, finally, HERA can handle them at 40 GeV. Usually the above steps are repeated a couple times in order to inject up to 210 of the 220 available bunches. In the main acceleration performed in HERA the protons reach the nominal energy of 820 GeV. The positrons[1] ejected from the LINAC I at an energy of 450 MeV are routed through the DESY II synchrotron (7.5 GeV) into PETRA. At 14 GeV they are transferred to HERA where they finally reach the nominal energy of 27.5 GeV. During a routine operation of HERA it takes typically a few hours to restore the luminosity conditions after a regular beam dump.

At present all four experimental halls are occupied. The two multipurpose detectors, H1 and ZEUS, make use of both available beams by looking at electron-proton scattering. First results on polarized cross-sections in a fixed target DIS experiment have been reported by the HERMES collaboration. The natural transverse positron polarization is converted into a longitudinal polarization by means of spin rotators. Gaseous polarized targets complement the HERMES setup. The fourth apparatus, HERA-B, is in its final stages of construction. It aims to exploit a high rate of bottom quark production in hadronic collisions at high center-of-mass energies to study CP symmetry violation in the $B^0\overline{B^0}$ system. At HERA this process can be produced by a controlled insertion of various targets into the

[1] HERA can operate with either electrons or positrons. The experience of the early running in 1992-94 showed significantly shorter electron beam average lifetime compared to positrons so the latter were used later on. Running with electrons was resumed in 1998 when the proton energy was raised to 920 GeV.

proton beam.

Over the 1995-1997 running periods HERA made a big progress concerning the efficiency of operation. It resulted in a threefold increase in the delivered luminosity per year (from 12 pb^{-1} to 36 pb^{-1}). The improvements included the increase of the average initial beam currents (I_P = 77 mA, I_e = 36 mA) and the average duration of a run (7.5 h) as well as the reduction of an average injection time. An average peak luminosity of $8.4 \cdot 10^{30}$ cm^{-2} s^{-1} was reached. A new schedule, introduced in 1997, of shortening every second winter maintenance shutdown made that season particularly successful. For ZEUS the useful luminosity was more than doubled compared to all the previous years combined. The present analysis is based on the data collected in 1995-1997 corresponding to an integrated luminosity of approximately 43 pb^{-1}.

3.2 The ZEUS Experiment Environment

Zeus, according to ancient Greeks, should be quite a good match for Hera. In the case of the ZEUS detector and the HERA accelerator the chances are even higher than in mythology as the former was designed and built with the features of the latter in mind.

The ZEUS collaboration, now consisting of over 410 physicists from 53 institutes and twelve countries, took an effort of designing, building and operating a multipurpose, almost hermetic detector for studies of lepton-proton scattering. The main components were finished on schedule in March 1992 and first *ep* events were registered in May. Since then the core setup has been stable although modifications and extensions take place during maintenance periods.

The ZEUS coordinate system is anchored at the nominal interaction point (IP). The Z-axis coincides with the outgoing proton direction and the Y-axis points vertically up. The right-handed orthogonal system is complemented by the X-axis towards the center of the HERA ring. For the angular quantities, the polar angle, θ, is measured with respect to the positive Z-direction whereas the azimuthal angle, ϕ, is defined to be 0 along the X-axis. The pseudorapidity is related to the polar angle, $\eta = -\ln \tan \theta/2$.

The layout of the main components is shown in Figure 3.2, in a longitudinal cut. A general overview of the detector given here is followed by a more detailed description of the components used for the analysis reported in this thesis. A complete report on the ZEUS detector can be found in [24].

The region immediately surrounding the beampipe is instrumented with the tracking devices meant for charged particle detection. The set includes going outwards: Vertex Detector (VXD), Central Tracking Detector (CTD) supplemented by forward (FTD) and rear (RTD) endcaps to improve the acceptance in directions less inclined with respect to the beam. A superconducting coil surrounding the CTD produces a magnetic field of 1.43 T to allow for a measurement of particle momentum from the track curvature. The tracking chambers are almost hermetically encapsulated, with the exception of the beampipe direction, inside a highly segmented Uranium Calorimeter (UCAL). The UCAL consists of three main parts: forward (FCAL), rear (RCAL) and barrel (BCAL) around the CTD. The calorimeter is in turn surrounded by the respective sections of the inner muon identification chambers (FMUI, RMUI, BMUI) and the iron yoke whose function is to return the magnetic field flux of the solenoid. It is also instrumented with the Backing Calorimeter (BAC) detector which provides the measurement of residual energy leakages from the UCAL for the most energetic jets. The outer muon system, consisting of the FMUON, RMUON and BMUON chambers, encloses the yoke. The VETOWALL, a set of two scintillator planes normal to the beam direction separated by an iron absorber, situated a few meters behind the rear part of the detector serves as a powerful rejection of the proton beam-gas background.

A number of smaller components, not shown in the overview figure, are located along the beampipe. In the rear (electron) direction, in increasing distance from the interaction point, they are: the Beam Pipe Calorimeter (BPC) for detection of electrons scattered at low angles to be found at $z = -3$ m, immediately following it the C5 lead/scintillator counter for beam monitoring and two LUMI calorimeters (photon at -104 m and electron at -34 m) used in the luminosity measurement as well as for tagging DIS radiative events. In the forward direction the first to be found is the Proton Remnant Tagger (PRT) at +5 m, a lead/scintillator counter tagging proton dissociation products. The Leading Proton Spectrometer (LPS) includes 9 pots grouped in 6 stations located between +24 and +90 m. Finally, at +102 m, the Forward Neutron Calorimeter (FNC) is situated exactly on the

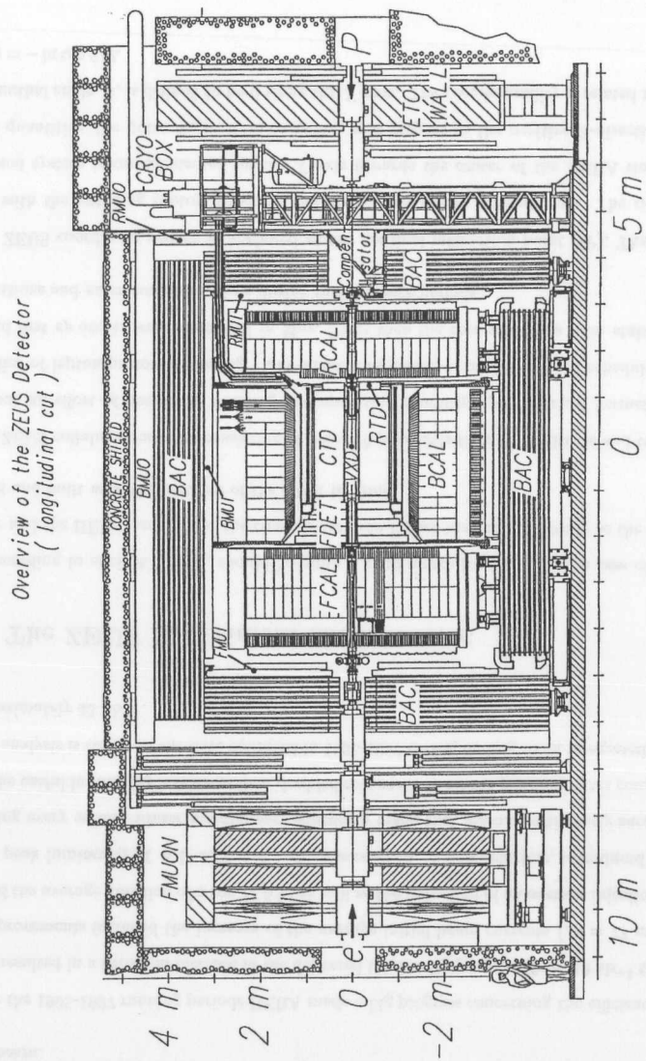

Figure 3.2: Overview of the ZEUS detector

line of neutral particles produced at zero degrees at the IP. The LPS and FNC are used to detect the leading baryon, if present, in the case of neutral and charge exchange reactions, respectively.

A brain of the detector, the trigger, has a three level structure, providing a consecutive reduction from the nominal frequency of 10 MHz to the acceptable rate of a few Hz at which the data are sent to the off-line environment.

3.2.1 The Uranium Calorimeter

In the design of the ZEUS calorimeter [25, 26, 27] the priorities were defined to be: hermeticity, high jet energy resolution, the possibility of an absolute energy scale calibration to within 1%, good resolving power between multiple jets and efficient electron identification. Achieving these criteria was extremely important to reach the physics goal of a reliable reconstruction of both the event kinematics and the structure of the final state, the former being crucial for this analysis as well.

An approximate hermeticity is obtained by covering 99.7% of the full solid angle with the exception of the forward (20×20 cm^2) and the rear (20×8 cm^2) holes to accommodate the beam pipe. Three parts overlapping in the polar angle: FCAL (2.2°–39.9°), BCAL (36.7°–129.1°) and RCAL (128.1°–176.5°), are about 7, 5 and 4 absorption lengths deep, respectively. This is enough to ensure that 90% of all observed jets are contained to better than 95%. The selection of the absorbing material, depleted uranium, allowed both full compensation of electromagnetic and hadronic energy deposits ($e/h = 1.0 \pm 0.05$) and an accurate (down to 1% level) in-situ calibration with the help of the low intensity natural radioactivity of U^{238} nuclei. The achieved energy resolution, measured under test beam conditions, was $\sigma(E_h)/E_h = 0.35/\sqrt{E(\mathrm{GeV})} \oplus 0.02$ and $\sigma(E_e)/E_e = 0.18/\sqrt{E(\mathrm{GeV})} \oplus 0.01$ for hadronic and electromagnetic showers, respectively. At the same time, the use of fast scintillators as the active material, made time measurements very accurate: $\sigma(t) < 1$ ns for energy deposits above 4.5 GeV, and helped to avoid pile-up effects between consecutive bunch crossings.

The longitudinal segmentation of the UCAL into inner electromagnetic (EMC) and two (one) outer hadronic (HAC) sections in FCAL and BCAL (RCAL) together with the transverse subdivision into

cells of approximate dimensions 20×20 cm^2 (HAC) and 20×5 cm^2 (EMC)[2] provides a reasonable spatial resolving power. Light from each cell is collected on both sides separately and transferred via wavelength shifters and light guides to photomultipliers. This doubling of the readout channels provides a further improvement in position resolution in the ϕ direction by considering an imbalance between the two signals. It has also a positive impact on the reliability: on the average there is less than one inactive cell out of the total 5918.

For each bunch crossing a fraction of the analog signal from the photomultipliers is sent to the Calorimeter First Level Trigger for inspection of the deposit patterns characteristic of various physical processes. The decisions are forwarded to the global trigger within 1 μs. The Fast Clear processor implementing a more detailed pattern recognition algorithm for selected class of events, is being incorporated into the ZEUS trigger but was not in operation during collection of the data reported here.

3.2.2 The Central Tracking Detector

The main tracking device at ZEUS, required by the trigger logic to be operational during the data taking, is the Central Tracking Detector [28, 29, 30]. It is a cylindrical chamber, immersed in a magnetic field of 1.43 T to allow a transverse momentum measurement, with a total of 4608 sense wires in 72 layers grouped in 9 superlayers. In the odd superlayers the wires are axial (parallel to the beam) whereas the even are tilted by approximately 5° to get a handle on the position in the Z-direction of the particle traversing the detector. For the same purpose but at a speed demanded by the trigger the three innermost layers are instrumented with a Z-by-timing system sensing the difference of signal arrival time at both ends of a wire. The active volume of a length of 205 cm, inner radius 18.2 cm and outer radius of 79.4 cm, with the nominal interaction point near its geometric center, provides an acceptance for charged particles emerging at polar angles within a range of $15° < \theta < 164°$. The chamber is flushed with a mixture of argon, carbon dioxide and ethane.

For the raw output of the chamber, a single hit efficiency better than 95% is reached. The pattern

[2] The cells are 10 cm high in the RCAL EMC section.

recognition is done offline[3] by a dedicated software package, VCTRAK. The output consists of tracks reconstructed with a resolution in the transverse momentum of $\sigma(p_t)/p_t = 0.0058p_t \oplus 0.0065 \oplus 0.0014/p_t$ (p_t in GeV), and the primary vertex position[4] determined with an accuracy $\sigma(z_{vtx}) \approx 1.5$ mm. The latter is calculated from a fit to a subset of all tracks. Particle identification, judged from the rate of energy loss along the track, dE/dx, is sensitive to the atmospheric pressure and therefore poses additional difficulties. In the present analysis no identification has been attempted and all the tracks were assumed to be due to charged pions.

The Vertex Detector (VXD), filling the gap between the CTD and the beam pipe, progressively died out in 1995 due to radiation damage and was removed afterwards. It provided, when active, additional constraints on the inner sections of the reconstructed tracks thus improving the resolving power for the vertex fitting algorithm.

3.2.3 The Forward and Rear Tracking Detectors

The track acceptance at low angles with respect to the beam is extended by a set of three planar drift chambers in the forward direction (FTD) [31] and a single chamber of the same design in the rear (RTD). Each chamber consists of three planes perpendicular to the beam, with sense wires rotated by 120° with respect to each other thus providing unambiguous reconstruction of a track segment in three dimensions. In the forward direction, where a majority of the hadronic activity is concentrated in a typical HERA ep event, the polar angle coverage reaches from 28° down to 7.5°, and is almost uniform for the three segments thanks to their radius increasing with the distance from the interaction point. The FTD, together with two pairs of transition radiation detector (TRD) sandwiched in between the chambers, forms a rigid pre-aligned structure of the forward detector (FDET). Separating the chambers expands the distance over which a particle is tracked and improves the accuracy of reconstruction. In the rear direction the RTD extends from 160° to 170°, *i. e.* matches the RCAL face around the beam pipe hole. Thanks to a single wire resolution around 130 μm the RTD can constrain the kinematics (Q^2) tighter than the RCAL energy measurement, if a

[3] A reduced version is executed online in the Third Level Trigger.

[4] At the end of 1997 reconstruction of secondary vertices was introduced. This feature is not used in the present analysis.

low angle electron track can be reconstructed. At both ends there is an overlap of the sensitive area between the CTD and the endcaps. However, optimization, calibration and operation of the chambers situated in a demanding region of a strong and non-uniform magnetic field (both FD and RTD) and densely populated with tracks (FTD only) are not trivial. There is the additional problem of the precise alignment with respect to the CTD. As a result the tracking package VCTRAK permits the reconstruction to be based on the CTD+VXD data only. The present analysis, nevertheless elected the inclusive mode of reconstruction with a cross check using the results obtained in the restricted mode.

3.2.4 Peripheral Subdetectors

A brief description of the components not discussed in the general overview of the detector yet important for the present analysis is given below for completeness.

The Small-angle Rear Tracking Detector

Completely shadowed by the RTD but still useful for the DIS electron detection at low Q^2 is the Small-angle Rear Tracking Detector (SRTD) [32] located on the front face of the RCAL. Its two planes of crossed (at right angles) scintillator strips of 1 cm width cover the 8 innermost RCAL towers (16 EMC cells) immediately adjacent to the beam pipe. A total of 272 strips is divided into 4 sectors. The most important contribution of the SRTD to the process of data taking is its excellent timing resolution, better than 1 ns, allowing for an efficient rejection of proton beam-gas background events taking place behind the RCAL and this can be done already at the first level of the trigger. The triggering based on an electron hit in a proper time window has been considered as well. For the off-line analysis, matching of the SRTD energy deposits with the RCAL electron becomes more important in the context of correcting for the preshowering effect due to inactive material in front of the calorimeter. Except in the case of an electron hitting two neighboring RCAL cells (when the energy imbalance permits very accurate position reconstruction) the SRTD is superior. Both correction methods mentioned above were exploited in the present analysis.

The Presampler

A dedicated tool for pinning down preshowering energy losses in inactive material in front of the calorimeter is the ZEUS presampler [33]. The faces of the forward and rear parts of the CAL are equipped[5] with an array of scintillator tiles each covering an area equivalent to a single HAC section (20×20 cm^2). Light left behind by charged particles is guided to photomultipliers by wavelengthshifters embedded in the scintillating material. Combination of the presampler signal, counting the multiplicity of minimum ionizing particles in the shower, with the energy measurement from the CAL allows a correction to be performed on an event-by-event basis. The coarse, compared to the SRTD, spatial resolution makes this correction slightly inferior so that it is only used for electrons outside of the SRTD acceptance.

The Luminosity Detector

The important issue of the luminosity monitoring at ZEUS is addressed by means of two calorimeters (LUMI) [34] dedicated to a continuous energy measurement of the electron and photon emerging from the bremsstrahlung (Bethe-Heitler) process, $ep \to e\gamma p$. The respective locations of the detectors are on the neutral line (γCAL) and behind the first station of bending dipoles used as a spectrometer for off-beam energy electrons (eCAL). A coincidence of signals from both devices is used as the signature. A background arises due to bremsstrahlung off residual beam gas present in the beam pipe. It is corrected for by subtracting the appropriately weighted rate observed for unpaired e bunches effectively measuring the quality of the vacuum in the interaction region. Contributions from pile-up of energetic synchrotron radiation and proton beam halo induced coincidences are removed by means of shielding and examination of unpaired p bunches, respectively.

Thanks to a lot of work invested in the understanding of the LUMI detector, the luminosity measurement at ZEUS is accurate to $\approx$ 1.4%. This result should be appreciated even though the present analysis deals only with normalized cross-sections and therefore is independent of the absolute luminosity calibration.

[5] The Barrel Presampler has been installed in 1998.

3.2.5 The Trigger

An extremely high rate of bunch crossings at HERA, 10 MHz, imposes tight requirements on the trigger. From the actual background rate varying (in broad range) up to 100 kHz, the *ep* physics events taking place a few times per second must be selected in real time and the corresponding data written to a permanent storage. At ZEUS, this is achieved in a three level system [35].

The first level trigger (FLT), operating at the nominal frequency of the HERA clock, is supposed to establish the output rate below 1000 Hz. It consists of the distributed (component) and the global (GFLT) part. Each subdetector contributing to the decisions made at this level extracts the general features of an event and makes them available to the GFLT no later than 2.5μs (26 bunch crossings) after the event took place. The programmable logic chip based GFLT has an additional 1.9 μs for processing after which the accept signal may be broadcast to the components accompanied by an identification of the event being triggered on. At this point a dead time is introduced to allow the subdetectors to forward the data to the Global Second Level Trigger (GSLT)[6]. The GSLT is a set of fast transputers executing a filter code with fully digitized, though still not complete, data as the input. The processing is done in an asynchronous mode, *i. e.* there is no time limit for the trigger decision to be made. A typical latency of the GSLT, however, is of the order of a few milliseconds and the accept rate reduced well below 100 Hz. Upon a positive decision the full event record is constructed by the Eventbuilder and passed to the Third Level Trigger (TLT), a farm of Silicon Graphics CPUs running a reduced version of the off-line reconstruction code to improve performance. The analysis at this stage is already physics-oriented and yields an output of a few accepted events per second. These data are then written to tapes managed by a robot in the DESY Computer Center.

The backbone of the detector, the central run control, communicates with the components and the trigger chain in a dedicated protocol broadcast over ethernet. In addition to taking regular data, dedicated calibration or test conditions may be administered through appropriate configuration files.

[6] There are a few component SLT processors as well.

Chapter 4

Reconstruction

The way from the raw output of the detector components to the physical quantities is long and time consuming. As an obvious prerequisite it requires a good understanding of the detector response to processes taking place inside its sensitive volume. This topic is discussed in more detail in chapter 5 in the context of the most common technique of detector simulation. There are, however, other issues involved in the reconstruction procedure, like the precise alignment of individual components, time dependent calibration as well as bookkeeping of hardware failures and configuration of the trigger. In addition, methods used for some aspects of reconstruction, for instance electron identification or charged track pattern matching, were considered both interesting and important enough to be presented here. They summarize a large amount of effort of people engaged in their development and refining.

This chapter, after an introduction to general ZEUS reconstruction issues, discusses certain details specific to the present analysis including justification of some data selection cuts on the grounds of known reconstruction defficiencies. Descriptions of the DIS electron identification, track reconstruction, methods of kinematics determination and, finally, the Breit frame reconstruction are given.

4.1 Reconstruction at ZEUS

The central point of the process of reconstruction is the ZEUS PHYsics Reconstruction (ZEPHYR) program. Starting from raw digitizations it produces output in terms of calibrated physical quantities and their relationships. Auxiliary information concerning current experimental conditions and detector status is taken into account as appropriate. The relative alignment of components as obtained from either *in-situ* surveying or from analysis of the raw data is used as well.

The reconstruction code consists of packages written and maintained by experts responsible for the individual subdetectors and merged on a high level by the centrally developed control routine. On the fly many distributions are histogrammed for data quality verification. Since 1998, a modern database containing a subset of the most commonly used data selection variables is also updated by ZEPHYR. Full output for a fraction of the reconstructed events is made available for immediate access on disk media. Typically the reconstruction process takes place a few days after the data have been collected and is limited by the availability of calibration data rather than computer performance.

The computing facility for the reconstruction used to be a dedicated multiprocessor SGI CPU but a smooth transition to a farm of PCs took place during 1997. Advantages of the new solution are lower cost, easier scalability and availability for general ZEUS analysis during between seasons of data taking. The storage system based on the magnetic tapes and robots to handle them is provided by the DESY Computing Center[1].

The execution is split into three phases in which gradually more general features of the data are uncovered; intermediate results from each phase are appended to the output. In the first phase the raw data from each component are transformed according to the calibration specific for the given subdetector. This auxiliary information is available either directly in the dataflow as provided by the ZEUS Slow Control system during the data collection or it is generated off-line in semi-automated processes. Typical information available after this stage consists of calorimetric energy deposits, tracks in chambers etc. At the second level an attempt is made to relate output from the different parts of the detector. This includes global matching of track segments found in different chambers, matching

[1] The volume of the data from a single year is of order of a few Terabytes.

energy leaks into BAC with UCAL clusters, and fitting the primary vertex. Finally, in the last stage of reconstruction particle identification is performed for instance by combining measurements from trackers, calorimetry and muon chambers.

The simulated data as generated by Monte Carlo programs are piped through the same chain of reconstruction after calculating the raw detector and trigger response in dedicated programs described in more detail in chapter 5.

Due to a growing number of subcomponents, especially those with a relatively long and complicated process of off-line calibration, a significant amount of reconstruction is left over for final analyses. This, together with a rapid increase of luminosity, encouraged efforts like event databases fed with prereconstructed observables used in data selection or even ntuples covering wide enough samples and containing nearly complete information to be usable in groups involved in related analyses.

4.2 DIS Electron

Neutral current DIS events at HERA bear a stamp of the scattered electron in the final state. Finding and measuring the electron is not only one of the means of determining the event kinematics but also serves as the primary method of discrimination against a huge photoproduction background. The design of the ZEUS calorimeter is aimed, among others, towards reliable identification of the electron. The separation of the relatively shallow and finely segmented EMC part of the calorimeter from the deeper hadronic part is used in the task of on-line DIS event selection right from the first level trigger where an electromagnetic shower is demanded. This is pushed further at the third level analysis by requiring, not neccessarily at a high probability, an identified electron. Given this DIS-enriched sample the final procedure is pursued as described below.

4.2.1 Electron Identification Algorithms

The primary information for electron identification is its signature in the calorimeter. It consists of a small number of cells, typically below 6, scattered over an array of a few neighboring towers of dimensions approximately 20×20 cm^2. Usually most of the energy is deposited in the EMC section

so that the leak to the HAC is limited. This pattern is to be confronted with and distinguished from a typical hadronic shower characterized in general by a wider spread in all three dimensions making the average total number of active cells higher. There are, however, neutral pions whose decays into collimated photon pairs mimic a positron. The distinction is especially difficult at small energies. In addition, at low x the directions of the electron and hadronic flow become poorly separated resulting in a significant probability for an overlap of the relevant signals in the RCAL. For these reasons electron identification is by no means a trivial task and quite a few different approaches to the problem have been developed.

Due to the fine segmentation of the calorimeter the electron shower, even though compact, is described by too large a number of parameters to be analyzed in a systematic and exclusive way. Therefore some projections of the information onto a set of a smaller number of better suited variables are needed. This can be achieved either by means of an *a priori* construction of such a subset of well understood quantities or by a novel method of neutral network analysis.

In wide use at ZEUS is a handful of algorithms based on the classic approach. They differ in the number and definition of explicitly analyzed variables and they can be traced with respect to strategies and calculations involved. On the other hand Sinistra [36], the ZEUS neural network electron finder, has its algorithm distributed over a set of parameters known as weights. This feature makes it impossible to "understand" its strategy. Even though not traceable, Sinistra's decision communicated in a form of a single probability can be checked against other algorithms or unbiased samples (cf. elastic Compton scattering events). The optimal values of the weights are established in the process of network training on a mixture of hadronic and electromagnetic showers. The process can be monitored in order to ensure quality of the convergence towards the final weights and to avoid overtraining. Because of its superior performance over the classic methods, Sinistra has been used in many analyses including the one reported in this thesis. Some additional details on the algorithm follow.

4.2.2 Sinistra: a Neural Network Electron Finder

Before a shower candidate is presented to the neural network some preselection cuts are imposed. As a prerequisite a clustering algorithm is used to identify seed calorimeter towers defined as local maxima of energy and to associate lower energy towers to the seeds. A seed tower with its 8 neighbors forms a window and it is required that not less than 90% of the total cluster energy is contained in the window. In addition, an absolute limit on the leakage outside the window is set at 1 GeV. It is further demanded that at least 80% of the energy is seen in the EMC section. The above preselection cuts, based on the known features of electromagnetic showers, significantly reduce hadronic background; the respective loss of efficiency is of the order of 1%.

The complete calorimeter data on the deposits within the window, 54 variables[2], complemented by a single geometric variable, the angle of incidence, are then fed into the neural network. As an output a single number, P, is given representing a measure of the probability of the shower to be of electromagnetic origin. From a list of all showers in an event the candidate maximizing the probability is chosen.

For the default cut on the output variable, $P > 0.9$, the efficiency of electron identification is typically above 98% for showers of energy exceeding 10 GeV maintaining at the same time high purity. Geometric regions of known defficiency of order of 1% exist for showers split into two parts of the calorimeter. There is a good agreement between the real and simulated data in this respect.

The core algorithm based solely on the calorimeter data has been supplemented with an electron candidate matched to a reconstructed track. This feature has been used in the present analysis in order to remove the electron track from the pool of hadron candidates. For a perfect reconstruction of the transformation to the Breit frame it would not be necessary as, due to kinematics, the scattered electron is bound to the target hemisphere but this may be readily spoiled by finite experimental resolution.

[2]There are two EMC and one HAC cell per RCAL tower each of them read out by two photomultipliers totalling 6 signals per tower. The electromagnetic part of showers in BCAL and FCAL is reduced to this format by summing pairs of EMC signals from neighboring cells.

4.2.3 Position Correction

The position of the electron candidate is calculated as an energy weighted average of the cells assigned to the accepted shower. An average resolution of approximately 1 cm improves by a factor of 10 in the case of showers split nearly evenly over two neighboring cells where the energy imbalance can be used to obtain a more accurate measurement. However, the SRTD scintillator strips provide a uniform resolution of 0.3 cm in the region immediately surrounding the RCAL beampipe hole [32]. For this reason, for an electron candidate traversing the sensitive area of the SRTD a correction was applied based on hits reconstructed in this detector.

The hits were searched for within a radius of 15 cm around the impact point predicted from the calorimeter position of the candidate. The closest match, if found with an energy deposit above 5 MIP, was accepted. In the absence of a high energy match the restriction was dropped and low energy hits were considered in the second pass. The algorithm is described in more detail in [37].

Finally a small correction of order of 1 to 3 mm was applied based on the alignment study done with the help of tracks extending from the CTD to RTD and traversing the SRTD. The correction evaluated for the detector setup as in 1995 was used. The data samples from 1996 and 1997 were restricted to a kinematic region where the electron typically is deflected at too large an angle to hit the SRTD.

4.2.4 Energy Corrections

The raw energy of an electron candidate measured in the calorimeter does not take into account losses in inactive material. Due to early showering, a significant fraction of the total energy will not enter the sensitive volume of the calorimeter. This effect is corrected for by inspection of the signals from the presampling detectors placed in front of the RCAL. The idea behind this is that the presampled energy is proportional to the charged particle multiplicity in the shower and therefore provides a handle on the energy loss in the inactive material. In general use is a linear formula

$$E_{corr} = E_{CAL} + aE_{shower} \tag{4.1}$$

with a free parameter a to be determined from a dedicated study. A typical value of a is of the order of 50 to 70 MeV/MIP. In this way both the systematic pulls and the resolutions are improved.

A vast majority of DIS electrons at low Q^2 can be matched to hits in the SRTD as described in the previous subsection. In this case the energy E_{SRTD} of the hit assigned to the calorimeter shower is used in the correction formula. In [38] true data and Monte Carlo are studied in three independent samples with an electron candidate constrained by other features of the event. These samples included elastic QED Compton events, elastic ρ^0 events in DIS and the so-called kinematic peak events where the electron energy is constrained by kinematics independent of the actual process taking place. The analysis led to consistent values of the parameter a in Eq. 4.1. They are found in the expected range but are different for data and MC, 56 and 79 MeV/MIP, respectively. It is also noted that in the latter case a relatively small quadratic term is needed to obtain a consistent correction applicable over a wide range of energies.

For electrons missing the SRTD fiducial region but still detected in the RCAL the energy correction is established with the help of the ZEUS presampler. Only parameters from a preliminary calibration based on kinematic peak events as well as test beam data were available at the time the present analysis was initiated. A simple linear correction acording to Eq. 4.1 was fit. Again differences were found between the coefficients for the true data and Monte Carlo, 72 and 51 MeV/MIP, respectively.

Recent progress in this subject has been summarized in [39]. Due to the larger geometric extent of the presampler compared to the SRTD a more sophisticated method is argued for. It takes into account the varying amount of the inactive material dependent on the impact point of the electron and results in a more accurate correction. The simplistic linear correction has been re-evaluated as well leading to a reduction of data-MC difference and general agreement with the parameters quoted above.

Up to 1997 there was no widely ackowledged method of refining the positron energy measurement in the central part of the calorimeter (BCAL). The FCAL in turn, is instrumented with a presampler, but it can be struck by the DIS electron only in the upper tail of the kinematic region considered in this thesis ($Q^2 \gtrsim 1500$ GeV^2). High Q^2 typically correlates with large electron energies so that within

the selection cuts used in this analysis there is barely a chance to find an electron of less than 10 GeV in the FCAL region. For completeness, however, a linear correction with the same parameter values as in the rear presampler was applied in such a case.

The directly measured electron energy, corrected or not, was only used in event selection as described in chapter 6. As described in section 4.5, the Breit frame was reconstructed using the energy recalculated from the Double Angle method to avoid systematic uncertainties of the absolute energy scale.

4.3 Charged Particles

4.3.1 The VCTRAK package

The process of charged particle reconstruction and fitting the interaction vertex position is performed in the first phase of the global ZEUS reconstruction. It is done by a software package VCTRAK also available in a user-customized form in the ZEUS analysis environment. A slightly reduced version, limited by the CPU throughput, is employed on-line in the third level trigger.

Operating on hits provided by the central trackers, CTD and VXD, and using additional 3D segments found in the endcap trackers, FTD and RTD, or 3D points reconstructed in SRTD, VCTRAK produces a list of tracks and a fit of the primary vertex. A restricted mode of reconstruction in which only the CTD and VXD data are accepted is executed as well. This is motivated by an expected reduction of uncertainties in the interalignment of the components at the cost of degraded acceptance.

The resolutions on the input hits in the $r\phi$ plane are conservatively assumed to be 0.015 and 0.03 cm in VXD and CTD, respectively. The Z-by-time instrumentation in the inner part of the CTD is accurate to about 5 cm but degrades the corresponding resolving power in the $r\phi$ direction by a factor of 3 for the involved wires.

The output tracks are parameterized as helices with a reference point in the neighborhood of the nominal beam line. The quality of a track is reported in terms of the χ^2 value of the helix fit as well as the covariance matrix between the 5 parameters. Assignment of individual hits to the recognized

tracks is given together with a summarized description of the track's topology. Extrapolation towards the calorimeter is attempted so as to allow matching to the energy deposits measured there.

The resolution in the transverse momentum for long tracks reaching the outermost superlayer is quoted to be $\sigma(p_t)/p_t = 0.0058p_t \oplus 0.0065 \oplus 0.0014/p_t$ for p_t in GeV and is modelled well in simulated data. The constant contribution is due to the rescattering inside the detector whereas the linear part reflects limitations of the chamber. The efficiency for track reconstruction is close to 98% for isolated particles whereas in densely populated environments a conservative estimate is 95%.

Well-behaved long tracks are eligible for a primary vertex fit. After the process of vertex finding is complete the participating tracks are refit with an additional constraint as given by the calculated vertex. The result is reported in terms of the vertex position and estimate of the fit quality. The resolution in the longitudinal position of the vertex has been estimated to be approximately 1.5 to 2 mm [40].

In addition, for simulated data the relationship between tracks and true charged hadrons is provided to support studies of the effects introduced by the reconstruction.

Some details of the algorithm are presented in the two following subsections. An extensive coverage of the subject with emphasis on numerical recipes can be found in [41].

4.3.2 Track Reconstruction Algorithm

The track finding is started in the outer part of the CTD where particles are more widely separated and pattern recognition is therefore less ambiguous. Seeds consisting of hits from a single axial superlayer are loosely bound by a "virtual" hit on the beam axis to guide the trajectory inwards. In the inclusive mode additional seeds partially taken from other tracker's data are considered as well. The seeds are then extended towards the center of the detector gathering consecutive axial hits in their trajectory. Hardware malfunction areas, as marked in calibration databases, are taken into account as inactive.

A reasonably well-constrained arc in the $r\phi$ plane is then merged with longitudinal measurements coming from axial or Z-by-timing wires. Either kind of information is required of any track candidate.

Due to relatively frequent problems with the Z-by-timing system misdirecting the pattern recognition, it is retried excluding this input in case of suspiciously poor hit efficiency. The process is continued down to the innermost superlayer of the CTD: the candidate is retained if it successfully matches hits in this region. A further attempt is made to match with the VXD, if data are available.

Tracks with too few unique hits (below 85%) are rejected. In a separate iteration shorter candidates are accepted but will not be assigned to the vertex later.

Raw knowledge of the trajectory at this stage allows one to correct the individual hits for multiple effects spoiling the planar drift approximation before the helix fit is performed. Shifts of more than a millimeter are possible but not common. The correction is particularly important for certain classes of otherwise strongly affected tracks but in general it mostly improves resolution without significant changes to values of the final helix parameters.

A 5-parameter helix is then fit to the candidate's set of corrected hits with an option of dropping or swapping some of them. At this time the helix is anchored around the most inward hit. An additional parameter, the $T0$ offset, can be introduced if it is revealed that a significant improvement of the fit quality can be achieved. Finally the helix is swum towards the perigee with the Z-axis. In preparation for a vertex fit the covariance matrix at that point is calculated.

4.3.3 Primary Vertex

The primary interaction vertex is sought in three steps. Firstly, otherwise well-measured tracks are rejected on grounds of incompatibility with the beamline in the $r\phi$ plane. Secondly, a fast linear fit provides a seed vertex position for the full fit to follow while possibly dropping additional tracks with too high a contribution to the χ^2. All tracks passing this test are definitely assigned to the vertex. Finally, the full fit is performed; minor adjustments to the helices of the remaining tracks are done in parallel. As a measure against secondary vertices a diffuse pseudo-proton along the beamline participates in this fit resulting in a possibility of constraining the vertex by a single track.

Subtle cases of vertex fitting failures are tackled by allowing each trajectory in turn to define a subset of other tracks compatible with it and passing it through the above steps in search for a vertex candidate with the highest multiplicity (and lowest χ^2 in case of a tie).

In a recent update of the VCTRAK package an upgrade of the vertexing algorithm has been introduced by implementing a search for secondary vertices. An improvement is seen in this case for the primary vertex reconstruction by means of an increased efficiency and decreased mean χ^2 of the fit. These effects are due to the weakened attraction of the tracks towards the only possible vertex as they can be claimed by secondary interaction points. This multiple vertex mode has not been applied in the regular ZEUS reconstruction to the data sample reported in this thesis.

It should be stressed here that the position of the primary vertex of an event must be established before the event kinematics can be calculated. This is because each of the methods discussed in the next section uses angular information about the DIS electron and/or the hadronic system and it is only defined with respect to the vertex.

4.4 Kinematic Variables

Knowledge of accurate values of the kinematic variables is of crucial importance for the correct classification of an event into an analysis bin. It is also a prerequisite for the reconstruction of the Lorentz transformation linking the ZEUS system to the Breit frame of reference.

Since there are only two independent kinematic variables, measurement of both the electron quantities and hadronic energy flows at the same time is overconstrained. Thus multiple methods exist using various subsets of the available information. They are introduced and discussed in the following subsections.

Each method uses calorimetric data with due suppression of the natural uranium radioactivity signal. Isolated cells in the EMC and HAC sections below 0.08 and 0.12 GeV, respectively, are rejected. This is done in parallel to cleaning up the cell pool of those problems detected only during systematic off-line analysis. It is estimated that per event no more than a few cells activated by uranium noise pass the above cuts.

4.4.1 Electron Method

The most straightforward approach is to establish the kinematics solely from the detected DIS electron, ignoring the hadronic part of the final state. The relevant formulae are in this case:

$$y_e = 1 - \frac{E'_e}{2E_e}(1 - \cos\theta'_e), \tag{4.2}$$

$$Q^2_e = 2E_eE'_e(1 + \cos\theta'_e), \tag{4.3}$$

and

$$x_e = E'_e\frac{1 + \cos\theta'_e}{2y_eE_p}, \tag{4.4}$$

where E'_e and θ'_e are the measured energy and scattering angle of the electron. The beam energies are denoted by E_e and E_p.

What was the main advantage of the method, inferring the measurement from a single object, becomes the weak point now because systematic uncertainties of the electron detection blow up in the kinematic variables. The accuracy of reconstruction may be calculated as

$$\frac{\Delta x_e}{x_e} = \sqrt{\left(\frac{1}{y}\right)^2\left(\frac{\Delta E'_e}{E'_e}\right)^2 + \left[-\tan\frac{\theta'_e}{2} + \left(\frac{1}{y_e} - 1\right)\cot\frac{\theta'_e}{2}\right]^2(\Delta\theta'_e)^2} \tag{4.5}$$

and

$$\frac{\Delta Q^2_e}{Q^2_e} = \sqrt{\left(\frac{\Delta E'_e}{E'_e}\right)^2 + \tan^2\frac{\theta'_e}{2}(\Delta\theta'_e)^2}, \tag{4.6}$$

and it evidently degrades at low y for the x variable and, in the case of an electron scattered at low angles ($\theta'_e \approx 180°$), for the Q^2 estimate. The latter region is naturally excluded by a fiducial cut on the electron impact position on the face of RCAL applied to reduce leaks of the electromagnetic shower outside the sensitive volume. In this analysis the cut was set as $\max(|x^e_{RCAL}|, |y^e_{RCAL}|) > 16$ cm.

The distributions of $(x_e - x_{true})/x_e$ and $(Q^2_e - Q^2_{true})/Q^2_e$ are plotted in Figs. 4.1 and 4.2, respectively. For the x_e variable it is seen that the resolution in terms of the RMS of the distribution remains at the 20 to 30% level for all kinematic bins with a weak tendency of the distribution to get narrower at high Q^2. A tail of events with a significant underestimation of x, however, persists in

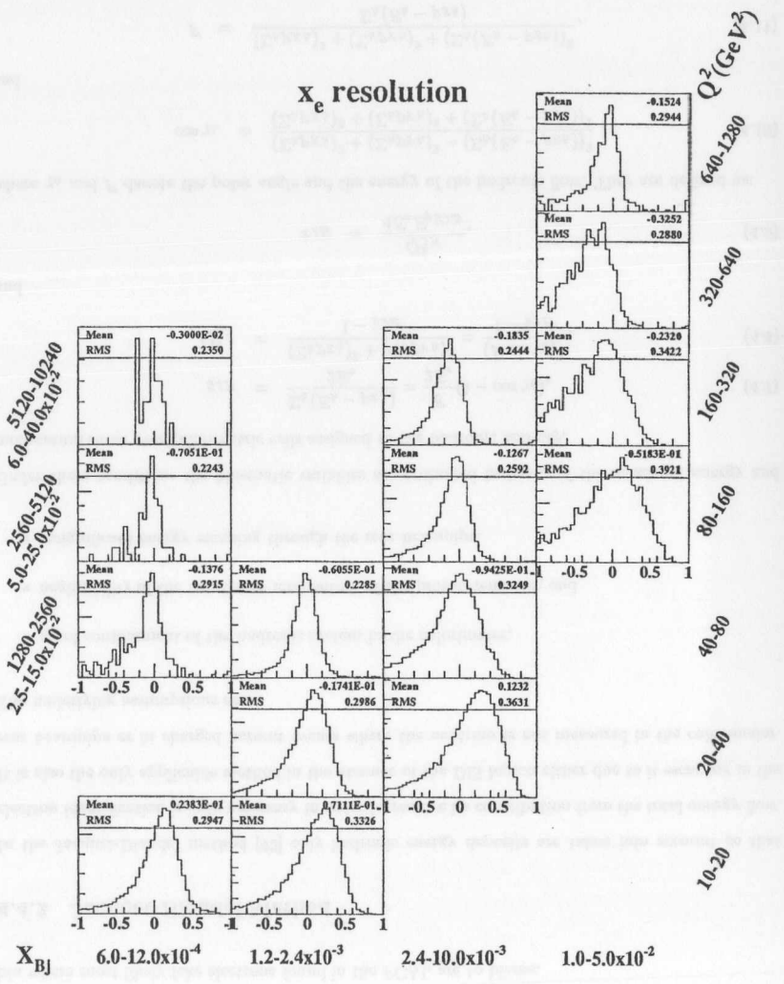

Figure 4.1: Resolution of x_e in kinematic bins.

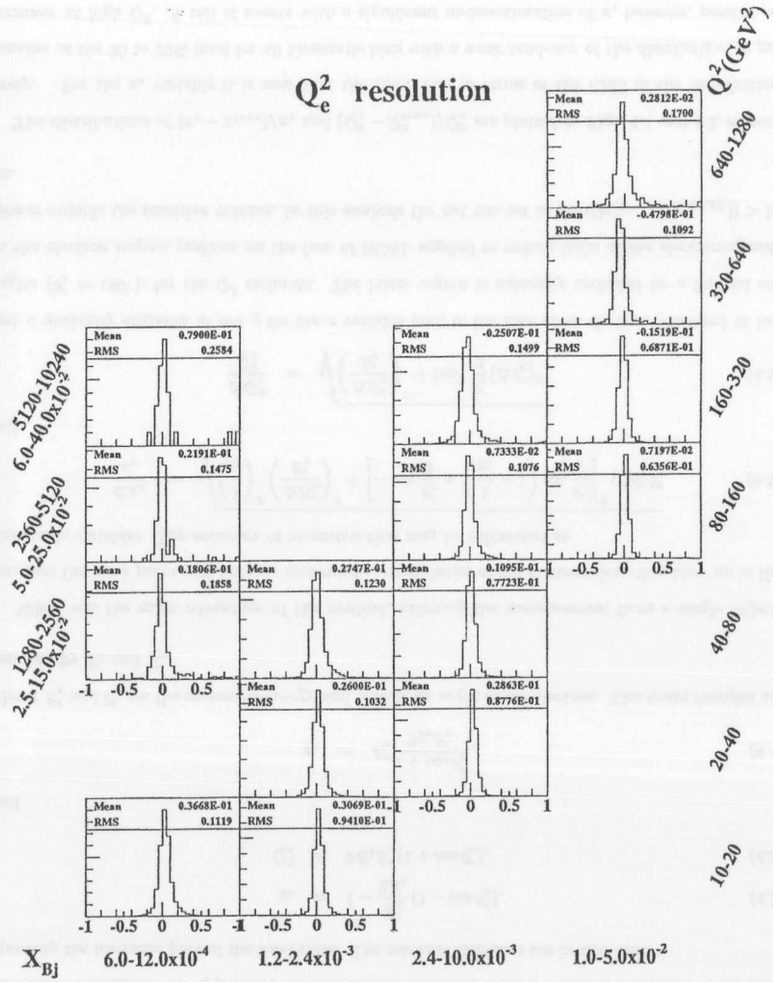

Figure 4.2: Resolution of Q_e^2 in kinematic bins.

that region as well. At low y a shift towards an overestimated x is observed. Q_e^2 in turn is resolved to about a 10% accuracy throughout the entire region studied with a deterioration in the highest Q^2 bin where most likely fake electrons found in the FCAL are to blame.

4.4.2 Jacquet-Blondel Method

In the Jacquet-Blondel method [42] only hadronic energy deposits are taken into account so that electron identification is only necessary in order to remove its contribution from the total energy flow. It is also the only applicable method in the absence of the DIS lepton either due to it escaping in the rear beampipe or in charged current events where the neutrino is not measured in the calorimeter. The underlying assumptions are:

- good containment of the hadronic system in the calorimeter,
- negligibility of the transverse momentum of the proton remnant, and
- insignificant energy escaping through the rear beampipe.

Under these conditions the kinematic variables are evaluated in terms of the measured energy and momentum sums over calorimetric cells assigned to the hadronic activity,

$$y_{JB} = \frac{\Sigma_h (E_h - p_{Zh})}{2E_e} = \frac{F}{2E_e}(1 - \cos\gamma_h), \tag{4.7}$$

$$Q^2_{JB} = \frac{(\Sigma_h p_{Zh})^2 + (\Sigma_h p_{Yh})^2}{1 - y_{JB}} = \frac{(F \sin\gamma_h)^2}{1 - y_{JB}}, \tag{4.8}$$

and

$$x_{JB} = \frac{Q^2_{JB}}{4E_e E_p y_{JB}}, \tag{4.9}$$

where γ_h and F denote the polar angle and the energy of the hadronic flow. They are defined as:

$$\cos\gamma_h = \frac{(\Sigma_h p_{Xh})^2 + (\Sigma_h p_{Yh})^2 - (\Sigma_h (E_h - p_{Zh}))^2}{(\Sigma_h p_{Xh})^2 + (\Sigma_h p_{Yh})^2 + (\Sigma_h (E_h - p_{Zh}))^2} \tag{4.10}$$

and

$$F = \frac{(\Sigma_h p_{Xh})^2 + (\Sigma_h p_{Yh})^2 + (\Sigma_h (E_h - p_{Zh}))^2}{\Sigma_h (E_h - p_{Zh})}, \tag{4.11}$$

and, in the naïve Quark Parton Model, they correspond to the scattering angle and the energy of the struck quark. Unlike F, γ_h is largely insensitive to the calorimetric energy scale uncertainty.

The size of the FCAL hole, $\approx 2.2°$, imposes a hard limit on the angle γ_h and this in turn translates into a kinematic restriction of $y \gtrsim 0.001$.

Within the region of validity of the method the propagation of errors is governed by the equations:

$$\frac{\Delta x_{JB}}{x_{JB}} = \sqrt{\left(\frac{1}{1-y_{JB}}\right)^2 \left(\frac{\Delta F}{F}\right)^2 + \left[-\tan\frac{\gamma_h}{2} + \left(\frac{1}{y_{JB}} - 1\right)\cot\frac{\gamma_h}{2}\right]^2 (\Delta\gamma_h)^2} \quad (4.12)$$

and

$$\frac{\Delta Q^2_{JB}}{Q^2_{JB}} = \sqrt{\left(\frac{2-y_{JB}}{1-y_{JB}}\right)^2 \left(\frac{\Delta F}{F}\right)^2 + \left[-2\cot\gamma_h + \left(\frac{1}{1-y_{JB}} - 1\right)\cot\frac{\gamma_h}{2}\right]^2 (\Delta\gamma_h)^2}. \quad (4.13)$$

It is seen that problems are encountered at high y. At low y, however, a good understanding of the calorimeter noise is required as a single cell in RCAL with an energy deposit of order of 100 MeV can bias the y_{JB} estimator by a few percent.

The resolution in y_{JB}, the only variable calculated with the JB method and used for the data selection in this analysis, is shown in Fig. 4.3 in the form of the $(y_{JB} - y_{true})/y_{JB}$ distribution for each kinematic bin. The JB method badly underestimates y, a shift of up to 30%, in the low x, Q^2 bins. In this region the current jet is well contained in the calorimeter but energy losses in the dead material reduce the measured hadronic deposits and thus, by Eq. 4.7, the value of y. This effect is compensated at higher x where a fraction of the current jet is lost in the forward beampipe. The resolution is of the order of 25 to 30%.

4.4.3 Double Angle Method

As an attempt to break free from the absolute energy scale uncertainty the Double Angle method [43] has been formulated. It uses both the scattered electron polar angle θ'_e and the hadronic flow angle γ_h (see Eq. 4.10). The relevant formulae are rather lengthy:

$$y_{DA} = \frac{(1-\cos\gamma_h)\sin\theta'_e}{\sin\gamma_h + \sin\theta'_e - \sin(\gamma_h + \theta'_e)}, \quad (4.14)$$

$$Q^2_{DA} = \frac{4E_e^2 \sin\gamma_h (1+\cos\theta'_e)}{\sin\gamma_h + \sin\theta'_e - \sin(\gamma_h + \theta'_e)}, \quad (4.15)$$

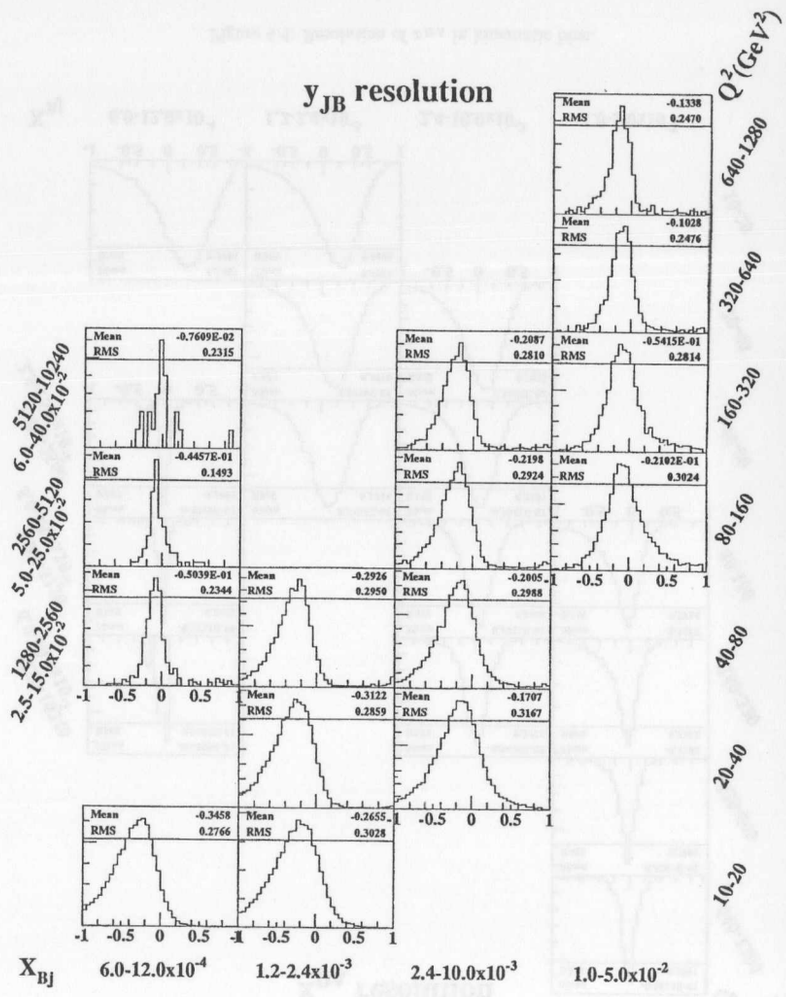

Figure 4.3: Resolution of y_{JB} in kinematic bins.

and

$$x_{DA} = \frac{E_e}{E_p}\frac{\sin\gamma_h + \sin\theta'_e + \sin(\gamma_h+\theta'_e)}{\sin\gamma_h + \sin\theta'_e - \sin(\gamma_h+\theta'_e)}. \quad (4.16)$$

It has been shown that the overall resolution of this method is superior in most of the kinematic plane compared to other recipes. Even more important, the use of Q^2_{DA} and x_{DA} has been observed to have a positive impact on the quality of the Breit Frame reconstruction. Nevertheless it is instructive to identify the regions where the accuracy may be degraded:

$$\frac{\Delta x_{DA}}{x_{DA}} = \sqrt{\left(\frac{-1}{\sin\theta'_e}\right)^2(\Delta\theta'_e)^2 + \left(\frac{-1}{\sin\gamma_h}\right)^2(\Delta\gamma_h)^2} \quad (4.17)$$

and

$$\frac{\Delta Q^2_{DA}}{Q^2_{DA}} = \sqrt{\left(\frac{y_{DA}-2}{\sin\theta'_e}\right)^2(\Delta\theta'_e)^2 + \left(\frac{-y_{DA}}{\sin\gamma_h}\right)^2(\Delta\gamma_h)^2}. \quad (4.18)$$

Due to its compact signature in the calorimeter the direction of the electron is in general reconstructed much more accurately than the hadronic energy flow direction so that the γ_h contributions are dominant in the above equation.

As discussed in the previous subsection the hadronic angle γ_h should not be too small in order to ensure good acceptance in the vicinity of the FCAL beampipe hole. In the same kinematic region the uranium noise signal passing the suppression cuts, typically at a rate of a few cells per event, results in a significant systematic overestimation of γ_h. For this reason a cut $y_{JB} > 0.04$ is imposed to exclude poorly reconstructed events. The box cut on the electron impact point in RCAL introduced in the context of the electron method is applicable in the DA approach as well.

In Figs. 4.4 and 4.5 the distributions of $(x_{DA}-x_{true})/x_{DA}$ and $(Q^2_{DA}-Q^2_{true})/Q^2_{DA}$ are shown. It is seen that the quality of reconstruction generally improves with increasing Q^2. This is not so readily reflected in the printed RMS widths of the distributions due to significant tails but the tendency is obvious in terms of the widths of the fitted gaussian curves (not shown). At the same time as the distributions get narrower the systematic shift in x_{DA} decreases from about 15% at low Q^2 to a few percent at high Q^2.

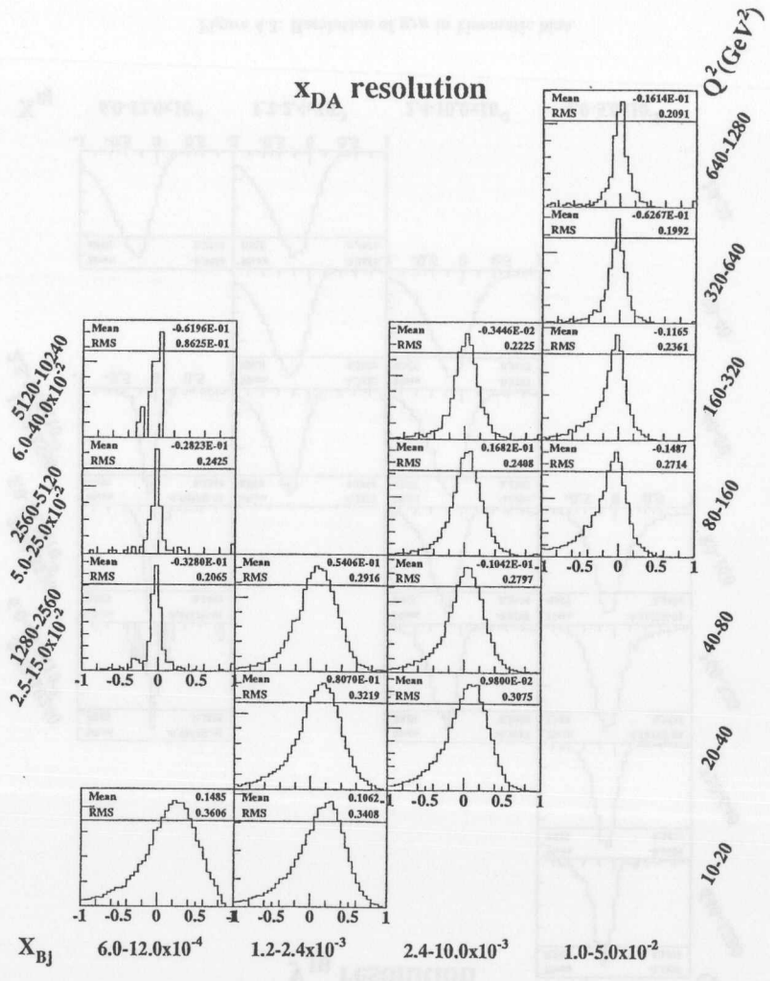

Figure 4.4: Resolution of x_{DA} in kinematic bins.

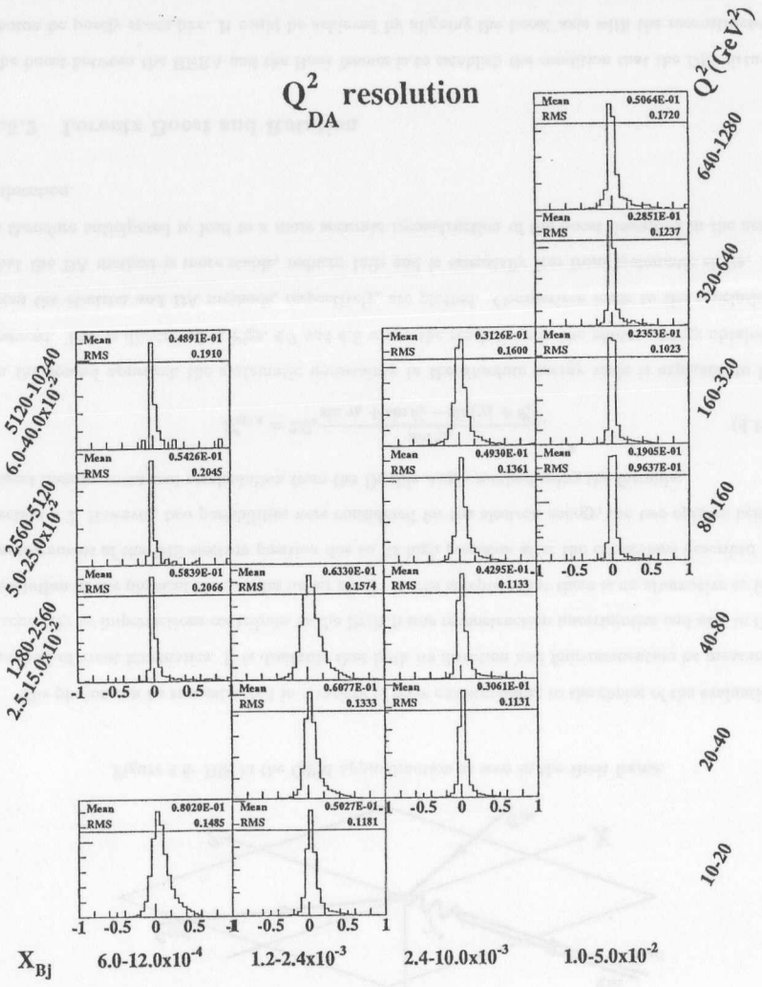

Figure 4.5: Resolution of Q^2_{DA} in kinematic bins.

4.4.4 Choice of the Method

The methods outlined in the previous subsections can be used simultaneously for selecting the data sample as concurrent cuts on variables reconstructed with different methods are possible. This is the case in this analysis where a low limit of 0.04 was demanded on y_{JB} and a high limit of 0.95 was imposed on y_e (to reduce photoproduction background). For the assignment of events to kinematic bins, however, a consistent set of quantities is preferable. In a dedicated investigation the Double Angle method has been found [43] to provide the best resolution over a wide kinematic range. This is consistent with the plots presented in previous subsections although there are still kinematic regions where the electron method is superior. A similar statement holds for the reconstruction of the Breit Frame described in the next subsection. For these reasons the Double Angle method has been chosen for this analysis.

4.5 Breit Frame

The Breit frame (BF), being neither the HERA frame[3] nor the center of mass system of the colliding particles, needs to be reconstructed from the measured kinematics of an event. The transformation from the HERA frame to the BF is a superposition of a Lorentz boost to make the virtual boson probing the proton purely space-like followed by a rotation to achieve proper alignment of the reference axes.

This section deals with details of the reconstruction of the Breit frame and gives the definition of the current fragmentation region in the Breit frame. A useful discussion accompanied by a derivation of formulae can be found in [44].

4.5.1 Virtual Photon

By definition the virtual photon[4] is completely space-like in the Breit frame, *i. e.* no energy is transferred to the hadronic system. Moreover, the photon and the proton are aligned and come into a head-on collision (see Fig. 4.6).

[3] The HERA frame coincides with the Breit frame for $x = E_e/E_p$ and $y = 1$.

[4] Neutral current DIS may also be mediated by a Z^0 boson. The generic term "photon" is used to denote this case as well.

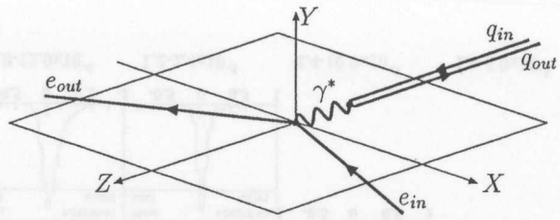

Figure 4.6: DIS in the QPM approximation as seen in the Breit frame.

The photon can be reconstructed in a variety of ways corresponding to the choice of the evaluation method of event kinematics. It is desirable that both its direction and four-momentum be measured accurately as imperfections contribute to the Breit frame reconstruction uncertainties and also to the resolution in the physical observables under study. It was accepted that there is no alternative to the measurement of the DIS electron position due to its high precision after the corrections described in section 4.2. However, two possibilities were considered for the electron energy, the two options being direct measurement and recalculation from the Double Angle method using the formula:

$$E'_{eDA} = 2E_e \frac{\sin\gamma_h}{\sin\gamma_h + \sin\theta'_e - \sin(\gamma_h + \theta'_e)}. \tag{4.19}$$

In the second approach the systematic uncertainty in the absolute energy scale is expected to be reduced. This is illustrated in Figs. 4.7 and 4.8 where the resolutions in the photon energy obtained from the electron and DA methods, respectively, are plotted. Comparison leads to the conclusion that the DA method is more stable, reduces tails and is essentially free from systematic shifts. It is therefore anticipated to lead to a more accurate reconstruction of the boost described in the next subsection.

4.5.2 Lorentz Boost and Rotation

The boost between the HERA and the Breit frames is to establish the condition that the DIS virtual photon be purely space-like. It could be achieved by aligning the boost axis with the reconstructed

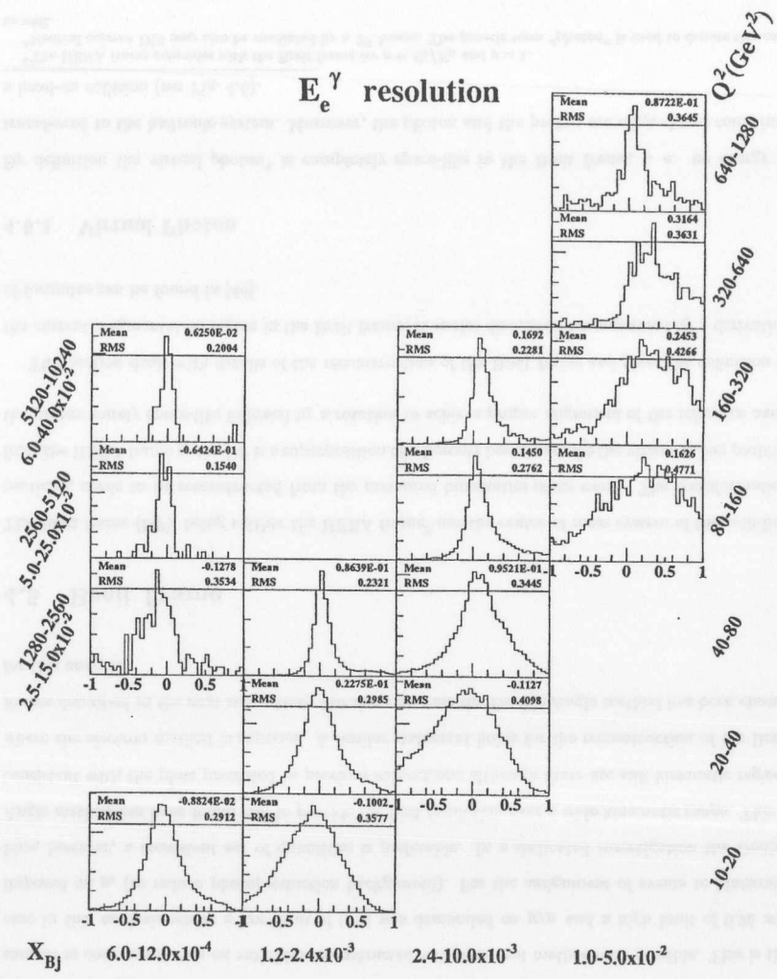

Figure 4.7: Resolution of E_e^{γ} from the electron method plotted in kinematic bins.

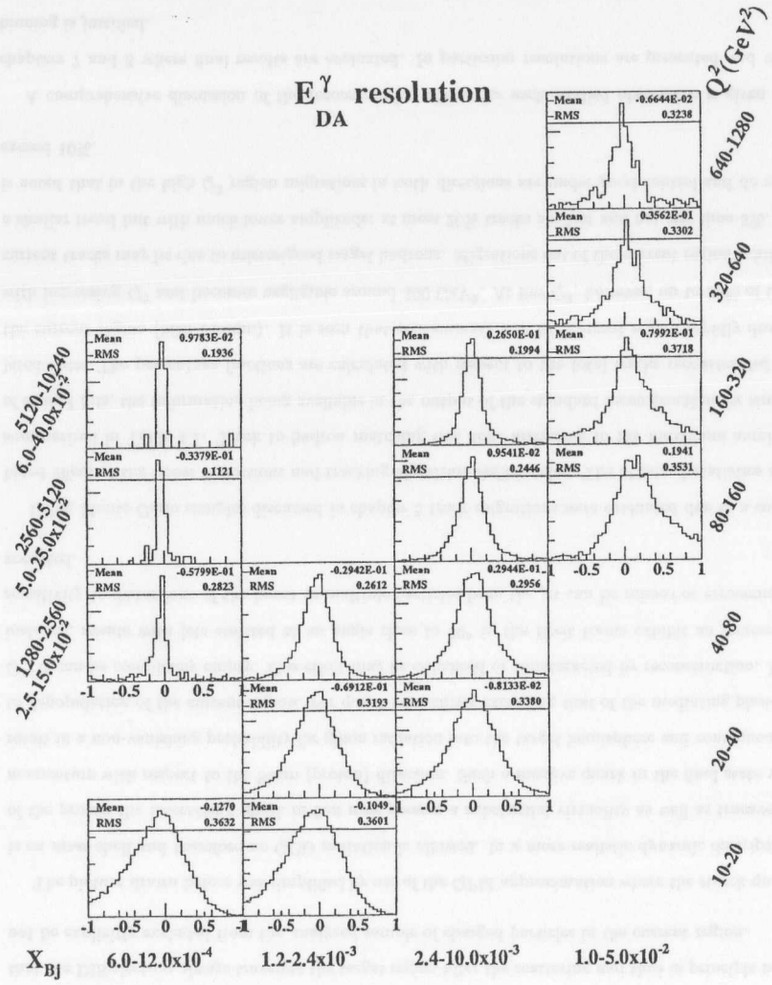

Figure 4.8: Resolution of E^{γ}_{DA} from the Double Angle method plotted in kinematic bins.

photon but a transverse component with respect to this direction is needed to force the photon to be parallel to the proton. As a result the velocity, $\vec{\beta}$, is given in terms of the virtual photon four-momentum, $q = (E_{\gamma^*}, \vec{q})$ and that of the proton $P = (E_p, \vec{P})$ measured in the laboratory frame:

$$\vec{\beta} = \frac{\vec{q} + 2x\vec{P}}{E_{\gamma^*} + 2xE_p}. \tag{4.20}$$

Actually, in the analysis region the velocity $\vec{\beta}$ is predominantly transverse with respect to the direction of the beams as the longitudinal momentum of the photon tends to balance the proton's contribution, $2xP_Z$, in the numerator.

Conventionally the proton direction is taken to be the positive Z-axis of the Breit frame. In addition the electron scattering plane contains the X-axis (in addition to the Z-axis). The rotation to satisfy these conditions may be calculated from the electron momentum boosted according to the Lorentz transformation employing $\vec{\beta}$ defined above. General formulae relating the four-momentum components in the two frames read:

$$\vec{p}_{BF} = \gamma(\vec{p}_{LAB} - \vec{\beta}E_{LAB}), \tag{4.21}$$

$$E_{BF} = \gamma(E_{LAB} - \vec{\beta}\cdot\vec{p}_{LAB}). \tag{4.22}$$

$\gamma = (1-\vec{\beta}^2)^{-\frac{1}{2}}$ is the Lorentz contraction coefficient.

Misreconstruction of the electron has therefore a twofold impact on the Breit frame. A residual boost exists between the true Breit frame and that calculated from the measured electron. This typically leads to distorted momenta of particles whereas displaced axes and therefore misidentified current region result from an inaccurate position of the reconstructed electron (*i. e.* the virtual photon direction). In fact these two effects are mixed and cannot be disentangled.

Charged particles were left unidentified in this analysis and consequently for the purpose of the transformation in Eq. 4.21 a pion mass assumption was made.

4.5.3 Current Fragmentation Region

The interest of this analysis concentrates in the current fragmentation region formed by particles with negative Z-components of momentum as measured in the BF. It is to be contrasted with the proton

remnant having positive p_Z and thus not expected to contribute significantly to the current hemisphere even after rearrangements of energy flows in the process of hadronization. A simple exercise shows that the DIS electron always traverses the target region after the scattering and thus in principle need not be explicitly excluded from the analyzed sample of charged particles in the current region.

The picture drawn before was simplified by use of the QPM approximation where the struck quark is on mass shell and therefore no QCD radiation is allowed. In a more realistic dynamic description of the proton the interacting quark in fact may possess a substantial virtuality as well as transverse momentum with respect to the beam (proton) direction. Such a massive quark in the final state will result in a non-vanishing probability for gluon radiation into the target hemisphere and consequently to depopulation of the current region. For quark virtualities exceeding that of the mediating photon, Q^2, it can be completely empty. This effect may be enhanced or counteracted by reconstruction. For instance, events with jets emitted at an angle close to 90° in the Breit frame exhibit an increased sensitivity to distortions of the boost as multiple particles from the jet can be missed or erroneously accepted.

Using Monte Carlo samples discussed in chapter 5 track migrations were evaluated due to a combined effect of the boost distortions and tracking algorithm defficiencies. The obtained statistics are summarized in Table 4.1. Track to hadron matching was done according to the maximum number of shared hits, the information being available in the output of the standard reconstruction for simulated data. The percentage fractions are calculated with respect to the total tracks *reconstructed* in the current region (fifth column). It is seen that contamination of the current region rapidly drops with increasing Q^2 and becomes negligible around 300 GeV^2. At low Q^2, however, up to 30% of the current tracks may be due to misassigned target hadrons. Migrations out of the current region exhibit a similar trend but with much lower amplitude: at most 20% tracks are lost and not less than 3%. It is noted that in the high Q^2 region migrations in both directions are under good control and do not exceed 10%.

A comprehensive discussion of the reconstruction effects for each studied observable is given in chapters 7 and 8 where final results are evaluated. In particular resolutions are presented and the binning is justified.

x	Q^2 GeV^2	Current tracks			Target tracks		
–	–	all	to TGT	%	all	to CUR	%
0.0006–0.0012	10–20	33403	3499	10.5	151207	10635	31.8
0.0012–0.0024	10–20	26875	3616	13.5	88344	6190	23.0
0.0012–0.0024	20–40	26901	2970	11.0	84783	5340	19.9
0.0024–0.0100	20–40	30561	6363	20.8	66670	3298	10.8
0.0012–0.0024	40–80	15462	1457	9.4	51881	2645	17.1
0.0024–0.0100	40–80	30658	3728	12.2	55714	2706	8.8
0.0024–0.0100	80–160	19375	1364	7.0	34654	1522	7.9
0.0100–0.0500	80–160	10301	1782	17.3	7699	176	1.7
0.0024–0.0100	160–320	32567	1617	5.0	53164	1968	6.0
0.0100–0.0500	160–320	41163	3902	9.5	26805	727	1.8
0.0100–0.0500	320–640	24649	1527	6.2	17689	472	1.9
0.0100–0.0500	640–1280	78155	2954	3.8	59343	1460	1.9
0.0250–0.1500	1280–2560	36434	1363	3.7	14714	236	0.6
0.0500–0.2500	2560–5120	12333	382	3.1	3461	44	0.4
0.0600–0.4000	5120–10240	4003	141	3.5	1017	10	0.2

Table 4.1: Track migration rates in the Breit frame.

Chapter 5

Monte Carlo and Detector Simulation

Reconstruction, described in the previous chapter, yields measurements biased by effects dependent on the particular setup of the ZEUS detector such as resolutions and acceptances. These effects need to be understood in order to correct for them or, at worst, accurately estimate the related experimental uncertainties. Otherwise uncorrected results would not address universal physics but rather specifics of the detector itself.

Being a complicated and extensive in space arrangement of subcomponents the ZEUS detector is not easily expressed in terms of the aforementioned performance parameters starting from first principles *i. e.* by taking its geometry and internal structure into account. A well established probabilistic method called Monte Carlo (MC) is used instead in high energy physics. The underlying assumption is that the good understanding of local properties can be transported to the macroscopic level through statistical analysis of simulated data. By averaging the calculated detector response over large samples of fake events and comparing it to the known input it is possible to learn about the systematic effects introduced by the apparatus. For the approach to work it is neccessary that the studied Monte Carlo sample is representative of the truly collected data set so that in the simulation the detector is probed in a similar manner.

In principle MC generators need not have any physics-related input to qualify for the correction

procedure provided the above condition is fulfilled. It seems, however, both easier and more practical to base them on our best knowledge of the studied high energy scattering process. This is probably the easiest way to tackle the obvious problem of the large variety of distributions that need to be modelled properly. Another advantage is that at the same price of extensive computer calculation, evaluation of the underlying physics assumptions is made. Hence, a comparison can be made and the MC model can be tuned or even revised using feedback from the experimental data.

This chapter deals with general issues related to the Monte Carlo methods and focuses on the details pertinent to the present analysis. The simulated samples of DIS events are described and their applicability to the collected DIS data is tested. The correction procedure is introduced but its application to the quantities measured in this analysis is postponed until chapters 7 and 8. Finally, the predictive power of the MC generators is addressed.

5.1 Event Generators

There can be different philosophies in implementing our knowledge of particle scattering at high energy in Monte Carlo generators. They share, however, a common pattern of separating the algorithm into distinct steps. The pattern reflects the theoretical understanding of interactions involving hadrons where perturbatively calculable hard processes at the partonic level are influenced by ingredients not calculable from first principles: hadronic structure in the initial state and hadronization in the final state. They are constrained by earlier measurements and phenomenological analyses. The validity of this separation is founded on the QCD factorization theorem.

Initially, a phase space point is chosen according to the differential cross-section for the process being simulated. This choice then constrains the configuration of the hard interaction for which the matrix element is calculated to a given accuracy in perturbation theory thus defining both the initial and final states of the interaction. The next step involves matching of the participating particles, in general deeply virtual, to an equivalent set of on-shell partons [1]. These are in turn either turned into observable hadrons by a phenomenological recipe for fragmentation (final state partons) or identified

[1] Actually, for partons which are never observable outside of hadrons the property of being on mass shell is not well defined. The evolution continues until the virtualities are of order of hadronic masses.

with partons resolvable in any hadron (initial state). This step concludes the procedure.

In the following subsections, generators relevant for this analysis are reviewed in the context of strategies adopted at each step of the calculation.

5.1.1 Parameterizations

For lack of a theoretical prediction, hadronic structure is parameterized in terms of parton distribution functions (PDF). They provide the probability of finding a parton carrying an arbitrary momentum fraction, x, inside a hadron moving fast with respect to the probe. It turns out that there is as well a dependence on the four-momentum transfer, Q^2, at which the hadron is being resolved into partons. This evolution, however, can be handled by perturbative QCD so that only the initial condition at an arbitrary scale is needed.

Extraction of PDFs from the experimental data is a subject by itself. Typically, one assumes a reasonable functional form of the x dependence at the starting scale, Q_0^2, and applies perturbative evolution Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [45, 46, 47] to obtain predictions at higher scales. These predictions are constrained by a chosen set of experimental data allowing for a fit to a few unknown parameters. In addition to the freedom of the choice of the initial parameterization, the scale and the data there are some ambiguities related to the evolution itself. Hence, a number of fits exists and, for the convenience of use, they are available in a form of a library of callable routines maintained at CERN [48]. Provision for interfacing to this library is made in event generators.

The actual parameterizations used in this analysis: GRV-94 HO [49], CTEQ4 [50] and MRSA low Q^2 [51], were picked in the DIS renormalization scheme. They all include data from HERA [52, 53] covering earlier unexplored kinematic regions of DIS and thus provide a good description of the cross-sections observed at ZEUS.

5.1.2 Hard Subprocess

The hard scattering in the case of DIS is an electromagnetic interaction between the virtual photon and a quark originating from the proton. It is typically calculated to first order in the electromagnetic coupling $\mathcal{O}(\alpha_{QED})$. The leading order includes no QCD vertices but corrections of order $\mathcal{O}(\alpha_S)$ are taken into account. They are the QCD Compton (QCDC) and the boson-gluon fusion (BGF) diagrams shown in Fig. 5.1 a) and b), respectively. The leading order diagram differs from the QCDC process

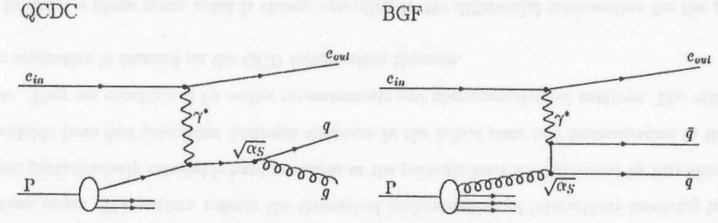

Figure 5.1: Feynman diagrams for the $\mathcal{O}(\alpha_s)$ processes in DIS.

only by the absence of gluon radiation in the final state.

The event generators used in this thesis, LEPTO [54] and HERWIG [55], support exact matrix element calculation for DIS to order $\mathcal{O}(\alpha_{QED}\alpha_s)$ and $\mathcal{O}(\alpha_{QED})$, respectively, *i. e.* HERWIG does not explicitly add contributions from BGF and QCDC. The former has the additional functionality of excluding individually QCDC and/or BGF contributions to facilitate interfacing to other generators.

5.1.3 Parton Showers

Particles involved in the hard scattering need to be dressed up with a shower of low virtuality partons. This process, in contrast to the hard matrix element calculation, is no longer of fixed order in the strong coupling constant as multiple parton branchings are needed. The splitting is discontinued as soon as the particles' virtuality approaches a low momentum transfer cut-off below which the perturbative calculation is not reliable any more and non-perturbative hadronization must take over in turn.

Two different philosophies of modelling the parton showers are employed in the MC programs referred to in this thesis. LEPTO and HERWIG use the DGLAP splitting functions, and in this way they take into account resummed to all orders leading logarithmic terms of the matrix element describing the dynamics of the branching. Effects due to the coherence of multiple parton emission are explicitly incorporated in both algorithms by requiring that emissions are in decreasing order in the relative angle between the two produced particles[2]. However, interference terms between the initial and final state showers are neglected. LEPTO supports matching of the phase space for the matrix element of the hard subprocess and the parton showering (MEPS) to avoid double counting or skipping of any regions.

An alternative approach is implemented in ARIADNE [56]. The algorithm relies on the assumption that gluon radiation can be calculated in a dipole approximation with dipoles stretched between color-charged partons. Partons multiply themselves in the branching process and so do the dipoles due to the non-abelian nature of QCD. Subsequent emissions are ordered in decreasing transverse momentum. By definition coherence is taken into account in the Color Dipole Model. In the collinear region, *i. e.* small p_T, the DGLAP splitting kernels are reproduced. An important feature is that the algorithm treats the initial and final state partons on equal footing so that this kind of interference is accounted for as well. Being a special case of gluon radiation, the QCD Compton contribution is included in the formalism but the BGF process has to be added explicitly. This is done through an interface to LEPTO that is used as the hard process provider for ARIADNE.

5.1.4 Hadronization

There are two competing models for turning low virtuality partons into hadrons. They are rather drastically different, yet due to lack of a perturbative solution to the problem of hadronization neither can be disfavored from the theoretical point of view. Actually, they incorporate a great deal of phenomenological work invested in the subject to date.

HERWIG pursues an observation of the local character of parton confinement into colorless hadrons.

[2] In the initial state showering the ordering is in the emission angle of the "soft" parton with respect to the incoming parton.

The locality is to be understood in color space: a set of partons in a color singlet state is disfavored to expand. To make use of this observation HERWIG first splits all the gluons into quark pairs and then pairs them again but this time based on the color and branching history criteria. Thanks to the locality the clusters are typically low-mass objects and can be decayed into hadrons without too much of an effort.

The cluster method may be confronted with the Lund string model [57] as implemented in JETSET [58]. In this model hadrons are produced by breaking overstretched color strings between partons. Each end of a string is attached to a quark whereas gluons are represented by kinks in the string. Partons moving in random directions stretch the string until it is energetically favorable to break it thus converting energy stored in the tubular color field into a quark-antiquark pair created off the vacuum. As the strings become shorter the relative motion of their ends slows down until no more breaking is possible. Colorless segments are then identified with hadrons after suitable rearrangements. The string hadronization model is used by LEPTO and ARIADNE.

5.1.5 QED Radiative Corrections

A convenient tool exists for calculating higher order electroweak corrections to the simulated events. It is provided by the DJANGO6 package [59] consisting of an interface of the HERACLES program [60] to the LEPTO event generator. Effects of electroweak virtual corrections are consistently evaluated in the one-loop approximation and combined with a calculation of real photon emission off either electron or quark. There is no facility for taking electroweak radiative effects into account in HERWIG so they were ignored in the case of this generator and therefore contribute to the systematic uncertainty related to the MC model dependence.

5.2 Detector Simulation

Events produced in the MC generators undergo simulation of the detector response. The program performing this task, MOZART [24], is based on the GEANT package [61] and incorporates our best knowledge of the apparatus supported by the test beam data as well as our experience of to-date data

collection. Decisions of the ZEUS trigger are calculated separately by the program ZGANA. Both simulation utilities are updated according to the detector configuration and therefore exist in versions valid for specific running periods.

The fully simulated events are passed through the reconstruction procedure in a manner identical to that applied to the real data described in chapter 4 and then they are subject to the data selection criteria to be discussed in chapter 6.

5.3 Monte Carlo Samples

Two Monte Carlo generators are chosen for the final analysis of the detector effects and evaluation of correction factors. They are ARIADNE and HERWIG. The former has been identified to provide a good description of the hadronic final state in the HERA kinematic range, in particular in terms of observables studied in this thesis [12, 62]. The latter represents in many aspects a radically different philosophy of event simulation including the hadronization model being one of the important topics of the present analysis. Hence, in comparison with ARIADNE it is taken to set a limit on the uncertainty related to our ignorance of the details of nonperturbative phenomena. Since in HERWIG no provision is made to calculate QED radiative effects they also contribute to the overall difference between the two generators.

The actual samples were generated with default settings of the tunable parameters. Kinematic ranges were bound by lower limits on the Q^2 variable. In order to reduce statistical uncertainties additional samples covering the high Q^2 regions were prepared. A total of 1.6 million events were available for the 1995 data sample; 0.3 and 0.45 million events were simulated for 1996 and 1997, respectively. Due to negligible differences in the detector setup in those two years these data were considered uniform and, consequently, merged after making consistency checks.

5.3.1 Description of the Kinematic Variables

As pointed out earlier in this chapter, it is desirable that the generated sample reflect properly basic features of the true data so that confidence can be gained in the correction procedure based on such

a Monte Carlo method. This is verified in this subsection for kinematic variables.

Three ways of measuring the scattered electron energy are compared in Fig. 5.2: the uncorected raw estimate from the calorimeter shown in a), the value after SRTD or presampler correction b), and as given by the DA method (see Eq 4.19) plotted in c). The distributions are normalized to unity and only events passing the final event selection as described later in section 6.5 contribute. An obvious deficiency of the simulation is observed in the first case; it is reduced by applying the correction discussed in subsection 4.2.4. The discrepancy is largely removed in the third distribution: this energy is actually used for reconstruction of the Breit frame considered in section 4.5.

$E - p_Z$ flows presented in Fig. 5.3 exhibit a similar improvement upon correction of the electron energy as can be inferred by comparing figures a) and b) showing the uncorrected and corrected distributions, respectively. Significant discrepancies remain, however, in the lower $E - p_Z$ region, especially for HERWIG. They are related to the deficiency in the simulation of the electron contribution as may be concluded from the nearly perfect agreement of the hadronic part of $E - p_Z$ plotted in c). The effect of the poorly simulated tail of the distribution on the event selection is addressed in one of the systematic checks where the cut is shifted from 35 to 40 GeV.

Distributions of the kinematic quantities involved in the event selection procedure are shown in Fig. 5.4. The contributing events are those accepted in the final analysis sample. It is seen that all the distributions are very well described by both MC models. Distributions of the basic variables used for classification of the data into bins, x_{DA} and Q^2_{DA}, are followed accurately through many orders of magnitude in Fig. 5.4 a) and b), respectively. In case of the energy transfer variable, y, there are small discrepancies seen in the intermediate range, but the rapid drop around $y_e \approx 0.6$ seen in Fig 5.4 c) is well reproduced. The drop is due to the cut on electron energy at 10 GeV. There are no deviations in the hadronic estimate, y_{JB} (Fig. 5.4 d).

An approximately 10% admixture of large rapidity gap events (LRG) has been observed in DIS at HERA [17]. The events are characterized by a lack of hadronic activity in the forward direction as measured in terms of the pseudorapidity, η_{max}, of the most forward cluster of energy in excess of 0.4 GeV. This class of events is not well modelled in the Monte Carlo used in this analysis (Fig. 5.5).

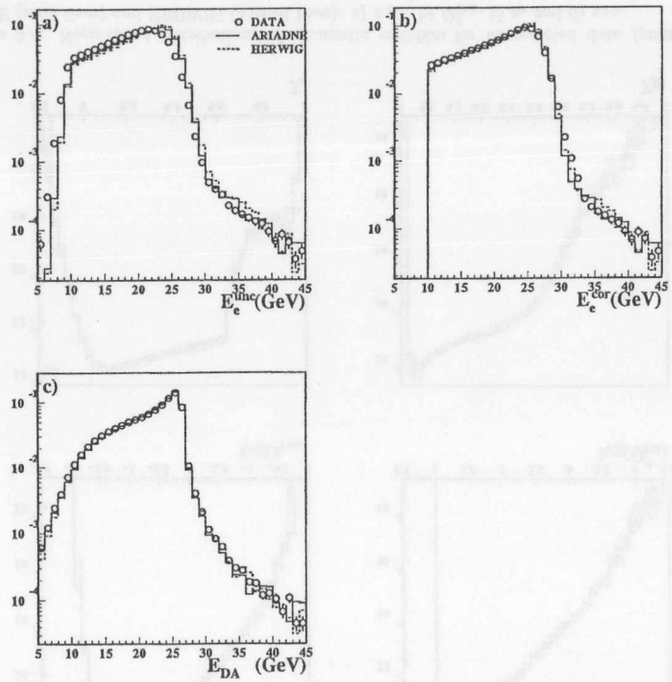

Figure 5.2: Electron energy distribution in uncorrected data (points), ARIADNE (solid lines) and HERWIG (dashed lines): a) before correction, b) after correction and c) from the Double Angle method.

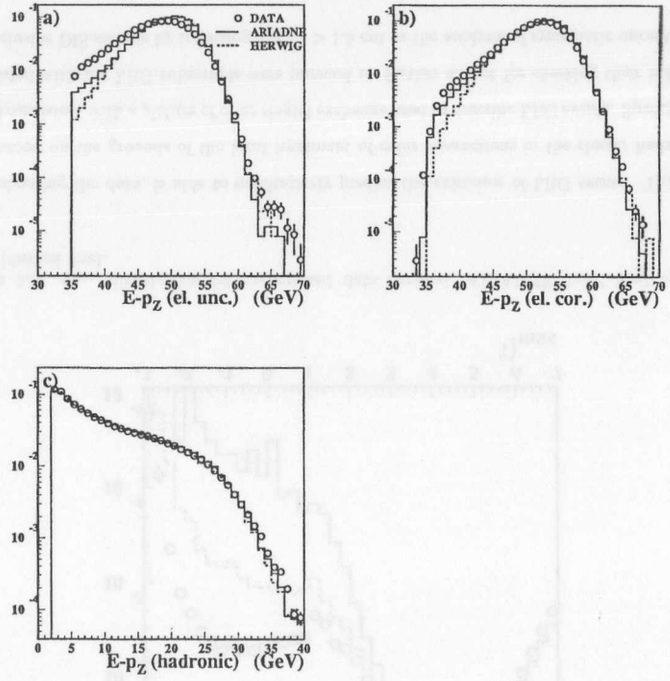

Figure 5.3: $E - p_Z$ distributions for uncorrected data (points), ARIADNE (solid lines) and HERWIG (dashed lines): a) with raw electron contribution, b) with corrected electron, c) hadronic contribution.

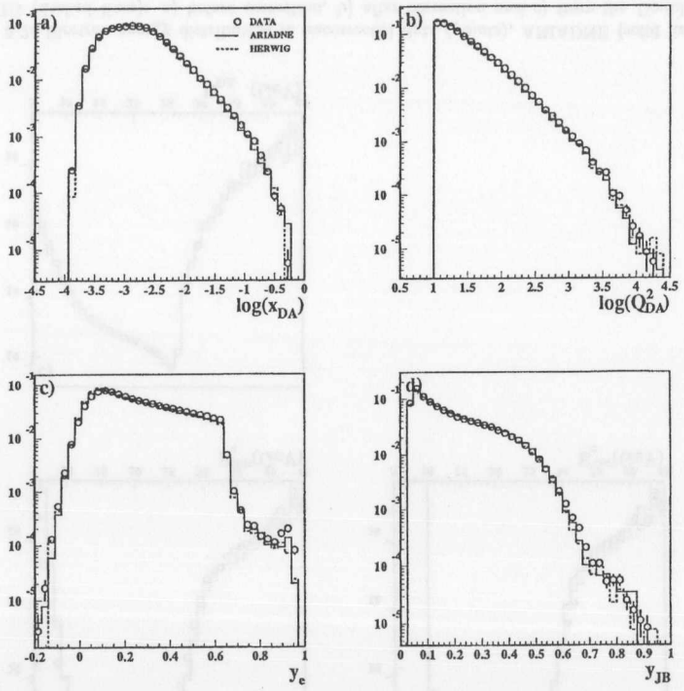

Figure 5.4: Normalized distributions of kinematic variables for uncorrected data (points), ARIADNE (solid lines) and HERWIG (dashed lines): a) x_{DA}, b) Q^2_{DA}, c) y_e and d) y_{JB}.

In particular, ARIADNE exhibits almost no sign of this type of events. HERWIG, while still strongly

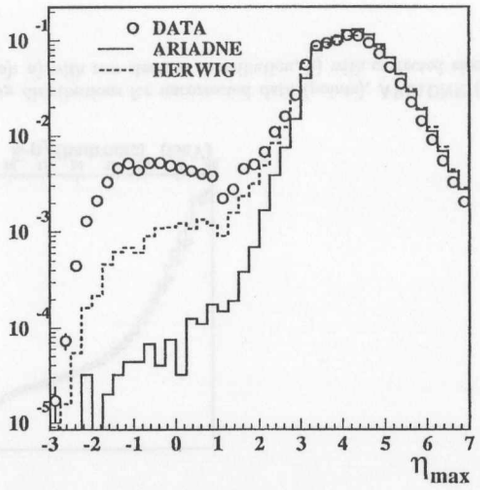

Figure 5.5: η_{max} distribution for uncorrected data (points), ARIADNE (solid line) and HERWIG (dashed line).

undershooting the data, is able to qualitatively predict the existence of LRG events. This may be understood on the grounds of the local treatment of color interactions in the cluster hadronization model consistent with a picture of color singlet exchange used to describe LRG events. Special features associated with the LRG subsample were pursued no further except for checking their influence on the inclusive DIS sample by imposing a $\eta_{max} > 1.5$ cut in the analysis of systematic uncertainties.

5.3.2 Description of the Tracking Variables

Simulation of the tracking system is studied next. A good agreement is seen in the distribution of the longitudinal position of the primary vertex, z_{vtx}, depicted in Fig. 5.6. The agreement is due to a good

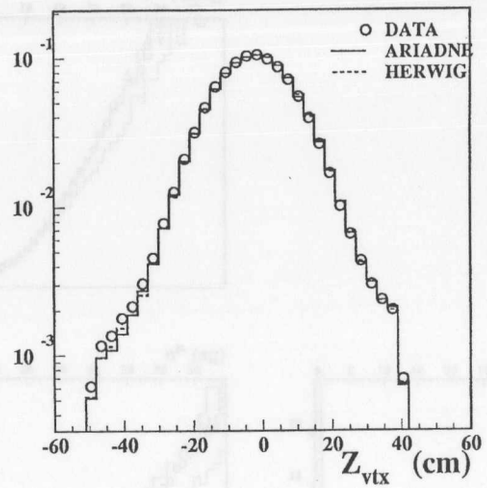

Figure 5.6: z_{VTX} distribution in uncorrected data (points), ARIADNE (solid line) and HERWIG (dashed line).

measurement of the vertex distribution which is one of the inputs to the detector simulation program. The distribution shown is averaged over the 1995 data taking period but there is no qualitative difference between the years. Even though it is centered at the origin an asymmetric cut is imposed in order to avoid a secondary peak due a satellite bunch associated with the proton beam.

Charged track multiplicities are shown in Fig. 5.7 for all reconstructed tracks, those attached to the primary vertex and a subset of vertex tracks reaching at least the third superlayer of the CTD in plots a), b), and c). It is seen that a slight underestimation of total multiplicity by ARIADNE vanishes as the assignment of a vertex requirement is imposed. HERWIG, however, displays a flatter distribution overestimating the fraction of high multiplicity events whereas the opposite is true for the intermediate range. The effect is reduced but still prominent for tracks satisfying topological cuts. A similar discrepancy although weaker is seen for the data collected in the following years. This is not currently understood.

Evolution, variation with the kinematic subregion defined by the kinematic binning, of the charged track pseudorapidity distribution, is shown in Fig. 5.8. The kinematic bins are defined and discussed in section 6.5. Comparison is made only to the ARIADNE simulation which gives a good description of the data in all the bins. The distribution conveniently illustrates the correlation of the hadronic activity (*i. e.* track pseudorapidity) with the y variable: the smaller the y (large x at fixed Q^2 or small Q^2 at fixed x) the more forward peaked is the η^{tr} distribution.

A number of other quantities including transverse and total momenta were checked and observed to be described at least reasonably well thus supporting the idea of using the generated MC samples to correct for detector effects.

5.4 Detector Effects

Distortions introduced to the measured quantities by detector imperfections and limitations can be quantitatively characterized in terms of experimental resolutions that in turn constrain the respective binning. It should not be finer than the resolution, otherwise significant migrations occur and more sophisticated approaches than the bin-by-bin method aimed at in this thesis are needed. Once the

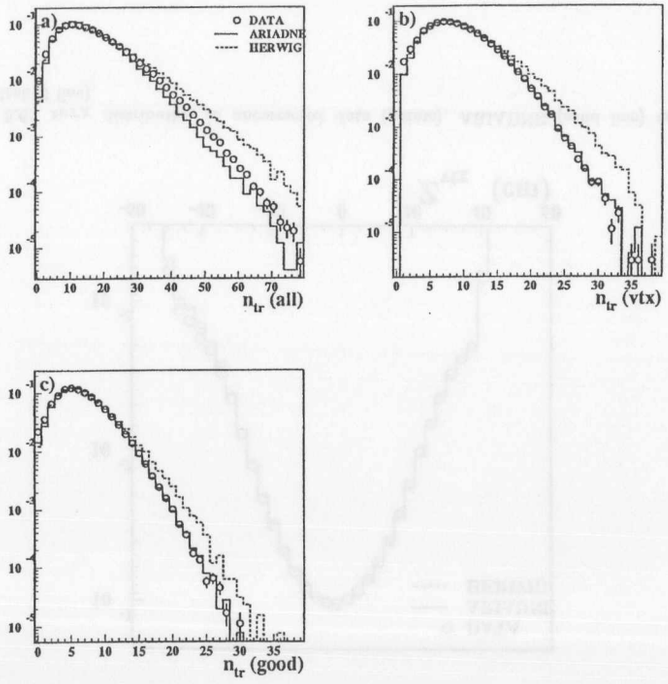

Figure 5.7: Track multiplicity distributions in uncorrected data (points), ARIADNE (solid lines) and HERWIG (dashed lines): a) all tracks, b) vertex tracks and c) vertex tracks passing the CTD superlayer cut.

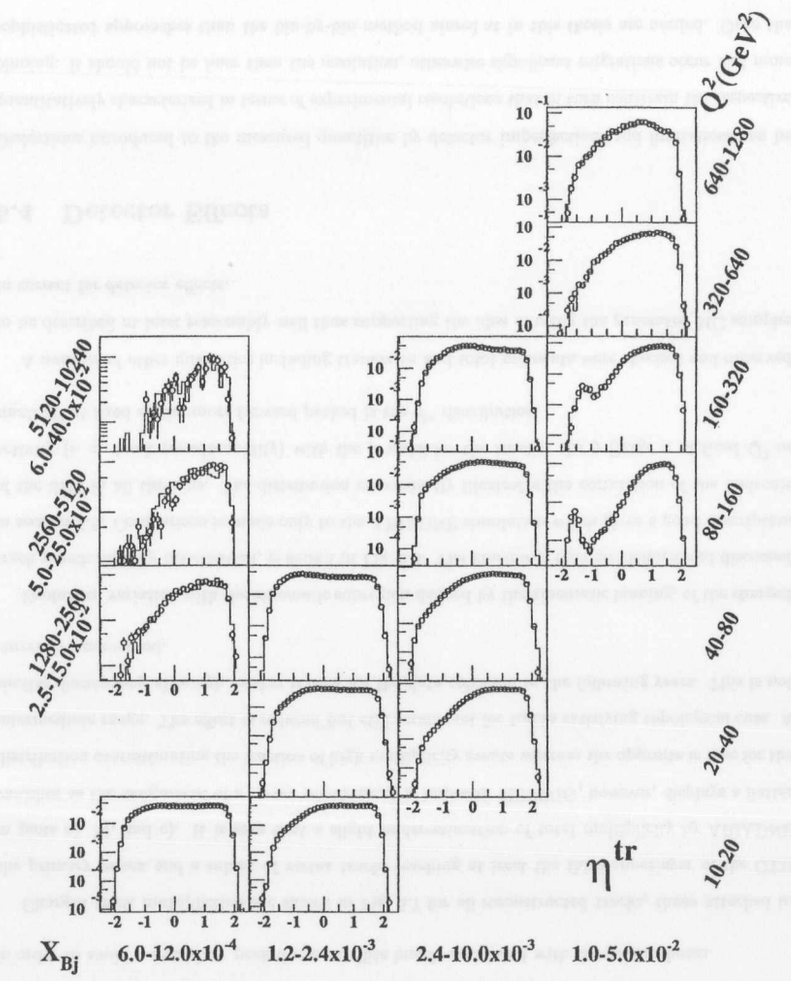

Figure 5.8: Track pseudorapidity (LAB frame) distributions in kinematic bins for uncorrected data (points) and ARIADNE (lines).

binning has been established its appropriateness can be verified by evaluating purities and efficiencies. Formulae relevant for these procedures are given in the following subsections.

5.4.1 Resolution

The resolution may be defined as the variance of the absolute difference, $\Delta = X_{meas} - X_{true}$, between the true and measured observable:

$$\text{Resolution} = \sqrt{<\Delta^2> - <\Delta>^2}, \tag{5.1}$$

or the spread of the relative difference, $\Delta_r = (X_{meas} - X_{true})/X_{meas}$:

$$\text{Resolution} = \sqrt{<\Delta_r^2> - <\Delta_r>^2}. \tag{5.2}$$

Being just one of the moments of the difference distribution, the resolution provides only a limited description: at least systematic shifts characterized by the mean value should be checked as well. Better yet, it is desirable to examine visually the distribution itself in order to get an insight to its behavior, the presence of non-gaussian tails or other possible complications.

Compatibility of binning with the resolution is typically expressed by a requirement that each bin be large enough to contain at least the mean±spread interval. For distributions exhibiting clear deviations from a gaussian shape it may be justified to adjust the size of the bin to accommodate an appropriate fraction of the tails of the distribution.

5.4.2 Efficiency and Purity

The correctness of the binning may be quantified in terms of purity and efficiency defined as:

$$\text{Purity} = \frac{\text{Number of entries generated } \textit{and} \text{ reconstructed in the bin}}{\text{Number of entries reconstructed in the bin}} \tag{5.3}$$

and

$$\text{Efficiency} = \frac{\text{Number of entries generated } \textit{and} \text{ reconstructed in the bin}}{\text{Number of entries generated in the bin}}. \tag{5.4}$$

Under the assumption of a bin size comparable to the resolution, in the absence of systematic shifts for a flat distribution both purity and efficiency of order 70% are expected. Influence of the neighboring bins may be enhanced in the case of rapidly varying distributions leading to decreased purities and efficiencies. Hence, values of order of 40% are still acceptable.

5.4.3 Correction Methods

A limited degree of migration between the bins justifies the use of a simple bin-to-bin correction method relying on the calculation of a multiplicative correction factor separately in each bin i as a ratio of the generated and reconstructed values

$$C_i = \frac{\mathcal{O}_i^{gen}}{\mathcal{O}_i^{rec}}, \tag{5.5}$$

with the statistical uncertainties evaluated under the assumption of independence of the generated and reconstructed distributions. This leads to a conservative estimate as obvious correlations between the two levels decrease the uncertainty.

When the factor C_i is applied to the uncorrected data,

$$\mathcal{O}_i^{cor} = C_i \mathcal{O}_i^{unc}, \tag{5.6}$$

it is believed to remove all the experimental bias and effects introduced by detector, reconstruction and experimental cuts unless the latter are used in the treatment of the generated MC sample as well. It is expected that the correction factors are neither significantly different from unity nor vary rapidly with abcissa.

The above formulae are used in chapters 7 and 8 to unfold the true values of observables studied in this thesis.

5.5 Monte Carlo Predictions

The statistics of the fully simulated samples are limited by the computing power at the disposal of the ZEUS Monte Carlo team. However, a large overhead of the detector response simulation may be saved if only the generated hadron level is desired. This is the case for obtaining the truth level predictions of the MC models to be compared with the corrected experimental data.

For this purpose the truth level was taken to be given by stable hadrons as defined by the default setting in JETSET *plus* the short-lived neutral kaon, K^0_S, and the neutral hyperon Λ^0.

Multiple calculations were performed using ARIADNE, LEPTO and HERWIG. Predictions are given in appropriate plots together with experimental results.

Chapter 6

Data Selection

In this chapter the path starting at the first level trigger and ending in the final data sample for the analysis is described. Selection criteria emerging from the on-line selection strategy as well as detector limitations are presented in a systematic manner. Kinematic bins are defined and characterized, and statistics of the data forming the basis for the analysis are given.

6.1 Trigger

Ideally, the trigger would be a black box doing fast and reliable filtering of the data as it pours from the detector. Obviously, there are lots of uninteresting interactions like proton beam-gas backgrounds that must be efficiently rejected except for dedicated studies of the trigger performance. This needs to be achieved without biasing the accepted part of the data sample and avoiding unnecessary losses as in many cases statistical uncertainties are still dominant. The high nominal bunch-crossing rate of 10 MHz makes this task particularly troublesome at HERA. In short, the aim of the trigger is to discard the backgrounds while keeping efficiency at a maximum for the selected classes of physics events.

The logic applied at each of the three trigger levels at ZEUS upon collection of the reported data sample is outlined in the following. A comprehensive description of the on-line DIS data selection may be found in [63].

6.1.1 First Level Trigger

Information about the event available at the FLT is limited. At the same time a substantial reduction of the rate must be achieved. The DIS selection is based on the detection of a scattered positron in the EMC section of the calorimeter. Two slots are in use as an alternative:

FLT30 A coincidence of the RCAL isolated electron with a low resolution sum of all energy deposits in the EMC section of RCAL exceeding 3.75 GeV.

FLT44 An alternative of total EMC deposits summed at high resolution exceeding 3.4 GeV in the RCAL or 4.8 GeV in the BCAL. For the RCAL the towers adjacent to the beampipe were skipped.

In both cases timing vetoes originating from SRTD, VetoWall or beamline monitors are applied. Additional demands on the tracking were studied but not implemented due to an excessive loss of physics events of the order of 3% especially in the case of FLT44. Due to an increased luminosity this condition had to be imposed starting in 1996.

Based on the data collected in 1993 the efficiency of the alternative of FLT30 and FLT44 was estimated to be around 98% [64] and considered adequate for the data taking in the following years.

6.1.2 Second Level Trigger

The Second Level Trigger operates in an asynchronous mode at an input rate of up to 1 kHz as provided by the GFLT. The availability of a more accurate overview of the event permits us to safely apply vetoes against the most severe backgrounds. It is done essentially independent of the physics-oriented filtering. Cuts on the global timing are very effective in the proton beam-gas events rejection as they result typically in deposits in the RCAL 10 ns early with respect to activity in the FCAL. Similarly the total calorimetric $E - p_Z$ of less than 75 GeV, the limit set safely above the kinematic bound of $2E_e = 55$ GeV, is aimed at the same type of events. The lower cut on this variable (shown below) was specifically required by all the DIS filters to suppress events with hard initial state QED radiation as well as to enhance the efficiency of the beam-induced background rejection. Cosmic rays,

mainly muons, initiated background is reduced by cutting on the difference of the up-down timing of the calorimeter. In addition, events with sparks in the calorimeter were identified and removed.

The data sample for this analysis was expected to satisfy each of the conditions of the DIS01 filter:

1. any of the DIS triggers at the FLT had fired *or* a high total $E_T > 25$ GeV in conjunction with a combined CAL+LUMI cut of $E - p_Z + 2E_\gamma > 29$ GeV;
2. either of the total deposits in the electromagnetic sections of RCAL, BCAL or FCAL or in the hadronic section of FCAL exceeded 2.5 GeV.

The E_T condition in 1 is supposed to increase the acceptance for high Q^2 events whereas the minimum deposit requirement in 2 discards events that are empty from the calorimeter point of view.

The events accepted by the DIS01 filter amounted to a cross-section of about 1μb in 1995. The logic and cuts were stable throughout the three years of data collection covered by this analysis.

6.1.3 Third Level Trigger

On-line selection is concluded at the Third Level Trigger where all the data pertaining to the considered events are at hand. A nearly complete analysis is thus performed, simplified slightly due to CPU time limitations.

Treatment of a DIS electron was to look for any algorithm returning a satisfactory candidate. The set of active electron finders varied between the years, the neural-network based Sinistra was implemented on-line beginning in 1996. The cuts on the quality of the candidate were suitably relaxed with respect to the final analysis. In such a relaxed mode, the overlap between the algorithms is very high and justifies the off-line usage of Sinistra for the 1995 data sample as well. The position of the electron was refined with the use of SRTD data if a good match was found.

Of the multiple DIS filters the so-called "medium Q^2" DIS03 filter was chosen for this analysis. The requirements consisted of:

- any DIS bit at FLT set;

- the cut on $E - p_z + 2E_\gamma$ tightened to 30 GeV;
- a redundant demand of $E - p_Z < 100$ GeV (looser than at SLT);
- an electron candidate with an energy of $E'_e > 4$ GeV from either of the algorithms;
- a cut on the electron impact position on the face of the RCAL evolving from 24×24 cm^2 to 28×28 cm^2 throughout the running period[1].

Events satisfying the above selection criteria were picked from the data sets and made available to the off-line analysis.

6.2 Off-line Preselection

In the process of the data reconstruction done off-line by ZEPHYR, where all the calibration information can be consistently used, events are classified according to their properties. Thus, for the reprocessed data, an additional preselection can be applied.

To obtain slightly overlapping but otherwise complementary samples, different criteria were adopted for the present analysis with respect to the data sets collected in 1995 and 1996/97. Due to the much increased total luminosity it was decided to use a more restrictive preselection in the latter case where the low Q^2 events had been already heavily prescaled by the trigger and therefore hardly usable. On the other hand the uncertainties of the measurements in this region based on the 1995 data were generally already dominated by systematics effects and therefore a further increase of statistics was not so important.

The preselection employed the so-called Data Summary Tape bits flagging particular classes of events. The bits relevant for this analysis are introduced below.

[1] As of late 1996 the box-cut was changed to a radial cut at 25 cm.

6.2.1 Low Q^2 Data Sample

Analysis of the 1995 data sample was meant to cover the lowest Q^2 region safely accepted by the trigger and extend up to where the measurement still made sense due to statistics decreasing with Q^2. It was therefore decided to use the DST bit 14 defined below.

- An electron candidate from any of the finders (including Sinistra) with energy $E'_e > 5$ GeV;
- SRTD-corrected position of the candidate not less than 13 cm from the beampipe in either the X or Y direction *or* no matching hit in the SRTD;
- $E - p_Z + 2E_\gamma > 35$ GeV;
- A timing cut at 8 ns on $|t_{global}|$, $|t_{FCAL}|$, $|t_{FCAL} - t_{RCAL}|$ and $|t_{RCAL}|$ (only for a total energy $E_{RCAL} > 5$ GeV).

The above requirements were imposed on top of the global DIS filter selection. This included demanding at least one DIS flag at the TLT to be fired or one of the complementary very high Q^2 flags. Vetoing events initiated by muons from either cosmic showers or the proton beam was activated as well.

Around 1.4 million out of the total recorded 15.6 million physics events corresponding to a total luminosity of 6.3 pb^{-1} survived up to this point.

6.2.2 High Q^2 Data Sample

Detector setup and trigger configuration did not undergo major changes between 1996 and 1997 so this data sample can be considered uniform. There were minor differences in the vertex distribution for example but their combination is not expected to lead to an inconsistency. This is addressed in more detail in chapters 7 and 8 where an explicit comparison between results on the data samples obtained each year is made.

The DST bit 18 adopted for both 1996 and 1997 imposed, *on top of the DST bit 14 requirements*, the following cuts:

- Tighter muon rejection;
- Calorimeter-only cut $E - p_Z > 35$ GeV;
- Primary vertex found with at least one track attached to it; the vertex position is in the range $-50 < z_{VTX} < 100$ cm;
- Tighter timing cut, $|t_{RCAL}| < 4$ ns;
- Electron candidate of energy above 10 GeV *found with Sinistra*;
- Box cut for the electron in RCAL at 32×32 cm^2;
- Kinematic bounds: $y_e < 0.95$, $y_{JB} > 0.04$ and $Q^2_{DA} > 10$ GeV^2.

The integrated luminosity analysed for 1996 was 10.5 pb^{-1} and the corresponding event yield reached almost 0.6 million events. The numbers for 1997 were 26.6 pb^{-1} and 1.4 million, respectively. These were selected from the total samples of 17.4 (27.8) million physics events collected in 1996 (1997).

Most of the cuts listed above were nearly identical to the final selection criteria described in the next section, except for the Q^2 range which was bounded by 80 GeV^2 from below.

6.3 Track Selection

The central analysis used a set of tracks reconstructed by the VCTRAK package as described in section 4.3. Criteria for the track selection for this analysis are in their majority governed by the chambers' acceptance and limitations of the reconstruction algorithm. As a result they have a lot in common with the cuts used in earlier analyses of charged particles [65, 66] although some differences are present.

The cuts were:

- Assignment to the primary vertex;

- Not matched to the Sinistra electron;
- Transverse momentum $p_t > 0.15$ GeV;
- Pseudorapidity $|\eta| < 2.0$;
- Hits in the three innermost superlayers of the CTD;
- Assignment to the current hemisphere of the Breit frame.

Both kinematic limits were imposed in the ZEUS frame, *i. e.* before boosting to the BF.

Since the tracks were sought for in the pool of hits not limited to the CTD and VXD, the polar angular acceptance was increased compared to regular CTD-limited analyses. Thus the pseudorapidity range was relaxed by 0.25 on each side. This range corresponds to trajectories inclined not less than 15° to the beam pipe. The superlayer requirement, however, reduced the acceptance near the edges of the angular range thus partially cancelling the loosened cut. As one of the systematic checks the pseudorapidity range was tightened to $|\eta| < 1.75$ and this effect was not found to be significant.

The limit on the transverse momentum corresponds to the curvature of radius of approximately 35 cm. Hence, such a particle emerging from the nominal interaction point can barely reach the third CTD superlayer. It is therefore seen that, again, cuts evaluated in the track selection are closely related and sometimes redundant.

6.4 Cleaning Cuts, Backgrounds

Since deep inelastic scattering contributes only a small fraction of the total ep scattering cross-section dominated by photoproduction ($Q^2 \approx 0$) it is important to control and understand the backgrounds. Contamination due interactions other than ep also need to be suppressed. Each of the cuts used in this analysis is given in turn below; the background aimed at and the workings of the method are also provided.

- Records are kept of hardware failures and other problems affecting the data quality. Only data satisfying minimum requirements on the detector setup (operational CTD and calorimeter, no

known bugs in the trigger logic etc) are accepted. This is achieved through a referral to a centrally maintained database.

- Momentum conservation requires that for a fully contained event the total $E - p_Z = 2E_e = 55$ GeV. This is used starting from the second level trigger onwards and the final off-line cut is $35 < E - p_Z < 65$ GeV with the quantities summed over the hadronic system *plus* electron contribution calculated after energy corrections.

  Photoproduced events where the electron is not detected peak at low values whereas beam-gas interactions depositing energy into the RCAL populate the high $E - p_Z$ tail, otherwise the latter exhibit a similar behavior to the photoproduction background. Radiative events with an energetic initial state photon are also suppressed. The cut is not sensitive to finding an electron[2] and thus is safe to use.

- The timing pattern of the activities in different parts of the calorimeter can distinguish between interactions taking place inside the detector and those happening outside, typically upstream in the incoming proton direction: the RCAL is then hit early with respect to the FCAL. To reduce the rate of this type of event it is required that $|t_{RCAL}| < 4$ ns provided a substantial energy deposit in RCAL is recorded. A huge background rate is present at $t_{RCAL} \approx -10$ ns corresponding to the time it takes for relativistic particles to traverse the detector in the longitudinal direction.

- The impact position of the electron in the RCAL is bound to be not closer than 16 cm in either the x or y direction (box cut). This not only improves the performance of the electron finding algorithms as discussed in section 4.2 but also suppresses beam-gas events emerging from the beampipe to hit the RCAL EMC section. The activity is then concentrated in the very inner part of RCAL. Hence, the box cut provides a good handle on this type of background.

- Demand for a vertex to be found is a must for a tracking analysis from the physics point of view alone as only hadrons produced in the primary hard interaction should be considered. However, the cut on the longitudinal position, $-50 < z_{VTX} < +40$ cm, does two additional things. It ensures that the emanating tracks can be safely measured within the CTD. At the same time it

[2] The electron energy corrections are limited and do not bias the cut.

removes vertexless events in which the calorimeter could be fooled due to its relative blindness to the direction from which the particles enter it.

The types of backgrounds affected by this cut are photoproduction, as neither electron nor other charged particles need to be produced in such a case, and again beam-gas events taking place inside the inner detector volume. Likely candidates for rejection are also muons of any kind traversing the detector at random locations.

- Muon events are more systematically searched for by algorithms aiming at such signatures as:
  - Calorimetric timing consistent with minimum ionizing particles traversing it downwards (cosmic muons);
  - High momentum tracks spanning the entire CTD in the transverse direction not compatible with the beamline (cosmic muons);
  - Sequence of longitudinal low energy deposits in the BCAL (halo muons).

  These criteria can be looked for in coincidence with activity in the muon chambers.

- Elastic QED Compton events can be rejected based on their overconstrained kinematics if both the electron and photon are detected.

- Residual calorimeter spark events not filtered out at earlier stages of the selection are suppressed.

Using Monte Carlo techniques it is assessed that the photoproduction background is eventually reduced below 2% in the integrated sample with an improvement to less than 1% at higher Q^2.

6.5 Kinematic Selection

In the final kinematic selection the imposed conditions were rather similar to those already present at the off-line preselection stage based on the DST bit 18 (subsection 6.2.2). In this way the data collected in 1995 were put on the equal footing as the samples from later years by requiring a universal set of cuts. They are briefly described in the following.

6.5.1 Phase Space Cuts

The scattered electron energy E'_e above 10 GeV was demanded. This stems from the need for good performance of the electron finding algorithm. In particular, at lower energies showers produced by neutral pion decays are able to mimick the electron signature in the calorimeter. Hence, both purity and efficiency of the algorithm are degraded. In order to reduce the probability of misidentification and stick to the region best understood in the Monte Carlo simulation a cut as above was imposed.

For small values of the energy transfer variable, y, a significant fraction of the hadronic activity may escape down the beam pipe in the forward direction thus remaining undetectable. An accurate determination of the kinematics is not possible for such events translating into a badly reconstructed Breit frame. This is true also for the method relying only on the scattered electron, as the resolution in the x variable is poor in this kinematic region. These effects may be minimized by cutting on the y estimator from the hadronic system (the Jacquet-Blondel method): $y_{JB} > 0.04$.

A characteristic feature of events with a misidentified electron is a high value of y obtained from the electron method. This is because typically such fake electrons are found in the forward calorimeter, where the majority of hadronic activity is usually observed, including neutral pions that are likely to mimic electrons. Due to the kinematics the DIS electron is bound to have an extremely high energy if detected in the FCAL, especially in the vicinity of the beampipe. The pion is not subject to such kinematic restrictions and, if taken as an electron candidate, leads to high values of y. These events are rejected if $y_e > 0.95$.

6.5.2 Binning

Kinematic bins for this analysis must conform to the resolutions in the reconstructed x and Q^2. At the same time the size of a bin is limited by the statistics needed for a measurement of a reasonable accuracy. Consideration of the acceptance for both entire events as well as particles produced in the current hemisphere constrains the kinematic region accessible in the ZEUS detector for studies in the Breit frame-of-reference.

The bins used in the present analysis have been defined and studied extensively in [66]. These in turn are based, with appropriate modification neccessitated by the tracking cuts, on the original bins developed in the proton structure function, F_2, measurement from the 1992 data [67]. These bins have been adopted as a standard for multiple analyses related to the Breit frame [12, 65, 68]. Guidance of the tracking detector acceptance has been used to extend the earlier definition to the high Q^2 region accessible only in the high statistics samples collected in 1996 and 1997.

Table 6.1 in addition to the kinematic ranges of the bins provides efficiencies and purities calculated

bin	x min | max	Q^2 GeV^2 min | max	efficiency %	purity %	track acceptance %
1	0.0006–0.0012	10–20	39	43	84
2	0.0012–0.0024	10–20	40	49	93
3	0.0012–0.0024	20–40	55	51	91
4	0.0024–0.0100	20–40	54	72	94
5	0.0012–0.0024	40–80	57	55	86
6	0.0024–0.0100	40–80	77	70	93
7	0.0024–0.0100	80–160	79	72	92
8	0.0100–0.0500	80–160	58	81	86
9	0.0024–0.0100	160–320	62	75	92
10	0.0100–0.0500	160–320	81	76	87
11	0.0100–0.0500	320–640	82	77	95
12	0.0100–0.0500	640–1280	83	78	95
13	0.0250–0.1500	1280–2560	85	78	92
14	0.0500–0.2500	2560–5120	82	74	91
15	0.0600–0.4000	5120–10240	79	82	93

Table 6.1: Kinematic bin definition, efficiency and purity (ARIADNE).

from Monte Carlo studies as well as the acceptance for hadrons produced in the current hemisphere in the Breit frame. The formulae relevant for the purity and efficiency evaluation are:

$$\text{Purity} = \frac{\text{Number of events generated } \textit{and} \text{ reconstructed in the bin}}{\text{Number of events reconstructed in the bin}} \tag{6.1}$$

and

$$\text{Efficiency} = \frac{\text{Number of events generated } \textit{and} \text{ reconstructed in the bin}}{\text{Number of events generated in the bin}}. \tag{6.2}$$

The same data are illustrated in Fig. 6.1, where the three numbers in each bin are, top to bottom, the kinematic efficiency, the purity and the tracking acceptance expressed in percent. There are a few

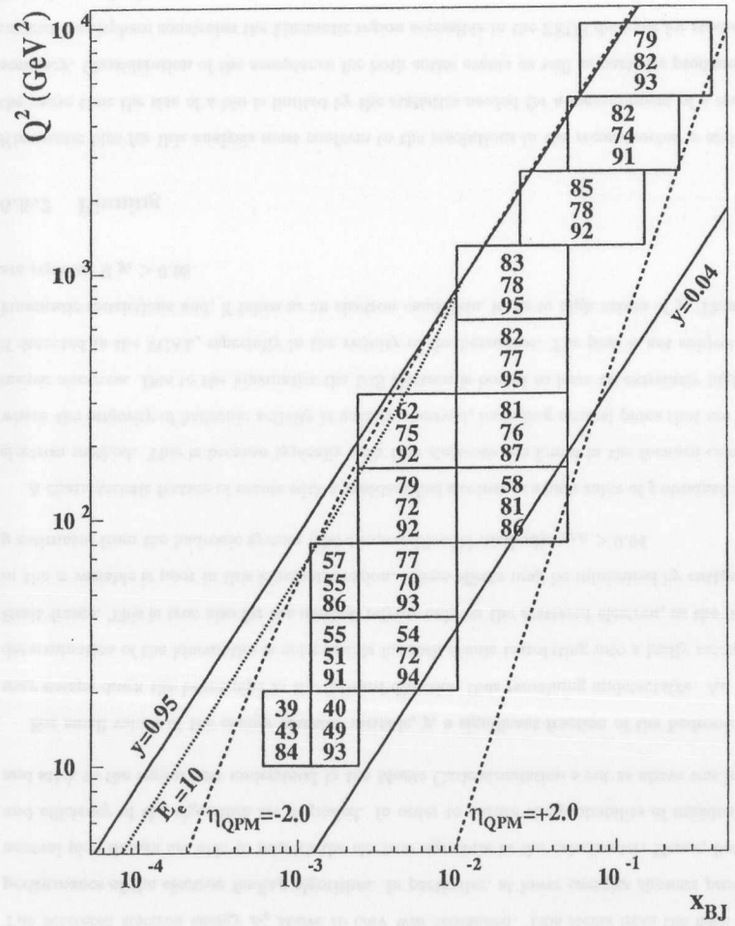

Figure 6.1: Analysis bins in the x–Q^2 plane.

things to be noted. Firstly, the layout of the bins in the x-Q^2 plane follows the boundaries given by the tracking chamber-imposed pseudorapidity cut, $|\eta| < 2.0$. They are drawn in the simplified Quark Parton Model approach where the direction of the struck quark, in the absence of QCD radiative corrections, defines the direction of the current jet. The phase space of good tracking acceptance is covered by the array of bins. No more than one charged hadron per six falls outside of the fiducial region defined collectively by the transverse momentum and pseudorapidity cuts. Secondly, the finer binning in the lowest x region together with a relatively poor resolution in that variable as seen in section 4.4 results in a degraded reconstruction showing up as low purities and efficiencies. Both quantities are affected by strong migrations between the bins and out of the accepted region. Thirdly, even at the cost of limited reconstruction accuracy it is worthwhile to have a handle on the x dependence of the studied observables. This justifies the aforementioned finer binning in the low x area. The combined effect of the limited statistics and tracking acceptance at high Q^2 does not permit a similar treatment of that corner of the phase space.

6.6 Final Data Sample

Distributions of the accepted events in the kinematic bins are summarized in Table 6.2. The 1995 data sample dominates statistically the total number of events due to the inclusion of low Q^2 bins. Above the threshold of 80 GeV^2, the 1997 data almost double both earlier years' samples combined. The region of overlap spans the range of Q^2 up to 2560 GeV^2, where the statistics for 1995 becomes unacceptably low.

The events counted in the table have been analysed with respect to the scaled momentum cross-sections and angular correlations in the current region of the Breit frame. The results are presented in the two chapters to follow.

Bin	x		Q^2 GeV^2		Events		
	min	max	min	max	1995	1996	1997
1	0.0006	0.0012	10	20	47068		
2	0.0012	0.0024	10	20	37675		
3	0.0012	0.0024	20	40	28023		
4	0.0024	0.0100	20	40	32397		
5	0.0012	0.0024	40	80	14424		
6	0.0024	0.0100	40	80	26889		
7	0.0024	0.0100	80	160	13805	22346	59323
8	0.0100	0.0500	80	160	7692	12071	32101
9	0.0024	0.0100	160	320	4406	6468	17102
10	0.0100	0.0500	160	320	5883	9555	25282
11	0.0100	0.0500	320	640	2511	4251	11433
12	0.0100	0.0500	640	1280	1009	1634	4463
13	0.0250	0.1500	1280	2560	466	720	2086
14	0.0500	0.2500	2560	5120	153	239	643
15	0.0600	0.4000	5120	10240	53	79	214
Total					222464	57377	152689
Luminosity (pb^{-1})					6.3	10.5	26.6

Table 6.2: Statistics of the data in kinematic bins.

Chapter 7

Fragmentation Functions

Measurements of the scaled momentum spectra are presented in this chapter as an approach toward fragmentation functions: a general and possibly universal description of the hadronization process. The scaled momentum, x_P, is evaluated by normalizing the magnitude of the particle momentum as measured in the current region of the Breit frame to the maximum kinematically allowed value of $Q/2$ [1]:

$$x_P = \frac{p^B_{tr}}{Q/2} = \frac{2p^B_{tr}}{Q}. \tag{7.1}$$

The procedure described in this chapter takes as the starting point charged hadrons reconstructed in the Breit frame and uses the resolution in x_P and $\ln(1/x_P)$ estimated from MC studies to justify the binning adopted in this analysis. Uncorrected data are then confronted with generated samples on the simulated level seeking support for the correction procedure. Correction factors (CF) are calculated and verified to be well behaved so that they are in turn applied to the data. An evaluation of systematic effects follows. The results obtained are briefly summarized; a longer discussion accompanied by an extraction of phenomenological QCD-motivated parameters is postponed until chapter 9.

[1] The limit neglects QCD radiation which, however, tends to decrease the total deposits in the current hemisphere.

7.1 Resolution and Binning

The measurement of normalized cross-sections for the scaled momentum x_P and $\ln(1/x_P)$ was done at ZEUS based on 1994 data. Here the resolutions on these variables are verified to be consistent with what was found in [66].

Figure 7.1 presents the resolution in x_P as derived from the joint MC samples generated for 1995 (bins with $Q^2 < 160$ GeV^2) and 1997 (higher Q^2). The difference $x_P^{rec} - x_P^{gen}$ is plotted versus the generated value and the assignment to the analysis bins is based on the true kinematics. Points are plotted at mean values of the shift in each bin and the bars indicate the respective spread (RMS). It is seen that the resolution improves with increasing Q^2 and typically remains within the bin sizes marked by the horizontal dashed lines. The bin sizes are not the same: this will be justified below by a uniform resolution in the logarithmic distribution. No significant systematic shifts are observed.

Similarly Fig. 7.2 shows the resolutions in the variable $\ln(1/x_P)$. The advantage of this variable is that it expands the soft region of the particle momentum spectrum. Unlike the x_P case above there is a systematic bias in the lowest Q^2 bins but it fades away and the distribution flattens out as higher scales are reached. [2] Apart from the hardest region of the spectrum there is a tendency for an improvement of the resolution toward higher Q^2. The persistently poor reconstruction of energetic particles is due to the limitation of the CTD to measure accurately high momenta. The horizontal dashed lines indicate the size of the binning taken to be 0.25; it accounts for the spreads and the shifts in most of the kinematic region. The observed uniform resolution in $\ln(1/x_P)$ calls for a non-equal binning for the simple x_P variable. The actual limits as in [65] are accepted: 0.02–0.05–0.10–0.20–0.30–0.40–0.50–0.70–1.00. An additional bin, 0.00–0.02, is to be used only for integrated distributions due to widely varying correction factors as discussed in section 7.3.

[2] The lowest y bin, $0.01 < x < 0.05$, $80 < Q^2 < 160$ GeV^2, will be discussed in the following.

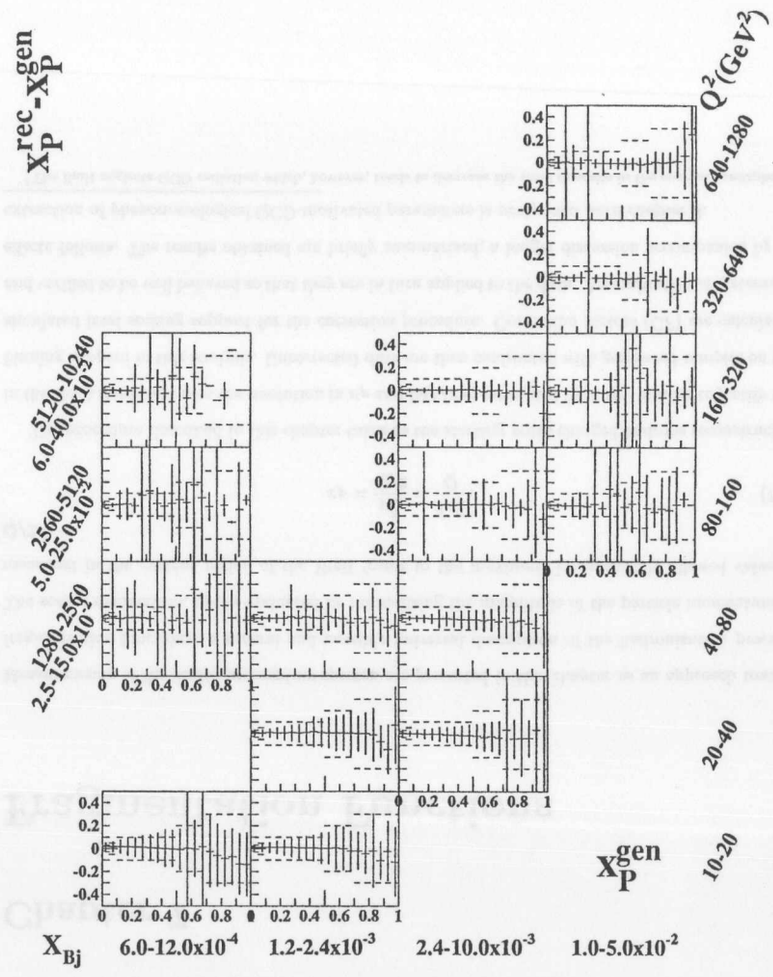

Figure 7.1: Resolution of x_P as a function of x_P^{gen} in bins defined by the true kinematics.

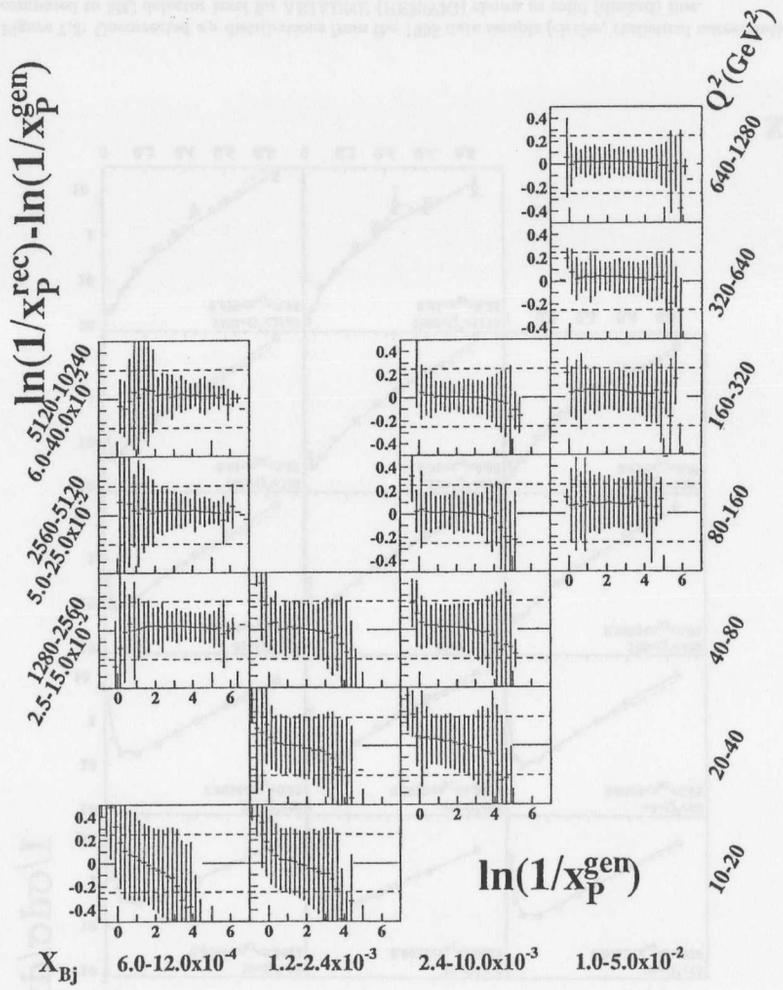

Figure 7.2: Resolution of $\ln(1/x_P)$ as a function of $\ln(1/x_P^{gen})$ in bins defined by the true kinematics.

7.2 Uncorrected Data

The distributions to be studied are experimentally evaluated as:

$$\frac{1}{\sigma_{tot}}\frac{d\sigma}{dx_P} = \frac{1}{N}\frac{dn_{ch}}{\Delta x_P} \tag{7.2}$$

and

$$\frac{1}{\sigma_{tot}}\frac{d\sigma}{d\ln(1/x_P)} = \frac{1}{N}\frac{dn_{ch}}{\Delta \ln(1/x_P)}. \tag{7.3}$$

They are presented in Figs. 7.3 and 7.4, respectively. The uncorrected data with the statistical uncertainties are compared to two MC models, ARIADNE and HERWIG, as calculated at the detector level. The statistical fluctuations of the MC curves are of the same order of magnitude as for the data points. All the quantities pertain to the 1995 sample. In the case of the x_P distributions a logarithmic scale is used to show the variation over several orders of magnitude. ARIADNE is noted to give an overall good description of the data with some discrepancies at high Q^2. HERWIG, while underestimating the x_P dependence, is still in reasonable agreement with the data.

For the uncorrected $\ln(1/x_P)$ distributions in Fig. 7.4 the bell-shaped curve is correctly reproduced by both MC samples. Statistical uncertainties on the data in most of the bins are too small to be seen. As Q^2 increases so do the magnitude and the abcissa of the peak. ARIADNE exhibits a tendency to undershoot the height of the distribution at low Q^2 but it is compensated for by a slightly faster evolution so that the discrepancy vanishes at the top Q^2 bins. A reasonable description is observed for HERWIG except for fluctuations at the highest Q^2 bins due to poor statistics.

With the aim of merging at the raw level, the data samples collected in 1996 and 1997 are compared in Figs. 7.5 and 7.6. Within the statistical uncertainties shown the ratios of the distributions are consistent with unity suggestive of a compatibility between the two subsamples. Hence, they are merged and corrected jointly.

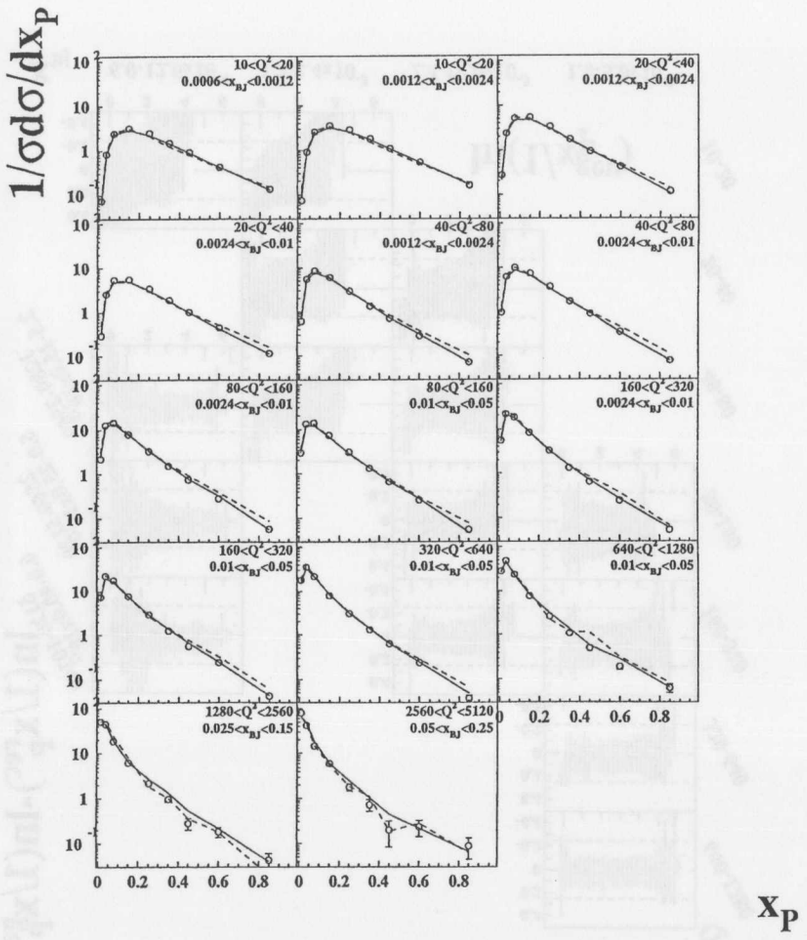

Figure 7.3: Uncorrected x_P distributions from the 1995 data sample (circles, statistical uncertainties) compared to MC detector level for ARIADNE (HERWIG) shown as solid (dashed) line.

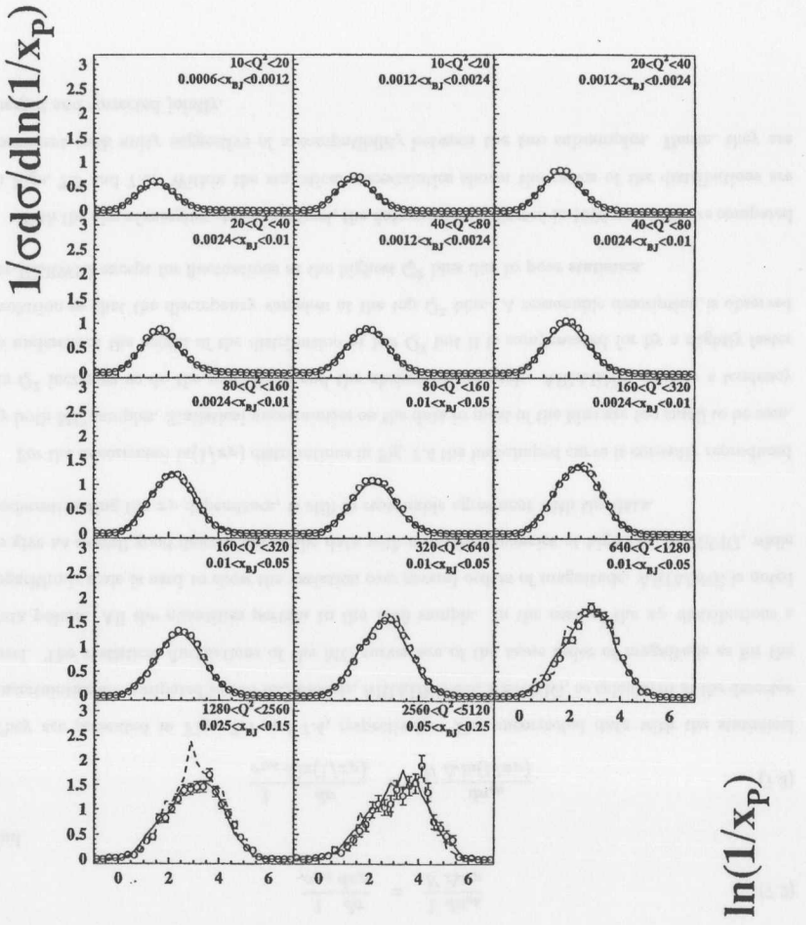

Figure 7.4: Uncorrected $\ln(1/x_P)$ distributions from the 1995 data sample (circles, statistical uncertainties) compared to MC detector level for ARIADNE (HERWIG) shown as solid (dashed) line.

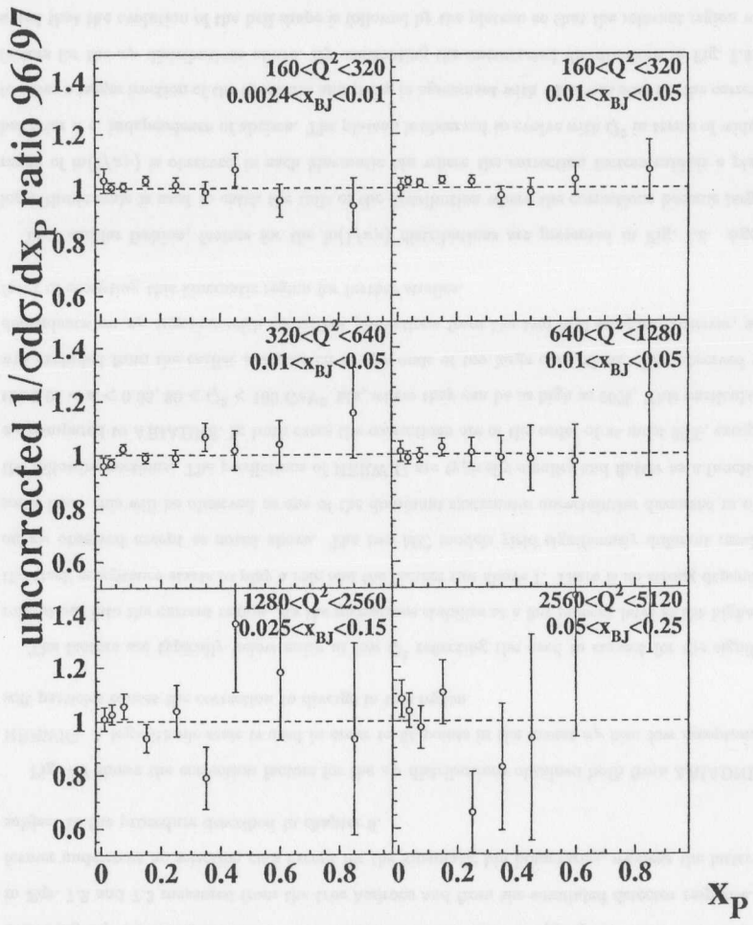

Figure 7.5: Ratio of the uncorrected x_P distributions from the 1996 and 1997 data samples.

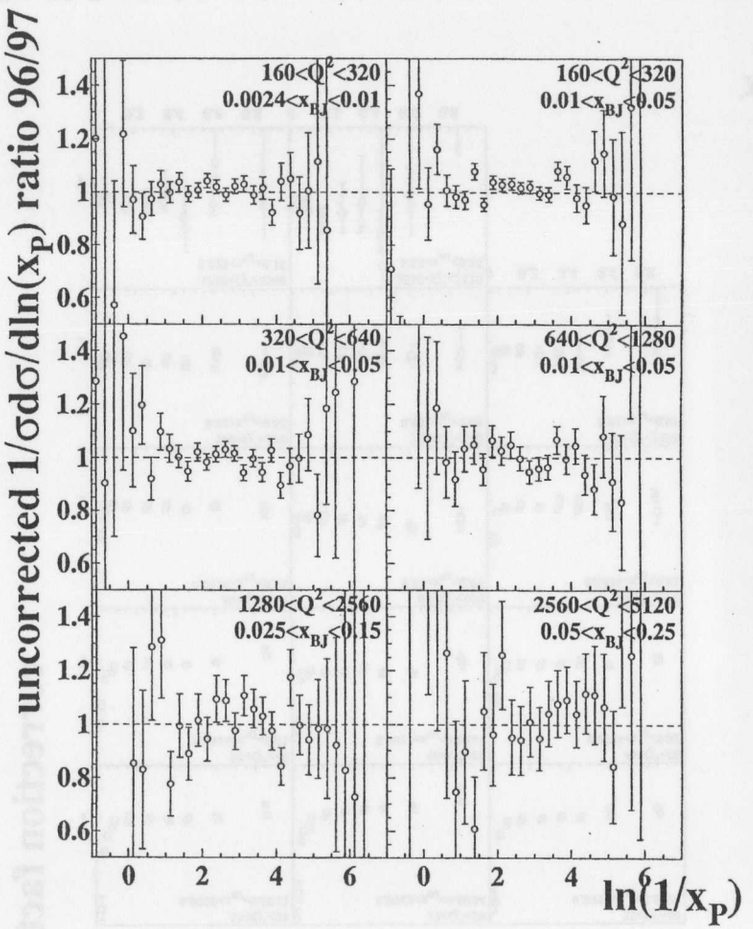

Figure 7.6: Ratio of the uncorrected $\ln(1/x_P)$ distributions from the 1996 and 1997 data samples.

7.3 Correction Factors

Following Eq. 5.5 correction factors are calculated as the ratio of appropriate distributions defined in Eqs. 7.2 and 7.3 measured from the true hadrons and from the simulated detector response. The former underwent no selection cuts except for the kinematic bin boundaries, whereas the latter were subject to the procedure described in chapter 6.

Fig. 7.7 shows the correction factors for the x_P distributions obtained both from ARIADNE and HERWIG. A logarithmic scale is used in order to fit points in the lowest x_P bin: low acceptance for soft particles causes the correction to diverge in this region.

The factors are typically below unity at low Q^2 reflecting the need to correct for the significant migrations into the current region. As the migrations stabilize at a few percent level at the higher Q^2, the track acceptance starts to play a role and the factors rise above 1. There is no strong dependence on x_P observed except as noted above. The two MC models yield significantly different results in some bins: this will be observed as one of the dominant systematic uncertainties discussed in one of the following sections. The predictions of HERWIG are typically smaller and flatter as a function of x_P compared to ARIADNE. In both cases the corrections are of the order of at most 20%, except for the $0.01 < x < 0.05$, $80 < Q^2 < 160$ GeV^2 bin, where they can be as high as 50%. This particular bin was excluded from the earlier analyses on the grounds of too large corrections. The observed weak dependence on x_P together with consistent predictions from the two MC samples, however, are in favor of accepting this kinematic region for further studies.

In a similar fashion, factors for the $\ln(1/x_P)$ distributions are presented in Fig. 7.8. Again a logarithmic scale is used to catch the tails of the distribution where the corrections become large. A range of $\ln(1/x_P)$ is observed in each kinematic bin where the correction factors exhibit a plateau behavior *i. e.* independence of abcissa. The plateau is observed to evolve with Q^2 in terms of widening to cover a larger fraction of the spectrum and rising in agreement with what was seen for the correction factors for the x_P distributions above. By confronting the uncorrected distributions in Fig. 7.4 it is noted that the evolution of the bell shape is followed by the plateau so that the relevant region where the distribution significantly deviates from zero matches the plateau where the correction factors are

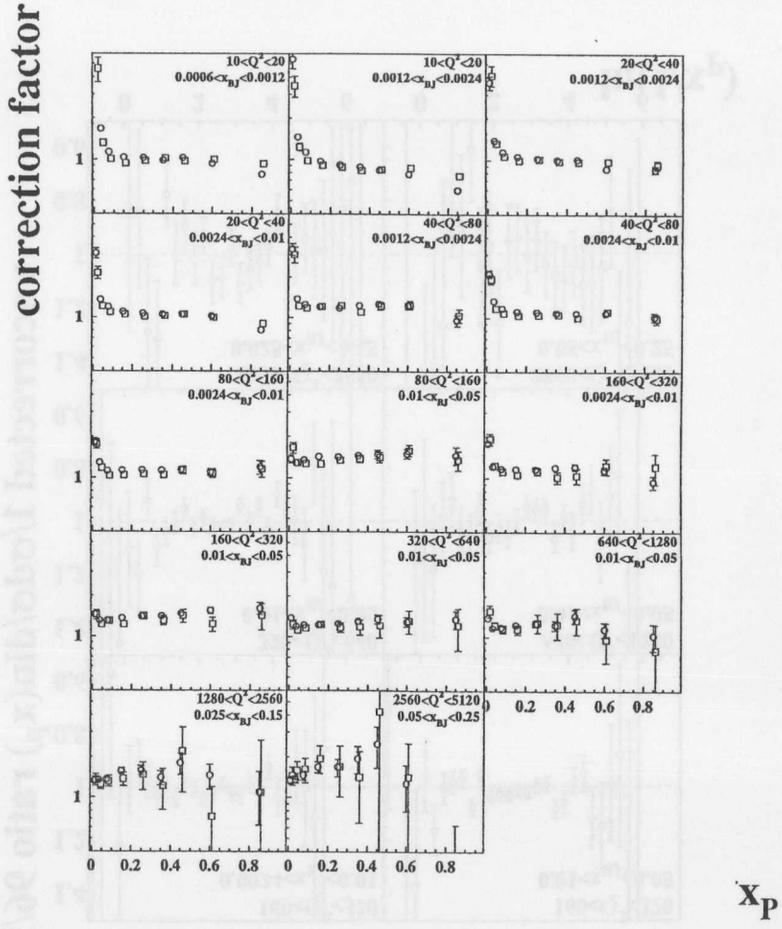

Figure 7.7: Correction factors to the normalied x_P distributions evaluated from the ARIADNE (circles) and HERWIG (squares) for the 1995 detector configuration. Statistical uncertainties are shown.

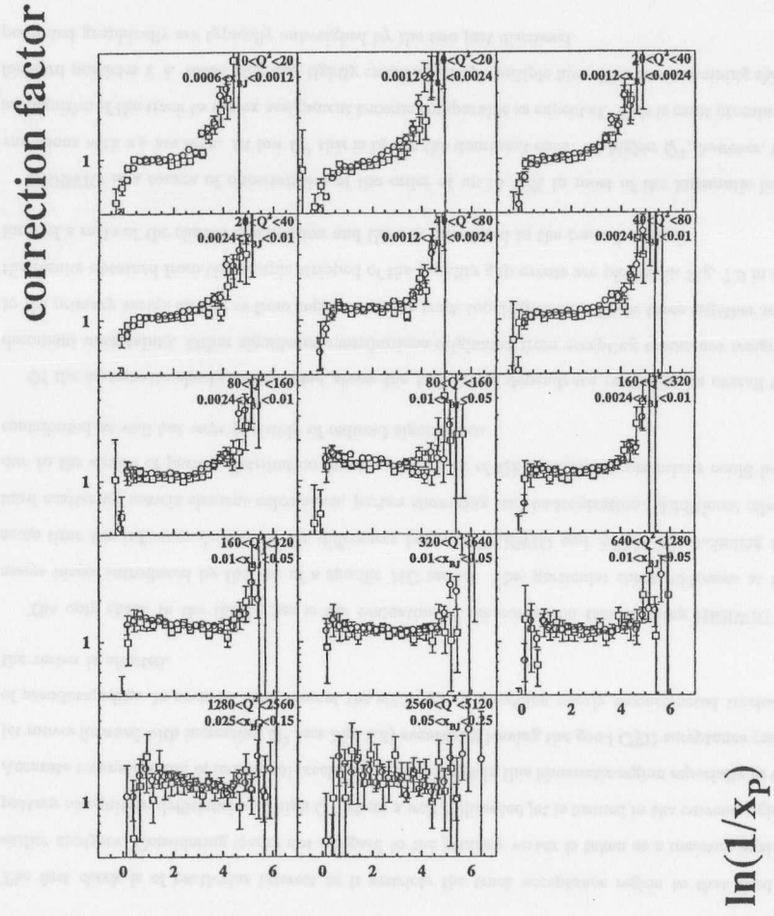

Figure 7.8: Correction factors to the normalized $\ln(1/x_P)$ distributions evaluated from the ARIADNE (circles) and HERWIG (squares) for the 1995 detector configuration. Statistical uncertainties are shown.

close to unity. The divergence in the soft particle region is especially prominent at low Q^2. ARIADNE typically yields larger and less flat corrections, a fact seen already in the x_P case above.

7.4 Corrected Distributions and Systematic Uncertainties

Having evaluated the correction factors in the previous section and having verified their good behavior, it is now straightforward to apply them to the raw data in quest for the unfolded results. Before doing that, however, a study of systematic uncertainties arising from various sources is summarized.

Logically, uncertainties beyond those due to the limited statistics can be subdivided into three classes. Firstly, the event selection cuts could have introduced an uncorrectable bias to the studied distributions. This issue is addressed by varying the thresholds of the respective criteria. In particular, the following checks have been performed one at a time(the standard cuts are given in parentheses):

- (35) $40 < E - p_Z < 60$ (65) GeV
- $y_{JB} > 0.05$ (0.04)
- $y_e < 0.80$ (0.95)
- $E'_e > 10.5$ (10.0) GeV
- $\max(|x_{RCAL}|, |y_{RCAL}|) > 17$ (16) cm
- $\eta_{max} > 1.5$ $(-\infty)$

The last item on the list is to clean up the sample from the contamination of large rapidity gap events not appropriately accounted for in the MC samples used for corrections, see Fig. 5.5.

The second class is related to similar effects but for the tracking-related requirements. The list of checks in this case includes:

- $|\eta^{tr}| < 1.75$ (2.0) and $p_T^{tr} > 0.2$ (0.15) GeV
- allow non-vertex tracks *but* require the above tighter kinematic cuts

- drop the demand on the number of CTD superlayers traversed
- (−50) $-40 < z_{vtx} < 40$ cm

The first check is of particular interest as it restricts the track acceptance region to that used in earlier analyses. Considering tracks not assigned to the primary vertex is taken as a measure against pattern recognition defficiencies at high Q^2 where a well collimated jet is formed in the current region. Accurate reconstruction of individual tracks becomes difficult in this kinematic region especially as the jet moves forward with increasing Q^2 (see Fig. 5.8) eventually leaving the good CTD acceptance range of pseudorapidity. In such an environment the efficiency of attaching poorly reconstructed tracks to the vertex is affected.

The only check in the third class is the evaluation of the correction factors using HERWIG to assess biases introduced by the use of a specific MC model. This particular check addresses at the same time the influence due to all the differences between HERWIG and ARIADNE including the hard scattering matrix element calculation, parton showering and hadronization. Additional effects due to the choice of parton distribution functions and lack of QED radiative corrections could have contributed as well but were probably of reduced significance.

Of the systematic checks enumerated above the MC model dependence proved to be overall the dominant uncertainty. Other significant contributions originated from accepting tracks not assigned to the primary vertex as well as from suppressing the track topological cut. These three together with the results obtained from the sample stripped of the rapidity gap events are plotted in Fig. 7.9 in the form of a ratio of the shifted distribution and the one calculated in the central analysis.

HERWIG is a source of uncertainties of the order of up to 20% in most of the kinematic bins; variations with x_P are seen. At low Q^2 this is by far the dominant shift. At higher Q^2, however, the ambiguities of the track to vertex assignment become comparable as expected. This is most prominent for hard particles *i. e.* those long and tightly constrained by multiple hits. The two remaining shifts presented graphically are typically outweighed by the two just discussed.

The checks omitted from Fig. 7.9 were examined as well and found to lead to non-negligible but

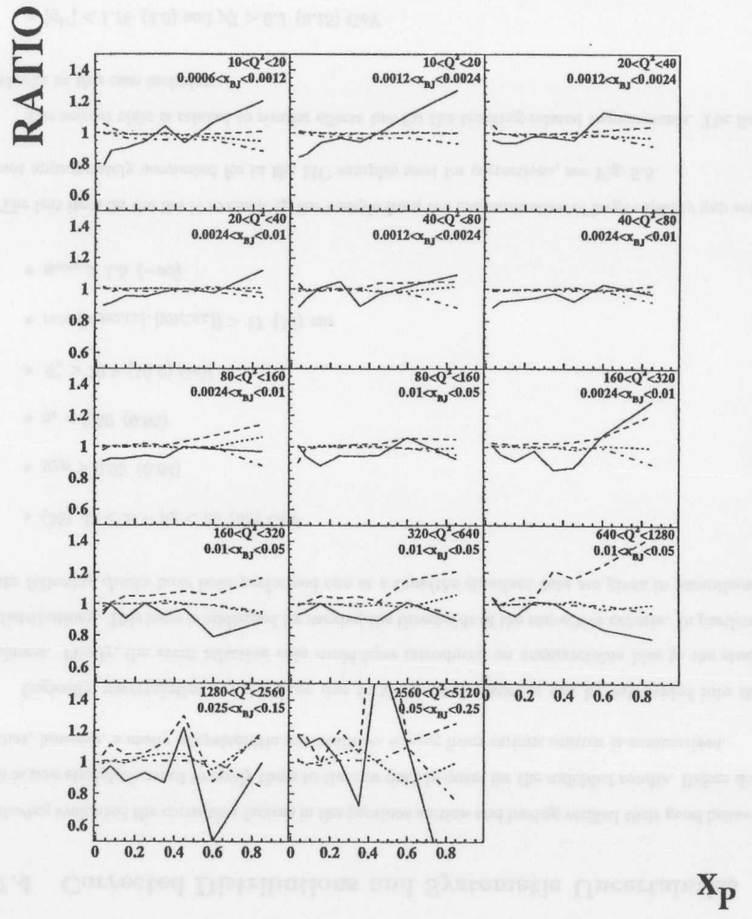

Figure 7.9: Dominant systematic shifts to the x_P distribution: HERWIG (solid), off-vertex tracks (dashed), track topology cut (dotted) and rejection of LRG events (dot-dashed).

small shifts in regions known to be affected by respective changes of cuts. In particular, the low y bin, observed to be problematic on the grounds of too large correction factor, suffers from a relatively high systematic uncertainty related to tightening the low y_{JB} cut. Still, it does not exceed the shift due to the model dependence.

All deviations from the central analysis were added in quadrature separately for shifts with positive and negative sign to produce (asymmetric) systematic uncertainties. Statistical and systematic contributions were combined in quadrature to obtain estimates of the total uncertainties.

Final results on the corrected x_P distributions measured from the 1995 data sample are presented in figure 7.10. Uncertainties, here and later on, are marked by bars (total) with small ticks (statistical). The distributions exhibit a clear evolution in that the peak present in low Q^2 bins moves towards the softer region and eventually cannot be observed any longer in the highest Q^2 region. The steepness of the fall-off from the peak towards the hard part of the spectrum is seen to increase, suggestive of the scaling violation. This will be discussed in more detail in chapter 9 in the context of QCD and MLLA calculations.

The data are compared to the hadron-level MC predictions of ARIADNE and HERWIG. The general features of the data, evolution of the peak and scaling violation, are seen in the MC simulations. ARIADNE gives a good account of the data whereas HERWIG underestimates the slope of the dependence on x_P.

In a similar manner the data are plotted in terms of $\ln(1/x_P)$ in Fig. 7.11. On a linear scale on the vertical axis the evolution of the distribution is more pronounced. The magnitude of the peak and its position are both observed to increase with Q^2. The former is related to the total multiplicity whereas the latter is characteristic of softening of the scaled momentum spectrum. Both topics are going to be quantitatively addressed in chapter 9. The overlaid predictions of MC models provide a reasonable description of the evolution. At low Q^2 ARIADNE undershoots the data but it is in good agreement in terms of the peak position. As the uncertainties increase towards the high Q^2 region the magnitude of the discrepancy decreases. HERWIG suffers from small random shifts around the measured distribution but still accounts for the main features of the data reasonably well.

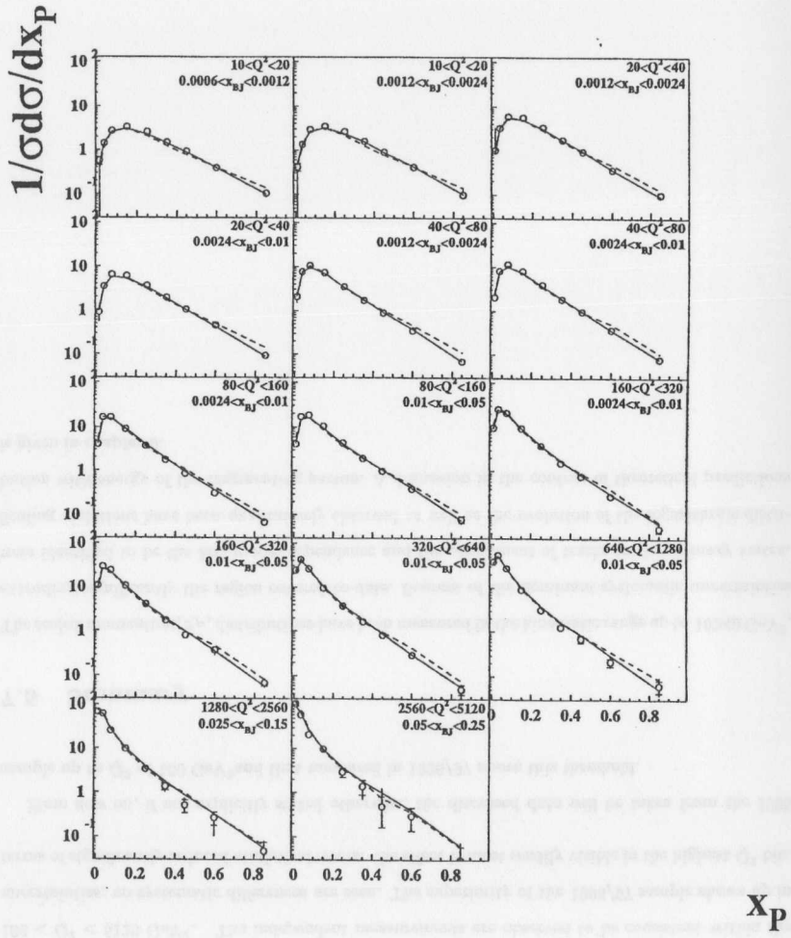

Figure 7.10: Corrected x_P distribution from the 1995 data sample (circles, statistical and total uncertainties) compared to ARIADNE (HERWIG) shown as solid (dashed) line.

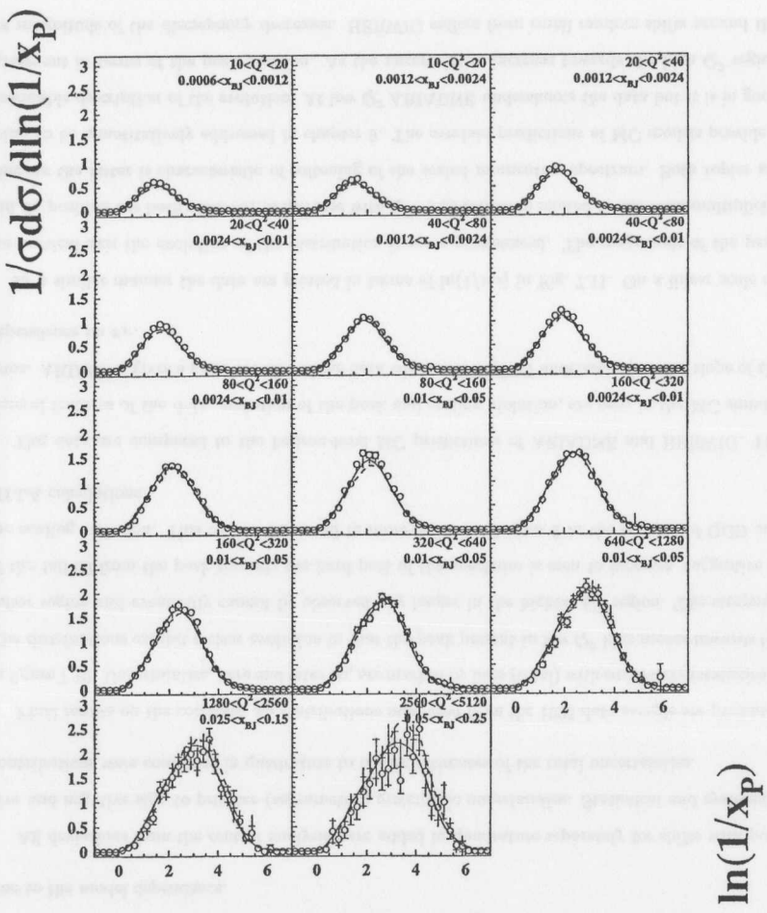

Figure 7.11: Corrected $\ln(1/x_P)$ distributions from the 1995 data sample (circles, statistical and total uncertainties) compared to ARIADNE (HERWIG) shown as solid (dashed) line.

Finally the samples of 1995 and 1996/97 are compared in the overlapping kinematic region. The corrected x_P and $\ln(1/x_P)$ distributions are given in Figs. 7.12 and 7.13, respectively, for the range $160 < Q^2 < 5120$ GeV^2. The independent measurements are observed to be consistent within the uncertainties; no systematic differences are seen. The superiority of the 1996/97 sample shows up in terms of significantly reduced statistical errors: the effect is most readily visible in the highest Q^2 bin.

From now on, if not explicitly stated otherwise, the discussed data will be taken from the 1995 sample up to Q^2 of 160 GeV^{2}and that measured in 1996/97 above this threshold.

7.5 Summary

The scaled momentum, x_P, distributions have been measured in the kinematic range up to 10240 GeV^2, extending significantly the region covered to-date. Sources of the dominant systematic uncertainties were identified to be the MC model dependence and the assignment of tracks to the primary vertex. Scaling violations have been qualitatively observed as well as the evolution of the logarithmic distribution with energy of the fragmenting parton. A discussion in the context of theoretical predictions is given in chapter 9.

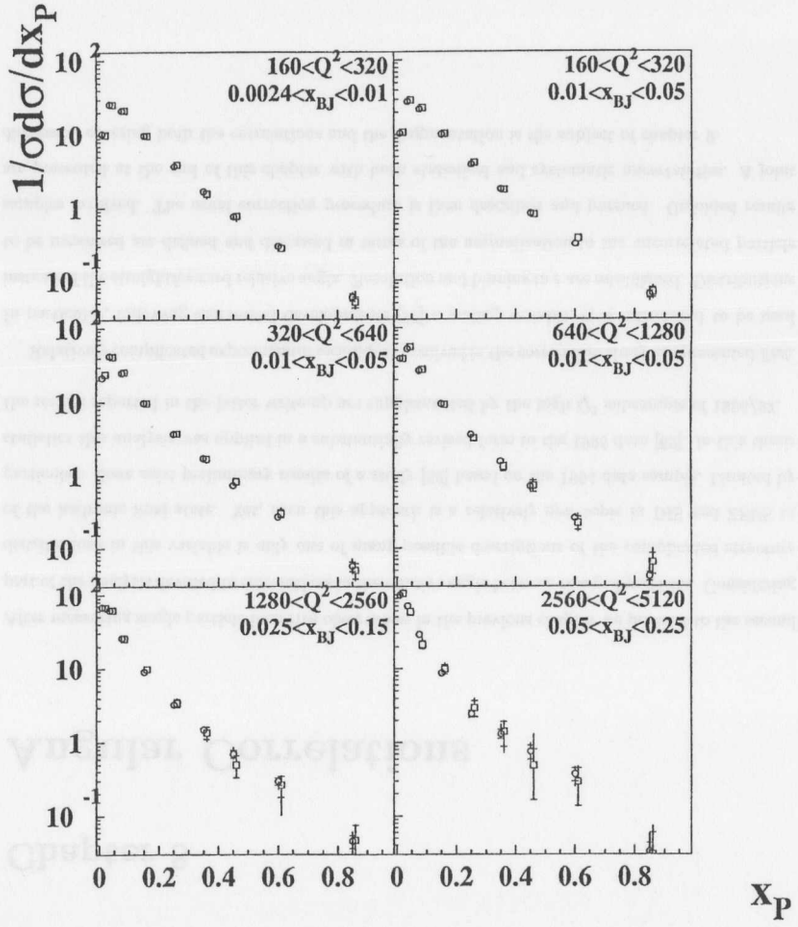

Figure 7.12: Corrected x_P distributions from the data samples of 1995 (squares) and 1996/97 (circles). Statistical and total uncertainties are shown.

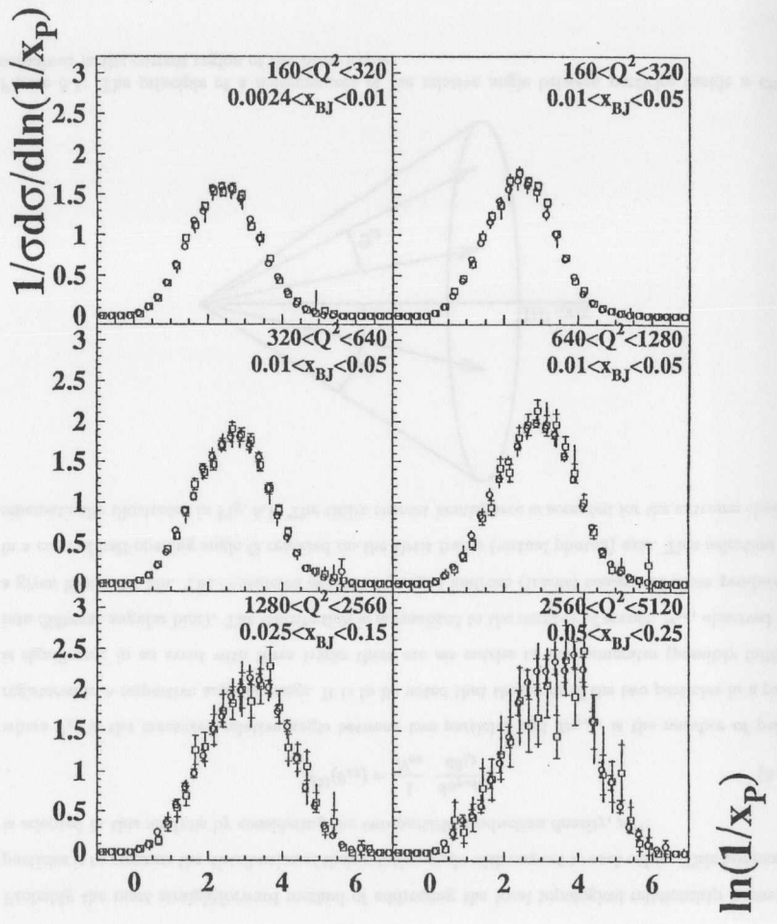

Figure 7.13: Corrected $\ln(1/x_P)$ distributions from the data samples of 1995 (squares) and 1996/97 (circles). Statistical and total uncertainties are shown.

Chapter 8

Angular Correlations

After measuring single particle inclusive observables in the previous chapter we proceed to the second part of the analysis devoted to correlations in the relative angle between charged particles. Considering distributions in this variable is only one of many possible descriptions of the complicated structure of the hadronic final state. Yet, even this approach is a relatively new topic in DIS and ZEUS in particular: there exist preliminary results of a study [69] based on the 1994 data sample. Limited by statistics this analysis was applied in a substantially revised form to the 1995 data [62]. In this thesis the results reported in the latter write-up are supplemented by the high Q^2 subsample of 1996/97.

Relatively complicated experimental techniques involved in the correlation study are presented first. In particular, following theoretical developments [70] a scaling variable, ε, is introduced to be used instead of the straightforward relative angle. Resolution and binning in ε are established. Distributions to be measured are defined and discussed in terms of the normalization to the uncorrelated particle samples involved. The usual correction procedure is then described and pursued. Unfolded results are presented at the end of this chapter with both statistical and systematic uncertainties. A joint discussion covering both the correlations and the fragmentation is the subject of chapter 9.

8.1 Measurement Techniques

8.1.1 Two-particle Production Density

Probably the most straightforward method of addressing the local topological relationship between particles is to measure the distribution of their relative angle with respect to each other. This approach is adopted in this analysis by considering the two-particle production density, $\rho^{(2)}$:

$$\rho^{(2)}(\theta_{12}) = \frac{1}{N_{ev}} \frac{dn_{pair}}{d\theta_{12}}, \tag{8.1}$$

where θ_{12} is the measured relative angle between two particles and dn_{pair} is the number of pairs registered in a respective angular range. It is to be noted that the order of the two particles in a pair is significant: in an event with three tracks there are *six* entries in the numerator (possibly falling into different angular bins). The distribution is normalized to the number of events, N_{ev}, observed in a given kinematic bin. The considered subset of charged hadrons (tracks) consists of those produced in a cone of half-opening angle Θ centered on the Breit frame (virtual photon) axis. This selection is schematically illustrated in Fig. 8.1. The entire current hemisphere is accepted for the extreme choice

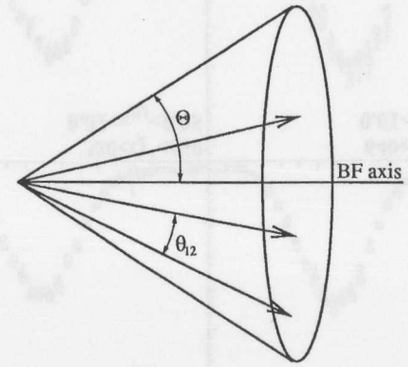

Figure 8.1: The principle of a measurement of the relative angle between particles inside a cone contained in the current region of the Breit frame.

of $\Theta = 90°$. The introduction of a new angular scale to the problem is motivated by a wish to study

the evolution of the correlations with the size of the cone. At present, theoretical predictions are calculated in a small angle approximation but the experimental statistics in the ZEUS data impose an effective lower limit of 45° on the currently accessible cone size. For this reason the analysis described below is repeated for three values of Θ: 90°, 60° and 45°.

8.1.2 Scaling Variable

Although the relative angle θ_{12} seems to be a natural variable to study correlations, its simplicity is outweighed by the theoretical advantages of using the scaling variable ε defined as [70]:

$$\varepsilon = \frac{\ln(\Theta/\theta_{12})}{\ln(P\Theta/\Lambda)}. \tag{8.2}$$

In addition to the angles described earlier, the above formula folds together the parent parton momentum, P, and the newly introduced effective scale parameter Λ. In this experiment P is identified with the current jet energy as calculated in the QPM approximation: $P = \frac{Q}{2}$. The domain for the ε variation is between 0 and 1 corresponding to *large* and *small* relative angles, θ_{12}, respectively. Both limits stem from theoretical uncertainties: the lower is due to neglecting recoil, an effect significant at large angles, and the upper is due to the breakdown of the perturbative approach at small angles *i. e.* small momentum transfers. Following [71] for the central analysis, a value for Λ of 0.15 GeV was assumed.

There is a threefold justification for using the more complicated variable ε:

- the small relative angle region is expanded allowing a more detailed insight to where the QCD effects are expected to be most prominent due to a running α_S;
- a theoretical prediction exists claiming correlations to be approximately independent of other variables: P, Θ and θ_{12} other than through the combination given in Eq. 8.2;
- a residual dependence leading to scaling violation in the correlation functions may be studied directly thanks to defining a common range for any kinematic bin and cone size.

In addition, ε emphasizes the logarithmic nature of the expressions present in the QCD description of gluon radiation.

Actually, taking into consideration the relatively large cones to be analysed, the suggestion [71] was followed to modify the experimental definition of the scaling variable. After the change it reads:

$$\varepsilon = \frac{\ln(\Theta/\theta_{12})}{\ln(P\sin\Theta/\Lambda)}. \tag{8.3}$$

The original definition in Eq. 8.2 was in fact a low angle approximation of the maximum transverse momentum reachable in a hadronic jet of opening angle Θ and energy P.

8.1.3 Resolution and Binning

The adopted scaling variable was then studied using MC samples to establish the experimental resolution. The same criteria for a generated hadron to match a reconstructed track as described in section 7.1 were applied. The averages and RMS spreads of the $\varepsilon^{rec} - \varepsilon^{gen}$ distribution are presented in Fig. 8.2. The chosen uniform binning of size 0.06 is indicated by the horizontal dashed lines. It is seen that the quality of reconstruction is relatively poor in the lowest Q^2 region below 40 GeV^2: both significant systematic bias and resolution exceeding the proposed bin size are visible. The situation improves with growing Q^2 and starting with bins between 40 and 80 GeV^2 a wide region of $\varepsilon \lesssim 0.5$ becomes compatible with the binning. Apart from the lowest y bin identified as problematic already in the context of the scaled momentum analysis in chapter 7, there is no additional indication of systematic shifts. Since, as will be seen shortly, the measurement of the correlation is statistically limited to $\varepsilon \lesssim 0.7$ and the high Q^2 region is of the main interest to this analysis the respective resolutions were taken to define the binning with the result as given above.

8.1.4 Normalization

There are two forms of presentation of the angular correlation data to be employed in this analysis, both motivated by theoretical predictions as discussed in chapter 9. They differ by the normalization applied to the distributions. Definitions of a couple of common parameters are in place:

$$\beta = \sqrt{12(\frac{11}{3}N_c - \frac{2}{3}n_f)^{-1}} \tag{8.4}$$

and

$$b = 2\beta\sqrt{\ln(P\sin\Theta/\Lambda)}. \tag{8.5}$$

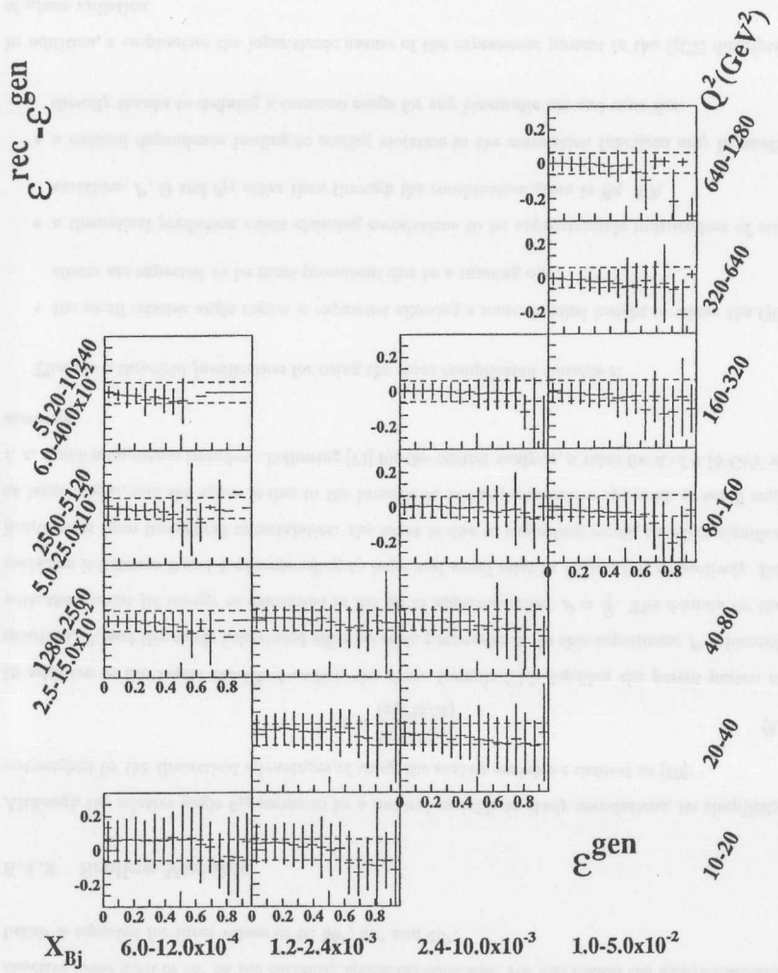

Figure 8.2: Resolution of the scaling variable ε as a function of ε^{gen} in bins defined by the true kinematics.

The first, β, absorbs the parametric dependence on the number of colors, N_c, and active flavors, n_f, taken here to be 3 and 5, respectively. It appears naturally in perturbative QCD calculations. The second, b, is to take care, when applied to the data, of the difference in phase space between the kinematic bins. The phase space factor involves the cone size, Θ, controlling the amount of gluon radiation accepted in the measurement.

Relative Angular Distribution

The angular distribution is normalized to the square of the mean total charged multiplicity counted in the studied cone, Θ:

$$\hat{r}_2^{raw}(\theta_{12}) = \frac{\rho^{(2)}(\theta_{12})}{< n >^2_\Theta}. \tag{8.6}$$

Being a differential distribution $\hat{r}_2$ needs to be multiplied by the respective Jacobian upon transformation to the scaling variable ε:

$$\hat{r}_2^{raw}(\varepsilon) = \frac{d\theta_{12}}{d\varepsilon}\hat{r}_2^{raw}(\theta_{12}) = \theta_{12}\ln(P\sin\Theta/\Lambda)\hat{r}_2^{raw}(\theta_{12}). \tag{8.7}$$

Finally, global rescaling involving the β and b factors defined above is performed to obtain:

$$\hat{r}_2(\varepsilon) = -\frac{\ln(\hat{r}_2^{raw}(\varepsilon)/b)}{2\sqrt{\ln(P\sin\Theta/\Lambda)}}. \tag{8.8}$$

The rather complicated form is driven entirely by theoretical considerations to be discussed in chapter 9; here it is only noted that the most remarkable prediction applicable to the rescaled distribution is the independence of any other quantity except ε. Note also the negative sign leading to an increasing $\hat{r}_2$ for a decreasing two-particle density $\rho^{(2)}$.

Correlation Function

There is at least a comparable amount of complication involved in the definition of the true angular correlation function r_2 to be described below. It assumes an uncorrelated angular distribution $\rho^{(2)}_{mix}$ as the reference. In this analysis such a distribution is modelled in an event-mixing technique. It amounts to considering particles originating from real but different events and treating them as if they had come from a single event. In analogy to Eq. 8.1 the mixed (uncorrelated) density is defined:

$$\rho^{(2)}_{mix}(\theta_{12}) = \frac{1}{N^2_{ev}}\frac{dn^{mix}_{pair}}{d\theta_{12}}. \tag{8.9}$$

Here, the normalization is to the *square* of the event multiplicity to maintain consistency; it must be as given above for every track now can be paired with not only members of the same event but with tracks from any other event in the sample. Hence, an additional factor, N_{ev}, in the denominator. In practice, the mixed distribution is probed using only a subsample of all possible pairs: in this analysis a pool of 50 tracks was used and refreshed using events as coming in according to a (random) sequence on the tape storage.

With the help of the unbiased density calculated as discussed above the correlation function is defined as the ratio of two differential distributions

$$r_2^{raw}(\theta_{12}) = \frac{\rho^{(2)}(\theta_{12})}{\rho^{(2)}_{mix}(\theta_{12})} \tag{8.10}$$

and therefore there is no numerical difference between the $r_2(\theta_{12})$ and $r_2(\varepsilon)$, *i. e.* no transformation involving a Jacobian is needed. Following that the final rescaling is applied according to

$$r_2(\varepsilon) = \frac{\ln r_2^{raw}(\varepsilon)}{\sqrt{\ln(P \sin\Theta/\Lambda)}}. \tag{8.11}$$

Here again, the actual rescaling is motivated by the theoretical expectation of a scaling behavior for an observable defined as above.

8.2 Uncorrected Data

Uncorrected rescaled results obtained from the 1995 data are compared to the reconstructed-level ARIADNE and HERWIG samples: Figs. 8.3 and 8.4 present the relative angular distribution, $\hat{r}_2$, and the correlation function, r_2, respectively, as functions of the scaling variable ε. Statistical uncertainties are shown for the data; they are of comparable magnitude for the MC curves.

The first distribution exhibits a steady rise with ε: this translates into a decrease of the probability for particle production with relative angle θ_{12} (Eq. 8.3). Qualitatively very weak evolution with the energy scale Q^2 is observed. A good description is noted in case of either MC model over the entire kinematic range.

The second distribution, the correlation function r_2 depicted in Fig. 8.4, proves to be more limited statistically. The low Q^2 bins exhibit a relatively weak dependence on ε; it is enhanced in the higher

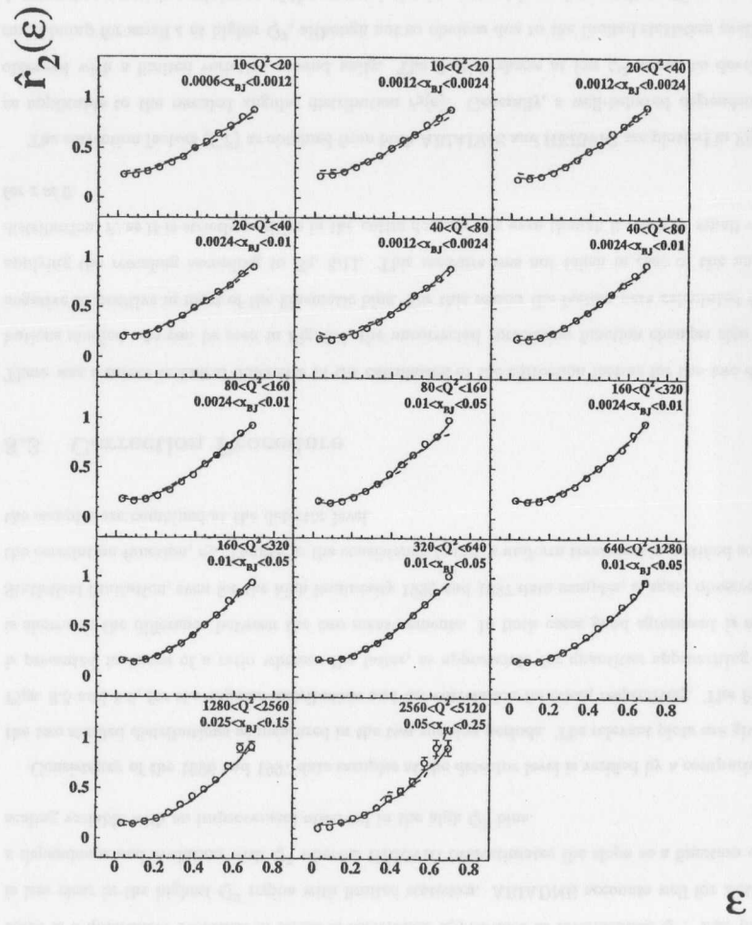

Figure 8.3: Uncorrected rescaled $\hat{r}_2(\varepsilon)$ distributions for the 1995 data sample (circles, statistical errors), ARIADNE (solid lines) and HERWIG (dashed lines). The cone size is $\Theta = 90°$.

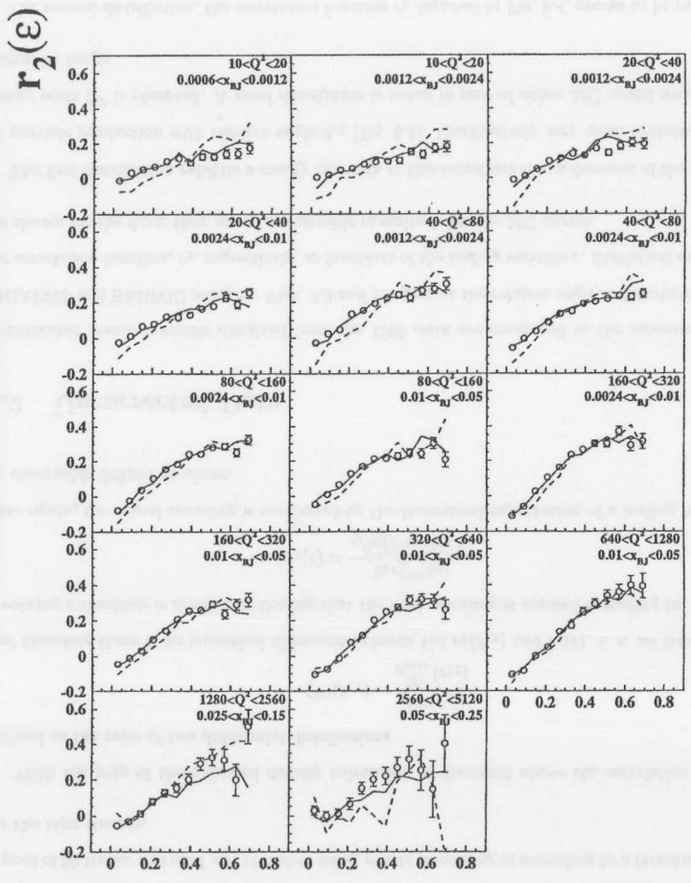

Figure 8.4: Uncorrected rescaled $r_2(\varepsilon)$ distributions for the 1995 data sample (circles, statistical errors), ARIADNE (solid lines) and HERWIG (dashed lines). The cone size is $\Theta = 90°$.

Q^2 region. For this observable any deviation from zero is a signature of interparticle correlations. Although not conclusive before applying corrections for detector effects a statement can be made that there is a qualitative evolution in terms of correlation appearance at intermediate Q^2. The picture is less clear in the highest Q^2 region with limited statistics. ARIADNE accounts well for both the ε dependence and evolution with Q^2 whereas HERWIG overestimates the slope as a function of the scaling variable with an improvement observed in the high Q^2 bins.

Consistency of the 1996 and 1997 data samples at the detector level is verified by a comparison of the two studied distributions as measured in the two running periods. The relevant plots are given in Figs. 8.5 and 8.6, for the angular distribution and the correlation function, respectively. The former is presented in terms of a ratio whereas the latter, as appropriate for quantities approaching zero, is shown as the difference between the two measurements. In both cases good agreement is noted. Statistical limitation, even for the high luminosity 1996 and 1997 data samples, is again observed for the correlation function, r_2. Thanks to the consistency found, a uniform treatment is justified so that the samples are combined at the detector level.

8.3 Correction Procedure

There was a minor technical difference in the calculation of the correction factors for the two distributions studied. As can be seen in Fig. 8.4, the uncorrected correlation function changes sign from negative to positive in most of the kinematic bins. For this reason the factors were calculated *before* applying the rescaling according to Eq. 8.11. This measure was not taken in case of the angular distribution, $\hat{r}$, as it is strictly positive in the entire domain of ε even though it assumes small values for $\varepsilon \approx 0$.

The correction factors (CF) as obtained from both ARIADNE and HERWIG are plotted in Fig. 8.7 as applicable to the rescaled angular distribution $\hat{r}_2(\varepsilon)$. Generally, a well-behaved dependence is observed with a limited variation around unity. The flattish shape at low Q^2 seems to develop a small bump for small ε at higher Q^2, although not so obvious due to the limited statistics available. A somewhat surprising minimum of the uncertainties is assumed in an intermediate Q^2 range. This

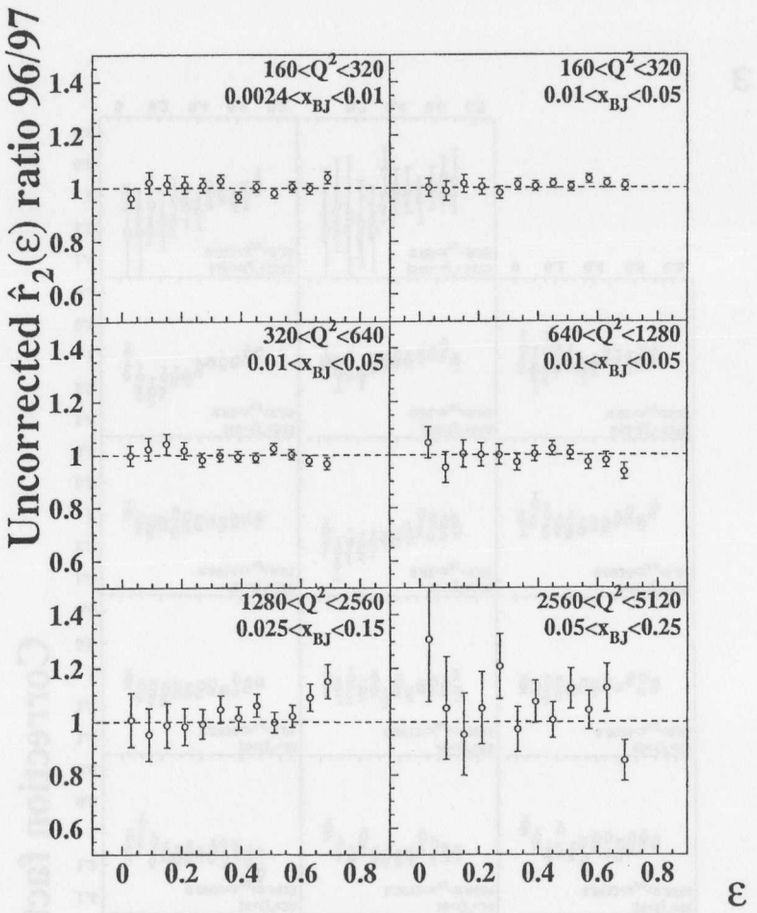

Figure 8.5: Ratio of the uncorrected $\hat{r}_2(\varepsilon)$ distributions from the 1996 and 1997 data samples. Statistical uncertainties are shown.

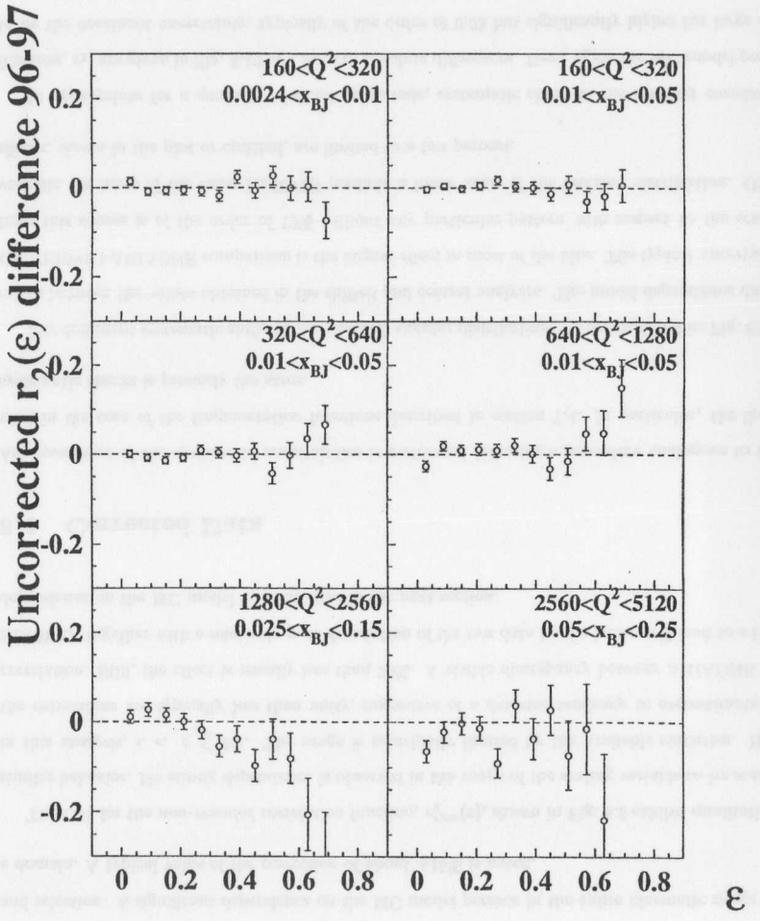

Figure 8.6: Difference of the uncorrected $r_2(\varepsilon)$ distributions from the 1996 and 1997 data samples. Statistical uncertainties are shown.

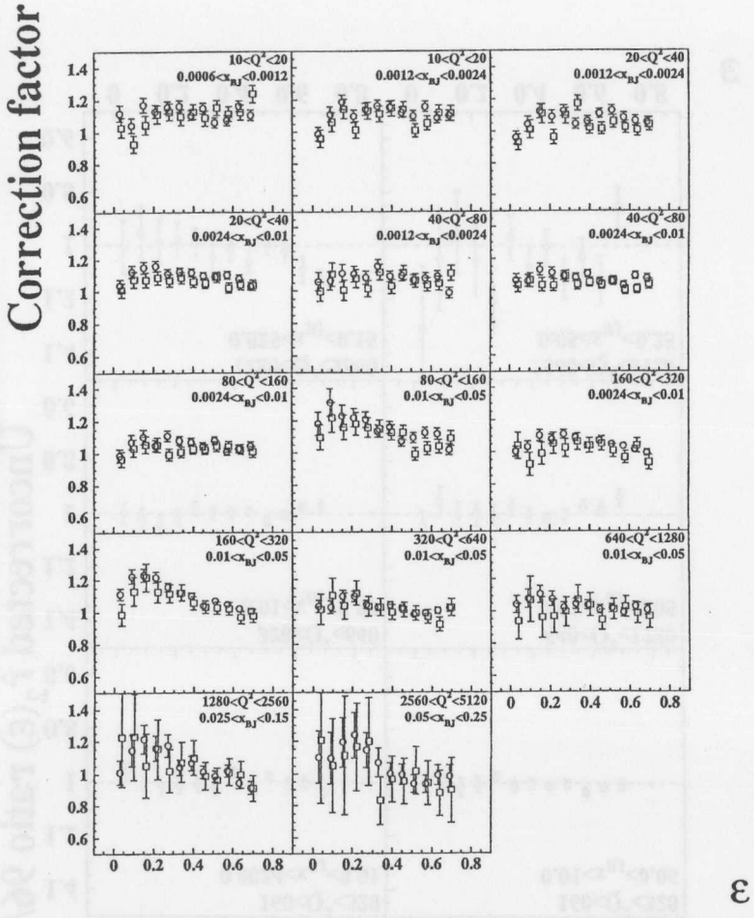

Figure 8.7: Correction factors to the rescaled $\hat{r}_2(\varepsilon)$ distributions evaluated for the 1995 detector configuration using ARIADNE (circles) and HERWIG (squares).

is due to the requirement of having at least a pair of particles in the current hemisphere (or even a narrower conical subregion) which is not likely to be satisfied by low multiplicity events at lower Q^2. The problematic kinematic bin of low y is again observed to be badly affected in the reconstruction and selection. A significant dependence on the MC model persists in the entire kinematic range and ε domain. A typical value of the correction of about +15% is noted.

The CFs for the *non-rescaled* correlation function, $r_2^{raw}(\varepsilon)$, shown in Fig. 8.8 exhibit qualitatively similar behavior. No strong dependence is observed in the range of the scaling variable to be studied in this analysis, *i. e.* $\varepsilon \lesssim 0.7$. The range is principally limited by the available statistics. Here, the corrections are typically less than unity, suggestive of a detector tendency to overestimate the correlation. Still, the effect is usually less than 20%. A visible discrepancy between ARIADNE and HERWIG, together with a relatively poor description of the raw data by the latter, will lead to a large dependence on the MC model to be studied in the next section.

8.4 Corrected Data

An assessment of the systematic uncertainties is performed following a procedure analogous to that used in the case of the fragmentation functions described in section 7.4. In particular, the list of systematic checks is precisely the same.

The dominant systematic shifts to the rescaled angular distribution, $\hat{r}_2$, are depicted in Fig. 8.9 as ratios between the values obtained in the shifted and central analyses. The model dependence due to the HERWIG-ARIADNE comparison is the largest effect in most of the bins. The typical uncertainty from this source is of the order of 10% without any particular pattern with respect to the scaling variable. In most of the cases HERWIG predicts a lower value of the rescaled distribution. Other effects, shown in the plot or omitted, are limited to a few percent.

As appropriate for a quantity of small magnitude, systematic shifts to the rescaled correlation function, r_2, are given in Fig. 8.10 in a form of absolute differences. Here, again the MC model proves to be the dominant uncertainty, typically of the order of 0.05 but significantly higher for large ε in some kinematic bins. There is a tendency for this shift to be positive. Other systematic effects are

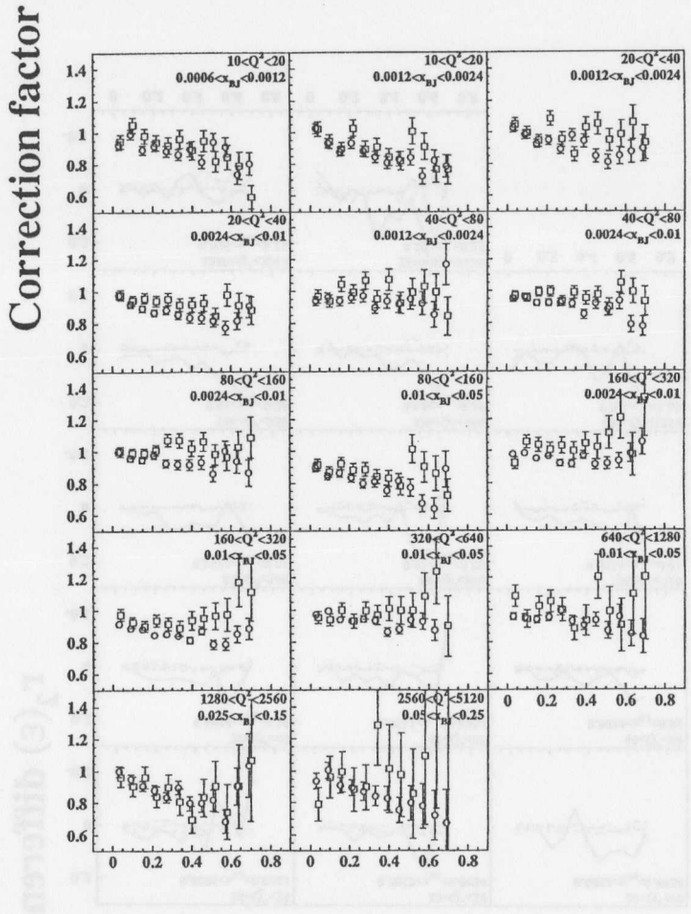

Figure 8.8: Correction factors to the *non-rescaled* $r_2^{raw}(\varepsilon)$ distributions evaluated for the 1995 detector configuration using ARIADNE (circles) and HERWIG (squares).

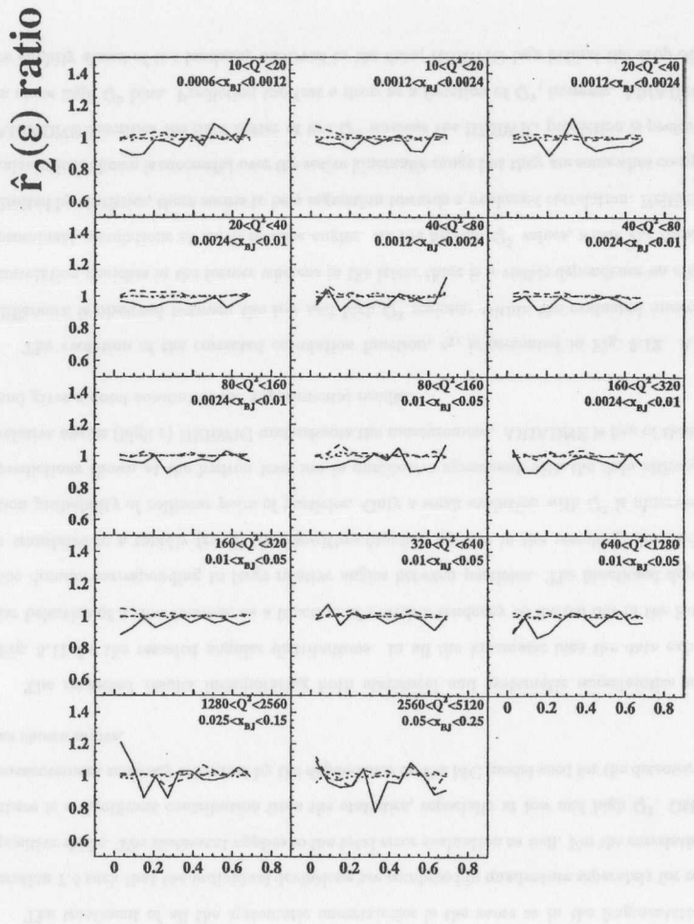

Figure 8.9: Dominant systematic shifts to the relative angular distribution: HERWIG (solid), off-vertex tracks (dashed), track topology cut (dotted) and rejection of LRG events (dot-dashed). Ratios of the shifted and central analyses are plotted.

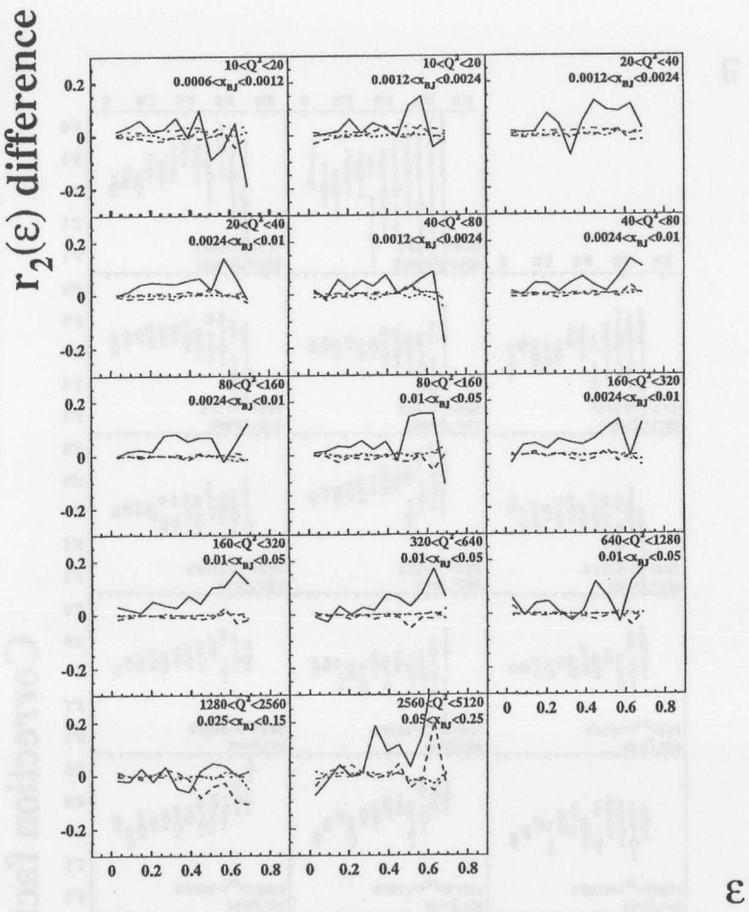

Figure 8.10: Dominant systematic shifts to the angular correlation: HERWIG (solid), off-vertex tracks (dashed), track topology cut (dotted) and rejection of LRG events (dot-dashed). Absolute differences between the shifted and central analyses are plotted.

strongly suppressed with respect to this uncertainty; this is particularly clear in the intermediate Q^2 range.

The treatment of all the systematic uncertainties is the same as in the fragmentation study in section 7.4 such that the individual deviations are combined in quadrature separately for negative and positive shifts. The statement applies to the total error evaluation as well. For the correlation function, there is a significant contribution from the statistics, especially at low and high Q^2. Otherwise, the measurement accuracy is limited by the dependence on the MC model used for the detector corrections as shown above.

The corrected results incorporating both statistical and systematic uncertainties are given in Fig. 8.11 for the rescaled angular distributions. In all the kinematic bins the data exhibit a similar behavior of a slow increase as a function of ε with a tendency to flatten out in the lowest part of the domain corresponding to large relative angles between particles. The functional dependence on ε translates to a rapidly (recall the logarithm function applied in the rescaling) decreasing production probability of collinear pairs of particles. Only a weak evolution with Q^2 is observed. The MC predictions shown at the hadron level are in qualitative agreement with the data although at small relative angles (high ε) HERWIG undershoots the measurement. ARIADNE is free of that defficiency and gives a good account of the experimental results.

The evolution of the corrected correlation function, r_2, is presented in Fig. 8.12. A qualitative difference is observed between the low and high Q^2 regions: within the evaluated uncertainties the correlation vanishes in the former whereas in the latter there is a visible dependence on ε indicative of prominent correlations at small relative angles. At the highest Q^2 values, where the measurement is limited by statistics, there seems to be a regression towards a weakened correlation. Neither of the MC calculations shown is successful over the entire kinematic range but they are somewhat complementary: ARIADNE describes the data better at low Q^2 whereas the HERWIG prediction is preferred at least in some high Q^2 bins. Predicting too fast a drop as a function of Q^2, however, ARIADNE seems to be slightly ahead of the tendency observed in the data; HERWIG lags behind the drop of the data.

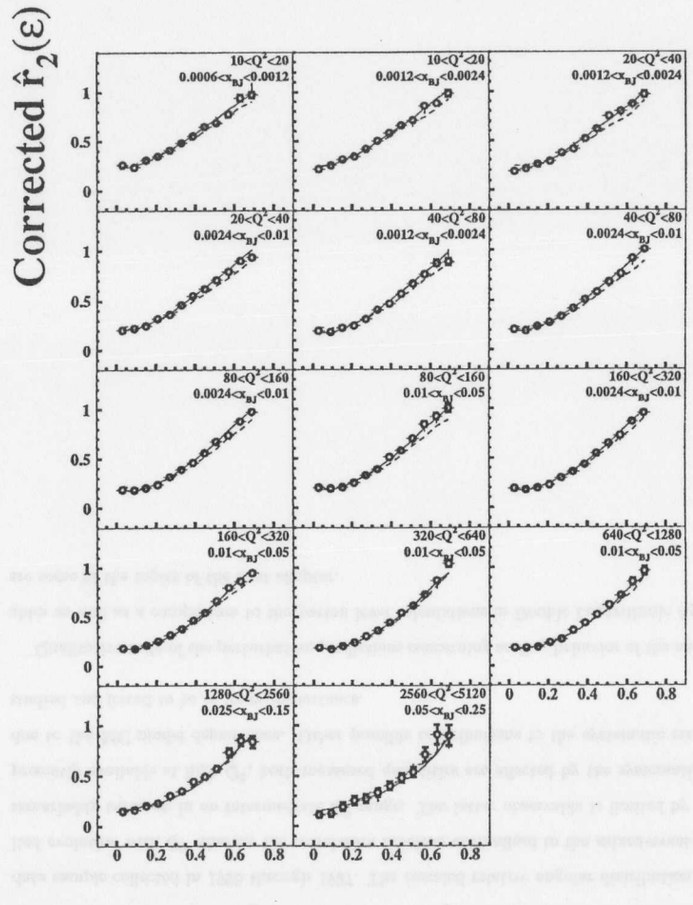

Figure 8.11: Corrected rescaled relative angular distribution $\hat{r}_2$ from the 1995 data sample (circles, statistical and total uncertainties) compared to ARIADNE (HERWIG) shown as solid (dashed) lines. The cone size is $\Theta = 90°$.

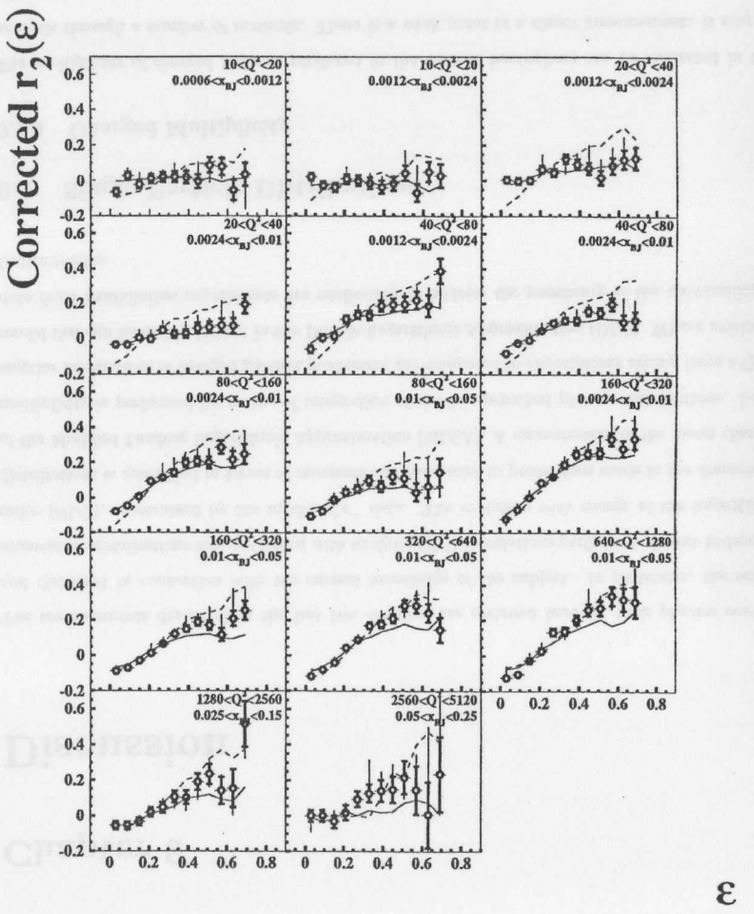

Figure 8.12: Corrected rescaled angular correlation function r_2 from the 1995 data sample (circles, statistical and total uncertainties) compared to ARIADNE (HERWIG) shown as solid (dashed) lines. The cone size is $\Theta = 90°$.

8.5 Summary

Correlations in the relative angle between pairs of charged hadrons have been measured for the DIS data sample collected in 1995 through 1997. The rescaled relative angular distribution exhibits limited evolution with Q^2 whereas the correlation function normalized to the mixed-event track sample remarkably turns on in an intermediate Q^2 range. The latter observable is limited by the statistics presently available at high Q^2; both measured quantities are affected by the systematic uncertainty due to the MC model dependence. Other possible contributions to the systematic error have been studied and found to be of lesser importance.

Qualitative tests of the perturbative predictions concerning scaling behavior of the angular observables as well as a comparison to the parton level calculations in Double Logarithmic Approximation are some of the topics of the next chapter.

Chapter 9

Discussion

The measurements described in the last two chapters are reviewed here for their physics content and discussed in connection with the current knowledge of the subject. In particular, the scaled momentum distributions are confronted with analytic QCD calculations performed at next-to-leading order (NLO), constrained by fits to the e^+e^- data. The evolution with energy of the logarithmic distributions is quantified in terms of moments and compared to predictions made in the framework of the Modified Leading Logarithmic Approximation (MLLA). A measurement of the mean charged multiplicity is performed by means of integration of the hump-backed plateau distributions. Local angular fluctuations of charged particle production are compared to expectations arising from a QCD model through formulae derived in the Double Logarithmic Approximation (DLA). Where available, data from annihilation experiments are confronted to address the possibility of the universality of fragmentation.

9.1 Single-Particle Distributions

9.1.1 Charged Multiplicity

The multiplicity of charged hadrons produced in the current hemisphere can be extracted in this analysis through a number of methods. There is a weak point in a direct measurement: it may be affected by a single correction factor to be applied to the mean value so that one is forced to trust

completely the MC description of this feature of the event. In addition, an approach through the multiplicity distribution itself is a difficult task due to unavoidable significant migrations between bins and so it is not attempted in this analysis. Alternatively, an integral of any single-particle inclusive distribution can be used instead permitting a more detailed insight into the quality of the MC simulation while keeping the corrections under control by appropriate binning. The $\ln(1/x_P)$ distribution was chosen for that purpose and a fixed integration range between -1 and 7 was used. Statistical uncertainties were evaluated assuming bin independence; systematic shifts were obtained by repeating the integration for distributions measured under alternative conditions. The results are quoted in Table 9.1 and compared to fixed target $\bar{\nu}_\mu p$ DIS [72, 73] data and e^+e^- measurements [74, 75, 76, 77] in Fig. 9.1. In addition, published HERA data are shown for ZEUS [12] and H1 [78]. The analysis reported in [79] based on an almost equivalent data set is also presented.

Over the kinematic range covered by this analysis the mean charged multiplicity spans almost an order of magnitude. The data provide no evidence for dependence on x_{Bj} (overlapping points in the lower part of the Q^2 range). There is a good agreement with earlier HERA data as well as with the new ZEUS analysis. The new results on the fixed target DIS follow the Q^2 evolution observed in the present analysis even though the respective x_{Bj} ranges differ by two orders of magnitude. Further support for the universality of fragmentation stems from the good agreement with the annihilation experiments for intermediate energies; the discrepancy at low energies can be attributed to the presence of additional processes in DIS leading to a depopulation of the current region [80].

Predictions of various Monte Carlo models are compared to the multiplicity data in Fig. 9.2. It is seen that calculations performed with the default settings of tuneable parameters provide a reasonable description of the data in case of ARIADNE and HERWIG, whereas LEPTO overestimates both the slope of the evolution and the magnitude of the charged multiplicity at high energies.

9.1.2 Scaled Momentum Distributions

The inclusive normalized cross-section for charged particle production, $1/\sigma\ d\sigma/dx_P$, is shown in Fig. 9.3 as a function of energy scale, Q, for different ranges of scaled momentum, x_P. For the sake of clarity the data for the two softest intervals have been scaled upwards; two types of solid symbols

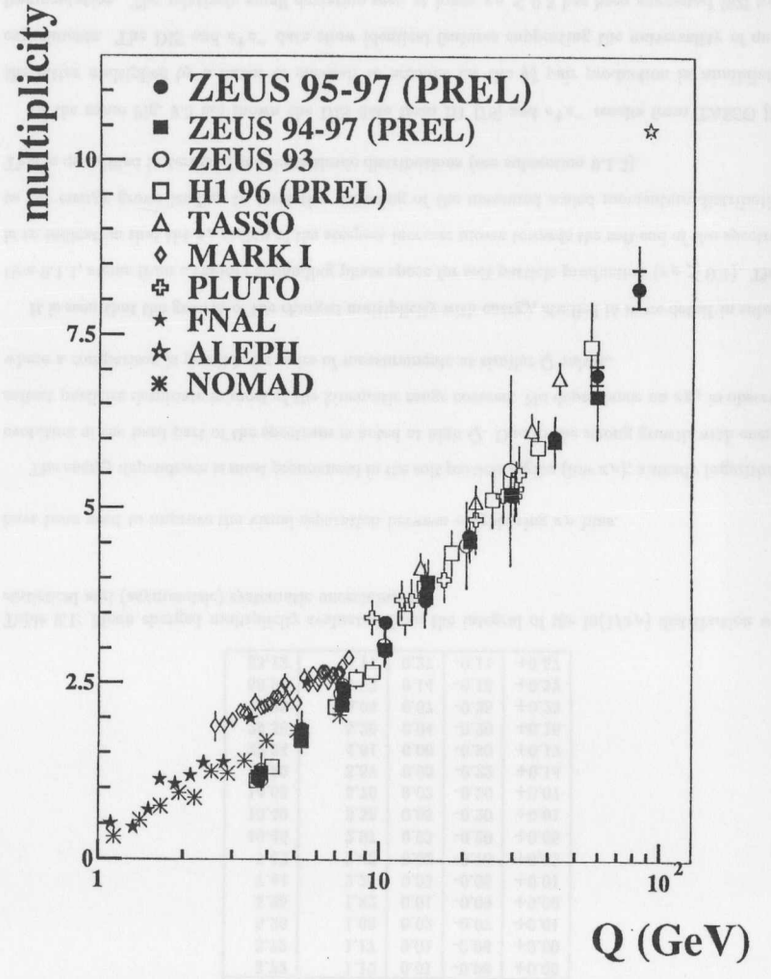

Figure 9.1: The mean charged multiplicity evaluated from the integral of the $\ln(1/x_P)$ distributions. The evolution with energy is shown for this analysis (solid points) compared to other sets of DIS and e^+e^- data.

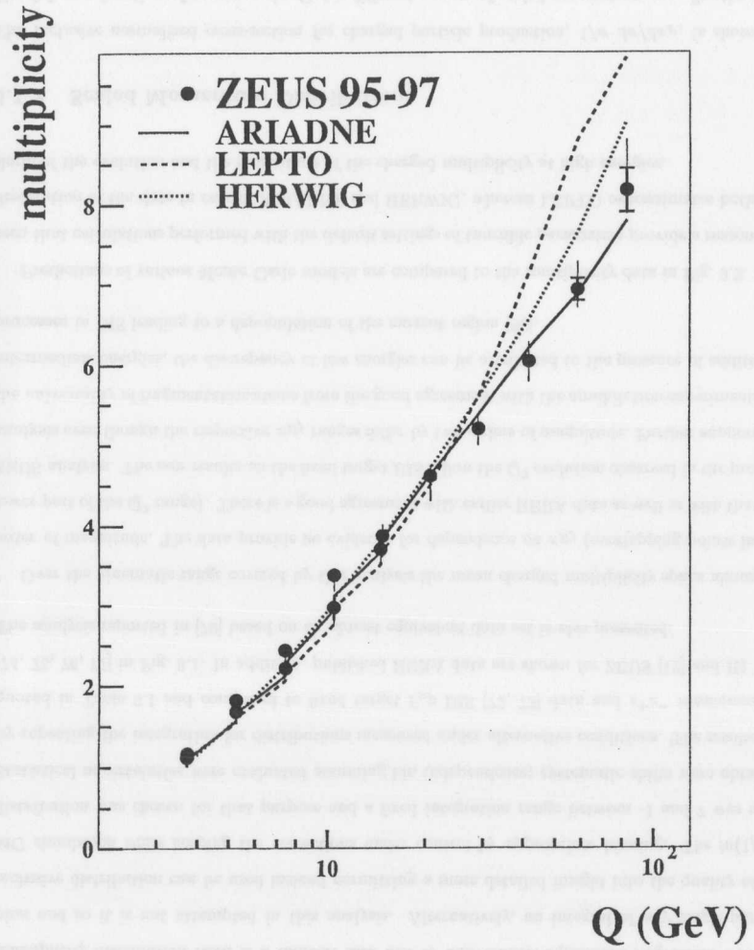

Figure 9.2: The mean charged multiplicity compared to MC calculations.

$< Q >$	$< N_{ch} >$	uncertainty		
		stat.	systematic	
3.72	1.12	0.01	-0.08	+0.00
3.72	1.17	0.01	-0.06	+0.00
5.26	1.68	0.02	-0.07	+0.01
5.26	1.82	0.01	-0.09	+0.00
7.44	2.21	0.03	-0.06	+0.01
7.42	2.43	0.02	-0.16	+0.02
10.46	2.97	0.03	-0.20	+0.05
10.49	3.38	0.05	-0.20	+0.01
14.60	3.70	0.03	-0.20	+0.07
14.82	3.87	0.03	-0.22	+0.14
20.84	4.61	0.06	-0.30	+0.12
29.30	5.20	0.04	-0.20	+0.18
41.83	6.04	0.07	-0.25	+0.23
58.95	6.93	0.14	-0.18	+0.32
83.12	8.17	0.27	-0.11	+0.57

Table 9.1: Mean charged multiplicity evaluated from the integral of the $\ln(1/x_P)$ distribution with statistical and (asymmetric) systematic uncertainties.

have been used to improve the visual separation between neighboring x_P bins.

The energy dependence is most pronounced in the soft particle region (low x_P); a steady logarithmic evolution of the hard part of the spectrum is noted at high Q. Due to the strong growth with energy, softest particles dominate in most of the kinematic range covered. No dependence on x_{Bj} is observed where a comparison is possible for pairs of measurements at similar Q values.

It is seen that the growth of the charged multiplicity with energy, studied in more detail in subsection 9.1.1, stems from a rapidly expanding phase space for soft particle production ($x_P \lesssim 0.1$). There is an indication that the x_P region of the steepest increase moves towards the soft end of the spectrum as the energy grows leading to a relative softening of the measured scaled momentum distribution. This is quantified in terms of the logarithmic distributions (see subsection 9.1.3).

In the same Fig. 9.3 are shown the DIS data from H1 [78] and e^+e^- results from TASSO [81], the latter multiplied by a factor of one-half to account for the $q\bar{q}$ pair production in annihilation experiments. The DIS and e^+e^- data show identical features supporting the universality of quark fragmentation. The relatively small deviation seen at lower $x_P \lesssim 0.3$ has been suggested [82] to be

related to the difference in quark flavor distribution between e^+e^- and DIS[1] rather than being a genuine fragmentation effect.

A simultaneous test of both the predictive power of QCD for single-particle inclusive observables and the concept of fragmentation universality has become possible once two milestones were achieved: the hard scattering matrix elements were calculated up to NLO [83] and the fragmentation functions were evaluated to the same order using e^+e^- data [84]. The convolution, resting on the assumption of the QCD factorization theorem, of the parton distribution functions, hard scattering matrix element and the fragmentation functions was performed using CYCLOPS [85]. The numerical results compared in Fig. 9.4 to the present analysis have been obtained with the MRSA' [51] parameterization of PDFs at $\Lambda_{QCD} = 0.23$ GeV.

It is seen that the calculation is able to follow the trends of the data apart from the low x_P, low Q region where the data exhibit a much stronger dependence on energy. Actually, this region is expected [86] to suffer from the limitations of the fragmentation function approach, namely from the neglect of hadron masses. It is encouraging to observe a good agreement with the data in the remaining part of the phase space, a strong support for fragmentation universality. To study the hadron mass effects, the calculation was repeated assigning the pion mass to the generated hadrons in the final boost to the Breit frame: although not self-consistent this approach helps to identify regions strongly influenced by the massless approximation. Not surprisingly, the resultant dashed lines indicate the importance of the mass effects for soft particle production: the calculation reproduces quantitatively the strong scaling violations present in the data.

Measurement of the scaling violations in the region of good theoretical understanding could be quantified in terms of an extraction of the strong coupling constant α_S. For that purpose a set of fragmentation functions fitted assuming different values of α_S must be used, like the one reported in [87], to allow for appropriate interpolation. It is disappointing that a change made to the procedure of e^+e^- data fitting rendered the new parameterizations marginally useful for DIS calculations, as can be inferred from the dotted curve of Fig. 9.4. Namely, a uniform treatment of the light quarks (u, d) and the strange quark led to a persistent overestimate of particle production up to $x_P \approx 0.3$

[1]There are about 0.2–0.3 fewer strange quarks produced in DIS compared to e^+e^- annihilation.

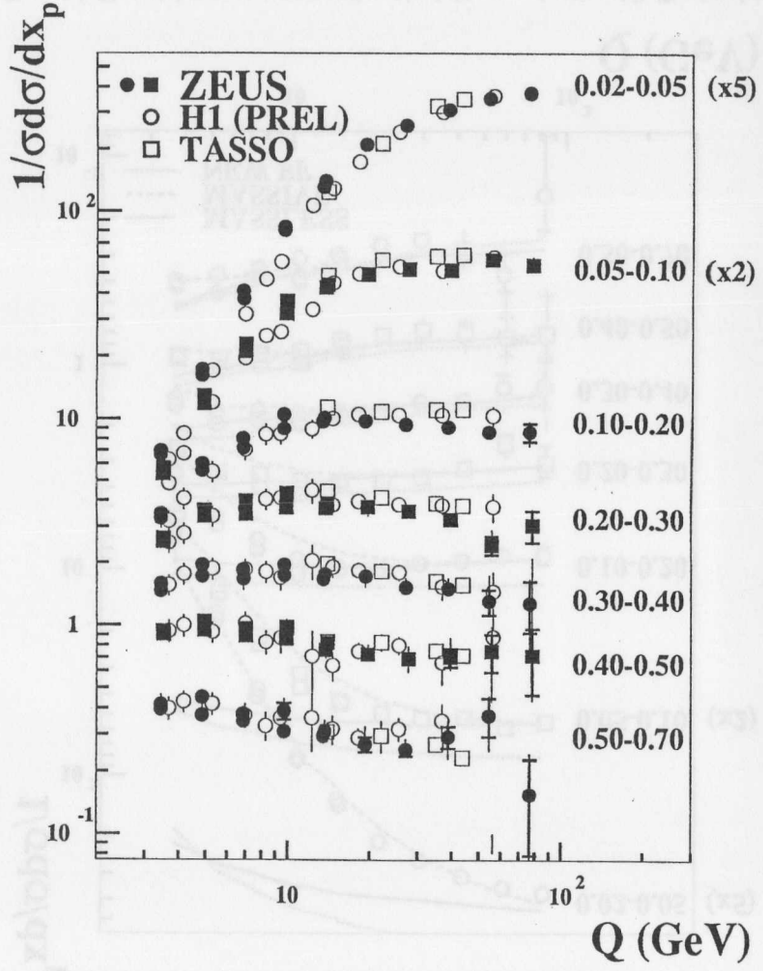

Figure 9.3: The scaled momentum cross-section, $1/\sigma\ d\sigma/dx_P$, as a function of Q. The data (solid points) are compared to H1 (open squares) and TASSO (open circles) results. To improve readability, for the two lowest x_P ranges multiplicative factors were used as indicated in the plot.

mirroring the feature of the e^+e^- data displayed in Fig. 9.3. The overestimate is present irrespective of the choice of the α_S value and PDF parameterization used.

Finally, the performance of the LO Monte Carlo algorithms is investigated in Fig. 9.5. Discontinuities in the lines are due to pairs of kinematic bins with equal Q. It is noted that all of the presented calculations reproduce the gross features of the data discussed above. Both ARIADNE and HERWIG provide a reasonably good description of the measurement. The soft end of the particle spectrum is identified as the main contribution to the LEPTO overestimate of the total charged multiplicity (see Fig. 9.2). Although only leading order QCD effects are consistently taken into account in the MC programs discussed it is seen that the parton shower treatment together with hadronization leads to a better description of the data than the fixed-order NLO calculations impeded by the massless hadron assumption.

9.1.3 $\ln(1/x_P)$ Distributions vs MLLA Predictions

Theoretical predictions exist within the MLLA framework of QCD [21] for a number of observables derivable from the logarithmic scaled momentum distributions, $1/\sigma\ d\sigma/d\ln(1/x_P)$. In particular, away from the soft end of the particle spectrum, for $\ln(1/x_P) < \ln(Q/2p_0)$, where p_0 is of the order of a few hundred MeV, the distribution is expected to follow the so-called limiting spectrum:

$$\frac{1}{\sigma}\ \frac{d\sigma}{d\ln(1/x_P)} = \kappa_{ch}\bar{D}^{lim}(\ln(1/x_P), \ln(Q/2\Lambda_{eff})), \quad 0 \lesssim \ln(1/x_P) \lesssim \ln(Q/2p_0). \tag{9.1}$$

Two parameters enter the analytic expression: an effective energy scale, Λ_{eff}, and a normalization constant, κ_{ch}, connecting the calculated partonic to measured hadronic observables. The results of a separate fit in each kinematic bin are shown in Fig. 9.6. The p_0 cut-off chosen to be 0.4 GeV severely restricts the range of the fit at low Q^2; it relaxes, however, faster than the distribution evolves so that for $Q^2 \gtrsim 80$ GeV^2 the peak region becomes covered by the fit. The visible quality of the fit improves with energy: χ^2/NDF falls below 1.5. Consistency of the MLLA predictions at parton level with the measured hadron distributions supports the validity of the LPHD hypothesis for single-particle observables.

Values of Λ_{eff} from the fit are plotted as a function of energy in Fig. 9.7 and compared to an earlier

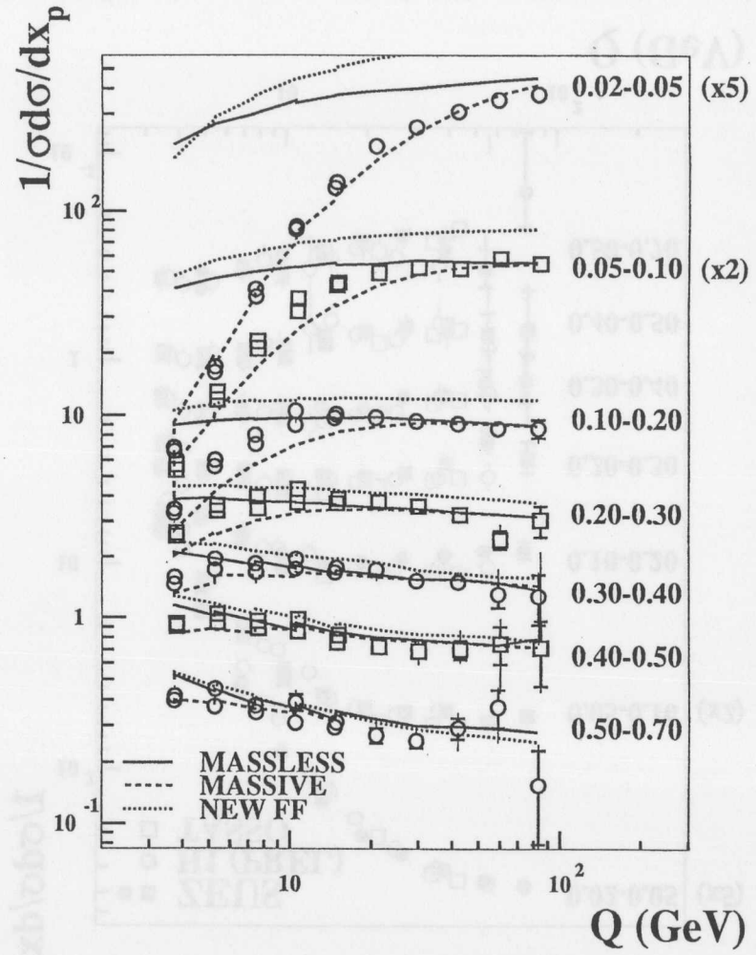

Figure 9.4: The scaled momentum cross-section, $1/\sigma\ d\sigma/dx_P$, as a function of Q. The data (points) are compared to QCD calculations. Also shown are predictions using as input different fragmentation functions (solid versus dotted curves) as well as the influence of non-vanishing hadron masses (dashed). To improve readability, for the two lowest x_P ranges multiplicative factors were used as indicated in the plot.

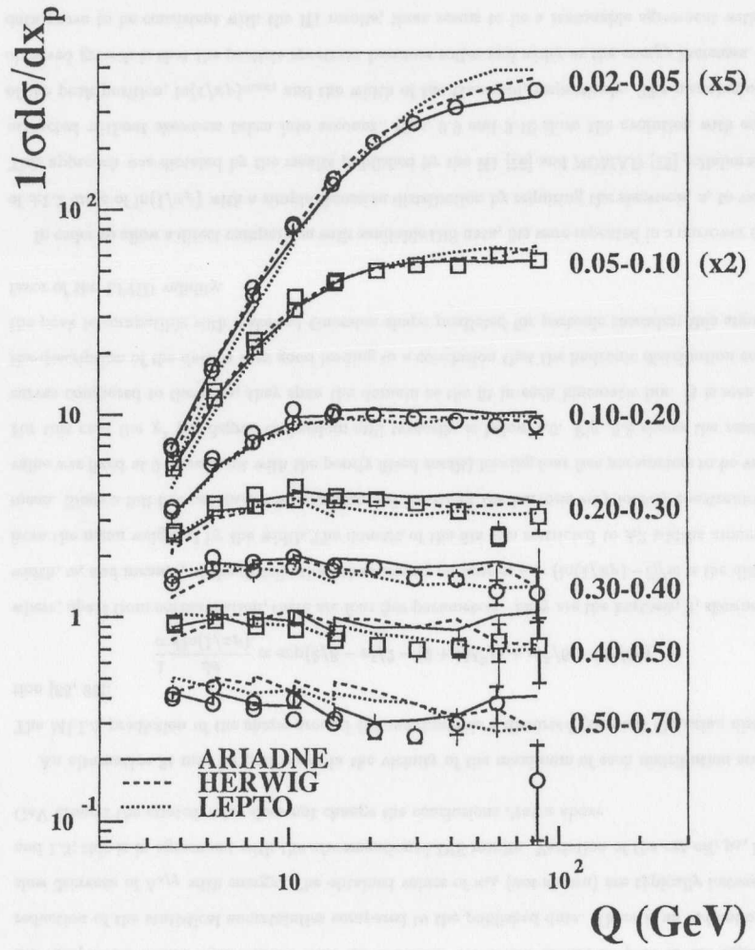

Figure 9.5: The scaled momentum cross-section, $1/\sigma \; d\sigma/dx_P$, as a function of Q. The data (points) are compared to three MC models: ARIADNE (solid line), HERWIG (dashed) and LEPTO (dotted). To improve readability, for the two lowest x_P ranges multiplicative factors were used as indicated in the plot.

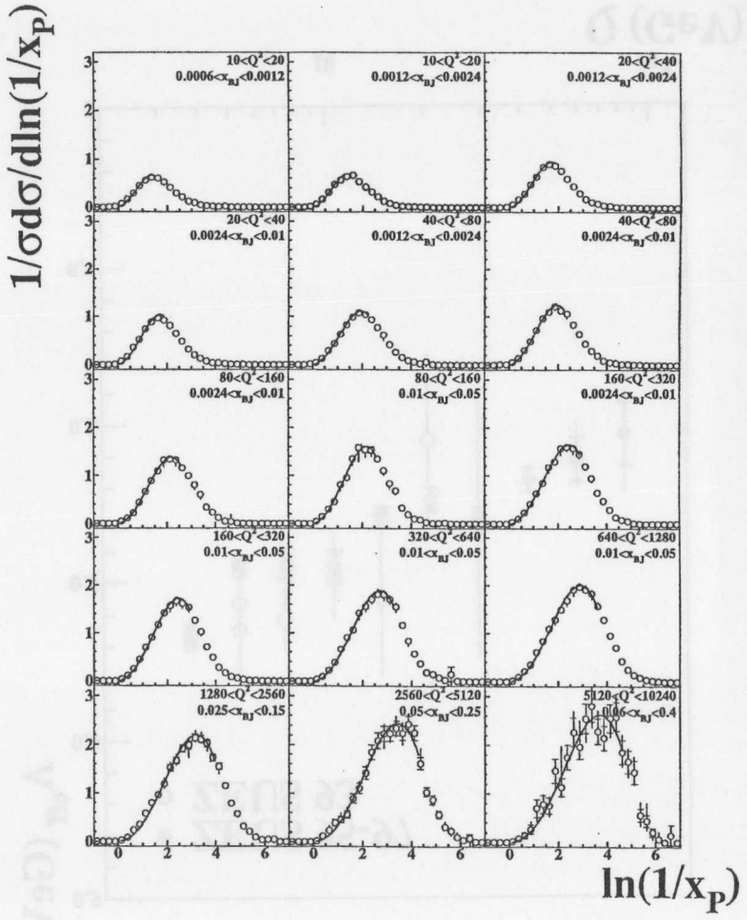

Figure 9.6: Limiting MLLA spectrum fit to the $\ln(1/x_P)$ distribution.

DIS measurement [12]. The new results fall in the range 0.2 to 0.3 GeV and, in the overlap region, are compatible within uncertainties with the old ones. Attention should be paid to the significant reduction of the statistical uncertainties compared to the published data. There is an indication of a slow decrease of Λ_{eff} with energy. The obtained values of κ_{ch} (not shown) are typically between 1.1 and 1.3; this is in agreement with the aforementioned DIS results. Variation of the cut-off, p_0, by 0.1 GeV around the quoted value does not change the conclusions drawn above.

An alternative fit may be performed in the vicinity of the maximum of each distribution studied. The MLLA prediction of the shape around the maximum is a distorted (skewed) Gaussian distribution [88, 89]:

$$\frac{1}{\sigma}\frac{d\sigma}{d\ln(1/x_P)} \propto \exp(k/8 - s\delta/2 - (2+k)\delta^2/4 + s\delta^3/6 + k\delta^4/24), \qquad (9.2)$$

where, apart from normalization, there are four free parameters. They are the kurtosis, k, skewness, s, width, w, and mean, l, of the distribution; the running parameter $\delta = (\ln(1/x_P) - l)/w$ is the distance from the mean weighted by the width.The domain of the fits was restricted to ± 2 widths around the mean. Since a full five-parameter fit was observed to render the kurtosis very loosely constrained, its value was fixed at 0 (consistent with the poorly fitted result) leaving four free parameters to be varied. For this case the χ^2 per degree of freedom still typically is below 1.0. Fig. 9.8 shows the resultant curves compared to the data; they span the domain of the fit in each kinematic bin. It is seen that the description of the data is very good leading to a conclusion that the hadronic distribution around the peak is compatible with a skewed Gaussian shape predicted for partonic cascades; this argues in favor of the LPHD validity.

In order to allow a direct comparison with available DIS data, fits were repeated in a narrower range of ± 1.2 units of $\ln(1/x_P)$ with a simple Gaussian distribution by requiring the skewness, s, to vanish. This approach was dictated by the results published by the H1 [78] and NOMAD [73] collaborations extracted without skewness taken into account. Figs. 9.9 and 9.10 show the evolution with energy of the peak position, $\ln(1/x_P)_{max}$, and the width of the Gaussian, respectively. The meaning of the observed growth is that the particle spectrum becomes softer and wider as the energy increases. The data prove to be consistent with the H1 results; there seems to be a reasonable agreement with the charged current, fixed target NOMAD measurement despite the difference in the x_{Bj} range.

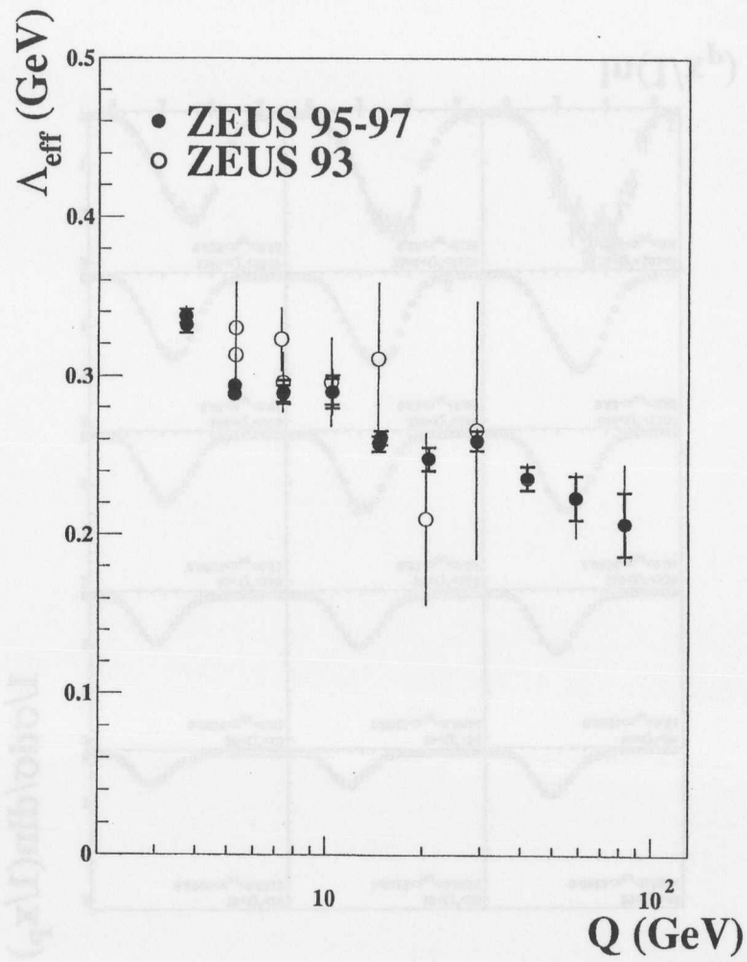

Figure 9.7: Evolution with energy of Λ_{eff} from the limitng spectrum fit to the $\ln(1/x_P)$ distribution. The data (solid points) are compared to published ZEUS results (empty points) showing statistical uncertainties only.

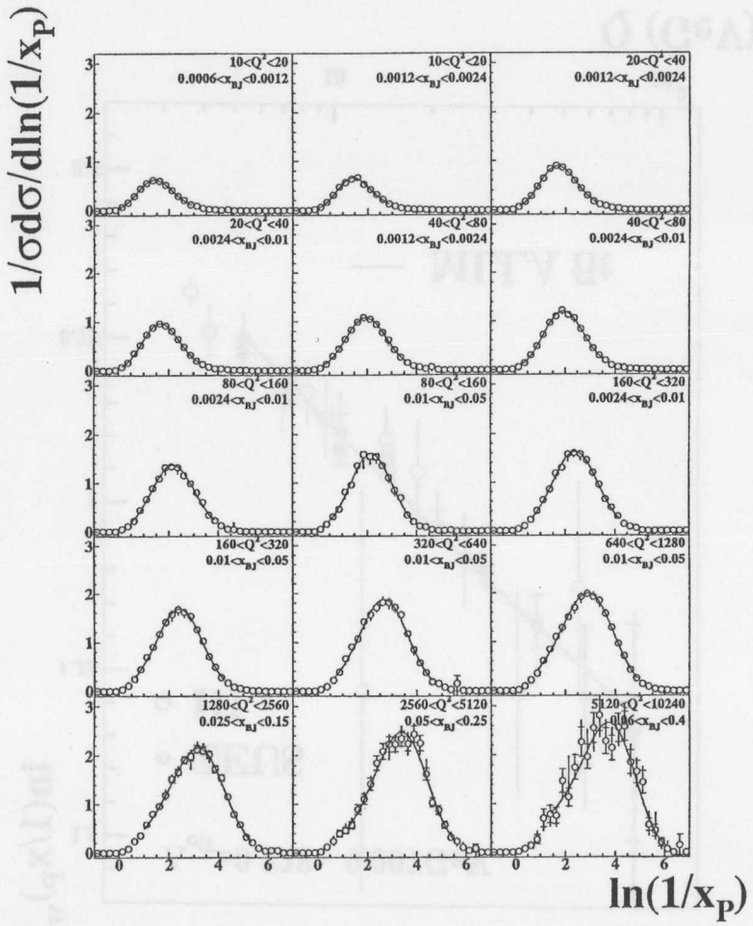

Figure 9.8: The Distorted Gaussian spectrum fit to the $\ln(1/x_P)$ distribution.

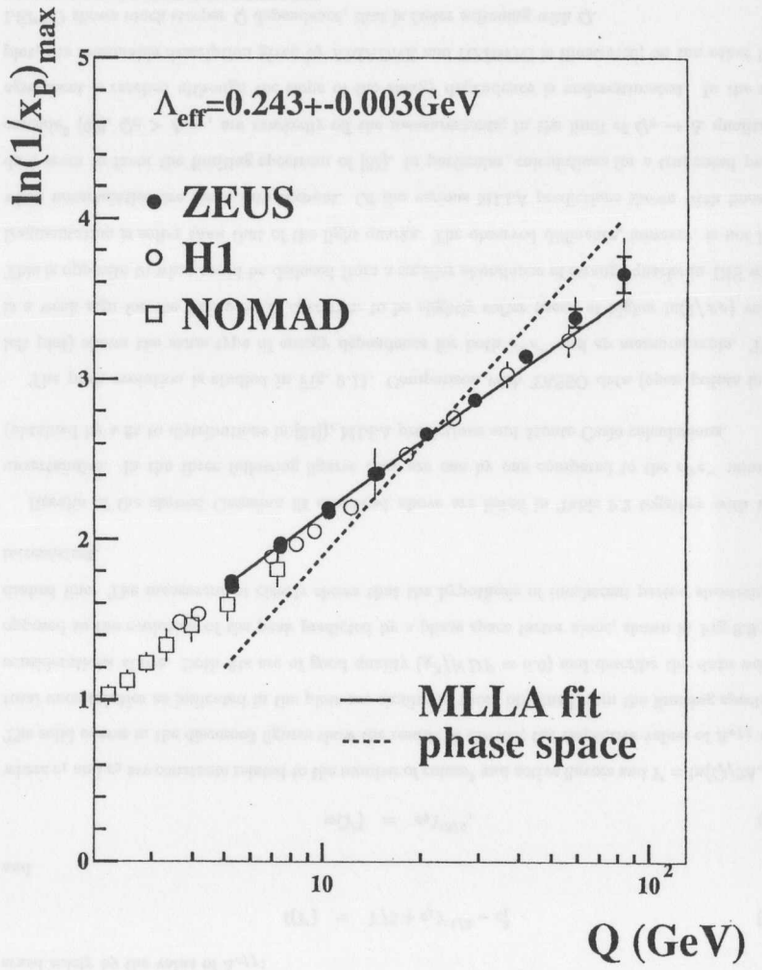

Figure 9.9: Energy evolution of the peak of the $\ln(1/x_P)$ distribution. A single-parameter MLLA fit is overlaid.

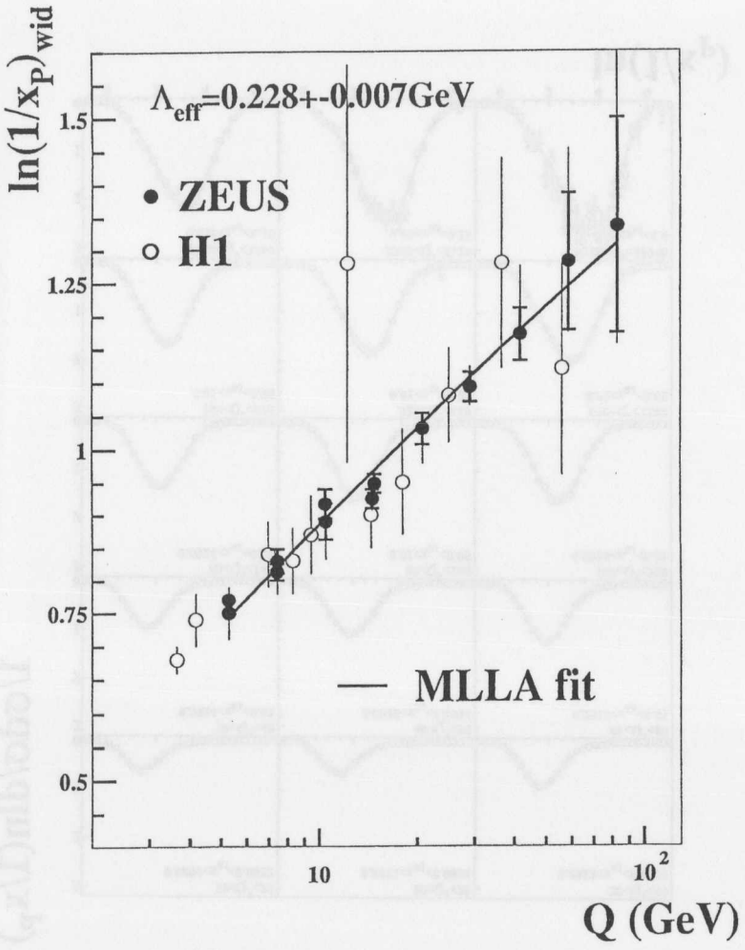

Figure 9.10: Energy evolution of the width of the $\ln(1/x_P)$ distribution. A single-parameter MLLA fit is overlaid.

Within the MLLA, the evolution of the Gaussian parameters with energy is predicted to be governed solely by the value of Λ_{eff}:

$$l(Y) = Y/2 + c_1 Y^{1/2} - c_1^2 \tag{9.3}$$

and

$$w(Y) = c_2 Y^{3/4}, \tag{9.4}$$

where c_1 and c_2 are constants related to the number of colors[2] and active flavors and $Y = \ln(Q/2\Lambda_{eff})$. The solid curves in the discussed figures show the results of the fits; the respective values of Λ_{eff} with total uncertainties as indicated in the plots are similar to those obtained from the limiting spectrum considerations above. Both fits are of good quality ($\chi^2/NDF \approx 0.6$) and describe the data well as opposed to the evolution of the peak predicted by a phase space factor alone, shown in Fig 9.9 as a dashed line. The measurement clearly shows that the hypothesis of incoherent parton showering is inconsistent.

Results of the skewed Gaussian fit described above are listed in Table 9.2 together with total uncertainties. In the three following figures they are one by one compared to the e^+e^- numbers (obtained by a fit to distributions in [81]), MLLA predictions and Monte Carlo calculations.

The peak evolution is studied in Fig. 9.11. Comparison with TASSO data (open points in the left plot) shows the same type of energy dependence for both e^+e^- and ep measurements. There is a weak sign for the DIS particle spectrum to be slightly softer (peak at higher $\ln(1/x_P)$ value). This is opposite to what could be deduced from a smaller abundance of strange quarks in DIS whose fragmentation is softer than that of the light quarks. The observed difference, however, is not large when uncertainties are taken into account. Of the various MLLA predictions shown with lines the data seem to favor the limiting spectrum of [88]. In particular, calculations for a truncated parton cascade[3] [90], $Q_0 > \Lambda_{eff}$, are markedly off the measurements; in the limit of $Q_0 \to \Lambda$ qualitative agreement is reached although the slope of the energy dependence is underestimated. In the right plot, the reasonable description given by ARIADNE and HERWIG is illustrated; on the other hand LEPTO shows much steeper Q dependence, that is faster softening with Q.

[2] A value of $n_f = 3$ was assumed to account for the studied kinematic range.
[3] Evolution of the parton cascade is terminated *before* the virtualities reach the cut-off Λ_{eff}.

$< Q >$	peak	width	skewness
3.72	$1.44^{+0.01}_{-0.05}$	$0.72^{+0.01}_{-0.01}$	$0.38^{+0.07}_{-0.04}$
3.72	$1.44^{+0.01}_{-0.03}$	$0.71^{+0.01}_{-0.02}$	$0.40^{+0.04}_{-0.11}$
5.26	$1.68^{+0.02}_{-0.02}$	$0.75^{+0.02}_{-0.01}$	$0.28^{+0.06}_{-0.07}$
5.26	$1.67^{+0.01}_{-0.03}$	$0.76^{+0.01}_{-0.01}$	$0.29^{+0.04}_{-0.04}$
7.44	$1.95^{+0.02}_{-0.04}$	$0.82^{+0.01}_{-0.03}$	$0.14^{+0.07}_{-0.06}$
7.42	$1.92^{+0.01}_{-0.03}$	$0.81^{+0.01}_{-0.01}$	$0.21^{+0.07}_{-0.04}$
10.46	$2.17^{+0.02}_{-0.03}$	$0.87^{+0.02}_{-0.01}$	$0.11^{+0.07}_{-0.07}$
10.49	$2.17^{+0.03}_{-0.05}$	$0.84^{+0.05}_{-0.01}$	$0.00^{+0.13}_{-0.12}$
14.60	$2.41^{+0.02}_{-0.03}$	$0.92^{+0.01}_{-0.01}$	$-0.02^{+0.10}_{-0.05}$
14.82	$2.42^{+0.02}_{-0.02}$	$0.92^{+0.01}_{-0.01}$	$-0.06^{+0.05}_{-0.04}$
20.84	$2.67^{+0.04}_{-0.02}$	$0.98^{+0.02}_{-0.01}$	$-0.13^{+0.05}_{-0.18}$
29.30	$2.91^{+0.03}_{-0.02}$	$1.07^{+0.01}_{-0.01}$	$-0.25^{+0.06}_{-0.11}$
41.83	$3.18^{+0.04}_{-0.04}$	$1.15^{+0.02}_{-0.02}$	$-0.36^{+0.09}_{-0.10}$
58.95	$3.37^{+0.09}_{-0.05}$	$1.12^{+0.05}_{-0.03}$	$-0.31^{+0.13}_{-0.26}$
83.12	$3.73^{+0.08}_{-0.15}$	$1.50^{+0.25}_{-0.46}$	$-1.41^{+1.39}_{-0.77}$

Table 9.2: Peak position, width and skewness of the $\ln(1/x_P)$ distribution from a distorted Gaussian fit. Total (asymmetric) uncertainties are given.

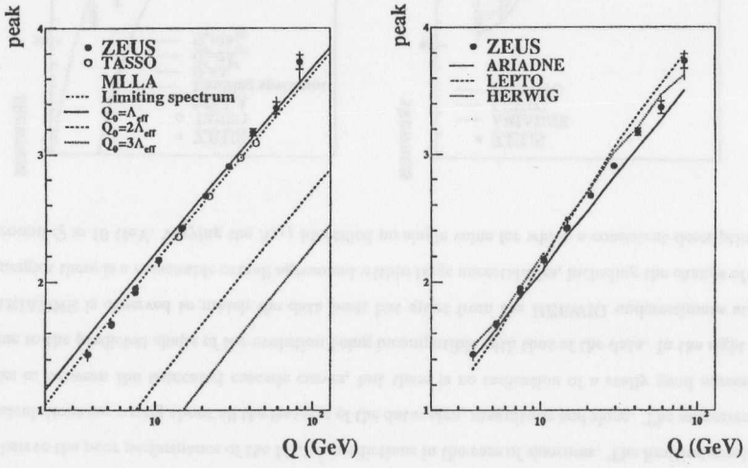

Figure 9.11: Evolution with energy of the peak of the $\ln(1/x_P)$ distribution from a distorted Gaussian fit. The data (solid points) are compared to the e^+e^- results and MLLA predictions (left plot) and Monte Carlo calculations (right plot).

Fig. 9.12 summarizes the width parameter in a similar manner to the peak treatment in the previous figure. The trend of a widening of the distribution with energy, observed already if a standard

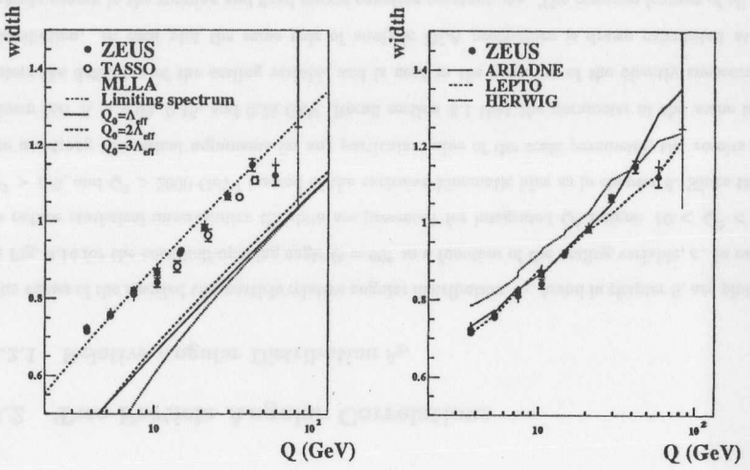

Figure 9.12: Evolution with energy of the width of the $\ln(1/x_P)$ distribution from a distorted Gaussian fit. The data (solid points) are compared to the e^+e^- results and MLLA predictions (left plot) and Monte Carlo calculations (right plot).

Gaussian shape is assumed, is not changed by inclusion of the skewness terms into the fit. Obviously, the width suffers from larger fluctuations compared to the peak value but this is not reflected in the uncertainties that remain small. The e^+e^- annihilation data are noted to be in reasonable agreement with the present analysis. Only the limiting spectrum MLLA prediction matches the magnitude and shape of the measured dependence; not even the $Q_0 = \Lambda_{eff}$ of the truncated cascade family of curves approaches the data. Monte Carlo simulations, shown at right, are far more successful in describing the data. Of the three models only HERWIG misses the absolute normalization but still accounts for the shape of the energy evolution. The other two calculations closely follow the experimental results.

Finally, the skewness of the $\ln(1/x_P)$ distributions is plotted in Fig. 9.13. An important feature

of the data is that the skewness decreases with energy and changes sign from positive to negative at $Q \approx 10$ GeV. This trend is present in the e^+e^- measurements as well; a good agreement is observed between the two data sets representing different types of hard scattering. It is disappointing to observe the poor performance of the MLLA predictions in the case of skewness. The limiting spectrum calculations are wrong about all the features of the data: sign, magnitude and slope. The measurement lies in between the truncated cascade curves, but there is no indication of a really good agreement due to the predicted shape of the evolution being incompatible with that of the data. In the right plot ARIADNE is observed to match the data best; but apart from the HERWIG underestimate at low energies there is a reasonable overall agreement within large uncertainties, including the change of sign around $Q \approx 10$ GeV. Varying the Λ_{eff} identified no single value for which a consistent description of

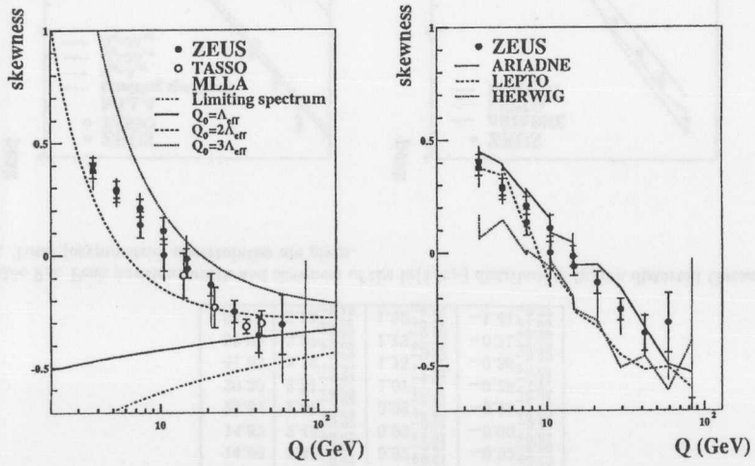

Figure 9.13: Evolution with energy of the skewness of the $\ln(1/x_P)$ distribution from a distorted Gaussian fit. The data (solid points) are compared to the e^+e^- results and MLLA predictions (left plot) and Monte Carlo calculations (right plot).

the three parameters studied could be achieved.

9.1.4 Summary

The increased statistics of the ZEUS DIS sample allowed an improvement in the precision of the measurement of single-particle inclusive distributions. Thanks to the reduced uncertainties and extended kinematic region stronger statements became possible concerning the quark fragmentation universality and the validity of joint MLLA and LPHD predictions. No significant indication was observed of the violation of the former concept. Deviations from an exact match between e^+e^- and ep measurements could be explained in terms of expected important differences between the two processes in question: DIS and e^+e^- annihilation. On the other hand, quantitative MLLA predictions of the shape of the $\ln(1/x_P)$ distributions could not be brought into agreement with the data whereas self-consistent fits to the limiting spectrum shape were possible leading to acceptable values of Λ_{eff} and κ_{ch}. It is not clear whether the failure of the skewed Gaussian related predictions is due to neglected higher order QCD corrections or the inapplicability of the LPHD to more subtle features of the single-particle spectra.

9.2 Two-Particle Angular Correlations

9.2.1 Relative Angular Distribution $\hat{r}_2$.

The values of the rescaled two-particle relative angular distribution, $\hat{r}_2$, found in chapter 8, are plotted in Fig. 9.14 for the cone half-opening angle $\Theta = 60°$ as a function of the scaling variable, ε. In order to reduce statistical uncertainties the data are presented for integrated Q^2 ranges: $10 < Q^2 < 20$, $Q^2 > 100$, and $Q^2 > 2000$ GeV^2 instead of the exclusive kinematic bins as in chapter 8. Since there are no strong theoretical arguments for any particular value of the scale parameter, the results are shown for: $\Lambda = 0.05$, 0.15, and 0.25 GeV. Recall section 8.1 that the parameter at the same time enters the definition of the scaling variable and is used in the rescaling of the directly measurable distribution. In each plot the same pair of analytic DLA predictions is drawn calculated at an *infinite* energy in the running and fixed strong coupling constant, α_S. The common feature of all the experimental distributions is a steep growth of $\hat{r}_2$ with ε. In the language of the relative angle, due to the normalization of the distribution, this dependence translates into a strong suppression of the

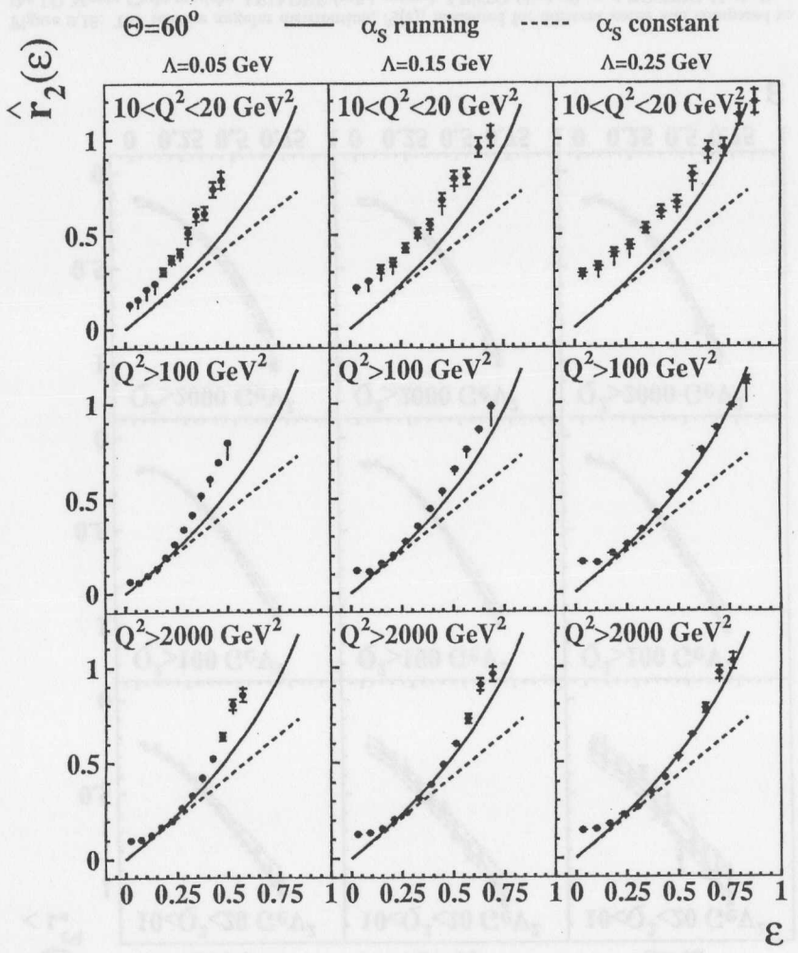

Figure 9.14: The relative angular distribution, $\hat{r}_2(\varepsilon)$, measured for cone $\Theta = 60°$ in integrated kinematic bins and for different choices of the scale parameter, Λ. The data are compared to the DLA predictions in the running (solid curve) and fixed (dashed) α_S regime.

rate of production of nearly collinear particle pairs. Unlike the angular correlation function r_2, the relative angular distribution $\hat{r}_2$ discussed here is mostly sensitive to the single particle density so that the geometric factor (phase space) determines the general behavior of the distribution. The evolution of the data with energy can be observed in terms of approaching from above the analytic curve representing the running α_S case except for the low ε range where the measured distribution flattens out and departs from the QCD prediction. This discrepancy can be attributed to the shortcomings of the theoretical model as the neglect of the recoil in gluon radiation mostly affects particles produced at large angles (*i. e.* small ε). It is seen that for $\Lambda \gtrsim 0.15$ GeV the high energy data are in reasonable agreement with the DLA prediction for running α_S; there is no indication of any compatibility with the prediction in the fixed α_S regime. The data seem to develop a bump below the analytic curve in the intermediate ε range as the energy increases but extension beyond the currently accessible kinematic range would be needed to confirm this observation. As the energy dependence tends to vanish at high Q^2, a scaling behavior of the rescaled relative distribution may be claimed.

The same observable for different cone half-opening angles is presented in Fig. 9.15. A value of $\Lambda = 0.15$ GeV was used for all plots. For any studied value of Θ between 90° and 45° the data exhibit similar evolution with energy as observed in the previous figure. At fixed energy, the rescaled distributions seem to become steeper as the cone narrows but the effect is not strong. In particular, there is hardly any difference between the high ε ends of the distributions measured at high energies for 60° and 45°; scaling with respect to the cone width is thus noted in addition to the approximate scaling with the energy discussed above. In the same figure, predictions of three LO Monte Carlo programs are given. The overall agreement is good in the case of ARIADNE, whereas at low Q^2 LEPTO and HERWIG over- and undershoot the data, respectively; at high energies and not too wide cones the data and the simulations nearly coincide.

A comparison with the e^+e^- annihilation measurement by DELPHI [91] is presented in Fig. 9.16. For compatibility with this and other correlation studies at LEP [92, 93], Λ was assumed to be 0.15 GeV; the ZEUS data for $\Theta = 45°$ were chosen for the comparison. A significant difference in the mean energy between the two experiments should be noted: $< Q > /2 = 30.6$ GeV and $\sqrt{s}/2 = 45.5$ GeV for DIS and e^+e^-, respectively. A very good agreement is observed between the two data sets in

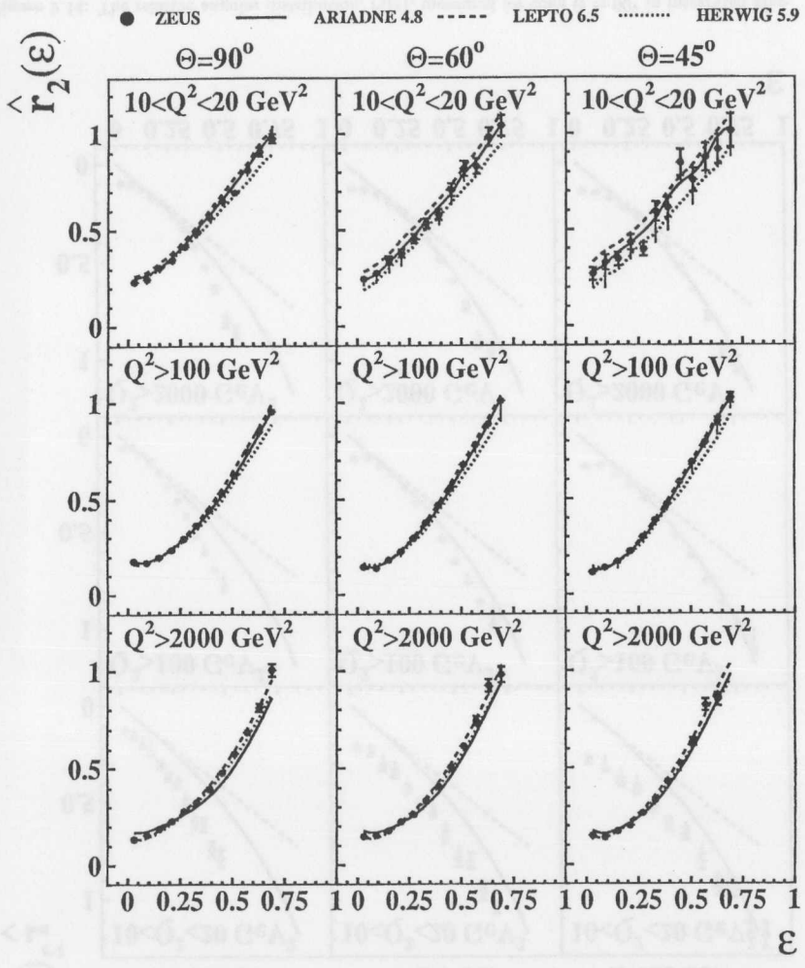

Figure 9.15: The relative angular distribution, $\hat{r}_2(\varepsilon)$, measured for different cones and compared to the LO Monte Carlo models: ARIADNE (solid curves), LEPTO (dashed), and HERWIG (dotted).

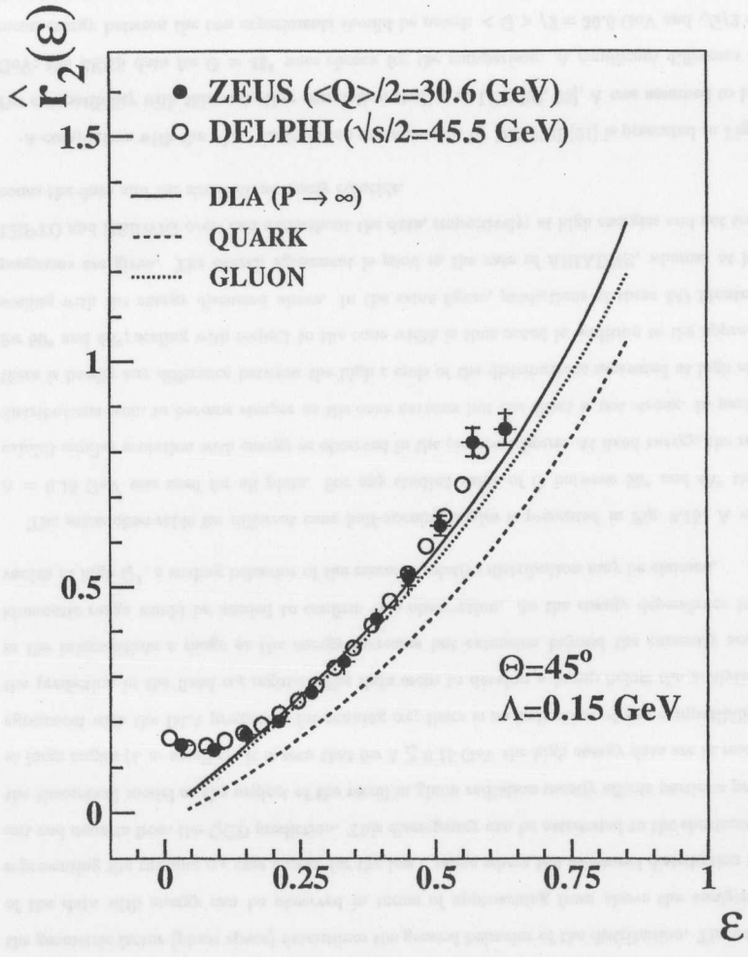

Figure 9.16: The relative angular distribution, $\hat{r}_2(\varepsilon)$, for cone $\Theta = 45°$ and $\Lambda = 0.15$ GeV. The data are compared to the DELPHI measurement and DLA calculations: asymptotic (solid curve) and at the finite energy of $P = 30.6$ GeV for the quark- (dashed) and gluon-originated (dotted) partonic cascade.

support of both universality of fragmentation and scaling in energy. In this figure, analytic curves are drawn for the quark- and gluon-originated cascades of finite (the ZEUS) energy as well an asymptotic, *i. e.* for an infinite energy, calculation; there is only a tiny splitting between the gluon and asymptotic predictions. The data, apart from the already discussed low ε region, follow the *gluon* curve: this is quite surprising since the current hemisphere of the Breit frame in DIS is populated predominantly by quark-originated cascades. In the DLA framework, however, a value 9/4 of the ratio of the mean parton multiplicities in gluon and quark cascades is used leading to an overestimated splitting between the two curves. A smaller value, obtained taking into account higher order corrections, would reduce this difference.

9.2.2 Angular Correlation Function r_2

The angular correlation function $r_2(\varepsilon)$ for $\Theta = 60°$ is tudied versus Q^2 in Fig. 9.17 for three different Λ values. A significant difference is noted between the low (10-20 GeV^2) and high Q^2 data. The choice of Λ is observed not to have a qualitative influence on the data in a way that the correlations are absent at low Q^2 but they do develop at higher energies with a steady dependence on the scaling variable: even there the correlation nearly vanishes for widely separated pairs of particles falling into the small ε region. In the highest Q^2 bin the data are even suggestive of a negative correlation in that angular region. On the transition between the two upper kinematic bins there isn't any sign of a strong energy dependence for not too low ε; the uncertainties become excessive though preventing a firm statement in that matter. On the grounds of the agreement with the asymptotic analytic predictions there is no reason to favor any value of Λ; the limited domain of ε does not allow experimental distinction between the running and fixed coupling scenarios whose predictions are nearly coincident.

The correlation function data, evaluated for the scale parameter $\Lambda = 0.15$ GeV, are examined for a possible dependence on the cone width in Fig. 9.18. In the lowest Q^2 region the distribution is flat irrespective of the Θ angle and similar (approximate) independence is observed in the highest measured Q^2 bin. The data in the intermediate kinematic range exhibit some sensitivity, concentrated near $\varepsilon \approx 0$: the narrower the cone the higher the (large-angle) correlations. Thus the support for the scaling behavior of the correlation function with respect to the energy and cone width is rather

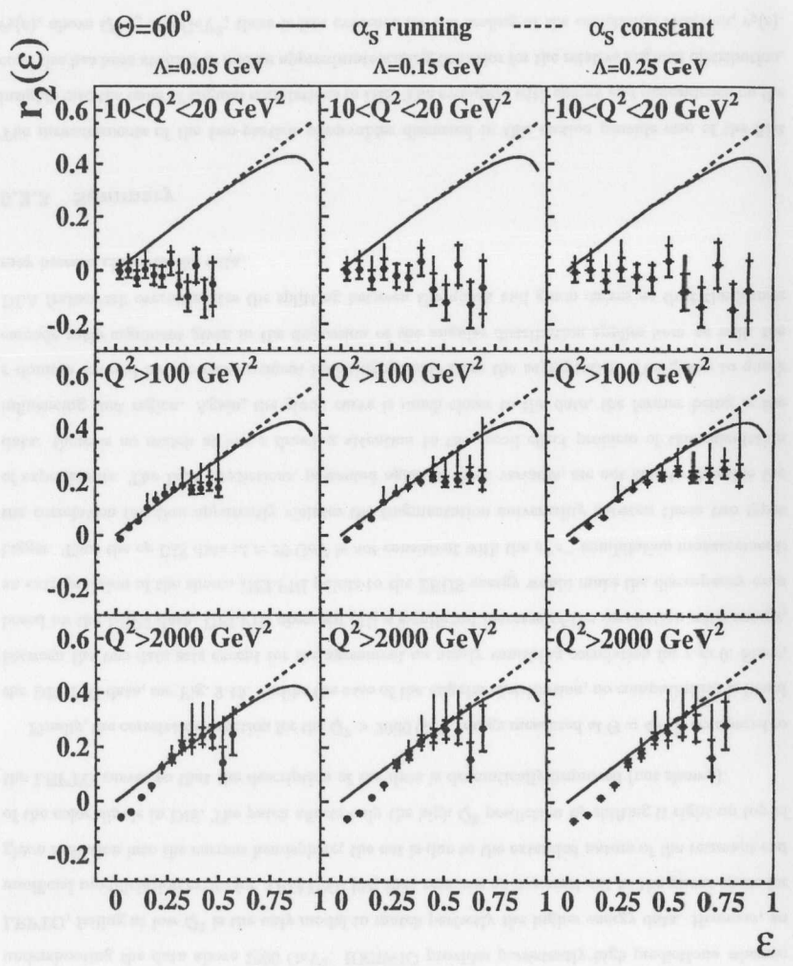

Figure 9.17: The angular correlation function, $r_2(\varepsilon)$, measured for cone $\Theta = 60°$ in integrated kinematic bins and for different choices of the scale parameter, Λ. The data are compared to the DLA predictions in the running (solid curve) and fixed (dashed) α_S regime.

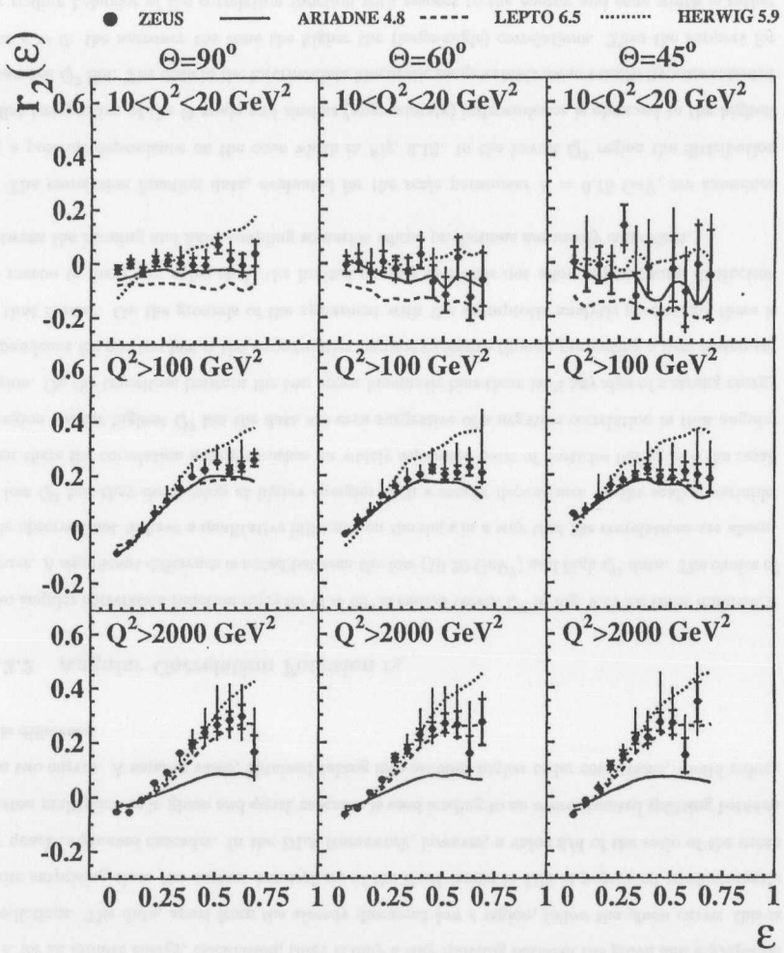

Figure 9.18: The angular correlation function, $r_2(\varepsilon)$, measured for different cones and compared to the LO Monte Carlo models: ARIADNE (solid curves), LEPTO (dashed), and HERWIG (dotted).

weak. Of the three Monte Carlo calculations shown in the same figure none is able to describe the measurement satisfactorily. ARIADNE, successful at low and intermediate energies, fails considerably undershooting the data above 2000 GeV^2. HERWIG provides persistently high predictions whereas LEPTO, failing at low Q^2 is the only model to match perfectly the higher energy data. However, an unofficial modificiation exists for ARIADNE [94] that removes an incorrect cut to the phase space for gluon radiation into the current hemisphere; the cut is due to the extended nature of the remnant end of the color dipole in DIS. The patch affects only the high Q^2 prediction by shifting it right on top of the LEPTO curves so that the description of the data is dramatically improved (not shown).

Finally, the correlation function for the $Q^2 > 2000$ GeV^2 range measured at $\Theta = 45°$ is compared to the DELPHI data, see Fig. 9.19. Unlike the case of the angular distribution, no compatibility is found between the two data sets except for the agreement on nearly vanishing correlation for $\varepsilon \approx 0$. Since, based on the LEP2 data, DELPHI observed [91] a significant increase of the correlation with energy, an extrapolation of the shown DELPHI points to the ZEUS energy would make the discrepancy even bigger. Thus the *ep* DIS data at ≈ 30 GeV is *not* consistent with the e^+e^- annihilation measurement: the correlation function apparently violates the fragmentation universality between these two types of experiments. The DLA predictions, presented again in three variants, are not able to describe the data: there is no match at low ε drawing attention to the recoil effect problem of the calculation influencing that region. Again, the gluon curve is much closer to the data, the former being in the ε domain covered by the measurement indistinguishable from the asymptotics. The gluon to quark cascade ratio argument given in the discussion of the angular distribution applies here as well: the DLA framework overestimates the splitting between the quark and gluon curves so that the former may become closer to the data.

9.2.3 Summary

The measurements of the two-particle observables discussed in this section provide one of the first insights into the issue of angular correlations in DIS. The evolution with energy and dependence on the cone size has been studied to find an approximate scaling behavior for the relative angular distribution, $\hat{r}_2(\varepsilon)$, above $Q^2 \gtrsim 100$ GeV^2; there is less evidence for the scaling of the correlation function, $r_2(\varepsilon)$.

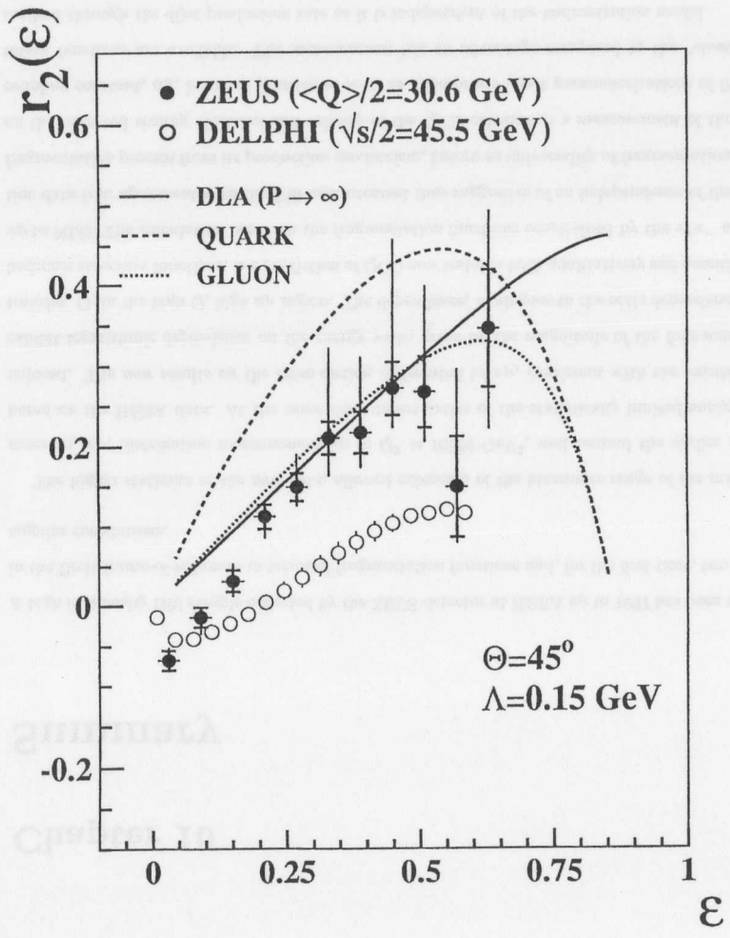

Figure 9.19: The angular correlation function, $r_2(\varepsilon)$, for cone $\Theta = 45°$ and $\Lambda = 0.15$ GeV. The data are compared to the DELPHI measurement and DLA calculations: asymptotic (solid curve) and at the finite energy of $P = 30.6$ GeV for the quark- (dashed) and gluon-originated (dotted) partonic cascade.

The status of experimental confirmation of the quantitative DLA predictions is positive for the former observable (the data prefer $\Lambda \gtrsim 0.15$ GeV) whereas the latter exhibits only a broad qualitative agreement with the running coupling regime of the analytic calculation. The quark-dominated data in both cases favor the prediction made for a *gluon*-originated cascade, an effect attributable at least in part to the DLA overestimation of the difference between the two species of parton cascades. Since the sensitivity to the single-particle distributions is mostly removed from the true correlation function compared to the angular distribution it is not clear whether the theoretical problems for the former are due to the approximations made in the calculation or a weaker stand of the LPHD concept for many-particle observables. Universality of fragmentation is examined through the comparison with the DELPHI data. There is a surprising mismatch between results of the two comparisons: a good agreement for the angular distributions is observed, whereas the correlation functions of DIS and annihilation are found to be incompatible. The latter result alone, being a sign of non-universality, is not completely surprising in connection to the similar difference of the long-range correlations reported recently in [80]. However, no particular reason for the discrepancy in the angular correlation has been identified to-date so this problem remains interesting but open. Differentiation between the Monte Carlo models is possible by means of the correlation function data: ARIADNE, especially with the improved remnant treatment, describes the data well; this is not the case for LEPTO (at low Q^2) and HERWIG.

Chapter 10

Summary

A high luminosity DIS sample collected by the ZEUS detector at HERA up to 1997 has been analysed in the Breit frame-of-reference in terms of fragmentation functions and, for the first time, two-particle angular correlations.

The higher statistics of the new data, allowed extension of the kinematic range of the scaled momentum, x_P, distribution measurement up to $Q^2 \approx 10000$ GeV^2, well beyond the earlier analyses based on the HERA data. At the same time uncertainties of the statistically limited analysis were reduced. The new results on the cross-section differential in x_P, consistent with the existing ones, exhibit logarithmic dependence on the energy scale, given by the magnitude of the four-momentum transfer, Q, in the high Q, high x_P region. The dependence, analogous to the scale dependence in the hadronic structure functions, is a prediction of QCD now testable both qualitatively and quantitatively up to NLO. The calculation based on the fragmentation functions constrained by the e^+e^- annihilation data is in agreement with the DIS measurement thus suggestive of an independence of the parton fragmentation process from its production mechanism, known as universality of fragmentation. Based on the observed scaling violation and validity of the QCD calculation a measurement of the strong coupling constant, α_S, becomes possible as soon as appropriate input parameterizations of fragmentation functions are available. The measurement has an advantage compared to the "classic" DIS method through the dijet production rate as it is independent of the hadronization model.

The study of the logarithmic scaled momentum distributions, $\ln(1/x_P)$, in terms of the moments

of a *skewed* Gaussian shape has become possible due to the extended kinematic domain and reduced uncertainties. This has brought into trouble the relevant QCD predictions made in the Modified Leading Log Approximation as no set of the two parameters entering the calculation, the effective energy scale, Λ, and the cut-off for parton cascade evolution, Q_0, was found to lead to a simultaneous description of the evolution with energy of the measured peak position, width, and skewness of the distorted Gaussian distribution. At the same time the agreement with the e^+e^- data supports the universality of fragmentation. The mean charged multiplicity has been calculated by integration of the logarithmic distribution and was found to be consistent with a wealth of available DIS data as well as the e^+e^- measurements away from the low energy region where the formerly observed deviations have been confirmed. These differences are attributable to the significant rate of DIS subprocesses, such as scattering off sea-quarks and boson-gluon fusion, not present in the annilation environment and thus constitute no argument to disqualify the universality concept.

A new field of interparticle correlations has been explored through a study of relative angular distributions of charged particles contained in conical subregions of the current hemisphere in the Breit frame. The angular distribution itself, driven mainly by the single-particle distribution, has been observed to conform to a QCD-predicted scaling behavior with respect to energy, Q, and the cone half-opening angle, Θ, in agreement with similar observations made for e^+e^- at LEP. The predictions, available only in the Double Logarithmic Approximation, neglecting among others the four-momentum conservation compared to the MLLA framework, are in a good agreement with the DIS and annihilation data; such a consistency is a sign for both the universality and the validity of the Local Parton-Hadron Duality hypothesis for this simple two particle observable. However, the picture becomes less clear when the study is done of the single-particle effects subtracted true correlation function. There is less evidence for the scaling behavior in the ZEUS data strengthening a similar statement drawn from the LEP results. However, it is not the case that at least the universality is preserved: the e^+e^- and DIS data are incompatible even when the energy dependence is taken care of; the latter results at least are in the qualitative agreement at high energy with the analytic calculation.

The weaker stand, compared to the single-particle inclusive observables, of the theoretical predictions for the angular correlations may have at least two reasons. Firstly, the fragmentation benefits

from the refined MLLA approach not yet available for the correlations and, secondly, it is possible, that the LPHD hypothesis originally stated for single-particle spectra is really not extendable to the many-particle case.

Since the quality of Monte Carlo simulations is essential for a high energy experiment it is worth pointing out here that the color-dipole model as implemented in ARIADNE has proved to provide a reasonable description of the quantities studied in this thesis thus increasing the confidence in this numerical tool.

A tempting possibility of the α_S fitting to the DIS fragmentation data was already mentioned. A progress in the phenomenological parameterizations is a prerequisite but it should not prevent a concurrent experimental work to improvement the measurement. It should be instructive to pursue, and the first steps have been done at ZEUS, the correlation study in terms of other DLA-calculable observables like multiplicity correlations; this could provide a deeper insight into the LPHD and constrain models used for the quantitative QCD predictions.

Bibliography

[1] ZEUS Collab., M. Derrick *et al.*, *Phys. Lett. B* **345** (1995) 576.

[2] ZEUS Collab., M. Derrick *et al.*, *Z. Phys. C* **72** (1996) 399.

[3] ZEUS Collab., J. Breitweg *et al.*, *Eur. Phys. J. C* **7** (1999) 609.

[4] ZEUS Collab., J. Breitweg *et al.*, *Phys. Lett. B* **407** (1997) 402.

[5] H1 Collab., C. Adloff *et al.* submitted to *Nucl. Phys. B.*

[6] ZEUS Collab., J. Breitweg *et al.*, *Eur. Phys. J. C* **6** (1999) 239.

[7] ZEUS Collab., M. Derrick *et al.*, *Z. Phys. C* **72** (1996) 47.

[8] ZEUS Collab., M. Derrick *et al.*, *Phys. Lett. B* **363** (1995) 201.

[9] H1 Collab., C. Adloff *et al.*, *Eur. Phys. J. C* **6** (1999) 575.

[10] ZEUS Collab., J. Breitweg *et al.* submitted to *Eur. Phys. J. C.*

[11] H1 Collab., C. Adloff *et al.*, *Phys. Lett. B* **406** (1997) 256.

[12] ZEUS Collab., M. Derrick *et al.*, *Z. Phys. C* **67** (1995) 93.

[13] ZEUS Collab., J. Breitweg *et al.*, DESY 99-041 (hep-ex/9903056).

[14] H1 Collab., C. Adloff *et al.*, *Z. Phys. C* **75** (1997) 437.

[15] ZEUS Collab., J. Breitweg *et al.*, *Z. Phys. C* **74** (1997) 207.

[16] ZEUS Collab., J. Breitweg *et al.* to be submitted to *Eur. Phys. J. C.*

[17] ZEUS Collab., M. Derrick *et al.*, *Phys. Lett. B* **315** (1993) 481.

[18] ZEUS Collab., J. Breitweg *et al.*, *Eur. Phys. J. C* **1** (1998) 81.

[19] M. Kuhlen, MPI-PhE/97-33 (hep-ph/9712505).

[20] R. P. Feynman *Photon-Hadron Interactions.* Benjamin, N.Y. (1972).

[21] Y. L. Dokshitzer, V. A. Khoze, A. H. Mueller, and S. J. Troyan *Basics of Perturbative QCD.* Editions Frontières, Gif-sur-Yvette (1991).

[22] Y. Azimov *et al.*, *Z. Phys. C* **27** (1985) 65.

[23] V. A. Khoze and W. Ochs, *Int. J. Mod. Phys. A* **12** (1997) 2949.

[24] ZEUS Collaboration, The ZEUS Detector, Status Report 1993, DESY 1993.

[25] A. Andresen *et al.*, *Nucl. Inst. and Meth. A* **309** (1991) 101.

[26] A. Bernstein *et al.*, *Nucl. Inst. and Meth. A* **336** (1993) 23.

[27] A. Caldwell *et al.*, *Nucl. Inst. and Meth. A* **321** (1992) 356.

[28] C. Brooks *et al.*, *Nucl. Inst. and Meth. A* **283** (1989) 477.

[29] N. Harnew *et al.*, *Nucl. Inst. and Meth. A* **279** (1989) 290.

[30] B. Foster *et al.*, *Nucl. Inst. and Meth. A* **338** (1994) 254.

[31] B. Bock *et al.*, *Design, Construction and Test Results of the ZEUS Forward Tracking Detector*, ZEUS note 93-110.

[32] A. Bamberger *et al.*, *Nucl. Inst. and Meth. A* **401** (1997) 63.

[33] A. Bamberger *et al.*, *Nucl. Inst. and Meth. A* **382** (1996) 419.

[34] J. Andruszków *et al.* DESY 92-066.

[35] W. H. Smith *et al.*, *The ZEUS Trigger System*, ZEUS note 89-084.

[36] H. Abramowicz, A. Caldwell, and R. Sinkus, *Nucl. Inst. and Meth. A* **365** (1995) 508.

[37] J. Ng and W. Verkerke, *An Overview of SRTD Analysis*, ZEUS note 95-037.

[38] J. E. Cole, *A Study of Positron Energy Correction in 1995 DIS Data using the SRTD*, ZEUS note 98-022.

[39] H.-J. Grabosch, A. Meyer, and S. Schlenstedt, X_0 *Determination and Electron Energy Correction with the Presampler*, ZEUS note 98-045.

[40] A. Quadt and O. Ruske, *A New Method to Measure Vertex Distributions*, ZEUS note 98-036.

[41] G. F. Hartner, *VCTRAK Briefing: Program and Math*, ZEUS note 98-058.

[42] F. Jacquet and A. Blondel in *Proc. of the study of an ep facility for Europe* (U. Amaldi, ed.) DESY 79-48 (1979) 391.

[43] S. Bentvelsen, J. Engelen, and P. Kooijman in *Proc. of the Workshop on Physics at HERA* (W. Buchmueller and G. Ingelman, eds.) vol. 1 DESY (1992) 23.

[44] B. Levtchenko, *The Breit System (Primer)*, ZEUS note 94-099.

[45] V. N. Gribov and L. N. Lipatov, *Sov. J. Nucl. Phys.* **15** (1972) 438.

[46] G. Altarelli and G. Parisi, *Nucl. Phys. B* **126** (1977) 297.

[47] Y. L. Dokshitzer, *Sov. Phys. JETP* **46** (1977) 641.

[48] H. Plothow-Besch, *Comp. Phys. Commun.* **75** (1993) 396.

[49] M. Glück, E. Reya, and A. Vogt, *Z. Phys. C* **67** (1995) 433.

[50] H. L. Lai *et al.*, *Phys. Rev. D* **55** (1997) 1280.

[51] A. D. Martin, W. J. Stirling, and R. G. Roberts, *Phys. Rev. D* **51** (1995) 4756.

[52] ZEUS Collab., M. Derrick *et al.*, *Z. Phys. C* **65** (1995) 379.

[53] H1 Collab., S. Aid *et al.*, *Nucl. Phys. B* **439** (1995) 471.

[54] G. Ingelman, A. Edin, and J. Rathsman, *Comp. Phys. Commun.* **101** (1997) 108.

[55] G. Marchesini *et al.*, *Comp. Phys. Commun.* **67** (1992) 465.

[56] L. Lönnblad, *Comp. Phys. Commun.* **71** (1992) 15.

[57] B. Andersson *et al.*, *Phys. Rep.* **97** (1983) 31.

[58] T. Sjöstrand, *Comp. Phys. Commun.* **82** (1994) 74.

[59] H. Spiesberger "DJANGO6 version 2.4." unpublished program manual (1993).

[60] H. Spiesberger "HERACLES version 4.5.2." unpublished program manual (1993).

[61] R. Brun *et al.*, CERN DD/EE/84-1 (1987).

[62] L. Zawiejski and J. Okrasiński, *Charged Hadron Angular Correlations in DIS 1995 Sample*, ZEUS note 98-053.

[63] E. Gallo, *The TLT and DST filters for the DIS group in 1995*, ZEUS note 96-001.

[64] W. H. Smith *et al.*, *Nucl. Inst. and Meth. A* **355** (1995) 278.

[65] ZEUS Collab., J. Breitweg *et al.*, *Phys. Lett. B* **414** (1997) 428.

[66] V. A. Noyes *Measurement of Scaled Momentum Distributions in the Breit Frame at HERA using the ZEUS Detector.* PhD thesis Glasgow University November 1994.

[67] ZEUS Collab., M. Derrick *et al.*, *Phys. Lett. B* **316** (1993) 412.

[68] R. Waugh. PhD thesis Glasgow University. In preparation.

[69] L. Zawiejski and J. Chwastowski, *A Study of Angular Correlations Between Hadrons in Deep Inealastic Scattering Neutral Current Events at HERA*, ZEUS note 96-033.

[70] W. Ochs and J. Wosiek, *Z. Phys. C* **68** (1995) 269.

[71] W. Ochs. Private communication.

[72] M. Derrick *et al.*, *Phys. Lett. B* **91** (1980) 470.

[73] NOMAD Collab., J. Altegoer *et al.*, *Phys. Lett. B* **445** (1999) 439.

[74] TASSO Collab., W. Braunschweig *et al.*, *Z. Phys. C* **45** (1989) 193.

[75] PLUTO Collab., Ch. Berger *et al.*, *Phys. Lett. B* **95** (1980) 313.

[76] MARK I Collab., J. L. Siegrist *et al.*, *Phys. Rev. D* **26** (1986) 969.

[77] ALEPH Collab., R. Barate *et al.*, *Phys. Rep.* **294** (1998) 1.

[78] H1 Collab. "Fragmentation Measurements in the Breit Frame at HERA," contributed to ICHEP'98, ref. 531.

[79] J. T. Bromley *Measurement of Fragmentation Functions in DIS at HERA.* PhD thesis Glasgow University September 1998.

[80] ZEUS Collab., J. Breitweg *et al.*, .

[81] TASSO Collab., W. Braunschweig *et al.*, *Z. Phys. C* **47** (1990) 187.

[82] G. Kramer. Private communication.

[83] D. Graudenz, CERN-TH/96-52.

[84] J. Binnewies, B. A. Kniehl, and G. Kramer, *Z. Phys. C* **65** (1995) 471.

[85] D. Graudenz. CYCLOPS program, v1.1 and private communication.

[86] D. Graudenz, PSI PR/98-29 (hep-ph/9903295).

[87] J. Binnewies *Fragmentation Functions in Next-To-Leading Order QCD.* PhD thesis University of Hamburg (1997). DESY 97-128 (hep-ph/9707269).

[88] C. P. Fong and B. R. Webber, *Phys. Lett. B* **229** (1989) 289.

[89] C. P. Fong and B. R. Webber, *Nucl. Phys. B* **355** (1991) 54.

[90] Y. L. Dokshitzer, V. A. Khoze, and S. I. Troyan, *Int. J. Mod. Phys. A* **7** (1991) 1875.

[91] DELPHI Collab., P. Abreu *et al.*, *Phys. Lett. B* **440** (1998) 203.

[92] L3 Collab., M. Acciarri *et al.*, *Phys. Lett. B* **428** (1998) 186.

[93] L3 Collab., M. Acciarri *et al.*, *Phys. Lett. B* **429** (1998) 375.

[94] L. Lönnblad. Private communication.